Energy, Economic Growth and Environmental Sustainability: Five Propositions

Steven Sorrell

Sussex Energy Group, SPRU—Science and Technology Policy Research, University of Sussex, Sussex House, Brighton, BN1 9QE, UK; E-Mail: s.r.sorrell@sussex.ac.uk; Tel.: +44-1-273-877-067; Fax: +44-1-273-685-865

Received: 1 May 2010; in revised form: 7 June 2010 / Accepted: 9 June 2010 / Published: 18 June 2010

Abstract: This paper advances five linked and controversial propositions that have both deep historical roots and urgent contemporary relevance. These are: (a) the rebound effects from energy efficiency improvements are significant and limit the potential for decoupling energy consumption from economic growth; (b) the contribution of energy to productivity improvements and economic growth has been greatly underestimated; (c) the pursuit of improved efficiency needs to be complemented by an ethic of sufficiency; (d) sustainability is incompatible with continued economic growth in rich countries; and (e) a zero-growth economy is incompatible with a fractional reserve banking system. These propositions run counter to conventional wisdom and each highlights either a “blind spot” or “taboo subject” that deserves closer scrutiny. While accepting one proposition reinforces the case for accepting the next, the former is neither necessary nor sufficient for the latter.

Keywords: rebound effect; steady-state economy; monetary reform

1. Introduction

This paper questions the conventional wisdom underlying climate policy and argues that some long-standing and fundamental questions regarding energy, growth and sustainability need to be reopened. It does so by advancing the following propositions:
1. The rebound effects from energy efficiency improvements are significant and limit the potential for decoupling energy consumption from economic growth.

2. The contribution of energy to productivity improvements and economic growth has been greatly underestimated.

3. Sustainability requires both improved efficiency and a principle of “sufficiency”.

4. Sustainability is incompatible with continued economic growth in rich countries.

5. A zero-growth economy is incompatible with a fractional reserve banking system.

Although none of these propositions are new, they each run counter to conventional wisdom and highlight either blind spots or taboo subjects that deserve much closer scrutiny. This applies in particular to Proposition 5 which addresses an important subject that is almost entirely ignored within the sustainability literature. While accepting one proposition reinforces the case for accepting the next, the former is neither necessary nor sufficient for the latter. Also, while the focus of this paper is energy use and carbon emissions, the conclusions are equally relevant to (and informed by) broader resource and environmental constraints. The following sections summarise the arguments for each proposition in turn.

2. Rebound Effects are Significant and Limit the Potential for Decoupling Energy Consumption from Economic Growth

It is commonly assumed that historical improvements in energy efficiency have reduced energy consumption below what it would have been without those improvements—although since the “counterfactual” cannot be observed we can never be sure. Nevertheless, such improvements have clearly failed to reduce energy consumption in absolute terms. While the aggregate energy intensity of industrial economies has steadily fallen over the last century, energy consumption has continued to rise, along with the associated carbon emissions.

The most common explanation for the failure to decouple energy consumption from economic growth is that we haven’t tried hard enough: energy and carbon prices remain too low and policies to encourage energy efficiency are often small-scale, under-funded, poorly designed and/or ineffectual. In this view, the appropriate solution is to reinforce these policies—namely, to introduce more regulations, standards and financial support alongside the pricing of carbon emissions.

However, an alternative explanation for the failure to reduce energy consumption is that many of the potential energy savings have been “taken back” by various behavioural responses which are commonly grouped under the heading of rebound effects. While generally neither anticipated nor intended, these effects reduce the size of the energy savings achieved. An example of a rebound effect would be the driver who replaces a car with a fuel-efficient model, only to take advantage of its cheaper running costs to drive further and more often. Some heretics even argue that these effects lead to increased energy demand over the long term—an outcome that has been termed “backfire” [1,2]. If this is the case, non-price measures to encourage energy efficiency could actually increase carbon emissions.

Since energy efficiency improvements reduce the marginal cost of energy services such as travel, the consumption of those services may be expected to increase, thereby offsetting some of the predicted reduction in energy consumption. This so-called direct rebound effect was first studied by
Khazzoom [3] and has since been the focus of much research [4-6]. But even if there is no direct rebound effect for a particular energy service (e.g., even if consumers choose not to drive any further in their fuel efficient cars), there are a number of other reasons why the economy-wide reduction in energy consumption may be less than simple calculations suggest. For example, the money saved on motor-fuel consumption may be spent on other goods and services that also require energy to provide. Depending upon where the energy efficiency improvement takes place, these so-called indirect rebound effects can take a number of forms, including increases in the output of particular sectors, shifts towards more energy-intensive goods and services and increases in energy consumption as a result of lower energy prices and more rapid economic growth [7]. While judged negatively from a climate policy perspective, such effects also increase real income and generally improve welfare.

The overall or economy-wide rebound effect from an energy efficiency improvement represents the sum of these direct and indirect effects and is normally expressed as a percentage of the expected energy savings. Hence, an economy-wide rebound effect of 20% mean that one fifth of the potential energy savings are “taken back” through one or more of the above mechanisms while a rebound effect that exceeds 100% means that the energy efficiency improvements increase overall energy consumption. This possibility was first suggested by Jevons [8] and is commonly termed “Jevons Paradox”.

Rebound effects need to be defined in relation to particular measures of energy efficiency (e.g., thermodynamic, physical, economic), to relevant system boundaries for both the measure of energy efficiency and the change in energy consumption (e.g., device, firm, sector, economy) and to a particular time frame [7]. Disputes over the size and importance of rebound effects result in part from different choices for each of these variables [7,9]. Rebound effects may be expected to increase over time as markets, technology and behaviour adjusts. For climate policy, what matters is the long-term effect on global energy consumption from the adoption of new technologies.

Quantification of rebound effects is hampered by inadequate data, unclear system boundaries, endogenous variables, uncertain causal relationships, transboundary effects and complex, long-term dynamics such as changing patterns of consumption. In a comprehensive review, Sorrell and Dimitropoulos [5,6] found that estimates of the direct rebound effect for household energy services in the OECD were typically less than 30% and were expected to decline in the future as demand saturates and incomes increase. However, these effects have only been studied over limited time periods and the methods used have only captured a portion of the relevant effects [5]. Direct rebound effects may also be larger for low-income groups, for households in developing countries and (most importantly) for producers.

