INSEAD

The Business School for the World®

Social Innovation Centre

Faculty & Research Working Paper

The Underestimated Contribution of Energy to Economic Growth

Robert U. AYRES
Jeroen C.J.M. VAN DEN BERGH
Dietmar LINDENBERGER
Benjamin S. WARR
2013/97/TOM/EPS/ISIC

The Underestimated Contribution of Energy to Economic Growth

Robert U. Ayres*

Jeroen C.J.M. van den Bergh**

Dietmar Lindenberger***

Benjamin Warr****

Acknowledgement. We thank Reiner Kümmel and two anonymous referees for valuable inputs to this paper. The usual disclaimer applies.

* Emeritus Professor of Economics and Political Science and Technology Management, The Novartis Chair in Management and the Environment, Emeritus at INSEAD, Boulevard de Constance 77305 Fontainebleau, France. Email: Robert.ayres@insead.edu

** ICREA Research Professor at Institute of Environmental Science and Technology (ICTA) Universitat Autònoma de Barcelona Edifici Cn - Campus UAB 08193 Bellaterra (Cerdanyola) Spain. Email: jeroen.bergh@uab.es

*** Director Applied Research at the Institute of Energy Economics, Cologne University, D-50923 Cologne, Germany, (Corresponding Author) Email: dietmar.lindenberger@uni-koeln.de

***** Executive in Residence, INSEAD Social Innovation Centre at INSEAD, Boulevard de Constance, 77305Fontainebleau, France. Email: benjamin.warr@insead.edu

Abstract

Standard economic theory regards capital and labour as the main factors of production that satisfy the "cost-share theorem". This paper argues that when a third factor, namely energy, is added physical constraints on substitution among the factors arise. We show that energy is a much more important factor of production than its small cost share may indicate. This implies that continued economic growth along the historical trend cannot safely be assumed, notably in view of considerably higher energy prices in the future due to peak oil and climate policy.

Keywords: Economic growth; economic theory; energy cost share; technological constraints; peak oil; climate policy

JEL codes: D24, O44, Q43

1. INTRODUCTION

Virtually all of modern economic growth theory assumes that GDP growth per capita is driven by technological progress and capital investment, including knowledge investment (Romer 1994; Aghion and Howitt 1998; Barro and Sala-i-Martin 2003). The current terminology is "technology-enhanced labour productivity". Assuming that capital and knowledge stocks accumulate over time, most growth economists as well as politicians take for granted that incomes will increase on average at 2-2.5 % p.a. whence our great-grandchildren 100 years from now would be up to 10-times richer than we are today. The standard theory of growth does not take into account energy availability or prices and cannot explain economic declines except as a consequence of reduced labour hours, which usually are a consequence, not a cause, of the decline.

Nonetheless, economic declines (recessions) are a fact of economic life. Most recessions, including the present one, have evidently been caused by a combination of factors, such as an oil price spike, hyper-inflation or the collapse of "bubbles" (Kindleberger 1989). Given that the "Dot-Com" bubble of 1998-2000 and the real estate bubble of 2003-2007 were not predicted, or even recognized, by most economists until after the collapse, how much credence can we give to another forecast of smooth growth independent of energy availability? Can we safely assume there are no limits to growth and that our grandchildren will be a lot richer than we are? We argue, on the contrary, that it would be prudent to assume that long-term growth, notably in industrialized countries, may not continue along the historical "track" but may be much slower than in the past.

Various types of evidence support the view that energy is important in economic growth and there is no substitute for energy, though some forms of energy can be replaced by others (e.g. wind power for nuclear power). Stern (2011) provides evidence that substitution between energy carriers has been an important driver of growth in the past but that such substitution has reached its limit. Many econometric studies of aggregate production functions find complementarity or weak substitutability between energy and capital for OECD countries (Berndt and Wood 1979; Koetse et al., 2008; Fiorito and van den Bergh, 2011). Further support for this comes from so-called biophysical approaches, including system models that give much attention to direct and indirect energy use (Giampietro et al., 2011).

It is fair to mention that there is a branch of growth theory that includes environmental and resource variables (an overview is Smulder 2005). However, this has not affected the core of growth theory and associated policy debate. Moreover, most growth models with resources exclude realistic constraints on the substitution possibilities between energy and capital. Notable exceptions are D'Arge and Kogiku (1973), Kümmel (1982, 2002), Gross and Veendorp (1990), van den Bergh and Nijkamp (1994), Lindenberger and Kümmel (2011). All of these arrive at the conclusion that future growth will be severely hampered by scarce energy and material resources.

Although natural resource prices have fallen more or less continuously since the industrial revolution (Barnett and Morse 1963; Smith and Krutilla 1979), there is no guarantee that discovery and technological progress will always compensate for geological scarcity and keep resource prices falling in the future. The price of oil, in particular, is unlikely to fall in the future, regardless of progress in off-shore drilling or extraction from tar sands or shale. Geological scarcity of some resources (notably rare metals available only as by-products of others) cannot always be overcome by increasing prices. Dramatic decreases in the energy return on energy investment (EROEI) since the 1930s suggest that further declines are inevitable as old oilfields are exhausted and drilling moves into ever-deeper offshore waters, and under the Arctic ice (Hall, Cleveland, and Kaufmann 1986). In addition, recent studies indicate that prices of resources change inversely with EROEI. As EROEI decreases for depleting fossil fuel production, the corresponding energy prices increase (King and Hall, 2011). All this implies that the resource extraction sector and all contributing activities (production of equipment, transport) must become progressively larger to compensate for this decline, thus “crowding out” other economic activities. By the way, when the EROEI declines to unity, the resource ceases to exist, as such. The changing geopolitical implications of these trends are likely to have significant economic implications, not excluding resource-wars.

Here we will offer additional theoretical and empirical arguments for the view that energy will constrain future economic growth. In particular we will argue that a production function with energy as third factor besides capital and labor is better capable of tracing the pattern of production relations over time. This will then lead to the insight that energy is more important than indicated by its cost share. The policy implication of this is that the likely future scenario of peak oil and stringent climate policy through carbon pricing will cause much pressure on economic growth, and may well mean that past rates of growth are not feasible at all in the future. This argument has received some attention in environmental science and energy studies, but is neglected in economic journals. Our insights suggest that economists should take this argument more seriously and respond to it adequately in a theoretical and empirical sense.

