8 Exergy economics

New insights into energy consumption and economic growth

Paul Brockway, Steve Sorrell, Tim Foxon and Jack Miller

Introduction

To be effective in mitigating climate change, technical and policy initiatives to reduce energy demand must have significant impacts at the aggregate level. This means they must contribute to the decoupling of primary energy consumption from economic output for both national economies and the world as a whole. However, the feasibility, difficulty and cost of decoupling are disputed.

The rate of growth of global primary energy consumption has been remarkably stable since 1850 (2.4 per cent/year ± 0.08 per cent) and shows little sign of slowing down (Jarvis et al., 2012). However, since primary energy consumption ($E$) has grown more slowly than gross domestic product (GDP) ($Y$), there has been a steady decline in global primary energy intensity ($E/Y$) – termed relative decoupling. Globally, primary energy intensity fell by ~1.3 per cent/year between 1980 and 2000, but progress has subsequently slowed to only 0.3 per cent/year. To achieve the goals of the Paris Agreement, global primary energy intensity must fall at least six times faster than this – a much faster rate than has previously been achieved (Loftus et al., 2014).

The historical decline in global energy intensity appears largely due to countries getting richer rather than from finding ways to produce particular levels of wealth with less primary energy consumption (Csereklyei et al., 2014). This in turn suggests that the technological changes that reduce energy intensity are strongly correlated with those that increase wealth – so the energy required to produce a unit of output has fallen, but the ‘energy savings’ have been partly offset by increased output. It is possible that there is a causal relationship between these trends (i.e. lower energy intensity leads to increased economic output), but this is difficult to establish empirically. Moreover, despite wide differences in energy intensity, very few countries have achieved absolute decoupling of primary energy consumption from GDP (i.e. GDP rising while energy consumption is falling) for more than short periods of time. Also, when absolute decoupling has been achieved (such as in the UK) it has partly been driven by the ‘outsourcing’ of energy-intensive manufacturing to other regions (Hardt et al., 2018). Such outsourcing is clearly not feasible at the global level.
This apparently strong link between energy consumption and economic activity raises important questions for both theory and policy. The orthodox view is that both increased energy consumption and improved energy efficiency provide a relatively small contribution to the growth in economic output. Consistent with this view, orthodox economists argue that technological change can reduce energy consumption with relatively little effect on economic growth. In contrast, some ecological economists claim that over the last century economic growth has largely been achieved by providing workers with increasing quantities of energy, both directly and indirectly, as embodied in capital equipment (Cleveland et al., 1984). Ecological economists therefore view energy as contributing more to economic growth than is suggested by its small share of total costs (5–10 per cent). They are correspondingly more sceptical about the feasibility of decoupling.

The success of climate policy depends in part on which of these views is correct – or more precisely, which more accurately describes the situation for different regions and stages of economic development (Foxon, 2017; Stern and Kander, 2012) But debates on this topic involve a host of theoretical and methodological issues that are both highly technical and difficult to resolve. For example, there have been several hundred studies that use sophisticated econometric techniques to explore the ‘causal’ relationships between GDP and energy consumption, but these have failed to reach a consistent conclusion (Kalimeris et al., 2014; Omri, 2014).

Recently, however, a new field of research has emerged which has the potential to throw new light on these long-standing questions. This approach hinges upon the thermodynamic concept of exergy (the portion of an energy flow that can be used to perform physical work), and the use of physical measures of energy efficiency that are based upon the second-law of thermodynamics rather than the first. The argument is that exergy is the preferred way to measure energy flows since it captures both the quantity and quality of energy, while second-law efficiency measures are preferred to first-law since they capture the distance from the theoretical maximum efficiency.

Underlying this new approach is the estimation of the useful exergy inputs into national economies – where useful exergy is defined as the exergy outputs of end-use conversion devices, such as the mechanical drive from an engine, the high-temperature heat from a furnace or the visible light from a lightbulb. Useful exergy, in turn, is the product of the exergy inputs to those conversion devices (which can be estimated from data on final energy consumption) and their second-law conversion efficiencies. Researchers in this field are beginning to construct consistent time series of the total useful exergy consumption of individual countries and regions (Brockway et al., 2014; Serrenho et al., 2016; Warr et al., 2010). These databases provide a measure that can be used alongside the more traditional measures of primary and final energy consumption to gain deeper insights into the role of energy in the economy.

The core claim of these researchers is that useful exergy is a key driver of economic growth – and that this contribution is not recognised by orthodox
economics (Warr and Ayres, 2012). Increases in economic output have historically depended upon increased supplies of useful exergy, achieved through a combination of increasing use of primary energy, shifting towards higher-quality energy carriers (e.g., from oil to electricity) and improving second-law efficiencies at all stages of the energy conversion chain. Warr and Ayres (2012) go so far to suggest that improvements in second-law conversion efficiency provide a quantifiable surrogate for the majority of technical change that contributes to economic growth. Hence, far from being a minor contributor to economic growth, the combination of increased energy inputs and improved energy efficiency becomes a key driver. One implication of this work is that energy efficiency improvements by producers can significantly boost economic output – thereby partly or wholly offsetting the energy savings per unit of output that result from the improved efficiency. In other words, rebound effects for producers could be large.