Quantification of indirect and economy-wide rebound effects is very challenging, but some insight may be gained from theoretical models [2,10-12] and from energy-economic models of the macroeconomy [13,14]. The available studies relate solely to energy efficiency improvements by producers and show that the economy-wide rebound effect varies widely depending upon the sector in which the energy efficiency improvement takes place. All the studies conducted to date estimate economy-wide effects in excess of 30% and several predict backfire [7]. Moreover, these estimates do not take into account the amplifying effect of any associated improvements in the productivity of capital, labour or materials, although in practice these appear to be very common [10,15]. Many, if not most improvements in energy efficiency are the byproduct of broader improvements in product and
process technology and even dedicated investments to improve energy efficiency frequently have wider benefits [9,10]. For example, Worrell et al. [15] found that the average payback period from 52 industrial energy efficiency projects fell from 4.2 years to 1.9 years when non-energy benefits were included. Since these additional cost savings will also contribute to additional energy consumption, this implies that the rebound effect from new technologies need not necessarily be small just because the share of energy in total costs is small—and that “win-win” opportunities will have the largest rebound effects [9]. Moreover, Brookes [1] has argued that improvements in energy productivity are normally associated with proportionally greater improvements in total factor productivity, with the result that energy consumption is increased.

A key question is whether economic growth is the cause of increased energy consumption and/or improved energy efficiency, or whether increased energy consumption and/or improved energy efficiency is a cause of the growth in economic output. These relationships are difficult to establish empirically and econometric studies of “Granger causality” give inconsistent results [9,16,17]. In practice, there is likely to be a synergistic relationship between these variables, with each causing the other as part of numerous positive feedback mechanisms [18,19]. Various historical examples can be cited in support of this [1,2,20], including the experience with steam turbines during the Industrial Revolution (Figure 1). Jevons [8] argued that the early Savory engine for pumping floodwater out of coal mines “…consumed no coal because it rate of consumption was too high”. It was only with the subsequent technical and efficiency improvements by Watt and others that steam engines became widespread in coal mines, facilitating greater production of lower cost coal which in turn was used by comparable steam engines in a host of applications. One important application was to pump air into blast furnaces, thereby increasing the blast temperatures, reducing the quantity of coal needed to make iron and reducing the cost of iron. Lower cost iron, in turn, reduced the cost of steam engines, creating a positive feedback cycle. It also contributed to the development of railways, which lowered the cost of transporting coal and iron, thereby increasing demand for both.

Rosenberg [21] cites the comparable example of the Bessemer process for steel-making which: “…..was one of the most fuel saving innovations in the history of metallurgy [but] made it possible to employ steel in a wide variety of uses that were not feasible before Bessemer, bringing with it large increases in demand. As a result, although the process sharply reduced fuel requirements per unit of output, its ultimate effect was to increase….the demand for fuel” [21].

The low cost Bessemer steel initially found a large market in the production of steel rails, thereby facilitating the growth of the rail industry, and later in a much wider range of applications including automobiles. However, the mild steel produced by the Bessemer process is a very different product to wrought iron and suitable for use in a much wider range of applications. Hence, for both steelmaking and steam engines, improvements in the energy efficiency of production processes were deeply and perhaps necessarily entwined with broader improvements in process and product technology.

Brookes [1] cites the example of US productivity growth during the 20th century. Energy prices were falling in real terms for much of this period with the result that energy substituted for other factors of production and increased aggregate energy intensity. But these substitution effects were more than outweighed by the technological improvements facilitated by the availability of high-quality energy sources which greatly improved the overall productivity of the US economy—for example, in transforming the sequence, layout and efficiency of manufacturing through the introduction of electric
motors [22]. This meant that economic output increased much faster than energy consumption, owing to the greater productivity of capital and labour. The net result was to produce *falling* energy intensity (as measured by the energy/GDP ratio) alongside *rising* energy consumption—as Jevons’ predicted. Polimeni [23] provides econometric evidence for this process for a number of countries and time periods, lending support to the argument that this is a universal phenomenon.

**Figure 1.** Energy efficiency, positive feedbacks and economic growth in the 19th century.

These historical examples relate to energy efficiency improvements in the early stages of development of energy-intensive process technologies (*i.e.*, steam turbines, Bessemer process, electric motors) that are producing goods that have the potential for widespread use in multiple applications. The same consequences may not follow for energy efficiency improvements in mature and/or non-energy-intensive process technologies that are producing goods that have a relatively narrow range of applications, or for energy efficiency improvements in consumer technologies. But backfire seems more likely to occur over the long-term following improvements in the energy efficiency of “general-purpose technologies” such as steam turbines, lighting, motor vehicles and computers—particularly when these are used by producers and when the improvements occur at an early stage of development and diffusion [24]. These technologies have transformational effects, such as the growth of existing sectors, the emergence of new processes, products and services and changes in infrastructure, employment and consumer preferences. Moreover, such “general-purpose technologies” dominate overall energy consumption.

In sum, rebound effects will make energy efficiency improvements less effective in reducing overall energy consumption than is commonly assumed. This could limit the potential for decoupling carbon
emissions from economic growth, since the contribution from improved energy efficiency will be less than expected—although by precisely how much remains unclear. In principle, increases in energy prices should reduce the magnitude of such effects by offsetting the cost reductions from improved energy efficiency. This leads to the policy recommendation of raising energy prices through either energy/carbon taxation or emissions trading schemes. Cap and trade schemes are particularly attractive since they focus directly upon the desired ends (e.g., reduced carbon emissions) rather than a potentially problematic means to achieve those ends (e.g., improved energy efficiency) [25]. But while such measures will induce substitution and technical change, their impact on total factor productivity and economic growth remains disputed [9,26,27]. This leads to the second proposition, discussed below.

3. The Contribution of Energy to Productivity Improvements and Economic Growth Has Been Greatly Underestimated

Many of the arguments in favour of Jevons Paradox focus on the source of productivity improvements and the relationship between energy consumption and economic growth [9]. Orthodox and ecological economics provide very different perspectives on this question with correspondingly different conclusions on the potential for decoupling.

