UKERC

UK ENERGY RESEARCH CENTRE

UKERC Review of Evidence for the Rebound Effect

Technical Report 5: Energy, productivity and economic growth studies

Working Paper

October 2007: REF UKERC/WP/TPA/2007/013

Steve Sorrell and John Dimitropoulos

Sussex Energy Group (SEG) - University of Sussex

Preface

This report has been produced by the UK Energy Research Centre's Technology and Policy Assessment (TPA) function.

The TPA was set up to address key controversies in the energy field through comprehensive assessments of the current state of knowledge. It aims to provide authoritative reports that set high standards for rigour and transparency, while explaining results in a way that is both accessible to non-technical readers and useful to policymakers.

This report forms part of the TPA's assessment of evidence for a **rebound effect** from improved energy efficiency. The subject of this assessment was chosen after extensive consultation with energy sector stakeholders and upon the recommendation of the TPA Advisory Group, which is comprised of independent experts from government, academia and the private sector. The assessment addresses the following question:

What is the evidence that improvements in energy efficiency will lead to economy-wide reductions in energy consumption?

The results of the project are summarised in a *Main Report*, supported by five in-depth *Technical Reports*, as follows:

1. Evidence from evaluation studies
2. Evidence from econometric studies
3. Evidence from elasticity of substitution studies
4. Evidence from CGE modeling studies
5. Evidence from energy, productivity and economic growth studies

A shorter *Supplementary Note* provides a graphical analysis of rebound effects. All these reports are available to download from the UKERC website at: www.ukerc.ac.uk/

The assessment was led by the Sussex Energy Group (SEG) at the University of Sussex, with contributions from the Surrey Energy Economics Centre (SEEC) at the University of Surrey, the Department of Economics at the University of Strathclyde and Imperial College. The assessment was overseen by a panel of experts and is extremely wide ranging, reviewing more than 500 studies and reports from around the world.

Each Technical Report examines a different type of evidence and assesses its relevance to the rebound effect. Each seeks in particular to clarify the conceptual issues underlying this debate and to make these issues as accessible as possible. *Technical Report 5* focuses upon the relationship between energy, productivity and economic growth and examines the claim that improved energy efficiency will increase economy-wide energy consumption - the so-called 'Khazzoom-Brookes postulate'.

THE UK ENERGY RESEARCH CENTRE

Operating at the cusp of research and policy-making, the UK Energy Research Centre's mission is to be the UK's pre-eminent centre of research, and source of authoritative information and leadership, on sustainable energy systems.

The Centre takes a whole systems approach to energy research, incorporating economics, engineering and the physical, environmental and social sciences while developing and maintaining the means to enable cohesive research in energy.

To achieve this we have developed the Energy Research Atlas, a comprehensive database of energy research, development and demonstration competences in the UK. We also act as the portal for the UK energy research community to and from both UK stakeholders and the international energy research community.

Acknowledgements

We have benefited enormously in the preparation of this report from our interactions with Harry Saunders (Decision Processes Inc). Also, Dennis Anderson (Imperial College) provided two extremely helpful pieces of analysis, elements of which have been incorporated into this report.

We are very grateful for the comments received from Len Brookes, Blake Alcott (University of Cambridge), John Feather, Manuel Frondel (ZEW), Karsten Neuhoff (University of Cambridge), Jake Chapman and Nick Eyre (Energy Savings Trust). Useful comments and advice have also been received from our Advisory Group, namely: Jim Skea (UKERC); Horace Herring (Open University); Hunter Danskin (DEFRA); Tina Dallman (DEFRA); Ken Double (Energy Saving Trust); Lester Hunt (Surrey Energy Economics Centre); and Paolo Agnolucci (Policy Studies Institute).

We are extremely grateful for the guidance, comments, patience and moral support offered by Jim Skea (UKERC Research Director), Rob Gross (Head of UKERC TPA function), Dennis Anderson (Senior Advisor, UKERC TPA) and Phil Heptonstall (UKERC TPA).

The above individuals represent a range of views about the size of the economy-wide rebound effect and none of them are responsible for the content of this report.

Executive Summary

Headline points

- One interpretation of the so-called 'Khazzoom-Brookes postulate' is that *all* cost-effective energy efficiency improvements will increase energy consumption above where it would be without those improvements. This is a counterintuitive claim for many people and requires strong supporting evidence if it is to gain widespread acceptance. The main conclusion from this review is that such evidence does not exist.
- The theoretical arguments for the K-B postulate rely upon a conceptual framework that is stylised and restrictive, while the empirical evidence cited in its favour is indirect and suggestive. A number of flaws have been found with both. Nevertheless, the arguments and evidence used to defend the K-B postulate deserve more serious attention than they have received to date.
- It is conventionally assumed that there is considerable scope for substituting capital and other inputs for energy consumption while maintaining the same level of economic output. It is also conventionally assumed that technical change has improved the energy efficiency of individual sectors and contributed to the decoupling of energy consumption from economic growth. However, the evidence reviewed in this report suggests that there is more limited scope for substituting other inputs for energy and that much technical change has acted to increase energy intensity. Also, once different fuels are weighted by their relative 'quality' or economic productivity, there is less evidence for decoupling. Overall, this evidence points to economy-wide rebound effects being relatively large and to energy playing a more important role in economic growth than is conventionally assumed.
- The possibility of large economy-wide rebound effects becomes more plausible if it is accepted that energy efficiency improvements are frequently associated with proportionately greater improvements in total factor productivity. If this is the case, then rebound effects need not necessarily be small just because the share of energy in total costs is small. But energy efficiency improvements may not necessarily be associated with such improvements. Instead, the link between the two seems more likely to be contingent upon particular technologies and circumstances.
- The debate over the K-B postulate would benefit from more careful distinctions between different types of energy efficiency improvement. For example, the K-B postulate seems more likely to hold for energy efficiency improvements associated with 'general-purpose technologies' (GPTs), particularly when these are used by producers and when the improvements occur at an early stage of development and diffusion. Steam engines provide a paradigmatic illustration of a GPT in the 19th-century, while electric motors provide a comparable illustration for the early 20th century. In contrast, the K-B postulate seems less likely to hold for dedicated energy efficiency technologies such as thermal insulation, particularly when these are used by consumers or when they play a subsidiary role in economic production.

Introduction

The economy-wide rebound effect from energy efficiency improvements may be expected to be larger than the direct rebound effect and could potentially be greater than unity: in other words, energy efficiency improvements may actually increase overall energy consumption ('backfire'). From a climate change perspective, the economy-wide effect is ultimately what matters. However, the mechanisms involved are complex, interdependent and difficult to conceptualise, and the magnitude of this effect is extremely difficult to estimate empirically.

The authors most closely associated with claims for backfire are William Stanley Jevons, Len Brookes and Harry Saunders. It was Saunders who introduced the term 'Khazzoom-Brookes (K-B) postulate', namely:

> 'with fixed real energy prices, energy efficiency gains will increase energy consumption above what it would be without these gains'. (Saunders, 1992)

These authors use a mix of theoretical arguments and 'suggestive' empirical evidence to support their case, both of which are explored in detail in this report. However, the report also investigates a number of other sources of evidence, drawn in particular from ecological economics that focus on the relationship between energy and economic growth. The dispute over the magnitude of the economy-wide rebound effect is argued to be closely related to this much broader question.

Despite the breadth of literature covered, very few of the studies reviewed in this report provide quantitative estimates for the size of the economy-wide rebound effect. Indeed, the great majority of the studies make no reference to the rebound effect at all and instead provide 'suggestive' evidence on issues such as the importance of energy in economic growth. Moreover, much of this evidence is at odds with the conventional assumptions held by policymakers and energy analysts alike. Hence, the aim of this report is to examine both the strengths and weaknesses of this literature and the degree to which it may be used, both individually and in combination, to support or contest the K-B postulate. The emphasis throughout is on clarifying the theoretical issues involved, as well as making the concepts accessible to a non-technical audience.

Measures of energy efficiency and productivity

Defining and measuring both the independent variable for the rebound effect (an improvement in 'energy efficiency') and the dependent variable (a change in energy consumption) is far from straightforward. For the independent variable, attention must be paid (amongst other things) to: the *definition* of energy efficiency (e.g. first law thermodynamic, second law thermodynamic, physical or economic measures); the *system boundaries* to which it applies (e.g. individual device, process, firm, sector, national economy, regional economy, global economy); and the appropriate methods for *aggregating* different energy types (i.e. whether and how differences in energy quality are accounted for). For example, the standard practice of aggregating different energy types according to their thermal content neglects the 'quality' of different energy types, such as their ability to perform useful work. When the changing quality of energy inputs are accounted for, aggregate measures of energy efficiency are found to be improving much more slowly than is commonly supposed.

Many commentators assume that the relevant independent variable for the rebound effect is improvements in the *thermodynamic* efficiency of individual conversion devices or industrial processes. But such improvements will only translate into comparable improvements in *different* measures of energy efficiency, or measures of energy efficiency applicable to *wider* system boundaries, if several of the mechanisms responsible for the rebound effect fail to come into play. For example, improvements in the number of litres used per vehicle kilometre will only translate into improvements in the number of litres used per passenger kilometre if there are no associated changes in average vehicle load factors.

Rebound effects may be expected to increase over time and with the widening of the system boundary for the dependent variable (energy consumption). For the K-B postulate the relevant system boundary is normally taken as the national economy. But energy efficiency improvements may also affect trade patterns and international energy prices, thereby changing energy consumption in other countries. For the purpose of assessing the contribution of energy efficiency to reducing carbon emissions, the relevant system boundary is the whole world.

To capture the full range of rebound effects, the system boundary for the independent variable (energy efficiency) should be relatively narrow, while the system boundary for the dependent variable (energy consumption) should be as wide as possible. However, measuring or estimating the economy-wide effects of micro-level changes in energy efficiency is, at best, challenging. For this reason, the independent variable for many theoretical and empirical studies of rebound effects is a physical or economic measure of energy efficiency that is applicable to relatively wide system boundaries – such as the energy efficiency of an industrial sector. But such studies may overlook the 'lower-level' rebound effects resulting from improvements in physical or thermodynamic measures of energy efficiency appropriate to narrower system boundaries. Also, improvements in more aggregate measures of energy efficiency are unlikely to be caused solely (or even mainly) by the diffusion of more thermodynamically efficient conversion devices.

Many commentators also assume that a change in energy consumption following an energy efficiency improvement can be solely attributed to that improvement. But improvements in energy productivity may often be associated with broader improvements in the productivity of other inputs, since new technologies frequently provide both. If the *full* impact on energy consumption of these new energy-efficient technologies is taken as the appropriate dependent variable, then backfire becomes more likely.

Energy productivity and economic growth – the contribution of Len Brookes

Len Brookes deserves credit for developing coherent arguments in favour of the K-B postulate and for defending these against a range of criticism. The three most important arguments may be characterised as follows:

- *The productivity argument*: The increased use of higher quality forms of energy (especially electricity) has encouraged technical change, substantially improved total factor productivity and driven economic growth. Despite the substitution of energy for other inputs, this technical change has stimulated a sufficiently rapid growth in economic output that aggregate energy efficiency has <u>improved</u> at the same time as aggregate energy consumption has <u>increased</u>. This pattern may be expected to continue in the future. The productivity argument rests upon two separate, but related sources of empirical evidence. First, the work of Sam Schurr and colleagues on the historical importance of changes in energy quality (notably electrification) in driving productivity growth; and second, the work of Jorgenson and others on 'energy-using' technical change.

- *The endogeneity argument*: A common approach to quantifying the 'energy savings' from energy efficiency improvements is to hold energy intensity fixed at some historic value and estimate what consumption 'would have been' in the absence of those improvements. The energy savings from energy efficiency improvements are then taken to be the difference between the actual demand and the counterfactual scenario. But if the energy efficiency improvements are a <u>necessary condition</u> for the growth in economic output, the construction of a counterfactual in this way is misconceived.
- *The accommodation argument*: Energy efficiency improvements are claimed to 'accommodate' an energy price shock so that the energy supply/demand balance is struck at a higher level than if energy efficiency had remained unchanged (Brookes, 1984). While not immediately obvious, this argument rests on the assumption that the income elasticity of 'useful' energy demand falls steadily as an economy develops, but is always greater than unity.

The dictionary definition of 'postulate' is a starting assumption from which other statements are logically derived and which does not have to be self-evident or supported by empirical evidence. This appears to be Brookes' perspective and he acknowledges that the available evidence provides only 'suggestive' support for the K-B postulate. The main conclusion from this review is that these arguments and evidence are insufficiently robust to support his case. Flaws can be found both in the evidence itself in the manner in which Brookes uses this evidence. Specific criticisms include the following:

- Brookes cites Schurr's work in support of the postulate, but this applies primarily to the causal effect of shifts to higher quality fuels, rather than improvements in thermodynamic efficiency. Also, the patterns Schurr uncovered may not be as 'normal' as Brookes suggests.
- Brookes also cites econometric evidence on 'energy-using technical change' in support of the postulate. But these empirical results vary widely between different sectors, countries and time periods and are sensitive to minor changes in econometric specification. Moreover, the assumption of a fixed bias in technical change is flawed and the failure to check for the presence of cointegration or to account for changes in energy quality means that the estimates could be either biased or spurious. Moreover, even if energy-using technical change were to be consistently found, the relationship between this finding and the K-B postulate remains unclear.
- Brookes' 'accommodation' argument is based upon a theoretical model that is unconventional in approach and difficult to interpret and calibrate. The model rests on the assumption that the income elasticity of 'useful' energy demand is always greater than unity, but the study on which this claim is based has not been updated. Contemporary research on Environmental Kuznets Curves has not tested this hypothesis, since useful energy consumption is not employed as the independent variable.

A key theme in Brookes' work is that improvements in energy productivity are generally associated with proportionally greater improvements in total factor productivity. While Schurr's work provides evidence for this at the level of the national economy, numerous examples from the energy efficiency literature provide evidence for this at the level of individual sectors or technologies. If energy efficient technologies boost total factor productivity and thereby save more than energy costs alone, the argument that rebound effects must be small because the share of energy in total costs is small is undermined. Much the same applies to the contribution of energy efficiency improvements to economic growth. But energy efficiency improvements may not necessarily be associated with such improvements. Instead, the link between the two seems more likely to be contingent upon particular technologies and circumstances.

Energy productivity and neoclassical growth theory – the contribution of Harry Saunders

Saunders has provided a significant theoretical contribution to the rebound debate and has raised important questions regarding the behaviour of commonly used neoclassical production functions. While this work relies in particular upon neo-classical growth theory and production theory, it also has important implications for CGE modelling and for the econometric investigation of individual sectors. These implications do not appear to be fully appreciated in the wider energy economics community.

Saunders has shown how backfire is the predicted outcome of neoclassical production functions that are used widely in theoretical and empirical research. If such functions are considered to provide a reasonable representation of real-world behaviour, Saunders' work suggests that 'pure' energy-efficiency improvements are likely to lead to backfire. Alternatively, if rebound effects vary widely in magnitude between different sectors, such functions cannot be used to represent them. In either case, the implications are far-reaching.

The above conclusions apply to 'pure' energy efficiency improvements. But Saunders also uses numerical simulations to demonstrate the potential for much larger rebound effects when improvements in energy efficiency are combined with improvements in the productivity of other inputs. Again, if the validity of the theoretical assumptions is accepted, these results suggest that backfire may be a more common outcome than is conventionally assumed.

Saunders work suggests that rebound effects are highly sensitive to the value of the elasticity of substitution between energy (or energy services) and non-energy inputs. While there is a large empirical literature measuring elasticities of substitution between different inputs, the results are confusing and contradictory and may in practice provide little assistance in determining the likely magnitude of rebound effects. These issues are explored in detail in *Technical Report 3*.

Saunders approach is entirely theoretical and therefore severely limited by the assumptions implicit in the relevant models. For instance, technology always comes free, there are only constant returns to scale in production, markets are fully competititve, there is always full employment, qualitative differences in capital and energy are ignored and so on. A key weakness is the limited capability to capture the complexities of technical progress, which is assumed to occur autonomously without explicit representations of the processes which influence its rate and direction. These are captured within contemporary models of economic growth, but unfortunately such models have yet to be used to explore the rebound effect. Also, since what is at issue is the consequences of energy efficiency improvements, the source of those improvements may be a secondary concern.

Overall, Saunders work suggests that significant rebound effects can exist in theory, backfire is quite likely and this result is robust to different model assumptions. Since these results are rooted in a contested theoretical framework, they are suggestive rather than definitive. But they deserve to be taken seriously.

Energy productivity and ecological economics

Brookes (1984) quotes Sam Schurr's observation that: "….it is energy that drives modern economic systems rather than such systems creating a demand for energy." This highlights an important theme in Brookes' work: namely that energy plays a more important role in economic growth than is conventionally assumed (or more specifically, a more important role than is suggested by the small share of energy in total costs). But precisely the same claim is made by ecological economists, who attribute a large component of the increased productivity over the past century to the increasing availability of high-quality energy sources. This contrasts with conventional economists, who focus instead upon increases in capital and labour inputs and 'technical change'.

It is possible that this broader claim is valid, even if the K-B postulate does not always hold. A number of sources of evidence may be used in support of this claim, and each of these run contrary to conventional wisdom. For example:

- Reductions in aggregate energy/GDP ratios appear to be largely explained by structural change, changes in relative prices and improvements in the quality of energy inputs. Energy saving technical change at the micro-level may have contributed much less to such reductions than is commonly assumed.
- Historically, much technical change appears to have been energy-using, in that it has acted to increase aggregate measures of energy intensity. The conventional assumptions regarding the sign and magnitude of the *AEEI* parameter in energy-economic modelling could therefore be misleading.
- There may be much less scope for substituting other inputs for energy at the level of the economy as a whole than at the level of individual sectors. This is partly due to the energy 'embodied' in capital and labour inputs.
- Once different fuels are weighted by their relative 'quality' or economic productivity, there is less evidence for improvements in aggregate measures of energy efficiency. The observed reduction in conventional energy/GDP ratios may therefore overstate the extent to which energy consumption has been decoupled from economic growth.
- When changes in energy quality are taken into account, there appears to be little evidence for a turning point in the relationship between GDP and energy consumption. Hence, historical experience provides no support for the claim that economic growth can be maintained alongside absolute reductions in energy consumption.

Strong links can be found between the arguments and evidence used to support the K-B postulate and those used to support the claim that energy is a primary driver of economic growth. The implication of both is that attempts to decouple increased energy consumption from economic growth will be more expensive than is commonly assumed.

But while the 'ecological' perspective is well articulated and persuasive, the empirical evidence remains patchy and in some cases flawed. For example

- Estimates of the indirect energy consumption associated with particular goods and services exhibit considerable diversity. Such studies are rarely detailed enough to allow the indirect energy consumption associated with energy efficiency improvements to be estimated. As a result, they provide few empirical estimates of the magnitude of the indirect rebound effect and provide an insufficient basis on which to draw any general conclusions.
- The results of econometric investigations of causality relationships between energy and GDP remain ambiguous and the policy implications that are drawn are oversimplified. Also, the relationship being investigated here ('Granger causality') is not the same as causality as conventionally understood.
- The alternatives to conventional models of economic growth that have been developed by ecological economists appear to suffer from a number of statistical problems. As a result, claims that the marginal productivity of energy is in order of magnitude larger than its cost share, or that improvements in energy conversion efficiency can act as a suitable proxy for all forms of technical change, must be treated with considerable caution.

Unfortunately, the different assumptions of neoclassical and ecological economists seem to have prevented an objective comparison of their methods and conclusions. A bridge between the two could potentially be provided by recognising that increased inputs of energy services may frequently *enhance* the productivity of capital and labour:

> "……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." (Toman and Jemelkova, 2003)

It is an empirical question as to whether such benefits apply in practice and to what extent. Ecological economists appear to claim that such a situation is the norm, with the result that the increased availability of high-quality energy has been a primary driver of economic activity. But if the increased availability of energy inputs has a disproportionate impact on productivity and economic growth, then improvements in energy efficiency may do the same, because the effect of both is to increase the output of energy conversion devices – or 'useful work'.

If it is useful work rather than raw energy (or exergy) inputs that drives economic activity, then improvements in thermodynamic conversion efficiency could mitigate the economic impact of increasing shortages of high-quality forms of energy. However, improvements in conversion efficiencies are necessarily associated with embodied energy and are ultimately constrained by thermodynamic limitations. Also, if these improvements have a disproportionate effect on economic output, they may also be associated with large rebound effects.

Summary and conclusions

The K-B postulate' states that cost-effective energy efficiency improvements will increase energy consumption above where it would be without those improvements. This is a counterintuitive claim for many people and requires strong supporting evidence if it is to gain widespread acceptance. The main conclusion from this review is that such evidence does not exist. The theoretical and empirical evidence cited in favour of the postulate is suggestive rather than definitive, only indirectly relevant to the rebound effect and flawed in a number of respects. Nevertheless, the arguments and evidence deserve more serious attention than they have received to date. Much of the evidence points to economy-wide rebound effects being significantly larger than is conventionally assumed and to energy playing a more important role in economic growth than is conventionally assumed.

The possibility of large economy-wide rebound effects has been dismissed by a number of leading energy analysts. But it becomes more plausible if it is accepted that energy efficiency improvements are frequently associated with proportionately greater improvements in total factor productivity. If this is the case, then rebound effects need not necessarily be small just because the share of energy in total costs is small. But energy efficiency improvements may not necessarily be associated with such improvements. Instead, the link between the two seems more likely to be contingent upon particular technologies and circumstances.

Future research should therefore investigate whether, how, to what extent and why improvements in different measures of energy efficiency are associated with broader improvements in economic productivity, and the circumstances under which economy-wide rebound effects are more or less likely to be large. For example, on the basis of this review we may speculate that rebound effects should be larger for energy efficiency improvements associated with:

- energy intensive production sectors compared to non-energy intensive sectors;
- energy supply industries compared to energy users;
- core process technologies compared to non-core technologies;
- technologies in the early stages of diffusion compared to those in the later stages; and
- technologies that improve capital and labour productivity, compared to those that do not.

Rebound effects may be particularly large for the energy efficiency improvements associated with 'general-purpose technologies', such as steam engines, railroads, automobiles and computers. General purpose technologies (GPTs) are those that have a wide scope for improvement and elaboration, are applicable across a broad range of uses, have potential for use in a wide variety of products and processes and have strong complementarities with existing or potential new technologies. Steam engines provide a paradigmatic illustration of a GPT in the 19th-century, while the introduction of electric motors into manufacturing provides a comparable illustration for the early 20th century. The former was used by Jevons to support the case for backfire, while the latter was to use by Brookes.

The key to unpacking the K-B postulate may therefore be to distinguish the energy efficiency improvements associated with GPTs from other forms of energy efficiency improvement. The K-B postulate seems more likely to hold for the former, particularly when these are used by producers and when the energy efficiency improvements occur at an early stage of development and diffusion of the technology. The opportunities offered by these technologies have such long term and significant effects on innovation, productivity and growth that overall energy consumption is increased, rather than reduced. In contrast, the K-B postulate seems less likely to hold for dedicated energy efficiency technologies such as improved thermal insulation, particularly when these are used by consumers or play a subsidiary role in economic production. These technologies have much smaller effects on productivity and economic growth, with the result that overall energy consumption is reduced.

The implication is that climate policy should focus on encouraging dedicated energy efficient technologies, rather than improving the energy efficiency of GPTs. However, these categories are poorly defined and the boundaries between them are blurred. Moreover, even if GPTs can meaningfully be distinguished from other forms of technology, continued economic growth is likely to depend upon the diffusion and improvement of new types of GPT that may necessarily increase aggregate energy consumption.

In conclusion, while it is unlikely that all energy efficiency improvements lead to backfire, we still have much to learn about the factors that make backfire more or less likely.

Contents

1 Introduction

The economy-wide rebound effect from energy efficiency improvements may be expected to be larger than the direct rebound effect and could potentially be greater than unity: in other words, energy efficiency improvements may actually increase overall energy consumption ('backfire'). From a climate change perspective, the economy-wide effect is ultimately what matters. However, the mechanisms involved are complex, interdependent and difficult to conceptualise, and deriving empirical estimates of the magnitude of this effect is challenging. The literature on this subject also demonstrates some ambiguity over key theoretical issues, such as the relationship between thermodynamic measures of energy efficiency and economic measures energy productivity as well as the appropriate boundaries of the effect in time and space. With the notable exception of CGE modeling, direct estimation of the size of the economy-wide rebound effect is rare. Instead, the literature comprises an eclectic mix of theoretical argument, mathematical modelling, anecdotal examples and 'suggestive' evidence from econometric analysis and economic history.

CGE modelling is covered in *Technical Report 4* of this study and provides suggestive evidence that economy wide rebound effect could be large - although not necessarily greater than unity. The present report reviews a much wider range of literature with the aim of identifying any lessons that may be learnt regarding the magnitude of the economy-wide rebound effect. The focus throughout is whether or not 'backfire' is a real possibility - from any or all types of energy efficiency improvement. A conclusion that it is would fly in the face of conventional assumptions, as well as having profound implications for climate policy.

The three authors most closely associated with the economy-wide rebound effect are William Stanley Jevons, Len Brookes and Harry Saunders. W.S. Jevons argued as far back as 1865 that improved energy efficiency may actually increase overall energy demand and the possibility of 'backfire' from energy efficiency improvements has subsequently been labelled 'Jevon's paradox' (Alcott, 2005). Brookes forceful arguments in favour of backfire were published in a series of papers from 1978 onwards (Brookes, 1978; 1984; 1990b; 2000), although these were informed by empirical work conducted in the early 1970's (Brookes, 1972). This work prompted a fierce response from critics (Grubb, 1990; Herring and Elliot, 1990; Toke, 1990; Grubb, 1992), to which Brookes provided a number of robust responses (Brookes, 1992; 1993). Brookes' arguments were placed on a more formal footing by Saunders (1992) who used neoclassical growth theory to suggest that backfire is a very likely outcome from improved energy efficiency. Saunders' results depend upon the assumed functional form and parameters of an economy-wide production function and he was careful to state that his results do not *prove* backfire, but merely provide suggestive evidence in its favour, given certain assumptions how the economy operates.

It was Saunders who introduced the term 'Khazzoom-Brookes postulate', namely: 'with fixed real energy prices, energy efficiency gains will increase energy consumption above what it would be without these gains' (Saunders, 1992). The term postulate indicates a starting assumption from which other statements are logically derived. It does not have to be self-evident or supported by empirical evidence. But since most commentators do not accept the postulate, this assessment treats it as a hypothesis and seeks out testable implications. In addition, the 'Jevons-Brookes postulate' would be a more accurate term, since Khazzoom's work (discussed in *Technical Report 2*) focuses entirely on the direct rebound effect.

The work of Jevons, Brookes and Saunders introduces range of themes, theoretical issues and sources of evidence that are in some way relevant to the economy-wide rebound effect. These include: the relationship between macro-level energy productivity and micro-level thermodynamic efficiency (Berndt, 1978; 1990; Ang, 2006); the relationship between energy productivity and total factor productivity (Schurr, 1983; Schurr, 1985; Jorgenson, 1996); the contribution of energy to economic growth (Toman and Jemelkova, 2003; Stern and Cleveland, 2004); and the insights offered by various approaches to growth theory (Saunders, 1992; 2006). Taken together, these different strands by no means 'prove' the K-B postulate, but do permit a greater understanding of the issues involved. Each will be explored in detail in the current report, together with a number of other issues that appear relevant but were not cited by Brookes or Saunders. These include: the use of decomposition analysis to investigate the rebound effect (Schipper and Grubb, 2000); econometric studies of the 'causal' relationship between energy consumption and economic growth (Kraft and Kraft, 1978; Stern, 1993); empirically validated alternatives to the neoclassical growth model that include energy (or some related measure such as exergy or useful work) as a factor of production (Kummel, *et al*., 2002; Ayres and Warr, 2005).

Despite the breadth of this literature, almost none of the studies discussed in this report provide quantitative estimates for the size of the economy-wide rebound effect. Indeed, the great majority of the studies make no reference to the rebound effect at all. Instead, they provide largely 'suggestive' evidence on issues such as the importance of energy consumption in driving economic growth. Some of this evidence has previously been used by authors such as Brookes in developing the case for the K-B postulate, while other more recent studies could potentially be used in the same way. Moreover, much of this evidence is at odds with the conventional assumptions held by policymakers and energy analysts alike. Hence, the aim of this report is to examine both the strengths and weaknesses of this literature and the degree to which it may be used, both individually and in combination, to support or contest the case for the K-B postulate. The emphasis throughout is on clarifying the theoretical issues involved, as well as making the concepts accessible to a non-technical audience.

The report is structured as follows. Section 2 describes some of the mechanisms responsible for the economy-wide rebound effect and provides a historical perspective on the debate, including the contribution of W.S. Jevons. Section 3 provides an overview of the theoretical and methodological issues involved in measuring energy efficiency and in aggregating different types of fuel into a single measure of energy inputs. Much of this chapter may be familiar to energy analysts, but it provides an essential basis for the remainder of the report. It also highlights an important ambiguity regarding the appropriate independent variable for the rebound effect and the extent to which the estimated magnitude of the effect may depend upon the choice that is made.

Section 4 provides an extended discussion of the contribution of Len Brookes to the rebound debate and uses this to introduce and evaluate a wide range of relevant evidence. Brookes' is argued to have contributed three main arguments in favour of the K-B postulate, namely: the 'productivity argument' (divided here into the 'energy quality' and 'biased technical change' arguments); the 'endogneity argument'; and the 'accommodation argument'. In each case, the report describes the historical research that forms the basis for the argument, summarises how Brookes uses this evidence to support his claims for backfire, identifies potential empirical and/or theoretical weaknesses and examines whether more recent research confirms or contradicts Brookes' claims.

Section 5 evaluates the contribution of Harry Saunders to the rebound debate and the insights that may be obtained from neoclassical production theory and neoclassical growth theory. Since this work is fairly technical, the discussion is preceded by a description of the Solow-Swan growth model and supported by a number of appendices. This Section describes and explains the results from growth models and related research, identifies several limitations that may reduce the degree of confidence in these results and highlights their sensitivity to the choice of particular parameters. It also develops an extension of the growth model that shows how declining energy productivity (relative to other inputs) could in theory reduce overall energy intensity - the 'opposite' of backfire.

Section 6 introduces broader evidence on the relationship between energy consumption and economic growth and highlights the potential relevance of this to the economy-wide rebound effect. This work is associated in particular with ecological economics and provides a coherent alternative to the neoclassical mainstream. The review encompasses: econometric studies of the 'causal' relationship between energy consumption and GDP; discussion of the potential for substitution between energy and other factors of production; empirical estimates of the 'indirect' energy consumption associated with energy efficiency improvements; econometric analysis of the determinants of energy/GDP ratios; and alternative models of economic growth that include 'useful work' as one of the factors of production. What links this research to the rebound effect is the argument that energy plays a more important role in economic growth than is commonly assumed - or more specifically, that would be suggested by its relatively small share of total costs. In particular, the contribution of energy efficiency to reduced energy consumption may have been overestimated because changes in energy quality have not been taken into account.

Section 7 provides a brief summary of the overall conclusions and implications. It is concluded that the available evidence is insufficiently robust to verify the K-B 'hypothesis'. Contrary to the claims of Brookes and others, the extent of rebound appears to vary widely between different technologies, sectors and time periods and is not necessarily greater than unity. Nevertheless, the evidence suggests that economy-wide rebound effects may be significantly larger than is conventionally assumed and that energy efficiency improvements may sometimes lead to backfire. This is particularly because improvements in energy efficiency are often associated with broader improvements in total factor productivity. The evidence also raises concerns about the potential for decoupling increased energy consumption from continued economic growth.

2 Causal Mechanisms and Historical Perspectives

2.1 Mechanisms

It is helpful to begin with a brief account of the relevant causal mechanisms for the economy-wide rebound effect.

The *rebound effect* is an umbrella term for a number of mechanisms which reduce the size of the 'energy savings' achieved from improvements in energy efficiency. *Direct rebound effects* relate to individual energy services, such as heating and lighting, and are confined to the energy required to provide that service. *Indirect rebound effects* relate to the energy required to provide other goods and services, the consumption of which is affected by the energy efficiency improvement. The *economy-wide rebound effect* represents the sum of direct and indirect rebound effects and is normally expressed as a percentage of the *expected* energy savings from an energy efficiency improvement. Hence, a rebound effect of 100% means that the expected energy savings are entirely offset, leading to zero net savings. Box 2.1 proposes a classification scheme for the economy-wide rebound effect, while Box 2.2 summarise some of the mechanisms responsible for the indirect rebound effect.

Box 2.1 Classifying rebound effects

The economy-wide rebound effect represents the sum of the direct and indirect effects. For energy efficiency improvements by consumers, it is helpful to decompose the direct rebound effect into:

a) a *substitution effect*, whereby consumption of the (cheaper) energy service substitutes for the consumption of other goods and services while maintaining a constant level of 'utility', or consumer satisfaction; and

b) an *income effect*, whereby the increase in real income achieved by the energy efficiency improvement allows a higher level of utility to be achieved by increasing consumption of all goods and services, including the energy service.

Similarly, the direct rebound effect for producers may be decomposed into:

a) a *substitution effect*, whereby the cheaper energy service substitutes for the use of capital, labour and materials in producing a constant level of output; and

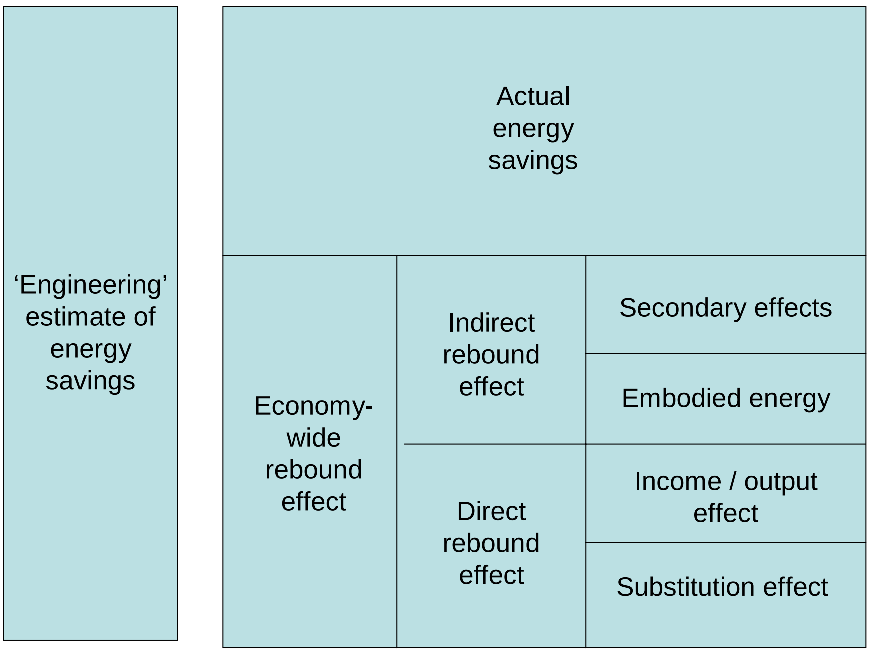
b) an *output effect*, whereby the cost savings from the energy efficiency improvement allows a higher level of output to be produced - thereby increasing consumption of all inputs, including the energy service.

It is also helpful to decompose the indirect rebound effect into:

a) the *embodied energy*, or indirect energy consumption required to <u>achieve</u> the energy efficiency improvement, such as the energy required to produce and install thermal insulation; and

b) the *secondary effects* that result as a <u>consequence</u> of the energy efficiency improvement, which include the mechanisms listed in Box 1.1.

A diagrammatic representation of this classification scheme is provided below (see also the *Supplementary Note*). The relative size of each effect may vary widely from one circumstance to another and in some cases individual components of the rebound effect may be negative. For example, if an energy service is an 'inferior good', the income effect for consumers may lead to reduced consumption of that service, rather than increased consumption. It is theoretically possible for the economy-wide rebound effect to be negative ('super conservation'), although this appears unlikely in practice.



Box 2.2 Indirect rebound effects

- The equipment used to improve energy efficiency (e.g. thermal insulation) will itself require energy to manufacture and install and this 'embodied' energy consumption will offset some of the energy savings achieved.
- Consumers may use the cost savings from energy efficiency improvements to purchase other goods and services which themselves require energy to provide. For example, the cost savings from a more energy efficient central heating system may be put towards an overseas holiday.
- Producers may use the cost savings from energy efficiency improvements to increase output, thereby increasing consumption of capital, labour and materials inputs which themselves require energy to provide. If the energy efficiency improvements are sector wide, they may lead to lower product prices, increased consumption of the relevant products and further increases in energy consumption.
- Cost-effective energy efficiency improvements will increase the overall productivity of the economy, thereby encouraging economic growth. The increased consumption of goods and services may in turn drive up energy consumption.
- Large-scale reductions in energy demand may translate into lower energy prices which will encourage energy consumption to increase. The reduction in energy prices will also increase real income, thereby encouraging investment and generating an extra stimulus to aggregate output and energy use.
- Both the energy efficiency improvements and the associated reductions in energy prices will reduce the price of energy intensive goods and services to a greater extent than non-energy intensive goods and services, thereby encouraging consumer demand to shift towards the former.