The organization of the remainder of this paper is as follows. Section 2 discusses the neglect of energy in growth studies. Section 3 explains the core problem, namely the misinterpretation of cost shares. Section 4 proposes an alternative production function approach. Section 5 shows that the cost-share theorem no longer holds, if technological constraints on factors are taken into account. Using this insight, Section 6 examines the impact of increasing energy scarcity and higher energy prices on economic growth. Section 7 concludes and offers policy-relevant insights relating to energy efficiency and energy-climate policy.

2. ENERGY: THE NEGLECTED FACTOR OF PRODUCTION

In standard economic growth theory following Solow, the crucial assumption was that there are only two factors of production that matter, capital and labour. This assumption can be traced to an identity, a stylized fact and an equilibrium condition arising from an old income allocation theorem. The identity is that the GDP is defined as the sum total of payments to capital (interest, dividends, rents and royalties) and payments to labour (wages, salaries). The stylized fact is that the cost shares of capital (30%) and labour (70%) in the US GDP have remained virtually constant for a long time. The income allocation (cost-share) theorem says that the output elasticity of each factor of production must be proportional to its cost share. Since primary energy still accounts for a very small fraction of total factor cost – even after oil price rises –

influential theorists have argued that energy cannot be an important source of productivity (Denison 1979). Moreover, the combination of the cost-share theorem and historically nearly constant cost shares seems to justify the simple standard Cobb-Douglas production function.

There is good reason to doubt that past GDP growth per capita is entirely explained by capital accumulation or non-specific knowledge accumulation, as most growth theorists seem to believe. In the first place, those factors of production rarely, if ever, decline. More important, the national accounts do not reflect payments to a category called "energy" (actually exergy[1]). It is obvious that neither labour nor capital can function without inputs of energy, either as food or animal feed, or as fuel for engines, or electric power for light, communication and appliances of all sorts. In other words, a flow of energy capable of doing work, in some form, is just as essential for economic output (production) as capital or labour. It should, logically, be regarded as a factor of production, along with labour and capital.

Yet energy has been largely ignored by economic theory, since the first attempts at quantification by the French physiocrats in the eighteenth century. They regarded agriculture as the basis of every economy, and agriculture was – for them – a function of land and labour (including labour by working animals). The energy inputs to the land by sun and rain were never considered separately. Their value was hidden in the value of arable land. Similarly, the value of the solar energy embodied in fossil fuels is implicitly assumed to be reflected in the cost of labour to cut the wood or dig the coal, and thus in the rents accruing to mine owners and, more recently, oil companies. In the past, these energy-related rents have been both hidden and very small, considered as a fraction of GDP, because the energy resources in question were abundant.

The neglect of energy as a factor has continued, even though coal and later oil and natural gas became increasingly important and even dominant inputs to economic activity during the past two centuries. Attempts to treat energy as an explicit factor of production next

[1] Exergy is defined as available energy, meaning energy that can do physical work. More precisely, it is the maximum work that can be done by a system reversibly approaching thermodynamic equilibrium. It is what most people mean when they speak of energy. Whereas energy is conserved and cannot be used up, exergy is destroyed by transformation processes (as entropy increases).

to capital and labour in quantitative analysis (Hudson and Jorgenson 1974; Allen and et al 1976; Jorgenson 1978) were a response to economic concerns raised by the Arab oil embargo and the accompanying "energy crisis" of 1973-74 and again by the Iranian Revolution in 1979-1980. In both cases associated oil price spikes were followed immediately by deep recessions. In fact, other lesser oil price spikes since WW II have also been followed by recessions (Hamilton 2003, 2005). In response to this apparent relationship, several economists introduced the KLEM production function, where K refers to capital, L to labour, E to energy and M to materials (Hudson and Jorgenson 1974; Jorgenson 1978, 1984; Berndt and Wood 1979). But the measurement difficulties were not resolved at the time, or since. Moreover, several critics, notably Denison (1979), the "dean" of growth accountants, argued that energy prices could not have a significant impact on GDP because the cost share of energy in the national accounts was so small – only 4 or 5 percent for most OECD countries at the time. To explain the recessions, many economists purported to find other causes, such as "tight money".

The energy return on energy invested (EROEI) is the ratio of the energy supplied by a process to the energy used directly and indirectly in that process. It is a useful indicator, since the cost of extraction – due to labour and capital inputs from the rest of the economy – is proportional to the energy required for mining, drilling, pumping, and distribution (Hall, Cleveland, and Kaufmann 1986). A declining EROEI means more and more primary energy is needed as time goes on just to service the rest of the economy. A declining EROEI therefore almost certainly portends rising energy prices (Hall and Klitgaard, 2012). The rising cost trend is not monotonic because of occasional significant discoveries on the one hand, and fluctuations in demand, exaggerated by geopolitics and speculation, on the other hand. But discoveries and technological improvements are less and less likely to compensate for the underlying geological reality.[2]

[2] One might assume that new supplies will smoothly and automatically appear as prices rise. But that assumption depends implicitly on another assumption about the distribution of resources in the earth's crust, namely that it is log-normal, as would be the case if resulting from random processes. If, in fact, the distribution function has a double peak (for instance) higher prices may not elicit greater output. This behaviour is characteristic of rare metals that are found as co-products or by-products of other metals (Skinner 1976; Hagelueken 2010).

What this means is that the 4-8 percent energy cost share that was typical of the US and the OECD in recent years is likely to increase significantly, whether the energy is obtained from domestic sources or imported. Energy costs will rise as output from old fields declines (at 7% per annum, on average) and new discoveries are increasingly remote, of poorer quality and more costly to exploit. The EROEI will continue to decline and the cost share for energy will increase in consequence. In addition, declining EROEI is typically associated with increasing capital cost and technology spending in the energy sector. This is consistent with a recent approach developing a thermodynamically founded production theory that takes into account fixed cost (Chen, 2005). When energy sources become scarcer, and capital and technological costs increase, more financial resources are consumed which crowds out other economic activity (Chen and Galbraith, 2011). We argue below that the economic impact is likely to be greater than standard theory predicts.