Orthodox economic models imply that the economy is a closed system within which goods are produced by capital and labour and exchanged between consumers and firms. While such models have been extended to include natural resources, ecosystem services and wastes, these commonly remain secondary concerns. Economic growth is assumed to derive from a combination of increased capital and labour inputs, changes in the quality of those inputs (e.g., better educated workers) and technical change [28,29]. Both increases in energy inputs and improvements in energy productivity are assumed to make only a minor contribution to economic growth, largely because energy accounts for only a small share (typically <5%) of total input costs. It is also assumed that capital and labour will substitute for energy should it become more expensive. From this perspective, improvements in energy productivity are unlikely to have a significant impact on overall productivity, so the corresponding rebound effects should be relatively small. Hence, there seems to be no reason why energy consumption could not be substantially decoupled from economic growth.

Ecological economists consider that the orthodox models ignore how economic activity is sustained by flows of high quality energy and materials which are then returned to the environment in the form of waste and low temperature heat [30]. The system is driven by solar energy, both directly and embodied in fossil fuels, and since energy cannot be produced or recycled it forms the primary input into economic production. In contrast, labour and capital represent intermediate inputs since they cannot be produced or maintained without energy. So far from being a secondary concern, energy becomes the main focus of attention.

Energy carriers differ in terms of their flexibility of use, amenability to storage, energy density, capacity to do physical work (captured by the thermodynamic concept of “exergy”) and economic productivity—as reflected by their differing price per kWh [31]. Electricity is generally considered the highest quality energy carrier, followed by natural gas, oil, coal, wood and other biomass. Since high-quality energy carriers are more productive, they should be given more weight in aggregate
measures of energy consumption [32]. When this is done, aggregate measures of energy efficiency are found to be improving more slowly than is commonly supposed [31,33,34]. This conclusion is reinforced when the energy embodied in traded products is accounted for [35], together with the energy used in international aviation and shipping [36]. Failure to make these adjustments can lead the extent of decoupling to be overstated and the source of that decoupling obscured [9].

Ecological economists claim that the massive improvements in labour productivity over the last century have largely been achieved by providing workers with increasing quantities of high quality energy, both directly and indirectly as embodied in capital equipment and technology [37-39]. This has facilitated technical change, enhanced the productivity of capital and labour, and allowed more economic output to be produced for each unit of energy input. They further claim that increases in energy inputs have contributed more to the growth in economic output than is suggested by the small share of energy in total costs. Such increases lead to more “useful work” being obtained from energy conversion devices, but this may also be achieved by shifting towards higher quality energy inputs and by improving thermodynamic conversion efficiency [9]. The implication of this line of argument is that improvements in energy efficiency should have a disproportionate impact on economic output which in turn will increase overall energy consumption. Hence, the contribution to economy-wide rebound effects from this source could be large.

Ecological economists also claim that the indirect energy consumption associated with capital and labour (e.g., the energy required to manufacture thermal insulation) limits the extent to which they can substitute for energy in economic production [40]. The energy embodied in capital goods is commonly overlooked by studies that estimate energy-saving potentials at the level of individual sectors and then aggregate the results to economy as a whole. Furthermore, many energy-economic models assume a greater potential for substitution that is allowed for by physical laws [41]. Hence, from an ecological perspective, the potential for decoupling energy consumption from economic growth appears more limited.

Empirical support for the ecological perspective may be derived from a variety of sources. For example, Kaufmann [42,43] uses econometric analysis to argue that historical reductions in energy/GDP ratios owe more to shifts towards high-quality fuels than to technological improvements in energy efficiency. He shows how these, together with changes in the proportion of household income spent on fuel, price-induced factor substitution and a shift towards a service-based economy, explain most of the reduction in US energy intensity since 1929. Once this is controlled for, Kaufmann finds no statistically significant evidence for the “energy saving technical change” that is traditionally assumed in energy-economic models [44,45]. Similarly, a number of studies suggest that the indirect energy consumption associated with labour and capital inputs constitutes a significant portion of the direct energy savings from energy efficiency improvements—for example, as much as 83% in the US forest products industry between 1954 and 1984 [46]. Historical experience also provides little support for the claim that increases in income will lead to declining energy consumption [47,48]. While the income elasticity of energy consumption may be both declining and less than one in OECD countries, there is no evidence that it is negative or soon to become negative. Again, neglect of changes in fuel mix and energy prices, together with the energy embodied in traded goods, may have led earlier studies to draw misleading conclusions regarding the source and extent of decoupling [42].
Table 1. Orthodox and ecological perspectives on energy, productivity and economic growth.

<table>
<thead>
<tr>
<th></th>
<th>Orthodox view</th>
<th>Ecological view</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main source of productivity</td>
<td>Exogenous or endogenous technical change</td>
<td>Increasing availability of high-quality energy, both directly and embodied in capital equipment and technology</td>
</tr>
<tr>
<td>improvements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marginal productivity of energy inputs</td>
<td>Proportional to share of energy in the value of output</td>
<td>Greater than the share of energy in the value of output</td>
</tr>
<tr>
<td>Input substitution in production</td>
<td>Scope for substitution indicated by substitution elasticities estimated at the sector level</td>
<td>Scope for substitution overestimated by substitution elasticities estimated at the sector level, since these neglect embodied energy</td>
</tr>
<tr>
<td>Decoupling of energy consumption from GDP</td>
<td>Decoupling has already occurred in OECD economies and there is considerable scope for further decoupling</td>
<td>Conventional measures of energy inputs overstate the amount of decoupling. A strong link exists between quality adjusted energy use and economic output and will continue to exist, both temporally and cross-sectionally.</td>
</tr>
<tr>
<td>Economy-wide rebound effect</td>
<td>Likely to be small</td>
<td>Likely to be large</td>
</tr>
</tbody>
</table>

Source: Clevleand et. al. [38]; Ayres and Warr [19].