This chapter provides an overview of this emerging field. The following section summarises the orthodox view of the relationship between energy consumption and economic growth, including the assumptions upon which this rests and the limitations of those assumptions. Next, the concept of useful exergy is introduced to show how this may help to improve our understanding of this relationship. The following section summarises some recent research that estimates the useful exergy inputs into national economies, explores the trends in these over time, includes useful exergy within economic models and uses those models to identify the drivers of economic growth. We highlight two claims: first, that energy efficiency improvement by UK producers have provided one-quarter of UK economic growth since 1971; and second, that corresponding improvements by Chinese producers have increased global energy consumption. The chapter concludes with future directions for this line of research.

The role of energy in the economy

In the models used by orthodox economists, firms combine primary inputs (capital and labour) and intermediate inputs (energy and materials) to produce goods and services. Primary inputs facilitate production but do not form part of the product and are not used up during production (although capital may depreciate). In contrast, intermediate inputs are ‘created’ by production and are either embodied in products or used up during production. Subtracting the purchases of intermediates from the value of output leads to a measure of value added, which is the income received by capital and labour.

Orthodox models attribute increases in economic output to increases in primary and intermediate inputs and improvements in total factor productivity (TFP) – where the latter is the portion of growth not explained by increases in inputs (OECD, 2001; Solow, 1956). Increases in value added (including, at the aggregate level, GDP) are attributed to increases in primary inputs and TFP – with the latter accounting for a significant proportion of the total. TFP can be estimated as a residual in growth accounting studies or as a parameter in econometric studies, but it has traditionally been treated as exogenous and equivalent
to technical change.\textsuperscript{1} However, more recent models make technical change endogenous and simulate the positive externalities from education and research and development (R&D) (Aghion et al., 1998; Romer, 1994). These models also attribute a portion of economic growth to improvements in the quality of capital and labour inputs – such as better-educated workers.

Central to orthodox economics is the specification of production functions for firms, sectors or the economy as a whole, indicating the maximum output that can be produced from different quantities of primary and intermediate inputs (OECD, 2001). Production functions can either be specified for gross output and include all inputs or specified for value added and only include primary inputs. Specifications typically include a TFP multiplier that 'shifts' the production function over time, thereby capturing technical and other changes that allow more output to be produced from the same quantity of inputs. Production functions can be defined for different levels of aggregation using different functional forms and with differing rates of productivity improvement for each input. But it is generally assumed that production exhibits constant returns to scale, that markets are competitive, that firms maximise profits and that inputs can be substituted for one another following a change in relative prices. Using these assumptions, it can be shown that the rate of growth of output over time is a weighted average of the rate of growth of each input and the rate of growth of TFP. The weight on each input is the ‘partial output elasticity’ for that input, or the percentage change in output following a percentage change in that input, holding other variables constant. With these assumptions, it can be shown that the partial output elasticity is equal to the share of that input in total costs. This result has been labelled the cost share theorem (Kümmel et al., 2010).

The cost share theorem, together with the assumption of input substitutability, has important implications for the role of energy in economic production. Since energy represents a small share of total costs for most producers (<5 per cent), the theory implies that increases in energy inputs and improvements in the productivity of those inputs should make only a minor contribution to economic growth. Similarly, constraints on energy supplies are unlikely to have a major impact on economic growth since it should be possible to substitute away from energy. Taken together, these assumptions imply that energy consumption can be substantially decoupled from economic output.

This approach has been criticised by ecological economists, who challenge the core assumption (derived from the national accounts) that capital and labour should be treated as primary inputs, and that energy and materials should be treated as secondary inputs that make no contribution to value added. This makes little sense from a physical perspective, since all physical, biological and economic activity depends upon flows of high-quality energy that are then returned to the environment in the form of low-temperature heat. Like the biosphere, the economy is driven by solar energy, both directly and as embodied in biomass and fossil fuels. Labour and capital are not productive on their own – they only add value by harnessing the ‘free’ energy flows provided by nature. The productivity of capital and labour therefore depends entirely upon the associated energy flows.
Linked to this, ecological economists question the treatment of energy as a ‘produced’ input that can be substituted by capital and labour. Technically, the scope for substitution will be constrained at the level of the economy as a whole since producing more capital requires more of the thing that it is substituting for (Stern, 1997). In addition, many production functions violate the laws of thermodynamics, since they allow output to be produced with little or no energy. More realistic constraints on the relative magnitude of different inputs could mean that economies are less flexible in adjusting to rising energy prices than is traditionally assumed (Berndt and Wood, 1979; Lindenberger and Kümmel, 2011). Such constraints may undermine the cost share theorem, meaning that the dependence of capital and labour on energy flows could magnify the economic impact of changes in those flows (Giraud and Kahraman, 2014).

