Reducing energy demand: A review of issues, challenges and approaches

Steve Sorrell *

Centre on Innovation and Energy Demand and Sussex Energy Group, Science Policy Research Unit, University of Sussex, Falmer, Brighton BN1 9QE, UK

A R T I C L E   I N F O

Article history:
Received 4 September 2014
Received in revised form
30 January 2015
Accepted 1 March 2015
Available online 21 March 2015

Keywords:
Energy efficiency
Energy demand
Barriers
Socio-technical transitions

A B S T R A C T

Most commentators expect improved energy efficiency and reduced energy demand to provide the dominant contribution to tackling global climate change. But at the global level, the correlation between increased wealth and increased energy consumption is very strong and the impact of policies to reduce energy demand is both limited and contested. Different academic disciplines approach energy demand reduction in different ways: emphasising some mechanisms and neglecting others, being more or less optimistic about the potential for reducing energy demand and providing insights that are more or less useful for policymakers. This article provides an overview of the main issues and challenges associated with energy demand reduction, summarises how this challenge is ‘framed’ by key academic disciplines, indicates how these can provide complementary insights for policymakers and argues that a ‘socio-technical’ perspective can provide a deeper understanding of the nature of this challenge and the processes through which it can be achieved. The article integrates ideas from the natural sciences, economics, psychology, innovation studies and sociology but does not give equal weight to each. It argues that reducing energy demand will prove more difficult than is commonly assumed and current approaches will be insufficient to deliver the transformation required.

© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Contents

1. Introduction .............................................................................................................. 74
2. System scale and energy demand ............................................................................ 75
   2.1. An understanding of energy demand must begin with the natural sciences .......... 75
3. Economic growth and energy demand .................................................................. 76
4. Energy efficiency and energy demand .................................................................. 77
5. Energy markets and energy demand .................................................................... 78
6. Energy policy and energy demand ....................................................................... 79
7. Sociotechnical systems and energy demand .......................................................... 80
8. Summary ............................................................................................................... 81
Acknowledgements .................................................................................................... 81
References .................................................................................................................. 81

1. Introduction

Improving energy efficiency and reducing energy demand are widely considered as the most promising, fastest, cheapest and safest means to mitigate climate change. Many opportunities appear to be cost-effective at current energy prices and can deliver additional benefits such as improved energy security, reduced fuel poverty and increased economic productivity. Reflecting this, the International Energy Agency (IEA) and other bodies are placing increasing priority on reducing energy demand, the European Commission has proposed long-term targets for energy demand reduction and countries throughout the world are introducing a range of policies to deliver those reductions.

But previous attempts at reducing energy demand have not always been successful. Frequently, the assumptions on which policy interventions are based do not adequately reflect either the challenges...
involved or the factors shaping individual and organisational decision-making. Moreover, the complexity of economic systems can lead to unintended and unanticipated consequences from those interventions that may undermine the original aims. Policies are usually informed to varying degrees by ideas from academic research, but different academic disciplines approach the challenge of reducing energy demand in different ways—emphasising some mechanisms and neglecting others, preferring some methodological approaches and sources of evidence over others, and providing competing recommendations. In turn, different disciplines are more or less optimistic about the potential for reducing energy demand and provide insights that are more or less useful for policymakers.

The aim of this article is to provide an overview of the issues associated with energy demand reduction, to summarise the ‘framings’ of this challenge by the academic communities that are most influential in this area and to argue that a ‘sociotechnical’ perspective can offer additional insights. The article integrates ideas from the natural sciences, economics, psychology, innovation studies and sociology but does not give equal weight to each—for example, it does not discuss the application of social practice theory to energy demand [1]. The aim is not to provide a comprehensive review of the literature, but does give equal weight to each—for example, it does not discuss the application of social practice theory to energy demand [1]. The aim is not to provide a comprehensive review of the literature, but does not discuss the application of social practice theory to energy demand [1].

2. System scale and energy demand

2.1. An understanding of energy demand must begin with the natural sciences

Energy is a mysterious property of objects and systems that can be neither created nor destroyed, but can be transferred from one system to another and converted from one form to another. Since not all forms of energy are equally useful, a more relevant quantity is exergy or the availability to perform physical work. Exergy is a measure of both the quantity and quality of energy and, unlike energy, can be destroyed during conversion processes (e.g. in the conversion of electricity to low temperature heat). Energy – or more precisely, exergy – is of unique importance in nature because nothing functions without it. Complex physical systems such as organisms, ecosystems and human societies exist far from thermodynamic equilibrium and can only be maintained in this state by a constant throughput of high quality energy from outside the system—with larger and more complex systems requiring larger energy flows.

Biologists and ecologists have identified remarkably consistent and apparently universal relationships between the physical scale of systems and the size of these of energy flows, based upon quarter power exponents [2,3]. For example, the metabolic rate of organisms scales with the three-quarter power of mass over 27 orders of magnitude, from the smallest microbes to the largest mammals [2,4]. There appear to be common principles underlying this universal relationship, linked to the evolutionary optimisation of the fractal-like branching networks that supply energy and materials to organic systems—such as the vascular system of plants and the circulatory system of mammals [4].