Energy efficiency improvements by both consumers and producers therefore initiate a chain of effects that have repercussions throughout the economy. The relevant mechanisms are individually complex and mutually independent and are likely to vary in importance from one type of energy efficiency improvement to another.

Energy efficiency improvements reduce the cost of the outputs from energy conversion devices (e.g. lighting, steam), which may be referred to generically as 'useful work'.[1] Energy efficiency improvements in consumer technologies may contribute to human labour being substituted by the 'useful work' derived from fossil fuels (e.g. dishwashers may replace washing dishes by hand). They may also increase the real income of consumers, allowing them to purchase additional goods and services that also require energy for their provision. In a similar manner, energy efficiency improvements in production technology may lead, over time, to useful work being substituted for other factors of production as well as reducing the unit cost of production. If the energy efficiency improvement is confined to an individual, price-taking firm, this should allow the firm to increase output and capture a larger share of the market. If the energy efficiency improvements are sector wide, they should lower output prices and increase demand for the relevant product – whether from domestic consumers or for export. Lower output prices, in turn, will lower the input costs for other production sectors with corresponding secondary effects. For example, a shift from blast furnaces to (more energy efficient) electric arc furnaces may reduce the cost of

[1] See the *Technical Report 2* for a discussion of the relationship between useful work and energy services.

producing steel and hence the price of steel, which in turn should lower the cost of producing cars and hence reduce car prices, thereby increasing the demand for cars – and so on.

Economy-wide improvements in the energy efficiency of production technology may also have a 'composition effect', By lowering the price of energy intensive goods and services to a greater extent than non-energy intensive goods and services, consumer demand may shift towards patterns that increase aggregate energy consumption. Moreover, in so far as energy efficiency improvements lead to a reduction in the price of energy, this may stimulate an offsetting increase in energy demand. The reduction in energy prices will also increase real income, thereby encouraging investment and generating an extra stimulus to aggregate output and energy use.

If energy forms a small part of total costs, these secondary effects may be relatively small. However, a crucial that frequently overlooked point is that improvements in energy efficiency rarely occur in isolation, but typically as part of (and arguably a necessary condition for) broader improvements in processes and products. For example, a desktop PC is more energy efficient than a mainframe computer, but also represents an entirely new product (Saunders, 2000b). Such innovations open the door for entirely new economic opportunities and industrial expansion that may itself increase energy use:

> "….The electric power required to operate a single computer chip nowadays is negligible... but there are now 50 million PCs in private households, each consuming a kilowatt of power, for an average of 12 hours a week. There are 150 million more PCs in businesses, probably being used even more intensively. Annual sales are now 36 million units. Chip manufacturing is also very power intensive: approximately 1000 kWh is needed to fabricate each PC. Amazingly, the silicon fabricators and their suppliers already consume 1% of the electric power consumed in the US. Electric power required to operate the Internet PCs and their more powerful cousins (workstations, routers, etc) and associated networks adds another 8%, while non-network computers brings the total to 13%. Evidently, the rebound effect in this case is surprisingly powerful." (Ayres, 1999)[2]

But while improved energy efficiency may form part of - or even provide a precondition for - innovations such as personal computers, it does not follow that all the subsequent increase in energy consumption can meaningfully be attributed to improved energy efficiency. An important area of ambiguity in the existing literature is whether the relevant independent variable for the rebound effect is improvements in the thermodynamic efficiency of individual conversion devices, improvements in more aggregate measures of energy productivity or improvements in overall (total factor) productivity for which the energy efficiency improvements were (perhaps necessarily) associated. These definitional issues are of crucial importance and are discussed at length in this report.

The net impact of a particular energy efficiency improvement in a particular sector will be mediated by a host of variables - including the scope for substitution between different factors of production, the price elasticity of outputs and the share of energy in total costs – making any general statements about the magnitude of economy-wide rebound effects questionable. However, rebound effects may be expected to be greater at the economy-wide level than at lower levels of aggregation, and greater over the long-term than over the short-term. In particular, some authors suggest that the contribution of energy efficiency

[2] The estimate for the energy consumption of a PC appears much too high. However, the general point remains unchanged.

improvements to economic growth could be the key to understanding their net impact on energy demand (Ayres and Warr, 2005)

The correlation between aggregate energy consumption and GDP in both industrialised and developing countries is undeniable. While industrial countries appear to have partly decoupled growth in energy consumption (measured in kWh heat content) from growth in GDP in recent years, the close correlation reappears when energy consumption is weighted by the 'quality' of different fuel types (Cleveland*, et al.*, 2000). When the shift towards higher quality forms of energy (e.g. electricity) is taken into account, energy use and the level of economic activity appear to be tightly coupled, despite ongoing improvements in energy efficiency (Kaufmann, 1992). The importance of such quality adjustments in aggregate measures of energy consumption and energy efficiency is a central theme of this report.

As will be discussed in detail below, the ratio of energy inputs to economic output – at any level of aggregation - may be influenced by a host of factors other than changes in the thermodynamic efficiency of individual energy conversion devices. For example, a shift towards a 'service' economy may lower the energy-GDP ratio, independently of any changes in thermodynamic efficiency. Nevertheless, work by both Ayres and Warr (2002b) and Dahmus and Gutowski (2005) has suggested a strong correlation between historical trends in aggregate energy consumption in the US and measures that are more closely related to thermodynamic efficiency. For example, Table 2.1 compares estimates of the rate of change of energy efficiency within four US industrial sectors with the rate of change in overall output. Here, energy efficiency for the first three sectors is measured in tonnes of output per kWh of energy inputs, while for electricity generation output is measured in kWh of electricity generation per kWh of energy inputs. In each case, the average annual increase in production has exceeded the average annual increase in energy efficiency by factors ranging from 1.6 to 11.4.

Table 2.1 Average annual increase in production compared to average annual improvement in energy efficiency in four US industrial sectors

Product	**Time period**	**Average annual** $\Delta Y / Y$	**Average annual** $\Delta\varepsilon / \varepsilon$	**Average annual** $(\Delta Y / Y)/(\Delta\varepsilon / \varepsilon)$
Pig iron	1800-1984	4.1%	1.1%	3.7
Aluminium	1900-1987	11.1%	1.0%	11.4
Nitrogen fertilizer	1930-1989	7.1%	4.4%	1.6
Electricity generation				
Coal	1920-2000	4.6%	1.4%	3.3
Oil	1920-2000	5.3%	1.7%	3.0
Natural gas	1920-2000	7.8%	1.8%	4.4

Source: Dahmus and Gutowski (2005)

A historical perspective on rebound effects is provided by Fouquet and Pearson (2006), who present some remarkable data on the price and consumption of lighting services in the UK over a period of seven centuries (Table 2.2). Per capita consumption of lighting services grew much faster than per capita GDP throughout this period, owing in part to continuing reductions in the price of lighting services (£/lumen hour). This, in turn, derived from continuing improvements in the thermodynamic efficiency of lighting technology, in combination with continuing reductions in the real price of lighting fuel (itself, partly a consequence of improvements in the thermodynamic efficiency of energy supply). In this case, improvements in lighting technology were substantially more important than improvements in energy supply (in the ratio of 180 to 1 over the period 1800 to 2000).

Per capita lighting consumption increased by a factor of 6566 between 1800 and 2000, largely as a consequence of the falling cost of lighting services relative to income, but also as a result of the boost to per capita GDP provided by the technical improvements in lighting technology. Since lighting efficiency improved by a factor of 1000, the data suggest that per capita energy consumption for lighting increased by a factor of six. In principle, the direct rebound effect could be estimated by constructing a counterfactual scenario in which lighting efficiencies remained at 1800 levels. But this would be a meaningless exercise over such a time interval, given the co-evolution and interdependence of the relevant variables and the interrelationship between energy consumption and economic growth. To the extent that the demand for lighting is approaching saturation in many OECD countries, future improvements in lighting efficiency may be associated with smaller rebound effects. Nevertheless, this historical perspective gives cause for concern over the potential of technologies such as compact fluorescents to reduce energy consumption in developing countries.

Table 2.2 Seven centuries of lighting in the UK

Year	**Price of lighting fuel**	**Lighting efficiency**	**Price of lighting services**	**Consumption of light per capita**	**Total consumption of light**	**Real GDP per capita**
1300	1.50	0.50	3.0	-	-	0.25
1700	1.50	0.75	2.0	0.17	0.1	0.75
1750	1.65	0.79	2.1	0.22	0.15	0.83
1800	1.0	1	1	1	1	1
1850	0.40	4.4	0.27	3.9	7	1.17
1900	0.26	14.5	0.042	84.7	220	2.9
1950	0.40	340	0.002	1528	5000	3.92
2000	0.18	1000	0.0003	6566	25630	15

Note: 1800=1.0 for all indices

Source: Fouquet and Pearson (2006)

Unfortunately, this type of time series of is difficult to construct, so relatively little research has investigated the causal links between improvements in thermodynamic efficiency and more aggregate measures of economic output and energy consumption. In the examples cited above, a key question is to what extent the growth in economic output is the *cause* of the increased energy consumption and/or improvements in energy efficiency and to what extent the falling cost of 'useful work' (i.e. the outputs from energy conversion devices) is a contributory or primary cause of the growth in economic output (Toman and Jemelkova, 2003). In practice, there is likely to be a synergistic relationship between the two, with each causing the other as part of a positive feedback mechanism (Ayres and Warr, 2002b). The falling cost of useful work derives in part from energy efficiency improvements in both energy supply and energy use. But the relative importance of such improvements compared to other factors remains an empirical question.

The economic measure of energy efficiency is 'energy productivity', defined as the ratio of a monetary measure of economic output to either a physical or monetary measure of energy inputs (see Section 3). Both neoclassical and ecological economists have explored the relationship between improvements in energy productivity and economic growth and they appear to have arrived at different conclusions (Stern, 1993). Generally speaking, neoclassical authors have concluded that improved energy productivity plays a relatively minor role in economic growth (Denison, 1985; Gullickson and Harper, 1987), while ecological authors have concluded that it plays a dominant role (Beaudreau, 1995b; Ayres, 2002; Ayres and Warr, 2005). While these differing conclusions partly reflect differing assumptions, they are also supported by detailed and conflicting empirical evidence. The quantification of economy-wide rebound effects may hinge in part upon a satisfactory resolution of this complex debate.

In the view of Ayres and Warr (2005), economic growth is best seen as a positive feedback cycle in which improvements in thermodynamic conversion efficiency play an important role. Cheaper factor inputs that result, in part, from energy efficiency improvements enable goods and services to be produced at less cost and lower prices, leading to higher demand. Since demand for goods and services correspond to the sum of factor payments, most of which go back to labour as wages, it follows that the wages tend to increase as output rises. This in turn stimulates the further substitution of capital and energy for labour in both manufacturing and the 'household' production of energy services. This substitution encourages scale and learning economies which lower costs further and serve to perpetuate the 'growth engine'. However, energy efficiency improvements are only one source of cheaper factor inputs and in practice are difficult to isolate from broader technical change. Hence, their relative importance in driving economic growth remains an empirical question.

2.2 Jevons' paradox

The point of reference for practically all discussion of the economy-wide rebound effect is W.S Jevons *The Coal Question*, published in 1863 (Jevons, 1865). In the course of his investigation of UK coal reserves, Jevons developed a pioneering and coherent argument in favour of 'backfire'. In an oft cited passage he argued that:

> "… it wholly a confusion of ideas to suppose that the economical use of fuel is equivalent to a diminished consumption. The very contrary is the truth.......Every…improvement of the engine when effected will only accelerate anew the consumption of coal…" (Jevons, 1865)

This increase in fuel demand derived in part from the expansion of existing uses and in part from the development of new uses:

> "… if the quantity of coal used in the blast furnace, for instance, is diminished in comparison with the yield, the profits of the trade will increase, new capital will be attracted, the price of pig-iron will fall, but the demand for it increase; and eventually the greater number of furnaces will more than make up for the diminished consumption of each" (Jevons, 1865)

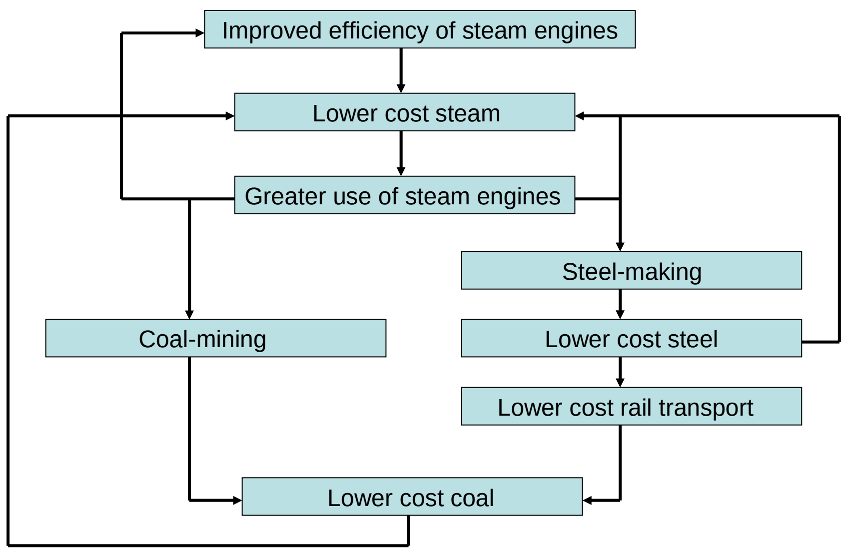
> "…..Whatever… conduces to increase the efficiency of coal, and to diminish the cost of its use, directly tends to augment the value of the steam engine, and to enlarge the field of its operations."

Jevons cites the example of the Scottish iron industry, in which:

> "..... the reduction of the consumption of coal, per ton of iron, to less than one third of its former amount, has been followed….by a tenfold increase in total consumption, not to speak of the indirect effect of cheap iron in accelerating other coal consuming branches of industry…" (Jevons, 1865)

A specific empirical example cited by Jevons was Savory and (later) Newcomen's development of an engine for pumping floodwater out of coal mines. This "…consumed no coal because its rate of consumption was too high." Jevons argues that it was only with the subsequent efficiency improvements by Watt and others that pumping 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 heated air into blast furnaces, thereby increasing the blast temperatures, reducing the quantity of coal needed to make iron and reducing the cost of iron (Ayres, 2002). Lower cost iron, in turn reduced the cost of steam engines, creating a positive feedback cycle (Figure 2.1). It also contributed to the development of railways, which lowered the cost of transporting coal and iron, thereby increasing demand for both.

Figure 2.1 Energy efficiency, positive feedbacks and economic growth



Jevons highlights the fact that improvements in the thermodynamic efficiency of steam engines were intertwined with broader technical change, including: "…. contrivances, such as the crank, the governor, and the minor mechanism of an engine, necessary for regulating, transmitting, or modifying its power…." (Jevons, 1865). Since such developments were essential to the increased use of steam engines as a source of motive power, it is misleading to attribute the increase in coal consumption to improvements in *thermodynamic* efficiency alone (Feather, 2005). But to the extent that improvements in thermodynamic efficiency are inseparable from other improvements in the relevant technology, attempt to draw such a distinction can also be misleading. This relationship between 'narrow' improvements in thermodynamic efficiency and 'broader' improvements in technology and productivity is a central theme of this report and appears crucial for the rebound debate.

The mechanism through which improved energy efficiency stimulates new uses for products has also been cited by Rosenberg (1989):

> "[the Bessemer process] 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."

The low cost Bessemer steel initially found a large market in the produc of tion of steel rails, thereby facilitating the growth of the rail industry, and later in a much wider range of applications including automobiles. It appears to have played an important role in driving economic growth in the 19th century. However, the mild steel produced by the Bessemer process is a very different product to wrought iron (which has a high carbon content) and is suitable for a much wider range of applications (Feather, 2005). Hence, once again, the improvements in the thermodynamic efficiency of production are necessarily entwined with broader developments in process and product technology.

Ayres (2002) provides several more recent examples of this process. For example, prior to the introduction of the Hall-Heroult process in 1887, aluminium was expensive and used solely for decorative uses. The process cut the price of aluminium by more than half in three years, thereby expanding the market, encouraging investment, scale economies and product innovation, and leading to a further increase in demand. The energy requirements for electrolytic smelting fell from 50kWh/kg in 1988 to around 13kWh/kg today and have been associated with continued growth in the demand for both aluminium and electricity for the smelting process (Ayres*, et al.*, 2003). This suggests that improvements in energy efficiency have increased rather than reduced overall energy consumption, as well as encouraging economic growth. But a host of other technical improvements have also contributed to the reduced cost of aluminium products, while advances in areas such as heat treatment have greatly expanded its range of uses. So once again, the contribution of improvements in thermodynamic efficiency appear difficult to isolate from broader technical changes

In addition to empirical investigations into the production and use of coal, Jevons based his conclusions on improvements in labour productivity during the preceeding century, which appeared to have increased, rather than reduced, total employment:

> "….The economy of labour effected by the introduction of new machinery throws labour out of employment for the moment. But such is the increased demand for the cheaper products, that eventually the sphere of employment is greatly widened. Often the very labourers whose labour is saved find their more efficient labour more demanded than before. Seamstresses, for instance, have perhaps in no case been injured, but have often gained wages before unthought of, by the use of the sewing machine..." (Jevons, 1865)

This echoes the views of both predecessors and contemporaries of Jevons, including Malthus:

> "When a machine is invented, which, by saving labour, will bring goods into the market at a much cheaper rate than before, the most usual effect is such an extension of the demand for the commodity, by its being brought within the power of a much greater number of purchasers, that the value of the whole mass of goods made by the new machinery greatly exceeds their former value; and, notwithstanding the saving of labour, more hands, instead of fewer, are required in manufacture." (Sherwood, 1985)

The analogy between contemporary debates about energy productivity and 19th-century debates about labour productivity has also been highlighted by Khazzoom (1980) - and it appears compelling. At first sight, it is not obvious why improvements in the productivity of one input (labour) are generally acknowledged to increase demand for that input (economy-wide and over the long-term), while improvements in the productivity of a second (energy) are generally expected to reduce demand. As a result, Alcott (2006) argues that the debate on the economy-wide rebound effect may simply be reinventing the wheel. However, improvements in efficiency/productivity are linked to rebound effects through the medium of costs and prices. The fact that labour has historically formed a much greater share of production costs than energy could be important, since the cost impacts of improvements in labour productivity are likely to be much greater than those following improvements in energy productivity.

In a similar manner, the examples cited above relate to efficiency improvements in the early stages of development of energy intensive process technologies and intermediate goods, thereby leading to significant reductions in the cost of both. It is possible that the same consequences may not follow for efficiency improvements in mature and/or non-energy intensive process technologies and goods that lead to only very small reductions in costs. Similarly, the same consequences may not follow from improvements in consumer technologies that supply energy services with a low own-price elasticity and where energy represents only a small share of total costs. (Ayres, 2002). Saunders' formulation of the Khazzoom-Brookes postulate implies that *all* economically justified energy efficiency improvements increase energy consumption. But it seems more likely that the size of the rebound effect depends upon the particular nature and location of the energy efficiency improvement.

3 Measures of energy efficiency and productivity

Many commentators on the rebound effect assume that the independent variable is some improvement in the thermodynamic efficiency of an individual energy conversion device. However, much of the evidence cited in support of the rebound effect refers to measures of 'energy efficiency' that relate to higher levels of aggregation, such as a sector or a national economy and which rely upon physical or economic measures of useful outputs. Such measures may be better described as measures of *energy productivity,* or its reciprocal *energy intensity,* and may be influenced by a host of factors other than improvements in the thermodynamic efficiency of individual devices. While thermodynamic measures of energy efficiency have their roots in physics and engineering, more aggregate measures of energy productivity are generally informed by economic theory. The relationship between these different measures appears to be a major source of confusion within the rebound debate.

The aim of this section is to clarify the theoretical and methodological issues underlying thermodynamic, physical and economic measures of energy efficiency and energy productivity. An overview of these topics is a necessary prerequisite for the more general discussion of the relationship between energy use, productivity and economic growth, contained in the remainder of this report. Section 3.1 introduces thermodynamic and physical measures of energy efficiency, including the concept of *exergy* and the use of decomposition analysis. Section 3.2 describes different approaches to aggregating diverse energy inputs and advocates a quality-adjusted measure based upon the Divisia index. It also shows how different approaches to measuring aggregate energy consumption can lead to quite different conclusions on the extent to which energy efficiency has improved over time. Section 3.3 introduces economic measures of energy productivity and total factor productivity and shows how these depend upon the prices of other inputs, the level of output and the current state of technology. It uses a simple neoclassical production function to distinguish between price induced substitution of energy for (or by) other inputs and 'autonomous' improvements in energy efficiency as a result of technical change. It also summarises different approaches to measuring 'neutral' or 'biased' technical change, including so-called 'energy-saving' technical change Finally, Section 3.5 summarises and highlights some implications of the above for the measurement of the rebound effect.

3.1 The concept and measurement of energy efficiency

The energy efficiency of a system may be defined broadly as the ratio of useful outputs to energy inputs. But the measures for the numerator and denominator in this expression may vary widely depending upon the purpose in hand, and analysts have frequently used competing definitions for similar purposes. This section summarises and compares the different definitions of energy efficiency and comments on the issues involved.

3.1.1 Thermodynamic measures

The most basic definition of energy efficiency derives from the first-law of thermodynamics and measures the ratio of 'useful' energy outputs (e.g. light energy from a lightbulb) to the heat content, or calorific value of fuel inputs.[3] A conventional lightbulb, for example, has a first-law efficiency of only 6%, since 6% of energy inputs are converted to light energy and the remainder is lost as 'waste' heat (Berndt, 1978; Patterson, 1996). Note that the first-law efficiency of a process depends upon how 'useful' is defined. When waste heat and other losses are taken into account, the first-law efficiency becomes 100%, since energy is not 'used up' but is merely transformed from 'available' to less available forms. Energy conservation is therefore assured by the first-law.

A drawback with 'first-law' efficiency measures is that they do not take into account the 'availability' of energy inputs or outputs, or their ability to perform *useful work* (see below) (Berndt, 1978). 'Work' may be broadly defined as an increase in the kinetic, potential, physical or chemical energy of a subsystem that is located within a larger system in which energy - according to the first-law - is always conserved (Ayres, *et al.*, 2003). As an example, the energy in high-pressure steam has a greater availability to perform work than the same amount of energy in the form of low temperature heat. Measuring energy inputs on the basis of their thermal content effectively implies that their ability to perform useful work is equivalent to their ability to generate low temperature heat. But this 'lowest common denominator' measure overlooks important qualitative differences between energy carriers.

A common measure of the ability to perform useful work is *exergy*, defined as the maximum amount of work obtainable from a system as it comes (reversibly) to equilibrium with a reference environment (Ahern, 1980). Exergy is only non-zero when the system under consideration is distinguishable from its environment through differences in either relative motion, gravitational potential, electromagnetic potential, pressure, temperature or chemical composition (Ayres, 1998a). Exergy therefore provides a general measure of the 'quality' of both energy and material inputs to production, as well as both 'useful' and waste outputs (Ayres, 1998a; Wall, 2004). Unlike energy, exergy is 'consumed' in conversion processes, and is mostly lost in the form of low temperature heat. A heat unit of electricity will be ranked higher on exergy basis than a heat unit of oil or natural gas, since the former can do more useful work. Similarly, a heat unit of oil or natural gas will be ranked higher than a heat unit of coal.

The notion of exergy leads to a second definition of energy efficiency, which compares the actual energy used for a task with the theoretical minimum energy required. This 'second-law' measure of energy efficiency is typically smaller than the first-law efficiency, suggesting a greater potential for efficiency improvement. For example, the first-law efficiency of electric resistance heating may exceed 99%, but this falls to around 5% when a second-law definition is used (Rosen, 2004). The difference arises because resistance heating converts high exergy electricity to low exergy space heat. If a high efficiency electric heat pump is used instead, the same amount of heat may be obtained from only 14% of the electricity

[3] The gross calorific value, of a fuel is defined as the amount of heat released by a specified quantity of fuel (initially at 25°C) once it is combusted and the products have returned to a temperature of 25°C.This takes into account the latent heat of vaporization of water in the combustion products, and is useful in calculating heating values for fuels where condensation is practical (e.g. space heating). The net calorific value, ignores the recovery of latent heat and is defined as the amount of heat released by combusting a specified quantity fuel and returning the temperature of the combustion products to 150 °C.

inputs (Rosen, 2004). Similarly, while the first-law efficiency of gas-fired space heating is approximately 60%, the second-law, or exergy efficiency is only around 10% (Hammond, 2004). However, the same conclusion does not apply to all conversion processes: for example, the first and second-law efficiencies of conventional coal-fired electricity generation are both around 35% (Hammond, 2004).

Most policy relevant literature uses first-law efficiency measures, but second-law efficiency measures may be preferable since they focus attention what needs to be conserved - namely exergy rather than energy per se (Berndt, 1978). However, while exergy analysis is increasingly common at the engineering process level, it has been slow to gain acceptance for sector or economy-wide efficiency comparisons - despite an increasing body of work in this area (Wall, 1990; Hammond and Stapleton, 2001; Dincer, 2002). Perhaps the most significant contributions have been by Ayres *et al* (2003), who have developed a time series for the second-law (exergy) efficiency of the US economy throughout the 20th century. Ayres *et al* (2003) use the term *useful work* for the exergy content of the useful outputs of conversion devices, and estimate time series for the second-law conversion efficiencies of electric power generation, transportation, high and medium temperature process heat, and low temperature space heat in the US. Table 3.1 illustrates the substantial improvement in these efficiencies that have occurred over the last hundred years. Ayres *et al* (2005) also estimate a comparable time series for the second-law conversion efficiency of electric power to useful work. They show that individual applications (e.g. lighting, motor drives) have become considerably more efficient during the past century, but the overall second-law efficiency of electricity use has remained almost constant owing to the least efficient applications (low temperature space heat and fractional horsepower motors) increasing their share of total electricity consumption.

Table 3.1 Average second-law efficiency of primary conversion processes in the US

Year	**Electricity generation and distribution**	**Transportation**	**High temperature process heat (steel)**	**Medium temperature process heat (steam)**	**Low temperature space heat**
1900	3.8	3.0	7	5	0.25
1970	32.5	8.0	20	14	2
1990	33.3	13.9	25	20	3

Source: Ayres *et al* (2003)

Whether first or second-law measures are used, thermodynamic indicators tend to be confined to direct energy use and do not take into account the indirect energy use in the provision of capital and labour. So, for example, energy efficiency measures for domestic heating systems do not generally take into account the energy required to manufacture, deliver and install the system. Also, maximising either first or second-law efficiency is inappropriate from an economic perspective, since it is necessary to take into account the costs associated with other inputs such as capital and labour (Berndt, 1978).

3.1.2 Physical measures

For many purposes, it is simpler to measure useful energy outputs in terms of tangible physical indicators for the relevant energy service, rather than heat content or exergy. For example, a suitable output measure for personal transportation by private car could be vehicle kilometres or passenger kilometres. A physical measure of energy efficiency could then be vehicle kilometres per litre of gasoline. This example employs a volumetric measure of energy inputs, but measures based on thermal content are more commonly applied in situations where multiple fuels are used: e.g. litres of beer produced per kWh of energy inputs. It is rare, however, to combine a physical measure of useful energy outputs with an *exergy* measure of energy inputs.

Physical measures are often specified in terms of *energy intensity* (i.e. the inverse energy efficiency) and referred to as measures of unit energy consumption (*UEC*) or specific energy consumption (*SEC*). Physical measures may be applied at the level of the individual energy conversion device, but are more commonly applied at higher levels of aggregation, such as industrial processes, individual firms or individual sectors. In each case, changes in these physical measures may result from factors other than improvements in the thermodynamic efficiency of conversion devices. For example, a common physical measure of energy intensity in freight transport is the ratio of amount of fuel used by freight vehicles (litres of diesel) to the weight of goods transported (tonnes). But changes in this ratio may result from changes in the type and mix of goods transported, the average distance travelled for each good, the amount of packaging used, the mix of vehicles within the overall fleet, the average load factors of vehicles and the degree of empty running, as well as from improvements in vehicle fuel efficiency. Techniques of decomposition analysis has been developed to estimate the relative contribution of different variables to aggregate changes in these physical efficiency indicators, and these are now both methodologically sophisticated and widely employed (Ang, 1999). Their potential contribution to estimates of the rebound effect is described further in Section 4.5.

Appropriate physical indicators for outputs are likely to vary from one product to another and one type of energy service to another, making an aggregate economy-wide physical measure inappropriate. Moreover, there may be difficulties in using these indicators at the sector level owing to problems of joint production: for example, energy inputs into sheep farming are used to produce wool and meat simultaneously, making it difficult to assign energy inputs to either (Patterson, 1996). Also, unlike thermodynamic measures of energy efficiency, physical measures are not constrained to be less than unity.

3.1.3 Economic measures

By replacing the numerator with an indicator of the economic value of output, the energy efficiency of different sectors can be compared. For example, the energy efficiency of both the brewing and dairy sectors can be measured in terms of value added per GWh of energy input. With economic indicators, it is more common to refer the ratio of output to energy inputs (*Y/E*) as *energy productivity* rather than energy efficiency, or to refer to its inverse (*E/Y*) as energy intensity. Again, unlike thermodynamic measures of energy efficiency, economic measures of energy productivity are not constrained to be less than unity

As with physical indicators, changes in economic indicators may result from a host of factors other than improvements in the thermodynamic efficiency of individual conversion devices. The move from physical to economic indicators increases the number of influencing factors, as does the use of such indicators at higher levels of aggregation. The indicator that is furthest from a thermodynamic measure of energy efficiency is therefore the ratio of GDP to total primary energy consumption within a national economy. Nevertheless, the inverse of this measure - the energy/GDP ratio - has been widely employed as an energy intensity indicator since the early 1970s (Ang, 2006). Typically, energy inputs are measured on a thermal content (rather than exergy) basis and, for cross-country comparisons, GDP is measured on either an exchange rate converted basis, or in terms of purchasing power parities.

The final, logical step is to measure the denominator (energy inputs) in terms of economic value, as well as the numerator, to give a purely economic ratio. The denominator may either reflect actual expenditures on energy commodities, or the total kWh consumption of each fuel may be multiplied by some estimate of their average or marginal productivities (Adams and Miovic, 1968b; Kaufmann, 1994). But to appreciate the issues involved here, it is first necessary to review how multiple energy inputs may be meaningfully aggregated.

3.2 The concept and measurement of aggregate energy

Measures of energy use and energy productivity at higher levels of aggregation need to combine the inputs of multiple energy carriers. For example, measures of total energy use by a national economy must combine the individual uses of different fossil fuels, together with nuclear and renewable electricity. The most common approach is to aggregate inputs according to their thermal content, or ability to generate heat, measured in kWh. As indicated above, this first-law measure is flawed since it neglects the relative ability of different energy sources to perform useful work. In addition, it implicitly assumes that one energy input can be perfectly substituted by another, which is not the case for the majority of end-uses (e.g. a television cannot run directly on coal) (Berndt, 1978).

A better approach, therefore, would be to aggregate energy carriers by their exergy content, or ability to perform useful work. This would give much greater weight to electricity inputs, for example, since they have higher exergy content than fossil fuels. A quality weighted measure of aggregate energy consumption, based upon exergy, would therefore be expected to depend very closely upon the particular fuel mix.

From an economic perspective, it would be better still to weight each unit (kWh) of energy input by some measure of its relative worth. This would give a further weight to electricity inputs, for example, given their relative flexibility, availability, controllability, ease of transport and corresponding higher price per kilowatt hour compared to other forms of energy. If energy markets are competitive, energy prices provide a fairly accurate means of reflecting the differences in the 'quality' of different energy inputs, indicating not just their exergy content (ability to do work) but also attributes such as weight, cleanliness, safety, ameanability of storage, flexibility of use, cost of conversion and so on (Cleveland*, et al.*, 2000). A measure of aggregate energy inputs could then weight each individual energy input (*i*) by its price (P_i) relative to a numeraire (Berndt, 1978). For example, if the price per kWh of the first energy type (P_1) is taken as a numeraire, a measure of aggregate energy inputs (E^*) may be computed from the individual (kWh) inputs of each energy type (E_i) as follows:

$$E^* = \sum_{i=1,n} \frac{P_i}{P_1} E_i \qquad (3.1)$$

This formulation is still flawed, however, since it assumes that a unit of energy type 1 is completely equivalent to (i.e. perfectly substitutable with) P_2/P_1 units of energy type 2 - which is an unrealistic assumption for many end uses (Berndt, 1978). It is also sensitive to the choice of numeraire (Stern, 1993). Analysts such as Berndt (1978) have therefore proposed an alternative aggregation formula, based upon a discrete approximation to the

Divisia index. This in turn is based upon more general work by Diewert (1976) and others in the area of index number theory that is widely used in the aggregation of diverse economic inputs (Jorgenson and Griliches, 1967).[4] For example, a Divisia index may be used to aggregate the contribution of low and highly skilled workers into a single measure of labour inputs.

The formula for constructing the discrete Divisia index of aggregate energy consumption (E^*) is as follows (Berndt, 1978):

$$\ln E_t^* - \ln E_{t-1}^* = \sum_{i=1,n} w_{it}[\ln E_{it} - \ln E_{it-1}] \qquad (3.2)$$

Where, the weight factors (w_{it}) for each energy type and year are given by:

$$w_{it} = \left[(\frac{1}{2}) \sum_{i=1,n} \left(\frac{P_{it} E_{it}}{\sum_{i=1,n} P_{it} E_{it}} + \frac{P_{it-1} E_{it-1}}{\sum_{i=1,n} P_{it-1} E_{it-1}} \right) \right] \qquad (3.3)$$

The interpretation of Equation 3.2 is that the percentage (logarithmic) change in the aggregate energy index is estimated from the weighted average of the percentage changes in the quantity of each individual energy type. The weights, in turn, are based upon the share of each energy type in total energy costs.

One advantage of the Divisia approach is that it allows for variable substitution between different energy types, without imposing any a priori restrictions on the degree of substitution. Since it is built up from weighted rates of change, it allows the quantities and prices of each energy carrier to be measured in different units. It is also closely related to a particular functional form for production and cost functions (the 'translog') that is widely used in empirical work (Christensen*, et al.*, 1975).[5] Furthermore, it is consistent with the methods used for aggregating other economic inputs that are widely employed in productivity analysis and other fields.

The Divisia index does have some drawbacks, however. It is more difficult to construct, since data on prices is required in addition to data on energy consumption. Energy prices vary widely between sector and application and may be distorted by market power, regulatory constraints and other factors. Similarly, the quality of some energy carriers, notably coal, may vary between different uses and over time. Also, the suitability of the Divisia index will depend in part upon the nature of the application. For example, it may be more appropriate for studies of the economic importance of energy in economic growth, than for studies of the relationship between energy use and carbon emissions.

The implications of using a 'quality weighted' measure of energy inputs are discussed further in Section 4. But it should be apparent that aggregate measures of energy efficiency (e.g. at the sector level) are likely to be significantly influenced by the aggregation method used. For example, the substitution of a kWh of coal with a kWh of electricity will not affect an aggregate measure of energy consumption that is based upon the thermal content of energy carriers, but since electricity as a higher exergy content than coal, such a substitution will

---

[4] Similar Divisia indices also play a prominent role in the decomposition analysis of energy demand trends (Ang, 1999)

[5] See Annex 2 for a discussion of different forms of production function.

increase an exergy measure of aggregate energy consumption. Similarly, since the market price of electricity is greater than that for coal, such a substitution will increase a Divisia index of aggregate energy consumption as the value of energy inputs has increased (Zarnikau, 1999). If output is unchanged, a measure of energy efficiency that aggregated inputs by thermal content will also be unchanged, while one that aggregated inputs using either exergy or a Divisia index will have reduced.

As an empirical illustration, Table 3.2 reproduces Hong's (1983) estimates of the annual rate of growth of the energy-GDP ratio of the US economy over the period 1950 to 1978. Prior to 1973, the estimates based upon the thermal content of energy inputs are broadly equivalent to those derived using a Divisia index. However, after 1973 the *thermal* energy-GDP ratio declined as an average annual rate of 2.11%, while the corresponding Divisia energy-GDP ratio declined at an annual rate of only 0.42%. The conventional method of energy aggregation may therefore overstate the improvement in economy-wide energy intensity that followed the energy price rises in 1973.