Ecological economists have also developed alternatives to conventional models of economic growth which depart from the traditional assumption that the productivity of each input is proportional to the share of that input in the value of output [18,19,37,39]. In contrast to orthodox theory, these models reproduce historical trends in economic growth extremely well, without attributing any role to technical change. Instead, increases in energy inputs provide the primary explanation for productivity improvements and economic growth [49]. Of particular interest is Ayres and Warr [19], who combine historical data on the “exergy” content of fuel inputs (i.e., the ability to perform useful work) and second-law thermodynamic conversion efficiencies to develop a time series of useful work inputs to the US economy over the past century. They show that the useful work obtained from fuel resources has grown much faster than the consumption of fuels themselves, owing to substantial improvements in thermodynamic conversion efficiencies. By including useful work in their aggregate production function, rather than primary energy, Ayres and Warr [19] obtain an extremely good fit to US GDP trends over the past century, thereby eliminating the need for a multiplier for technical change. In this model, improvements in thermodynamic conversion efficiency have a dramatic effect on economic output and act as the primary “engine” of economic growth [50]. This in turn implies that rebound effects may be larger than orthodox theory suggests [9].

These studies challenge conventional theory and raise concerns about the potential for decoupling carbon emissions from economic growth since the contribution from improved energy efficiency may be less than expected. However, the empirical evidence remains patchy and sometimes contradictory [9]. For example, the estimated magnitude and sign of “energy-saving technical change” varies between different sectors, countries, time periods, technologies and econometric specifications and the results are sensitive to whether and how energy quality is accounted for [9,51-53]. Similarly, the different variants of “ecological growth models” rely upon an unusual and oddly behaved production function, provide results that are difficult to reconcile with each other and appear vulnerable to bias from a number of sources [9]. While the orthodox approach implies that capital,
labour and energy inputs have independent and additive effects on economic output, with any residual increase being attributed to technical change, the ecological approach implies that capital, labour and energy are interdependent inputs that have synergistic and multiplicative effects on economic output and that the increased availability of high-quality energy sources has provided a necessary condition for most historical improvements in economic productivity. As Toman and Jemelkova [54] observe: “……when the supply of energy services is increased, there is not just more energy to be used by each skilled worker or machine; the productivity with which every unit of energy is used also rises. If all inputs to final production are increased in some proportion, final output would grow in greater proportion because of the effect on non-energy inputs” [54].

In sum, orthodox analysis implies that rebound effects are small, improvements in energy productivity make a relatively small contribution to economic growth and decoupling is both feasible and cheap. In contrast, the ecological perspective suggests that rebound effects are large, improvements in energy productivity make an important contribution to economic growth and decoupling is both difficult and expensive. While the empirical evidence remains both suggestive and equivocal, the ecological perspective highlights some important blind spots within orthodox theory that are reflected in the design of economic models used to underpin climate policy. If this perspective is correct, both the potential for and continued reliance upon decoupling needs to be questioned.

4. Sustainability Requires Both Improved Efficiency and a Principle of “Sufficiency”

The achievement of much higher levels of energy efficiency seems essential if carbon emissions are to be radically reduced. But given the difficulties highlighted above, this may not be sufficient to meet ambitious carbon targets.

The preferred strategy to achieve sustainability is consuming more efficiently, which implies reducing the environmental impacts associated with each good or service. But the extent to which this is successful will depend upon the size of any associated rebound effects. If these are significant, energy use may not be reduced as much as expected and in some circumstances could increase. Despite this, most OECD countries pay little attention to such possibilities and offer few options for mitigating the undesirable consequences.

A related and complementary strategy is consuming differently, which implies shifting towards goods or services with a lower environmental impact. This could involve purchasing “greener products”, increasing expenditure on “services” rather than manufactured goods, or entering into arrangements such as energy service contracting and car sharing schemes. These strategies are frequently cited as having environmental benefits, although the empirical evidence to support such claims is often lacking. For example, a shift towards services and away from products may increase energy use, particularly if it involves higher standards of service, the extensive use of transport, or the construction of resource-intensive infrastructure such as telecommunications networks. In a recent review, Heiskanen and Jalas [55] concluded that the environmental benefits of product-to-service arrangements are modest at best while Suh [56] estimates that a shift to service-oriented economy could actually increase carbon emissions (although reduce the carbon intensity of GDP) owing to the heavy reliance of services upon manufactured commodities. Also, much of the observed decoupling in developed countries has been achieved by outsourcing manufacturing to developing countries. For
example, official figures indicate a 5% reduction in the UK’s carbon emissions between 1990 and 2004, but this changes to a 15% increase when the emissions embodied in international trade are accounted for [57].

Given the potential limitations of consuming efficiently and consuming differently, it seems logical to examine the potential of a third option—simply consuming less [58-60]. The key idea here is **sufficiency**, defined by Princen [60] as a social organising principle that builds upon established notions such as restraint and moderation to provide rules for guiding collective behaviour. The primary objective is to respect ecological constraints, although most authors also emphasise the social and psychological benefits to be obtained from consuming less.

While Princen [60] cites examples of sufficiency being put into practice by communities and organisations, most authors focus on the implications for individuals. They argue that “downshifting” can both lower environmental impacts and improve quality of life, notably by reducing stress and allowing more leisure time. This argument is supported by an increasing number of studies which show that reported levels of happiness are not increasing in line with income in developed countries [61,62]. As Binswanger [63] observes: “…the economies of developed countries turn into big treadmills where people try to walk faster and faster in order to reach a higher level of happiness but in fact never get beyond their current position. On average, happiness always stays the same, no matter how fast people are walking on the treadmills”.

It is possible that an ethic of sufficiency could provide a means of escaping from such treadmills while at the same time contributing to environmental sustainability.

According to Alcott [64], sufficiency implies both environmental motivation and purchasing power: “….those who are to alter their behaviour towards less consumption must be able to consume. Their purchasing power either remains unused or is itself reduced through working and earning less”. Hence, the concept appears primarily applicable to the wealthy and is of little relevance for those suffering from absolute or relative poverty. But with increasing income inequality, spiraling levels of personal debt and the fallout from the 2008 economic recession, the proportion of people able to exercise such a choice is likely to be falling.