Since human societies rely upon analogous networks for distributing energy, water, food and other materials they may be subject to comparable constraints and hence exhibit comparable relationships between system scale and energy flows [5]. While drawing analogies between physical and human systems can be problematic, it is demonstrably the case that larger, wealthier, more populous and more complex societies require larger energy flows. Such societies evolved by accessing progressively larger energy flows and they cannot be sustained in the absence of those flows [6]. The massive increases in global population and wealth since the beginning of the 19th century have been associated with equally massive increases in energy consumption, derived largely from the ‘energy surplus’ provided by fossil fuels (i.e. the energy available from those fuels after subtracting the energy used to obtain them). Specifically, the sevenfold increase in global population since 1800 has been paralleled by a four-fold increase in per capita primary energy consumption (eight-fold in the industrialised world), leading to a 27-fold increase in global energy consumption [7].

The rate of growth of global primary energy consumption has been remarkably stable since 1850 (2.4%/year ± 0.08%) and shows no sign of slowing down [8]. Hence, if energy demand reduction is to be measured as a departure from this 150-year trend, there appears to be little sign of it yet at the global level. However, since primary energy consumption (E) has grown more slowly than gross domestic product (GDP) (Y), there has been a steady decline in global energy intensity (E/Y) and hence a steady increase in energy productivity (YE), with the precise trend depending upon how these variables are measured. Aggregate energy consumption is commonly expressed as the product of population (P), per capita wealth (YP) and energy intensity (E/Y), but many factors contribute to these aggregate ratios and more disaggregated breakdowns are required to understand their trends. The IEA [9] estimate that global energy intensity declined by 1.3%/year between 1990 and 2000, but this slowed to 0.4%/year after 2000 as a consequence of emerging economies (notably China) accounting for a larger proportion of global GDP. These economies are more energy intensive than the global average, but they are also reducing their energy intensity (and growing their economies) at a more rapid rate.

Regional trends in economic growth and energy consumption underpin this global picture and are consistent with it. For example, Brown et al. [10] examined 220 countries over 24 years and found that, on average, every 1% increase in per capita wealth was associated with a 0.76% increase in per capita energy consumption. As the authors observe, the closeness of this result to the three-quarter power relationship observed in natural systems may not be a coincidence. Similarly, Cserkelyei et al. [11] analysed 99 countries over the period 1971–2010 and found a comparable elasticity of approximately 0.7—implying that energy intensity is lower in richer countries and that, on average, a 1% increase in per capita income is associated with a 0.33% decrease in per capita energy intensity. This result was derived from repeated cross-sectional analyses of national data and indicates that the per capita energy use associated with any given level of per capita income has not changed for four decades. The long-term decline in regional and global energy intensity is therefore due to countries getting richer rather than from producing particular levels of wealth with less energy. This in turn suggests that the technological changes that have reduced energy intensity are strongly correlated with the technological changes that have increased per capita wealth.
The stability of the aggregate relationship between per capita wealth and per capita energy consumption obscures the significant and important variations from one country to another. Countries with similar per capita GDP can differ by an order of magnitude in their per capita energy consumption (and vice versa) as a consequence of differences in energy resources, energy prices, fuel mix, climate, population density, economic structure, regulatory standards and other variables. Notable outliers include energy-producing countries such as Saudi Arabia, Qatar, Australia and the US, where fossil fuels are subsidised or only lightly taxed. The US, for example, requires 50% more energy than the EU to produce a unit of GDP. Similarly, while per capita energy use increases with per capita income, the relative rates of growth of these two variables vary from one country to another and over time as a consequence of varying rates, types and directions of change in economic structure, fuel mix, technological efficiency and other variables.

It is critically important to understand the reasons for these spatial and temporal variations and to identify whether, how and to what extent patterns in one region and time period can be reproduced in different regions and time periods. Of particular interest are countries such as the UK and Denmark which have succeeded in growing their economies over periods of a decade or more with primary energy consumption declining in per capita and sometimes in absolute terms. But to date, these countries are too few in number and too small in size to have any noticeable influence on the global trends. Also, since this success has partly been achieved by scaling down domestic manufacturing, it may have been offset by increased energy use elsewhere. For example, while the UK’s territorial GHG emissions fell by 27% between 1990 and 2008, it’s “consumption-based” emissions increased by 20% as a consequence of imported consumer goods displacing (more energy efficient) domestically produced goods [12]. Such displacement is clearly not possible at the global level.

With populations growing and per capita energy consumption rising, global energy consumption continues its upward trend—growing by 2.3% in 2012–2013. Population growth has slowed in the industrialised world, but the increased wealth driving this ‘demographic transition’ is closely linked to higher per capita energy consumption. Analysis by DeLong et al. [13] suggests that national populations may stabilise when the rate of energy consumption reaches approximately 13 kW per capita which is at least five times the current global average. If correct, this suggests a huge potential for increased energy demand unless the dominant and long-established relationships between energy, wealth and population can be modified in a significant way. Hence, an understanding of these complex relationships must form the basis of any meaningful discussion of energy demand reduction.

