How Dependent is Growth from Primary Energy ? Output Energy Elasticity in 50 Countries (1970-2011)*

Gaël Giraud
CNRS, CES, PSE, REFI†

Zeynep Kahraman
TSP‡

April 10, 2014

Abstract

Except for specialized resource economics models, neoclassical economics pays little attention to the role of energy in growth. This paper examines the basic flaws behind the mainstream analytical arguments for this neglect, and provides an empirical reassessment of this role. We use an error correction model in order to estimate the long-run output elasticity of primary energy use in 50 countries between 1970 and 2011. By contrast with mainstream macroeconomics, our findings show that this elasticity lies between 0.6 and 0.7. This estimation is robust to the choice of various samples of countries and subperiods of time. In addition, we show that energy and growth are cointegrated and primary energy consumption univocally Granger causes GDP growth. This confirms the results on cointegration and causality between energy consumption and growth already obtained elsewhere (Stern (2010)).

JEL codes: N70, O40, Q43.

1 Introduction

Most of the mainstream economic models used to explain the growth process (Aghion and Howitt (2008) or Blanchard et al. (2010)) do not include energy as a factor that could foster economic growth. Because thermodynamics implies that energy should

*We wish to thank The Shift Project for financial support, and especially Jean-Marc Jancovici for fruitful discussions, as well as participants of seminars at Paris-1 university, Paris School of Economics, Veblen institute, and Dauphine university. The usual *caveat* applies.

†CNRS, Centre d'Economie de la Sorbonne, Paris School of Economics, Labex Financial Regulation. E-mail: gael.giraud@normalesup.org

‡The Shift Project. E-mail: zeynep.kahraman@theshiftproject.org

be essential to all economic processes, ecological economists, on the other hand, often ascribe to energy the central role in economic growth (see Stern (2010) for a survey). Ecological economists derive their view of the role of energy in economic growth from the biophysical foundations of the economy discussed, e.g., by Georgescu-Roegen (1971), Costanza (1980), Cleveland et al. (1984), Hall et al. (2003), Ayres and Warr (2009), Murphy and Hall (2010). Some geographers (e.g., Smil (1994)) and economic historians (e.g. Wrigley (1988), Allen (2009), Arnoux (2014)) also argue that energy played a crucial role in economic growth, as well as in the industrial revolution.

Is energy an important driver of economic growth ? And, if so, what is the magnitude of the dependency of growth from energy ? Or should the obvious instrumentality of energy be understood as translating itself into, say, the driving force of capital, whose accumulation is thought of by most economists, at least since Smith and Ricardo, as being the secret of growth ? This paper attempts to pave the road towards an answer to these questions by empirically estimating the long-run output elasticity of primary energy use per capita in 50 countries (including most OECD countries and energy exporting countries) between 1970 and 2011. In order to remain as close as possible to data to which mainstream economists are accustomed to, we measure energy in million tons of oil equivalents, and refrain from using the (otherwise quite interesting) indexing methods that account for differences in quality among fuels (Stern (1993), Ayres and Warr (2009)). On the other hand, whenever important variables are omitted from the analysis, it is known that no cointegration emerges between the variables under study, while a spurious regression will result. We therefore include energy efficiency and capital in our analysis. As for capital, we test whether adding it, or not, as an explanatory variable significantly changes the elasticity of energy use. Finally, as observed in Stern's (2010) synthesis, in many cases, results on the relationship between energy and output differ depending on the samples used, the countries investigated etc. In order to check whether our finding is robust to the choice of countries and time periods, we repeat the whole exercise on a subsample of 15 OECD countries. This choice was also motivated by the idea of testing whether more "advanced" countries are less energy-dependent than emerging ones. Eventually, we end up with three Protocols depending upon the country list and the variables under scrutiny: Protocol 1 includes 50 countries but no capital; Protocol 2, 48 countries[1] and capital; Protocol 3: 15 OECD countries and capital.

To put it in a nutshell, our findings amply confirm the standpoint defended by ecological economists: primary energy is a key factor that drives GDP growth. For the countries listed in Protocol 1, and during the last 4 decades, its long-run output elasticity evolved between 0.6 and 0.7. This means that, *ceteris paribus*, an increase (resp. decrease) of 10% of energy use per capita induces, on average, an increase (resp. decrease) of about 6 to 7% of GDP per capita. This is best illustrated by Figure **??**, where x-axis reports GDP-growth at the world level, and the y-axis, the growth of energy use. A simple regression suffices to provide the right magnitude of the output elasticity of primary energy. Our paper can be viewed as an attempt to evaluate the robustness of this seemingly obvious order of magnitude.

By contrast, we find a long-run elasticity for capital around 0.2, suggesting that, at least in the recent decades, capital accumulation has played a minor role compared to energy. Adding capital (Protocol 2) does not alter this robust long-run relationship. Moreover, the subsample of OECD countries (Protocol 3) turns out to

[1]For reason of lack of trustable data availability, we had to discard 2 countries from the first list.

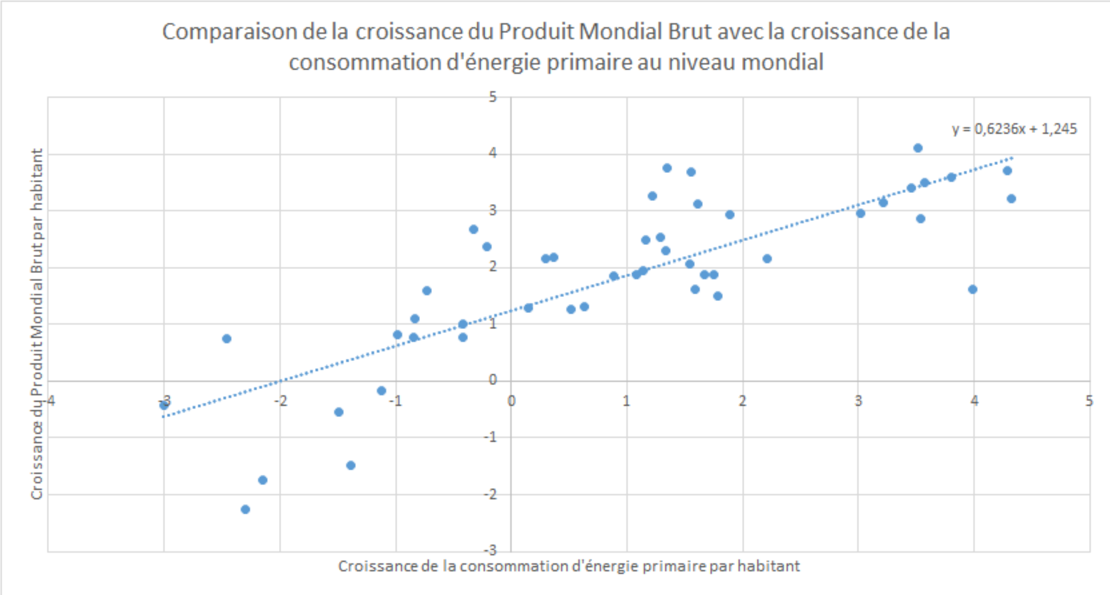

Figure 1: **A rough estimation of energy elasticity around 0.6.**
Source: EIA, http://www.eia.gov/totalenergy/data/annual/pdf/sec1_13.pdf

be slightly less dependent from primary energy consumption but still the estimated average output elasticity of energy is 0.6.[2] In addition, these estimations are also robust to the choice of various subperiods of time.