The common treatment of energy as an undifferentiated input is also problematic. Energy carriers differ in quality on multiple dimensions, including their flexibility of use, amenability to storage, energy density, economic productivity and capacity to do work (Cleveland et al., 2000). Since high-quality energy carriers are more productive, they should be given more weight in aggregate measures of energy consumption. When this is done, aggregate energy intensity is found to be declining more slowly than is commonly assumed (Cleveland et al., 2000; Berndt, 1978). Studies that neglect changes in energy quality may therefore overlook an important contributor to economic growth (Gentvilaitė et al., 2015; Stern, 2010).

In contrast to the neglect of energy by orthodox economists, economic historians attribute a central role to energy in explaining previous long-term surges in economic growth – and in particular the nineteenth-century industrial revolution (Allen, 2009; Kander et al., 2014; Pomeranz, 2009; Wrigley, 2013). The continuing importance of energy is also suggested by the large impact of energy price shocks on economic output (Kilian, 2008), and by the limited decoupling that has been achieved to date. For example, Csereklyei et al. (2014) analysed 99 countries over the period 1971–2010 and found that, on average, every 1 per cent increase in per capita wealth was associated with a 0.7 per cent increase in per capita energy consumption. But it is not clear whether this strong correlation is due to economic growth causing increased energy consumption (the orthodox view), increased energy consumption causing economic growth (the ecological view), or a combination of the two. While it is possible to test this econometrically, the results are ambiguous and sensitive to the method, data and specification employed (Kalimeris et al., 2014; Omri, 2014).

These various strands of theory and evidence raise concerns about the validity of orthodox models, the accuracy of the cost share theorem and the feasibility of absolute decoupling. If economic growth depends upon increased energy consumption, it may be difficult to reduce global carbon emissions while at the same time increasing global GDP. However, the studies arguing for the importance of energy are limited in number, variable in quality and inconsistent in approach – and have largely been ignored by mainstream economists. The approach described below – termed ‘exergy economics’ – represents an attempt
to build a bridge between these two communities. The distinctive features of this approach are: the use of exergy as a thermodynamic measure of energy quality; the focus on the ‘useful’ stage of the energy conversion chain, rather than the primary or final stages; and the willingness to challenge key assumptions of orthodox economics, such as the cost share theorem. These are summarised next.

**Foundations of exergy economics**

*The concept of useful exergy*

Exergy is a measure of the portion of an energy flow that can be used to perform work, i.e. the portion that is ‘available’ or ‘useful’. As with energy, exergy is measured in joules, takes a variety of forms (e.g. kinetic, electrical or chemical) and can be converted from one form to another. But while energy is solely a measure of quantity, exergy is a measure of both quantity and quality – where quality is defined as the capacity to perform work. So, for example, 1 kWh of electricity has the same energy as 5 kg of water at 20°C, but the electricity has more exergy (i.e. is of higher quality) owing to its greater potential to be converted to physical work. The biggest difference between energy and exergy arises when considering thermal energy (heat). For any given quantity of heat within a particular environment, a temperature-dependent portion constitutes low-grade heat that has little or no ability to perform work. This represents a portion of an exergy flow that has been dissipated, or ‘destroyed’.

The concept of exergy derives from the second-law of thermodynamics, which (in one form) states that every energy conversion process involves the loss of some measure of energy quality – which means that some exergy is necessarily destroyed. Energy, on the other hand, is always conserved and cannot be destroyed, as per the first-law of thermodynamics.

The term ‘exergy’ was first introduced by Zoran Rant (1956), although the principles on which it is based date back to the nineteenth century (Anderson, 1887). Exergy may be formally defined as the maximum physical work that can be extracted from a system as it reversibly comes into equilibrium with its environment. The exergy of a system depends upon the differences between that system and its environment, which may be in terms of kinetic energy, potential energy, temperature, pressure or chemical potential (Baierlein, 2001; Romero and Linares, 2014). Exergy can be defined for materials as well as for energy flows, and has been proposed as a global sustainability indicator (Romero and Linares, 2014) and a universal measure of resource availability (Valero and Valero, 2011).

Here we focus upon the use of exergy as an alternative measure of the chain of energy flows within a national economy. At the top of this chain is primary exergy, derived from fossil fuels, nuclear fission and renewables. The exergy of fossil fuels differs from their energy content by a so-called ‘exergy factor’ (Szargut et al., 1987), while the exergy of nuclear and renewables depends upon
how they are measured. Sources of primary exergy are processed and converted into commercial energy carriers (secondary exergy) such as electricity, gasoline and diesel which are ultimately delivered to consumers (final exergy) with some losses along the way (e.g. resistive losses in electricity grids). The last stage of exergy conversion takes place within end-use devices such as engines, boilers, furnaces, motors and light bulbs which convert final exergy into useful exergy, such as low- and high-temperature heat, mechanical power and electromagnetic radiation.