3. Economic growth and energy demand

The previous section has shown that aggregate energy demand is closely correlated with the population and wealth of human societies, but there is considerable variation from one country to another. A key question is whether energy consumption increases because wealthier people demand more goods and services, or whether the availability of this energy drives the increase in wealth. In practice, both mechanisms coexist and are interdependent, with changes in energy availability, technology, income and population reinforcing each within positive feedback loops [14]. Establishing the relative importance of different variables is correspondingly difficult, with sophisticated econometric studies providing contradictory results [15,16]. The mainstream view is strongly influenced by orthodox economics which argues that the small share of energy in total input costs implies that increased energy consumption can only provide a small contribution to the growth in economic output. Robert Solow was awarded the Nobel Memorial Prize in Economics for developing a model of economic growth that ignored energy altogether [17] and it wasn’t until the 1970s that orthodox ‘production functions’ routinely included energy as one of their inputs. In contrast many economic historians [18], natural scientists [19] and ecological economists [20,21] consider energy to be critical for economic growth— with the massive energy surplus from fossil fuels liberating humanity from prior biophysical constraints:

... As long as supplies of both mechanical and heat energy were conditioned by the annual quantum of insolation and the efficiency of plant photosynthesis in capturing incoming solar radiation, it was idle to expect a radical improvement in material conditions of the bulk of mankind... [18]

...economic growth in the past has been driven primarily not by ‘technological progress’ in some general sense but specifically by the availability of cheaper energy...from coal, petroleum and gas... As Alvin Weinberg once said, energy is the ultimate resource. It is essential. It is needed for every economic sector and activity and there is no substitute... [22].

Ecological economists claim that the improvements in labour productivity over the last century have largely been achieved by providing workers with increasing quantities of energy, both directly and indirectly as embodied in capital equipment and technology [21,23]. There has also been a progressive shift towards higher quality energy carriers (e.g. more energy dense, more flexible, more transportable) which has allowed more economic output to be produced from each unit of energy input. Capital equipment embodies technological innovations, but since the primary use of that equipment is to transform energy flows into energy services, and since the equipment would be useless without those energy flows, technical change and increased energy use are closely linked. Ecological economists therefore view energy as contributing more to economic growth than is suggested by its small share of total input costs. More primary energy inputs leads to more useful energy and more final energy services, but this may also be achieved by shifting towards higher quality energy inputs and by improving thermodynamic conversion efficiency [24].

The feasibility of significant energy demand reduction may depend on which of these views is correct—with the orthodox perspective suggesting more scope for reducing energy demand than the ecological perspective. But the relationship between increased energy inputs, technical change and economic growth is complex and the empirical evidence is open to competing interpretations [15]. In developed economies, there is some evidence to suggest that the historical link between energy flows and economic output has weakened since the 1970s, perhaps as a consequence of the diffusion of information and communication technologies (ICT) [25]. Kander et al. [25] claim that ICT enables a less energy intensive form of economic development and that reductions in European energy intensity since 1970 owe more to the diffusion of ICT than to the shift towards a service-based economy. Similarly, Stern and Kander’s analysis of Swedish economic growth over the past century indicates that the contribution of energy to economic growth has declined over time [26]. These studies, together with the inter-country variation in energy intensity noted above, provide some support for the claim that a greater decoupling of national and global energy demand from GDP may be expected in the future. But on present evidence, expectations of rising incomes alongside falling energy consumption appear optimistic. Such an outcome would amount to an Environmental Kuznets Curve (EKC) for energy [27], but to date the evidence for such curbs is confined to a limited number of short-term and local environmental problems and appears largely absent for energy consumption [15].
4. Energy efficiency and energy demand

Despite the strength of these historical correlations, improved energy efficiency and reduced energy demand are widely expected to provide the dominant contribution to reduced carbon emissions in the short to medium term—and to do so at little or possibly negative cost. For example, the IEA’s ‘450 scenario’ has improved energy efficiency accounting for 71% of emission reductions (relative to the baseline scenario) in the period to 2020 and 48% in the period to 2035 [9]. But the link between improved energy efficiency and reduced energy demand (and hence emissions) is not straightforward: the first need not necessarily lead to the second, and both can be interpreted and measured in multiple ways.

Energy efficiency is the ratio of useful outputs to energy inputs for a specified system – such as a motor, a machine tool, an industrial process, a firm, a sector or an entire economy – while energy intensity is the inverse of this measure. In all cases, the measure of energy efficiency will depend upon how inputs and outputs are defined and measured. Depending upon the system, outputs may be measured in energy terms, such as heat content or physical work; physical terms, such as vehicle kilometres or tonnes of steel; or economic terms such as value-added or GDP [28]. Different measures may be more or less appropriate for different systems and purposes, and no measure will capture everything of value. For example, vehicle kilometres and passenger kilometres both measure the quantity of mobility, but the former does not capture load factors, the latter does not capture passenger comfort and neither are necessarily correlated with the frequency and ease of access to relevant destinations.

In many cases the most relevant output of a system is an energy service of some form, such as motive power, thermal comfort and accessibility. But energy services are difficult to measure, dependent upon social context and partly subjective, so a different definition, interpretation or understanding of the relevant energy service may lead to a different judgement on the energy efficiency of a particular system. For example, thermal comfort depends upon internal air temperature, but also upon radiant temperature, air velocity, humidity, activity levels, clothing, external temperature, behavioural norms and sociocultural conditions, so the thermodynamic efficiency of the boiler and the thermal performance of the building envelope are far from the only considerations.

The measurement of energy inputs also raises issues, especially when different energy carriers are combined [29]. The most common approach is to sum the thermal content of each energy carrier (in joules), but this amounts to summing apples and oranges. Energy carriers vary on multiple dimensions (e.g. volumetric energy density, gravimetric energy density, ease of storage, ease of transport, cleanliness) and they are only partially substitutable (e.g. try running a truck on battery-stored electricity). Higher quality energy carriers receive a higher price since they are more flexible, suitable for a wider range of end uses and produce more economic output per joule [29]. Price-based weighting schemes should therefore be (but rarely are) used to account for the different quality of energy carriers and when this is done aggregate measures of energy efficiency are found to be improving more slowly than is commonly supposed [29–31]. For example, Kaufmann [32] shows that much of the reduction in US energy intensity between 1950 and 1990 was linked to the shift towards higher quality and hence more productive energy inputs – such as from coal to oil – rather than from efficiency improvements per se.