Adopting sufficiency as a guiding principle would require a major change in lifestyles. While most people would acknowledge that “quality of life” is not solely about material consumption, numerous psychological, economic and cultural obstacles can make it difficult for individuals to reduce current levels of consumption. For example, many people are partially “locked-in” to current consumption patterns owing to factors such as land-use patterns and physical infrastructures (which constrain choice in areas such as travel), the rapid obsolescence of consumer goods and the difficulty of reducing the number of hours of work. Consumption habits are also shaped by factors such as the search for status through the acquisition of symbolic “positional goods” (which creates a never-ending zero sum game), the rapid adaptation of aspirations to higher income levels (thereby reducing the happiness associated with that income), and what Jackson [65] calls the “almost perfect fit” between the search by producers for newer, better and cheaper products and the corresponding desire by consumers for novelty [58,63]. In this context, the practice of sufficiency requires a minimum level of financial security, deeply held values and considerable determination. If adopted successfully by enough individuals, it could demonstrate a viable and attractive alternative to consumerism and begin to shift social attitudes in a number of areas. But since “luxury” consumption fulfils so many deep
psychological needs, the widespread adoption of sufficiency appears unlikely to develop through voluntary action alone.

If the balance of factors encouraging or discouraging sufficiency were to change in favour of the former, it should become complementary to efficiency as a means of both improving quality of life and adapting to tightening ecological constraints. But to move the “sufficiency ethic” from the marginal to the mainstream is likely to require collectively agreed objectives, priorities, procedures and constraints that are institutionalised through government action in some form. This means that the most important agent of change is likely to be individuals acting as citizens in the political process rather than as “downshifting” consumers. Also, as Alcott [64] has pointed out, sufficiency is not immune from rebound effects, even when pursued at a national level. A successful “sufficiency strategy” will reduce the demand for energy and other resources, thereby lowering prices and encouraging increased demand by others which will partly offset the energy and resource savings. While this “sufficiency rebound” would improve equity in the consumption of resources, it would nevertheless reduce the environmental benefits of the sufficiency measures. But since the global “ecological footprint” already exceeds sustainable levels in many areas the global consumption of resources needs to shrink in absolute terms [66]. To achieve this and to effectively address problems such as climate change, will require collective agreement on ambitious, binding and progressively more stringent targets at both the national and international level. This raises a further and fundamental challenge for rich countries which is addressed by the next proposition.

5. Sustainability Is Incompatible with Continued Economic Growth in Rich Countries

The preceding arguments highlight a conflict between reducing energy consumption and carbon emissions in absolute terms while at the same time continuing to grow the economy. Recognising the importance of rebound effects and the role of energy in driving economic growth therefore re-opens the debate about limits to growth. This debate is long-standing and multifaceted, but a key point is that the goal of economic development should not be to maximise GDP but to improve human well-being and quality of life. Material consumption is merely a means to that end and GDP is merely a measure of that means. GDP measures the costs rather than the benefits of economic activity and is only a proxy for welfare under very restrictive conditions [67]. It is likely that, beyond a certain level increased material consumption will reduce well-being, since the (typically unmeasured) social and environmental costs will exceed the benefits. As Ekins argues [68], human well-being is not determined solely by the consumption of goods and services but also by “human capital” (e.g., health, knowledge), “social capital” (e.g., family, friends and social networks) and “natural capital” (i.e., ecosystems and the services they provide)—none of which are necessarily correlated with GDP. Attempts to value these contributions through the use of alternative measures of economic progress typically find that “well-being” is not improving or even declining in rich countries, despite increases in GDP [69,70]. Hence, while growth in per capita income is likely to improve well-being in developing countries, the same may not be true for developed countries.

Table 2 compares this ecological perspective on economic development with the orthodox model. While many elements of this perspective are now influencing mainstream policy (for example, the pricing of environment externalities), the goal of maximising GDP largely remains unchallenged.
Table 2. Different models of economic development.

<table>
<thead>
<tr>
<th>Primary policy goal</th>
<th>Orthodox view</th>
<th>Ecological view</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Economic growth as measured by GDP. Growth should allow the solution of other problems.</td>
<td>Development in the sense of improved quality of life. Growth has negative side-effects.</td>
</tr>
<tr>
<td>Primary measure of progress</td>
<td>Gross Domestic Product (GDP)</td>
<td>Index of Sustainable Economic Welfare (ISEW) or some comparable indicator</td>
</tr>
<tr>
<td>Scale/carrying capacity</td>
<td>Not an issue because it is assumed that market could overcome resource limits via substitution and technical change</td>
<td>Primary concern since there is limited scope for substituting natural for man-made capital</td>
</tr>
<tr>
<td>Income distribution</td>
<td>Secondary concern. A “trickle down” policy (a rising tide lifts all boats)</td>
<td>Primary concern. Directly affects quality of life and is often made worse by economic growth</td>
</tr>
<tr>
<td>Economic efficiency/allocation</td>
<td>Primary concern, but generally including only marketed goods and services</td>
<td>Primary concern, but including both market and non-market goods and services. Human, natural and social capital must be valued</td>
</tr>
<tr>
<td>Role of “sufficiency”</td>
<td>Not recognised. More is always better.</td>
<td>Congruent with overall aims. More is not always better</td>
</tr>
</tbody>
</table>

Source: Adapted from Costanza [71].

Over the long term, continued economic growth can only be reconciled with environmental sustainability if implausibly large improvements in energy and resource efficiency can be achieved. This point is easy to demonstrate with the $I = P*A*T$ equation, which represents total environmental impact ($I$) as the product of population ($P$), affluence or income level ($A$) and technological performance or efficiency ($T$) [72]. In the case of climate change, $I$ could represent total carbon emissions, $A$ GDP per capita and $T$ carbon emissions per unit of GDP (itself a product of energy consumption per unit of GDP and carbon emissions per unit of energy consumption). The decoupling strategy seeks reductions in $T$ that will more than offset the increases in $P$ and $A$, thereby lowering $I$.