Table 3.2 Annual percentage growth rates in the energy productivity of the US economy

Period	**Energy input Thermal input basis**	**Energy input Divisia index basis**	**Output**
1950-1965	2.86	2.83	3.76
1965-1973	4.07	4.34	4.04
1973-1978	0.73	2.42	2.84

Source: (Hong, 1983)

These post-1973 figures are illustrative of a more general result: when the quality of energy inputs are accounted for (in whatever way), energy intensities are found to be declining slower than is commonly supposed (Cleveland*, et al.*, 2000). This is because technical progress in energy use is not confined to improvements in thermodynamic efficiency, but also includes the substitution of low quality fuels by high quality fuels - notably electricity. This substitution adds value to consumed energy, thereby both increasing the value of energy inputs and increasing the amount of output obtained (in both useful work and value terms) from the same heat content of input. Aggregate measures of energy efficiency/intensity should therefore be interpreted with considerable caution. To understand further the issues involved, it is necessary to review the basic economic concepts underlying the measurement of productivity.

3.3 The concept and measurement of energy productivity

3.3.1 Individual factor productivity

The economic notion of energy productivity - or its inverse, energy intensity - is based upon neoclassical production theory (Beattie and Taylor, 1993). The starting point is an assumed *production function* which indicates the maximum possible flow of output (*Y*) obtainable from the flow of capital (*K*), labour (*L*), energy (*E*) and materials (*M*) inputs, given the current state of technology, denoted by *A* (Berndt, 1990):

$$Y = f(K, L, E, M; A) \qquad (3.4)$$

This production function may represent an individual process, a firm, a sector or a national economy. It represents an 'efficiency frontier' and many studies implicitly assume that firms or economies are operate at this frontier, despite the multiple examples of real-world inefficiencies (DeCanio, 1997; 2003; Sorrell*, et al.*, 2004). Furthermore, the measurement of all types of input and output raises aggregation and quality issues comparable to those discussed above for energy. For example, labour inputs may be composed of employees of different 'quality' (e.g. low and high skill), making a simple aggregation of total hours/year arguably inappropriate. Hence, in discussing *all* measures of productivity, careful attention must be paid to how the relevant variables are measured.

The economic productivity of an individual factor (θ_i) is then given by the ratio of output to input for that factor, while the factor intensity (τ_i) is given by the ratio of input to output. For example, the productivity of energy inputs is given by:

$$\theta_E = \frac{1}{\tau_E} = \frac{Y}{E} = \frac{f(K, L, E, M; A)}{E} \qquad (3.5)$$

This shows that energy productivity depends upon the level of each input, the current state of technology and the level of output. Under certain assumptions, a *cost function* can also be defined which indicates the minimum possible total cost (*C*) of producing a given level of output, given the prices of each input and the current state of technology:[6]

$$C = g(P_K, P_L, P_E, P_M; Y; A) \qquad (3.6)$$

Empirical studies tend to estimate cost functions more frequently than production functions, since the relevant independent variables (factor prices) can usually be assumed to be exogenous. A fundamental result from production theory (termed 'Shephard's Lemma') is that the cost minimising demand for any input can be obtained from the partial derivative of the cost function with respect to the price of that input (Beattie and Taylor, 1993). So, the cost minimising demand for energy is given by:

$$E = \frac{\partial g(P_K, P_L, P_E, P_M; Y; A)}{\partial P_E} \qquad (3.7)$$

This expression is particularly useful for empirical research. It shows that the optimal demand for energy depends upon the price of each input (P_i), the current state of technology (*A*) and the level of output (*Y*). Hence, a change in any one of these could change the optimal demand for energy and therefore the energy productivity.[7]

This simple neoclassical framework allows us to distinguish between two sources of improvement in energy productivity (i.e. reductions in energy intensity). Improvements following increases in energy prices are likely to result from the *substitution* of labour, materials or (most likely) capital inputs for energy inputs. This may improve energy productivity (or reduce energy intensity - *E/Y*), but at the same time may reduce total output (*Y*) since, if the prices of other inputs are unchanged, the total cost of producing a

---

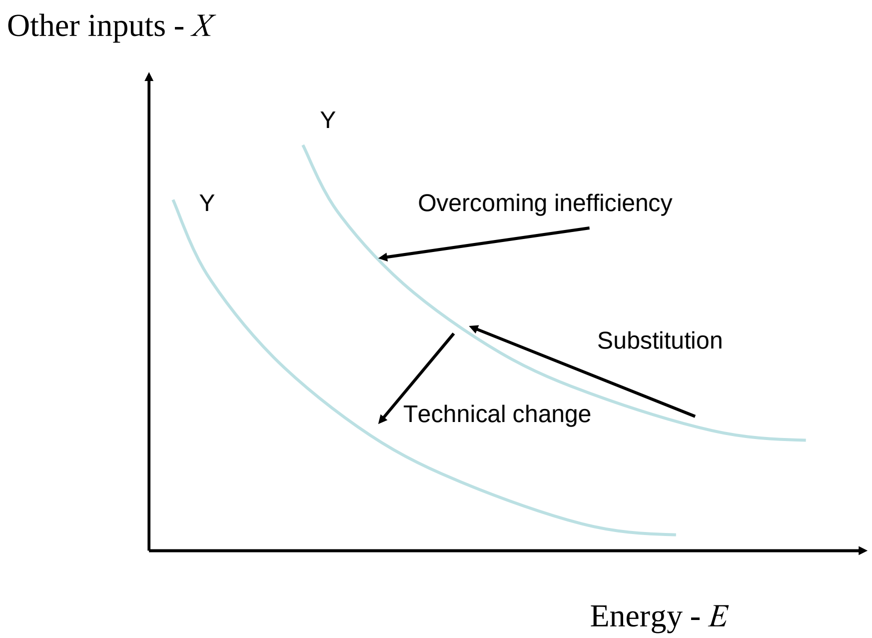
[6] Total costs may also be expressed as the product of output and a unit cost function (*c*). Under the assumption of competitive markets, unit costs should be equal to the unit price of outputs (P_Y): $P_Y = C / Y = c(P_K, P_L, P_E, P_M; Y; A)$

[7] Berndt (1978) shows that the elasticity of energy productivity with respect to a change in the price of the *i*th input is equal to the simple cross price elasticity for that input $\partial \ln \theta_E / \partial \ln P_i = \partial \ln E / \partial \ln P_i$. In the case of a change in the price of energy, the elasticity of energy productivity is equal to the own-price elasticity of the demand for energy. Since there are numerous empirical estimates of own and cross price elasticities for energy demand in different sectors, the corresponding elasticities of energy productivity can easily be estimated. For example, if the own-price elasticity for the demand for energy is equal to -0.5, then a 1% increase in the price of energy should reduce average energy productivity by 0.5% (Berndt, 1978).

given level of output will have increased. Conversely, improvements in energy productivity may also result from *technical change*, represented by changes in the factor *A*. These improvements are assumed to occur independently of any changes in relative prices and are desirable since they occur without any reduction in economic output.

As illustrated in Figure 3.1, substitution may be represented a movement *along* an isoquant of a production function, in which the level of output is held constant. Substitution may require investment in technologies that can combine factors in different ways (e.g. energy-efficient motors), but these are assumed to be chosen from a set of existing technologies. In contrast, 'technical change' refers to the development of new technologies and methods of organisation that *shift* the isoquant to the left, allowing the same level of output to be produced from a lower level of inputs. The measure relevant to the first source of (single factor) productivity improvement is the *elasticity of substitution* between two factor inputs (examined in detail in *Technical Report 3*), while the measures relevant to the second source of productivity improvement are *total factor productivity* and the observed *bias* in technical change. Both of these are introduced below.

Figure 3.1 Substitution, neutral technical change and overcoming inefficiency



The term 'technology' is not being used here to refer to an individual device, but instead to represent the set of possible combinations of factor inputs that may be theoretically combined to produce a given level of output - as represented by the isoquant of the production function. Moreover, the production function represents the most efficient combination of factor inputs and in practice firms may use relatively inefficient combinations. This is therefore an abstract approach and the neat conceptual distinction between substitution and technical change can be difficult to maintain in practice (Sue Wing, 2006).[8] In practice, much investment may represent a move from less efficient to more efficient combinations, while still remaining within the production function 'frontier'- as indicated by the 'overcoming inefficiency' arrow in Figure 3.1.

Much of the energy efficiency literature focuses on the scope for public policy to overcome non-price barriers to energy efficiency and thereby achieve the third type of improvement indicated in Figure 3.1 (DeCanio, 1997; Sorrell, *et al*., 2004). However, many conventional approaches (e.g. CGE modelling) neglect such opportunities and implicitly assume that the economy is working at the 'efficiency frontier' represented by the isoquant of the production function. Hence, while the neo-classical approach is insightful, it also has some important conceptual limitations that are only addressed in more recent literature.

3.3.2 Total factor productivity

The cost function implies that firms minimise the total cost of inputs for a given level of output, with the optimal input mix depending upon relative prices. This means that maximising the productivity of an individual input, such as energy, is inappropriate, since it does not take into account the costs associated with increasing the level of other inputs. A measure of the productivity with which all factor inputs are used is termed total factor productivity (*TFP*), and defined as the rate of growth of output minus the weighted sum of the rate of growth in inputs:[9]

$$TFP_f = \frac{\partial \ln Y}{\partial t} - \frac{\partial \ln I}{\partial t} \quad (3.8)$$

Where I represents an aggregate measure of total inputs, formed in a similar manner to the aggregate measure of energy inputs discussed in section. Typically, each input is weighted by its share in the value of output.[10] Under certain assumptions, a corresponding equation derived from the cost function may also be employed:

$$TFP_g = \frac{\partial \ln X}{\partial t} - \frac{\partial \ln P_Y}{\partial t} \quad (3.9)$$

Where X represents an aggregate measure of total input costs and P_Y represents the output price. Note that the term total factor productivity is normally used to refer to the *rate of growth* of output minus the growth of inputs, rather than the ratio of output to aggregate inputs. The interpretation is therefore different to that given above for measures of single factor productivity.

Unlike changes in individual factor productivity, improvements in total factor productivity are always desirable, since they indicate that more output is being obtained from the same quantity of inputs (Berndt, 1990). Fabricant (1954) and others have developed 'growth accounting' techniques to allow the contribution of growth in factor inputs to increases in

---

[8] The notion of substitution implies a 'frictionless' move from one existing technique to another, but in practice this will take time. Also, to classify an observed shift in technique as substitution, rather than technical change, it is necessary to discern whether the technique that is used was available before the price change, and to clarify what 'available' means - both of which may be problematic (Sue Wing, 2006).

[9] Conventional economic accounting may overlook the contribution of some inputs, such as natural resources. In recognition of this, the term multifactor productivity (MFP), rather than total factor productivity, is sometimes employed. The omission of such inputs can bias empirical estimates of *TFP*.

[10] For example: $\frac{\partial \ln I}{\partial t} = \sum_{i=1,n} \xi_i \frac{\partial \ln i}{\partial t}$, where $\xi_i = iP_i / YP_Y$

output to be estimated (Denison, 1985). These demonstrate that growth in factor inputs is insufficient to explain the phenomenal growth in economic output that has been observed through time at different levels of aggregation. For example, early growth accounting exercises found that the rate of capital accumulation per person accounted for little more than one eighth of the GDP growth rate in the United States and other industrialised countries. Solow (1957) was the first to attribute the residual increase in output to 'technical change', and to represent it by a function of time (*A(t)*) that served as a exogenous multiplier to the production function. Technical change is a misleading term however, since it is a shorthand expression for any change that leads to a shift in the production function (Figure 3.1) and therefore may also include social, organisational and managerial factors (Solow, 1957).

The aggregation of diverse capital and labour inputs raises similar issues to those described above for energy, and later growth accounting exercises have shown how the proportion of output growth that is attributed to 'technical change' depends upon how the factor inputs are measured (Jorgenson and Griliches, 1967).[11] Such 'quality adjustment' is now a standard feature of productivity analysis, but appears to have been applied less frequently to energy inputs, where measurement in terms of thermal content remains common.

Standard growth accounting techniques estimate total factor productivity as the *residual* output growth that is not explained by the growth of factor inputs. But the choice of weightings for aggregating input quantities depends upon implicit assumptions about the underlying production function, and hence can be sensitive to the specification that is used (Mawson*, et al.*, 2003).[12] For estimating productivity trends at the sector level, a preferred approach is to use econometric techniques to estimate the parameters of a production or cost function directly. In addition to obtaining estimates of total factor productivity, the econometric approach can provide valuable information on the possibility for substitution between factor inputs, the existence of scale economies and the nature and direction of technical change.[13] This approach leads to an alternative definition of total factor productivity, namely the change in output resulting from 'technical change', keeping other input quantities fixed:

$$TFP_f = \frac{\partial \ln Y}{\partial \ln A(t)} = \frac{\partial \ln f(K, L, E, M, A(t))}{\partial \ln A(t)} \qquad (3.10)$$

Technical change will also reduce total costs for a given level of output. A 'dual' expression may therefore be derived from the cost function (Berndt, 1978):

$$TFP_g = -\frac{\partial \ln C}{\partial \ln A(t)} = -\frac{\partial \ln g(P_K, P_L, P_E, P_M; Y; A)}{\partial \ln A(t)} \qquad (3.11)$$

---

[11] For example, Jorgenson and Giriliches (1967) developed a quality adjusted measure of labour inputs that accounted for changes over time in educational attainment, and a quality adjusted measure of capital inputs that accounted for utilisation levels and shifts away from long-lived physical structures and towards durable equipment. Since the economic life of buildings is longer than that for equipment, a pound spent on acquiring new equipment should yield more services in the year of acquisition than a pound spent on acquiring new buildings. Hence, the latter should be given greater weight within the aggregate of capital inputs. Taken together, these adjustments reduced the proportion of US economic output growth attributed to technical change from 90% to 50%.

[12] An alternative approach that does not require specification of a production function is to divide an output quantity index by an input quantity index, but this approach can be sensitive to the method of indexing used (Mawson*, et al.*, 2003).

[13] The 'translog' functional form (see Annex 2) is commonly employed for this purpose, since it provides flexibility in the extent to which one input can substitute for another.

Here, input prices and the level of output are held constant. These two measures of total factor productivity can be shown to be related as follows (Ohta, 1974; Berndt, 1978):

$$TFP_g = \frac{1}{(\partial \ln C / \partial \ln Y)} TFP_f \qquad (3.12)$$

The first term on the right hand side in this equation is a measure of the returns to scale in the cost function: i.e. the proportional change in input costs (*C*) following a proportional change in output (*Y*). For example, increasing (decreasing) returns to scale means that a doubling of output requires less than (more than) a doubling of input costs. Under the standard neoclassical assumptions of perfect competition, instant adjustment to equilibrium and constant returns to scale, the rate of growth in output should be equal to the rate of reduction in costs ($\partial \ln C / \partial \ln Y = 1$). In these circumstances, the two measures of total factor productivity are equal. It is common to assume constant returns to scale in empirical work since this greatly simplifies the analysis – although the accuracy of this assumption is open to question.

The econometric estimation of *TFP* from production functions is prone to bias because producers may adjust to improvements in *TFP* by increasing output, thereby introducing a correlation between *TFP* and input usage. To avoid this problem, it is more common to estimate *TFP* from cost functions, since the relevant independent variables (factor prices) can usually be assumed to be exogenous.

The above analysis is microeconomic and applies best to an individual firm or homogeneous sector. It is less applicable to a national economy that is composed of diverse sectors. Hence, econometric approaches to estimating productivity trends are largely pursued at the sector level.

3.3.3 Biased technical change

While 'neutral' technical change increases the productivity of all inputs by a comparable amount, 'biased' technical change increases the productivity of some inputs more than others (Berndt, 1990). The notion of biased technical change is central to the empirical investigation of energy productivity in general and the rebound effect in particular. However, some confusion may be created by the existence of three competing definitions of biased technical change, namely: the *Hicks* definition; the *factor price bias* definition; and the *factor augmenting* definition. These are briefly described below.

3.3.3.1 Hicks' definition

Hicks (1932) introduced the first definition of biased technical change for the simple case of a production function with only capital and labour inputs (*KL*). In Hicks' definition, technical change is said to be *neutral*, *labour saving* or *capital saving* depending on whether, at a constant capital-labour ratio (K/L), the growth rate of the marginal productivity of labour relative to that of capital either stays constant, decreases, or increases:[14]

$$\frac{\partial \ln(\partial Y/\partial L)}{\partial t} - \frac{\partial \ln(\partial Y/\partial K)}{\partial t} = 0$$ neutral technical change

$$\frac{\partial \ln(\partial Y/\partial L)}{\partial t} - \frac{\partial \ln(\partial Y/\partial K)}{\partial t} < 0$$ labour saving (capital using) technical change

$$\frac{\partial \ln(\partial Y/\partial L)}{\partial t} - \frac{\partial \ln(\partial Y/\partial K)}{\partial t} > 0$$ capital saving (labour using) technical change

3.3.3.2 Binswanger's definition

Hicks' definition of biased technical change is difficult to measure empirically and also difficult to generalise to multifactor production functions - including those that use energy as one of the inputs. Binswanger (1974) therefore proposed an alternative definition of biased technical change, that is suitable for production functions with many inputs, and also more suitable for empirical research. This is the *factor price bias* (ψ_i): a measure of the rate of change in the share of factor costs in the value of output, holding input prices constant:

$$\psi_i = \left.\frac{\partial s_i}{\partial t}\right|_{P_i = \bar{P}_i} \qquad (3.13)$$

Where:

$$s_i = \frac{iP_i}{YP_Y} \qquad (3.14)$$

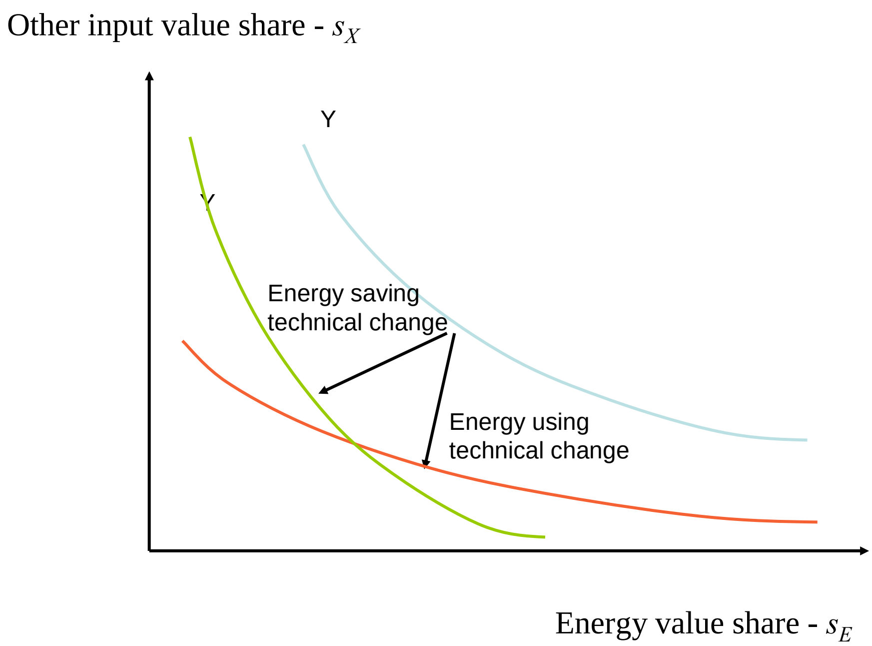
Under the assumption of constant returns to scale and competitive markets, s_i is also equal to the share of factor *i* in total input costs ($s_i = iP_i / C$).

Improvements in total factor productivity may be expected to result in cost savings on all inputs. The factor price bias is therefore a *relative* measure, comparing the cost savings for factor *i* that result from technical change to the corresponding savings on all input costs. If the factor price bias is negative (positive) for factor *i* it implies that the share of *i* in the value of output decreases (increases) over time, irrespective of changes in relative factor prices (Berndt and Wood, 1986). If the factor price bias is negative, technical change is commonly labelled as 'factor-saving' for factor *i* while if the factor price bias is positive, technical change is labelled as 'factor-using' (Jorgenson, 1984). For our purposes, we are particularly interested in whether technical change (under this definition) is 'energy-saving' or 'energy-using'.

[14] An alternative, and equivalent definition, is that technical change is said to be neutral, labour saving or capital saving depending on whether, at a constant capital-labour ratio, the marginal rate of technical substitution between capital and labour either stays constant, decreases, or increases.

A simple diagrammatic interpretation of energy-saving/using technical change is given in Figure 3.2. Improvements in total factor productivity should reduce the use of all factor inputs per unit of economic output. Hence, the amount of energy required to produce a unit of output should reduce over time. If technical change has an energy-saving bias, the degree of reduction should exceed that of other factor inputs, while if it has an energy-using bias, the degree of reduction should be less than for other factor inputs. The change in overall energy productivity will depend upon the relative sign and magnitude of total factor productivity growth and the energy price bias. In some circumstances, the amount of energy required to produce a unit of output may increase over time. This may happen, for example, if improvements in total factor productivity are small, but the energy price bias is positive and large.

Figure 3.2 Energy-saving/using technical change



Binswanger (1974) proposed two ways of estimating factor price biases: one for short-run estimations that assumed that the bias was constant, and a second for long-run estimations that assumed it was not. The first approach is easier to employ, since it allows technical change to be represented by a simple time trend. This approach has been used by Jorgenson and colleagues to explore how changes in energy and other prices may have influenced total factor productivity (Jorgenson, 1981; Hogan, 1991; Jorgenson, 1998).[15] This is examined in detail in Section 4.4.

3.3.3.3 Factor augmenting definition

The third definition of biased technical change is based upon the notion of *augmenting* the inputs to production. This is normally represented through the use of exogenous, time-

[15] With this approach (which is discussed further in Section 4.4) the effect of a change in energy prices on total factor productivity is equal to the energy price bias: $\partial TFP_g / \partial P_E = \theta_E = \partial s_E / \partial t$. Hence, under these assumptions, the magnitude of the energy price bias will influence the impact of increased energy prices on total factor productivity and economic growth.

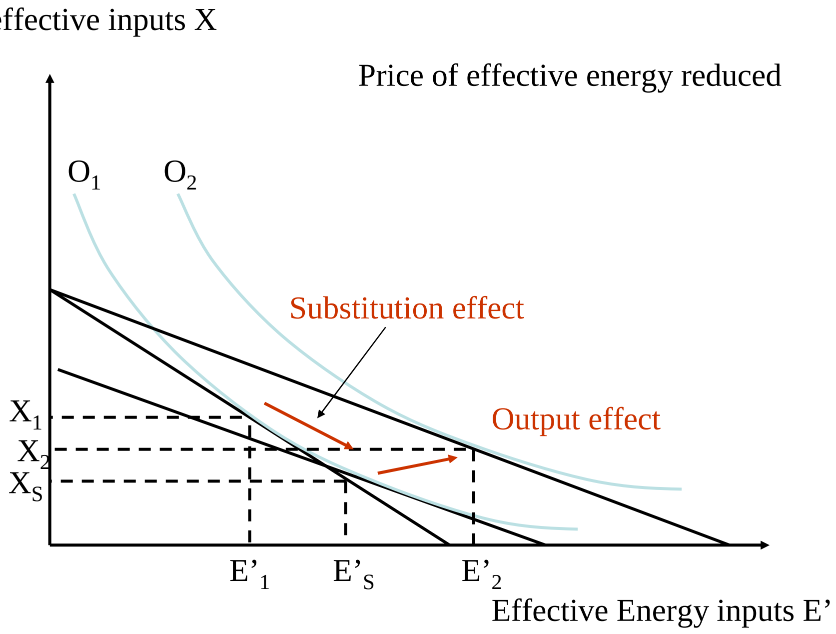
dependent multipliers ($\upsilon_i(t) \geq 1$) on each input factor (*i*) in the production function. For example, the introduction of a multiplier on energy inputs ($\upsilon_E(t)E$) implies that the economic productivity of energy inputs has increased. This means that the same output (*Y*) can now be obtained with less energy inputs, or alternatively that more output can be obtained from the same quantity of energy inputs. In this case, technical progress is said to be 'energy-augmenting'. For a three factor (*KLE*) production function, technical progress may be capital augmenting, labour augmenting, energy augmenting or a combination of all three. This may be represented as follows:

$$Y = f(\upsilon_K(t)K, \upsilon_L(t)L, \upsilon_E(t)E) \qquad (3.15)$$

Frequently, the product $\upsilon_i(t)i$ is referred to as an 'effective' factor input ($\tilde{i}$). The production function then translates 'effective' factor inputs into economic output: $Y = \tilde{f}(\tilde{K}, \tilde{L}, \tilde{E})$ - where $\tilde{K} = \upsilon_K(t)K$, $\tilde{L} = \upsilon_L(t)L$ and $\tilde{E} = \upsilon_E(t)E$. Normally, the multipliers are assumed to be exponential functions of time ($\upsilon_i(t) = e^{-\lambda t}$) with a fixed growth rate ($\lambda_i = \partial \ln v_i / \partial t$).

This framework provides two different ways of viewing the process of neutral or biased technical change. The isoquants of a production function that relates real inputs to economic output (*f*) *shift their position* over time as a consequence of neutral or biased technical change - as in Figure 3.1. In contrast, the isoquants of a production function that relates effective factor inputs to economic output ($\tilde{f}$) remains *unchanged* over time. Technical change reduces the *cost* of effective factor inputs and, if technical change is biased, there will be *substitution* between effective factor inputs since the cost of one effective input reduces faster than the other (for example, energy augmenting technical change will reduce the cost of effective energy inputs to encourage substitution towards effective energy). The reduced cost of effective factor inputs should also lead to an increase in economic output. Figure 3.3 shows that net change in demand for effective factor inputs will be given by the sum of these substitution and output effects. The net change in demand for real factor inputs will be given by the net result of the reduction in real inputs per unit of effective factor input and the change in demand for effective factor inputs.

Figure 3.3 Factor augmenting technical change leads to substitution between 'effective' factor inputs and increases in output



The factor augmenting perspective therefore allows technical change to be viewed as substitution between effective inputs (i.e. movement along the isoquant of production function $\tilde{f}$, as opposed to a leftwards shift of the isoquant of production function *f*). If production function $\tilde{f}$ is homogeneous, neutral technical change may be represented by $\upsilon_K = \upsilon_L = \upsilon_E$. If the production function is not homogeneous, a combination of neutral and biased technical change may be represented as: $Y = \upsilon_N(t)\tilde{f}(\tilde{K}, \tilde{L}, \tilde{E})$, or $Y = \upsilon_N(t) * f(\upsilon_K(t)K, \upsilon_L(t)L, \upsilon_E(t)E)$. However, from the discussion above it should be clear that the neutral technology multiplier ($\upsilon_N(t)$) plays a different role from the other technology multipliers, in that it shifts the isoquants of production function $\tilde{f}$ to the left.

In the factor augmenting framework, the overall growth rate of total factor productivity is given by the weighted sum of the growth rates of the individual factor productivities, with weights being the share of each input in the value of output.[16] Since energy inputs typically represent a small share of total costs, this approach suggests that improvements in energy productivity (υ_E) should, in principle, have a relatively small impact on *TFP*.

16 With this framework, *TFP* may be defined as: $TFP_f = \frac{\partial \ln Y}{\partial t} = \sum_{i=K,L,E} \frac{\partial \ln f(\tilde{i})}{\partial \tilde{i}} \frac{\partial \ln \tilde{i}}{\partial t} = \sum_{i=K,L,E} s_{\tilde{i}} \lambda_i$

Under competitive market conditions, the first term in the summation represents the cost share of factor input *i* in the value of output, while the second term represents the growth rate of the relevant factor augmentation multiplier.

The relationship between these different approaches to defining biased technical change can be a source of confusion. In particular, *energy saving* technical change under the first (Hicks) definition may not necessarily be the same as energy-saving technical change under the second (Binswanger) definition and neither may be equivalent to *energy augmenting* technical change under the third definition (David and van der Klundert, 1965; Acemoglu, 2002). While Binswanger's definition of energy saving technical change is most commonly used in econometric investigations of energy productivity, the notion of energy augmenting technical change plays a more prominent role in CGE modelling (*Technical Report 4*) and neoclassical growth theory (see Section 5). Both will be referred to later in this report.

3.3.4 Autonomous energy efficiency improvements

Of particular importance for our purposes is the rate of growth in energy productivity over time. The measure of the rate of growth of energy productivity, holding relative input prices constant, is commonly termed the 'autonomous energy efficiency index' (*AEEI*) and is normally defined as follows (Sanstad*, et al.*, 2006):

$$AEEI = \left.\frac{\partial \ln \theta_E}{\partial t}\right|_{P_i = \overline{P}_i} = \left.\frac{\partial \ln (Y/E)}{\partial t}\right|_{P_i = \overline{P}_i}$$
$$AEEI = -\left.\frac{\partial \ln \tau_E}{\partial t}\right|_{P_i = \overline{P}_i} = -\left.\frac{\partial \ln (E/Y)}{\partial t}\right|_{P_i = \overline{P}_i} \quad (3.16)$$

The level of aggregation to which the *AEEI* parameter refers varies with application. When first introduced, the *AEEI* parameter applied to the level of the national economy and hence to the rate of change of the energy/GDP ratio holding relative prices constant. But subsequently, the *AEEI* parameter has come to be used for non-price induced energy efficiency improvements at the sector or industry level.

Under this definition, the *AEEI* parameter is <u>not</u> equivalent to the energy augmenting multiplier (υ_E), since labour or capital augmenting technical change will also affect energy productivity (θ_E). It is also not equivalent to the Binswanger definition of energy saving technical change, as explained below. However, the terminology in this area is inconsistent and some authors have identified the energy augmenting multiplier as the *AEEI* (Saunders, 1992).

Also, under this definition an improvement in energy productivity over time (i.e. a reduction in energy intensity) leads to a *positive* value for the *AEEI*. However, the convention used for the *sign* of the *AEEI* parameter does not appear to be standardised either.

The *AEEI* parameter is intended to incorporate all 'non-price-based' changes in energy productivity: i.e. those that derive from technical change as defined above, as opposed to the price-induced substitution of other factor inputs for energy. A positive *AEEI* is frequently interpreted as technical progress that improves energy productivity (reduces energy intensity) as a consequence of the diffusion of more energy-efficient technologies. However, increases in *AEEI* may also reflect: shifts in the composition of output towards less energy intensive products and services, both within and between sectors; the removal of non-price barriers to the diffusion of energy efficient technologies (the 'overcoming inefficiency' arrow in Figure 3.1); and behavioural changes that reduce energy service demand (e.g. lower thermostat settings). Since all of the above changes may be induced and encouraged by policy, the term 'autonomous' is a misleading. Also, the relative importance of structural changes, as compared to the diffusion of more energy efficient technologies, may be expected to increase with the level of aggregation.

Manne and Richels (1990) showed that the value assumed for the *AEEI* within top-down energy models has a dramatic impact on the estimated cost of reducing CO_2 emissions over the long term - thereby triggering a controversy that continues to this day (Kaufmann, 2004; Sue Wing and Eckaus, 2006b). But while the *AEEI* is a commonly used parameter in energy modelling, it is difficult to estimate empirically – especially at the aggregate level (Sue Wing and Eckaus, 2006a). Instead, many empirical studies estimate the energy price bias (ψ_E) at level of individual sectors. Holding both output and input prices fixed, the following relationship can be derived between the *AEEI* and the energy price bias (Hogan and Jorgenson, 1991; Sanstad*, et al.*, 2006):

$$\psi_E = s_E(TFP_g - AEEI) \qquad (3.17)$$

In other words, the energy price bias is the 'share weighted' deviation of the autonomous energy efficiency trend from the trend in total factor productivity.[17] If energy productivity is improving at the same rate as total factor productivity ($TFP_g = AEEI$), then the energy price bias is zero and technical change is 'neutral' (under Binswanger's definition). If energy productivity is improving faster (slower) than total factor productivity, then the energy price bias is negative (positive) and technical change is energy-saving (energy-using).

Normally, we would expect the *AEEI* and the energy price bias to be opposite in sign. For example, if energy productivity is improving ($AEEI > 0$), we would expect the value share of energy to be falling ($\psi < 0$), or energy-saving technical change. But equation 3.18 suggests that this may not necessarily be the case. For example, energy-saving technical change ($\psi < 0$) may result from falling total factor productivity ($TFP_g<0$), even if energy productivity is improving ($AEEI>0$) (provided $|TFP_g| > \| AEEI|$). Empirical estimates of the energy price bias are therefore not necessarily a good guide to the magnitude or sign of the *AEEI* parameter (Sanstad*, et al.*, 2006).

The above definitions of *AEEI* and the energy price bias offer little insight into the source and nature of technical change. For example, they fail to distinguish between *embodied* technical change that requires new vintages of tools, machinery and other forms of capital equipment, and *disembodied* technical change, that derives from improvements in the knowledge and skill required to use, maintain and adapt that equipment (Berndt, 1990). They also assume that the rate and direction of technical change is fixed and therefore independent of price changes and policy interventions (Easterly and Levine, 2001). But in practice, technical change is driven by regulation, investment in R&D and a range of other factors and is clearly influenced by changes in relative prices (Löschel, 2002). This recognition has underpinned a growing volume of research into induced or endogenous technical change, including numerous attempts to incorporate endogeneity into energy-economic models (Grubb*, et al.*, 2002; Kohler*, et al.*, 2006). Hence, while the neoclassical approach to defining and measuring energy productivity provides some valuable insights, it is also subject to some important limitations.

[17] Remember that positive values for TFP_g imply improvements in total factor productivity (declining costs per unit of output), while positive values for *AEEI* imply improvements in energy productivity (declining energy intensity) over time.

3.4 Summary and implications

3.4.1 Summary

This section has summarised the different definitions of energy efficiency, energy intensity and energy productivity, identified their relationship to standard concepts in engineering and economics and clarified the relationship between them. It has shown that energy efficiency may be defined using thermodynamic, physical or economic measures that may be applied to widely different system boundaries (e.g. an individual motor, industrial process, firm, sector, or national economy). These choices are frequently linked: for example, thermodynamic measures are much more applicable at the level of individual conversion devices, while economic measures are more applicable at the sector or economy-wide level. However, competing choices are available for many system boundaries and may affect the conclusions that are drawn. For example, a first-law thermodynamic measure may suggest that there is little scope for improving the energy efficiency of electrical resistance space heating, while a second law measure may suggest that there is a considerable scope. Similarly, a measure of energy efficiency for travel by private car may exhibit very different trends depending upon whether the relevant measure of output is passenger kilometres, vehicle kilometres or tonne kilometres (see *Technical Report 2*).

Measures of energy efficiency will also depend upon how different types of energy input are aggregated. The most common approach is to aggregate different energy types according to their thermal content, but this neglects the 'quality' of different energy types, such as their ability to perform useful work (exergy) or their relative economic productivity (Divisia index). When the changing quality of energy inputs are properly accounted for, aggregate measures of energy efficiency are found to be improving much more slowly than is commonly supposed. The continuing neglect of energy quality in the energy policy literature is surprising and may potentially lead to erroneous conclusions being drawn.[18]

The review of economic approaches to measuring productivity demonstrates that energy productivity should not be pursued in isolation, but instead as part of a general effort to improve total factor productivity. Concepts from neoclassical production theory have been shown to provide a useful basis for exploring the source and direction of changes in energy efficiency (e.g. distinguishing between price-induced factor substitution and neutral/biased technical change) as well as providing methods for estimating the relevant parameters empirically. However the conventional neoclassical approach assumes that firms are operating at the efficiency frontier and that technical change is exogenous – both of which are implausible. The implications of more realistic assumptions are now the subject of active research.

3.4.2 Implications for the rebound effect

The implication of this review is that defining and measuring both the independent variable for the rebound effect (an improvement in energy efficiency) and the dependent variable (a change in energy consumption) is far from straightforward. By implication, the conclusions

[18] Giampietro (2006) comments: "…..when we decide to sum apples and oranges the chosen protocol will define the final number and its usefulness. That is, if we decide to calculate their aggregate weight, we will get a number that is not relevant for nutritionists, but for the truck driver transporting them. On the other hand, if we sum them by using their aggregate nutritional content, we will get a number that is not relevant for either a truck driver or an economist studying the economic viability of their production. The more we aggregate items that have to be described using different attributes.....using a single category of equivalence, the more we increase the chance that the final number generated by this aggregation will be irrelevant for policy discussions." (Giampietro, 2006).

drawn about the magnitude of the rebound effect will depend upon the particular choices that are made.