These findings sharply contrast with the custom, popular in macroeconomics, to calibrate the output elasticity of energy according to the cost share of energy. In most countries, this practice leads to the postulate that energy elasticity should be close to 0.08 on average —at least a factor 8 lower than what data tell us. Given the inevitable scarcity of natural resources that most countries will presumably encounter in the coming decades, such a factor matters. We therefore begin this paper with an internal critique of the conventional argument (known as the cost-share theorem) which underlies the popular calibration. Our conclusion is that there are good reasons to believe that, in general, the output elasticity of energy is decoupled from its GDP share. This opens the door for an empirical reassessment of its output elasticity, to which the rest of the paper is dedicated. The possibility of decoupling between elasticity and share holds for any production factor. It should therefore also encourage the reexamination of many other input factors, such as land. The classical factor of land, including all natural resource inputs, gradually diminished in importance in economic theory as its value share of GDP fell steadily in the 20th century (cf. Schultz (1951)) and today is usually subsumed as a subcategory of capital. The analysis provided in this paper suggests that this might be an important mistake. Here, we focus on energy but subsequent work will be devoted to extending our inquiry to other natural resources.

On the empirical side, our analysis addresses several controversial issues, and provides new striking results. In recent decades, indeed, significant reductions in energy intensity have been achieved in many developed and some developing countries

[2]Further work on Western European countries (not reported here) would show that it can be even close to 0.3, in ecologically more "forward-looking countries". But in any case, we never found an output elasticity of energy below 0.3.

(Gales et al., (2007), Stern (2010)). Some key factors could reduce or strengthen the linkage between energy use and economic activity over time. Among them, one can think of substitution between energy and other inputs (within an existing technology), technological change, shifts in the mix of energy input, etc. This evolution might mitigate the influence of energy use on growth in the long-run. As we introduced energy efficiency in our analysis, this possible effect will be captured by the long-run output elasticity of energy efficiency. This latter parameter turns out to lie, as well, between 0.6 and 0.7. Thus, our inquiry does *not* suggest that energy use be the sole first-order factor driving growth. Efficiency plays a dual, almost comparable, role.

Our result therefore balances the opinion defended by some ecological economists according to whom substitution between capital and resources and technological progress can only play a limited role in mitigating the scarcity of resources. Cleveland et al. (1984), Hall et al. (2003), for instance, downplay the role of technological change, arguing that either increased energy use accounts for most apparent productivity growth, or that technological change is real but innovations mainly increase productivity by allowing the use of more energy. Therefore, it is argued, increased energy use should be the main cause of economic growth. Our impression is that the primary role attributed to energy use should not, however, hidden the factor of efficiency.

Very much like the mere correlation concept, elasticity does not say anything about the possible causal relationship between growth and energy. Our last contribution consists in reexamining the empirical evidence on this long-standing and intricate question: is it the ability to consume abundant resources of energy that fosters growth or, on the contrary, is it growth that, being driven by another engine (e.g., capital), mechanically increases the use of energy? Tests for causality between energy, GDP, and other variables started as early as in the late 1970's (Kraft and Kraft (1978)). While early studies relied on Granger causality tests on unrestricted vector autoregressions (VAR) in levels of the variables, we follow more recent studies by using cointegration methods in a multivariate framework (see Yu and Jin (1992) for the first cointegration study of the energy/GDP relationship, and Stern (2000) for a more recent study).[3]

The results of the early studies that tested for Granger causality using a bivariate model were generally inconclusive (Stern (1993)). Where nominally significant results were obtained, they mostly indicated that causality runs from output to energy. Stern (1993) tested for Granger causality in a multivariate setting using a VAR model of GDP, capital and labor inputs, and a Divisia index of quality adjusted energy use in place of gross energy use. It then was found that energy Granger causes GDP. Stern (2000) estimated a dynamic cointegration model for GDP, quality weighted energy, labor, and capital, using the Johansen methodology. The analysis showed that there is a cointegrating relation between the four variables and that energy Granger causes GDP either unidirectionally or possibly through a mutually causative relationship. 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.

[3]The multivariate methodology is important because reductions in energy use are frequently countered by the substitution of other factors of production for energy and vice versa, resulting in an insignificant overall impact on output.

Oh and Lee (2004) and Ghali and El-Sakka (2004) applied Stern's methodology to Korea and Canada, respectively, coming to the same conclusions, hence extending the validity of Stern's results beyond the United States. 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 a three and four decade period 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 bi-directional relationship in the OECD sample.

This body of work suggested that the inconclusive results of earlier work are probably due to the omission of non-energy inputs. Here, our conclusion is that energy and GDP cointegrate and energy use univocally Granger causes GDP in the long-run. We reach this outcome on a larger sample of countries and a longer time period than most of the earlier studies, without even having to employ a quality adjusted energy index.

The paper is organized as follows. Section 2 provides the critical evaluation of the cost-share theorem. We believe that the focus on prices, rather than quantities, is responsible for a large part of the controversies that have long prevailed on the role of energy in growth. The empirical estimation of energy output elasticity is presented in section 4. Section 5 is devoted to the causality issue. Our findings have a number of implications for the future and for policy-making. Some of them are discussed in the last section.

2 The analytical arguments

In this section, we briefly review the standard argument which, within the neoclassical literature, explains why energy output elasticity should allegedly be low (at least below 0.1). We then proceed with three different lines of criticism that show that this argument contains basic flaws. This will open the door for a reexamination of the empirical estimation of primary energy output elasticity, which is pursued in the next section.

2.1 The cost share theorem

This textbook result argues that, in perfectly competitive markets, under constant returns to scale and absent any externality of omitted variables, output elasticity of any production factor should equal its cost share. Denoting $x = (x_i)_{i=1,\dots,n}$ the input vector, $Y(\cdot) \in \mathcal{C}^1(\mathbb{R}^n)$, the production function, and $p = (p_i)_i \in \mathbb{R}^n_+$, the price of inputs, the profit maximization program of the representative producer reads:[4]

$$\max_x \ PY(x) - p \cdot x, \tag{1}$$

where $P \in \mathbb{R}_{++}$ is the output price. Under the above mentioned regularity and convexity conditions, its first-order condition leads to:

$$\varepsilon_i := \frac{x_i}{Y(x)} \times \frac{\partial Y}{\partial x_i}(x) = \frac{p_i x_i}{p \cdot x} \tag{2}$$

[4]Again, mainstream economics often assumes that a country's production sector can be best represented by a unique producer behaving as a single individual.

where ε_i is the output elasticity of the production factor, i.

One way to derive equation (2) from economic primitives is to think of it as equivalent to assessing that, at equilibrium, the output market price, P, should be equal to the marginal cost of each firm populating the industry sector. Suppose, indeed, that industry is made of $n \geq 1$ firms, where the output of the j^{th} firm is y_j. The price scalar, P, arises, in fact, from the inverse excess demand, $D^{-1}(Y)$ (e.g., Mas-Colell et al. (1995), p. 580 sq), at the aggregate output $Y := \sum_{j=1}^{n} y_j$. And the mapping $D(\cdot)$ itself stems from the utility-maximization programme of households, $h = 1, ..., m$, provided their utility functions be all sufficiently regular, so that $D(\cdot)$ is invertible:

$$\forall h, \quad d^h(P) := \text{Argmax}_{z_h \geq 0} u_h(z_h) \text{ s.t. } Pz_h \leq w_h$$

where $w_h \geq$ is the wealth of household h. Clearly, $D(P) := \sum_h d^h(P)$. Thus, given some aggregate output, Y, $P(Y)$ denotes the market-clearing price, i.e., is such that, whenever every consumer chooses optimally her consumption bundle, the resulting aggregate excess demand equals Y.