3. THE COST SHARE CONFUSION

The Denison critique, mentioned above, was widely accepted. But it was based on a drastically oversimplified model of the economy: namely an economy consisting of many small profit maximizing firms producing a single product serving both as consumption and capital good. In his Macroeconomics text-book, Mankiw (1997) imagines an economy consisting of a large number of small bakeries producing bread from rented capital and rented labour, but without any need for energy, flour and yeast. In such an economy, it is an easy textbook exercise to show that each input factor (capital and labour) will be used in proportion to its marginal productivity. In this simple model economy it follows that the output elasticity of capital and labour, respectively, will be exactly equal to its cost share. Even if a third factor (say, energy) would be included, *but without any constraints on the allowed combinations of factors*, the Euler-Lagrange optimization calculus would yield an extremum where the output elasticities of the factors are equal to their cost shares. We call this the “cost-share theorem”.

The logic of the simple “bakery model” might have been challenged long ago, except for the fact that the cost shares of capital and labour for the US economy have been relatively constant at approximately 0.3 and 0.7 respectively throughout the twentieth century. This empirical observation has seemed like confirmation both of the over-simplified model-economy used to derive the cost-share theorem, and the use of the C-D function in economic forecasting models.

However the two-factor single-good bakery model is clearly unrealistic. When the economy is considered as a multi-product, multi-sector, multi-factor system, it becomes obvious – even without mathematical proof – that the impact of a cut in one essential input can have a much bigger effect on the whole economy than its cost share. Fresh water may account for a very small fraction of the GDP, but without it the food production system would fail. Food production may be only 4% of the economy, but without food labour cannot function and there can be no economy. Schelling's famous statement that a 50% decline in agricultural output would result in only a 2% decline in GDP is simply wrong (Schelling 2002). In the real multi-sector economy, a cut in the output of the agricultural sector would also be felt by downstream food and beverage processors, truckers, wholesalers, retailers, hotels and restaurants, not to mention consumers. There are no substitutes for food, so the idea that downstream users would buy "something else" instead makes no sense. The overall impact of a 50 percent drop in agricultural output would be a lot greater than 2 percent of the GDP. This is also, because, as the history of famine teaches, food scarcity would result in sharply rising prices of agricultural products so that the relative contribution of agriculture to GDP would rise correspondingly.

Raw material input costs apply only to the primary extractive sectors, even though value is added by a sequence of downstream sectors. If an essential input to the primary sector is absent, and if there is no substitute, the whole system must fail. More generally, materials composition is limited, that is, substitution between materials is subject to hard technical constraints, which affects production relationships (van den Bergh, 1999).

Much the same argument applies to energy. A few years ago petroleum inputs to the economy also accounted for around 4 percent of the US GDP. Schelling's argument, as applied to agriculture would apply equally to oil. Suppose oil inputs to the economy were cut by a factor of two, as in Schelling's agricultural example. The overall impact in that hypothetical case would also be 2 percent of US GDP. But in the real, multi-sector economy, a 50 percent cut in crude oil supplies would also cut oil refinery output and petrochemical output. Car, truck and air transportation would also be cut by virtually the same amount, because there is no immediate substitute for liquid hydrocarbon fuels, and all other sectors depending on non-electrified transportation services would be affected to a similar degree. The overall impact on the economy, as captured by a good general equilibrium model would be far greater than 2 percent. The multiplier effect in this case would probably be closer to a factor of ten, if not more. It follows that *the output elasticity of petroleum must be far greater than its small cost share* (Ayres and Warr 2009).

It turns out that the traditional cost share theorem, as taught in textbooks, is not even true for a single sector economy if there are constraints on input factor combinations. In the overwhelming majority of economic models, notably of economic growth, labour and capital are assumed to be easily substitutable (in a Cobb-Douglas or CES function sense), which means that the output could be generated by labour with very little capital, or by capital with very little labour. But, in reality, the range of short-term factor substitutability is fairly narrow. There is an optimal operating point for the economy and an optimal combination of capital and labour: too much labour, or too little, will result in under-utilization of labour or under-utilization of machines (capital), respectively. CES captures this better than Cobb-Douglas but still not well, and as we will see below another class of production functions is required to provide a more realistic picture.

This is equally true in the three-factor case. We know that capital would be totally unproductive without a flow of energy (actually exergy) to drive the machines and heat or cool the factories and office buildings, which protect the machines and the people (labour) who handle them. For instance, in the economy-as-a-set-of bakeries case, we postulate a need for fuel gas for the bakery ovens. And suppose we insist that there is a certain fixed requirement of gas fuel by each oven. If the flow of fuel is too great it will be wasted (or even harmful – the baker will be poisoned or the bread will be burnt). On the other hand, for given output (bread) there is a certain potential of substituting capital for energy by increasing the energy efficiency of machinery (the oven), albeit within technical and thermodynamic limits with respect to the increase of conversion efficiencies and reduction of heat losses. Again, this implies limited substitutability.

To make it more precise and complicated, one could talk of hard and soft constraints. The hard ones are technological constraints which require that capacity utilization and automation – given the state of technology – are incomplete. Soft constraints are of a social, organizational, financial and legal nature. They relate to such varied issues as labour, health and environmental regulations, pressure on management exerted by labour unions, and vested interests that influence internal firm politics. Financial constraints depend on the fact that economic restructuring of the capital stock takes time, which effectively limits the range in factor space available to profitable substitution. All these constraints vary over time which contributes to the variability of substitution opportunities. Calculating these opportunities with a simplified production model that excludes the various constraints or assumes a CES functional form is perhaps nice as a first toy model in an introductory economics textbook, but as we will see later it cannot serve as the basis for credible statements about the relationships between production factors, and especially the significance of the factor energy for economic growth.

4. SPECIFYING THE PRODUCTION FUNCTION

The next step is to specify an appropriate production function. Although not all economists accept the idea of production functions (Klein 1946; Robinson 1971; Harcourt 1972; Felipe and Fisher 2003), the concept is central to modern growth theory since Solow (Solow 1956, 1957). However, if the output elasticities of the factors cannot be assumed to be equal to the respective cost shares, due to technological and other constraints (as discussed below), it follows that one cannot parameterise the production function, such as simply choosing exponents for capital, labour, and energy adding up to one as in a constant-returns-to-scale Cobb-Douglas function. To make a long story short, if the output elasticity of each of the three factors K,L and E depends not only on the cost share of the factor but also on shadow prices of constraints that limit the range of substitution, a different form of production function and a different method of parameterisation are needed.