This equation can be used to examine the goal of keeping the long-term equilibrium rise in mean global temperature to below two degrees centigrade. According to Meinhausen et al. [73] a 75% probability of limiting global warming to 2 °C requires cumulative CO$_2$ emissions in the period 2000–2050 to be less than 1000 GtCO$_2$. Since ~350 GtCO$_2$ was emitted between 2000 and 2009, the remaining carbon budget for the period to 2050 is only 650 GtCO$_2$ (compared to emissions of ~35.7 GtCO$_2$ in 2006). Assume for the purpose of illustration that global emissions increase at 1%/year to 2015 and then decline exponentially to 2050. Then the required annual rate of emission reduction to meet this cumulative budget is 4.9%/year (for an 82% reduction on 2006 emissions). For comparison, the French nuclear programme and the UK’s “dash for gas” both achieved emission reductions of only 1%/year—excluding emissions from land-use change, international shipping and aviation as well as those embodied in traded goods. Delaying the emissions peak drastically increases the required speed and depth of emission reduction, as does anticipated changes in the global carbon cycle [36,74].
By 2050, the UN projects that the world’s population will be between 7.8 and 10.8 billion. Taking the mean estimate of nine billion (an average growth of 0.7%/year), this means that average per capita emissions in 2050 will need to be around 0.75 tCO$_2$/year implying an average reduction of ~5.6%/year. This compares to a global average of ~5.4 tonnes per capita in 2006 and 19 tonnes per capita in the US. If global per capita GDP continues to grow at approximately 1.4%/year, this means that emissions per unit of GDP will need to fall by at least 6.9%/year over this period (a total reduction of 92%). This compares with a 1.3%/year reduction over the period 1970–2000 and a 0.3%/year increase since 2000 [75]. If instead, carbon intensity continued to fall at only 1.3%/year, global emissions would be 55% higher by 2050. Emissions may peak earlier in rich countries, but if they are required to contribute a proportionally greater share of total reductions, rich countries may need to reduce their carbon intensities by more than 6.9%/year.

Emission reductions of this speed and scale were not adequately explored by the Stern Review, in part because the rapid increase in emissions since 2000 was ignored [36]. But there are good reasons to question their plausibility. While global carbon emissions per unit of GDP have fallen by around 30% since 1970, the rate of improvement has been declining and has recently reversed [76]. The historical improvements in this ratio owe much to the shift towards higher quality fuels such as natural gas, but accelerating resource depletion could lead to an increased reliance on lower quality, carbon-intensive fuels such as coal and non-conventional oil [77]. These have a lower “energy return on investment” than conventional fossil fuels, as do many renewable technologies [78], which implies that an increasing proportion of productive resources will be required to obtain energy, leaving less available for non-energy goods and services [79]. Such trends make the combination of high GDP growth and rapid emission reductions even less plausible.

The required speed of emission reduction could be slowed if the carbon budget was increased, but this would greatly increase the probability of serious and irreversible consequences, such as the loss of the Greenland and West Antarctic ice sheets [80]. Alternatively, the scenario could be made more plausible if lower or zero levels of per-capita GDP growth were considered acceptable for developed countries, but this runs counter to the objectives of all governments.

The required changes look even more challenging when rebound effects are considered. The $I = P*A*T$ equation implies that the right-hand side variables are independent of one another—or at least if any dependence is sufficiently small that it can be neglected. But in practice the variables are endogenous. So while a reduction in the economy-wide emission intensity ($T$) may have a direct effect in lowering emissions ($I$), it will also encourage economic growth ($A$), which in turn will increase emissions. Over the long term and up to a certain level of income, rising affluence ($A$) encourages higher population levels ($P$), which will further increase emissions ($I$). Hence, a change in $T$ will trigger a complex set of adjustments and the final change in emissions is likely to be lower than the IPAT identity suggests. This in turn, implies that greater changes in $T$ may be required to achieve a particular reduction in $I$.

Hence, in an increasingly “full” world, the goal of continued economic growth in the rich countries deserves to be questioned. This seems unlikely to improve the quality of life in those countries and is increasingly incompatible with resource and environmental constraints. But proposing the alternative of low, zero or even negative growth raises monumental problems of how such an economy could function. For example, how could rising aspirations be met with or without a major redistribution of
wealth? How could sufficient investment in areas such as low carbon technology be achieved? How could poverty be reduced in the South without growing markets for exports in the North? And so on. But if the notion of limits to growth remains taboo, these crucial questions will remain unexplored.

6. A Zero-Growth Economy Is Incompatible with a Fractional Reserve Banking System

Proposals for a zero-growth economy have had practically no influence on public policy because modern economies are structurally dependent upon continued GDP growth—and hence upon continued increases in the consumption of goods and services. As Hubbert [81] observes:

“…..During the last two centuries we have known nothing but exponential growth and in parallel we have evolved what amounts to an exponential growth culture, a culture so heavily dependent upon the continuance of exponential growth for its stability that it is incapable of reckoning with problems of non-growth”.

This “growth addiction” has social, psychological, cultural and economic dimensions that both reflect and reinforce the power of dominant interest groups. But as economies are presently structured, any slowing of the rate of economic growth has serious consequences for businesses (e.g., reduced profits, bankruptcies), individuals (e.g., foreclosures, unemployment), public services (loss of tax revenue, cuts) and politicians (e.g., loss of office), that all sectors of society have strong incentives to maximise economic growth. Set against this, concerns about long-term environmental sustainability and quality of life are easily overridden. Hence, the challenge is not simply to demonstrate the un-sustainability of the present model of economic development and the benefits of alternative models (Table 2) but also to propose ways in which the dependence of modern economies upon continued economic growth can be broken.

Recent studies by Victor [59] and Jackson [52] have explored the structural drivers of growth and the manner in which a zero-growth economy could function. While both provide valuable contributions, they also ignore what appears to be a key explanation of the “growth imperative”—namely the nature of modern monetary systems and the fact that most money is created by commercial banks as interest-bearing debt [82-84]. The implications of this are obscured by orthodox macroeconomics, although they are more prominent within the Austrian and Post Keynesian traditions [85-87] and were recognised by Schumpeter [88] and Keynes [89].

A purely private enterprise system can only function if companies can obtain sufficient profits which in turn requires that the selling price of goods exceeds the costs of production. This means that the selling price must exceed the spending power that has been distributed through payments to factor inputs. Hence, to ensure sufficient “aggregate demand” to clear the market, additional spending power is required from some other source [84]. In a purely private enterprise system, this normally derives from investment in new productive capacity which will increase the amount or quality of goods supplied, but only after some interval. Investment therefore serves the dual role of increasing productive capacity and creating additional demand to clear the market of whatever has already been produced [83,84].

Aggregate demand is commonly expressed as the product of the amount of money in circulation and the speed with which that money circulates through the economy [90]. Hence increases in aggregate demand require increases in the money supply or the speed of circulation or both. Increases
in the money supply contribute, in turn, to increases in aggregate output, the average price of goods and services or both.