Importantly, improvements in one measure of energy efficiency may not be reflected in improvements in a second measure, or in measures appropriate for a different spatial or temporal boundary. Indeed, it is entirely possible for an improvement in one measure to be associated with deterioration in another. For example, an electric heat pump is more energy efficient than a gas boiler when energy inputs are measured at the building level, but may be less energy-efficient when those inputs are measured at the source level (e.g. the fuel into the power station) or on a life cycle basis. Similarly, Kaufmann [33] shows how the energy savings in the US forest products industry over the period 1958–1984 were largely offset by the energy used to produce the relevant capital equipment.

In a similar manner, improvements in energy efficiency (however measured) may not always reduce energy demand and reductions in energy demand may result from something other than improved energy efficiency. To claim ‘energy-savings’ or ‘demand reduction’ it is necessary to specify the reference against which those savings are measured or estimated. That involves specifying the relevant spatial and temporal boundary and unit of measure, as well as invoking ceteris paribus assumptions. The reference may be historical energy consumption or a counterfactual scenario of what energy consumption ‘would have been’ in the absence of specified changes. But since data on energy consumption is not always available (or accurate), counterfactuals are unobservable and countervailing variables are difficult to control (for), the causal link between specific changes and the resulting ‘energy savings’ can be hard to establish. Most approaches rely upon decomposition or econometric analysis of secondary data at the aggregate level and the results are frequently lacking in resolution and sensitive to model specification. Experimental or quasi experimental studies can control for confounding variables at the micro level, but these are costly to conduct, comparatively rare and restricted in their spatial and temporal boundaries. As a result, the literature is replete with unreliable estimates of historical energy savings and questionable claims about future energy savings—both in relation to specific technologies and policies and in relation to the determinants of aggregate trends. California, for example, is hailed as an energy efficiency success story since per capita electricity consumption has remained fairly constant since the 1970s and is more than 40% below the US average. But a careful analysis of the contributory factors finds that California’s ambitious energy efficiency policies account for less than one third of this difference [34].

The link between improved energy efficiency and reduced energy demand is further complicated by the presence of multiple rebound effects. For example, since fuel-efficient cars make travel cheaper, consumers may choose to drive further and/or more often, thereby offsetting some of the energy savings achieved. Drivers may use the savings on fuel bills to buy other goods and services which necessarily require energy to provide—such as laptops made in China and shipped to the UK. Reductions in fuel demand will translate into lower fuel prices which in turn will encourage increased fuel consumption elsewhere. Similar mechanisms exist in industry, where cost-effective energy efficiency improvements allow firms to expand output, lower product prices and increase market demand which in turn stimulates economic growth and aggregate energy consumption. In some cases, energy efficient innovations may lead to new, unforeseen energy-using applications, products and industries [35,36]. For example, the Bessemer process greatly improved the energy efficiency of steel-making, but also produced cheaper and higher quality steel suitable for a wider range of uses, thereby increasing demand for both steel and coal. Rebound is therefore an emergent property of complex economic systems, with the multiple mechanisms and effects being difficult to isolate and measure, especially over the longer term. But a growing body of evidence suggests that these effects are larger than was previously thought and frequently offset or even eliminate the energy savings from improved energy efficiency [36–39].

From an engineering perspective, energy demand may be reduced by: improving the thermodynamic efficiency of energy conversion devices such as boilers and engines; preserving, heat, light, momentum or materials in passive systems, such as houses, cars and steel bars; or reducing demand for final energy services such as thermal comfort and mobility [40]. For example: gas use for home heating may be reduced by installing a more efficient boiler, insulating the walls or roof, or accepting lower internal temperatures; petroleum use for car travel...
may be reduced by improving the efficiency of the engine, reducing the size, weight, rolling resistance and/or air resistance of the vehicle; or simply driving less; and coal use for steel manufacture may be reduced by improving the efficiency of blast furnaces, increasing scrap recovery and product life, or designing buildings and products to use less steel.

Globally, Cullen et al. [41,42] estimate that that global average conversion losses could be reduced by a theoretical maximum of 89% and passive systems losses by a practical maximum of 73%, implying that current demand for energy services could be provided with much lower energy consumption. But this is a theoretical potential, so the technical and (especially) economic potential is likely to be much less. Also, the rate at which improvements in conversion efficiency or passive systems can be achieved is constrained by the rate of turnover of the relevant capital stock. For example, cars and white goods have an average lifetime of 10 years, while the lifetime of power stations, blast furnaces, ships and aircraft can easily exceed 30 years [43]. Premature replacement of existing equipment can accelerate the rate of efficiency improvement, but this involves trade-offs between the energy used in constructing new equipment and the energy used in operating the old equipment [44].