For the independent variable, attention must be paid (amongst other things) to:

- the *definition* of energy efficiency (first law thermodynamic, second law thermodynamic, physical or economic measures);
- the *system boundaries* to which it applies (individual device, process, firm, sector, national economy, regional economy, global economy);
- the appropriate methods for *aggregating* different energy types (i.e. whether and how differences in energy quality are accounted for); and
- the extent to which the energy efficiency improvements are considered *independently* of associated improvements in the productivity of other factor inputs.

Similar considerations apply to the dependent variable (changes in energy consumption), where attention must be paid in particular to the *system boundaries,* the method of *aggregating* different energy types and the *timeframe* to which it applies (e.g. short, medium or long term, however defined).

Many commentators assume that the relevant independent variable for the rebound effect is improvements in the *thermodynamic* efficiency of individual conversion devices or industrial processes. But such improvements will only translate into comparable improvements in *different* measures of energy efficiency, or measures of energy efficiency applicable to *wider* system boundaries, if several of the mechanisms responsible for the rebound effect fail to come into play. For example, improvements in the number of litres used per vehicle kilometre will only translate into improvements in the number of litres used per passenger kilometre if there are no associated changes in average vehicle load factors.

Rebound effects may be expected to increase over time and with the widening of the system boundary for the dependent variable (energy consumption). For example, the energy savings for manufacturing as a whole may be expected to be less than the energy savings for an individual firm that invests in an energy efficient technology. For the K-B postulate the relevant system boundary is normally taken as the national economy. But energy efficiency improvements may also affect trade patterns and international energy prices, thereby changing energy consumption in other countries. For the purpose of assessing the contribution of energy efficiency to reducing carbon emissions, the relevant system boundary is the whole world. But assessing the contribution of energy efficiency improvements to trade patterns is methodologically challenging.

To capture the full range of rebound effects, the system boundary for the independent variable (energy efficiency) should be relatively narrow, while the system boundary for the dependent variable (energy consumption) should be as wide as possible. For example, the independent variable could be the energy efficiency of an electric motor, while the dependent variable could be economy-wide energy consumption. However, measuring or estimating the economy-wide effects of such micro-level changes effects is, at best, challenging. For this reason, the independent variable for many theoretical and empirical studies of rebound effects is a physical or economic measure of energy efficiency that is applicable to relatively wide system boundaries – such as the energy efficiency of an industrial sector. But such studies may overlook the 'lower-level' rebound effects resulting from improvements in physical or thermodynamic measures of energy efficiency appropriate to narrower system boundaries. For example, improvements in the energy efficiency of electric motors in the engineering sector may lead to rebound effects within that sector, with the result that the energy intensity of that sector is reduced by less than it would be in the absence of such effects. But if the energy intensity of the sector is taken as the independent variable, these lower-level rebound effects will be overlooked. Also, improvements in more aggregate measures of energy efficiency are unlikely to be caused solely (or even mainly) by the diffusion of more thermodynamically efficient conversion devices. On the contrary, as the level of aggregation increases, the link between changes in physical or economic measures of energy intensity and improvements in thermodynamic efficiency at the micro-level becomes increasingly tenuous.

Many commentators also implicitly assume that changes in the quality of energy inputs can be neglected when exploring the implications of changes in thermodynamic, physical or economic measures of energy efficiency - at any level. But any substitution from low to high quality fuels increases the amount of useful work obtainable from the same heat content of input, as well as increasing the value of energy inputs. Measures of energy efficiency may therefore improve, without any improvement in the thermodynamic efficiency with which individual fuels are used. The rebound effects associated with fuel substitution may also be different from those associated with improvements in thermodynamic efficiency. Separation of these effects would greatly aid understanding of the processes involved, but the persistent neglect of energy quality means that this is rarely done.

Finally, several commentators implicitly assume that a change in energy consumption following an energy efficiency improvement can be solely attributed to that improvement. But improvements in energy productivity may often be associated with broader improvements in the productivity of other inputs, since new technologies frequently provide both. If the *full* impact on energy consumption of these new energy-efficient technologies is taken as the appropriate dependent variable, then backfire becomes more likely. Conversely, if only a portion of this impact is attributed specifically to the energy efficiency improvement, then backfire becomes less likely. However, it may be both difficult and misleading to isolate the impact on energy demand of the energy efficiency improvements. What matters for climate policy is how a new, energy-efficient *technology* affects overall energy demand. If energy demand is increased, redefining the independent variable to demonstrate that backfire has not occurred simply misses the point.

In summary, the literature on the rebound effect exhibits considerable ambiguity with regard to appropriate definition of the independent and dependent variables. This review of the definitions of energy efficiency has shown how the choices made can greatly influence the conclusions that are drawn. It has also provided the conceptual basis for approaching the wide range of literature to be reviewed in the remainder of the report.

4 Energy productivity and economic growth

4.1 Introduction

Len Brookes has been the most persistent and coherent advocate of the K-B postulate and a long-term critic of government energy efficiency policy (Brookes, 1978; 1984; 1990b; 2000; 2004). While arguing that "the claims of what might be called the 'Jevons school' are susceptible only to suggestive empirical support…", Brookes has marshalled a range of arguments and evidence to bolster his case, including some interesting empirical work (Brookes, 1972). Brookes arguments for backfire are mixed up with criticisms of energy efficiency policy that, while reasonable (e.g. advocating a focus on economic efficiency rather than energy efficiency alone), are not always directly relevant to the rebound effect (Brookes, 2004). Also, Brookes retired some years ago so his writings do not take into account more recent research, such as the growing body of work on Environmental Kuznets Curves (EKC) for energy use and carbon emissions (Dinda, 2004).

This section evaluates Brookes' arguments, both on their own terms and in the light of more recent research. These arguments are shown to hinge upon several of the issues introduced in the previous section, including the relationship between energy and total factor productivity and the importance of energy quality. Following a general introduction, the discussion is organised around three broad arguments made by Brookes, which are termed here the: the 'productivity' argument (Sections 4.3 and 4.4); the 'accommodation' argument (Section 4.5); and the 'endogeneity' argument (Section 4.6). In each case, we assess both the quality of the evidence itself and its relevance to the rebound effect. The discussion of the productivity argument is subdivided into an evaluation of the historical studies of Schurr and colleagues (Section 4.3), and an evaluation of the econometric evidence for 'energy-using' technical change (Section 4.4).

The general conclusion is that: first, the evidence base for each of these arguments is flawed; and second, they are only indirectly relevant to the rebound effect. As a result, they cannot be said to provide a convincing case in favour of the K-B 'hypothesis'. At the same time, each of these arguments highlights some important issues that are frequently neglected by other authors. Taken together, they provide a case for energy playing a more important role in economic growth than is commonly assumed.

4.2 The key arguments

A key argument of Brookes (2000) runs as follows:

> "…it has been claimed since the time of Jevons (1865) that the market for a more productive fuel is greater than for less productive fuel, or alternatively that for a resource to find itself in a world of more efficient use is for it to enjoy a reduction in its implicit price with the obvious implications for demand."

However, the use of the term 'implicit price' here is confusing. Individual energy efficiency improvements do not change the price of input energy, or energy commodities (P_E), but instead lower the effective price of output energy, or useful work (P_S). For example, gasoline prices (P_E) are unchanged following an improvement in vehicle fuel efficiency, but the price per vehicle kilometre (P_S) is reduced. The 'obvious implications' therefore relate to the demand for useful work (*S*), and not to the demand for energy commodities themselves (*E*).

While the former may be expected to increase, energy demand may either increase or decrease depending upon the price elasticity of demand for useful work (see *Technical Report 2*).

Of course, the combined impact of multiple energy efficiency improvements could potentially lower energy demand sufficiently to reduce energy prices and thereby stimulate a corresponding increase in economy-wide energy demand.[19] This forms one component of the economy-wide rebound effect, described in section 2.1. But while it is obvious that the overall reduction in energy consumption will be less than microeconomic analysis suggests (even when direct rebound effects are allowed for), theoretical arguments alone appear to be an insufficient basis for claiming that backfire is inevitable (Allan*, et al.*, 2006). This is the point at issue and is not resolved by calling the implications 'obvious'.

Brookes also highlights a 'lump-of-energy-dependent-activity-fallacy', where it is assumed that the level of 'energy-dependent activity' will remain substantially fixed while the 'implicit price of energy' falls under the influence of raised energy efficiency. At the level of individual energy services, this amounts to assuming that the demand for useful work (S) will remain unchanged following a reduction in the price of useful work (P_S) - or in other words, that the direct rebound effect is zero. It also amounts to assuming that the indirect and economy-wide effects are zero - for example, that the cost reductions in the production of good do not translate into increased demand for that good or increased demand for products that use that good. Brookes also criticises a 'fallacy of composition' - assuming that individual energy savings can be added together to produce an estimate of what can be saved over the economy as a whole. If the 'energy savings' here are taken to be those net of direct rebound effects, this fallacy again amounts to assuming that the indirect and economy-wide rebound effects are zero.

In both cases, Brookes is highlighting the persistent and pervasive neglect of rebound effects in the conventional assessment of energy efficiency opportunities. While Brookes first made these points in the 1970s, the majority of policy evaluations continue to embody both fallacies. Hence, Brookes is performing a valuable service by questioning conventional assumptions regarding the energy savings achievable from energy efficiency policies. However, arguing that the rebound effect is greater than zero is quite different from arguing that it is greater than one – as the K-B postulate suggests. Hence, convincing empirical support for the K-B postulate is still required.

Brookes marshals a number of other arguments in support of the K-B postulate that appear more amenable to empirical test. In doing so, he highlights some crucial issues regarding the relationship between energy consumption and economic growth that appear to be both controversial and insufficiently researched. The three most important arguments may be characterised as follows:

- *The productivity argument*: The increased use of higher quality forms of energy (especially electricity) has encouraged technical change, substantially improved total factor productivity and driven economic growth. Despite the substitution of energy for other inputs, this technical change has stimulated a rapid growth in economic output which has both reduced aggregate energy intensity while at the same time

[19] For example, Kydes (1997) found that accelerated improvements in the energy intensity of the US economy (24% over 20 years compared to a base case of 17.5%) lowered world oil prices by 15.7% compared to the base case.

increasing aggregate energy consumption. This pattern may be expected to continue in the future.

- *The endogeneity argument*: A common approach to quantifying the 'energy savings' from energy efficiency improvements is to hold energy intensity fixed at some historic value and estimate what consumption 'would have been' in the absence of those improvements. The energy savings are then taken to be the difference between the actual demand and the counterfactual scenario. But if the energy efficiency improvements are a necessary condition for the growth in economic output, the construction of a counterfactual in this way is misconceived.
- *The accommodation argument:* Energy efficiency improvements 'accommodate' an energy price shock so that the energy supply/demand balance is struck at a higher level than if energy efficiency had remained unchanged. While not immediately obvious, this argument rests in part on the assertion that the income elasticity of 'quality adjusted' energy demand falls steadily as an economy develops, but is always greater than unity.

The productivity argument rests upon two separate, but related sources of empirical evidence: the work of Sam Schurr and colleagues on the importance of *energy quality* in productivity growth (Schurr*, et al.*, 1960); and the work of Jorgenson and others on *biased technical change* (Jorgenson, 1984; Hogan and Jorgenson, 1991). These will be considered separately and in detail below. The endogeneity argument is not developed in detail by Brookes, but appears relevant to the use of decomposition analysis to explore the rebound effect (Schipper and Grubb, 2000). The accommodation argument appears to be based upon some original empirical work conducted by Brookes in the early 1970s (Brookes, 1972), together with a theoretical model that was inspired by the experience of the oil price shocks (Brookes, 1984). While each of these arguments provides some valuable insights, they also have a number of important weaknesses. Moreover, they have not been updated to take account of more recent empirical research.

The following four sections examine the energy quality, biased technical change, endogeneity and accommodation arguments in detail. Each section describes the historical research that forms the basis for the argument, summarises how Brookes uses this research to support the case for backfire, identifies potential empirical and/or theoretical weaknesses and examines whether more recent research confirms or contradicts Brookes' claims.

4.3 Productivity and energy quality

4.3.1 Schurr and the importance of energy quality

Brookes claims that improved energy efficiency will increase overall energy demand. Empirical support (but not proof) for this claim could be provided by demonstrating a positive correlation between a particular measure of energy efficiency and a corresponding measure of energy consumption. The most aggregate measure of energy efficiency is the energy/GDP ratio for a national economy, but - as discussed in section 3 - this is also the measure that is most weakly related to the thermodynamic efficiency of individual energy conversion devices.

Schurr was one of the first economists to explore historical trends in the energy/GDP ratio for the US economy and to compare these with historical trends in energy consumption and total factor productivity (Schurr*, et al.*, 1960). The results of Schurr's work have been repeatedly cited by Brookes as being '… consistent with the contentions of the Jevons School…without necessarily proving that it is due to the Jevons effect' (Brookes, 2000). Moreover, Schurr and colleagues have provided plausible explanations for their results that contain some important insights into the relationship between energy consumption and economic growth (Schurr, 1982; 1983; 1984; 1985).

Schurr collected data on energy consumption, total factor productivity and energy productivity for the US economy over the period 1880 to 1981. Energy productivity was defined as the ratio of GDP to total primary energy consumption, with energy intensity as the inverse of this measure. Importantly, Schurr did not 'quality adjust' any of the measures of factor inputs, so energy was measured on a thermal input basis. Table 4.1 summarises the main results. Over the period 1920 to 1953, Schurr's measures of energy, labour and total factor productivity were all found to be growing, while during the period 1953 to 1969, energy productivity was relatively unchanged while total factor productivity continued to grow rapidly. Both periods exhibited falling energy prices relative to other inputs and large increases in energy consumption, and were characterised by a decreasing share of coal in final energy consumption and an increasing share of oil and electricity. Also, in both of these periods, Schurr's estimates of total factor productivity grew significantly faster than those of energy productivity.

The oil price shocks of 1973 and 1979 led to a break in this historical trend. While US energy intensity (on a thermal input basis) fell by 2.4% per year over the period 1973-1981, the growth in total factor productivity slowed to only 0.4% per year and total energy consumption fell slightly. While the coincidence of higher energy prices and slower productivity growth suggests a causal relationship between the two, the relationship is more complex than it first appears (Norsworthy*, et al.*, 1979; Berndt and Wood, 1986; Schurr and Sonenblum, 1986; Olson, 1988; Berndt, 1990). Also, while the post-1973 decline in energy intensity represents a break in the post-war trend, comparable rates of decline in this indicator occurred at various intervals during the period 1920 to 1953 - when energy prices were falling in relative terms.

Table 4.1 Historical trends in energy intensity and total factor productivity in the US economy (% annual growth rates)

Period	**Energy intensity**	**Total Factor Productivity**	**Relative energy costs**	**Total energy consumption**
1920-1953	-1.3%	+2.3%	-1.0	+1.90
1953-1969	0.0%	+2.5%	-0.6	+3.60
1973-1981	-2.4	+0.4	+11.21	-0.03

Source: Schurr (1982; 1985)

Notes: Primary energy inputs measured by thermal content. Total factor productivity measured as output relative to labour and capital inputs combined. Relative energy costs measured as wholesale price index for fuel and power relative to the wholesale price index for all commodities.

Schurr sought explanations for the pre-1973 pattern of rising economic productivity and declining energy intensity. He examined whether structural change in the economy and improvements in the thermodynamic efficiency of energy conversion devices (particularly within electricity generation) had contributed to the reduction in energy intensity during this period. Importantly, he concluded that while both these factors were important, they provided only a partial explanation of the observed trends. His focus, in other words, was on the contribution of factors *other* than structural change and thermodynamic efficiency to the observed trends in energy use and productivity. Indeed, he sought to explain how the energy productivity of the US economy could improve *independently* of improvements in thermodynamic efficiency. As the rebound effect is normally understood in relation to improvements in thermodynamic efficiency, rather than more aggregate measures of energy intensity, Brookes' use of this evidence to support the case for economy-wide backfire appears, at first sight, to be rather odd.

Since energy prices were falling in relative terms during the pre-1973 period, economic theory predicts that - ceteris paribus - energy should have substituted for other factors of production, thereby *increasing* energy intensity (reducing energy productivity) and improving capital and labour productivity. Schurr *et al*'s data indicate that substitution did take place, with energy inputs doubling relative to labour inputs and increasing by 50% relative to capital inputs. But the substitution effects (movements along an isoquant) were outweighed by technological improvements (shifts of the isoquant) which greatly improved the overall productive efficiency of the US economy. 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 alongside *rising* total energy consumption – which is consistent with the K-B postulate, while not necessary demonstrating that is correct.

Schurr argued that the technological improvements which drove output growth depended crucially upon the increased availability of more 'flexible' forms of energy (oil and electricity) at relatively low costs. These contributed to changes in industrial processes, consumer products and methods of industrial organisation that were quite revolutionary. Schurr placed particular emphasis on the role of electric motors in improving productive efficiency (Schurr, 1982). Industrial drives had previously been based upon complex and unwieldy systems of shafts and belts linked to a single prime mover. By replacing these with light, flexible and highly controllable electric motors mounted on individual machines, it was possible to considerably improve the sequence, layout and efficiency of industrial production (Schurr, 1982). The greater flexibility of oil products also had an transformational role: by allowing the development of the internal combustion engine, they facilitated the mechanisation of agriculture, the relocation of labour to other sectors, the increased movement of labour across the US, and the development of spatially dispersed systems of production and distribution.

Schurr's argument was that these changes were only possible because oil and electricity were *qualitatively* different from the forms of energy they replaced. Their advantages in terms of flexibility, controllability, ease of transport and other factors led to increasing returns to energy inputs. Schurr (1982) notes that:

> "…. I not saying that innovative energy using technologies were the sole cause of rapid improvements in the overall productive efficiency, however, I do believe they were a major cause. Furthermore, I am not saying the energy supply developments with the sole cause of the emergence of innovative energy using technologies. What I am saying is that energy supply developments were *essential* features of this process."

Schurr's pioneering contribution, therefore, was to highlight the importance of energy quality for productivity growth. He notes that:

"…. energy efficiency can be improved (i.e. energy conservation can be achieved) either by *reducing* the amount of energy consumed in particular processes of production through the substitution of other input factors for energy, as during the post-1973 period, or by *expanding* the quantity and value of goods and services produced through the leverage exercised on the overall efficiency of production, as during the earlier periods.

Greater use of higher quality forms of energy leads to the second type of improvement in 'energy efficiency', rather than the first. While the conversion of primary fuel to electricity may be thermodynamically inefficient on a first-law basis, the economic usefulness of electricity improves both total factor productivity and energy productivity. Schurr is therefore using an economic definition of energy efficiency in the above paragraph, although his use of the term 'energy conservation' is arguably inappropriate since energy consumption continues to rise.[20]

4.3.2 Schurr and the rebound effect

How do these results support the K-B postulate? Brookes makes it clear that his relevant independent variable for the rebound effect is an improvement in *thermodynamic* efficiency - although, importantly, this is qualified by the requirement that the improvement be cost effective:

"…. when the author refers to measures to raise energy efficiency or energy productivity he is referring to deliberate actions to raise the cost effectiveness of the use of fuel and electrical energy by such means as: raising the engineering efficiency of conversion of fuels to useful heat or work; or increasing the effectiveness of the associated energy service by, for example, higher standards of insulation" (Brookes, 2000).[21]

But as indicated above, Schurr's primary concern is to explain how and why factors *other* than improvements in thermodynamic efficiency reduce aggregate measures of energy intensity. So why does Brookes play so much weight on this evidence to support his case for backfire? The key appears to lie in Brookes' observation that improvements in energy productivity rarely occur in isolation from improvements in the productivity of other factors.[22] For example:

"….Elliott *et al* (1997) have drawn attention to what they call the 'Cashmir effect',[23] under which a technical fix to improve energy productivity produced a bonus in materials productivity. Sutherland (1998) has argued……that energy efficiency increases, not as a direct effort to reduce energy use, but as a result of overall productivity improvements in all inputs…...Rosenberg (1983) shows how two advances in steelmaking - one aimed at achieving economies of scale and the other raising ore utilisation - produced energy efficiency bonuses" (Brookes, 2000)

---

[20] The term 'energy conservation' lacks precise definition and is less commonly used now than in the 1970's. To 'conserve' energy implies comparison against a baseline. This could either be historical (using less energy now than in the past) or counterfactual (using less energy now than what 'would have been used' had some action not been taken).

[21] Brookes (2000) also notes that: "…In this paper, a change in the productivity of fuel should be taken to mean a change that can be traced to an improvement in the engineering efficiency with which fuel is converted into useful heat or work".

[22] Brookes (2000) calls this the 'principle of indivisibility of economic productivity'.

[23] Properly 'cashmere effect'. Elliot *et al* do not say that improved energy productivity was the main aim of the change in drying technique. Instead, they argue that energy productivity is rarely pursued for its own sake, but arises out of more general improvements technology - which is precisely Brookes' point.

A very similar point is made by Saunders (2000b), who cites examples of technological improvements that both increase energy productivity and improve the productivity of other factors. For example, electric arc furnaces for steelmaking are not just energy efficient: they also allow scrap steel to be recycled, thereby bypassing the energy intensive processing of iron ore in the blast furnace (Saunders, 2000b). But since the blast furnace is also the most *capital* intensive component of steelmaking, this innovation also increases the productivity of capital in the steel sector, contributing to cost reductions and output growth. If the increase in output is sufficiently large, overall energy consumption may increase – despite a reduction in energy consumption per tonne of steel.

Numerous examples of how energy efficiency improvements can increase overall productivity can be found within the energy efficiency literature (Box 4.1). Such examples are commonly used to support the case for improved energy efficiency, with the objective of delivering economic benefits and reduced energy consumption. The potential for economy-wide rebound effects is almost invariably ignored. But in Brookes' (and Saunders') view, it is precisely this type of opportunity that is most likely to lead to backfire (Saunders, 2000b).

Box 4.1 Examples of the link between improved energy efficiency and improved total factor productivity

- Lovins and Lovins (1997) used case studies to argue that better visual, acoustic and thermal comfort in well-designed, energy efficient buildings can improve labour productivity by as much as 16%. Since labour costs in commercial buildings are typically 25 times greater than energy costs, the resulting cost savings can potentially dwarf those from reduced energy consumption.
- Pye and McKane (1998) showed how the installation of energy efficient motors reduced wear and tear, extended the lifetime of system components and achieved savings in capital and labour costs that exceeded the reduction in energy costs.
- Sorrell *et al.* (2004) found a host of examples of the 'hidden benefits' of energy efficiency improvements within 48 case studies of organisational energy management. For example changes to defrosting regimes at a brewery led to energy savings, water savings, reduced maintenance and reduced deterioration of building fabric.
- Worrell *et al.* (2003) analysed the cost savings from 52 energy efficiency projects, including motor replacements, fans/duct/pipe insulation, improved controls and heat recovery in a range of industrial sectors. The average payback period from energy savings alone was 4.2 years, but this fell to 1.9 years when the non-energy benefits were taken into account.
- Using plant-level data, Boyd and Pang (2000) estimated fuel and electricity intensity in the glass industry as a function of energy prices, cumulative output, a time trend, capacity utilisation and overall productivity. Their results show that the most productive plants are also most energy efficient and that a 1% improvement in overall productivity results in a more than 1% improvement in energy efficiency.

Brookes' argument, therefore, appears to be as follows

1. Most improvements in energy productivity will be associated with improvements in the productivity of other factors.
2. As a result, improvements in energy productivity are normally associated with *proportionally greater* improvements in total factor productivity.

3. Improvements in total factor productivity will increase economic output, leading to a corresponding increase in demand for factor inputs.
4. The resulting increase in demand for energy inputs will more than offset the reduced demand in energy per unit of economic output.

But there are a number of potential flaws in this argument. First, while it may be true that many improvements in energy productivity will be associated with improvements in the productivity of other factors, it has not been demonstrated that this will always be the case, or even that this will be the case in the majority of instances. Moreover, little guidance has been provided on which type of improvement in which type of sector is more or less likely to fall into this category.

Second, the link between improvements in energy productivity and total factor productivity remains unclear. Schurr's focus was the revolutionary technical change facilitated by the increased availability of low-cost oil and electricity. These productivity improvements may or may not have been associated with improvements in thermodynamic or physical measures of energy efficiency. While Schurr's evidence applies to the impact of energy quality improvements on total factor productivity, the issue for the rebound effect is the impact of improvements in various measures of energy efficiency on total factor productivity. Improvements in energy efficiency that are <u>not</u> associated with changes in energy quality may have a smaller impact on total factor productivity and hence may be less likely to lead to backfire. A large number of energy efficiency improvements would appear to fall within this category.

Third, Schurr's results may be less applicable to rebound effects from energy efficiency improvements by households, since the 'multiplier' effect from improvement in total factor productivity is less relevant.[24] Brookes (1990a) acknowledges this limitation, but points out that Schurr's data applies to the US economy as a whole - and therefore includes household energy consumption. He also refers to 'unpublished research' that suggests that domestic consumers spend a constant proportion of their income on energy. But more recent research suggests that the proportion of household income spent on energy has varied widely over time.

Fourth, the patterns that Schurr found may not be reproduced in all countries and in all time periods. This point deserves further research, but as an illustration Table 4.2 summarises the trends in energy and total factor productivity in the US economy over the period 1982 to 1999. This suggests that total factor productivity grew much slower than in period analysed by Schurr, and the annual fall in energy intensity *exceeded* the annual growth in total factor productivity – a situation that Brookes considered unlikely (Brookes, 1990b).[25] Table 4.3 illustrates that similar trends are observable in the UK. However, overall energy consumption still increased in both countries over this period – as Brookes would have predicted.

---

[24] Brookes' arguments in favour of backfire appear most relevant to energy efficiency improvements by producers. However, his criticisms of government policy appear to be mostly directed at energy efficiency improvements by households, where the rebound effect may be smaller.

[25] Brookes (1990b) argues that: ".....If energy productivity were to exceed productivity of the economy as a whole in conditions where energy supply is not constraint, surpluses of energy would result and its price would fall. If energy productivity were to continue to exceed multifactor productivity it would imply a totally implausible world in which consumers and producers continue to strive to give priority to energy economies over the economy in the use of other resources in the face of ever increasing surpluses of ever cheaper fuel and electrical energy."

Table 4.2 Recent trends in energy intensity and total factor productivity in the US (% annual growth rates)

Period	Energy intensity	Total Factor Productivity	Total energy consumption
1982-1991	-1.90%	0.95%	1.04%
1992-1999	-1.92%	1.36%	1.68%

Source: National Institute of Economic and Social Research; Department of Trade and Industry (UK); Department of Energy (US)

Notes: Primary energy inputs measured by thermal content. Total factor productivity measured as output relative to labour and capital inputs combined.

Table 4.3 Recent trends in energy intensity and total factor productivity in the UK (% annual growth rates)

Period	Energy intensity	Total Factor Productivity	Total energy consumption
1982-1991	-1.53%	1.45%	1.08%
1992-1999	-1.94%	1.54%	0.78%

Source: National Institute of Economic and Social Research; Department of Trade and Industry (UK); Department of Energy (US)

Notes: Primary energy inputs measured by thermal content. Total factor productivity measured as output relative to labour and capital inputs combined.

Finally, the patterns that Schurr found may not necessarily continue in the future. Schurr (1983) anticipated that they would, pointing to the importance of electricity in driving the ICT revolution. However, as high quality forms of energy provide a greater proportion of the overall energy mix, they may become used for tasks that are less and less able to make use of their 'quality' attributes – such as the application of electricity for household heating (Kaufmann, 1992). This may lead to diminishing returns, with a correspondingly smaller impact on total factor productivity. It is also possible that there could be a switch back to 'lower quality' fossil fuels, such as coal and tar sands, since the resource base for these is much greater than for conventional oil and gas (Bentley, 2002).

In summary, Brookes is correct when he states that Schurr's work offers only 'suggestive' support for the K-B postulate. First: this work applies primarily to the causal effect of shifts to higher quality fuels, rather than improvements in thermodynamic measures of energy efficiency; second, the link between energy efficiency improvements and improvements in total factor productivity may vary greatly, both over time and between different sectors and energy services; and third, the patterns Schurr uncovered may not be as 'normal' as Brookes suggests. There is clearly a need for econometric analysis to explore these issues further. Nevertheless, if Schurr is correct, his work does draw attention to the importance of the increased availability of high quality energy as a driver of productivity improvements and economic growth - a point which continues to be widely overlooked. This is explored further in Section 6.

4.4 Productivity and biased technical change

4.4.1 Introduction

Schurr's work offers some valuable insights into the relationship between aggregate energy consumption and total factor productivity, but does not quantify the contribution of different variables. To establish greater confidence in these results, it is necessary to go beyond a comparison of growth rates and to develop some econometric estimates of the relative importance of factor substitution and technical change. Jorgenson and colleagues have developed a substantial body of work in this area and their results offer some support for Schurr's 'electrification hypothesis' (Jorgenson and Fraumeni, 1981b; Jorgenson, 1984). Moreover, both Brookes and Saunders argue that Jorgenson's work provides 'suggestive' support for the K-B postulate (Brookes, 1990b; Saunders, 1992).

Jorgenson and Fraumeni (1981b) develop estimates of the direction and magnitude of technical change in 35 US industrial sectors over the period 1958 to 1974. For Brookes, the finding that technical change was *energy-using* in the majority of sectors is suggestive of backfire:

> "Jorgenson subjected Schurr's findings to econometric analysis and found, sure enough, that multi factor productivity growth was both electricity-using and total-energy using - despite parallel improvement in energy productivity at the whole economy level." (Brookes, 1990b)

Saunders (1992) provides a similar interpretation of a more recent study by Hogan and Jorgenson (1991):

> "One surprising finding is that the technical bias for energy[26] appears to be positive. That is, with a fixed energy price, Hogan and Jorgenson measure a trend of increasing value share for energy....this suggests the presence in the US economy of conditions that favour the Khazzoom-Brookes postulate.." (Saunders, 1992)

As with Schurr, the nature of the link between these findings and the K-B postulate is both indirect and unclear. Also, the studies cited by Brookes and Saunders have been superseded by more recent research that provides more ambiguous results. The following explores these issues in more detail.

4.4.2 Jorgenson and energy-using technical change

Following the energy price shocks of the 1970s and the associated productivity slowdown, a number of researchers turned their attention to the impact of higher energy prices on total factor productivity. Within the prevailing neoclassical paradigm, this involved distinguishing between price-induced factor substitution and the effect of autonomous (i.e. non price-induced) technical change (i.e. changes represented as 'overcoming inefficiency' in Figure 3.1 were ignored). Jorgenson and Fraumeni (1981b) used the econometric approach to estimating total factor productivity that was introduced in Section 3.3.2. With this approach, the effect of a change in energy prices on total factor productivity (TFP_g) (holding technical change and other prices constant) is estimated to be equal to the *energy price bias* (θ_E) – or the effect of technical change on the value share of energy (s_E), holding input prices constant:

[26] The 'technical bias for energy' is the same as the 'energy price bias' introduced in Section 3.3.3.

$$\frac{\partial TFP_g}{\partial P_E} = \frac{\partial s_E}{\partial t} \quad (4.1)$$

Jorgenson and Fraumeni chose to explore productivity growth at a relatively disaggregated level, because more aggregate econometric studies (e.g. of the energy/GDP ratio) were prone to numerous difficulties in estimation.[27] Their starting point was an assumed functional form[28] for the unit cost function for each of 35 US industrial sectors. This gave the price of output of each sector as a function of the prices of capital, labour, energy and materials, as well as the state of technology in the sector which was represented by a simple time trend. For each sector, the cost function led to a set of equations for the cost share (s_i) of each factor input (*i*), which took the following general form (Jorgenson and Fraumeni, 1981b):[29]

$$s_i = \beta_i + \sum_j \beta_{ij} \ln p_j + \beta_{it} t \quad (4.2)$$

Here p_i represents the price of each input and *t* represents time. In this framework: the parameter β_i measures the 'base' cost share of input *i* in the sector, independent of time and relative factor prices; the parameters β_{ij} measure how the cost shares change in response to changes in input prices (factor substitution); and the parameter β_{it} measures how the cost share changes over time as a result of technical change. It is the parameter β_{it} that determines the estimated bias in technical change that was introduced in Section 3.3.3. Holding input prices constant, a negative value for this parameter implies that the share of factor *i* in total costs will fall over time, while a positive value implies that the share will increase. In the case of energy, a negative value implies energy-saving technical change, while a positive value implies energy-using technical change. In the Jorgenson and Fraumeni formulation, these biases are assumed to be fixed and not influenced by changes in relative prices. This is an important limitation and is discussed further below.

The cost function also led to an equation for total factor productivity in each sector:

$$TFP_g = \beta_t + \sum_i \beta_{it} \ln p_j + \beta_{tt} t \quad (4.3)$$

Here, parameter β_t measures the 'base' *TFP* in the sector; the parameters β_{it} measure how *TFP* changes in response to changes in input prices; and the parameter β_{tt} measures how *TFP* changes over time. Unlike the bias in technical change, this formulation makes *TFP* dependent upon both time and relative prices. The parameters that determine the effect of changes in relative prices on *TFP* (β_{it}) are the <u>same</u> parameters that appear in the cost share equations as representing the bias of technical change. So, for example, the existence

---

[27] Hogan and Jorgenson (1991) argue that: "…The pervasive time series analysis of energy/GDP ratios provides few, if any of the necessary controls. Hence, most of the statistical analyses of factor biases in long-term trends in productivity, as separate from the effect of substitution among inputs, are hopelessly muddled." As an example Moroney (Moroney, 1992b) sought to estimate an aggregate production function for the US economy using the same 'translog' functional form employed by Jorgenson. This approach failed, because multicollinearity between the independent variables made all coefficients statistically insignificant (despite a R^2 of 0.99). As a result, Moroney was forced to adopt a simpler and more restrictive form for the aggregate production function (a Cobb-Douglas).

[28] Jorgenson assumed the 'translog' functional form, discussed in Annex 2. Jorgenson was one of the originators of this form of production function (Christensen*, et al.*, 1969).

[29] As described in Section 3.3.1, the cost minimising demand for any input is obtained from the partial derivative of the cost function with respect to the price of that input ('Shepard's Lemma'). Additional restrictions are normally imposed on the values taken by various parameters, in order to ensure that the cost function satisfies various conditions, including linear homogeneity (when input prices double, total costs double) and monotonicity (costs are increasing function of input prices).

of energy-saving technical change ($\beta_{Et} < 0$) implies that *TFP* will increase as energy price increases, while the existence of energy-using technical change implies that *TFP* will decrease as energy prices increase. This effect of prices on *TFP* is in addition to any price-induced substitution of capital or labour for energy.

Jorgenson and Fraumeni fitted this model to time series data for 35 US industrial sectors over period 1953 to 1973. A critical result, that forms the basis of Saunders and Brookes' arguments, is that technical progress was found to be *energy-using* ($\beta_{Et} > 0$) in 29 of the 35 sectors. The implication of this result is that, in the absence of changes in relative prices, energy forms an increasing proportion of total costs in these sectors. While technical change may reduce the amount of energy required to produce a unit of output, the percentage reduction will be less than for an aggregate measure of all inputs. Depending upon the relative magnitude of *TFP*, the energy cost share (s_E) and the energy price bias (β_{Et}), it is also possible that the energy intensity ($\tau_E = E/Y$) of these sectors *increased* over time (i.e. *AEEI<0*), independently of any change in relative prices. This is the opposite of what is conventionally assumed within energy-economic modelling.[30] Very similar results were reported by Jorgenson (1984), who adjusted the model to account separately for electric and non-electric energy inputs. In this case, Jorgenson found technical change to be *electricity-using* in 23 out of the 35 industries, and non-electric energy-using in 28 industries.

Jorgenson (1984) claims that the influence of the (predominantly positive) energy price bias (β_{Et}) on *TFP* provides a partial explanation for Schurr's empirical findings. For example, over the period 1922 to 1953, falling real energy prices encouraged the substitution of energy for other inputs but also had a sufficiently positive influence on *TFP* that overall energy intensity fell. In contrast, the large increase in energy prices after 1973 both encouraged the substitution of other factors for energy and reduced *TFP*. These combined to produce a large reduction in overall energy intensity. However, Schurr's emphasis on the role of electricity in productivity growth is only partially supported, since Jorgenson and Fraumeni found that utilisation of non-electrical energy increased *TFP* in a wider range of industries than did electrification (Jorgenson, 1984). The large shift towards oil between 1953 and 1973 may partly explain this finding, but since Jorgenson did not disaggregate 'non-electric' energy inputs further, or employ a quality weighted index of energy inputs, the evidence is inconclusive. However, it is notable that the sectors that show significant non-electric energy-using technical change include several, such as agriculture and transport, where the increased availability of liquid fuels may be expected to have had a large impact. Furthermore, in a useful commentary on Jorgenson's results, Waverman (1984) refers to unpublished research that suggests that technical change was electricity and gas using, but coal saving in 19 out of 20 US manufacturing sectors. In other words, technical change in US manufacturing appears to be biased towards the increasing use of higher quality forms of energy.