Following end-use conversion, useful exergy is preserved or trapped within passive systems for a period of time to produce energy services (Cullen and Allwood, 2010; Cullen et al., 2011). So, for example, the heat delivered from a boiler (conversion device) is held within a building (passive system) for a period of time to provide thermal comfort (energy service). Unlike useful exergy, energy services cannot be measured in common units and hence cannot be aggregated. Useful exergy is eventually dissipated as low-temperature heat, but improvements in second-law conversion efficiency (e.g. more efficient boilers) or the ability of passive systems to trap exergy (e.g. more insulation) will allow more energy services to be provided per unit of useful exergy.

**Useful exergy and economic production**

The main reason for thinking about useful exergy in this context is that it provides a more relevant measure of the contribution of energy to economic production and human welfare. The exergy that is destroyed at each stage of the conversion chain contributes no economic value and no energy services, but the useful exergy at the final end of the chain contributes to the production of marketable goods (e.g. the heat used to manufacture steel) and to the energy services required by consumers (e.g. the light energy used for illumination). Improvements in second-law conversion efficiency at all stages of the energy chain allow more useful exergy to be delivered from the same amount of primary energy – and this is, rather than the total energy outputs of conversion devices, that has economic value. Hence, it is the productive part of energy flows – useful exergy – that should be the focus of attention within economic models.

In a series of papers, Ayres and Warr estimated the useful exergy flows within national economies over the past century and included these within simple models of economic growth (Ayres and Warr, 2010; Ayres et al., 2003; Warr and Ayres, 2012; Warr et al., 2008). Their results suggested that the output elasticity of useful exergy was at least ten times greater than the cost share of energy, and much larger than the output elasticity of labour. Moreover, when energy was replaced with useful exergy, the estimated contribution of TFP to economic growth largely disappeared – at least for the period prior to 1970. Their explanation for these results was that energy (cf. exergy) is more productive than orthodox economists assume, and that the productivity of primary exergy has increased over time owing to continuing improvements in second-law conversion efficiency.
The studies by Ayres and Warr embody a number of innovations, including: estimating aggregate time series of useful exergy; using this data within economic models (despite the fact that useful exergy is not a traded commodity); and employing an unconventional ‘linear exponential’ (LINEX) production function. The latter has not been accepted by orthodox economists because it violates some of the standard assumptions of neoclassical production theory. However, the theoretical arguments in favour of employing useful exergy are persuasive and the concept is not wholly unfamiliar to economists since it amounts to ‘quality-weighting’ energy inputs.

Comparable quality-weighting of labour and capital inputs is now an standard feature of orthodox growth accounting (OECD, 2001). As an illustration, Figure 8.1 shows time series of both standard and quality-weighted capital, labour and energy inputs for Portugal over the period 1960–2010, where the latter is defined as useful exergy. Since quality-weighted inputs grow faster than the standard input measures, they can explain a larger proportion of economic growth, leaving less to be attributed to a residual (TFP). The next section summarises the contribution of researchers who building upon Ayres and Warr’s work.

Figure 8.1 Normalised time series of inputs and outputs to production in Portugal.
Source: Santos et al. (2018), with permission.

Note
Capital inputs measured as a stock of assets and a flow of services. Labour inputs measured as total hours worked and total hours worked adjusted with a human capital index. Energy inputs measured as primary energy supply and useful exergy consumed.
Research findings and insights

Exergy accounting

A first step in understanding the contribution of useful exergy to economic growth is to map the exergy flows through a national economy. An estimate of these flows for a single year allows the locations and magnitudes of exergy losses to be identified, the relative efficiencies of different processes to be compared and the technical (but not necessarily economic) potential for improvements to be highlighted. Similarly, an estimate for several years allows the trends in exergy use and efficiencies to be identified, while estimates for several countries allow their relative efficiencies to be compared. Estimating useful exergy involves: (a) collecting data on primary and final energy consumption and converting these to an exergy basis; (b) mapping final exergy flows onto different

Table 8.1 Breakdown of end-uses, by useful exergy category and energy carrier group

<table>
<thead>
<tr>
<th>Energy carrier group</th>
<th>Category of useful exergy</th>
<th>End-use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combustible fuels</td>
<td>Heating</td>
<td>High-temp. industrial heat</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Med-temp. industrial heat</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Water heating</td>
</tr>
<tr>
<td></td>
<td>Mechanical drive</td>
<td>Space heating/cooking</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gasoline road vehicles</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diesel automobiles</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diesel goods vehicles</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Aviation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diesel marine transport</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diesel rail</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Coal rail</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Industrial static motors</td>
</tr>
<tr>
<td></td>
<td>Lighting</td>
<td>Non-electric lighting</td>
</tr>
<tr>
<td>Electricity</td>
<td>Heating</td>
<td>High-temp. industrial heat</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Med-temp./industrial heat</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Water heating</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Space heating</td>
</tr>
<tr>
<td></td>
<td></td>
<td>‘Wet’ appliances (e.g. dishwashing)</td>
</tr>
<tr>
<td></td>
<td>Cooling</td>
<td>Space cooling (AC)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Refrigeration</td>
</tr>
<tr>
<td></td>
<td>Mechanical drive</td>
<td>Electric motors</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Electric rail</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Electric road vehicles</td>
</tr>
<tr>
<td></td>
<td>Lighting</td>
<td>Electric lighting</td>
</tr>
<tr>
<td></td>
<td>Electronics</td>
<td>Consumer electronics</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Computing</td>
</tr>
<tr>
<td>Food and feed</td>
<td>Muscle work</td>
<td>Draught animal work</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Human work</td>
</tr>
</tbody>
</table>