Following the Marshallian tradition pervasive in the entire neoclassical micro- and macro-economics, a corporate in a competitive industry will not react strategically to the behavior of its competitors. Thus, firm j will only consider the market price vector (P, p), when choosing an input vector, $x^j = (x_i^j)_i$, so as to maximize its individual profit:

$$\text{Max}_{\ x^j} P(Y) y_j(x^j) - p \cdot x^j. \tag{3}$$

Hence, it is argued, j should choose the level of each input, x_i^j, so as equalize its marginal revenue, $\partial(P(Y)y_j(x^j))/\partial x_i^j$ with its marginal cost, p_i. As we have assumed that j is negligible and takes prices as given,

$$\frac{\partial P(Y)}{\partial x_i^j} = \underbrace{\frac{\partial P(Y)}{\partial y_j}}_{=0} \times \frac{\partial y_j}{\partial x_i^j} = 0. \tag{4}$$

This leads to

$$P(Y) \frac{\partial y_j}{\partial x_i^j} = p_i. \tag{5}$$

On the other hand, constant returns to scale mean that the function $Y(\cdot)$ is 0-homogeneous, which yields the Euler equation:

$$Y(x) = \sum_i x_i \frac{\partial Y}{\partial x_i}(x) = \sum_i x_i \frac{\partial y_j}{\partial x_i}(x) = \sum_i x_i \frac{p_i}{P(Y)}.$$

This implies

$$P(Y)Y = \sum_i p_i x_i = p \cdot x. \tag{6}$$

An individualized version of the cost-share formula then follows from (5) and (6):

$$\frac{x_i}{Y(x)} \times \frac{\partial Y}{\partial x_i^j}(x) = \frac{p_i x_i}{p \cdot x}, \tag{7}$$

from which (2) follows once one realizes that a variation of the aggregate output, $Y(x)$, must arise from the change of *some* individual output, so that $\partial Y(x)/\partial x_i^j = \partial Y(x)/\partial x_i$.

When applied to primary energy as an input factor, and provided output stands for GDP, this argument readily implies that the GDP elasticity of energy should lie between 0.08 and 0.1 on average. This is indeed the range of values most often taken by the cost share of energy in rich countries, within the last decades. To take but an example, Figure 2 provides the evolution of the primary energy share in the U.S. GDP, between 1970 and 2010.

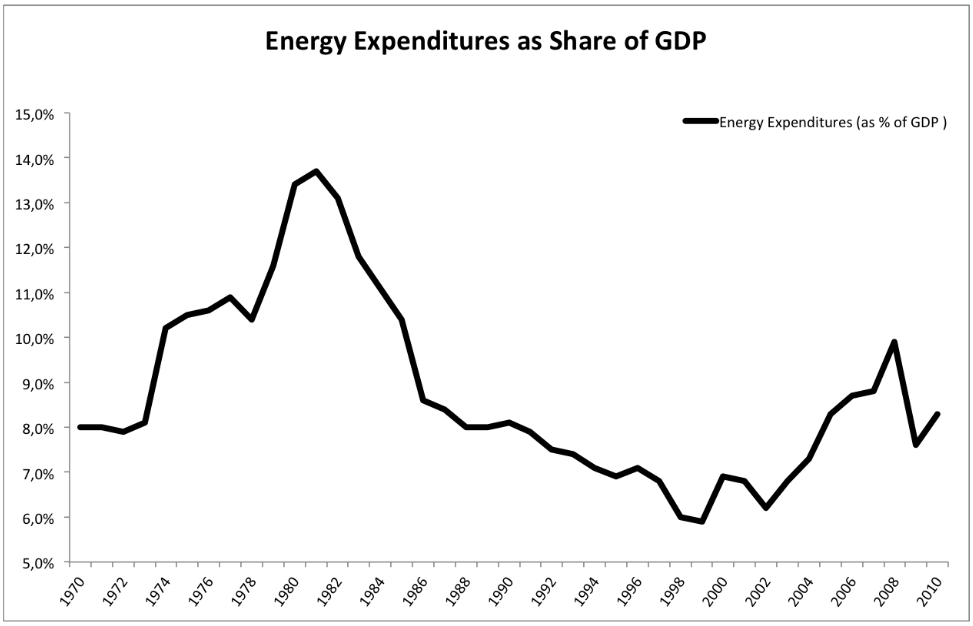

Figure 2: **The GDP share of primary energy, U.S., 1970-2010.**
Source: EIA, http://www.eia.gov/totalenergy/data/annual/pdf/sec1_13.pdf

More precis

In the following subsections, we review three elementary arguments explaining why, in spite of its popularity,, the equality (2) may permanently fail to be empirically satisfied.

2.2 A flaw in the concept of perfect competition

The first one concerns the basic assumption of "perfect competition" underlying the optimization programme (1) itself, and the way we derived (2). As already emphasized, this formulation assumes that every producer takes prices as exogenously given. Indeed, being "negligible", each firm's behavior has no effect on the market price vector. (1) assumes that this remains true at the aggregate level, when the production sector is approximated by the behavior of a representative firm. By contrast, in a monopolistic set-up (where the whole production sector can indeed be represented by a single firm), the producer must take into account in its maximization problem the influence of its own production plan on prices.

There is, however, a fallacy in the argument above. Indeed, for the inverse demand function, $P(\cdot)$, to be defined, we need the aggregate excess demand, $D(\cdot)$, to be (at least locally) invertible, hence to have a non-zero derivative at Y : $\partial D(y)/\partial y \neq 0$

at least in (a neighborhood of) the point $y = Y$. Hence $\partial D^{-1}(p)\partial p \neq 0$ at $p = P(Y)$. Since most of the neoclassical literature presumes that the "law of demand" is fulfilled in any well-functioning market,[5] the sign of this nonzero derivative should be clear: $\partial D^{-1}(p)/\partial p < 0$ at $p = P(Y)$. But, at the same time, (4) requires that $\partial P(Y)/\partial y_j = 0$ for every j. Notice that the contradiction is independent from the size of the integer n. A similar argument had been first sketched by Stigler (1957), and formulated by Keen (2006).

The difficulty arises from the erroneous interpretation of perfect competition underlying the reasoning above, where it is implicitly and abusively assumed that, as $n \to +\infty$, the economy behaves as an economy where the set of firms is an atomless measure space $(I, \mathcal{I}, \lambda)$.[6] In the later case, indeed, since $\lambda(\{j\}) = 0$, and $Y = \int_I y_j d\lambda(j)$, then $\partial P(Y)/\partial y_j = 0$. But, as long as n is finite, *whatever being its size*, one must have, under the conditions stated *supra*, $\partial P(Y)/\partial y_j = \partial D^{-1}(Y)/\partial y_j \neq 0$.