Energy-using technical change was also a key finding of a subsequent and widely-cited study by Hogan and Jorgenson (1991). This covered the period 1953-79 and compared econometric estimates of economy-wide parameters with the results of simulations from the

---

[30] The *AEEI* parameter is commonly assumed to be positive, uniform across sectors and in the range 0.4 to 1.5%/year. Long-range projections of energy demand, carbon emissions and abatement costs are very sensitive to small differences in the magnitude of this parameter (Löschel, 2002).

ETA-MACRO model (Manne and Richels, 1990). The point at issue here was the appropriate assumption for the *AEEI* parameter within energy-economic models and the impact of rising energy prices on future productivity growth. Hogan and Jorgenson argued that conventional assumptions may substantially underestimate the long term cost of reducing CO_2 emissions.

Hogan and Jorgenson (1991) weighted their parameter estimates for each sector by the sector's contribution to GDP to obtain economy-wide estimates for *TFP*, *AEEI*, the energy price bias and the energy/GDP ratio. The *AEEI* parameter was calculated using the relationship introduced earlier in Section 3.3.3, namely:

$$AEEI = TFP_g - \frac{\theta_E}{s_E} \qquad (4.4)$$

With this approach, the *AEEI* parameter depends upon the estimated values of total factor productivity and the energy cost share - both of which vary with relative prices and over time. The results for two different base years are summarised in Table 4.4. Both columns fix factor prices and economic structure (i.e. the relative contribution of each sector to GDP) at 1972 levels, but the first column sets the time (*t*) variable in Equations 4.2 and 4.3 to zero, while the second represents the cumulative effect of 17 years of biased technical change.

Table 4.4 Hogan and Jorgenson's estimates of the value of key parameters for the US economy (annual % change)

	Base year 1972	**Base year 1989**
TFP_g	0.846	-0.013
AEEI - electric	-0.088	-0.970
AEEI - non-electric	+2.246	+4.845
Energy price bias – electric	0.05	0.05
Energy price bias - non-electric	0.042	0.042
Energy price bias – total	0.092	0.092
Electric energy/GDP	-0.366	0.493
Non-electric energy/GDP	-0.334	0.525
Total energy/GDP	-0.339	0.520

Source: Hogan and Jorgenson (1991)

Notes: Positive TFP_g implies improving total factor productivity. Positive *AEEI* implies falling energy intensity. Positive *energy price bias* implies increasing share of energy the value of output (energy-using technical change). Positive *energy/GDP* ratio implies increasing energy intensity.

The model estimates economy-wide technical change to be both electric energy-using and non-electric energy using. In 1972, technical change was estimated to increase the electricity intensity of the economy by approximately 0.09%/year, but to decrease non-electric energy intensity by 2.25%/year - independently of changes in relative prices.[31] Despite the bias towards increasing electricity intensity, the 0.85% improvement in total factor productivity reduced the overall energy/GDP ratio by 0.3%/year. However, since the overall rate of productivity growth is estimated to be falling over time, a projection to 1989 (at constant prices) leads to quite different conclusions. Overall productivity growth is now found to be declining, the electricity intensity of the economy is increasing by 0.1%/year; and the energy/GDP ratio is increasing. This is, of course, different from what actually

[31] This latter result demonstrates how the *AEEI* and the energy price bias can have the same sign - as discussed in Section 2.3.3.

occurred, partly because changes in relative prices and structural change in the economy also had significant effects and partly because of the limitations of the model. But what it shows is the potential impact over the long-term of the estimated direction and magnitude of technical change (assuming that this remains fixed).

For Hogan and Jorgenson, the most interesting implication of these results was that higher energy prices could have a much greater impact upon long-term productivity growth than conventionally assumed. But for Brookes and Saunders, the most interesting implication is that, with relative prices and economic structure held constant, the electricity intensity of the economy is estimated to be *increasing* (*AEEI<0*), together with the share of electricity in total costs. However, while the share of non-electric energy in total costs was also found to be increasing, the non-electric energy intensity of the economy was estimated to be falling (*AEEI>0*). While this difference may again be suggestive of the greater importance of high quality electricity in productivity growth, it creates some difficulty in using the results to support claims for backfire.

4.4.3 More recent findings on energy-using technical change

A small number of more recent studies take a similar approach to Hogan and Jorgenson. The results suggest that the findings may be specific to individual sectors and countries as well as sensitive to the specification used.

Roy *et al* (1999) fit an identical econometric model to time series data for seven energy intensive industries in India. They note that the model may be less applicable in a developing country context owing to price regulation of outputs and the relative importance of technology transfer compared to endogenous technical change. This may be one reason why, in contrast to Hogan and Jorgenson, they fail to find a statistically significant time trend for *TFP* growth. However, apart from the iron and steel sector, all sectors are found to exhibit energy-using technical change.[32] In contrast to the US, this is in a context in which energy forms a much greater share of total costs than either capital or labour.

Sanstad *et al* (2006) present comparable results for energy intensive industries in India, South Korea and the US. They estimate *TFP* and the energy price bias for each sector and calculate the implied rate of improvement in energy productivity (*AEEI*) - again assuming that the energy price bias is fixed. The results are very heterogeneous, with some sectors exhibiting declining *TFP*. In the three US sectors, technical change is found to be energy-using, while energy intensity is increasing over the period 1958-1996 - consistent with the earlier results of Hogan and Jorgenson. In South Korea over the period 1980-1997, two out of four sectors exhibit energy saving technical change and declining energy intensity, while the other two exhibit the opposite. Overall technical progress is energy saving in South Korean manufacturing, with declining energy intensity. In India over the period 1973-1994, technical change is found to be energy-using in four out of seven sectors, while energy intensity is falling in only two sectors. Overall, technical progress is energy-using in Indian manufacturing, and acts to increase energy intensity.

Welsch and Oschen (2005) fit a comparable econometric model to time series data for aggregate West German manufacturing over the period 1976-1994. Given the high level of

[32] The magnitude, however, was relatively small. For example, assuming constant energy prices and a fixed energy price bias, it would take about 100 years to double the 1993-94 energy cost share in Indian aggregate manufacturing.

aggregation, they also measure the effect on energy consumption of changes in the pattern of imports and exports. The significance of this variable in changing the energy cost share suggests that earlier models may potentially suffer from omitted variable bias - although this will be less important for more disaggregated models.[33] Their preferred model finds technical change in West German manufacturing to be energy saving. Factor substitution, biased technical change and trade effects are estimated to have contributed to the year-to-year variation in energy intensity in the proportion 66:30:4.

The diversity of results in these studies suggests that it would be inappropriate to assume that Hogan and Jorgenson's findings can be generalised to different sectors, countries and time periods - as would be expected if the K-B postulate holds. At the same time, energy-using technical change and/or negative values for the *AEEI* parameter would appear to occur much more often than is commonly assumed - thereby calling into question some standard assumptions of energy-economic modelling. The relatively small number of studies in this area, together with their apparent sensitivity to econometric specification, suggests the need for further work.

4.4.4 Limitations of the Jorgenson approach

All the above studies suffer from four key limitations. First they assume that the bias of technical change is fixed over time, and therefore not influenced by changes in relative prices and other variables. Second, they implicitly assume that factor shares can adjust instantly to changes in factor prices, and hence that the process of investment and stock rotation can be neglected. Third, they assume a deterministic time trend for technical change and do not employ more sophisticated econometric techniques to test for the presence of 'cointegration'. Finally, and most importantly, they measure energy consumption on a thermal input basis and do not adjust for changes in energy quality. A more realistic treatment of each of these issues could add considerable complexity to the models, but also promises more policy-relevant insights.

4.4.4.1 Exogenous versus endogenous technical change

The assumption of a fixed technical bias runs counter to intuition. It seems reasonable to assume that technical change will be biased against those factors whose real prices are increasing and towards those factors whose prices are falling. In the case of Jorgenson and Fraumeni's study, labour prices were increasing over the period in question (1953 to 1973) while energy prices were falling. Since these conditions changed after 1973, and may be expected to change further in the future, it may be inappropriate to use these findings as a basis for long-term projections. In a seminal paper on biased technical change, Binswanger (1974) suggested that models that assume a fixed bias are only suitable for short time periods.[34] Hogan and Jorgenson's approach therefore runs counter to this recommendation.

One relatively simple approach to accommodate the potential effect of relative prices is to allow costs to be a function of the *rate of change* of prices, as well as their level. Norsworthy (1981) applied this approach to the same dataset as Jorgenson and Fraumeni and also included the rate of change of output, to accommodate the fact that more rapid output

---

[33] Problems of multicolliniearity made it difficult to separate the effects of changes in trade patterns from those of biased technical change. These were addressed through the imposition of additional equality constraints.
[34] However, Binswanger's empirical results suggest that (in the case of agriculture) relatively large changes in prices are required to change the rate or direction of bias.

growth may be more costly. Norsworthy's found that the bias on technical change was approximately neutral for energy. In other words, Jorgenson and Fraumeni's results were not robust to slight changes in econometric specification. Very similar results are reported by Berndt and Wood (1982)

In practice, the rate and direction of technological change may be expected to be influenced by: research and development activities by the public and private sectors; by learning economies, where productivity improves over time and with increasing scale of activities; and by spillovers from one sector to another - such as when developments in the aerospace industry encourages the use of high efficiency CCGTs for electricity generation. More recent research focuses increasingly upon understanding these processes and upon making the rate and direction of technical change *endogenous* within economic models (Grubb*, et al.*, 2002; Löschel, 2002; Kohler*, et al.*, 2006; Sue Wing, 2006). The assumption of a fixed and exogenous *AEEI* and/or energy price bias is therefore both dated and misleading.

4.4.4.2 Embodied versus disembodied technical change

Economists frequently distinguish between embodied and disembodied technical change (Berndt, 1990). Embodied technical change refers to improvements in the design and performance of technologies that can only be embodied in new plant or equipment. This type of technical change depends upon new investment and the pace of change may vary widely between different sectors and energy services, depending upon equipment lifetimes and the speed of capital stock rotation. In contrast, disembodied technical change refers to advances in knowledge that make more effective use of all inputs, independently of the age of the capital stock. This includes processes such as learning-by-doing and learning-by-using (Arrow, 1962). For example, embodied technical change in manufacturing could be represented by new machinery, while disembodied technical change could be represented by improved methods of operating and maintaining existing machinery.

Each of the above models assumes implicitly that technical progress is disembodied, since capital is assumed to adjust instantaneously to changes in prices. This assumption may be justified in the case of Jorgenson and Fraumeni, since their objective was to examine the short-term impact of energy price increases on total factor productivity. However, it is much less suitable for the analysis and projections of long run demand patterns.

To accommodate embodied technical change, it is necessary to use a fully dynamic model that reflects the vintage of the capital stock, the process of stock rotation and investment and the rate of adjustment to price changes. For instance, Berndt and Hesse (1986) specify a model that allows for long run adjustment of capital, hence relaxing the disembodied technical progress assumption of the Jorgenson models. Although their specification yields similar biases for energy in the US, almost all the labour biases are different - suggesting again that different specifications can lead to very different results.

Sue Wing and Eckhaus (2004) provide a particularly good example of a model that separates the effect of embodied and disembodied technical change. Based on the work of Berndt, Morrison and Watkins (1981), they distinguish between variable inputs (labour, energy materials) and quasi-fixed capital inputs (ICT, electrical equipment, machinery, vehicles and structures) – each with a different rate of depreciation. In this model, disembodied technical change is interpreted as the short run rate of energy intensity improvement, while the long-run average *AEEI* is given by the sum of the short run rate and the effect of innovations that are embodied in new vintages of quasi-fixed capital inputs.

Sue Wing and Eckhaus apply their model to Jorgenson's dataset of 35 US industrial sectors over the period 1958-1996. To facilitate comparison with Jorgenson's (1984) results, they also estimate over the shorter time period 1958-1979. Their results indicate substantial heterogeneity across industries in estimates of the *AEEI* parameter, as well as differing effects over different time periods. For the period 1958-1979, the long run *AEEI* - reflecting the joint influence of embodied and disembodied technical change – is significantly negative in eight sectors (increasing energy intensity) and significantly positive in two (declining energy intensity) - with the rest of the results being statistically insignificant. For the period 1958-1996, only five sectors provided significant results - with a positive *AEEI* in one and a negative in four. In only one sector do the significant estimates of *AEEI* have the same sign in both time periods.

Over the full time period, disembodied technical change was found to lead to increasing energy intensity in ten sectors and decreasing energy intensity in eight. Relatively few sectors provided significant estimates of embodied technical change, but these suggest that the embodiment of energy using innovation occurs in machinery, vehicles and structures, while that of energy saving innovation occurs in ICT and electrical equipment.

Between 1958 and 1996, the aggregate energy intensity of US manufacturing fell by approximately one third (32%). Sue Wing and Eckhaus are able to decompose this change into the relative contribution of structural change (-12%), price induced substitution of variable inputs (-6%), disembodied technical change (+15%) and changes in the level and composition of quasi-fixed inputs (-32%). The last of these represents both changes in the mix of capital (with some types of capital being more energy intensive than others) and changes in the energy intensity of each type of capital over time, as a consequence of embodied technical change. The latter was estimated to have increased the energy intensity of machinery (+9%), vehicles (+9%) and buildings and structures (+30%), while at the same time reducing the energy intensity of electrical equipment (-10%) and ICT (-38%).

Hence, the picture presented by Sue Wing and Eckhaus (2004) is much more complicated and heterogeneous than that suggested by Jorgenson's earlier work. Technical change appears to be energy using for some types of capital and energy saving for others. The shifts in these trends over time, and particularly after the first energy price shock, are suggestive of price-induced technical change. In contrast, disembodied technical change appears to be energy using and largely unaffected by changes in relative prices. The engineering interpretation of such disembodied changes remains unclear. Over the full period, the effects of embodied and disembodied technical change largely offset each other, resulting in a small net increase in energy intensity.

4.4.4.3 Deterministic versus stochastic time trends

Standard regression analysis of time series data can lead to misleading results when the time series are 'non-stationary' – that is, growing over time and without a fixed (stationary) mean. This is a common occurrence with time series data when the magnitude of a variable in one period depends closely on its value in a previous period – a so-called 'stochastic' process. Even when a time trend has been removed from the data, it is possible to obtain spurious correlations between two variables that are in fact unrelated (Granger and

Newbold, 1974). However, tests are now available to eliminate this possibility and assess whether there is a statistically significant, long-run relationship between two or more variables. If so, the variables are said to be 'cointegrated' and the relationship can be explored through techniques such as error correction models.

Since technical change is a cumulative process, it may be expected to take a stochastic form (Kaufmann, 2004). Hence, models which employ a deterministic time trend to simulate technical change are likely to be flawed. For example, Clark and Youngblood (1992) show how including a deterministic time trend in a translog cost model (such as used by Jorgenson) can lead to results which indicate biased technical change, where none in fact exists. Similarly, Lim and Shumway (1997) show how failing to allow for cointegration can allow the precision of relationships to be overestimated and may also lead to incorrect signs. An alternative approach is illustrated by Hunt *et al.* (1999), who employ a structural time series model (Harvey, 1989) to explore trends in UK energy demand. Their results show that the rate and direction of technical progress has varied stochastically over time and between fuels as a result of a variety of observed and unobserved factors.

Since none of the studies cited above check for the presence of cointegration or use more sophisticated econometric techniques, their estimates for the rate and direction of technical change could potentially be either biased or spurious.

4.4.4.4 Energy quality

The final flaw in the studies reviewed is the failure to take into account changes in energy quality. While Jorgenson and Fraumeni (1981a) do distinguish between electrical and non-electrical energy, most of other studies aggregate all energy forms into a single measure, based upon their thermal content. Hence, contrary to Brookes (1990) claims, this work does little to substantiate Schurr's hypothesis regarding the importance of energy quality.

The potential flaws in this approach are highlighted by Kaufmann (1992) in a recent study of the determinants of changes in the energy/GDP ratio of several OECD countries. This study, which will be discussed in greater detail in Section 6, found that changes in this ratio were largely explained by changes in economic structure, energy prices and energy quality – leaving essentially no role for either autonomous or biased technical change (*AEEI*). More recently, Kaufmann has updated this work for the US economy to take into account co-integration between the variables (Kaufmann, 2004). The results are broadly the same. At the level of the macro-economy, once changes in energy quality are taken into account there is no deterministic time trend that can be attributed to energy saving technical change (*AEEI*). This either implies that there are correspondingly no comparable trends at the sector level, or that improvements in energy efficiency in some sectors are largely offset by declining energy efficiency in others. The results suggest the neglect of energy quality in conventional models could be very misleading, in that changes in energy intensity that derives from shifts to higher quality fuels could instead be attributed to technical change. Jorgenson's finding of energy-using technical change is therefore again called into question.

4.4.5 Biased technical change and the rebound effect

What are the implications of this review for the K-B postulate? If we accept for the moment that a finding of energy-using technical change provides support for the postulate, there are a number of difficulties with the available evidence base.

First, neither Jorgenson's work itself, nor those of comparable studies such as Sanstad (2006) consistently find energy-using technical change. Instead, the empirical results vary widely between different sectors, countries and time periods and frequently find energy-saving technical change. Second, the results from these studies appear sensitive to quite minor changes in econometric specification (Norsworthy, 1981; Berndt and Wood, 1982). Third, the assumption of a fixed bias in technical change is flawed, as demonstrated by more recent work on price induced technical change (Kohler*, et al.*, 2006). Fourth, the implicit assumption that technical change is disembodied is unrealistic, and when embodied technical change is allowed for in more sophisticated models the results suggest that the magnitude and sign of technical change varies between sectors and over time, as well as between different types of energy-using technology (Sue Wing, 2004). Fifth, since none of the studies check for the presence of cointegration, their estimates for the rate and direction of technical change could potentially be either biased or spurious (Clark and Youngblood, 1992). Finally, since none of the studies allow for changes in energy quality, it is possible that changes in energy intensity that derive from shifts to higher quality fuels could wrongly be attributed to technical change (Kaufmann, 1992).

Given these criticisms, it is difficult to claim that this evidence base provides suggestive support for the K-B postulate. Given the 'absolute' nature of that postulate, we would expect clear and consistent evidence in its favour from repeated studies. In practice, the evidence is very mixed - as would be expected if the pattern of technical change varied in magnitude and sign between different energy services, sectors and time periods and was influenced by variables such as relative prices. At the same time, however, this evidence base presents an equally difficult challenge to conventional assumptions regarding the *AEEI* parameter.[35] If technical change is not consistently energy-using, neither does it appear to be consistently energy-saving. To date, the comprehensive study, that takes each of the above criticisms into account has yet to be conducted.

Even if there were strong evidence for energy-using technical change, the relevance of this for the K-B postulate remains unclear. First, energy-using technical change (in Jorgenson's definition) implies that the value share of energy increases over time, independently of changes in relative prices. As indicated by Equation 4.4, this need not necessarily mean that energy intensity increases over time (i.e. a negative *AEEI*). Indeed, while Hogan and Jorgenson (1991) found energy-using technical change for both electric and non-electric energy, electric intensity was found to be increasing while non-electric intensity was found to be decreasing.

Even if technical change were consistently found to increase energy intensity, there is still difficulty in linking this finding to the K-B postulate. The central claim of the K-B postulate is that an improvement in a particular measure of energy efficiency that is relevant to one system boundary leads to an increase in energy consumption within the same or a wider system boundary. But the evidence from the studies reviewed here points to something different - namely that the contribution of technical change has sometimes been to *increase* measures of energy intensity (i.e. reduce energy efficiency) and thereby increase overall energy consumption, even while other factors (such as structural change) are acting to decrease it. In other words, Saunders and Brookes have highlighted studies in which a

[35] It is worth noting that in Jorgenson's model, the technical biases are fixed and total factor productivities are allowed to vary with time. But in many neoclassical partial and general equilibrium models, technical biases are allowed to vary with time and total factor productivity is held fixed. Therefore, Jorgenson's results are obtained under different assumptions to the ones used in commonly used in energy-economic models.

measure of energy efficiency (*AEEI*) has the opposite sign to what is conventionally expected, but *also* to what appears to be required for an empirical estimate of the rebound effect.

This puzzle may potentially be resolved by returning again to the definition of the appropriate independent variable for the rebound effect. It is clearly the case that technical change has improved the thermodynamic conversion efficiency of individual devices, such as motors and boilers. What Jorgenson's work suggests is this that has not necessarily translated into improvements in more aggregate measures of energy intensity at the level of industrial sectors. Similarly, what Sue Wing and Eckhaus' results suggest is that this has not necessarily translated into improvements in more aggregate measures of energy intensity for particular types of capital (e.g. machinery). We could envisage, for example, a situation in which the motors within individual machines were becoming more energy efficient, but the machines themselves were becoming less energy efficient (i.e. more energy use per unit of output), perhaps because they were becoming heavier and more complex. As a result, technical change acts to increase overall energy consumption for this type of capital. This finding runs counter to conventional wisdom as well as to the findings of decomposition studies, which generally find declining trends in energy intensity in the majority of sectors (Schipper and Grubb, 2000). A key difference is that decomposition studies do not separate the effect of price-induced factor substitution from that of technical change. Such studies are discussed further in the following section, but a clear implication is that more research is required to reconcile these two perspectives.

In summary, the only econometric evidence that has been put forward to support the "Jevons school" is suggestive at best and the manner in which this evidence supports the K-B postulate is far from clear. Moreover, the robustness of the evidence itself is open to question.

4.5 Endogeneity and decomposition analysis

4.5.1 Introduction

A common approach in the energy policy literature is to estimate the 'energy saved' by energy efficiency improvements over a particular period by comparing current energy consumption with an estimate of what energy consumption 'would have been' had particular measures of energy intensity remained unchanged. For example, the IEA analysed data from 11 OECD countries over the period 1973 to 1998 to suggest that energy use would have been nearly 50% higher in 1998 if end-use intensity had remained at its 1973 level (Geller*, et al.*, 2006b). But this begs the question of whether the improvements in energy efficiency were themselves a necessary condition for the increase in output. Brookes argues strongly that they were:

> "….it is inconceivable that populations of today could be maintained with the technology of 500 years ago. According to Cipolla (1962), energy is all that stands between us and "grinding agrarian poverty": inanimate energy allied to man's ingenuity is what has permitted the very large increase in output in the last 200 years without which the increase in population would not have occurred. Would this increase (and the associated increase in energy consumption) have occurred if conversion efficiencies had stayed at the abysmally low levels (between less than 1% and a few per cent) prevailing in the early years of the 19th century?" (Brookes, 2000)

A similar argument was originally made by Jevons (1865), who considered that the population and affluence of the mid-19th century was inconceivable at the conversion efficiencies of Savory's steam engine. Alcott (2006) takes this argument further and points out the parallels with 19th debates on labour productivity. The central issue, again, is the contribution of energy efficiency improvements to the growth in economic output. While it may be 'inconceivable' that populations of today could be maintained with the conversion efficiencies of 500 years ago, it may be equally inconceivable that they could be maintained with the agricultural and medical technology of 500 years ago. What is at issue is the relative importance of energy efficiency improvements compared to other types of technical change and the extent to which the former may be a precondition for the latter. This critical question is very difficult to assess empirically, but does appear particularly relevant to the use of decomposition analysis within energy studies. This section discusses some of the issues involved and evaluates the approach taken by Laitner (2000) and Schipper and Grubb (2000) to the rebound effect.

4.5.2 Decomposition analysis

The starting point with decomposition analysis is to express trends in aggregate quantities as the product of a number of different variables. For example, economy-wide energy consumption (*E*) may be expressed into the product of population (*P*), GDP per capita (*A=Y/P*) and energy use per unit of GDP (*T=E/Y*) or:

$$E = PAT \qquad (4.5)$$

An additive decomposition expresses the change in energy use ($\Delta E = E_T - E_0$) over a particular period as the sum of the change in each of the right-hand side variables ($\Delta E = \Delta P + \Delta A + \Delta T$), while a multiplicative decomposition expresses the ratio of energy use at the end of the period to that at the beginning of the period ($R_E = E_T / E_0$) as the product of comparable ratios for each of the right-hand side variables ($R_E = R_P R_A R_T$). Similar expressions can be developed at varying levels of detail for energy use within individual sectors. Thanks in part to the work of Lee Schipper and colleagues (Schipper and Meyers, 1992), decomposition analysis has become a widely used tool within energy economics (Ang, 1999).

Decomposition analysis implicitly assumes that the variables on the right-hand side of Equation 4.5 are independent of one another, or at least that any dependence is sufficiently small that it can be neglected. But Jevons and Brookes argue that improved energy efficiency enables both higher affluence (*A=f(T)*) and higher population (*P=f(T))*. The causality is also likely to operate in reverse, such as when greater affluence permits more investment in research and development (*T=f(A)*) (Alcott, 2006). If the dependent variables are endogenous, the relationships may perhaps be better expressed as a system of simultaneous equations (Alcott, 2006):

$$\begin{aligned} E &= f(P, A, T; X_E) \\ P &= g(E, A, T; X_P) \\ A &= h(E, P, T; X_A) \\ T &= i(E, P, A; X_T) \end{aligned} \qquad (4.6)$$

Here, each X_i represents a vector of exogenous variables that also influence *i*. For example, we would expect the level of population to be influenced by medical knowledge and the standard of health care and sanitation facilities. These, however, are also likely to be endogenous in that they may be influenced by population, affluence and energy use. Assuming, for the moment, that an equilibrium can be defined and a clear separation between exogenous and endogenous variables can be maintained, each endogenous variable may in principle be expressed as a function of the exogenous variables; for example:

$$E = \hat{f}(X_E, X_P, X_A, X_T) \qquad (4.7)$$

The use of a simultaneous equation framework permits a clearer understanding of the implications of changes in energy efficiency. For example, regulatory interventions (X_E) to encourage improvements in the energy efficiency of new conversion devices will have a direct effect on the economy-wide energy/GDP ratio (*T*) through the fourth of the equations in 4.6. However, such improvements may also encourage economic growth (*A*), which in turn will increase the total demand for energy (*E*). Over the long term, rising affluence may encourage higher population levels (*P*), which in turn will increase energy consumption (*E*). Each of these changes may influence the energy/GDP ratio (*T*). Hence, a change in an exogenous variable may trigger a complex set of changes, and the total change in energy consumption following the regulatory intervention may be greater or less than that created by the direct change alone.

While rhetorically persuasive, a simultaneous equation framework cannot be used to trace the dynamic, economy-wide and highly complex changes relevant to the K-B postulate. More insightful approaches may instead be provided by growth theory (discussed in Section 5) and CGE modelling (discussed in *Technical Report 4*). But here we summarise two alternative and much simpler approaches to estimating the contribution of particular measures of energy efficiency improvements (*T*) to particular measures of economic output, or activity (*A*). Both of these have been widely cited in support of the argument that economy-wide rebound effects must be small.

4.5.3 'Back of the envelope' estimates

4.5.3.1 Laitner's Approach

Laitner (2000) has provided a simple but revealing approach to estimating the potential magnitude of the growth effects from energy efficiency improvements, together with the additional effect of reductions in energy prices. This is based upon an equation of the form:[36]

$$E_t = E_o * Y_t^{\eta_Y(E)} * P_{E_t}^{\eta_{P_E}(E)} * \Delta\varepsilon \qquad (4.8)$$

Where E_t = primary energy consumption in year *t; Y_t* = economic output; P_{Et} = primary energy prices; $\eta_Y(E)$ = the income elasticity of primary energy demand; $\eta_{P_E}(E)$ =the own-price elasticity of primary energy demand (measured in kWh thermal content); and $\Delta\varepsilon$ = the proportional reduction in the economy-wide energy/GDP ratio (*E/Y*) assumed to be achieved by energy efficiency policies between year *0* and year *t.*

[36] Laitner actually uses an equation for carbon emissions, but little is lost in confining attention to energy.

Based upon data from the US Annual Energy Outlook 1999, Laitner chooses indicative values of 0.82 for the income elasticity of primary energy demand and -0.30 for the own-price elasticity. With a projected (exogenous) 31% increase in GDP between 1998 and 2010, and a 12% increase in energy prices, the equation projects a 20.6% increase in energy consumption. Assuming that energy efficiency policies lead to an additional 30% reduction in the economy-wide energy/GDP ratio compared to the baseline scenario ($\Delta\varepsilon = 0.7$), the net increase in energy consumption is reduced to 15.6%.

However, these energy efficiency policies also increase the productivity of energy inputs and may therefore lead to additional increases in economic output. Laitner therefore asks to what level would GDP need to rise in order to completely offset the reduction in projected energy consumption achieved by the (assumed) 30% reduction in the energy/GDP ratio? Holding other variables constant and solving for GDP (*Y)*, he concludes that GDP would need to increase by 55% more than what would have happened otherwise. Laitner considers a rebound of this size to be implausible.

An additional contribution to the economy-wide rebound effect may result from the increase in demand stimulated by the reduction in energy prices compared to the business as usual scenario. As an illustration, Laitner shows that a combination of a 22% increase in GDP compared to the base scenario, combined with a 42% reduction in energy prices, would be sufficient to offset the reduction in energy consumption achieved by the reduced energy/GDP ratio. Again, changes of this magnitude are considered to be implausible.

4.5.3.2 Evaluation and critique

Laitner's view that such growth effects are implausible may be compared with Saunders' comment that it *may* be possible for an X% improvement in 'energy efficiency' to result in a greater than X% increase in GDP (Saunders, 2000b). The key requirement is that energy services provide a highly attractive substitute for other factors of production. In this case, "….fuel becomes so attractive in replacing capital and labour its use expands radically, thus substantially expanding the economy" (Saunders, 2000b). However, while this is a theoretical possibility, the required magnitude of the elasticity of substitution between energy and other factors of production appears to be much greater than empirical estimates suggest (see *Technical Report 3*). Using a theoretical analysis based upon an economy-wide production function, Saunders (2000) concludes that a 20% increase in the productivity of energy inputs alone should only increase GDP by some 2.3%. Wei (2007) finds a minor error in Saunders' calculations and estimates a slightly higher value of 3.6%. Both analyses assume a Cobb Douglas form for the economy-wide production function, which has a unitary elasticity of substitution. Since this is larger than many empirical estimates suggest, this estimate could potentially represent an upper bound on the impact of economy-wide energy efficiency improvements on GDP (Saunders, 2000b).

Hence, Laitner and Saunders appear to agree that the growth effects of a 20% improvement in some economy-wide measure of 'energy efficiency' are likely to be relatively small - of the order of 2-3%. But while Laitner uses this result to argue that backfire is unlikely, Saunders position is that backfire is likely to be the norm. Indeed, Saunders (2007) shows that, if the economy-wide production function is assumed to take a Cobb-Douglas form (as in the example above), energy efficiency improvements *always* leads to backfire.[37]

These conflicting conclusions result in part from differences in approach, and in part from different definitions of the independent variable. Laitner is using a highly simplified relationship between energy consumption and economic output, incorporating income and price elasticities derived from empirical studies. For Laitner, the relevant independent variable is a change in the economy-wide energy/GDP ratio, which is *assumed* to have been brought about by energy efficiency policies. As a result, the only rebound effects that are relevant are the growth and price effects. Saunders, in contrast is assuming a particular form for the economy-wide production function to represent a theoretical relationship between energy consumption and economic output. For Saunders, the relevant independent variable is an improvement in productivity of energy use within this function. This allows the rebound effect to be derived analytically, while its magnitude may be estimated by incorporating data on the share of different factors in total input costs.

Some details on Saunders approach for short run rebound effects (where capital and labour are assumed to be fixed) are provided in Box 4.2. This shows that, with a Cobb-Douglas production function, improvements in the productivity of energy use leave the economy-wide energy/GDP ratio <u>unchanged</u> in the short term. This means that energy productivity improvements <u>increase</u> energy consumption by as much as the increase economic output. The rebound effect then derives solely from the increase in economic output, and is given by $s_E/(1-s_E)$, where s_E represents the value share of energy. Saunders (2007) derives comparable results for the long-term (where capital is allowed to adjust),[38] and shows that, again, that the energy/GDP ratio remains unchanged . The rebound effect is now given by $1+s_E/s_L$, where s_L represents the share of labour in total costs. In both cases, the rebound effect exceeds unity.

The contrast between the two approaches now becomes clearer. Saunders is using an improvement in energy productivity as the independent variable and deriving a result in which the economy-wide energy/GDP ratio remains unchanged (i.e. energy consumption and economic output increase by the same amount). Laitner, in contrast, is simply assuming that improvements in energy productivity will reduce the economy-wide energy/GDP ratio by 30%. A criticism of Saunders approach could be that the real-world economy is unlikely to behave in the manner suggested by a Cobb Douglas production function! This will be discussed further in Section 5, but it is worth noting that a number of other assumptions for the functional form of the economy-wide production function produce comparable results (Saunders, 2007). In contrast, a criticism of Laitner's approach could be that the assumption that energy efficiency policies will reduce the energy/GDP ratio is flawed.

The reduction in the energy/GDP ratio is assumed to result from the cumulative effect of energy efficiency improvements at the micro level (Brown*, et al.*, 2001). But such improvements may be expected to: a) reduce the price of useful work and encourage the substitution of useful work for other factors of production; b) encourage structural adjustments, with energy intensive goods and sectors gaining at the expense of less energy

[37] Saunders results depend solely upon the assumed form of the production function and hence may be applied to an individual firm, a sector or an entire economy - provided one accepts that the behaviour of these can be represented in that way.

[38] This uses the simplifying assumption that the cost of capital is unchanged. For a discussion, see Section 5 and Saunders (2007).

intensive ones; and c) increase economic growth and aggregate consumption (Birol and Keppler, 2000).[39] The $\Delta\varepsilon$ parameter in Laitner's equation effectively subsumes the first two effects and assumes that the energy efficiency policy leads to a net 30% reduction in aggregate energy intensity. Some justification for the 30% figure may be derived from a variety of modelling exercises, including the study by Koomey *et al*. (1998) cited by Laitner. But these may potentially neglect some of the effects represented by (a) and (b) above. As discussed in Section 3.4, if measures of energy efficiency at a more disaggregate level are taken as the relevant independent variable, then the total rebound effect is given by the sum of (a), (b) and (c) for each particular improvement, which may be larger than suggested by the growth effects alone. The relevance of Laitner's result therefore depends in part on the robustness of the underlying modelling of the impact of energy efficiency policy on energy/GDP ratios. The earlier discussions of energy quality and biased technical change in Sections 4.3 and 4.4 suggests caution in assuming that energy efficiency improvements at the micro level will necessarily translate into reductions in more aggregate measures of energy intensity, such as the energy/GDP ratio. However, to explore this issue further would require an investigation of the relative appropriateness of different modelling techniques, which is beyond the scope of this report.

[39] The first impact will depend upon the elasticity of substitution between factors, as well as the share of energy in total costs. The second impact will depend upon the elasticity of substitution between products.

Box 4.2 Proof that a Cobb-Douglas production function leads to backfire in the short-term.

The economy-wide production function is assumed to take a 'Cobb Douglas' form, with constant returns to scale:

$$Y = aK^{\alpha}L^{\beta}(\tau E)^{1-\alpha-\beta}$$

Where: *K*=capital, *L*=labour, and *E*=energy. The multiplier τ ($\tau \geq 1$) increases the productivity of energy inputs, so that the product τE represents 'effective' energy inputs (see Section 3.3.3 and 5.3.1). In the short-term, *K*, *L* and the real price of energy (P_E /P_Y) are fixed and the marginal productivity of energy equals its real price (P_E/P_Y):

$$\frac{\partial Y}{\partial E} = P_E = (1-\alpha-\beta)\frac{Y}{E}$$

So:

$$\frac{E}{Y} = \frac{(1-\alpha-\beta)}{P_E}$$

Differentiating this expression with respect to the energy productivity multiplier (τ) gives the effect of improvements in energy productivity on the aggregate energy/GDP ratio with real energy price (P_E/P_Y) fixed:

$$\frac{\partial}{\partial \tau}\left(\frac{E}{Y}\right) = 0$$

Hence, improvements in the productivity of energy inputs leave the aggregate energy/GDP ratio unchanged. So the effect of τ on energy use is solely via output.