Further reductions in energy demand may be achieved by reducing demand for the relevant energy services (‘sufficiency’), but growing incomes create strong pressures in the opposite direction. This is particularly the case for countries at earlier stages of industrial development, but also applies more generally: for example, an analysis of lighting demand over three centuries and six continents finds no evidence of saturation even in the wealthiest countries [45]. Changes in demand for energy services can often occur fairly rapidly, but these too may be constrained by the lifetimes of relevant technologies and infrastructures [46]. For example, the physical characteristics and spatial location of houses, workplaces and other assets can lock-in heating, cooling and mobility needs for decades. More generally, voluntary actions to reduce any form of consumption face multiple obstacles within a growth-based economy [47].

In sum, it can be misleading to equate improved energy efficiency with reduced energy demand. The definition and measurement of these terms deserves more careful attention, as does the complex relationship between them. Similar comments apply to behavioural change, or ‘sufficiency’, since this can also have unintended consequences and necessarily involve swimming against a strong tide. A failure to acknowledge these complexities may partly account for the accumulation of estimates of ‘energy savings’ from specific interventions, while aggregate energy consumption continues to rise. This does not mean that energy demand cannot be reduced, but does imply that it may be more challenging than many analyses, policy documents and political statements suggest (e.g. [48]). Amongst other things, it points to the necessity of energy and carbon pricing as part of a portfolio of measures to reduce energy demand. This portfolio is the subject of the following three sections.

5. Energy markets and energy demand

Most modern societies are market economies of various forms, and since tradable energy commodities dominate total energy flows, energy demand is manifest as the aggregate purchases of those energy commodities in various markets. A distinction is commonly made between primary and final energy consumption, with conversion and distribution losses making the former larger than the latter. With energy suppliers facing rising marginal costs, energy consumers facing declining marginal utility, and government and regulatory bodies intervening in various ways, the demand for these energy commodities is in large part the emergent outcome of multiple economic decisions by multiple economic actors in multiple markets. Hence, energy demand is fundamentally an economic concept that can be investigated with the tools of economics. Energy demand responds to changes in energy prices, although to different degrees in different markets and over different periods of time. However, the simple textbook model of rational decision-making by well-informed actors in well-functioning markets provides a poor approximation to the markets for energy commodities and energy-using products, and an even poorer approximation to the ‘market’ for energy efficiency. This model therefore needs to be modified for more useful insights to be obtained.

First, it is commonplace to observe that people do not ‘demand’ energy for its own sake, but in combination with energy-using products that are used to provide final energy services—such as heating, lighting and motive power. Energy demand is therefore a derived demand, driven in part by the demand for energy services and influenced by the cost and efficiency of energy-using equipment as well as the price of energy. Since efficiency improvements reduce the effective price of energy services, they also increase demand for those services (rebounce). And since energy efficiency has improved continuously over the past 150 years, the demand for energy services has increased more rapidly than the demand for energy [49,50].

Second, while most debates about energy demand focus upon household consumption of electricity and fuels, this only accounts for around one quarter of total energy consumption. In practice, most of the energy ‘consumed’ by households is ‘embodied’ in non-energy goods and services, with the energy being consumed directly by upstream producers and distributors at different stages of the global supply chains. It is therefore essential that analysts and policymakers pay equal attention to energy demand in industry, commerce and freight transport.

Third, the idealised model of well-informed rational economic actors making ‘optimal’ decisions on the basis of known, invariant and consistent preferences is an acknowledged fiction with strictly limited usefulness. While economic incentives and cost-benefit trade-offs influence decisions on the purchase and use of energy and energy-using products, they explain only a portion of observed behaviour—even for large scale investment decisions by energy-intensive industrial firms [51]. Most decisions relevant to energy consumption are taken in the context of limited or asymmetric information about energy (service) costs by actors who are greatly constrained in their ability to respond to those costs even in circumstances where they are aware of and paying close attention to them—which is rarely the case. Behavioural economics provides a more robust understanding of economic decision-making and provide strong evidence for multiple, consistent and predictable departures from the orthodox model even in situations where the economic incentives are strong, actors are well-informed and conscious deliberation is involved [52,53]. Social psychology, in turn, shows how broader values, attitudes, beliefs and social norms profoundly influence different types of economic and energy-related decision [54]. With energy costs being small, largely invisible and poorly understood in most relevant situations, the more common situation is unreflective, habitual energy consumption in which energy costs are secondary to other factors such as convenience and symbolism and where energy-using behaviours exhibit considerable inertia. This leads authors such as Schwanen et al. [55] to emphasise the collective determinants of habitual behaviour rather than the deliberate decisions of utility-maximising individuals.

Fourth, the focus of the orthodox model upon market equilibria is of limited value in explaining market dynamics, including in particular the source, rate and direction of technical change and the parallel evolution of industrial structure, user preferences and market institutions. A key factor driving energy consumption is the speed of diffusion of energy using technologies, including both the replacement of existing stock (e.g. cars, fridges, motors) and the development and diffusion of new, energy-using capital goods and consumer products (e.g. mobile phones, laptops, space cooling.
equipment). The former changes the efficiency with which current energy services are provided—thereby stimulating increased demand for those services—while the latter may create entirely new markets, industries, consumer ‘needs’ and associated energy service demands. Diffusion is influenced by multiple economic, social and political factors and since energy efficiency represents a secondary and largely invisible attribute of most technologies, it plays only a limited role in the decisions of producers, consumers and other relevant actors.