Source: Miller et al. (2016), with permission (and without changes).
categories of useful exergy (such as mechanical power and heat) and different end-uses within each category (e.g., cars trucks); and (c) estimating average exergy efficiencies for those end-uses (see Table 8.1).

The first national exergy accounts were compiled by Reistad (1975), whose study coincided with rising oil prices and increasing concern about energy efficiency. Interest declined following the oil price collapse of 1981, but was reinvigorated after 2000 by Ayres and Warr (2005; Warr and Ayres, 2012). Other authors have since refined the methodology, both in terms of the level of disaggregation and the accuracy of efficiency estimates (Brockway et al., 2014; Miller et al., 2016; Serrenho, 2013). There are a growing number of single and multi-country studies within the OECD (Ayres et al., 2003; Brockway et al., 2014; Hammond and Stapleton, 2001; Serrenho, 2013; Serrenho et al., 2016; Warr et al., 2010; Williams et al., 2008) and the methodology has now been extended to Mexico (Guevara et al., 2016), India (Magerl, 2017) and China (Brockway et al., 2015). Lack of data remains a serious obstacle, but there have been significant improvements over the last few years with increasing efforts towards standardisation (Sousa et al., 2017).

**Exergy efficiency**

A key outcome of useful exergy accounts are time series estimates of primary to useful exergy efficiency – that is, the ratio of useful to primary exergy consumption ($\epsilon_{PU} = \frac{B_U}{B_P}$) – as well as final to useful exergy efficiency ($\epsilon_{FU} = \frac{B_U}{B_F}$). These estimates reflect both improvements in conversion efficiency and structural change within the economy. As an example, Serrenho et al. (2016) estimate that $\epsilon_{FU}$ in Portugal increased from 6 per cent in 1850 to 20 per cent in 2010, with most of this improvement occurring during the post-war period of electrification and industrialisation. Similarly, Brockway et al. (2014) estimate that $\epsilon_{PU}$ in the UK increased from 9 per cent to 15 per cent between 1960 and 2000, but has since remained stable (Figure 8.2). Closer examination reveals efficiency improvements in all end-use categories in the UK, with primary to useful heating efficiency increasing from 8 to 12 per cent, electricity efficiency from 8 to 14 per cent and mechanical drive from 11 to 21 per cent.

While $\epsilon_{PU}$ in Portugal and the UK stabilised only recently, it has remained around 11 per cent in the US for half a century (Figure 8.2). The tendency for the rate of improvement in exergy efficiency to slow in advanced economies was first observed by Williams et al. (2008), who termed it ‘efficiency dilution’. The reason is the increasing proportion of exergy being used in less efficient end-uses (e.g., air-conditioning, car travel, space heating), combined with a slowdown in the rate of efficiency improvement for individual end-uses. Most advanced economies are ‘outsourcing’ heavy industry to emerging economies and since heavy industry is relatively exergy efficient (although exergy intensive), this reduces the aggregate exergy efficiency of those economies. In contrast, the exergy efficiency of emerging economies is improving rapidly (Figure 8.2).
Improvements in exergy efficiency mean that the useful exergy inputs to national economies are growing faster than either primary or final exergy inputs (Figure 8.3). For example, the 10-fold growth in useful exergy use in China between 1971 and 2010 was supplied by a 4-fold increase in primary exergy combined with a 2.5-fold improvement in primary to useful exergy efficiency.

Table 8.2 decomposes the trends in useful exergy consumption in the US, UK and China to identify the relative contribution of increases in primary exergy, changes in the relative importance of different end-uses (structure) and changes in the efficiency of those end-uses. If emerging economies follow the same pattern as the Organisation for Economic Co-operation and Development

---

**Table 8.2** Decomposing the drivers of useful exergy consumption in the UK, US and China over the period 1971–2010

<table>
<thead>
<tr>
<th>Country</th>
<th>Primary exergy ($D_p$)</th>
<th>Structure ($D_s$)</th>
<th>Efficiency ($D_e$)</th>
<th>Useful Exergy ($D_T = D_pD_sD_e$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>China</td>
<td>3.96</td>
<td>1.66</td>
<td>1.48</td>
<td>9.76</td>
</tr>
<tr>
<td>US</td>
<td>1.32</td>
<td>0.90</td>
<td>1.29</td>
<td>1.53</td>
</tr>
<tr>
<td>UK</td>
<td>1.01</td>
<td>0.90</td>
<td>1.58</td>
<td>1.43</td>
</tr>
</tbody>
</table>