Taking into account the (possibly arbitrarily large but) finite number of firms, the exact first-order condition of (3) reads:[7]

$$P(Y)\frac{\partial y_j}{\partial x_i^j} + \frac{\partial D^{-1}}{\partial y_j}(Y) \times \frac{\partial y_j}{\partial x_i^j} = p_i.$$

The correct cost-share theorem is therefore:

$$\frac{x_i}{Y(x)} \times \frac{\partial Y}{\partial x_i}(x) = \frac{p_i x_i}{p \cdot x} - \frac{\partial D^{-1}}{\partial y_j}(Y) \times \frac{x_i}{p \cdot x} \times \frac{\partial y_j}{\partial x_i^j} \tag{8}$$

Keeping in mind that $\partial D^{-1}(Y)/\partial y_j < 0$, the consequence is twofold: 1) the maximal profit for an individual firm cannot be attained unless "marginal revenue" *exceeds* its "marginal cost". 2) Equivalently, at equilibrium, the output elasticity ε_i will exceed the cost-share.

2.3 Shadow prices

Let us now accept, for the sake of the discussion, the neoclassical way to formalize perfect competition. We next show that, even then, the cost-share theorem relies on very vulnerable grounds. The textbook argument recalled in subsection 2.1 rests, indeed, on the assumption that the representative producer's maximization program (1) faces *no* constraint apart from the very definition of $Y(\cdot)$. Suppose, on the contrary, that (1) must be written, somewhat more realistically:

$$\max_x \; Y(x) - p \cdot x \quad \text{s.t. } f(x) = 0 \tag{9}$$

where $f \in \mathcal{C}^1(\mathbb{R}^n)$ is some smooth function. Whenever the input, x_i, is interpreted as fossil energy, we can think of $f(\cdot)$ as capturing geological resource restrictions on fossil energies, geopolitical or climatic constraints, the bargaining power of labor forces, institutional rigidities of the labor market, etc. Even if one accepts the (self-contradictory) postulate that $\partial P(Y)/\partial y_j = 0$, the cost-share identity (2) now

[5] See Giraud and Quah (2004) for a discussion of this point.

[6] Cf. Aumann (1964) for a seminal formulation of an atomess economy.

[7] The gap between (2.2) and (5) is related to the lack of upper-hemi-continuity of the equilibrium correspondence with respect to the number of firms —a problem of which neoclassical economists are perfectly aware, cf. Mas-Colell (1982).

involves a shadow price given by the (normalized) Lagrange multiplier, λ, of the additional constraint, $f(x) = 0$:

$$\varepsilon_i = \frac{x_i\big(p_i - \lambda \frac{\partial f(x)}{\partial x_i}\big)}{p \cdot x - \lambda x_i \frac{\partial f(x)}{\partial x_i}}. \tag{10}$$

To our knowledge, this observation has been made in **?** and **?**. It follows that shadow prices may be responsible for the decoupling between the energy share, $p_i x_i / p \cdot x$, and its output elasticity, ε. Suppose, for instance, that the cost share remains small, while $\lambda \to +\infty$. Then, $\varepsilon_i \to 1$. Similarly, ε may take any real value between $x_i p_i / x \cdot p$ and $-\infty$ whenever $0 < \lambda < (p \cdot x)\frac{\partial x_i}{\partial f(x)}$. So that a large share $x_i p_i / x \cdot p$ is compatible with a small ε_i. At variance with the flaw underlined in the previous subsection, the latter argument for decoupling prevents us from concluding that one factor's return is underpaid (when the profit share is below its output elasticity) or overpaid (in the opposite situation): both might well exhibit a "fair return" once all the constraints in the production sector have been taken into account.

2.4 GDP versus gross output

Last, suppose not only that the conventional model of perfect competition is valid (i.e., $\partial P(Y)/\partial y_j = 0$) and that the representative producer's behavior can be fairly modeled as an unconstrained optimization program. Even in such a framework, a decoupling between the energy cost share and its GDP elasticity may appear once due account is taken of the difference between gross output and GDP. This is best seen by considering, e.g., the Solow growth model with energy introduced in Stern and Kander (2011). Omitting time indices for simplicity, the model consists of a specific production function

$$Y = \Big[(\gamma_V^{\frac{1}{\sigma}}\big(A_L^{\alpha} L^{\beta} K^{1-\beta}\big)^{\phi} + \gamma_E^{\frac{1}{\sigma}}(A_E E)^{\phi}\Big]^{\frac{1}{\phi}} \tag{11}$$

and a capital accumulation dynamics:

$$\dot{K} = sY - \delta K. \tag{12}$$

Equation (11) embeds a Cobb-Douglas function of capital (K) and labor (L) in a CES function[8] of value added, $\big(A_L^{\beta} L^{\beta} K^{1-\beta}\big)^{\phi}$, and energy, E, to produce gross output, Y. The parameter $\phi := (\sigma - 1)/\sigma$, where $\sigma \in \mathbb{R}_+$ is the elasticity of substitution between energy and value added aggregate. A_L and A_E are the augmentation indices of labor and energy, which can be interpreted as reflecting both changes in technology that augment the effective supply of the respective factor and changes in the quality of this factor.[9] Equation (12) is the standard ODE for capital that assumes, as in most of the Solowian macro-economic literature, that a fixed proportion, $s \in (0, 1)$, of gross output is saved while capital depreciates at a constant rate, $\delta \in (0, 1)$.

The elasticity of gross output, Y, with respect to energy is given by:

$$\frac{\partial \ln Y}{\partial \ln E} = \gamma_E^{\frac{1}{\sigma}}\big(A_E \frac{E}{Y}\big)^{\phi},$$

[8] Constant Elasticity of Substitution, cf. Mas-Colell et al. (1995) , Ex. 3C6, p. 97.
[9] An identical production function is used in Acurio ?? and Benasy-Fontagné.

and is equal to the cost-share of energy in terms of gross output, $p_E E/Y$. Under the constant returns to scale hypothesis, the GDP, G, is

$$G = Y\big(1 - \frac{\partial \ln Y}{\partial \ln E}\big) = \gamma_V^{\frac{1}{\sigma}} Y^{1-\phi} \Big[(A_L L)^\beta K^{1-\beta}\Big]^\phi,$$

so that its elasticity with respect to energy is

$$\frac{\partial \ln G}{\partial \ln E} = (1-\phi)\frac{\partial \ln Y}{\partial \ln E}$$

.

On the other hand, the share of energy in GDP is

$$\sigma_G^E := \frac{\frac{\partial \ln Y}{\partial \ln E} Y}{G} = \frac{\gamma_E^{\frac{1}{\sigma}}\big(A_E E\big)^\phi}{\gamma_V^{\frac{1}{\sigma}}\Big[(A_L L)^\beta K^{1-\beta}\Big]^\phi}.$$

As a consequence, the GDP elasticity of energy can be expressed as a function of σ_G^E:

$$\begin{aligned}\frac{\partial \ln G}{\partial \ln E} &= \sigma_G^E \times \frac{(1-\phi)}{Y^\phi} \gamma_E^{\frac{1}{\sigma}} \big(A_E \frac{E}{Y}\big)^\phi \\ &= \sigma_G^E \times \underbrace{\frac{(1-\phi)}{\beta}\big(\frac{\gamma_E}{\gamma_V}\big)^{\frac{1}{\sigma}} \frac{\partial \ln Y}{\partial \ln L}}_{\to +\infty \text{ as } \phi \to -\infty}.\end{aligned}$$

It follows that, whenever the elasticity of substitution between energy and value added, ϕ, is far below zero, the GDP elasticity of energy becomes *larger* than the share of energy in GDP. This occurs, in particular, when energy and value added are poorly substitutable.