Finally and most importantly, the focus on autonomous decision-making by individuals neglects how preferences, attitudes, expectations and behaviours are embedded in and shaped by broader physical and social systems that both enable and constrain individual choice. Consider, for example, the travel choices available to households living in low-density suburbs, with inadequate public transport and facing a 20-mile commute to work each day. As described in Section 7, energy services such as mobility are provided through large-scale, capital intensive and long-lived infrastructures that co-evolve with associated technologies, institutions, skills, knowledge and behaviours to create ‘sociotechnical systems’ [56]. While energy demand depends upon economic decisions informed by individual preferences, those preferences are to a large degree shaped, stabilised and maintained by the sociotechnical system themselves.

6. Energy policy and energy demand

Deliberate attempts by governments to reduce energy demand (‘framed as ‘energy conservation’) date back to the oil price rises of the 1970s and the subsequent concerns about energy security. Attention waned after the oil price falls of the 1980s when market liberalisation dominated the agenda, but since then demand reduction has gained prominence as a means to mitigate climate change—now framed as ‘promoting energy efficiency’. Policy in this area is growing in breadth and ambition and is having a demonstrable effect in many areas (e.g. white goods, buildings, vehicles, motors), but it remains primarily informed by orthodox economics. Other perspectives such as social psychology and innovation studies are less influential but may complement the dominant approach. This section summarises the key insights offered by these perspectives and how they may be used.

Despite the caveats presented in Section 5, it is clear that energy prices matter a great deal—especially over the longer term when organisations and households have time to adjust. A mountain of evidence demonstrates that higher energy prices lead to lower energy demand at all levels—as simple theory would predict. Trying to reduce energy demand without raising energy prices is akin to pushing a rock up a hill—it is hard work and at some point the rock will roll back down. The orthodox view is that prices should fully reflect costs, including the external costs of pollution—although many economists and most non-economists would consider the latter to be non-quantifiable. Nevertheless, there is a powerful rationale for raising the price of fossil-based energy through some form of carbon pricing. This should be as comprehensive as possible, increasing in stringency over time in line with long-term emission targets, and with the revenues being used to reduce labour and income taxes and/or support other policy objectives. But while carbon/energy pricing is necessary, it is far from sufficient and the achievable level of pricing is politically constrained [57]. Carbon pricing can have significant impacts on income distribution and economic competitiveness, and while revenue recycling can make such policies both distributionally progressive and economically beneficial at the aggregate level, some industrial sectors and social groups could still be hard-hit. Complementary policies are therefore required to reduce other obstacles to demand reduction, as well as to improve the political acceptability of carbon pricing itself—for example by exempting certain sectors or protecting low-income groups.

Even without carbon pricing, many studies suggest that there are numerous cost effective opportunities to improve energy efficiency in the majority of end-use sectors—as exemplified by the marginal abatement cost curves developed by McKinsey and others [58]. The dominant question then becomes ‘why are these opportunities not being taken up?’—with the dominant answer being ‘there are multiple, overlapping non-price barriers to improved energy efficiency’ [51]. For example, organisational staff may lack the time to identify and implement energy efficiency opportunities; they may employ simple payback calculations or rules of thumb rather than detailed investment appraisals; they may lack information on current energy consumption, energy efficiency opportunities and/or the energy costs of different investment options; they may face difficulties in raising capital, either internally or externally; they may be concerned about production interruptions, or the poor performance of energy-efficient technologies; they may lack the incentive to improve energy efficiency—for example, if energy use by their department is not directly metered or billed; and so on. These obstacles are typically much larger for households than for firms and commonly act in combination to block apparently cost effective investments. While many of these obstacles represent real costs that McKinsey and other analysts routinely overlook, there are often ways of reducing these costs through organisational initiatives or policy interventions—such as minimum efficiency standards, labelling schemes, subsidised energy audits, building regulations, reporting requirements, targeted information programmes and the like. The impact of these policies is hard to assess, but several studies indicate that they can be cost-effective [59,60]. These approaches should therefore play a central role in attempts at demand reduction.

Given the limitations of orthodox economics identified in the previous section, policy interventions informed solely by this framework are likely to be insufficient. Behavioural economics and social psychology provide deeper insights into the reasons why apparently cost-effective opportunities are neglected and can inform the design of more effective interventions that encourage energy and economically efficient choices. For example: people focus on losses more than gains, so energy efficiency information is best framed as preventing a loss rather than providing a gain; people tend to choose the default option, so steps should be taken to ensure that this is energy-efficient; people care about relative rather than absolute performance, so information on household energy consumption should be benchmarked against immediate neighbours; people are heavily influenced by who communicates information, so the benefits of energy efficiency should be communicated by trusted individuals; people are influenced more by salient than accurate information, so visual cues and vivid descriptions should be employed; and so on [61–66]. Governments are increasingly using these ideas to inform policy design, but they have yet to have much influence on energy efficiency policy. While we are unlikely to ‘nudge’ our way to a low energy future, such approaches have the potential to complement more traditional policies.