Source: Brockway et al. (2015), with permission (and without changes).
Economic insights

Exergy intensity of national economies

Once estimates of primary, final and useful exergy are available, their relationship to GDP \( (Y) \) can be examined. The primary exergy intensity of a national economy \( (\theta_P = B_P/Y) \) can be decomposed as:

\[
\theta_P = \frac{B_P}{Y} = \frac{B_F}{B_F} \cdot \frac{B_U}{B_U} \cdot \frac{B_F}{Y}
\]

Or:

\[
\theta_P = \varepsilon_{PF} \varepsilon_{FU} \theta_U
\]

Hence, reductions in primary exergy intensity may result from either improvement in primary to useful exergy efficiency \( (\varepsilon_{PF} = \varepsilon_{PF} \varepsilon_{FU}) \) or reductions in useful exergy intensity \( (\theta_U) \). As an illustration, Figure 8.4 illustrates the long-term trends in primary exergy and useful exergy intensity in five countries (Serrenho et al., 2016). This period includes transitions from agricultural to industrial societies (which occurred somewhat later in Portugal), together with two world wars. Although the picture is complex and there are important differences

Figure 8.3 Normalised trends in primary exergy, useful exergy and primary to useful exergy efficiency in China between 1971 and 2010.

Source: Brockway et al. (2015), with permission (and without changes).
Figure 8.4 Primary exergy (top) and useful exergy (bottom) intensity of Portugal, UK, US, Austria and Japan, 1850–2010.

Source: Serrenho et al. (2016), with permission.
between countries, one notable feature stands out: primary exergy intensity has declined over time while useful exergy intensity has remained relatively stable. This implies that most of the reduction in primary exergy intensity has derived from improvements in primary to useful exergy efficiency. While all five countries have experienced a relative decoupling of primary exergy from GDP there

Figure 8.5 Final exergy (top) and useful exergy (bottom) intensities in the EU-15, 1960 to 2010.

Source: Serrenho et al. (2014), with permission.
has been little or no decoupling of useful exergy. Indeed, the useful exergy intensity of the modern Portuguese economy is comparable to what it was in 1860.

Figure 8.5 provides more recent estimates of final and useful exergy intensities for the EU 15. Again, the picture is complex, but there is little evidence of a long-term trend towards lower useful exergy intensity. Serrenho et al. (2014) show that, once differences in residential energy use (linked to average temperatures) and high-temperature heat (linked to heavy industry) are taken into account, useful exergy is statistically constant across time and equal for the EU 15. Taken together, this evidence is suggestive of a strong link between useful exergy and economic output. However, the trend is not universal (e.g. useful exergy intensity is rising in Mexico (Guevara et al., 2016) and falling in China (Brockway et al., 2015)) and closer investigation is required to understand the underlying determinants.

Useful exergy as a factor of production

A second insight is gained by replacing primary energy with useful exergy within models of economic production. As noted, this was first done by Ayres and Warr (2005) who employed a LINEX production function and were able to explain US economic growth over the period 1900–1973 without the need for a TFP multiplier. However, there seems little prospect of this function being accepted by mainstream economists, so researchers are seeking more conventional ways to include useful exergy.

One approach is to estimate standard, three-input (capital, labour and energy – KLE) ‘constant elasticity of substitution’ (CES) production functions, with primary energy (E) being replaced by useful exergy ($B_U = \varepsilon PU BP$). CES functions form the foundation of many macroeconomic models, but (remarkably) the parameter values tend to be assumed rather than estimated (Sorrell, 2014). Traditional approaches to estimating substitution elasticities are problematic since they rely upon the cost share theorem, but estimating the CES function requires non-linear techniques that can be unreliable (Henningsen and Henningsen, 2012; Prywes, 1986). Further, it is necessary to impose assumptions about the ‘separability’ of inputs that typically lack empirical justification (Sorrell, 2014). Heun et al. (2017) estimate aggregate CES functions for the UK and Portugal using data on useful exergy and other inputs over the period 1960–2009. They find that the partial output elasticity of each input varies over time and differs from the cost share – thereby questioning the validity of the cost share theorem. However, they also find the results are sensitive to the specification used and the estimated output elasticities change rapidly over short periods of time.

Santos et al. (2018) take an alternative approach, using ‘co-integration’ techniques to test for the existence of an aggregate production function for Portugal over the period 1960 to 2009. If time series of labour, capital, energy (or useful exergy) and GDP are found to be co-integrated, this suggests there is a stable,
long-term relationship between them (Dickey et al., 1994). This relationship may in turn be interpreted as an aggregate production function, and the estimated parameter values can be used to derive the partial output elasticities. Santos et al. test 32 different specifications that vary in terms of whether a TFP multiplier is included, whether capital and labour inputs are quality-weighted, whether energy inputs are included, and whether these are measured as primary energy or useful exergy.