The three arguments listed above provide a strong case for believing that, in general, there is no reason to postulate the identity (2). Several other points are worth being mentioned, which further fragilize the faith in the cost-share theorem: so far, we have assumed that increasing returns to scale are constant. That most empirical studies lead to this conclusion is merely an artifact from

Moreover, they show that, in general, the GDP elasticity of energy should be larger than the share of energy in GDP. Let us now confirm this empirically, and measure the size of the gap. As we shall see, we find that, for the chosen panel of countries and within the time period 1970-2011, there is, on average, a factor close to 8 between the GDP elasticity of energy and its GDP share. According to the viewpoint sketched in subsections **??** , this . Following the standpoint of the present subsection, this suggests that the elasticity of substitution between energy and value added might have been close to -7. This last conjecture is consistent with the findings obtained by **?** with completely different methods.

3 An empirical assessment

Classical panel data estimation methods as Fixed Effects (FE) and Random Effects (RE) can produce inconsistent and potentially very misleading estimates of the average values of the parameters in dynamic panel data models when the slope coefficients are not identical. And in most panels, these parameters differ significantly across groups. To deal with this matter, Pesaran and Smith (1995) suggest a mean group estimator (MG) based on average of the estimated coefficient of each cross section. However this estimator does not take into account of the fact that certain parameters may be the same across groups.

Alternatively, an intermediate estimator, the Pooled Mean Group (PMG) estimator has several advantages for the purpose of our analysis as it combines the characteristics of efficiency of the pooled estimators with those of the mean group estimator. This method based on maximization of the log-likelihood function by means of the Newton-Raphson algorithm. The main advantage of the PMG method is it only constraints the long-run coefficients to be the same for the cross-sectional units but allows the short-run coefficients, speed of adjustment and error variances to differ among groups. This weak homogeneity assumption characteristic of this method makes it attractive over the traditional methods.

The PMG estimator generates consistent estimates of the mean of short-run coefficients across countries by taking the simple average of individual country coefficients. It can be argued that country heterogeneity is particularly relevant in short-run relationships, given that countries are affected by over lending, borrowing constraints, and financial crises in short-time horizons, albeit to different degrees. On the other hand, there are often good reasons to expect that long-run relationships between variables are homogeneous across countries.

Several practical points on the PMG estimation are worth noting. First, the time dimension has to be long enough to allow estimation of the model for each of the cross-sections separately. Second, the lag order has to be long enough to ensure that the residuals of the error correction model are serially uncorrelated but not too long to cause a serious loss of degrees of freedom. In this respect, there is a trade of between loss of degrees of freedom when including too many lags (relative to time series dimension) and loss of consistency when including too few lags. The optimal number of lags is best chosen according to an information criterion such as Akaike Information Criterion (AIC) or the Schwarz Bayesian Criterion (SBC).

Another advantage of the PMG estimator is that it is consistent when data have complex country-specific short-term dynamics which cannot be captured applying the same lag construction for all groups. Furthermore, as long as the PMG estimator does not impose any restriction on short—term coefficients, it provides information on country specific values of the speed of adjustment to the long-run relationship.

The restrictions of homogenous long-run coefficients and the error correction term in each model can be tested by Hausman test as proposed by Pesaran et al. (1999). The PMG parameter estimates are consistent and efficient only if homogeneity holds. Otherwise, the MG estimation method is preferred. Thus, it can be evaluated whether imposing long-run homogeneity helps to disclose significant adjustment of the factor demands to long-run equilibrium.

Moreover, we apply the Common Correlated Effects Mean Group (CCE-MG) methodology of Pesaran (2006) to the PMG estimator to correct for the cross-sectional dependencies arising from omitted common factors (such as common shocks), as we assume that countries are affected in different ways and to varying degrees by these shocks. CCE-MG estimator is similar to the Mean Group estimator but includes cross-section means of the independent variables as regressors, capturing cross-section dependence.

In addition, we perform a Granger causality tests to investigate the casual relationship between energy consumption and economic growth.

3.1 The Econometric Model

Most of the earlier studies on the energy consumption and growth nexus are evaluated within a bivariate framework. To avoid the omitted variable issue this study examines the relationship within a multivariate framework by including the energy efficiency to measure the technology progress as in (2) and the gross capital formation.

For each country, the long-run relationship under scrutiny is:

$$y_{it} = \alpha_{0i} + \alpha_{1i}c_{it} + \alpha_{2i}e_{it-1} + \alpha_{3i}k_{it} + \varepsilon_{it}, \tag{13}$$

where $i = 1, ..., 15$ refers the countries, $t = 1970, ..., 2011$ is the time period, y_{it} stands for the logarithm of the GDP per capita, c_{it} is the logarithm of the energy consumption per capita, e_{it} is the logarithm of energy efficiency and k_{it} is the logarithm of the gross capital formation per capita. The reason for the one-period lag in energy efficiency and the addition of capital, e_{it-1}, is that, without these two features, (13) coincides with the tautology (**??**), hence becomes statistically trivial. The presence of capital, in addition, enables to measure, as a by-product, the output elasticity of this traditionally important input factor.

The equation estimated through our error correction model will enable us to quantify the speed at which the long-run relationship (13) is restored after an exogenous shock:

$$\Delta y_{it} = \beta_{1i}\Delta c_{it} + \beta_{2i}\Delta e_{it-1} + \beta_{3i}\Delta k_{it} + \gamma\Big[y_{it} - \big(\alpha_{0i} + \alpha_{1i}c_{it} + \alpha_{2i}e_{it-1} + \alpha_{3i}k_{it}\big)\Big] + \varepsilon_{it}, \tag{14}$$

3.2 The Data

The analysis is based on a panel data covering the period from 1970 to 2011 for 15 countries (Table 3.2). The data set used in the analysis is gathered from different sources. The annual data on primary energy consumption (million tons of oil equivalents) obtained from the BP Statistical Review of World Energy 2012. GDP (in 2000 U.S dollars), Gross Fixed Capital Formation (in 2000 U.S dollars) and Population data are provided by World Bank, World Development Indicators. Prior to empirical research, all data were converted into logarithms. Therefore, the estimated coefficients reflect constant elasticities.

Table 1: Country List 1

Austria	Grece	Slovak Pepublic
Belgium	Ireland	Spain
Finland	Italy	United Kingdom
France	Netherlands	Japan
Germany	Portugal	United States

The first list contains all the current OECD countries with the exception of Estonia, Iceland, Israel, Korea, Luxembourg, and Slovenia. On the other hand, it

Table 2: Country List 2

Austria	Grece	Slovak Pepublic
Belgium	Ireland	Spain
Finland	Italy	United Kingdom
France	Netherlands	Japan
Germany	Portugal	United States

includes the following non-OECD countries: Algeria, China, Ecuador, Egypt, Hong-Kong, India, Indonesia, Iran, Russia, Saudi Arabia, Singapore, South Africa, South Korea, Kuwait, Malaysia, Pakistan, Philippines, Qatar. The advantage of considering a more comprehensive sample of countries, in addition to the mere gain in generality, is that it enables to reduce the bias induced by the outsourcing of energy use: indeed, the dependency of a country with respect to primary energy can be underestimated due to the fact that the energy needed for the production of its imports is accounted out of this very country. Thus, big importers are likely to exhibit an underestimated output elasticity of energy (this is especially the case for the Old World). The unique way to circumvent this difficulty would consist in considering all the countries together. Our first list is a proxy of this strategy.