Both economic and social psychological approaches focus primarily upon encouraging the adoption of existing cost-effective technologies or the efficient use of those technologies. But much larger opportunities to reduce energy demand lie in encouraging the diffusion of novel, energy-efficient technologies and systems, such as heat pumps, solid wall insulation and hybrid vehicles. Technologies such as these typically face an uphill struggle when first introduced, since they perform relatively poorly compared to established technologies and are more expensive since they have yet to benefit from scale and learning economies. New technologies may also lack established supply chains and market structures, fit poorly within existing infrastructure, be unfamiliar to key intermediaries such as contractors and suppliers, require significant changes in user practices and face political
opposition and regulatory obstacles (see Section 7). While such problems are common to many new technologies, the ‘market pull’ for energy efficient technologies is weak unless they simultaneously offer other user benefits—not least because public good benefits such as climate protection and energy security are not captured by individual adopters [67]. There is therefore a case for targeted support for the development and diffusion of new, energy-efficient technologies by such means as R&D subsidies, demonstration programmes, technology incubators and accelerators, public procurement programmes, tax breaks, technology-forcing standards and so on [67]. These policies exist to varying degrees in different jurisdictions, but at present innovation support is strongly biased towards energy supply rather than energy demand [68].

An effective policy approach to reducing energy demand can therefore draw upon all these perspectives, combining revenue-neutral carbon pricing with policies to reduce non-price barriers to improved energy efficiency, behavioural interventions to encourage more energy efficient choices and support for energy-efficient technologies at each stage of the innovation chain. With careful design, these approaches can be mutually supportive: higher energy prices can induce low-energy innovation and bring forward the date at which new, energy-efficient technologies become competitive; behavioural interventions and removal of non-price barriers can encourage cost-effective investment, ease the economic impact of higher energy prices and reduce the political opposition to those prices; recycling of carbon tax revenues can provide funds for innovation support, adoption subsidies and protection of vulnerable sectors and groups; and so on [67]. These elements are visible in the policy mix of most OECD countries but typically lack coherence and adequate political support. As a result, their net impact has been rather limited to date and biased towards incremental rather than radical change. How these more radical changes can be achieved is the key question for energy efficiency policy in the next decade. Pushing harder with existing policy instruments will certainly be necessary [67], but it may also be helpful to frame the challenge in a different way; namely, as requiring more far-reaching changes in the sociotechnical systems that provide energy services. The next and final section summarises the additional insights that this perspective may provide.

7. Sociotechnical systems and energy demand

Economics and social psychology share a common focus on individual choice and provide policy prescriptions which seek to influence those choices. But a contrary viewpoint is that effective change requires a focus on the collective influences on individual choice and the systems that both enable and constrain those choices. The core insight here is that energy-related decisions are structured by the systems that provide energy services such as heating and personal mobility. These systems are termed ‘sociotechnical’ since they involve interlinked social and technical elements that co-evolve over many decades [69–72]. These sociotechnical systems may be only dimly understood by the actors involved in reproducing them, but they constrain the agency of those actors and strongly influence the type and quantity of energy required to provide particular energy services. Hence, it is important to understand how these systems function, how they can change and how these changes can be directed and accelerated by public policy.

A particularly important and illuminating example is our car-based system of personal transport. This is centred on an individual artefact – the automobile – but this artefact is linked to and dependent upon multiple social and technical elements at a variety of levels. These include, but are not confined to: the global car industry and its many associated supply chains; the car maintenance and distribution network; the global oil industry and the associated infrastructure of oil wells, pipelines, refineries, distribution networks and fuel stations; the road infrastructure and traffic system; the patterns of land use that have developed around that road infrastructure, including amenities and workplaces that are only accessible by car; the multiple institutions, regulations and policies associated with the production and use of cars; the engineering skills and knowledge built up over decades in a variety of domains; the political power of relevant interest groups; the daily travel routines, behaviour and expectations of millions of car owners; and the symbolism and cultural norms that have become associated with car-based mobility (‘car culture’) [72]. These different elements act together to shape the level and pattern of personal mobility and hence the energy use for that mobility. While reductions in energy demand can be achieved through incremental changes in individual system elements (e.g. by using lighter and more fuel efficient vehicles) larger reductions are likely to require more radical changes in the system itself, such as a shift away from internal combustion engines to battery-electric motors, or the large-scale substitution of personal car travel by mass transit systems [71].

Many authors have described how these systems evolve and become established, how they encourage incremental change along predictable trajectories and how their stability can obstruct more radical change [72–75]. Sociotechnical systems frequently rest upon core technologies such as the internal combustion engine whose early evolution involves considerable uncertainty. Historical experience suggests that (contrary to the predictions of orthodox economics) apparently inferior technologies can become dominant when they obtain an early advantage that allow them to benefit from various positive feedback mechanisms—such as scale economies that reduce costs, lower prices and increase demand; learning economies that improve performance, increase product attractiveness and further reduce costs; and network economies that enhance value through the development of complementary goods and services [72,76–78]. As the core technology diffuses, other mechanisms come in to play to reinforce its dominant position, such as: investments in supporting infrastructure (e.g. roads, pipelines, garages); increased knowledge and capabilities in relevant areas (e.g. motor engineering); the growing economic and political power of relevant groups (e.g. the car industry); the establishment of supportive legal, and institutional frameworks (e.g. professional institutions, labour unions); and the evolving habits, norms and aspirations of consumers [72]. These interdependent and co-evolving elements combine to form geographically extensive systems that becomes increasingly entrenched, or locked-in, making it difficult for technologies and behaviours that diverge in various ways from the dominant system (e.g. electric vehicles, light rail, car sharing) to become established [76,79].