The first derivation of a production function that does not presume equality of output elasticities and cost shares and takes into account substitution possibilities specific to the (asymmetric) relation between the factor energy and the traditional factors capital and labor appeared in Kümmel et al. (Kümmel 1982; Kümmel et al. 1985; Kümmel, Henn, and Lindenberger 2002). They pointed out that the twice differentiability conditions on macroeconomic production functions result in a set of partial differential equations for the output elasticities, and computed a number of approximate solutions. It is easily shown that the simplest solution with no constraints on the variables is the Cobb-Douglas function. But as noted already, that solution requires constant (over time) output elasticities, whereas we want to allow for the possibility of output elasticities that depend on the three variables and that can change over time. The result is the so-called Linear-Exponential (LINEX) production function. Kümmel (1982) suggested that a mathematical expression for the output elasticities with appropriate asymptotic behaviour can be introduced and that the parameters of the expressions have clear economic interpretations. One parameter can be interpreted as the capital effectiveness and the other one as the energy demand of the fully utilized capital stock. The inverse of the latter measure is energy efficiency.[3]

The LINEX function takes into account that production possibilities, as a limiting case, include a state of total automation. Thus it is especially suited for the modelling of industrial production systems. Based on the same methodology, i.e. by specifying output elasticities that satisfy appropriate asymptotic technological boundary conditions, production functions for the service sector – with limited automation potential, through deployment of IT technology in trade, banking, insurance, or public administration – have also been derived (Lindenberger 2003).

[3] Just like the translog production function, the Linex functions belong to the category of variable elasticities of substitution (VES) functions. The relation of linex to translog functions is discussed in Kümmel et al. (1985), and the linex VES is discussed in Lindenberger and Kümmel (2011).

A modified version of the LINEX production function, replacing energy as such by "useful work" (done by energy) was introduced by Ayres and Warr (2005, 2009). Their logic is that total energy consists of both useful and non-useful (including waste) components, of which only the useful component is productive. Moreover, with this formulation the efficiency of converting energy inputs into useful work is a key driver of growth.

Regarding growth and automation, it is obvious that installed capital equipment has been more and more automated as time passes, and that this automation process involves an increasing amount of energy (directly and indirectly) and capital, and – by definition – less labour inputs. But the replacement of obsolescent physical capital and the addition of new capital equipment are limited not only by the state of technology but also by the availability of finance. Hence it is plausible that, if the capital stock at a given time were based on the best available technology, the economy would be more productive than it is, while it would use more energy and capital (and less labour) to produce the same output. But that hypothetical economy would be different from the present, existing economy. Besides technological constraints, this is another element supporting our hypothesis that the output elasticity of energy is greater than its cost share.

In spite of these considerations, most economists reject the argument that output elasticity for energy need not be equal to its cost share. Why is this so? The standard response seems to be that, if the output elasticity (marginal productivity) of energy is really much greater than its cost share (price), it follows that the economy would be more productive if more energy (and less labour) were employed. In other words, it seems to follow that *the economy is using too little energy, not too much*.

At first glance, this argument seems hard to refute, especially in view of empirical evidence implying that the economy is using more energy than necessary (IEA, 2011). The counter-argument is quite simple, however. In the first place, there is also evidence that the economy utilizes more labour than would be optimal, as suggested by the fact that payrolls increase much more slowly than corporate expenditures on cost-saving hardware and software. In the second place, if there were no constraints on energy and labour requirements for the existing capital stock, it would mean that there is no limit to substitutability among the factors. In that case energy could completely replace human labour, or capital or labour could completely replace labour or capital. These implications are unrealistic and can be rejected a priori. Again, it follows that the range of substitutability must be limited.

Summing up, there are technological constraints that limit the substitution between the production factors, especially if energy is taken into account as third factor besides capital and labour. “Hard” constraints result from two facts. First, the degree of capacity utilization is always less or equal to one, because one cannot feed more energy into the machines than technically possible at their full utilization. Second, the degree of automation cannot exceed the technologically possible degree of automation, whose maximum value is again unity. In addition there are “soft” constraints pertaining to social, financial, organisational, or legal restrictions, which also limit substitution possibilities over time. For example, social laws and regulations or constrained financial liquidity may limit the speed at which investments in increasingly automated production facilities are realized. The “hard” constraints are in an approximate way taken into account by linex production function.[4] The “soft” constraints also matter. Since empirical studies show that the output elasticity of energy is typically far above its cost share, constraints seems to be binding empirically. Associated shadow prices drive a wedge between output elasticities and factor cost shares (equivalently between factor productivities and prices). This is formally illustrated in the the next section.

5. FORMAL ILLUSTRATION OF THE IRRELEVANCE OF THE COST SHARE THEOREM

Going from qualitative considerations to quantitative mathematical analysis, Kümmel, Ayres and Lindenberger (2010) showed, by optimization of profit or time-integrated utility subject to technological constraints, that the usual equilibrium conditions must be replaced by conditions where shadow prices, due to technological constraints, add to factor prices. This destroys the

[4] Implemented through asymptotic properties of output elasticities and additional constraints, see e.g. Kümmel et al (1985), Lindenberger and Kümmel (2011).

cost-share theorem, which states the equality of output elasticities and cost shares. Empirical studies show that the output elasticity of energy is far above its cost share (see Section 6). If firm behaviour is cost-minimizing, there must be at least one binding constraint among the set of hard (technological) and soft (other) constraints that exist in the real world.