3.3 Cross section dependence tests

Traditional panel data estimation methods assume the independence of the cross-sections. However it is well known that presence of common shocks or spillover effects can cause correlations across countries. Consequently, the assumption of cross-sectional independence in the existence of the omitted common factors in the error terms can lead to inconsistent and misleading estimates. Hence, before examining the order of integration of our series and testing for co-integration, we test the hypothesis of cross sectional independence.

The cross-section dependence (CD) test proposed by **?** tests the null hypothesis of independence across the cross sections. One of the key features of this test is its robustness to structural breaks. The CD test is simply based on an average of all pairwise correlations of the ordinary least squares (OLS) residuals obtained from the individual regressions in the panel data model. The CD statistic can be defined as:

$$CD = \sqrt{\frac{2T}{N(N-1)}} \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{ij} \right) \longrightarrow \mathcal{N}(0,1). \tag{15}$$

where $\hat{\rho}_{ij}$ is the estimate of the pairwise correlation:

$$\hat{\rho}_{ij} = \hat{\rho}_{ji} := \frac{\sum_{t=1}^{T} \hat{\varepsilon}_{it}\hat{\varepsilon}_{jt}}{\left(\sum_{t=1}^{T} \hat{\varepsilon}_{it}^2\right)^{1/2} \left(\sum_{t=1}^{T} \hat{\varepsilon}_{jt}^2\right)^{1/2}} \tag{16}$$

Table 3.3 reports the CD test statistics and the corresponding p-values. The results of the test indicate that the null hypothesis of cross-section independence is rejected for all of the variables. Therefore, cross section dependence should be taken into account during following steps.

Table 3: Pesaran cross-section dependence (CD) test results

Variable	**CD-test**	$\hat{\rho}$	**p-value**
Y	61.37	0.964	0.000
E	26.81	0.438	0.000
C	31.17	0.487	0.000
K	51.59	0.806	0.000

3.4 Unit root tests

It is widely known that panel-based unit root and co-integration tests perform better than the tests based on individual time series by including the additional information that comes from the presence of the cross sectional dimension. However the literature of panel unit root and co-integration tests differentiates as first and second generation tests where the first group developed on the assumption of the cross-sectional independence. As the cross-section independence is rejected in our study, we will implement second generation unit root tests which take into account that the variables can be represented by a common factor, along with five commonly used first generation panel unit root tests, namely Levin, Lin and Chu test (2002), Breitung test (2000), Im, Pesaran and Shin test (2003), ADF-Fisher test and Philips Perron — Fisher test. Levin, Lin and Chu (LLC) and Breitung tests assume a common unit root process across the relevant cross-sections while the other tests allow for an individual unit root process.

Table 4: First generation panel unit root test results (with trend)

Unit Root Test	**Y**	**ΔY**	**C**	**ΔC**	**E**	**ΔE**	**K**	
Levin, Lin & Chu t*	0.759	−12.287***	2.678	−16.611***	−0.563	−13.76***	−0.198	−12.
Breitung t-stat	3.913	11.542***	3.641	−6.021***	−1.030	−7.678***	0.744	−8.2
IPS W-stat	0.678	−12.841***	1.768	−15.517***	−0.616	−16.00***	−2.358***	−10.
ADF-Fisher Chi-square	27.02	197.55***	16.47	253.87***	33.92	263.65***	46.98**	148
PP-Fisher Chi-square	14.68	227.98***	17.26	302.07***	34.13	432.72***	15.72	135

As we can see from Table 3.4, first generation unit root tests fail to reject the null hypothesis of the existence of a unit root process for the levels of the variables.[10] But the failure of this rejection may be due to the presence of the cross section dependence or potential presence of structural breaks. Therefore applying second generation tests would provide more robust results.

To convey a panel unit root test with cross-sectional dependence, Pesaran (2007) considers a statistic which is constructed from the following Cross-Sectionally Augmented Dickey-Fuller (CADF) regression and estimating the OLS method for the ith cross-section in the panel:

[10] The Null hypothesis is: All individuals follow a unit root process. The choice of lag length for the Breitung, IPS and Fisher-ADF test are determined by Schwarz Information Criterion. The LLC and Fisher-PP tests were computed using the Bartlett kernel with automatic bandwidth selection. Probabilities for Fisher tests are computed using an asymptotic χ^2-distribution. All other tests assume asymptotic normality. The asterisks represent significance at the 10% (*), 5% (**), and 1% (***) confidence levels.

$$\Delta y_{it} = \alpha_i + \rho_t y_{i,t-1} + c_t \bar{y}_{t-1} + \sum_{j=0}^{k} d_{tj} \Delta \bar{y}_{t-j} + \sum_{j=0}^{k} \delta_{tj} \Delta \bar{y}_{i,t-j} + \varepsilon_{it} \tag{17}$$

where $y_{t-1} = \frac{1}{N} \sum_{i=1}^{N} y_{i,t-1}$. The CIPS statistic isbased on the average of individual CADF statistics:

$$CIPS = \frac{1}{N} \sum_{i=1}^{N} t_t(N, T).$$

where is the t-statistic of the estimate of in the above equation. The results are presented in Table 3.4. All variables are integrated of order one.[11]

Table 5: CIPS test results

Variables	**With trend**	**Without trend**
Y	4.709(1.000)	2.701(0.997)
E	0.476(0.683)	2.036(0.979)
C	0.651(0.742)	−1.701(0.044)
K	2.182(0.985)	1.239(0.892)

Since our variables are non-stationary but have unit root, we can proceed by testing whether they follow the same path in the long-run, in other words, whether they are co-integrated.

3.5 Co-integration tests

After defining the stationary level of the data, we can apply co-integration tests to investigate the existence of a long run relationship. Similar to the first generation unit root tests, the first generation panel co-integration tests may not be able to reject the null hypothesis as a result of omitting possible structural breaks and cross-sectional dependence. Both types of tests have been applied in this study for co-integration: Pedroni (1999) first generation tests and a second generation test proposed by Westerlund (2007). Pedroni proposes seven test statistics that can be distinguished in two types of residual based tests. Four tests are based on pooling the residuals of the regression along the within-dimension of the panel, while three are based on pooling the residuals along the between-dimension.

Table 3.5 shows Pedroni's co-integration tests results. All of the within- and between- dimension statistics indicate a strong and robust evidence of co-integration between our variables under scrutiny.[12]

We also perform the Westerlund (2007) co-integration test which delivers robust critical values through bootstrap approach even under the assumption of cross-section dependence. The test checks whether an error correction model has or not an error correction (individual group or full panel) based on the following equation:

[11] Null hypothesis: series are $I(1)$. p-test are in parenthesis.

[12] Null Hypothesis: No cointegration. Automatic lag length selection based on SIC with a max lag of 9. Newey-West automatic bandwidth selection and Bartlett kernel.