Geels’ ‘Multi Level Perspective’ (MLP) provides one way of understanding the stability of these sociotechnical systems, as well as the processes through which they may be transformed (‘transitions’) [56]. Informed by a series of historical case studies; (e.g. [70,73,80,81]), Geels explains radical change as the result of interactions between three levels: the system itself which may be encountering internal difficulties (or more specifically the conventions, rules and norms that guide the actors in the system, termed the ‘regime’); the ‘niches’ in which radical innovations are being developed and protected in some way from the dominant system; and the exogenous socio-economic ‘landscape’ that is imposing pressures on the system. Relevant low-energy niches for personal mobility may include hybrid and battery electric vehicles, urban mass transit systems, communal and commercial car sharing and sustainable urban planning, while relevant landscape pressures may include rising oil prices, international agreements on climate change and the rapid development of ICT.

Based upon analyses of previous ‘socio-technical transitions’ (e.g. sailing ships to steam ships [70]; Dutch water supply [73]; US factory organisation [81]), Geels proposes a variety of pathways through which systems may be transformed [82]. A common theme is that niche technologies breakthrough when their internal momentum combines with growing tensions within the existing socio-technical
system and growing external pressures upon it. Together, these create ‘windows of opportunity’ for radical change. If successful, the breakthrough of new innovations trigger a series of inter-related technical, industrial, social and cultural changes that combine to create a new and different sociotechnical system based around a different set of core technologies.

Framing the challenge of energy demand reduction in this way encourages a move away from the individualist and incremental focus of current policy approaches and towards more overarching visions of long-term, systemic change. For example, visions of a future, low energy mobility system may build on a range of existing niche activities to include: land use planning that minimises travel needs; multimodal, mobility-on-demand urban transport systems; electric vehicles that are fully integrated into a smart grid and providing back-up power; prioritisation of cycling and walking in urban developments; and so on. Given the complexity of personal mobility, a future sustainable transport system seems more likely to encompass a range of (currently niche) innovations acting in synergy, rather than being centred upon a single dominant technology like the automobile. To deliver significant demand reductions, these need to reverse the trend of the last two centuries—specifically by enabling and encouraging people to travel less often over shorter distances at slower speeds in more efficient and fuller vehicles using lower carbon energy sources.

The MLP has been developed and applied as an analytical framework to explain historical transitions that took decades to unfold and were largely autonomous, unplanned and energy-using. But to be useful for energy policy, it needs to be remodelled into a prescriptive framework that is suitable for ‘steering’ future transitions that are directed, rapid and energy-saving. This represents a challenge and one that is not always consistent with the intentions and interests of the researchers involved, many of whom are primarily interested in historical explanations [83]. However, an increasing number of researchers have taken up this challenge and developed ideas on ‘strategic niche management’ [84], ‘transition management’ [85] and related approaches [86] and have begun to apply these to specific policy areas. Reflecting the complexity and unpredictability of the mechanisms involved, these approaches emphasise things like: the role of shared, long-term visions; the importance of coordinating policies both horizontally (e.g. between sectors) and vertically (e.g. between ministries and agencies); the need for continuous monitoring and adaptation; the need to keep options open and encourage parallel developments; the importance of integrating producers and end users in the innovation process; the need for a portfolio of instruments, including technology-specific policies; the complementary role of social and organisational innovations; and so on.

Despite this, the means by which large-scale transitions may be ‘steered’ still remains rather vaguely specified and the extent to which they can be steered remains contested [83,87]. Hence, while the sociotechnical perspective helps to focus attention on systemic determinants of energy demand and the need for radical system change, its practical implications for policy require much more development.

8. Summary

This article has reviewed the issues associated with energy demand reduction, the dominant framings of that challenge, the current policy approaches to meeting that challenge and the insights offered by a broader socio-technical perspective on that challenge. There are three broad conclusions.

First, reducing energy demand may prove more difficult than is commonly assumed. Complex economic systems necessarily require large energy flows, the correlation between energy consumption and economic growth is very strong and orthodox economics is likely to have underestimated the dependence of that growth upon increased energy consumption. The common expectation of energy efficiency improvements leading to proportional reductions in energy demand is misconceived—the linkages between the two are complex and rebound effects are frequently large. Much analysis and policy debate continues to oversimplify the issues involved.

Second, energy demand is an economic concept and that demand will not be reduced in the absence of rising energy prices and policies to reduce the economic barriers to improved energy efficiency. But although this framing is necessary and dominant, it has important limitations and the recommended policy approaches are insufficient. Behavioural interventions to encourage more energy-efficient choices and support for new, energy efficient technologies at each stage of the innovation chain are also required. These measures are underrepresented in the current policy mix which remains largely focused upon energy supply and incremental changes within existing systems.

Third, an effective response to climate change is likely to require larger and more rapid reductions in energy demand than have been achieved in the past. This in turn implies large-scale transformations in the ‘sociotechnical’ systems that provide energy services such as thermal comfort and personal mobility. Most historical transformations in systems such as these have not been deliberately directed by governments, and the nature, direction, speed and ultimate destination of these transformations is difficult to envisage. A key challenge for policy is not just to provide that direction, but to do so in a way that overcomes the inertia of sunk investment and delivers the speed of change required to avoid dangerous climate change. The growing literature on sociotechnical transitions could provide a source of ideas to assist in this task.

Acknowledgements

This research was funded by the Research Councils UK through their support for the Centre on Innovation and Energy Demand (Grant no. EP/K011790/1). A version of this paper was presented in Lancaster in July 2014 at a meeting of the six RCUK Centres on End Use Energy Demand (EUED). The author would like to thank Florian Kern, Ed Steinmüller and anonymous reviewers for their comments. The usual disclaimers apply.

References