The output Y of the physical basis is produced by the factors capital K, labour L, and energy (exergy) E, which perform physical work and process information. Entrepreneurial decisions, aiming at producing a certain quantity of output Y within the technology that exists at a given time t, determine the absolute magnitude of the total capital stock, its degree of capacity utilization, and its degree of automation. The capital stock consists of all energy-converting and information-processing machines together with all buildings and installations necessary for their protection and operation. Machines are designed and built for specific energy inputs and require a certain amount of labour for handling, supervision, and maintenance. The quantities of labour and energy that are combined with the capital stock of a fixed degree of automation determine the degree of capacity utilization. The degree of automation at time t is represented by the ratio of the actual capital stock to the capital stock that would be required in order to produce the actual output of goods with the actual technology in the state of maximum possible automation. This state is characterized by a combination of capital and energy such that adding one more unskilled worker adds virtually nothing to gross economic output so that the output elasticity of routine labour would be vanishingly small. In some manufacturing sectors of industrialized countries this point actually does not seem to be far away.

It is obvious from an engineering point of view and by definition that both the degree of capacity utilization and the degree of automation are functions of capital, labour and energy, while each degree cannot exceed the number 1,0. In other words, a production system cannot operate above design capacity, and the maximum degree of automation cannot be exceeded. These are two fundamental technological constraints on the combinations of capital, labour and energy in modern economies. They drastically change the conditions for economic equilibrium that result from the behavioural assumptions of standard economics. One such assumption is profit maximization, according to which the actions of all economic agents are supposed to move the economy into a point of factor space where the difference between output and total factor cost is maximum. Consider an economic system that produces output Y with three factors of production X_1, X_2, X_3, whose combinations are subject to technological constraints, labelled by the index a and expressed by the equations $f_a(X_1,X_2,X_3,t) = 0$ with the help of slack variables. There are two of such constraint equations. Their explicit forms have been calculated in Kümmel et al. (2010). Then profit maximization under constant returns to scale results in the following three equilibrium conditions:

$$\varepsilon_i = \frac{X_i}{Y}\frac{\partial Y}{\partial X_i} = \frac{X_i(p_i + s_i)}{\sum_{i=1}^{3} X_i(p_i + s_i)} \qquad i = 1,2,3. \tag{1}$$

These conditions relate the output elasticities ε_i of factors X_i to market prices p_i per factor unit and the factor shadow prices

$$s_i = -\sum_a \frac{\mu_a}{\mu}\frac{\partial f_a}{\partial X_i} \qquad i = 1,2,3. \tag{2}$$

Here, μ_a and μ are the Lagrange multipliers of the technological and the fixed-cost constraint equations in the optimization calculus. Thus, the output elasticities in eq. (1) are equal to "shadowed" cost shares.[5]

If there were no technological constraints, all Lagrange multipliers μ_a would be zero, the shadow prices s_i would vanish, and one would have the usual factor cost shares on the r.h.s. of eq. (1). That's why standard economics assumes that in economic equilibrium output elasticities equal factor cost shares. This would also justify the neoclassical-economic duality of production factors and factor prices, which is often used in traditional growth analyses. The

[5] Intertemporal optimization of utility U as a function of consumption C yields that the shadow price of capital contains an additional term proportional to the time derivative of dU/dC. This term is small for weakly decreasing U(C). If one does profit maximization without the fixed-cost constraint, the Lagrange multiplier μ in the shadow prices is replaced by 1.

essential information on production would be contained in the profit function as the Legendre transform of the production function. In the presence of technological constraints and non-zero shadow prices, however, the Lagrange multipliers μ_a are finite and functions of the output elasticities ε_i, so that the cost-share theorem and duality are not valid. For an understanding the role of energy in the economy, prices are not enough.

6. ENERGY AND ECONOMIC GROWTH REVISITED

As we have mentioned earlier, shortages and price spikes must have a negative impact on economic growth. A suitably modified theory of growth should be able to explain the many observations of growth slowdowns following price (Hamilton 2009).

More important, in the long run, is the forthcoming advent of "peak oil" (Hubbert 1969; Sorrell 2010). Whether it has already happened or whether it occurs ten or twenty years in the future, peak oil will be followed by a decline in the output of (cheap) petroleum. This must have a significant negative impact on future global economic growth. The reason is that energy in general, and oil in particular, are essential to virtually all economic activity. Hence, energy and oil have an output elasticity far greater than their small (though increasing) cost shares. As the prices of oil and oil substitutes rise, the demand for energy intensive products will fall, as happened in the autumn of 2008. That brings the price of oil temporarily back down, which encourages renewed consumption but discourages investment in energy conservation measures that depend on higher prices. This, in turn, delays needed economic adjustment while accelerating the onset of the next crisis.

Based on the LINEX production function with three factors, we have been able to explain past economic growth for the US, Germany, UK and Japan, without the need for a multiplier to explain most of the growth such as is needed in the Solow theory. We find that the output elasticity of an essential (non-substitutable) input, like petroleum, or more generally, energy, tends to be much larger than its cost-share, whereas the output elasticity of labour tends to be much smaller than its cost share (Kümmel et al. 1985, 2002); Lindenberger 2003; Ayres and Warr 2005; Warr et al. 2010; Warr and Ayres 2010; Kümmel 2011; Lindenberger and

Kümmel 2011).[6] When growth theory is expressed in terms of the LINEX function, without a priori assumptions about output elasticities (which are instead determined econometrically) it turns out that the estimate of output growth is improved and that the primary drivers of growth are capital deepening, and the increasing supply of energy or “useful work” (defined as primary energy input multiplied by the efficiency of conversion) in the economy. This is consistent with two past trends: (1) the discovery and exploitation of large oil (and gas) reserves, and (2) the increasing efficiency of conversion of primary energy (fossil fuels) into various forms of useful work, such as electric power and motive power.

The discrepancy between calculated output elasticities and factor cost shares, which results from technological constraints on factor substitution, can be associated with another issue: It can be argued that labour is “over-priced” in modern (western) economies whereas energy, especially petroleum, has been relatively under-priced in relation to its physical productivity.[7] This gives further support to a policy proposal which has been around for some time: namely, to shift taxes from labour to energy, partly as a means of environmental regulation, but with an additional benefit, namely reducing labour market distortions and thus lowering average unemployment rates (von Weizsaecker and Jesinghaus 1992; Goulder 1995; de Mooij 1999; Patuelli 2005). Indeed, past studies may have significantly underestimated the potential benefits of such a tax policy due to underrating the role of energy in the first place.