To interpret the results as an aggregate production function, the variables must be co-integrated, the parameters must be non-negative and there must be evidence that the inputs ‘cause’ the output. Santos et al. find that none of the specifications incorporating a TFP multiplier meet these criteria – suggesting that standard formulations are incorrect. Instead, the only specification that meets these criteria is when both capital and labour are quality-adjusted and useful exergy is included – again suggesting that standard formulations are incorrect. Moreover, this specification includes two co-integrating relationships, one of which is interpreted as a constraint on input combinations as a consequence of the essential contribution of useful exergy to economic production. Overall, this rigorous study provides strong support for the inclusion of useful exergy as a factor of production.

Rebound effects from improved exergy efficiency

Improvements in exergy efficiency make useful exergy cheaper, thereby boosting economic output and encouraging the substitution of useful exergy for capital and labour. This in turn reduces the exergy savings from those improvements – a form of rebound effect.

There is a large and growing literature on rebound effects, but most studies focus upon energy efficiency improvements by consumers rather than producers since these are easier to estimate. This is unfortunate, since rebound effects for producers may potentially be larger (Saunders, 2013; Sorrell, 2007). The development of economic models incorporating useful exergy opens up a new route for investigating such effects. Following Heun et al. (2017), Brockway et al. (2017) estimate aggregate three-input CES production functions for the US, UK and China over the period 1980–2010, replacing primary energy with useful exergy. Following Saunders (2008), Brockway et al. (2017) derive an expression for the rebound effect that relies upon the cost share theorem. This leads to a mean estimate of the rebound effect 13 per cent for the US and UK, and 208 per cent for China. Or in other words, Brockway et al.’s (2017) results suggest improved exergy efficiency leads to significant exergy savings in the US and UK, but increased exergy consumption in China.

The confidence intervals on Brockway et al.’s (2017) estimates are large, and the method has other limitations such as the arbitrary choice of nesting structure and the continued reliance upon the cost share theorem. However, the use of an exogenous, thermodynamic measure of energy efficiency ($\varepsilon_{PU}$) represents an important step forward – and points the way to further work in this area.
Including useful exergy within energy–economy models

The next stage is to incorporate useful exergy into whole-systems models of the economy and to use these to develop projections of future economic growth and energy consumption. Such models can overcome some of the limitations of aggregate and sectoral production functions and can better capture the dynamic feedbacks that drive economic growth (Ayres and van den Bergh, 2005). Widely used computable general equilibrium (CGE) models are insufficient for this purpose, since they embody many of the problematic assumptions of orthodox theory and assume rather than estimate many of the parameters (Sorrell, 2014). A more promising approach is to employ macroeconometric models, consisting of a group of simultaneous equations that represent key macroeconomic relationships. These include identities (such as GDP being equal to the sum of public and private consumption, investment and exports) and behavioural relationships estimated from historical data (Fair, 1984).

The MAcroeconomic Resource CONsumption model (MARCO–UK) is the first attempt to integrate useful exergy within such a model and currently consists of 30 identities and 27 behavioural equations estimated from UK data over the period 1971–2013 (Sakai et al., 2018). Both useful exergy ($B_U$) and final to useful exergy efficiency ($\varepsilon_{FU}$) are endogenous variables, with the former being estimated from its lagged value, GDP and quality-adjusted labour and capital. Useful exergy is also an explanatory variable for other variables, including consumption, investment, exports, labour supply and final energy consumption. Capital and exergy are specified as complementary, with one being required to activate the other. Constructing the model in this way captures both the drivers of improvements in $\varepsilon_{FU}$ and the contribution of those improvements to output growth – which occur through both the consumption and production side of the economy. For example, lower-priced useful exergy improves productivity and stimulates increased production through additional capital investment.

MARCO-UK allows the development of counterfactual scenarios in which the values of key variables are held at their base year values – thereby allowing the contribution of those variables to economic growth to be estimated. Initial results (Figure 8.6) suggest that improvements in final to useful exergy efficiency have contributed one-quarter of UK economic growth since 1971 – comparable in scale to that contributed by capital investment. In contrast, increases in labour inputs are estimated to have contributed only 10 per cent of the observed growth. Put another way, the results suggest that improved energy efficiency has played a far more important role in driving UK economic growth than is traditionally assumed.

These results are provisional and require further analysis and development. But the MARCO-UK framework is flexible and can be extended in a variety of ways, including to other countries.
Conclusions, future directions and policy recommendations

Orthodox economics provides the lens through which most researchers and policymakers view the world, but that lens may obscure or distort some important features – such as the critical role of energy in economic production. The neglect of energy derives from the foundational assumptions of orthodox economics and could lead to misleading policy recommendations.