Substituting $E = Y(1-\alpha-\beta)/P_E$ into the production function:

$$Y = aK^{\alpha}L^{\beta}\tau^{1-\alpha-\beta}Y^{1-\alpha-\beta}\left(\frac{1-\alpha-\beta}{P_E}\right)^{1-\alpha-\beta}$$

Rearranging:

$$Y = a^{\frac{1}{\alpha+\beta}}K^{\frac{\alpha}{\alpha+\beta}}L^{\frac{\beta}{\alpha+\beta}}\tau^{\frac{1-\alpha-\beta}{\alpha+\beta}}\left(\frac{1-\alpha-\beta}{P_E}\right)^{\frac{1-\alpha-\beta}{\alpha+\beta}}$$

Differentiating this with respect to the energy productivity multiplier (τ) gives the effect of improvements in energy productivity on economic output. Multiplying through by τ / Y converts this into an expression for the elasticity of output with respect to changes in energy productivity:

$$\eta_{\tau}(Y) = \frac{\tau}{Y}\frac{\partial Y}{\partial \tau} = \frac{1-\alpha-\beta}{\alpha+\beta} = \frac{s_E}{1-s_E}$$

Where s_E represents the share of energy costs in the total value of output.

Source: Saunders (2007); Wei (2007)

4.5.4 Feedback between energy intensities and energy use

4.5.4.1 Schipper and Grubb's approach

Schipper and Grubb (2000) provide a wide-ranging and insightful analysis of rebound effects, based upon a comprehensive decomposition analysis of energy use patterns in twelve IEA economies over the period 1973 to 1994. The dataset includes information on 'activity' levels (*A*) and final energy consumption per unit of activity (I) for approximately 30 types of activity (*i*) in these economies.[40] For example, vehicle kilometres provide the measure of activity for car travel, while fuel use per kilometre provides the measure of energy intensity. Together, the activities account for 85-90% of final energy consumption in these countries. While data on fuel mix is available for many activities, the aggregate energy consumption for activity *i* in year *t* is measured in heat equivalent: in other words, shifts in the quality of energy inputs are ignored. Energy consumption for activity *i* in year *t* is then given by: $E_{it} = A_{it}.I_{it}$.

This dataset does not allow rebound effects to be measured directly. Instead, Schipper and Grubb *infer* the existence or otherwise of significant rebound effects from the apparent degree of correlation between different variables. These are generally identified visually from graphs rather than being quantified. In particular, they look for an increase in activity levels relative to GDP ($\partial(A_i / Y)/\partial t > 0$) during periods when energy intensity is falling ($\partial I_i / \partial t < 0$) and take this as a very approximate indicator of the elasticity of activity levels with respect to changes in energy intensity ($\eta_{I_i}(A_i)$), which is expected to be negative if there is a significant rebound effect. For example, they look for an increase in passenger vehicle kilometres per unit of GDP during periods in which fuel use per kilometre is falling.

While Laitner (2000) explored the potential impact of changes in an economy-wide measure of energy intensity (*E/Y*) on an economy-wide measure of economic activity (*Y*), Schipper and Grubb (2000) explore the potential impact of more disaggregated measures of energy intensity on more disaggregated measures of economic activity. Moreover, while Laitner used an economic measure of activity as the dependent variable (GDP), Schipper and Grubb use a mix of physical (e.g. vehicle kilometres) and economic measures (e.g. value-added). They use this data to make a host of observations on the likely magnitude of rebound effects in different sectors, and generally conclude that these are either small or not observable. However, the lack of quantification in Schipper and Grubb's approach (to rebound effects), together with the lack of control for several variables that could influence energy intensity or activity levels, limits the degree of confidence in the results.

[40] Namely, seven manufacturing sectors, ten end-uses in households, five modes each of personal and freight transport, fuel and electricity use in the service sector and a number of miscellaneous activities.

The manufacturing sector provides an illustration of their approach. For the 12 countries as a whole, manufacturing output (in value-added) increased by 50% between 1973 and 1994, final energy consumption reduced by 9% and final energy consumption per unit of output fell by 40%. As suggestive evidence of small rebound effects in this sector, Schipper and Grubb note that:

- Energy intensities continued to fall even when energy prices were falling (e.g. after 1986), although the rate of change did slow somewhat. This is claimed to demonstrate that there is been no significant substitution towards energy and that technical progress has achieved savings on all inputs. But the energy price trends are presented in absolute terms, rather than relative to other inputs, and the decomposition methodology does not permit substitution effects to be separated from those of technical change. For example, the database includes the sectors and time period for which Hogan and Jorgenson (1991) found energy using technical change.
- Output from sectors that showed the greatest decline in energy intensity grew slightly more than output from other sectors and faster output growth was associated with more rapidly falling intensities. Both these observations could be taken as evidence of a small rebound effect, but the direction of causality is unclear.
- In six countries, energy intensive industries increased their share of overall output during the period, even though these industries exhibited relatively modest declines in their energy intensity. This is opposite to what would be expected if there were significant rebound effects and suggests that many factors other than energy efficiency affect output growth.

For passenger transport Schipper and Grubb found that fuel use per kilometre (I) fell by some 30% in the US and Canada and also fell relative to other measures of activity such as GDP. Also, the US experienced the smallest increase in vehicle km per capita (*A*) and the greatest decline in intensity. These observations are claimed to demonstrate that the rebound effect in this sector is 'unimportant' (i.e. $\eta_I(A)$ is small), despite the fact that more disaggregated studies[41] indicate direct rebound effects of up to 30% (see *Technical Report 2*). They also find that the ratio of vehicle kilometres per capita to GDP per capita has varied little within each country during the period in question, despite large fluctuations in fuel costs. This suggests a low elasticity of vehicle kilometres with respect to fuel prices, and by implication a low elasticity with respect to energy intensity, since both change the cost per kilometre of driving (although *Technical Report 2* provides some caveats to this assumption). It also suggests that income, population and transport infrastructure play a more important role than fuel prices (and by implication, energy efficiency) in determining the amount of car travel. However, the data shows substantial differences in vehicle kilometres per capita, fuel use per kilometre and fuel prices between North America and Europe. This could partly reflect a long-term rebound effect, but it is difficult to distangle the separate influence of these co-evolving factors. Schipper and Grubb argue that "….wide-open spaces seem a more plausible explanation" and that "…..this is hardly the rebound effect at issue here." But this leaves open the question of where exactly the boundary for the rebound effect should be drawn. For Brookes and Jevons, the long-term effect is precisely what matters.

[41] Including a study by Johansson and Schipper (1997).

By weighting the individual activity and intensity indices by the share of each activity in total energy consumption in the base year,[42] Schipper and Grubb are able to develop aggregate indices for energy intensity (I_t^*) and activity (A_t^*) in each country. The aggregate energy intensity of all countries fell by 30-40% over the period, with most of the reductions between 1977 and 1986. The rate of improvement levelled off during the 1990s (when energy prices were low), but did not reverse. Aggregate activity levels increased in all countries, but grew less rapidly than GDP in all but three countries with no surge in activity levels following the energy price reductions of 1986, suggesting that $\eta_{P_E}(A)$ is small and by implication that $\eta_I(A)$ is small. These observations are taken as evidence for induced technical change, the irreversibility of energy efficiency investment (Dargay, 1992) and small rebound effects (i.e. small $\eta_{I_i}(A_i)$).

In all countries, the index of activity levels per unit of GDP (A^*/Y) decreased by less than the aggregate index of energy intensity (I^*), which in turn decreased by less than the energy/GDP ratio (E/Y) - suggesting that structural change reduced aggregate energy use.[43] Also, greater reductions in the energy intensity index were not associated with smaller reductions in the energy/GDP ratio, as may be expected if rebound effects encouraged both increased activity and structural change towards more energy intensive activities. Schipper and Grubb interpret this as evidence for the limited influence of falling energy intensity on both the level of each activity and the relative mix of activities, with the result that the reductions in energy intensity appear to have reduced energy consumption below what they would have been otherwise.

4.5.4.2 Evaluation and critique

Schipper and Grubb's approach allows them to demonstrate that the energy/GDP ratio is a poor proxy for changes in economy-wide energy intensity. Nevertheless, while their approach provides a substantial advance upon the use of energy/GDP ratios, it is still subject to some weaknesses. First, the total rebound effect from improvements in the thermodynamic efficiency of individual technologies may be greater than suggested by the observed growth effects (i.e. $\eta_{I_i}(A_i)$), for the reasons given in Section 4.5.3. Second, their decomposition methodology does not allow them to identify the relative contribution of structural change at the 'sub-activity' level, changes in product mix (in manufacturing) and changes in fuel mix, as well as the relative importance of price induced factor substitution and energy-saving technical change. If one or more of these is important for a particular activity, the implicit assumption that changes in aggregate measures of energy intensity result largely from underlying improvements in thermodynamic efficiency may be called into question. In practice, it seems likely that such factors will have considerably more influence

---

42 For example: $I_t^* = \sum_i w_i \frac{I_{i,t}}{I_{i,0}}$, where $w_i = \frac{E_{i,0}}{E_0}$. In this 'Laspeyres' index the weights are fixed and defined in relation to the base year *t=0* (Ang, 1999)

43 In general, an energy intensity index calculated as $I_t = \sum_i E_i / \sum_i A_i$ (e.g. the energy/GDP ratio) will only be equivalent to a weighted energy intensity index such as $I_t^* = \sum_i w_i (I_{i,t} / I_{i,0})$ when changes in the individual component intensities are the same and when the share of each activity remains the same. The difference between the two can be regarded as a measure of structural change in the economy.

on economic measures of energy efficiency in manufacturing than on physical measures of energy efficiency in transport and households.

To resolve these issues empirically, a decomposition analysis would need to use the highest level of disaggregation that the data permits, attempt to separate price induced factor substitution from technical change and apply some form of energy quality weighting when different fuels are combined. Recent work by Sue Wing and Eckhaus (2006b; 2006a) (discussed in Section 4.4.4), represents an important step in this direction, although they still neglect changes in fuel mix. Sue Wing and Eckhaus (2006a) show that using higher levels of aggregation in the definition of activities leads to a downward bias in estimates of structural change and a upward bias in estimates of intensity change as a result of the misattribution of shifts in the mix of 'sub-activities' to energy efficiency improvements. They also show that conventional assumptions for the magnitude of the *AEEI* parameter (e.g. 1%/year) may be seriously in error. Similarly, Kaufmann (1992) demonstrates that economy-wide improvements in energy intensity may be explained quite independently of energy-using technical change, provided changes in energy quality are taken into account (Section 4.3.3). While these observations do not invalidate Schipper and Grubb's conclusions, they do suggest the potential importance of variables that their study neglects.

Schipper and Grubb acknowledge that rebound effects may be large in particular situations:

> "… where energy costs are a key input or even a constraint on output - the iron/coal example of Jevons... or space heating for low income families - there can indeed be significant rebounds and even increases in energy use as a direct result of greater efficiency." (Schipper and Grubb, 2000)

However, these cases are considered to be "rare exceptions" in developed economies: first, because there is little evidence for such effects in their data; and second, because energy forms a small share of total costs, leading to low price elasticities. However, the same conclusion may not follow for energy efficiency improvements in developing countries:

> "…..In low income countries, energy and energy costs is often a constraint on industrial activity…..And for more than a billion consumers who rely on gathered firewood and other renewables and no electricity whatsoever, the time alone required to gather energy and to carry out tasks without any mechanical assistance means little time for participation in the commercial economy. As households hook up to the formal economy, we would expect a significant burst of commercial energy use both because of its efficiency in saving time and because higher incomes permit purchase of more appliances that use this energy. This must also be true of a billion or so households in the formal economy with a minimum of appliances and a clear income constraint on commercial energy use......In short, the shadow of Jevons lurks here for the same reason the more efficient coal use did not save coal." (Schipper and Grubb, 2000)

In other words, rebound could easily become backfire for energy efficiency improvements affecting one to three billion households in the developing world for the foreseeable future. This conclusion is of fundamental importance for global climate policy, but remains largely overlooked in the existing literature since this is focused almost entirely upon rebound effects in developed economies.

Schipper and Grubb cite the low cost share of energy in developed countries as the primary reason for the apparent absence of significant rebound effects in their data:

> "….[in manufacturing] the cost share of energy in total output on average is under 6%, hence even a cost-free 35% reduction in energy intensities, implying a 2% reduction in costs on average, could hardly lead to booming growth in output." (Schipper and Grubb, 2000)

> "….if economies are now 75% as energy intensive as they were in 1973, what could eat up the savings? It would require an additional 33% growth in GDP... It is fanciful to suppose that the 20% energy savings, together with the restraining effect of structural changes, would boost GDP by 33%." (Schipper and Grubb, 2000)

But these arguments are potentially flawed for the reasons given in the preceding sections. The diffusion of energy efficient technologies *may* have contributed substantially to the observed GDP growth *if* they also boosted the productivity of other factors. This appears likely for the energy efficiency improvements that are associated with general improvements in process and product technology (e.g. the move from desktop PCs to laptops), rather than dedicated investments to improve energy efficiency (e.g. insulation). To the extent that the observed improvements in energy intensity can be attributed to technical change, the technologies involved may often have been of this more general, productivity-enhancing form. By implication, the observed change in activity levels and overall GDP may partly be the result of the diffusion this type of technology. If this is the case, the overall boost to economic activity achieved by such technologies may be larger than implied by the share of energy in total costs - especially in manufacturing.

As discussed in Section 3.4, whether any increase in energy consumption from the diffusion of such technologies can meaningfully be labelled as a rebound effect depends upon the appropriate choice for the independent variable. Since their study is confined to more aggregate measures of energy intensity and activity, Schipper and Grubb overlook the possibility that rebound effects could be larger if a different independent variable was chosen. Ultimately, this issue can only be settled empirically, but it does suggest that Schipper and Grubb's failure to find significant rebound effects in their data may not necessarily mean that they are not there.

4.5.5 Summary

This section has explored Brookes' argument that improvements in energy efficiency have provided a necessary condition for the historical growth in economic output and human population. Such an argument may be interpreted as implying that the increased availability of 'useful work' (i.e. the output from energy conversion devices) has provided a necessary condition for the historical growth in economic output. This, in turn, may derive from either improvements in thermodynamic conversion efficiency or the increased availability of low-cost energy supplies - especially of higher quality energy forms. Hence, the so-called 'endogneity argument' represents a particular way of expressing the strong relationship between energy consumption and economic growth – which is argued to be the underlying theme of Brookes' work.

While it is extremely difficult to evaluate this claim empirically, this section has explored its implications for the widely used technique of decomposition analysis in energy studies. In particular, it has the highlighted the potential limitations of two studies by Laitner (2000) and Schipper and Grubb (2000), which both claim that economy-wide rebound effects are small. These studies neglect the potential influence of variables such as changes in energy quality, which may play a key role in achieving aggregate reductions in energy intensity (Section 4.4). They also confine attention to relatively aggregate measures of energy intensity, thereby overlooking the possibility that rebound effects could be larger if a different independent variable was chosen. They also assume that rebound effects must be small because the share of energy in total costs is small, thereby overlooking the possibility that new, energy-efficient technologies may significantly improve total factor productivity. As a result, neither study can be said to provide a definitive argument against either the 'endogneity argument' or the K-B postulate.

However, since Brookes has not provided a definitive, empirically-based argument in support of the 'endogneity argument', the question remains open. Some progress could potentially be made by applying a combination of decomposition and econometric techniques to relatively disaggregated data, as illustrated by Sue Wing and Eckhaus (2006b). An alternative approach, taken by Saunders (1992) is to use growth theory, discussed later in Section 5. But Brookes has pointed to another source of evidence, which can be related to more recent research on 'Environmental Kuznets Curves'. This will be examined in the next section.

4.6 Accommodation and Environmental Kuznets Curves

4.6.1 Introduction

Brookes commonly distinguishes between two situations: first, where high energy prices are a 'constraint' on the level of economic activity, and second where they are not. He argues that energy efficiency improvements lead to backfire in both cases.

Brookes considers the first situation to be the most common. Following Schurr, he argues that this situation is characterised by rapid growth in total factor productivity, facilitated by the substitution of energy for capital and labour inputs. The second situation corresponds to the period immediately after an energy supply shock, such as occurred in 1973 and 1979. During these (relatively short and uncommon) periods, Brookes argues that the 'normal' situation is reversed, with falling total factor productivity and with capital and labour substituting for energy.

Brookes' observations of the economic effects of the first oil shock led him to the following conclusion - versions of which appear several times in his writings:

> "….the first OPEC price hike had resulted in energy price being a major constraint on economic activity. The effect of relieving such a constraint by raising the level of energy efficiency is to *accommodate* the price rise, shifting the demand curve so that the balance between supply and demand is struck at a higher price and hence higher level of production and consumption than if no energy efficiency response had taken place. Consumption is lower than before the price rise but not so low as in the absence of a conservation response……This is not to say that it was wrong to respond to energy price constraints by attempts to raise energy efficiency, only that the effect of doing so at the highest level of aggregation is for total energy consumption to be higher than without such a response…" (Brookes, 2000)

When explored in a partial equilibrium framework, with the level of economic activity fixed, this argument is puzzling. For example, Brookes talks about energy efficiency investment "…. shifting the demand curve of oil to the right, causing the world balance between supply and demand for oil to be struck at a higher price." (Brookes, 1990b). But conventionally, energy efficiency investment would be interpreted as shifting the demand curve to the left, reducing energy demand and with it the price of oil. While rebound effects could shift the demand curve to the right again, this would only result in backfire if the elasticity of aggregate energy service demand exceeded unity – which is an unproven assumption. However, Brookes is *not* assuming that the level of economic activity is fixed, but instead that it depends closely upon the level of energy prices.

Brookes (1984) presents an analytical model that make sense of the accommodation argument - a step that is missing from his later papers.[44] The important step is to consider the impact of energy price rises on GDP as well as on investment in energy efficiency. Brookes (1984), in turn, draws upon Brookes (1972) and (1978). Oddly, practically all commentators discuss Brookes' later papers (1990b; 1992; 1993; 2000) but ignore the original analysis. However, without the latter, you can't make sense of the former.

This section summarises and critiques Brookes' analytical model, together with the empirical research upon which it is based. It also summarises more recent research on Environmental Kuznets Curves that appears relevant to this argument.

4.6.2 The income elasticity of useful energy demand

4.6.2.1 Brookes' hypothesis

Brookes 'accommodation argument' hinges upon the hypothesis that the income elasticity of *useful* energy consumption (*U*) for the economy as a whole is greater than or equal to one ($\eta_Y(U) \geq 1.0$) (Brookes, 1972).[45] The notion of 'useful' energy demand is based upon Adams and Miovic (1968a), which represents an early attempt to derive an aggregate measure of final energy consumption that takes energy quality into account.

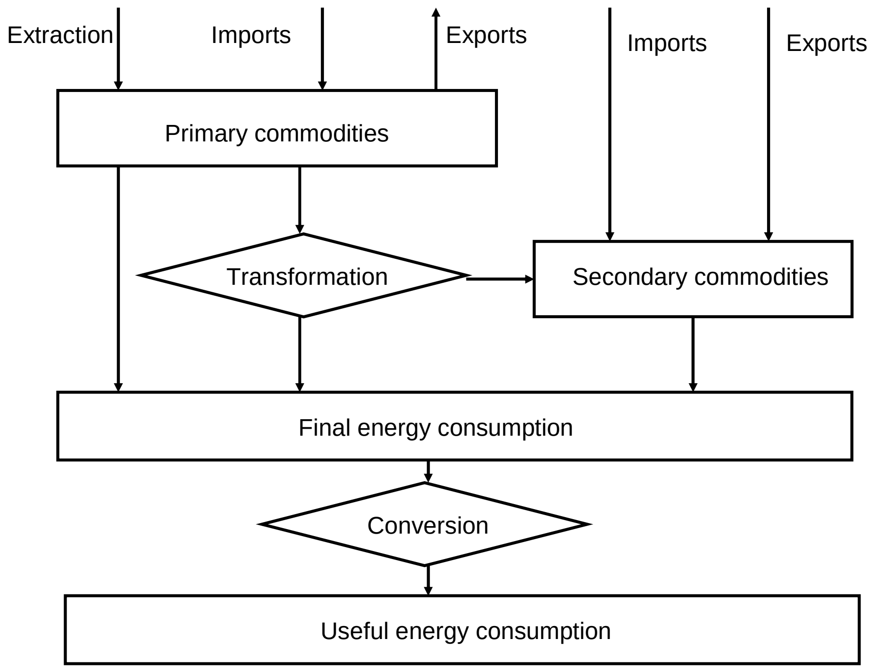
Adams and Miovic argue that measuring energy consumption on a kWh basis is inadequate because the thermodynamic efficiency with which different fuels are utilised varies greatly from one application to another. This suggests a distinction between final and useful energy consumption as indicated by Figure 4.1, with useful energy being interpreted as the output from the conversion devices employed by end users. However, Adams and Miovic do not use engineering estimates of these conversion efficiencies, but instead estimate the *relative productivities* of different fuels using pooled annual cross-section data of final energy consumption (*E*) and GDP (*Y*) for EU Member States over the period 1950-1962. Specifically, they estimate GDP as a function of the weighted final consumption of different energy carriers: $Y = \sum_{i=1,n} w_i E_i$ - where E_i represents the final consumption of energy carrier *i* in kWh and the coefficients w_i are a measure of the average productivity of that energy carrier, or

[44] Brookes (1984) is a book chapter and therefore has not been subject to peer review.

[45] Brookes sometimes talks about the 'useful energy coefficient' and sometimes about the income elasticity of useful energy demand (Brookes, 1972). These are not the same thing (Ang, 2006). The useful energy coefficient for a given period is defined as the ratio of the average annual growth rate of useful energy consumption (*U*) to the average annual growth rate of GDP (*Y*). Thus, the useful energy coefficient between year 0 and year *n* is given by: $c_{0,n} = [(U_n / U_0)^{1/n} - 1]/[(Y_n / Y_0)^{1/n} - 1]$. But when the growth of one variable is positive and that of the other is negative, the interpretation of this indicator becomes complicated. In this context, the income elasticity of useful energy demand - the ratio of the proportionate change in *U* to the proportionate change in *Y* – is a better measure: $\eta_Y(U) = (\partial U / U)/(\partial Y / Y) = \partial \ln U / \partial \ln Y$.

its contribution to GDP.[46] The estimates of w_i are then used to weight the final consumption of different energy carriers to obtain an aggregate measure of 'useful energy' consumption (*U*): $U = \sum_{i=1,n} w_i E_i$.

Figure 4.1 Primary, secondary, final and useful energy



With this approach, Adams and Miovic find that the productivity of electricity is 9.9 times greater than that of coal, while that of petroleum is 1.6 times greater - implying that a shift from coal to electricity or petroleum in final consumption would increase conventional measures of energy productivity (reduce energy intensity) as well as increasing economic output. Hence, while their measure of energy quality is cruder than the Divisia index (Berndt, 1978), the net result is to give a greater weight to higher quality fuels.

Over the period 1950-1962, the consumption of useful energy in the EU grew much faster than the consumption of final energy (Adams and Miovic, 1968a). While the income elasticities of final energy consumption were less than one, the income elasticities of useful energy consumption were greater than one, suggesting that useful energy consumption increases *more* than proportionally with GDP. Brookes (1972) takes these results further by developing a simple hypothesis regarding the relationship between useful energy consumption per capita (*u=U/P*) and GDP per capita (*y=Y/P*). (note that Adams and Miovic did not make the per capita adjustment and did not claim any long-term relationship between these variables). Brookes hypothesises that as an economy moves through various stages of economic development, the income elasticity of *useful* energy demand steadily falls from a high-value tending asymptotically to unity. This implies that useful energy demand can never grow slower than GDP. Brookes does not provide a rigorous theoretical basis for this hypothesis, but instead points to correlations between GDP per capita and

[46] Productivities are measured relative to the coal. The consumption and productivity of oil is not differentiated by product type, and (reflecting the age of the study) the productivity of natural gas is not measured separately.

'energy' consumption per capita[47] and argues that '…..there is no evidence that countries with a high proportion of their output devoted to services are able to reduce their energy inputs per unit of output'. This assertion is repeated in subsequent papers (Brookes, 1990b) and runs counter to conventional wisdom.

Brookes (1972) tests this hypothesis with an analysis of panel data on useful energy consumption per capita (*u*) and GDP per capita (*y*) for 22 selected OECD countries over the period 1950 to 1968. Useful energy consumption is calculated from data on final energy consumption using the weights derived by Adams and Miovic - although in principle these should change with time. A plot of the annual, cross country means of the income elasticity of useful energy consumption is downward sloping, beginning around 1.4 in 1950 and finishing around 1.25 in 1968. Also, the slope of mean values for the less developed countries is steeper than that for the more developed countries. Both of these observations appear consistent with Brookes' hypothesis. A non-linear curve with an asymptopic value of ~1.0 proves consistent with the data and provides a better fit than a straight line.

This analysis appears to form the basis of the analytical model in Brookes (1984), which in turn forms the basis of the subsequent assertions that energy efficiency 'accommodates' an energy price rise, leading to greater energy consumption. The argument could therefore be challenged if more recent research provided evidence that the income elasticity of useful energy consumption was less than one, or was even negative.

4.6.2.2 Environmental Kuznets curves

Since the early 1990s, there has been an explosion of empirical research into so-called Environmental Kuznets Curves (EKC). This is a hypothesised, inverted U-shape relationship between various indicators of environmental degradation and income per capita, first popularised by the World Bank (IBRD, 1992). The EKC hypothesis suggests that, beyond a certain level of income, the income elasticity of environmental degradation will become negative – implying that economic growth is necessary to reduce environmental impacts. Typically, the indicator is modelled as a quadratic function of the logarithm of income. 'Proximate' explanations for EKC include structural changes, changes in inputs, higher productivity and emission technologies, while 'underlying' explanations include environmental regulations and increased preferences for environmental quality (Stern, 2004c).

The EKC literature includes a large number of studies on energy use and/or carbon emissions (see the review in Richmond and Kaufmann (2006). However, none of these studies use a quality adjusted measure of final energy consumption. Instead, most studies focus on primary energy consumption (measured in thermal content) and/or carbon emissions. As a result, they do not provide a direct test of Brookes' hypothesis.

Earlier studies do appear to find evidence for an EKC for primary energy consumption and/or carbon emissions (Tucker, 1995; Schmalensee*, et al.*, 1998). However, Stern (2004c) is one of a number of authors to criticise the methodologies used:

> "… most of the EKC literature is econometrically weak. In particular, little or no attention has been paid statistical properties of the data used - such as serial dependence of

[47] Brookes is not consistent in his use of the term 'energy' in either the 1972 or the 1984 paper. Sometimes he appears to be talking about useful energy consumption and sometimes about final, or even primary, energy consumption. In this case, he appears to mean final energy consumption.

> stochastic trends in time series - and little consideration has been paid to issues of model adequacy such as the possibility of omitted variable bias....When we do take diagnostic statistics and specification tests into account and use appropriate techniques, we find that the EKC does not exist (Perman and Stern, 2003)." (Stern, 2004c)

Similar points are made by Richmond and Kaufmann (2006), who also suggest that most existing studies are biased by the omission of energy prices. Their rigorous study of energy use in OECD countries is also one of the few that takes energy quality into account (Richmond and Kaufmann, 2006). However, rather than using a quality adjusted measure of energy consumption as the dependent variable, they use primary energy consumption calculated in heat units. Changes in energy quality are allowed for by including three dependent variables representing the share of coal, oil and hydro/nuclear in final energy consumption.[48] With this framework, they find *no* evidence for a turning point in the relationship between income and energy use. The income elasticity of primary energy consumption is less than one, but not negative, implying that the relationship may better be represented by monotonically decreasing curve rather than an EKC.

Brookes' hypothesis for the income elasticity of useful (i.e. quality-adjusted final) energy consumption is not necessarily inconsistent with an EKC for primary (i.e. non-quality adjusted) energy consumption. Negative income elasticity for the latter implies that primary energy consumption is falling as income increases. This may result from efficiency improvements in the transformation of primary fuels, shifts towards higher quality fuels, or a reduction in final energy consumption. Similarly, reductions in final energy consumption may result from improvements in the productivity of final energy use, shifts towards higher quality fuels or a reduction in useful energy consumption. Hence, for falling primary energy consumption (an EKC) to coincide with increasing useful energy consumption (Brookes' hypothesis), offsetting shifts towards higher quality fuels and/or improvements in conversion efficiency and energy productivity are required.[49] This is entirely possible, but arguably becomes less likely as the income elasticity of primary energy consumption becomes more negative. Thus, while the existing EKC evidence cannot be used to reject Brookes hypothesis, it seems reasonable to argue that stronger (weaker) evidence for an EKC in primary energy consumption makes Brookes hypothesis less (more) plausible. Since the most recent research cast doubt on the existence of an EKC for primary energy consumption, Brookes' hypothesis remains plausible. What is needed, of course, is an EKC study that employs useful energy consumption (i.e. quality adjusted final energy consumption) as the dependent variable.

4.6.3 The analytical model

Brookes (1984) uses this hypothesis as the basis of a simple theoretical model, in which the relevant behaviour of the world economy is represented by three simple equations. To avoid complications of fuel substitution and differences in fuel quality, he assumes a single homogeneous source of energy. The model is used to explore the economic impact of an exogenous energy price shock. The assumption underlying this model is that energy prices influence the long-term level of useful energy consumption, which in turn strongly influences GDP per capita. Brookes criticises conventional approaches for taking GDP growth as largely exogenous and modelling energy demand as a consequence of that growth, albeit modified

[48] The approach is therefore very similar to Kaufmann (1992), discussed in Section 4.3.3.

[49] Although the analysis in Brookes (1972) assumes that the productivity of individual fuels is fixed.

by responses to changes in energy prices. Instead, his focus is the feedback upon the original assumption of GDP growth that might follow from a reduction in useful energy demand following an increase in energy prices.

To back up his empirical work on the income elasticity of useful energy demand, Brookes refers to Schurr's (tentative) conclusion that:

> "….it is energy that drives modern economic systems rather than such systems creating a demand for energy." (Brookes, 1984).

This claim appears to be key to understanding Brookes' work and will be returned to in Section 6. Brookes also quotes empirical work[50] by Kouris suggesting that national income alone is sufficient to explain long-term changes in the demand for energy, implying that there is no role for energy prices (Brookes, 1984). This seems hard to square with Kouris' published work, as well as contradicting conventional assumptions (Kouris, 1983).[51] Nevertheless, it appears central to Brookes' model:

> "The fact that the long-term trend [in energy consumption][52] can be wholly explained by national income changes, without the intervention of energy price, may lead some people to conclude that energy price is not an important factor in determining the level of energy consumption. It leads me to the opposite conclusion... that it may be the level of energy consumption that is greatly influencing the level of output per capita (hence national income) with energy price influencing the long-term level of energy consumption. This means putting the causality the opposite way round from the familiar assumption, and it means that (at one remove) energy prices may vary well be a major influence on the equilibrium level of economic activity..." (Brookes, 1984)

These considerations lead to a model that abstracts from capital and labour inputs and assumes a linear relationship between GDP (*Y*) and useful energy consumption (*U*):

$$Y = (A \cdot P_E^c) \cdot U + B \qquad (4.9)$$

The model effectively divides the economy into two components: a fixed 'non-energy dependent' component (*B*) and a growing energy dependent component with additional units of output requiring uniform increments of useful energy consumption. Energy saving technical change (i.e. a positive *AEEI*) is subsumed within the model as part of the definition of useful energy. Hence, improvements in the efficiency of energy use that take place as part of this underlying trend are treated as an addition to useful energy inputs. In contrast, price induced substitution changes the slope of the relationship (*A*) by an amount determined by the 'price elasticity of conservation response', represented by *c.*

The relationship of useful energy demand to energy prices (P_E) is assumed to take the following form:

$$U = K \cdot Y \cdot P_E^{-a} \qquad (4.10)$$

---

[50] The reference is to comments by Kouris at a conference.

[51] To our knowledge, Kouris has never argued that energy prices do not influence the long-term level of energy demand. Instead, he has argued that: "….a long-term income elasticity does not have a direct impact on energy demand and hence is of doubtful value in energy demand projections." (Kouris, 1983).

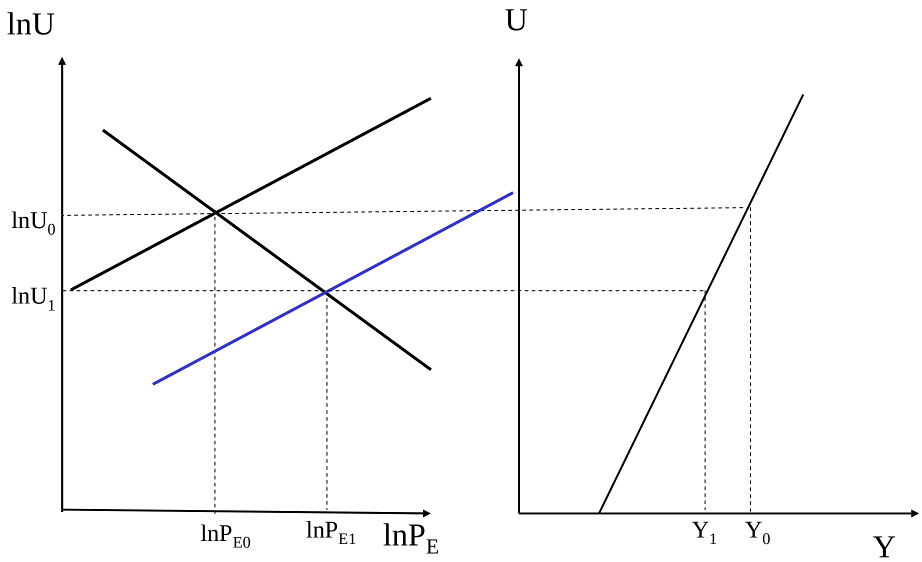
[52] Again, it is not clear whether Brookes is talking about primary, final or useful energy consumption here.

Where *a* represents the price elasticity of demand for energy 'subject to income remaining constant'. The final element of the model is an equation for energy supply:

$$U = C \cdot P_E^b \tag{4.11}$$

Given this framework, Brookes' argument appears to run as follows.[53] First, an exogenous price shock shifts the supply curve to the right, increasing energy prices from P_{E0} to P_{E1} and reducing useful energy demand from U_0 to U_1 (Figure 4.2). Given the assumption of a linear relationship between economic output and useful energy inputs (Equation 4.9), this leads to a corresponding reduction in economic output from Y_0 to Y_1.

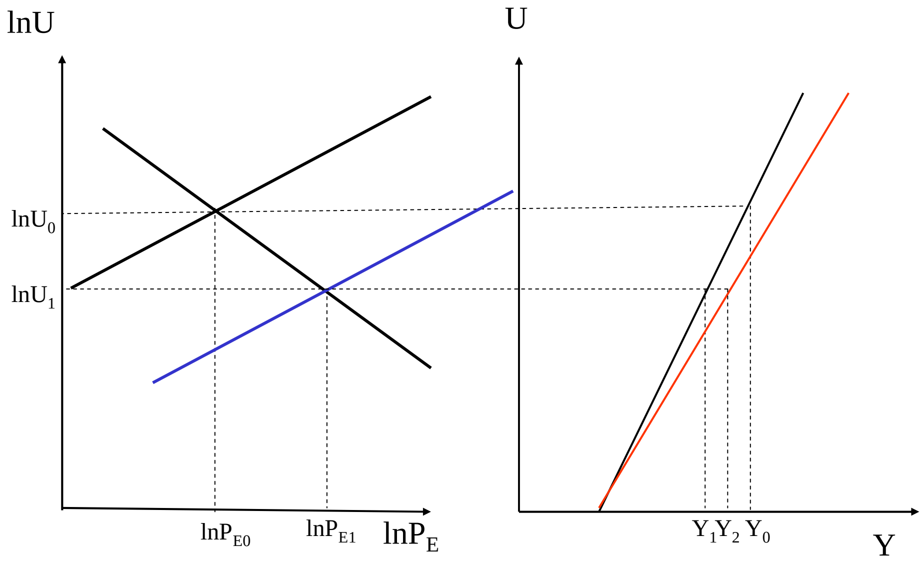
Figure 4.2 Adjustment to a supply shock in Brookes' model - 1



Given the assumption of a positive 'price elasticity of conservation' ($c>0$), substitution of capital and labour for energy inputs acts to *recover* some of the lost output, which increases to Y_2. (Figure 4.3). However, this is not sufficient to compensate for the initial reduction in activity ($Y_1<Y_2<Y_0$). The interpretation of this process is discussed further below.