The non-equality of output elasticities and cost shares has important consequences for the standard theory of economic growth. The first implication is that the standard Cobb-Douglas production function must be used with caution, especially in combination with the assumption that its (constant) output elasticities equal factor cost shares. This assumption ignores important technological constraints associated with shadow prices that modify the “cost share theorem”

[6] In fact, the Cobb-Douglas function, which can be considered as a log-linear local approximation of any functional form, also reproduces historical economic growth relatively well if its (constant) output elasticities are not set equal to the factors’ cost shares but estimated econometrically. Then the CD output elasticities turn out to be approximately equal to the time averages of the (factor-dependent) output elasticities of the linex production function (Kümmel, 2011).

[7] The reason why the cost share of energy underestimates the value of energy is related to paying only for the extraction cost of energy, not for the value of the resource itself, which we simply exploit from nature. This is analogous to not paying the full value of human labour in times of slavery simply because slaves were, by definition, taken and used by force (Hall, 2012).

(see Section 5 above). The CD function also assumes that the elasticity of substitution between capital and labour is unity. Dropping the latter assumption leads to a more complicated functional form (CES) that permits other values of the elasticity of substitution (Arrow et al. 1961). The functional form we would favor is of the so-called LINEX (LINear EXponential) type, which belongs to the category of variable elasticity of substitution (VES) functions. In this case, both the elasticity of substitution and the output elasticities of factor inputs are functions of all the input variables, namely capital, labour and energy or energy services (work), which will vary over time. The LINEX function captures the technological fact that capital is productive only to the extent it is utilized by labour and energy. It further captures the production possibilities associated with increasing automation, where value added is generated more and more by automated, energy-driven capital, substituting for human labor. The empirical results of applying the LINEX function to observed growth trajectories provide evidence that the contribution of energy to economic growth is considerably higher than the relatively low share of energy in factor cost might indicate.

7. CONCLUSIONS

The non-equivalence of cost shares and output elasticities has important consequences for the analysis of economic growth in the presence of energy and environmental constraints. For instance, the under-pricing of energy (exergy) resources accounts in part for our addiction to oil and our over-dependence on certain key technologies such as the internal combustion engine. This, in turn, accounts for much of the atmospheric pollution, especially of greenhouse gases (GHGs), that has accompanied our global industrialization process. It further accounts for the logical inference that energy taxes or pollution taxes (or both) may be the way forward in terms of confronting the challenge of climate change by reducing GHG emissions.

The widespread idea of perpetual GDP growth suggests that "our great grand-children will be a lot richer than we are". This proposition must be challenged for at least two reasons. In the first place, it depends on the assumption that growth will continue as in the past, even if energy becomes scarce and ever more expensive. The rather standard assumption, that economic growth is independent of energy availability must be discarded absolutely. It is not tenable. It implies, wrongly, that energy-related emissions (GHGs) can be reduced or eliminated without consequences for growth.

The advent of "peak oil" means that the supply of easily recoverable oil and gas cannot be expected to continue to increase in the future and thus drive energy prices down – as it did for most of the last two centuries. Hence future economic growth will depend more than in the past on technological progress, especially in the area of increasing energy (exergy) efficiency in the economy. Unfortunately the rate of exergy efficiency increase (in the US, at least) has been slowing down since the 1970s. The implication of this trend is that either US economic growth will slow down permanently (with global consequences) or effective measures to increase the rate of increase of exergy efficiency must be undertaken to compensate for the coming decline in natural resource availability. A complication is that improved efficiency will not necessarily cut energy consumption (as much as we would like) due to so-called "rebound effects" from both consumers and producers, which will increase overall demand and partly undo some of the benefits of efficiency gains (Sorrell, Dimitropoulos, and Sommerville 2009). Only with adequate CO_2 pricing, and thus considerably higher energy prices, serious rebound effects can be avoided and considerable energy conservation can be realized.

The second reason for challenging the proposition (that economic growth is independent of energy) is that it implies that investments to ameliorate future environmental problems can safely be delayed (or avoided altogether) on the grounds that "our grandchildren will be much richer than we are" and hence, presumably better able to pay the bills. Slower or uncertain growth changes the balance between current and future costs and benefits.

A third point worth considering is that the view that GDP is a measure of human welfare is quickly losing support (Daly and Cobb 1989; Stiglitz 2009; van den Bergh 2009). So if because of increasing energy scarcity and higher energy prices we will experience less growth, it need not immediately translate into less welfare. This ameliorates the negative consequences of peak oil and climate protection, although it is not certain to what extent precisely. Of course, this story applies especially to rich countries; growth will still be needed to allow considerable welfare improvement in poor countries.

Finally, it is clear already that climate change is likely to impose huge environmental damages, due to more intense storms, floods, rising sea levels, the onset of new diseases (because micro-organisms evolve faster than large animals can develop immunity in a context of changing climate) and mass migrations from threatened areas (Stern et al. 2006). Moreover, the economic activities undertaken to prevent repair or compensate for damages – fighting new diseases, building fences to keep out refugees or dikes to protect coastal cities, for example – will increase the GDP, without increasing anyone's welfare. Indeed, there is now a very real potential for resource wars, as nations try to secure long-term energy supplies for themselves (Klare 2001; Friedrichs 2010). Many believe that the Iraq war was actually such an attempt. In summary, we believe that the present era is at a "tipping point".

In view of the foregoing, economic policies need to focus much more directly on increasing efficiency (conservation) and energy-related innovation to reduce carbon dioxide emissions from burning fossil fuels and fostering a transition to renewable energy (rather than to coal and non-conventional fossil fuels). This will be impossible without a considerable increase of fossil fuel prices to cover both scarcities and climate externalities (Sinn, 2009). Unfortunately, rather than raising prices and cutting supply of, and demand for, conventional fossil fuels, the body politic as well as many economists still seem focused on increasing the supply and keeping energy prices low.