The key issue is how the increase in the money supply is brought about. Governments could create the new money interest-free and spend it in to circulation in much the same way as coins and notes are created [91]. But instead, the bulk of the money supply is created by commercial banks who print credit entries into the bank accounts of their customers in the form of interest-bearing loans [82,92,93]. The loan appears as an asset on the bank’s balance sheet while the deposit appears as a liability. This system of “fractional reserve banking” has its origins in part in the practices of the early goldsmiths who made “loans” of a far greater quantity of gold that they actually held in their vaults (Box 1). This gave them substantial profits and allowed them to increase their claims on wealth, but also served the essential function of increasing purchasing power in a growing economy. This and similar practices (e.g., the use of bills of exchange for trade) contributed to the evolution of modern banking in which central banks control credit creation through imposing minimum reserve requirements, varying the total volume of bank reserves and acting as lender of last resort (Box 2).

**Box 1. The origins of modern banking.**

In mediaeval Europe, gold and silver coins were the primary medium of exchange. Merchants and landowners commonly deposited their coins with goldsmiths for safekeeping, receiving a receipt of deposit to certify ownership. To avoid the hassle of taking out gold to make a purchase, only to have it re-deposited later by the seller, the deposit receipts became accepted as a means of payment—*i.e.*, they effectively became a circulating medium of exchange. As a result, the gold tended to remain in the goldsmiths’ vaults for extended periods of time. The goldsmiths realised that they could make extra profits if they lent out the gold in the interim—an arguably fraudulent practice, since the receipts guaranteed that the gold was deposited whereas in fact it had been lent to a third party. The practice was extended further by lending out deposit receipts instead of gold—effectively printing money. This greatly increased profits, since goldsmiths could charge interest on the “loans” and be paid back in real purchasing power (or collateral), although they had nothing of substance to lend. This and similar practices contributed to the evolution of fractional reserve banking with individual banks issuing paper money with only a fraction of the deposits backed by commodities such as gold. Such systems were vulnerable to periodic losses of confidence and “runs on the bank” as depositors sought to redeem their receipts as commodities.

This situation appeared to change with the establishment of central banks that were given monopoly rights to print paper money. But banks continued to create credit, or “book money” in the form of deposit entries when providing customers with a loan. Effectively, these loans expanded the money supply while repayments contracted it. Since checkbook and subsequently electronic money has increasingly supplanted commodity or paper money, banks now create and allocate the majority of purchasing power in modern economies.

Source: Werner [82,94]; Hixson [95].

A crucial consequence of this system is that most of the money in circulation only exists because either businesses or individuals have gone into debt and are paying interest on their loans. While
Individual loans may be repaid, the debt in aggregate can never be repaid because this would remove more than 90% of the money supply from circulation [96]. The health of the economy is therefore entirely dependent upon the continued willingness of businesses and consumers to take out loans for either investment or consumption. Any reduction in borrowing therefore threatens to tip economies into recession.

**Box 2. Central bank control of credit creation.**

The volume of credit creation is influenced by central banks which require commercial banks to hold a percentage of their *demand deposits* in the form of cash or deposits at the central bank. These *reserves* are created by the central bank and form the *monetary base* or *high powered money*. Savings deposits are not subject to a reserve requirement. Credit creation is also influenced by international rules on the minimum ratio of a bank’s capital (*i.e.*, equity and retained earnings) to the risk-adjusted value of its assets (*i.e.*, reserves, government securities and loans). For example, banks may be required to have a reserve ratio of 10% and a capital ratio of 8%.

To illustrate the effect of minimum reserve requirements, assume Bank A receives a new deposit of £100. With a 10% reserve requirement, it is able to loan £90 and deposit the remaining £10 with the central bank as reserve. But since the loan does not remove purchasing power from the depositor, it is more accurately described as the creation of new purchasing power. The £90 will be re-deposited in Bank B, who will also be able to loan out 90% (£81) and keep the remainder as reserve. This process can continue through successive waves of lending, depositing and re-lending until up to £900 of new purchasing power is created [92]. An alternative way of viewing the process is to envisage Bank A depositing £100 as reserves at the central bank and granting £900 of new loans [82]. A less stringent reserve requirement would allow more purchasing power to be created and vice versa, provided the banks are able to meet their capital adequacy requirements. Repayment of the loans removes this purchasing power from circulation, although defaults do not.

The central banks control the level of reserves by buying or selling government securities from either the banks or the public using newly created, high powered money. For example, the central bank purchase of government securities should increase commercial banks deposits and hence reserves and thereby expand the money supply through the process indicated above. The ratio of demand deposits to high-powered money is termed the *money multiplier*. However, evidence suggests that changes in bank credit generally *precede* changes in high powered money, the reverse of what the multiplier theory suggests [97-100]. This is emphasized by the “endogenous money” school which claims that, in order to maintain stability in the financial system, central banks are unable to deny banks the level of reserves that they require [99-103]. Instead, central banks adjust the supply of reserves in order to control the interest rate at which banks borrow reserves from each other which in turn influences interest rates throughout the economy [104]. Hence, central banks focus primarily upon controlling the price rather than the quantity of money.

Individual loans need to be repaid with interest which provides banks with their main source of revenue. A large part of these interest payments will be recycled in the form of wages, dividends, and investments, thereby putting the money back into circulation. However, a portion will be retained as
bank capital in order to reduce the risk of insolvency, underpin further loans and meet capital adequacy requirements (Box 2). This money will be removed from circulation and will therefore be unavailable for the repayment of loans [83]. As a result, the only way that firms can make profits and borrowers can repay their loans in aggregate is if the volume of new borrowing exceeds the amount that is being withdrawn through both principal repayments and additions to bank capital [93]. In other words, total debt must increase. Binswanger [83] has explored the implications of this using a simple “circular flow” model of the macoconomy that builds upon a tradition of work in Post Keynesian economics [105-107]. Using a number of standard assumptions about the behaviour of banks, households and firms, he shows that:

“….In the long run, abstracting from business cycle fluctuations, capitalist economies can either grow (at a sufficiently high rate) or shrink if the growth rate falls below a positive threshold level. A zero growth economy is not feasible in the long run” [83].