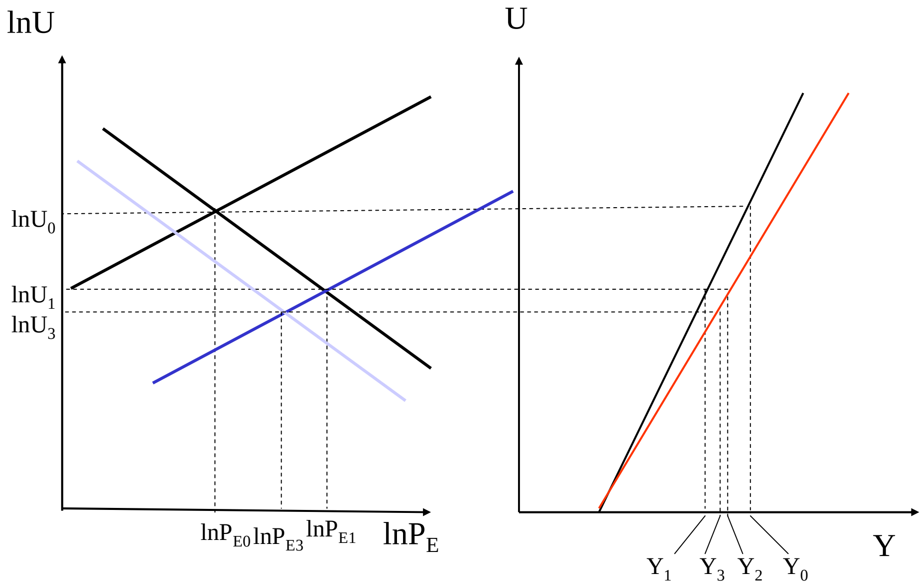
[53] We have found Brookes' exposition of this model very difficult to interpret. This account is our own interpretation and includes some intermediate steps that are missing from Brookes (1984)

Figure 4.3 Adjustment to a supply shock in Brookes' model - 2



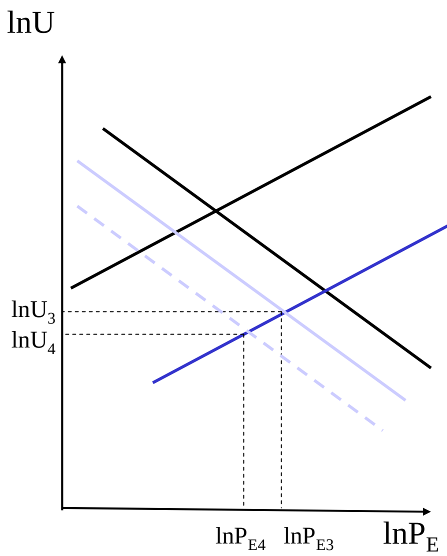
From equation 4.10, the net reduction in economic activity produces a leftwards shift of the demand curve (Figure 4.4). This leads to a new equilibrium, with a level of economic output (Y_3) which is *lower* than before the price shock ($Y_3 < Y_0$) but *higher* than it would have been without any price-induced energy efficiency improvements - represented by the change in slope of the energy consumption/economic activity sub-model ($Y_3 > Y_1$). Similarly, energy prices are *higher* than before the price shock ($P_{E3} > P_{E0}$) but *lower* than they would have been without the price-induced substitution ($P_{E3} < P_{E1}$). The results are sensitive to the particular values assumed for the individual parameters. Brookes conducts sensitivity tests for a range of assumptions, which suggests that a doubling of energy prices may reduce economic activity by between 6 and 16%

Figure 4.4 Adjustment to a supply shock in Brookes' model - 3



The key point for the rebound debate is that, while useful energy consumption is lower than before the price shock ($U_3<U_0$), it is *higher* than it would have been without the price-induced energy efficiency improvements (i.e. substitution of capital for energy). This reason for this is illustrated in Figure 4.5. Without the efficiency improvements, economic activity would be lower ($Y_1<Y_3$) and the leftwards shift of the demand curve would be greater. This leads to a *lower* equilibrium level of energy demand (U_4) than would have been the case without the efficiency improvements ($U_3>U_4$). In other words, improvements in energy efficiency have acted to increase aggregate energy consumption (backfire) by restoring some of the lost economic output. This is what Brookes means by such improvements 'accommodating' an energy price rise.

Figure 4.5 Adjustment to a supply shock in Brookes' model - 4



4.6.4 Problems with the analytical model

The model is very difficult to interpret, partly because it adopts an unconventional approach. For example, a leftwards shift of an energy demand schedule is normally used to represent energy efficiency investment, but in this case it represents a reduction in economic activity - with energy efficiency investment being represented instead by a change in the slope of the energy consumption/economic activity sub-model. Conventionally, price induced factor substitution would be expected to reduce economic output since, if the price of other inputs remained unchanged and if substitution possibilities are less than perfect, it will be more costly to produce the same level of output (Hogan and Manne, 1970). This implies that the parameter *c* in Equation 4.9 should be negative. But in Brookes' model, any reductions in economic output derive solely from the price-induced reduction in useful energy demand (Equation 4.10). The parameter *c* is assumed to be positive, with a positive impact on economic activity.

It is difficult to develop a clear interpretation of parameter *c*. Price induced substitution could be interpreted as increasing conversion efficiency and hence the amount of useful energy (*U*) from each unit of final energy (*E*). This would increase economic output (*Y*) per unit of final energy input (*E*). But Brookes does not distinguish clearly between useful and final energy. Instead, the process appears to increase the productivity of useful energy. This follows from the original definition of useful energy in Adams and Miovic (1968b), in which the average productivity of individual fuels is estimated econometrically.

The model is only stable if the short-term price elasticity exceeds the 'price elasticity of energy conservation response' ($a>c$).[54] If this condition is not met, the model explodes. In fact, stability is only achieved through the introduction of the constant term B. Without this, the solution is trivial, since the E's cancel and we are left with a relationship between the parameters, with no relationship between the variables. However, the practical interpretation of the constant term - the 'non-energy dependent component of the economy - is unclear. One possible interpretation could be a component that is not dependent upon the energy derived from commercial fuels.[55] However, this is not expanded upon by Brookes and the empirical evidence for this parameter derives solely from Brookes (1972). Similarly, the behaviour of the model is highly sensitive to the value of *c*, but since this represents something different from the elasticities of substitution conventionally measured in empirical studies, the model is difficult to calibrate.

Brookes claims that empirical support for the model could be provided if the proportion of national income spent on energy could be shown to keep within a relatively narrow band. He presents data showing that UK energy expenditure without taxation accounted for a mean of 8.26% of GDP over the period 1955 to 1976, with a standard deviation of 0.27%. Corresponding figures for the period 1980 to 2005 are 5.92% and 0.73%. This suggests a slight decline in the economic importance of energy, but whether this undermines or supports the model depends upon how 'narrow' is defined.

Ultimately, the validity of the model hinges upon the assumed causal relationship between useful energy consumption and economic output represented by Equation 4.9. But the cited evidence for a positive correlation between these two variables (Brookes, 1972) does not provide a sufficient basis for causal analysis, since it is equally possible that useful energy consumption provides an instrumental variable for the growth of other factors contributing to economic growth, such as education, telecommunications, infrastructure and so on. Brookes states that capital and labour inputs are implicit in slope of the model, but the framework does not allow for changes in these, or indeed for any form of technical change other than energy efficiency improvements. It therefore provides a very dubious basis for drawing conclusions about aggregate economic behaviour.

4.6.5 Summary

Brookes' has repeatedly argued that energy efficiency 'accommodates' an energy price shocks and increases energy consumption above what it would have been in the absence of the price shock. This argument is based upon a highly simplified theoretical model of the world economy, which is both unconventional in approach and difficult to interpret and calibrate. The model rests on the assumption that economic output can be adequately represented as a linear function of useful energy inputs, with no allowance for technical change or increases in capital and labour inputs. Since this would be unacceptable to most economists, the model provides a questionable basis for drawing general conclusions about economic behaviour.

---

54 Substituting Equation 4.10 into Equation 4.9 leads to the following relationship: $U = (KBP^{-a})/(1 - KAP^{c-a})$. Here, the numerator represents the normal price effect while the denominator represents the 'multiplier' effect of increases in economic output

55 For example, a subsistence economy using fuel wood is still using energy, but this not conventionally measured.

In support of the model, Brookes (1972) provides evidence that the income elasticity of useful energy demand in EU countries has historically exceeded unity. Quoting Schurr's observation that "….it is energy that drives modern economic systems rather than such systems creating a demand for energy", Brookes anticipates that this relationship will continue into the future (Brookes, 1984). Contemporary research on Environmental Kuznets Curves has not rejected this hypothesis, since useful energy consumption is not employed as the independent variable. Given the weakness of the evidence for an EKC for primary energy consumption, Brookes' hypothesis remains plausible.

Ultimately, the theoretical coherence of the model may matter less than the assumptions and empirical observations that lie behind the model, as well as Brookes' work more generally. Our interpretation of Brookes' work suggests that there are two underlying themes: first, that energy plays a more important role in economic growth than is commonly assumed; and second, changes in energy quality (especially electrification) have played a crucial but neglected role. Perhaps surprisingly, these ideas have more in common with ecological economics than with conventional neoclassical economics. The far-reaching implications that follow are discussed further in Section 6.

4.7 Summary and implications

This section has evaluated the arguments and evidence put forward by Len Brookes in favour of the K-B postulate. It is also used these arguments as a basis for evaluating a wider range of literature that appears relevant to the K-B postulate, such as that on Environmental Kuznets Curves. Brookes is seen to have provided three arguments in favour of the K-B postulate, namely:

- *The productivity argument*: The increased use of higher quality forms of energy (especially electricity) has encouraged technical change, substantially improved total factor productivity and driven economic growth. Despite the substitution of energy for other inputs, this technical change has stimulated a sufficiently rapid growth in economic output that aggregate energy efficiency has improved at the same time as aggregate energy consumption has increased.
- *The endogeneity argument*: A common approach to quantifying the 'energy savings' from energy efficiency improvements is to hold energy intensity fixed at some historic value and estimate what consumption 'would have been' in the absence of those improvements (Geller*, et al.*, 2006a). The energy savings from energy efficiency improvements are then taken to be the difference between the actual demand and the counterfactual scenario. But if the energy efficiency improvements are a necessary condition for the growth in economic output, the construction of a counterfactual in this way is misconceived.
- *The accommodation argument*: Energy efficiency improvements are claimed to 'accommodate' an energy price shock so that the energy supply/demand balance is struck at a higher level than if energy efficiency had remained unchanged (Brookes, 1984). While not immediately obvious, this argument rests on the assumption that the income elasticity of ' useful' energy demand falls steadily as an economy develops, but is always greater than unity (Brookes, 1972).

Brookes deserves credit for developing these and other arguments in favour of the K-B postulate and for defending these against a range of criticism. However, he acknowledges that the available evidence provides only 'suggestive' support for the postulate. The main conclusion from this review is that the arguments and evidence he provides are insufficiently robust to support his case. Flaws can be found both in the evidence itself in the manner in which Brookes uses this evidence to support the postulate. Specific criticisms include the following:

- *Productivity and energy quality*: Schurr's work applies primarily to the causal effect of shifts to higher quality fuels, rather than improvements in different measures of energy efficiency. Since the effect of those shifts on total factor productivity and aggregate energy consumption may not be the same as the effect of improvements in energy efficiency, this evidence is only indirectly relevant to the rebound effect. Also, the patterns Schurr uncovered may not be as 'normal' as Brookes suggests and the link between energy efficiency improvements and improvement of total factor productivity may vary greatly, both over time and between different sectors and energy services.
- *Productivity and biased technical change:* Neither Jorgenson's work itself, nor those of comparable studies consistently find energy-using technical change. Instead, the empirical results vary widely between different sectors, countries and time periods and are sensitive to minor changes in econometric specification. The assumption of a fixed bias in technical change is flawed, and more sophisticated models suggest that the magnitude and sign of technical change varies between sectors and types of capital as well as over time. Also, the failure to check for the presence of cointegration or to account for changes in energy quality in these studies means that many of the estimates could be either biased or spurious. Moreover, even if energy-using technical change were to be consistently found, the relationship between this finding and the K-B postulate remains unclear.
- *Endogeneity and decomposition analysis:* The endogeneity argument is rhetorically persuasive but lacks a firm empirical basis. The relative importance of energy efficiency improvements compared to other forms of technical change in encouraging economic growth remains to be established.
- *Accommodation and Environmental Kuznets Curves*: The 'accommodation' argument is based upon a simplified theoretical model of the world economy, which is both unconventional in approach and difficult to interpret and calibrate. The model appears to rest on the assumption that the income elasticity of 'useful' energy demand is always greater than unity, thereby allowing economic output to be represented as a linear function of useful energy inputs. A 1972 study by Brookes provide some support for this hypothesis, but this has not been updated. Contemporary research on Environmental Kuznets Curves has not tested this hypothesis, since useful energy consumption is not employed as the independent variable.

A key theme in Brookes' work is that improvements in energy productivity are generally associated with (proportionally greater) improvements in total factor productivity. While Schurr's work provides evidence for this at the level of the national economy, numerous examples from the energy efficiency literature provide evidence for this at the level of individual sectors or technologies If energy efficient technologies boost total factor productivity and thereby save more than energy costs alone, the argument that rebound effects must be small because the share of energy in total costs is small is undermined. Much the same applies to the contribution of energy efficiency improvements to economic growth. But this leaves open the question of whether energy efficiency improvements are *necessarily* associated with improvements in total factor productivity, or whether this contingent upon particular technologies and circumstances.

While Brookes fails to provide a convincing case in favour of the K-B postulate, he does highlight some important issues that are frequently neglected in conventional energy analysis. The assumption underlying Brookes' argument seems best summed up in Schurr's observation that: "….it is energy that drives modern economic systems rather than such systems creating a demand for energy" (Brookes, 1984). But in attributing a central role to energy in driving economic growth, Brookes has more in common with contemporary ecological economists than with conventional growth economists. In turn, the work of ecological economists may be as challenging to conventional wisdom as the K-B postulate. Three claims in particular have echoes in Brookes work and in the evidence reviewed above:

- The increased availability of higher quality forms of energy has been a necessary condition for the economic growth of the past two centuries, and will remain a condition for future economic growth (Hall*, et al.*, 1986).
- Much technical change is energy-using, in that it acts to increase aggregate measures of energy intensity. The observed reduction in these aggregate measures owes more to structural change and improvements in energy quality than it does to technological improvements (Kaufmann, 2004). Conventional assumptions about the magnitude and sign of the *AEEI* parameter in energy models are therefore flawed.
- The failure to find evidence of an Environmental Kuznets Curve for primary energy consumption suggests that decoupling energy consumption from economic growth may be more difficult than is conventionally assumed.

This review suggests that the case for the K-B postulate is rooted in a broader claim that energy plays a more important role in driving economic growth than is commonly assumed (or more specifically, a more important role than is suggested by the small share of energy in total costs). It is possible that this broader claim is valid, even if the K-B postulate does not hold. Unfortunately, the evidence for this broader claim is no easier to assess than that for the K-B postulate itself. Section 6 will review a number of approaches to this question that are largely taken from ecological economics. But prior to this, Section 5 will review some important and influential arguments in favour of the K-B postulate that are taken from neoclassical theory.

5 Energy productivity and neoclassical growth theory

5.1 Introduction

Harry Saunders introduced the notion of the 'Khazzoom-Brookes postulate' and has single-handedly brought a new level of sophistication to the rebound debate (Brookes, 2000). While Brookes advances a largely qualitative thesis, Saunders arguments are firmly grounded in neoclassical production theory and neoclassical growth theory. The relatively abstract and mathematical nature of this theory can be an obstacle to those who lack the relevant background, while the restrictive assumptions involved may be a reasonable focus of criticism. Saunders' work may also fail to convince those who seek empirical evidence for the K-B postulate since it is wholly theoretical. However, Saunders takes care to state that his results do not *prove* backfire, but merely provide suggestive evidence in its favour, given certain assumptions about how the economy operates. He also provides qualitative arguments and illustrative examples that clarify and reinforce his more formal derivations (Saunders, 2000b), and highlights the empirical work by Hogan and Jorgenson (1991) that was reviewed earlier in Section 4.4. Most importantly, he shows that backfire is the predicted outcome of *standard* economic models using *standard* assumptions about the functional form of production functions. This is a challenging result which needs to be taken seriously.

This section provides an account of Saunders' work, highlights some weaknesses and limitations and develops an extension of this work in relation to Constant Elasticity of Substitution (CES) production functions. The text is intended to be accessible to those unfamiliar with neoclassical production theory and neoclassical growth theory and uses appendices to introduce key concepts and provide derivations. Section 5.2 introduces the basic neoclassical model, including the role of technology. Section 5.3 summarises Saunders' use of this model to explore the rebound effect as well as his more recent work on 'fuel conserving' production functions and technology simulations. Section 5.4 discusses several theoretical considerations that are relevant to assessing this type of evidence and examines the conditions under which growth models with CES production functions may predict energy savings. Finally, Section 5.6 concludes with some implications of this work for the rebound debate.

5.2 The Solow model of economic growth

In 1956, Solow published a seminal paper on economic growth and development (Solow, 1956). For this work and his subsequent contributions to the understanding of economic growth, Solow was awarded the Nobel Prize in economics in 1987. The model provides one framework for exploring the determinants of long-run economic growth and for explaining the differences in output levels and growth rates across countries and over time. It is closely linked to the various approaches to measuring total factor productivity, discussed in Section 2, including the identification of the relative importance of growth in inputs and technical change in increasing economic output.

The basic structure of the Solow model is outlined in Annex 1. In its simplest form, the model assumes a single closed economy producing a single good that is used for both consumption and investment. The behaviour of the model is determined by two equations: an economy-wide production function and a differential equation describing the process of capital accumulation.[56] The basic model suggests that, under certain assumptions, an economy will move steadily towards a 'balanced growth' path - defined as a situation in which the rate of growth of national income is equal to the rate of growth of population which is assumed to be constant. In this 'steady-state', output per worker, capital per worker and consumption per worker are all constant. The model predicts that increases in the propensity to save and invest will increase the *level* of output, capital and consumption per worker, but will not change the long-run *rate* of growth. While the original version of the model did not include energy among the production inputs, it may be extended to do so. The inclusion of additional inputs does not change the main result, however, namely that in the steady state capital, labour and energy should all grow at the same natural rate (Saunders, 1992).

There are several intuitive theoretical insights that may be drawn from the growth model, without having to accept that the underlying assumptions simulate reality to any great extent. Moreover, although it rests upon a number of extremely simplifying assumptions, the model does fit some of the 'stylised facts' of economic growth (Kaldor, 1968). For example, it predicts that countries that have a higher savings/investment rate will tend to be richer and countries that have a higher population growth rate will tend to be poorer. In the case of the latter, a higher fraction of savings is required simply to keep the capital-labour ratio constant in the context of an increasing population. While the actual capital stock is growing, the capital stock per worker remains constant. Mankiw *et al* (1992) found that half of the cross-country variation in income per capita could be explained by variations in population growth and the savings rate alone.

However, the model does not predict the key empirical fact of economic growth, namely sustained increase in per capita income. The basic Solow model allows economies to grow for while, but not for ever. To generate sustained growth in per capita income, it is necessary to introduce technological progress into the model, which can be achieved in a variety of ways (Annex 1). As discussed in Section 3, the traditional approach is to introduce exogenous, time-dependent multipliers into the economy-wide production function. This form of exogenous technological progress has been called 'manna from heaven', since it is implicitly assumed to 'decend' upon the economy automatically and costlessly, at a constant rate. As with the other assumptions of the basic Solow model, this conception of technological progress is flawed and more recent approaches to growth theory have sought to make the rate and direction of technical change endogenous to the model (Aghion and Howitt, 1998).

The inclusion of technical progress into the neoclassical growth model leads to a higher rate of growth for economic output than the natural growth rate of population. Hence, the model with exogenous technology allows a closer representation of the long-run trajectory of national economies. Solow was among the first to notice that only by including technology in

[56] In the original formulation (and in keeping with standard neoclassical theory), capital (*K*) and labour (*L*) provide the only inputs to the production function. The apparent implication that output can be produced with no raw material inputs has attracted strong criticism from ecological economists (Daly, 1997), but this is only one of number of 'heroically' simplifying assumptions that underpin the basic Solow model. Whether such assumptions provide a convenient simplification for the purpose of isolating relevant causal mechanisms or a misleading abstraction from reality is a moot point. However, much of the subsequent development of growth theory consists of relaxing and modifying these assumptions, examining the implications and comparing these against empirical data.

the model can historical growth rates be replicated. Technical progress is nowadays considered to contribute a very large proportion of GDP growth.

To use the model, it is necessary to assume an appropriate functional form for the economy-wide production function. This is entirely analogous to the assumption of functional forms within CGE modelling (*Technical Report 4*), or in the econometric analysis of individual sectors (Section 4.4 and Annex 2). Functional forms can be estimated, but there is no universally accepted functional form, especially at the economy-wide level. The functional form assumption is of vital importance for insights from the Solow model, since different functional forms can yield radically different results - especially with regard to the impact of technical change on energy consumption. Whether this is a shortcoming or a useful insight is a moot point, but it appears especially important to the application of the model to the rebound effect.

5.3 Saunders' investigation of the rebound effect

Saunders (1992) used the neoclassical growth model to argue that backfire is a likely outcome of energy efficiency improvements. This conclusion was subsequently challenged by Howarth (1997) who argued that Saunders' failure to distinguish between energy and energy services led to the probability of backfire being overestimated. However, Saunders (2000a) subsequently demonstrated that backfire is still predicted by the neoclassical model when an alternative choice is made for the production function for energy services. In two more recent contributions, Saunders has focused on the potential of different types of production function to generate backfire, thereby contributing insights that are relevant to partial and general equilibrium analysis more generally, as well as to the neoclassical growth model specifically (Saunders, 2005; 2007). This section summarises the general approach and key insights from each of these studies. A discussion of the potential limitations of this work is postponed to Section 5.4.

5.3.1 Khazzoom-Brookes and neoclassical growth

Saunders (1992) argued that the neoclassical growth model is well suited to exploring the rebound effect, because it allows a theoretical examination of the effect of both neutral and biased technical progress on the economy-wide use of inputs over a period of several decades. This involves the joint effect of the overall increase in economic output and (in the case of biased technical change) changes in the mix of inputs as a result of changing relative marginal productivities.

In the absence of any form of technical change, the steady-state solution of the neoclassical growth model simply predicts that energy, capital and labour inputs and economic output will all grow at the same rate (Annex 1). By implication, the energy/GDP ratio will remain unchanged. As discussed in Section 3.3.3 and Annex 1, technical progress is conventionally represented in the growth model by exogenous, time-dependent multipliers ($\upsilon_i \geq 1$)[57] on the inputs (*i*) in the economy-wide production function. For example, the introduction of a multiplier on energy inputs ($\upsilon_E E$) implies that the economic productivity of energy inputs has increased. This means that the same output (*Y*) can now be obtained with less energy inputs, or alternatively that more output can be obtained from the same quantity of energy

[57] In full, this should be $\upsilon_i(t)$, but υ_i is is used for simplicity.

inputs. In this case, technical progress is said to be 'energy-augmenting'. The rate of technical progress is conventionally assumed to be constant (λ_i)[58] in each case and hence unaffected by developments within the economy.

For a three input (*KLE*) production function, technical progress may be neutral, capital augmenting, labour augmenting, energy augmenting or a combination of all four. Saunders represents this general form as follows:

$$Y = \upsilon_N F(\upsilon_K K, \upsilon_L L, \upsilon_E E) \qquad (5.1)$$

Where:

$\upsilon_N = e^{\lambda_N t}$ = neutral technical progress

$\upsilon_K = e^{\lambda_K t}$ = capital augmenting technical progress

$\upsilon_L = e^{\lambda_L t}$ = labour augmenting technical progress

$\upsilon_E = e^{\lambda_E t}$ = energy augmenting technical progress

Provided the production function is homogenous, an alternative formulation may be defined using only the input specific technology multipliers:

$$Y = F(\upsilon_K K, \upsilon_L L, \upsilon_E E) \qquad (5.2)$$

In practice, individual technologies (including energy efficient technologies) are likely to augment several inputs at the same time and so could be represented by a combination of the above. As argued in Section 4.2.3, if an energy efficient technology also augments capital and labour inputs, its contribution to economic growth may be significantly greater than suggested by the cost share of energy alone.

Saunders calls energy augmenting technical progress a 'pure' energy efficiency improvement and argues that it is closely related to improvements in 'engineering' (i.e. thermodynamic) measures of energy efficiency (ε) (Saunders, 2005). The terminology here can be misleading, however. The energy multiplier (υ_E) is a measure of the economic productivity of energy inputs (in this case, applied to the level of the macroeconomy) and could therefore change quite independently of any improvements in thermodynamic efficiency. Also, a thermodynamic measure of energy efficiency could not exceed unity ($\varepsilon \leq 1$), whereas there are no such restrictions on economic measures of energy productivity.

In a later paper, and in response to a critique by Howarth (1997), Saunders refers to the product $\upsilon_E E$ as 'energy services' (Saunders, 2006). In the terminology used in this project, it may be better to refer instead to 'useful work' (*S*). But, again, this is potentially misleading since, in thermodynamic terms, the outputs from an energy conversion device cannot exceed the energy inputs (i.e. $S = \varepsilon E$ and since $\varepsilon \leq 1$, *S<E*). In contrast, in Saunders' formulation, 'energy services' ($\upsilon_E E$) are greater in magnitude than energy inputs ($\upsilon_E E \geq E$ since $\upsilon_E \geq 1$). An alternative terminology, frequently used with neoclassical

[58] $\upsilon_i = e^{\lambda_i t}$, so the growth rate is given by: $\partial \ln \upsilon_i / \partial t = \lambda_i$

growth theory, is to refer to the product $\upsilon_E E$ as 'effective energy' inputs ($\tilde{E}$). The production function then translates 'effective' inputs into economic output (*Y*):

$$Y = \tilde{F}(\tilde{K}, \tilde{L}, \tilde{E}) \tag{5.3}$$

Where $\tilde{K} = \upsilon_K K$, $\tilde{L} = \upsilon_L L$ and $\tilde{E} = \upsilon_E E$.[59]

As described in Section 3.3.3, there is an important difference between a production function that maps the relationship between *real* inputs and economic output (Equation 4.2) and one which maps the relationship between *effective* inputs and economic output (Equation 4.3). Since technical progress changes the amount of real inputs required to produce a unit of output, this may be represented by a change in the location and/or shape of the individual isoquants of the first production function (*F*). However, the location and shape of the individual isoquants of the second production function ($\tilde{F}$) are assumed to be *fixed* over time: i.e. the amount of effective input required to produce a unit of output is assumed to be fixed. Technical progress instead changes the amount of real inputs required to produce a unit of effective input.

It should also be noted that the multiplier ($\upsilon_E = \tilde{E}/E$) on energy inputs is different from the aggregate measure of energy productivity ($\theta_E = Y/E$), defined earlier in Section 2. The latter may change as a result of price-induced substitution of one input for another, as well from any of the three types of technical change represented by the multipliers. It is also not the same as the *AEEI* parameter, defined in Section 2 as the growth rate of θ_E holding input prices constant:

$$AEEI = \left.\frac{\partial \ln \theta_E}{\partial t}\right|_{P_i = \overline{P_i}} = \left.\frac{\ln(Y/E)}{\partial t}\right|_{P_i = \overline{P_i}} \tag{5.4}$$

The *AEEI* isolates the contribution of technical change from the substitution effect that follows changes in relative prices. However, *any* form of non-neutral technical change may lead to a change in an aggregate measure of average energy productivity (*Y/E*) and hence in the *AEEI*. This is because (assuming prices are exogenous and fixed) increases (decreases) in the marginal productivity ($\partial Y / \partial i$) of one input (*i*) relative to another should lead to substitution towards (away from) that input. Hence, changes in the *AEEI* may not derive from energy augmenting technical change alone - or even at all. However, the terminology in this area appears to be inconsistent and the energy augmenting multiplier (υ_E) is often *identified* as the *AEEI* parameter.

Using a simple neoclassical growth model, Saunders (1992) simulated what would happen to output and energy consumption in the economy (by the year 2100) as a result of all four types of technical progress, with each assumed to be proceeding at the rate of 1.2% per year. The results were shown to depend closely upon the assumed functional form of the economy-wide production function, represented by Equation 4.3. The functional form describes the 'shape' of the production surface, representing the different optimal combinations of *effective* inputs that may be used to produce a given level of output using existing technology. The form places restrictions on this shape, while the particular

[59] See section 3.3.3 for a discussion of the relationship between this approach and total factor productivity.

parameter values determined it precisely.[60] The most important feature defining the different functional forms is the scope they provide for the substitution of one effective input for another, holding output constant (defined by the Hicks elasticity of substitution - σ - see Annex 2). As indicated above, this functional form is assumed to remain fixed over time.

In his 1992 paper, Saunders employed the widely used 'Cobb Douglas' and nested 'Constant Elasticity of Substitution' (CES) production functions (see Annex 2). More specifically, he employed a particular formulation of the CES that was first introduced by Hogan and Manne (1977) and which nests a Cobb Douglas production function for capital and labour within a CES function for energy and non-energy inputs (designated *(KL)E*).[61]

Using the Cobb Douglas production function within the neoclassical growth model, Saunders demonstrated that *all* forms of technical progress (i.e. including energy augmenting) lead to energy, capital and output growing faster than population. This implies a fixed energy/GDP ratio and growing consumption per worker. Most importantly, all forms of technical progress lead to economic output and energy consumption growing *faster* than without technical progress. In other words, with a Cobb Douglas production function, all forms of technical progress *increase* overall energy consumption. If energy augmenting technical progress is taken as the relevant independent variable for the rebound effect, then the neoclassical model with a Cobb Douglas production function inevitably leads to backfire.

The results for the CES specification were more ambiguous. Capital augmenting, labour augmenting and neutral technical progress were all found to increase overall energy consumption compared to a simulation without technical progress. But energy-augmenting technical progress only increased energy consumption if the Hicks elasticity of substitution between (effective) energy and (effective) non-energy inputs was greater than unity. The Hicks elasticity of substitution is a measure of the ease with which a decrease in non-energy inputs can be compensated by an increase in energy inputs - or vice versa - while output is held constant (Annex 2). Higher (lower) values of the elasticity of substitution mean that substitution is easier (more difficult). Two inputs are often said to be weak (strong) substitutes when elasticity of substitution is less than (greater than) unity (see Annex 2).

Saunders also found that energy augmenting technical progress *always* led to backfire for the two alternative nesting schemes in the Hogan-Manne CES (*(K(LE)*) and (*(KE)L*). This suggests that the one exception found to the K-B postulate[62] in Saunders' simulations could be relatively unique.

The finding of backfire with a Cobb-Douglas production function could perhaps have been anticipated. This production function was developed to model a situation where capital and labour earn a constant share of national income.[63] With a three input Cobb-Douglas, energy also earns constant share of national income. Hence, if real energy prices are fixed, it

---

[60] The different functional forms in common use are introduced in Annex 2.

[61] The CES took the form: $Y = \upsilon_N \left\{ a\left[(\upsilon_K K)^{\gamma}(\upsilon_L L)^{1-\gamma}\right]^{\rho} + b(\upsilon_E E)^{\rho} \right\}^{\frac{1}{\rho}}$, which represents a Cobb Douglas production function for capital and labour inputs, nested within a CES function ((*KL)E*). The Hicks elasticity of substitution between energy and non-energy inputs (σ) in the CES is given by $\sigma = 1/(1-\rho)$.

[62] Namely, energy augmenting technical progress in the Manne Richels form of nested CES production function, with a Hicks elasticity of substitution between energy and non-energy inputs of less than unity.

[63] This behaviour is consistent with one of the stylised facts of economic growth: the apparent constancy of the income share of labour and capital despite a steady increase in capital intensity (*K/Y*).

follows that the energy/GDP ratio is constant and that energy consumption grows at the same rate as GDP (Feather, 2005). This suggests that simulations are redundant for a Cobb Douglas production function (Wei, 2007). The implications of a CES are more complex and are discussed further below.

5.3.2 Howarth's critique

Saunders approach is consistent with the standard assumptions of the neoclassical growth model, in that improvements in input productivity ($\Delta\upsilon_i$) are assumed to simply 'appear' at zero cost and at a constant rate. This assumption was challenged by Howarth (1997), who argued (quite reasonably) that the provision of 'effective' energy inputs ($\tilde{E} = \upsilon_E E$) requires capital and labour inputs and hence have an associated cost.[64] Howarth (1997) explored the implications of this by simulating the provision of effective energy inputs with a Leontief production function: i.e. assuming a fixed ratio of inputs (see Annex 2). This was incorporated within a standard neoclassical growth model employing a Cobb Douglas form for the aggregate production function.

Using this approach, Howarth demonstrated that backfire only occurs when the elasticity of effective energy demand with respect to the energy productivity multiplier ($\eta_{\upsilon_E}(\tilde{E})$ is greater than unity.[65] In turn, Howarth shows that the magnitude of this elasticity depends upon the share of effective energy in total output costs, and the share of energy costs in the total cost of effective energy. In the majority of circumstances, *both* of these are likely to be small, so their product is smaller still. As a consequence, the demand for effective energy may be expected to be relatively inelastic with respect to changes in the energy productivity multiplier, with the result that the direct effect of improvements in energy efficiency will be greater than the feedback effects that result from the increased demand for effective energy. Howarth therefore concludes that backfire is unlikely in practice - directly contradicting Saunders' findings.

However, these criticisms were subsequently challenged by Saunders (2000a), who showed that Howarth's theoretical results stem entirely from his assumption of a Leontief (i.e. fixed-proportions) production function for the provision of 'effective energy' inputs (Saunders, 2000a). This functional form fails to represent the substitution possibilities that are likely to be available in practice and is poorly suited for use in growth models (Saunders, 2000a). Saunders showed that if the production function for energy services is assumed instead to take a Cobb Douglas form, the conclusion that improvements in energy productivity inevitably lead to backfire re-emerges. Saunders' arguments in this case again demonstrate the limited modeling capability of the Solow model, since theoretical results appear to depend so closely on the assumed functional forms.

---

[64] In a similar manner, Khazzoom (1980) neglected the capital costs associated with energy efficient equipment in his original study of the direct rebound effect. This neglect was subsequently challenged by several authors (Besen and Johnson, 1982; Einhorn, 1982; Henly*, et al.*, 1988; Lovins*, et al.*, 1988) who argued that it may lead to the the direct rebound effect being overestimated. Further details are provided in *Technical Report 2*, while useful simulations demonstrating the importance of capital costs can be found in Mizobuchi (2006).

[65] This approach to the economy-wide rebound effect has much in common with Khazzoom's approach to the direct rebound effect (for a discussion of the latter, see Technical Report 2). Although the argument is somewhat microeconomic, following the same aggregation rationale for production functions, we can say that this elasticity ($\eta_{\upsilon_E}(\tilde{E})$ provides a very rough approximation of the rebound at the macro level.

5.3.3 Fuel conserving production functions

Saunders original study found that that the simulated impact of improvements in energy productivity depends in large part on the assumed functional form for the economy-wide production function. However, this conclusion is not confined to the neoclassical growth model, but also extends to the use of such functions within partial equilibrium analysis. As a result, Saunders more recent work does not use the neoclassical growth model directly, but instead analyses the theoretical behaviour of the production functions themselves (Saunders, 2007).[66] The aim is to derive the exact mathematical conditions that are needed for different functional forms to yield energy savings, following energy-augmenting technical progress. The results provide an indication of the relative usefulness of different functional forms for exploring the rebound effect.

Saunders' (2007) defines a *short run* 'fuel conserving condition' for a production function as follows:

$$\left.\frac{\partial}{\partial \upsilon_E}\right|_{E=E^0}\left[\frac{\partial Y}{\partial E}\right] < 0 \qquad (5.5)$$

Here, $\partial Y / \partial E$ defines the marginal productivity of energy use, or the increase in output for a unit increase in energy inputs, *holding other inputs constant*. In equilibrium, this should be equal to the real market price of energy (p_E/p_Y), which is assumed to be fixed in order to isolate the impact of technical change. If the marginal productivity of energy use falls as a consequence of energy augmenting technical change ($\upsilon_E > 0$), then energy consumption must fall (thereby increasing the marginal productivity of energy use again) in order to restore the equilibrium (i.e. keep the marginal productivity of energy fixed). This condition applies to the short run because it is assumed that capital and labour inputs are held fixed. In the long run, energy augmenting technical change may lead to capital and labour being either drawn into or away from the sector, thereby either expanding or contracting output further.[67]

Saunders works through the relevant calculus to determine the fuel conserving conditions for four general forms of production function (Leontief, Cobb-Douglas, Generalised Leontief, and the Hogan-Manne version of the nested CES[68] - see Annex 2 for definitions). The results are summarised in Table 5.1. This shows that three commonly used production functions are either always or never fuel conserving, implying that the choice of these can completely pre-determine the outcome of model simulations.[69] One interpretation of this result could be that these three production functions are unsuitable for empirical investigation of the rebound effect. An alternative interpretation is that, if suitably parameterised Cobb Douglas or Generalised Leontief production functions are considered to provide a good approximation of real-world behaviour, then backfire is a likely outcome of energy augmenting technical change.