This neglect of energy has long been criticised, but alternative approaches lack a coherent theoretical framework and methodological approach. The developments described in this chapter could provide a more robust alternative, based around the concept of useful exergy. As illustrated above, a number of research topics are currently being pursued and the initial results suggest that improved exergy efficiency is a key driver of economic growth. However, the research is at an early stage and is handicapped by lack of data, and the unfamiliarity of the exergy concept inhibits wider acceptance.

Nevertheless, research is progressing in a number of directions, including: disaggregating the MARCO-UK model to the sector level; extending the co-integration approach to other countries and functions (Santos et al., 2018); extending the exergy framework to incorporate passive systems and energy services (Cullen et al., 2011); investigating the relationships between useful exergy, energy services and human needs (Brand-Correa and Steinberger, 2017); and incorporating useful exergy with input–output frameworks (Heun et al., 2018).
Taken together, these have the potential to significantly improve our understanding of the relationship between energy use and economic activity, and thereby the feasibility of absolute decoupling.

The policy implications of this work are mixed: if improved energy efficiency is essential to economic growth it should be given much greater policy support; but if rebound effects are large the environmental benefits of those improvements may be less than anticipated. But such conclusions are likely premature: what is more important – given the imperative of accelerated energy-GDP decoupling – is the willingness to question established assumptions, and to explore alternative ways of understanding the role of energy in the economy. Exergy economics is a step in that direction.

Acknowledgements

This research was funded by the United Kingdom’s Engineering and Physical Sciences Research Council (EPSRC) through grants to the Centre on Innovation and Energy Demand (EPSRC award EP/K011790/1) and the Centre for Industrial Energy, Materials and Products (EPSRC award EP/N022645/1). In addition, Paul Brockway’s time was funded as part of the research programme of the UK Energy Research Centre (UKERC), supported by the UK Research Councils under EPSRC award EP/L024756/1.

Notes

1 However, when TFP is estimated as a residual it also reflects measurement error and other factors such as omitted variables.

2 The exergy content of a quantity of materials is the amount of exergy required to produce the material from a reference environment by reversible processes.

3 The exergy and energy content of electricity are identical. But there are different ways of accounting for the primary energy/exergy content of nuclear and renewable electricity sources and no consensus on the preferred approach (Johansson et al., 2012).

4 Such as the requirement that the marginal productivity of an input should decline when the use of that input increases (Saunders, 2008).

5 A closer examination reveals an increase in useful exergy intensity after the Second World War in all countries except Portugal, followed by a modest decline after 1970. The first period coincides with the ‘golden years’ of post-war economic growth, while the second begins around the time of the first oil crisis. But taking the period as a whole, Serrenho et al. (2016) find no evidence of a time trend at the 5 per cent significance level.

6 The LINEX function was first introduced by Kümmel (1982) and imposes additional constraints on the allowed combinations of inputs.

7 A (KL)E CES function is specified as: $Y = \varphi e^{\int \delta_i \left[ \delta K_i^{\rho_i} + (1 - \delta_i) L_i^{\rho_i} \right] \rho_i + (1 - \delta_i) E_i \delta_i^{-1},$ where $\rho$ and $\rho_i$ define the ease of substitution between inputs, $\delta$ and $\delta_i$ define the contribution of each input to economic output, $\lambda$ defines the rate of productivity growth and $\varphi$ is a scaling factor. Estimation involves obtaining values for these six parameters.
8 Santos et al. (2018) assume a simpler ‘Cobb Douglas’ production function of the form: \( Y = ae^{\alpha_i K} K^\lambda K^\eta L^\eta \). The \( \alpha_i \) terms define the partial output elasticity of each input and \( \lambda \) defines the rate of productivity growth. The Cobb Douglas was chosen because it can be straightforwardly related to the co-integration specification, but it is restrictive because it assumes a unitary elasticity of substitution between each variable.

9 With time series data it is common for one or more of the variables to be non-stationary, creating the risk of ‘spurious regressions’. But it is possible for two or more non-stationary variables to be co-integrated, meaning that certain linear combinations of these variables are stationary and that there is a stable long-run relationship between them.

10 This relies upon Granger causality tests. A time series \((x_t)\) is said to ‘Granger cause’ another time series \((y_t)\) if the prediction of \(y\) is improved by the inclusion of past values of \(x\) in addition to past values of \(y\). Granger causality tests are designed to show whether one variable can meaningfully be described as dependent variable and the other as independent, or whether the relationship is bidirectional, or whether no relationship exists (Stern, 2011). This is test of ‘statistical precedence’ rather than causality as normally understood, since the fact that A precedes B need not necessarily mean that A causes B. For example, a meteorological forecast of rain can be shown to Granger cause rain!

11 Brockway et al. (2017) define energy rebound as: \( R = 1 + \eta_{ex} (B_P) \), where \( \eta_{ex} (B_P) \) is the elasticity of primary exergy consumption with respect to primary to useful exergy efficiency. They use the implicit function and cost share theorems to derive an expression for \( \eta_{ex} (B_P) \).

References