The basic finding of this paper is that in industrial economies for energy the output elasticity is significantly larger than its cost share, whereas for labour the opposite is the case. An objection to this finding might be: “If this were true, money would be lying on the street. One only has to increase the input of cheap energy and decrease the input of expensive labour until output elasticities and cost shares are equal.” However this objection overlooks the barriers that block access to that side of the street where the money lies. These barriers reflect technological constraints of industrial production systems, resulting in factor shadow prices: One cannot increase the input of the production factor energy and correspondingly reduce the input of labour to handle machines beyond the (automation) limit defined by the actually designed capacity of machines. Another potential objection is that econometrically estimated output elasticities might be biased because the state of knowledge, which is an unobservable variable, affects the input quantities. However, models of innovation diffusion, where explicit time-dependences of technology parameters according to logistic differential equations take into account the evolution of the state of knowledge, and the application of these models to US, Japanese, and German production systems, confirm the mentioned discrepancies between output elasticities and cost shares (Kümmel et al., 2002). These discrepancies simply reflect technological restrictions, i.e. limits to factor input substitutability.

Acknowledgement. We thank Reiner Kümmel and two anonymous referees for valuable inputs to this paper. The usual disclaimer applies.

References

Aghion, Philippe, and Peter Howitt (1998). *Endogenous growth theory.* Cambridge MA: The MIT Press.

Allen, Edward L., and et al. (1976). US energy and economic growth, 1975-2010. Oak Ridge TN: Oak Ridge Associated Universities Institute for Energy Analysis.

Arrow, Kenneth J., H. B. Chenery, B. S. Minhas, and Robert M. Solow (1961). Capital-labor substitution and economic efficiency. *Review of Economics and Statistics* 43 (August): 225-250.

Ayres, Robert U., and Benjamin S. Warr (2005). Accounting for growth: The role of physical work. *Structural Change and Economic Dynamics* 16 (2): 181-209.

Ayres, Robert U., and Benjamin S. Warr (2009). Energy efficiency and economic growth: The "rebound effect" as a driver. In: *Energy Efficiency and Sustainable Consumption*, edited by H. Herring and S. Sorrell. London: Palgrave Macmillan.

Ayres, Robert U., and Benjamin S. Warr (2009). *The economic growth engine: How energy and work drive material prosperity*. Cheltenham, UK and Northampton MA, US: Edward Elgar Publishing.

Barnett, Harold J., and Chandler Morse (1963). *Scarcity and growth: The economics of resource scarcity*. Baltimore MD: Johns Hopkins University Press.

Barro, Robert J., and Xavier Sala-i-Martin (2003). *Economic growth*. 2nd Edition. The MIT Press, Cambridge. MA.

Berndt, Ernst R., and David O. Wood (1979). Engineering and econometric interpretations of energy-capital complementarity. American Economic Review 69 (3): 342-354.

Chen, Jing (2005). *The Physical Foundation of Economics: An Analytical Thermodynamic Theory*. World Scientific Press, Hackensack, NY.

Chen, Jing, and James Galbraith (2011). Institutional structures and policies in an environment of increasingly scarce and expensive resources: A fixed cost perspective, *Journal of Economic Issues* 45 (2): 301-308.

Daly, Herman E., and John Cobb (1989). *For the common good: Redirecting the Economy toward Community, the Environment, and a Sustainable Future*. Boston MA: Beacon Press.

d'Arge and Kogiku (1973). Economic growth and the environment. *Review of Economic Studies* 15(1): 61-77.

de Mooij, R.A. (1999). The double dividend of an environmental tax reform. Chapter 20 in: J.C.J.M. van den Bergh (ed.). *Handbook of Environmental and Resource Economics*. Edward Elgar, Cheltenham.

Denison, Edward F. (1979). *Accounting for slower growth*. Washington DC: Brookings Institution Press.

Felipe, Jesus, and Franklin M. Fisher (2003). Aggregation in production functions: What applied economists should know. *Metroeconomica* 54 (2 & 3): 208-262.

Fiorito, G., and J.C.J.M. van den Bergh (2011). Capital-Energy Substitution for Climate and Peak Oil Solutions? An International Comparison Using the EU-KLEMS Database. Working Paper, ICTA-UAB.

Friedrichs, J. (2010). Global energy crunch: How different parts of the world would react to a peak oil scenario. *Energy Policy* 38(8): 4562-4569.

Giampietro, M., Mayumi, K. and Sorman A.H (2011). *The metabolic pattern of society*. Routledge, London.

Goulder, Lawrence H. (1995). Effects of carbon taxes in an economy with prior tax distortions: An intertemporal general equilibrium analysis. *Journal of Environmental Economics and Management* 29 (3): 271-297.

Gross, L.S. and E.C.H. Veendorp (1990). Growth with exhaustible resources and a materials balance production function. *Natural Resource Modeling* 4: 77-94.

Hagelueken, Christian and Meskers, Christina M. (2010). Complex Life Cycles of Precious and Special Metals. In *Struengmann Forum Report: Linkages of Sustainability*, edited by Thomas E. Graedel and Ester van der Voet. MIT Press, Cambridge, MA.

Hall, Charles A. S., Cutler J. Cleveland, and Robert K. Kaufmann (1986). *Energy and resource quality: The ecology of the economic process.* New York: Wiley-Interscience.

Hall, Charles A.S. (2012), private communication.

Hall, Charles A.S., Klitgaard Kent A. (2012). *Energy and the Wealth of Nations*. Springer, New York.

Hamilton, James D. (2003). What is an oil shock? *Journal of Econometrics* 113: 363-398.

Hamilton, James D. (2005). Oil and the macroeconomy. In: *The New Palgrave: A Dictionary of Economics*, edited by J. Eatwell, M. Millgate and P. Newman. London: Macmillan.

Hamilton, James D. (2009). Causes and consequences of the oil shock of 2007-08. *Brookings Papers on Economic Activity*. Spring 2009: 215-259.

Harcourt, G. C. (1972). *Some Cambridge controversies in the theory of capital*. Cambridge, UK: Cambridge University Press.

Hubbert, M. King (1969). Energy resources. In: Resources and Man, edited by Cloud. San Francisco: W. H. Freeman and Company.

Hudson, Edward, and Dale W. Jorgenson (1974). US energy policy and economic growth, 1975-2000. Bell Journal of Economics and Management Science 5: 461-514.

IEA (2011). *25 Energy Efficiency Policy Recommendations*. International Energy Agency, Paris.

Jorgenson, Dale W. 1978. The role of energy in the US economy. *National Tax Journal* 31: 209-220.

Jorgenson, Dale W. (1984). The role of energy in productivity growth. *The Energy Journal* 5 (3): 11-26.