Slow or negative growth will leave firms with lower profits and unused capacity, discouraging them from investing. Less investment will mean fewer loans being taken out and thus insufficient money entering into circulation to replace that being removed. And less money in circulation will mean that there is less available for consumers to spend, which will exacerbate the economic slowdown and cause more bankruptcies and unemployment. Hence, to avoid a damaging downward spiral, total debt and the total amount of money in circulation needs to rise each year which means that the value of goods and services bought and sold must also rise—either through inflation or higher consumption. The monetary system therefore creates a structural requirement for continued growth and increased consumption, whether or not this is consistent with the objectives of governments and the central bank.

One implication of this system is that economic growth depends upon ever increasing debts to banks, making the economy vulnerable to recession as a consequence of over-indebtedness [84]. Extension of credit to consumers or for speculative purposes encourages consumer or asset price inflation (“bubbles”) and reinforces the business cycle by expanding the money supply during booms and contracting it during slumps [108,109]. In contemporary circumstances, this process is magnified by the incentive structure of modern banking and the complex mechanisms to ensure against financial risk [110,111].

A second implication is that there is an increasing disconnect between debt and real physical wealth, leaving the system vulnerable to crisis due to the inevitable mismatch in the growth rates between the two. This point was first highlighted by Frederick Soddy [112] and has been revived by ecological economists: “….Debt grows at compound interest and as a purely mathematical quantity encounters no limits to slow it down. Wealth grows for a while at a compound interest, but having a physical dimension its growth sooner or later encounters limits…..Since wealth cannot continually grows as fast as debt, the one-to-one relation between the two will at some point be broken…….” [113].

But the most important implication is that a zero-growth or even a low growth economy appears incompatible with a fractional reserve banking system.

Credible alternatives to the fractional reserve system have been proposed by Fisher [96,114,115] and more recently by Robertson and Huber [91] and Kotlikoff [110]. A common aim is to separate the lending function of banks from the creation of money and to place the latter more fully under government control. Fisher proposed establishing a Currency Commission that would issue money and use it to purchase bank assets in sufficient quantities to allow their demand deposits to be fully backed
by the newly created currency. The money supply would remain unchanged during this process and the deposit banks would subsequently be required to maintain 100% reserves—thereby preventing them from increasing the money supply through issuing credit. Robertson and Huber [91] show how the same outcome could be achieved by declaring demand deposits as legal tender and removing current accounts from the banks’ balance sheets. The central banks would decide how much the money supply needs to be increased each year and would credit this to governments as public revenue. This could then be used to increase public expenditure, reduce taxes, lower government borrowing or fund a basic income scheme.

The claimed advantages of a 100% reserve scheme include simplification of the financial system, reduction of government debt, elimination of the need for deposit insurance, greater control over the money supply and a reduced risk of bank runs and boom-bust cycles [110]. It would also allow the profit from creating and spending new money (estimated to be around £50 billion/year in the UK [91]) to accrue to the public rather than the private sector and potentially restrain the “growth imperative” of the existing monetary system [83].

A related approach, termed “Neo-Chartalism” focuses upon the government’s role in supplying high-powered money to the economy (Box 2) [116-118]. Neo-Chartalists argue that modern “fiat currency” systems operate in a fundamentally different way from the commodity-backed systems they replace and that such currencies gain value by being required for the payment of taxes. A key claim is that the aggregate level of bank reserves is largely a consequence of government deficit spending, with central banks acting largely to modify the level of bank reserves on a short-term basis in order to meet interest rate targets [116]. As Tcherneva [119] notes: “…. this completely reverses conventional wisdom. Governments do not need the public’s money to spend: rather the public needs the government money to pay taxes. Government spending always creates new money, while taxation always destroys it”. Hence, governments do not need to raise tax revenue or borrow money in order to finance spending and a “balanced budget” is an inappropriate goal.

Lawn [118] shows how this perspective could be used to facilitate the transition to a steady state economy. He proposes raising the reserve requirement for banks, capping the amount of credit that they can create and using high-powered money to finance investment in public goods and critical infrastructure. Governments would not use taxation to raise revenue but instead to achieve environmental and distributional objectives and to reduce the inflationary effect of government spending.

In combination, the 100% reserve and Neo-Chartalist proposals provide a fertile source of ideas for a monetary system that is compatible with sustainability. The details of these proposals deserve more serious scrutiny and the extent to which they can break the “growth imperative’ identified by Binswanger remains unclear. But the issue is not so much the credibility of individual proposals, but the recognition that a transition to a zero-growth economy is unlikely to be achieved without fundamental reforms to the monetary system. To the extent that this issue is overlooked, a key part of the jigsaw is missing.
7. Conclusions

This paper has advanced five linked propositions regarding energy consumption, economic growth and environmental sustainability. While not new, these run counter to conventional wisdom and/or highlight blind spots within orthodox theory. Each raises numerous theoretical and empirical questions that deserve detailed and critical investigation. This will take time, but unfortunately that commodity is becoming increasingly scarce.

A sustainable economy needs to have much higher levels of energy and resource efficiency than exist today and policies to encourage this will have a crucial role to play. The objective should be to channel the benefits of improved efficiency into low carbon energy supply and improved quality of life rather than further growth and increased consumption. Options include green fiscal reform [120], progressive efficiency standards [121], caps on emissions and resource use, support for low carbon technologies, measures to encourage flexible working arrangements and reduced working hours, redistribution of income and so on. But such measures are likely to fail if the structural factors that make economies dependent upon continued economic growth remain unaddressed. In particular, the fractional reserve banking system needs to be replaced with one more consistent with the goal of sustainability. It is hoped that this paper will at least stimulate some thinking in that direction.

Acknowledgements

The early part of this paper is based upon a review of the evidence for rebound effects, conducted by the UK Energy Research Centre [5]. The financial support of the UK Research Councils is gratefully acknowledged. The author’s thinking on these topics has benefited from interactions with John Dimitropoulos, Harry Saunders, Blake Alcott, Paul Ekins and James Robertson and from the comments and criticisms of anonymous reviewers. The usual disclaimers apply.

References


© 2010 by the authors; licensee MDPI, Basel, Switzerland. This article is an Open Access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).