Table 6: Pedroni Residual Cointegration Test

Deterministic intercept and trend			**No deterministic intercept and trend**		
Alternative hypothesis: common AR coefs. (within-dimension)					
	Statistic	Prob.		Statistic	Prob.
Panel v-Statistic	2.716057	0.0033	Panel v-Statistic	6.801555	0.0000
Panel rho-Statistic	−6.842150	0.0000	Panel rho-Statistic	−9.435575	0.0000
Panel PP-Statistic	−18.61345	0.0000	Panel PP-Statistic	−14.51132	0.0000
Panel ADF-Statistic	−15.77301	0.0000	Panel ADF-Statistic	−13.23211	0.0000
Alternative hypothesis: individual AR coefs. (between-dimension)					
	Statistic	Prob.		Statistic	Prob.
Group rho-Statistic	−5.538399	0.0000	Group rho-Statistic	−10.06613	0.0000
Group PP-Statistic	−26.62993	0.0000	Group PP-Statistic	−24.58705	0.0000
Group ADF-Statistic	−17.06919	0.0000	Group ADF-Statistic	−18.39500	0.0000

$$\Delta Y_{it} = c_i+\alpha_i(Y_{it-1}-\beta_{1i}E_{it-1}-\beta_{2i}C_{it-1}-\beta_{3i}K_{it-1})+\sum_{j=1}^{pi}\alpha_{ij}\Delta Y_{i,t-1}+\sum_{j=1}^{pi}\gamma_{1ij}\Delta E_{i,t-j} + \sum_{j=0}^{pi}\gamma_{2ij}\Delta C_{i,t-j} + \sum_{j=1}^{pi}\gamma_{3ij}\Delta K_{i,t-j} + e_{it} \quad (18)$$

where α_i is the speed of adjustment term. If $\alpha_i = 0$, there is no error correction and the variables are not co-integrated. If $\alpha_i < 0$, the model is error correcting implying that the variables are co-integrated. Westerlund, developed four new panel co-integration tests without any common-factor restriction. P_t and P_a tests are designed to test the alternative hypothesis that the panel is co-integrated as a whole, whereas the two other test, G_t and G_a test whether at least one element in the panel is co-integrated.[13]

Table 7: Westerlund panel cointegration test results

	Value	Z **value**	p **value**	**Robust** p **value**
Gt	−4.234	−9.548	0.000	0.000
Ga	−35.266	−17.024	0.000	0.000
Pt	−16.989	−9.071	0.000	0.000
Pa	−34.030	−18.251	0.000	0.000

Westerlund test results strongly reject the null hypothesis of no co-integration. We therefore conclude that there exists a robust long-run relationship between growth and primary energy use.

[13]Null hypothesis: No co-integration.

3.6 Long run relationship estimation

Based on the existence of a long-term relationship we can make an estimation of this relationship in cooperation with the short-term dynamics by an error correction model (ECM). A general dynamic specification is represented by an auto-regressive distributed lag model of order p and q, ARDL$(p, q, q, , q)$:

$$y_{it} - \sum_{j=1}^{p} \lambda_{ij} y_{i,t-j} + \sum_{j=0}^{q} \delta'_{1j} c_{i,t-j} + \sum_{j=0}^{q} \delta'_{2j} e_{i,t-j} + \sum_{j=0}^{q} \delta'_{3j} k_{i,t-j} + \mu_i + \alpha_i t + \varepsilon_{it} \quad (19)$$

where t is a linear time trend and the general lag structure is meant to control for di?erent short-run output dynamics across countries. It can be re-written in the following error correction model form (Pesaran et al.(1999)):

$$\Delta y_{it} = \phi_i y_{i,t-1} + \beta'_1 c_{it} + \beta'_2 e_{it} + \beta'_3 k_{it} + \sum_{j=1}^{p-1} \lambda^*_{ij} \Delta y_{i,t-j} + \sum_{j=0}^{q-1} \delta^*_{1j}{}' \Delta c_{i,t-j}$$
$$+ \sum_{j=0}^{q-1} \delta^*_{2j}{}' \Delta e_{i,t-j} + \sum_{j=0}^{q-1} \delta^*_{3j}{}' \Delta k_{i,t-j} + \mu_i + \alpha_i t + \varepsilon_i \quad (20)$$
$$i = 1, 2, \ldots, N \text{ and } t = 1, 2, \ldots, T$$

with

$$\phi_i = -(1 - \sum_{j=1}^{p} \lambda_{ij})$$
$$\beta_i = \sum_{j=0}^{q} \delta_{ij}$$
$$\lambda^*_{ij} = - \sum_{m=j+1}^{p} \lambda_{im}, \quad j = 1, 2, \ldots, p-1$$
$$\delta^*_{ij} = - \sum_{m=j+1}^{q} \delta_{im}, \quad j = 1, 2, \ldots, q-1$$

where the error correction speed of adjustment parameter, ϕ_i, and long run coefficients,β_i, are of primary interest. The long-run coefficient δ incorporates short-run information, is an unobserved country-specific effect and ε_{it} is the error term. When the ARDL$(p, q, q, , q)$ is stable (i.e., error correcting), the adjustment coefficient is negative and less than 1 in absolute value. In this case, the long-run relationship is defined by:

$$y_{it} = -\frac{\beta_i}{\phi_i} x_{it} + \eta_{it}.$$

where (x_{it}) is a vector of explanatory variables, and η_{it} is a stationary process. In the steady-state, (x$_{it}$) and (y_{it}) are related to each other, with a long-term elasticity of $-\beta_i/\phi_i$. An important assumption for the consistency of the ARDL model is that the resulting residual of the error-correction model be serially uncorrelated and the explanatory variables can be treated as exogenous. PMG estimator constrains the long-run elasticities to be equal across all panels. This pooling across countries yields efficient and consistent estimates when the restrictions are true. Often, however, the hypothesis of slope homogeneity is rejected empirically. If the true model is heterogeneous, the PMG estimates are inconsistent; the MG estimates are consistent in either case. The test of difference in these models is performed with the familiar Hausman test. For comparison purposes we present three different panel estimators, including two estimators without cross-sectional dependence (PMG and MG) and the other one with cross sectional dependence (CCE-MG).

Table 8: Selection of the estimation method

Model	**PMG**	**MG**	**CCE-MG**
Dependent variable: $\mathbf{\Delta Y_{it}}$			
Energy consumption per capita	0.6740	0.6075	0.6102
(C_{it})	(0.062)***	(0.126)***	(0.165)***
Energy efficiency	0.6036	0.5833	0.4864
($E_{i,t-1}$)	(0.070)***	(0.279)**	(0.183)***
Capital formation per capita	0.1244	0.1972	0.1643
(K_{it})	(0.022)***	(0.082)***	(0.036)***
Convergence coefficient	-0.6230	-1.6958	-0.9554
($\mathbf{Y_{i,t-1}}$)	(0.124)***	(0.814)**	(0.189)***
Hausman test p value	0.0419		

The results of the error correction model for long-run estimates are reported in Table 3.6.[14] As noted previously, the estimated coefficients are elasticity estimates. A long-run equilibrium (co-integrated) relationship exists, implying meaningful long-run estimates. The estimated error correction coefficients are negative and highly significant indicating that the system moves toward equilibrium. Moving from MG to PMG (i.e. imposing long-run homogeneity) reduces the standard errors and reduces significantly the measured speed of convergence. This restriction cannot be rejected at the 1% level by the Hausman test statistics. Hence, the PMG estimators are consistent and more efficient than the MG estimators. However our results are not likely to vary significantly with respect to the estimation method. Estimated long-run elasticities of energy consumption, energy efficiency and capital formation per capita are positive and statistically significant.

The reader noticed, of course, that the sum of our two estimated output elasticities is greater than 1. Adding the elasticity of capital (Protocols 2 and 3) even leads to a total return to scale about 1.5. This suggests that global returns to scale with respect to energy use, energy efficiency and capital are strictly increasing. It will come as a surprise to some mainstream economists who are used to think of production as being characterized by constant returns to scale. Notice, first, that, given the finiteness of resources, *some* non-convexity of the production sector (i.e., increasing returns to scale) must be present, if our economies are to experience any long-standing growth as they did in the past.[15] Moreover, depending upon the way it

[14] Standard errors are given in parentheses. The lag structure is ARDL$(1, 1, 2, 1)$.