Kindleberger, Charles P. (1989). *Manias, panics and crashes: A history of financial crises*. second (revised) ed. New York: Basic Books.

King, Carey W., Charles A.S. Hall (2011). Relating Financial and Energy Return on Investment. *Sustainability* 2011, *3*(10), 1810-1832.

Klare, Michael T. (2001). *Resource wars: The new landscape of global conflict*. New York: Henry Holt & Co.

Klein, Lawrence R. (1946). Remarks on the theory of aggregation. *Econometrica* 14: 303-312.

Koetse, M.J., De Groot, H.L.F., Florax, R. (2008): Capital-energy substitution and shifts in factor demand: A meta-analysis. *Energy Economics* 30: 2236–2251.

Kümmel, Reiner (1982). The impact of energy on industrial growth. *Energy* 7: 189-203.

Kümmel, Reiner (2011). *The Second Law of Economics – Energy, Entropy, and the Origins of Wealth.* New York, Dordrecht, Heidelberg, London: Springer (The Frontiers Collection).

Kümmel, Reiner, Wolfgang Strassl, Alfred Gossner, and Wolfgang Eichhorn (1985). Technical progress and energy-dependent production functions. *Journal of Economics* 45: 285-311.

Kümmel, Reiner, Julian Henn, and Dietmar Lindenberger (2002). Capital, labor, energy and creativity: modeling innovation diffusion. *Structural Change and Economic Dynamics* 13: 415-433.

Kümmel, Reiner, Robert U. Ayres, and Dietmar Lindenberger (2010). Thermodynamic laws, economic methods and the productive power of energy. *Journal of Non-equilibrium Thermodynamics* 35: 145-181.

Lindenberger, Dietmar (2003). Service production functions. *Journal of Economics* 80: 127-142.

Lindenberger, Dietmar, and Reiner Kümmel (2011). Energy and the state of nations. *Energy* 36: 6010-6018.

Mankiw, N. Gregory (1997). *Macroeconomics*. London: Worth Publishing.

Patuelli, R. Nijkamp, Peter, Pels, E. (2005). Environmental Tax Reform and the Double Dividend: A meta-analytical performance assessment. *Ecological Economics* 55 (4): 564-583.

Robinson, Joan (1971). The measure of capital: The end of the controversy. *The Economic Journal* 81:597-602.

Romer, Paul M. (1994). The origins of endogenous growth. *Journal of Economic Perspectives* 8 (1): 3-22.

Schelling, Thomas C. (2002). “Greenhouse Effect.” In: The Concise Encyclopedia of Economics. http://www.econlib.org/Library/Enc/GreenhouseEffect.html

Sinn, Hans W. (2008). Public policies against global warming: A supply-side approach. *International Tax and Public Finance* 15(4): 360-394.

Skinner, Brian J. 1976. *Earth resources*. Englewood Cliffs NJ: Prentice-Hall.

Smith, V. Kerry, and John Krutilla, eds. (1979). *Scarcity and growth reconsidered*. Baltimore MD: Johns Hopkins University Press.

Smulders, Sjak (2005). Endogenous technological change, natural resources and growth. In: D. Simpson and M. Toman (eds). *Scarcity and Growth in the New Millennium*. RFF Press, Boston.

Solow, Robert M. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics* 70: 65-94.

Solow, Robert M. (1957). Technical change and the aggregate production function. *Review of Economics and Statistics* 39 (August): 312-320.

Sorrell, Steve (2010). Global oil depletion: a review of the evidence. *Energy Policy* 38(9): 5290-5295.

Sorrell, Steve, John Dimitropoulos, and Matt Sommerville (2009). Empirical estimates of the direct rebound effect: A review. *Energy Policy* 37 (4): 1356-1371.

Stern David I. (2011). The role of energy in economic growth. *Annals of the New York Academy of Sciences* 1219: 26-51.

Stern, Nicholas, S. Peters, A. Bakhshi, C. Bowen, S. Cameron, D. Catovsky, S. Crane, S. Cruikshank, N. Dietz, S. -L. Edmondson, I. Garbett, G. Hamid, D. Hoffman, B. Ingram, N. Jones, H. Patmore, R. Radcliffe, M. Sathiyarajah, C. Stock, T. Taylor, H. Wanjie Vernon, and D. Zeghelis (2006). The Stern review: The economics of climate change. London: HM Treasury.

Stiglitz, Joseph E. (2009). GDP Fetishism. *The Economists' Voice* 6(8), Article 5. http://www.bepress.com/ev/vol6/iss8/art5 van den Bergh, J.C.J.M. (1999). Materials, capital, direct/indirect substitution, and materials balance production functions. *Land Economics* 75 (4), 547-561.

van den Bergh, Jeroen C.J.M. (2009). The GDP Paradox. *Journal of Economic Psychology* 30(2): 112-135.

van den Bergh, Jeroen C.J.M., and Peter Nijkamp (1994), Dynamic macro modelling and materials balance. *Economic Modelling* 11: 283-307.

von Weizsaecker, Ernst Ulrich, and Jochen Jesinghaus (1992). *Ecological tax reform: A policy proposal for sustained development*. London: Zed Books.

Warr, Benjamin S., and Robert U. Ayres (2010). Evidence of causality between the quantity and quality of energy consumption and economic growth. *Energy* 35: 1688-1693.

Warr, Benjamin S., Nina Eisenmenger, Friedrich Krausmann, Heinz Schandl, and Robert U. Ayres. (2010). Energy use and economic development; A comparative analysis of useful work supply in Austria, Japan, the United Kingdom and the US during 100 years of economic growth. *Ecological Economics* 69(10): 1904-1917.

Europe Campus
Boulevard de Constance
77305 Fontainebleau Cedex, France
Tel: +33 (0)1 60 72 40 00
Fax: +33 (0)1 60 74 55 00/01

Asia Campus
1 Ayer Rajah Avenue, Singapore 138676
Tel: +65 67 99 53 88
Fax: +65 67 99 53 99

Abu Dhabi Campus
Muroor Road - Street No 4
P.O. Box 48049
Abu Dhabi, United Arab Emirates
Tel: +971 2 651 5200
Fax: +971 2 443 9461

www.insead.edu

INSEAD

The Business School
for the World®

Printed by INSEAD