[15] Hurwicz and Reiter (1973) proved, indeed, within a general equilibrium setting, that the phase

is interpreted, energy efficiency may well be a (major) component of the Total Factor Productivity. That conventional macroeconomic estimations of factor productivity invariably suggest constant returns to scale comes from a well-known accounting artifact —which has been documented as early as in Samuelson (1979). Consequently, standard estimations tell us little about "real" returns to scale. On the other hand, at the microeconomic level, increasing returns to scale are consistent with empirical verifications of "every day experience" (see, e.g., Blinder (1988)). [16]

4 Dynamic Panel Granger-Causality

It is now understood that in the absence of cointegration between the variables a Granger causality test on a VAR in levels is invalid. Ohanian (1988) and Toda and Phillips (1993) showed that the distribution of the test statistic for Granger causality in a VAR with nonstationary variables is not the standard chi-square distribution. This means that the significance levels reported in the 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 χ^2-distributions apply.

Cointegration tests can be used to test for omitted nonstationary variables. A lack of cointegration implies that variables essential to cointegration are omitted from the model. Therefore, testing for cointegration is still a necessary prerequisite to causality testing on data with potential unit roots.

Given the co-integration relationship between variables, we then examine the causality between variables using PMG estimator. The error-correction model to be estimated is given by following equations:

$$\begin{aligned}\Delta y_{it} = \phi_i y_{i,t-1} + \beta_1' c_{it} + \beta_2' e_{it} + \beta_3' k_{it} + \sum_{j=1}^{p-1} \lambda_{ij}^* \Delta y_{i,t-j} + \sum_{j=0}^{q-1} {\delta_{1j}^*}' \Delta c_{i,t-j} \\ + \sum_{j=0}^{q-1} {\delta_{2j}^*}' \Delta e_{i,t-j} + \sum_{j=0}^{q-1} {\delta_{3j}^*}' \Delta k_{i,t-j} + \mu_i + \varepsilon_{it} \quad (21)\end{aligned}$$

$$\begin{aligned}\Delta c_{it} = \omega_i c_{i,t-1} + \alpha_1' y_{it} + \alpha_2' e_{it} + \alpha_3' k_{it} + \sum_{j=1}^{p-1} \vartheta_{ij}^* \Delta c_{i,t-j} + \sum_{j=0}^{q-1} {\gamma_{1j}^*}' \Delta y_{i,t-j} \\ + \sum_{j=0}^{q-1} {\gamma_{2j}^*}' \Delta e_{i,t-j} + \sum_{j=0}^{q-1} {\gamma_{3j}^*}' \Delta k_{i,t-j} + \mu_i' + \varepsilon_{it}' \quad (22)\end{aligned}$$

space of economic growth must be compact unless the production sector exhibits some form of non-convexity. This result is independent from the neoclassical treatment of economic dynamics.

[16] Nor does the presence of increasing returns to scale contradict the second law of thermodynamics, since neither GDP nor efficiency are expressed in physical units (but in monetary and mixed units respectively): Polluting a river increases the GDP *and* entropy together !

$$\Delta e_{it} = \varphi_i e_{i,t-1} + \sigma_1' y_{it} + \sigma_2' c_{it} + \sigma_3' k_{it} + \sum_{j=1}^{p-1} \partial_{ij}^* \Delta e_{i,t-j} + \sum_{j=0}^{q-1} {\tau_{1j}^*}' \Delta y_{i,t-j} + \sum_{j=0}^{q-1} {\tau_{2j}^*}' \Delta c_{i,t-j} + \sum_{j=0}^{q-1} {\tau_{3j}^*}' \Delta k_{i,t-j} + \mu_i'' + \varepsilon_{it}'' \quad (23)$$

$$\Delta k_{it} = \chi_i k_{i,t-1} + \eta_1' y_{it} + \eta_2' c_{it} + \eta_3' e_{it} + \sum_{j=1}^{p-1} \phi_{ij}^* \Delta k_{i,t-j} + \sum_{j=0}^{q-1} {\rho_{1j}^*}' \Delta y_{i,t-j} + \sum_{j=0}^{q-1} {\rho_{2j}^*}' \Delta c_{i,t-j} + \sum_{j=0}^{q-1} {\rho_{3j}^*}' \Delta e_{i,t-j} + \mu_i''' + \varepsilon_{it}''' \quad (24)$$

The PMG estimator allows determining the direction of causality by testing the significance of the coefficients of related variables. Short-run causality can be determined by the statistical significance of the coefficients of each explanatory variable. Long-run causality can be examined by testing the significance of speed of adjustment. All the null hypotheses can be tested using standard Wald restriction test with a χ^2-distribution.[17]

Table 9: Panel causality test results

Dependent Variable	**Sources of causation (independent variables)**				
	Short run				*Long run*
	ΔY	**ΔE**	**ΔC**	**ΔK**	**ECT**
ΔY	–	1.26	23.03***	190.86***	−0.325***
ΔE	1228.4***	–	8702.6***	9.53***	−0.959***
ΔC	0.41	0.16	–	0.44	−0.628
ΔK	47.59***	2.35	39.13***	–	−0.158***

The short-run and long-run causality tests reveal several interesting results. First, the results show that energy consumption is exogenous to the other variables in the model. There is an umambiguous unidirectional causality from energy consumption to economic growth in both the short and long-run. Energy consumption also indirectly effects the economy growth through its positive impact on capital formation.

Sensitivity analysis: Robustness of results

1) Sub-Samples:

It could be argued that, one individual country could significantly affect the estimated parameters, even when the Haussman tests do not reject the hypothesis of common long-run coefficients. A sensitivity analysis is thus performed in order to assess the robustness of results to variation of country coverage, by eliminating one country at a time from the original sample and re-estimating the PMG procedure.

[17] Wald χ^2-test statistics for short-run causality. The lag length is one. ECT represents the coefficient of the error-correction terms.

The estimated coefficients are shown in Figures 4a, 4b, 4c after arranging the estimates in decreasing order across sub- samples. Although the width of confidence intervals is somewhat affected for Netherlands, for all estimated coefficients, the sample composition does not make a significant difference (all long-run coefficients remain statistically significant at the 1 per cent level).

2) Lag Structures:

We have also conducted a sensitivity analysis of the PMG results to changes in the lag structure of the dependent and independent variables by re-estimating the regression with different ARDL specifications, imposing a maximum lag order of 3 in order to maintain a reasonable number of degrees of freedom. Among the possible combinations of lags for the four variables, we have adopted criteria of keeping the number of lags for energy efficiency being more or equal that for the other variables. Figure 5a 5b 5c shows the results for this specification with the different estimation procedures.

PMG estimates of long-run coefficients do not seem to be strongly affected by the choice of the lag structures.

3) Sub sample stability:

Another sensitivity analysis is performed in order to assess the robustness of results to variation of the time period, by eliminating five-year period at a time from the original sample and re-estimating the PMG procedure. We evaluated the 1975- 2011, 1980 — 2011, 1985 -2011 and 1990 - 2011 sub-periods. The estimated coefficients are shown in Table 9.

References

Kraft, J. and Kraft, A. (1978). On the relationship between energy and GNP. *Journal of Energy and Development*, 3:401–403. 4