---

[66] Saunders (2006b) is the most recent version of a working paper that has been in circulation for a number of years.

[67] Depending upon the value of $\left.\frac{\partial}{\partial \upsilon_E}\right|\left[\frac{\partial Y}{\partial K}\right]$ and $\left.\frac{\partial}{\partial \upsilon_E}\right|\left[\frac{\partial Y}{\partial L}\right]$

[68] Namely, a Cobb Douglas production function for capital and labour inputs, nested within a CES function - *(KL)E*.

[69] This is unfortunate in the case of the Generalized Leontief, since it allows for considerable flexibility in substitution between inputs and is relatively easy to estimate.

Table 5.1 Fuel conserving conditions for common forms of production function

Production Function	Fuel Conserving Condition
Leontief	*Always Conserving*
Cobb-Douglas	*Never Conserving*
Generalised Leontief	*Never Conserving*
CES	$\sigma < 1 - s_E$

Notes:

- σ is the Hicks elasticity of substitution between energy and non-energy inputs ,
- $s_E = (P_E E / P_Y Y)$ is share of energy costs in the value of output,

Source: Saunders (2007)

The particular form of CES function examined by Saunders is found to be fuel conserving if the Hicks elasticity of substitution between effective energy and effective non-energy inputs is less than (1-s_E), where s_E represents the share of energy costs in the total value of output. Since the latter is normally small, the behaviour of this function largely depends upon the magnitude of this particular elasticity of substitution compared to unity - as suggested by Saunders (1992). The 'nesting structure' for the Hogan-Manne CES function is commonly represented as *(KL)E* and Saunders finds that the two alternative 'nesting structures' - *(KE)L* and *(LE)K* - invariably lead to backfire. This leads Saunders to comment that the choice of *(KL)*E to depict reality 'leaves one with an uncomfortable feeling of arbitrariness'. Saunders does not explore alternative representations of the nested CES, although these are widely used within energy-economic models.[70]

Empirical research tends to use cost functions rather than production functions and these generally take a more 'flexible' form than the production functions analysed above. Saunders (2007) therefore analyses the behaviour of five general types of cost function, namely the Hogan-Manne CES, the Translog, the Symmetric Generalized Barnett (SGB), the Symmetric Generalized McFadden (SGM) and the Gallant (Fourier). The second of these is widely used in empirical studies (see Annex 2), while the remainder are obscure and unknown to most practitioners in energy economics.

Saunders finds that the results for the Hogan-Manne CES cost function are the same as those for the corresponding production function. In contrast, the SGM can only depict backfire, while the SGB and Gallant are able to reproduce the full range of rebound effects.[71] This suggests that the latter may be suitable for empirical research into the rebound effect, but these remain largely absent not just from the rebound literature but from energy economics more generally.

The results for the Translog cost function have evolved over several versions of Saunders working paper and are important since this function is very widely used.[72] Saunders originally found that the Translog may be either fuel saving or fuel using, depending upon

---

[70] Possibilities include using a CES function rather than a Cobb-Douglas for the capital and labour composite, and extending the function to include materials inputs as well - thereby creating scope for additional nesting structures.

[71] Owing to the complexity of the calculus, Saunders uses simulations to explore the behaviour of these three functions

[72] Most modern empirical papers use the translog cost function, mainly due its greater flexibility. However, the highly flexible nature of this function is not suitable for all datasets (Berndt, 1979; Moroney, 1992a) and many analysts prefer a Leontief form for the short run or a multilevel CES for the long-run (Manne, 1990; Kemfert, 2000; Van der Zwaan, 2002).

the magnitude of the coefficient on (*lnE*)^2 in the cost function relative to the square of the value share of energy (s_E^2). But the square of the value share of energy depends upon both the own-price elasticities of capital, labour and energy *and* the elasticities of substitution between them. Hence, the fuel conserving condition for the Translog may be represented by a function (*g*):

$$g[s_K, s_L, s_E, \sigma_{KK}, \sigma_{KL}, \sigma_{KE}, \sigma_{LL}, \sigma_{LE}, \sigma_{EE})] \qquad (5.6)$$

In other words, the magnitude of *each* elasticity of substitution play a role in determining the behaviour of the Translog - in contrast to the Hogan-Manne CES, where only the elasticity between effective energy and other effective inputs appears relevant. Similar results apply to the SGB, SGM, and Gallant (Fourier) functions. This suggests that Saunders early results for the Hogan-Manne CES may have led researchers to focus inappropriately upon one particular parameter (see *Technical Report 3*).

Restrictions normally have to be imposed upon the parameter values in a Translog cost function to ensure that its behaviour is consistent with basic economic theory. In particular, the cost function must be *concave*[73] - implying that the marginal product of each input declines with increasing use of that input. In many applications, such as CGE modelling, these conditions need to be satisfied for all input combinations, but empirically estimated cost functions sometimes violate these conditions (Diewert and Wales, 1987). In the most recent version of his working paper, Saunders finds that imposing a global concavity restricition means that the Translog production function *always* leads to backfire. As a result:

> "Unless one is prepared either to surrender concavity or to embrace the belief that the real world only allows backfire, the Translog function must be taken off the table for rebound analysis." (Saunders, 2007)

However, Ryan and Wales (2000) show that if concavity is imposed *locally* at a suitably chosen reference point, the restriction may be satisfied at most all of the data points in the sample. Under these circumstances, the Translog may be able to represent different types of rebound effect for particular data sets – but only if it can be empirically verified that concavity is honoured across the domain of measurement. Hence, it is possible that the Translog may still be useful in some circumstances.

Using the identity *E=(E/Y)Y,* Saunders develops expressions for the elasticity of energy demand with respect to energy augmenting technical progress ($\eta_{\upsilon_E}(E)$), again holding other inputs fixed:

$$\eta_{\upsilon_E}(E) = \frac{\upsilon_E}{E}\frac{\partial E}{\partial \upsilon_E} = \frac{\upsilon_E}{(E/Y)}\frac{\partial (E/Y)}{\partial \upsilon_E} + \frac{\upsilon_E}{Y}\frac{\partial Y}{\partial \upsilon_E} \qquad (5.7)$$

Saunders calls the first term on the right-hand side the 'substitution' effect and the second term the 'output' effect. However, a better term for the first may be 'intensity effect', since it represents the change in the aggregate energy/output ratio as a result of the technical improvements, holding output (*Y*) fixed. The source of the change is the use of less energy inputs (*E*) to provide the same quantity of effective energy inputs ($\tilde{E}$). Since this derives

[73] A function is concave if it lies above the line between any two points, implying that: $f[(x_1 + x_2)/2] \geq f(x_1/2) + f(x_2/2)$

from a technical change in the production of effective energy, labelling this a substitution effect could potentially be misleading.

The net result of energy augmenting technical change is to increase the marginal productivity of effective energy relative to other effective inputs and thereby encourage substitution of effective energy for other effective inputs (i.e. to use more $\tilde{E}$ and less $\tilde{K}$ and $\tilde{L}$ to produce the same *Y*). Since the input augmentation terms for capital and labour remain unchanged, this also means that less 'real' capital (*K*) and labour (*L*) inputs are required to produce the same quantity of output. But Saunders definition of the short run fuel conserving condition requires that capital and labour inputs be held constant. Hence, output must increase - as represented by the second term in 5.7.[74] To visualise this in practical terms, one could imagine a situation in which an increase in 'effective energy' (or useful work) following a technology improvement is used to replace labour by re-configuring an assembly line. The freed-up labour is then re-deployed elsewhere to increase overall output, perhaps further increasing the consumption of effective energy as a result. The net impact on real energy consumption depends upon the characteristics of the production process, as represented by the production function.

The results provide greater insight into the behaviour of each function. For example, with the Leontief production function, the energy/output ratio is found to decline in direct proportion to the technical improvements (i.e. intensity effect = -1), with no increase in economic output. This behaviour reproduces the simple 'engineering' view of energy efficiency improvements that fails to take into account either substitution or output effects. In contrast, with a Cobb Douglas production function, there is no change in the energy/output ratio, while output increases in proportion to the share of energy in the value of output (see Box 4.2). With a translog production function, the intensity and output effects can take on a range of positive or negative values depending upon the magnitude and sign of the coefficient on $(lnE)^2$ and the value share of energy. What Saunders' results imply is that the choice can have a major impact on the results.

Since capital and labour inputs are not fixed in the long-run, the long-run rebound effect may be greater than suggested by the expressions in Table 5.1. This is because energy augmenting technical progress also increases the marginal productivity of capital and labour, thereby causing capital and labour to be drawn in to the affected sectors at the expense of others.[75] If the production efficiency of other sectors improves at the same time, the net result needs to be assessed with the aid of a CGE model. However, Saunders develops simplified estimates of the long-run rebound effect for each function under the assumption that labour is fixed and the cost of capital is fixed, but capital is mobile.[76] The results demonstrate that, even in the long-run, the choice of a functional form predetermines the amount of rebound expected. Only the SGB and Gallant functions appear

---

[74] In other words, the substitution and output effects relate to the behaviour of production function $\tilde{F}$, following a change in the energy augmenting multiplier (υ_E) and subject to the constraint that real capital and labour inputs remain fixed.

[75] i.e. $\frac{\partial}{\partial \upsilon_E}\left\|\left[\frac{\partial Y}{\partial K}\right]\right. > 0$ and $\frac{\partial}{\partial \upsilon_E}\left\|\left[\frac{\partial Y}{\partial L}\right]\right. > 0$

[76] This assumption derives from the neoclassical growth model, where in equilibrium the long-run growth rate is fixed and equal to the cost of capital.

flexible enough to honour the fuel conserving condition, under a wider set of standard assumptions.

While the above arguments may seem abstract and technical, the implications are important. Saunders has demonstrated that *most of the production and cost functions used within theoretical and empirical research are useless for investigating the rebound effect*. Most of these functions predict that energy augmenting technical change will lead to backfire. The only exceptions are SGB and Gallant (Fourier) which are rarely used and the Hogan-Manne version of the nested CES which appears both arbitrary and restrictive. While the Translog costs function may still be useful in some circumstances, few researchers have conducted the appropriate tests to see whether this is the case.

This points to two possibilities. If empirically estimated neoclassical production and cost functions are considered to provide a reasonable representation of real-world behaviour, Saunders' work suggests that energy augmenting technical change is likely to lead to backfire. Alternatively, if rebound effects vary widely in magnitude between different sectors, such functions cannot be used to represent them.

5.3.4 Technology simulations

Having established short and long-run fuel conserving conditions for various forms of production function, Saunders (2005) takes the natural step of incorporating econometric estimates of the parameters of such functions. Specifically, he develops a spreadsheet tool, named CECANT[77], for exploring how individual energy efficient technologies may potentially affect energy consumption at either the sector or economy-wide level. The spreadsheet model requires the user to input the parameters of an econometrically estimated production or cost function, such as those provided for US manufacturing sectors by Jorgenson and Fraumeni (1981a) (see Section 4.4). It also requires the user to specify the performance of the assumed technology in terms of the input augmenting multipliers indicated above. For example, an individual technology could be assumed to improve the productivity of capital, labour and material inputs by 5% ($\upsilon_K = \upsilon_L = \upsilon_M = 1.05$) and that of energy inputs by 20% ($\upsilon_E = 1.2$). These performance assumptions may need to be weighted to reflect the assumed contribution of the new technology to the value of output of an individual sector or the whole economy.

Using a similar approach to Saunders (2007), the spreadsheet calculates the short run rebound effect from the introduction of the assumed technology, under the assumption that capital and labour inputs remain fixed, together with energy prices. By assuming that the consumer utility function takes a Cobb Douglas form, the own-price elasticity of the demand for output from the relevant sector is effectively constrained to be minus one: implying that the percentage increase in output following the productivity improvement will equal the percentage reduction in production costs.

The CECANT model approach has the attractive feature that the total change in fuel consumption from introducing a technology can be subdivided into the individual changes resulting from improvements in the productivity of each input ($\Delta\upsilon_i$). Using the translog cost functions estimated by Jorgenson and Fraumeni (1981a), Saunders provides an illustrative example for introducing the technology indicated above into five US manufacturing sectors.

[77] Denoting 'Calculator for Energy Consumption Changes Arising from New Technologies'.

Each individual technology multiplier, including the energy augmenting multiplier, is shown to increase overall energy consumption in each sector.[78] The estimated rebound effect resulting from the improvement in energy productivity *alone* ranges from 106% in agriculture to 207% in communications. However, the estimated *total* rebound effect from introducing the technology (i.e. allowing for the productivity improvements in all inputs) ranges from 174% in agriculture to as much as 862% in communications. Hence, this example graphically suggests that technologies that improve the productivity of inputs other than energy may potentially lead to very large rebound effects. And since the energy efficiency characteristics of a new technology cannot be separated from its productivity-enhancing effects on other inputs, it is the rebound effect from the technology as a whole that matters

The model also uses a number of simplifying assumptions to estimate the long-term effect of a technology, when capital and labour inputs are no longer assumed to be fixed. This suggests that the short-term rebound effects estimated above may provide a lower bound to the long term effect.

The CECANT model was developed prior to the most recent version of Saunders' paper on fuel conserving production functions. The findings summarised above regarding concavity restrictions with translog functions therefore have some important implications for the use of this model. In particular, Jorgenson and Fraumeni (1981a) imposed global concavity restrictions when estimating their function. In these circumstances, it is perhaps not surprising that the CECANT model predicts backfire. The model could potentially be used with empirically estimated SGB, Gallant (Fourier) or Hogan-Manne CES cost functions or with Translog functions if they can be shown to be concave over the domain of interest. However, to date Saunders has not investigated this.

5.4 Limitations of Saunders' approach

Saunders acknowledges a variety of limitations to his theoretical approach and suggests that all results require 'cautious interpretation'. These limitations include some which are generic to neoclassical production theory, some which are relevant to neoclassical growth theory alone and some which are solely relevant to Saunders' use of that theory. Several of the more important limitations will be highlighted here, although none will be treated in any depth (i.e. this is a cursory skim over a vast literature). The treatment of technical change in CES production functions will also be briefly discussed (an in-depth treatment of this important issue is provided in Annex 3).

5.4.1 Generic limitations of the neoclassical growth model

Given the 'heroic' assumptions which underlie it, it is not surprising that the neo-classical growth model has some important and well-known limitations. One of the most important is the implicit assumption that capital can be costlessly transferred from one use to another. This implies that "….a certain tonnage of steel which has been constructed into a machine of a given sort…..can at a moments notice and without cost be re-moulded into another form of machine" (Jones, 1975). While this may be a useful theoretical simplification for some purposes, it also appears to bypass most of the real problems of an economy.

[78] Improvements in the productivity of capital, labour or materials inputs also increase the marginal productivity of energy. Energy consumption must then rise to preserve equality between marginal productivity and energy price.

A second difficulty is the time required to adjust to the steady state. As Sato (1966) has demonstrated, the implied adjustment time could be very long, even following a relatively modest change in the savings ratio. It is important to note that the Solow model was originally developed to address the macro-economy over the very long-run. Simulations in the model that assume improvements in energy efficiency throughout the economy fail to account for the interactions between sectors and the structural changes that may be expected to result. Therefore simulations over a period of 100 years (as in Saunders (1992)) are at best only very indicative of the potential growth rates of output, capital and energy.

The single sector neoclassical model also works with an all-purpose investment and consumption good. This highly abstract approach completely obscures all the distinctions between energy use in production and consumption and between different types of technology that are relevant to real-world investigations of energy demand (Howarth, 1997). Wei (2007) shows a possible way forward by developing a two sector neoclassical growth model that distinguishes between energy and non-energy goods and between productivity improvements in energy production and in energy consumption. This model suggests that the long-term impact of productivity improvements in energy production may be very much greater than productivity improvements in energy use. But while intuitively plausible, the high level of abstraction remains.

There are also major inconsistencies between the predictions of relative income shares from the neoclassical growth model and real-world data. For example, the neoclassical model predicts the capital share (from estimates of its marginal product) to be around 60%, yet observed capital shares are around 25-35% (Mankiw*, et al.*, 1992). Capital (labour) appears to be much more (less) important for growth than the Solow model would suggest, in part because differences in the quality of labour inputs are neglected.

Other microeconomic assumptions of the neoclassical growth model further restrict its relevance to the appraisal of actual energy efficiency investments. For instance, energy prices are held fixed, consumers' utility is represented through a single function (usually Cobb-Douglas), technology always comes free, there are only constant returns to scale in production, there is always full employment, qualitative differences in capital and energy are ignored, environmental externalities are neglected and so on. Such assumptions may be less important for theoretical growth economics, but they may be crucial for investigating the impact of specific energy efficiency improvements.

Perhaps the most serious difficulty with the neoclassical growth model lies with the very notion of an aggregate production function. A considerable literature challenges the idea that this concept is meaningful at all. For example, it is relatively easy to show that if two sectors each have a Cobb Douglas production function , and if the exponents on the inputs differ between these two sectors, there cannot be a Cobb Douglas aggregate production function (Temple, 2006). Fisher (1993) and others have demonstrated that the conditions for successful aggregation of micro production functions into an aggregate production function are extremely stringent. Temple (2006) suggests that a reasonable interpretation of this work is that aggregate production functions (and associated notions such as the aggregate elasticity of substitution) cannot be meaningfully defined in any circumstances that might apply to real-world economy. This does not appear to have stopped economists from using aggregate production functions or from engaging in debates over their appropriate functional form. Moreover, this criticism is not confined to the neoclassical growth model itself: the aggregation of 'sub-production' functions is also a central feature of

CGE modelling, for example. This is therefore a profound and far-reaching criticism of neoclassical theory that raises particularly serious concerns over the practical usefulness of results from both neoclassical growth and CGE models.

It is important to note that most of the above criticisms are less applicable to be more disaggregated approach adopted in Saunders (2007) and Saunders (2005). In particular, by incorporating empirical estimates of cost functions for individual sectors, Saunders (2005) gets much closer to linking theoretical propositions to empirical data. Moreover, by recommending the use of more flexible functional forms, Saunders avoids some of the difficulties associated with the Cobb Douglas and CES functions which are more commonly used in CGE models. At the time, Saunders demonstration of the propensity of such functions to backfire raises questions over the extent to which the results from CGE models may actually be 'hardwired' into the functional forms employed.

5.4.2 Elasticities of substitution

Saunders' results for the Hogan-Manne CES production function suggest that the Hicks elasticity of substitution between effective energy and a composite of (effective) non-energy inputs (σ) could influence the magnitude of the rebound effect. Put simply, a high elasticity of substitution could lead to a large rebound effect, while a small elasticity of substitution could lead to a small rebound effect. This contrasts with the widely known result that a high elasticity of substitution between energy and other inputs could decrease the cost of reducing carbon emissions (Hogan and Manne, 1977). As Saunders (2000b) has pointed out, this suggests a potential trade-off in climate policy:

> "…If one believes σ is low, one worries less about rebound and should incline towards programmes aimed at creating new fuel efficient technologies. With low σ carbon taxes are less effective in achieving a given reduction in fuel use and would prove more costly to the economy. In contrast, if one believes σ is high, one worries more about rebound and should incline towards programmes aimed at reducing fuel use via taxes. With high σ, carbon taxes have more of an effect at lower cost to the economy." (Saunders, 2000b)

Saunders' theoretical result for the Hogan-Manne CES relates to the elasticity of substitution between *effective* energy inputs and *effective* non-energy inputs (i.e. a composite of labour and capital inputs) in the aggregate production function. In contrast, econometric studies typically measure the elasticity of substitution between individual pairs of *real* inputs (e.g. between energy and capital, or between energy and labour). While related, these are not the same thing.

Saunders response to Howarth (1997) also draws attention to the potential importance of the elasticity of substitution between energy and non-energy inputs (σ). (Saunders, 2000a). While Howarth's use of Leontief production function for effective energy inputs (with $\sigma = 0$) suggests a relatively small rebound, Saunders use of an alternative Cobb Douglas production function (with $\sigma = 1$) turns this into backfire (Saunders, 2000b). In this case, however, σ relates to the production function for effective energy, rather than the aggregate production function for the sector or economy. It is therefore a *different* parameter to that discussed above.

Despite the definitional issues, these results do suggest that empirical estimates of the elasticity of substitution between energy and non-energy inputs at different levels of aggregation could provide some information on the likely magnitude of rebound effects. This assumption forms the basis of *Technical Report 3*, which examines the empirical evidence for the elasticity of substitution between energy and capital (σ_{EK}) in some detail. The Report concludes that the relationship between empirical estimates of elasticities of substitution and the magnitude of rebound effects is more complex than is generally assumed. Saunders' statement that "…the ease with which fuel can substitute for other factors of production (such as capital and labour) has a strong influence on how much rebound will be experienced" is potentially misleading. A more precise statement would:

- refer to 'effective energy' inputs ($\upsilon_E E$) rather than fuel;
- clarify that the elasticity in question is the Hicks elasticity of substitution, rather than alternative measures such as the Allen-Uzawa elasticity of substitution that are commonly used in empirical work;
- clarify that the measure relates to substitution between effective energy and a *composite* of capital and labour inputs, rather than each input individually;
- Include the qualification that this only applies when effective energy is 'separable'[79] from this composite, meaning that that the marginal rate of technical substitution between capital and labour is unaffected by the price of energy; and
- clarify that this conclusion derives from the Hogan-Manne version of the nested CES production function, and therefore does not necessarily hold for other production and cost functions.

As shown in *Technical Report 3*, the majority of empirical studies use Translog rather than CES cost functions; measure energy rather than effective energy; do not test for or impose any seperability restrictions; estimate Allen Uzawa rather than Hicks elasticities of substitution; and measure elasticities of substitution between pairs of inputs, rather than between energy and a composite of inputs. As a result, they do not provide a direct test of Saunders proposition.

In addition to these reservations, Saunders more recent work on fuel conserving production function has shown that the magnitude of the elasticities of substitution between *each pair* of inputs may play an important role in determining the magnitude of any rebound effects. But not only does this describe a more complex situation than suggested by the above quote, it also suggests that an empirical finding that energy is a 'weak substitute' for another factor (i.e. σ <1) is not necessarily inconsistent with the potential for large rebound effects. This is entirely consistent with Berndt and Wood's (1979) explanation of how energy and capital may be *complements* rather than substitutes. Although not previously recognised as such, this explanation effectively describes how an energy efficiency improvement stimulated by an investment credit may lead to backfire (see *Technical Report 3* for an explanation).

---

[79] Separability of inputs in production functions is commonly used within production theory to justify the grouping, or *nesting*, of different inputs. The assumption is that producers in engage in a two-stage decision process: first optimising the combination of inputs within each nest, and then optimising the combination of nests required to produce the final output. Two factors may only be legitimately grouped within a nest if they are separable from factors outside of the nest. For example, a *(KL)E* nesting structure requires that capital and labour are 'separable' from energy. This means that the marginal rate of technical substitution between capital and labour is unaffected by the price of energy. In practice, the assumption of seperability is frequently not supported by empirical data.

These conclusions suggest the discussion about elasticities of substitution may have obscured the real issue, which is the own-price elasticity of energy services in different contexts. While these are determined by elasticities of substitution, the relationship is far from straightforward. Also, the discussion regarding substitution elasticities may have obscured the important point that rebound effects are *also* determined by the price elasticity of output in the sector in which the energy efficiency improvement is achieved.

The implications of Hicks elasticities of substitution for the use of CES production functions within neoclassical growth models are discussed further in Annex 3. It should be noted that almost all theoretical models of economic growth (and most energy-economy integrated assessment models) use a value for the Hicks elasticity of substitution between energy and non-energy inputs that is *less* than unity - so that energy and non-energy inputs are 'weak' substitutes. This common assumption has some support from empirical studies. For example, Hogan and Manne show that, under certain assumptions, the long-run value of the Hicks elasticity of substitution between energy and non-energy inputs is broadly comparable to negative of the own-price elasticity of energy demand (Hogan and Manne, 1977) – which normally considered to be less than unity. Similarly, Kemfert (1998) and Kemfert and Welsch (2000) estimate elasticities of substitution between energy and non-energy inputs using a similar functional form to that employed by Saunders and reach similar conclusions regarding the value of this parameter – which is again less than unity. This suggests that the exception to backfire that Saunders found with a CES production function could in fact be more representative of reality.

5.4.3 Technical change

Another critical assumption of the neoclassical approach is how factor specific and neutral technical progress is included. There is no universally accepted way of doing this, and different types of energy-economic models use a range of approaches according to objectives, level of aggregation, theoretical assumptions and data availability (Löschel, 2002). As discussed in Section 4.4, technical progress is usually separated into exogenous and endogenous and into disembodied and embodied. The neoclassical growth model assumes technical progress to be exogenous and disembodied and models this through an exponential function of time that is assumed to capture all the costless, non-price-induced changes that reduce the energy/output ratio. In this, it has similarities with earlier approaches to energy-economic modelling which use comparable parameters to represent the *AEEI*.

But in practice, technical change is also embodied and endogenous and influenced by investment patterns, policy interventions and changes in relative prices. A simple parameter to represent the *AEEI* is therefore an inadequate basis for policy analysis (Azar and Dowlatabadi, 1999). Both modern growth theory and more recent approaches in energy-economic modelling seek to endogenise technical change. In endogenous growth models, technology does not just 'appear', but is generated from within the model. Energy efficiency is not "autonomous" but induced through investment in R&D, while knowledge is assumed to accumulate. Changes in relative prices affect investment in R&D and processes such as learning-by-doing, spillovers, innovation diffusions are directly modeled.

In a rare example, Smulders and de Nooj (2003) model the impact of energy efficiency improvements with an endogenous growth model. In this paper, induced energy innovation, deriving either from high-prices, supply shortages or conservation policies (all exogenously given), is simulated to offs*et al*most all national income losses from reduced energy use. This example suggests that a single exogenous change in constant input augmentation growth rates, as in the traditional neoclassical model, is insufficient to capture the complex responses of technical progress to changes in various parameters. Unfortunately, endogenous growth models have yet to be used to investigate the rebound effect.

5.4.4 The representation of technology in CES production functions

While econometric studies tend to work with flexible translog production or cost functions, neoclassical growth models (and CGE models) more frequently use the nested, multi-input CES production function. This is despite the fact that the parameters of such functions are difficult to estimate empirically and the assumptions upon which they are based (notably with regard to 'separability') are generally incorrect (Frondel and Schmidt, 2004).

Annex 3 explores the behaviour of these functions in neoclassical growth models in more detail, particularly with regard to their representation of technical change. In particular, it:

- develops a proof for the condition for backfire with a nested CES production function of the Hogan-Manne form;
- clarifies the similarities and differences between variants of the CES function and compares two different approaches to representing improvements in 'energy efficiency' in these functions; and
- establishes the conditions under which a nested CES function of the Manne Richels form can model declining energy intensity (i.e. a positive *AEEI*).

Following Saunders, Annex 3 shows that, in a nested CES production function, energy augmenting technical progress $(\partial \ln \upsilon_E / \partial t \geq 0)$ can only lead to 'energy saving' ($AEEI \geq 0$) when the Hicks elasticity of substitution between effective energy and effective non-energy inputs is less than unity (i.e. effective energy and effective non-energy inputs are 'weak' substitutes). When the Hicks elasticity of substitution between effective energy and effective non-energy inputs is greater than unity (i.e. they are 'strong' substitutes), energy augmenting technical progress increases overall energy intensity ($AEEI \leq 0$). This is because energy augmenting technical progress leads to effective energy substituting for other inputs, and the resulting increase in effective energy demand is more than sufficient to offset the reduction in energy inputs required to produce one unit of effective energy. The net result is that demand for other inputs falls while demand for energy increases.

However, Annex 3 shows that energy savings *are* possible when effective energy is a weak substitute for effective non-energy inputs ($\sigma < 1$), *provided* that the energy augmentation multiplier has a negative growth rate $(\partial \ln \upsilon_E / \partial t \leq 0)$. This departs from conventional assumptions and appears odd, since it seems to imply technical 'regress' rather than progress (i.e. input productivity declining over time). However, what matters is the growth rate of the energy multiplier (υ_E) *relative to the other multipliers* (i.e. υ_L and υ_K). In practice, all the multipliers are likely to have positive growth rates, but if the growth rate of energy is *less* than the others, the *relative* productivity of energy is declining. Saunders fuel conserving condition for the Hogan-Manne version of the nested CES production function can then be re-phrased as:

A CES production technology is fuel conserving when a positive (negative) growth rate in energy augmenting technology relative to other input augmenting technologies combines with a Hicks elasticity of substitution between effective energy and effective non-energy inputs that is less than (greater than) unity.

These different outcomes derive from the fact that a reduction in the relative price of effective energy, makes effective energy either more economically attractive or less attractive with respect to other inputs, depending upon whether it is weak or strong substitute (i.e if $\sigma<1$ or $\sigma>1$). The growth model with a Hogan-Manne CES production function therefore provides two alternative ways of stimulating a situation of declining aggregate energy intensity, depending upon the appropriate values of the Hicks elasticity of substitution between effective energy and effective non-energy inputs and the way in which input augmenting technical change is specified.

5.5 Summary and implications

Saunders has provided a significant theoretical contribution to the rebound debate and has raised important questions regarding the behaviour of commonly used neoclassical production functions. While this work relies in particular upon the neo-classical growth model, it also has important implications for CGE modelling and for the econometric investigation of individual sectors. These implications do not appear to be fully appreciated in the wider energy economics community.

Saunders has shown how the predicted magnitude of the rebound effect depends almost entirely on the choice of functional form for the relevant production function –whether at the firm, sector or economy-wide level. Most of the production and cost functions used within theoretical and empirical research are effectively useless for investigating the rebound effect. Most of these functions predict that energy augmenting technical change will lead to backfire and the functions which do not are either very rarely used (SGB and Gallant) or overly restrictive (Hogan-Manne CES). While the standard 'Translog cost function' may still be useful in some circumstances, few researchers have conducted the appropriate tests to see whether this is the case. This therefore suggests two possibilities. If empirically estimated neoclassical production and cost functions are considered to provide a reasonable representation of real-world behaviour, Saunders' work suggests that energy augmenting technical change is likely to lead to backfire. Alternatively, if rebound effects vary widely in magnitude between different sectors, such functions cannot be used to represent them.

Saunders' work also highlights the potential for very large rebound effects when improvements in energy efficiency are combined with improvements in productivity of other inputs. This suggests that, if this situation is the norm, then backfire may be a more common outcome than is conventionally assumed. However, the robustness of this result must be questioned in the light of Saunders recent work on concavity restrictions within Translog cost functions (Saunders, 2007).

With the standard Hogan-Manne version of the nested CES production function, the results are highly sensitive to the value of the Hicks elasticity of substitution between energy and non-energy inputs. This result has led Saunders to suggest a possible trade-off between the size of the rebound effect and the economic impact of carbon/energy taxes. However, this conclusion appears to be oversimplified and potentially misleading. Saunders more recent work suggests that the magnitude of rebound effects may depend upon the elasticity of substitution between *each pair* of inputs. Moreover, the existing empirical work on elaticities of substitution may provide relatively little guidance on rebound effects because what is being measured is quite different from what is being assumed within theoretical models. In particular, that an empirical finding that energy is a 'weak substitute' for another factor (i.e. $\sigma < 1$) is not necessarily inconsistent with the potential for large rebound effects.

Saunders approach is entirely theoretical and therefore severely limited by the assumptions implicit in the relevant models. A key weakness is the limited capability to capture the complexities of technical progress. Neutral and input specific technology is modelled, at best, exogenously without explicit representations of endogenous processes that affect energy and output. This characteristic limits the capacity of such models to address policy relevant objectives. The technology myopia of the traditional growth model is exposed when compared with more recent developments in endogenous growth theory. Unfortunately, to date no authors appear to have used endogenous growth models to explore the rebound effect.

Overall, Saunders work suggests that significant rebound effects can exist in theory, backfire is quite likely and this result is robust to different model assumptions. Since these results are rooted in a contested theoretical framework, they are suggestive rather than definitive. But they deserve to be taken seriously.

6 Energy productivity and ecological economics

6.1 Introduction

The analysis of the preceding sections is largely based upon the assumptions and approach of neoclassical economics. While Brookes and Schurr depart from this tradition to some extent by insisting upon the importance of energy in economic growth, they do not locate this argument within a broader theoretical framework that challenges conventional assumptions. However, such a framework does exist. Beginning with the work of Georgescu-Roegen (1971), ecological economists have developed a comprehensive and coherent alternative to the neoclassical mainstream in which energy (or more precisely, exergy) plays a central role.[80] While this tradition of work has not explicitly investigated the rebound effect, it offers a number of insights that appear very relevant to it. The aim of this section is to survey and critique some of the key studies in this area and to highlight their potential relevance to the rebound effect in general and the Khazzoom-Brookes postulate in particular.

Section 6.2 introduces the ecological perspective on energy and economic growth and highlights two important hypotheses that follow directly from it and which contradict conventional neoclassical assumptions. Section 6.3 discusses the scope for substitution between energy and other factors of production and shows how, from an ecological perspective; this may be more constrained than is commonly assumed. This section also describes how the 'embodied energy' of goods, services and factors of production contributes to the indirect rebound effect and reviews empirical estimates of the magnitude of this effect in particular cases. Section 6.4 examines how improvements in the quality of energy inputs may be an important factor in driving economic growth and presents empirical results which suggest that, when such changes are taken into account, there is little evidence for energy-saving technical change at the economy-wide level.

Section 6.5 examines the evidence that increases in energy consumption can be considered a *cause* of GDP growth (rather than vice versa as is commonly assumed) and suggests that these results depend upon whether and how the quality of different energy inputs is accounted for. Section 6.6 describes a number of ecological alternatives to neoclassical and endogenous growth models and evaluates their claim to fit historical data on GDP growth extremely well, while at the same time eliminating the need for a separate multiplier for technological progress. It also discusses whether the inclusion of useful work (i.e. exergy inputs multiplied by second law conversion efficiency) as a factor of production can improve the explanatory power of such models, and whether improvements in exergy conversion efficiency can provide a suitable proxy for technical change. Finally, Section 6.7 summarises the overall implications of this work for the economy-wide rebound effect.

6.2 Ecological perspectives on economic production and growth

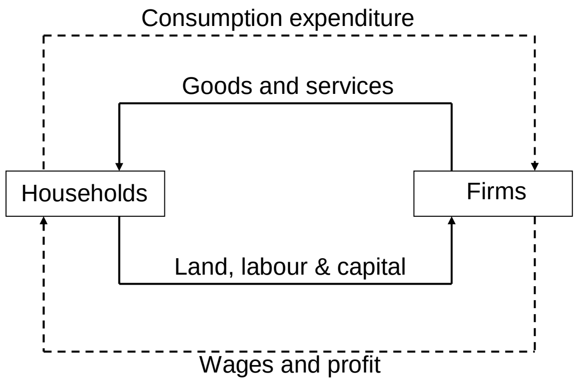
Ecological economics may be defined in a variety of ways, but a central theme is to ground economic theory and practice in physical reality, and especially in the laws of thermodynamics, the constraints imposed by nature and the contribution of 'natural

[80] The boundary between ecological economics and environmental economics is blurred and only a subset of ecological economists actively research energy issues. Those that do frequently use the term 'biophysical economics' (Cleveland, 1999). The term 'ecological economics' is therefore used fairly loosely here.

capital'[81] and associated 'ecosystem services' to human wealth and well-being (Common and Stagl, 2006). While neoclassical theory is modeled upon classical mechanics, ecological economics takes its inspiration from ecology and systems theory. And while neoclassical economics is primarily concerned with the efficient allocation of resources, ecological economics is also concerned with optimum *scale* of the economy and with the distinction between quantitative growth and qualitative development (Daly and Cobb, 1989).

Figure 6.1 illustrates the standard conceptual model of neoclassical economics, which focuses on the exchange of goods and services between households and firms. Primary attention is paid to consumer preferences, the role of technology and the conditions for market equilibrium. In this 'circular flow' model, goods and factors of production appear to flow 'endlessly' between firms and households, taking no account of natural resources, ecosystem services and the production of waste. Extensions to the model can begin to take such factors into account, but they generally remain a secondary concern - implying that they have little relevance to traditional questions such as the source and stability of economic growth. While such a model may have been appropriate when environmental resources were plentiful, it seems increasingly inappropriate at a time when human activities are exceeding the carrying capacity of the planet (Wackernagel and Rees, 1996).

Figure 6.1 Neoclassical circular flow model of economic production

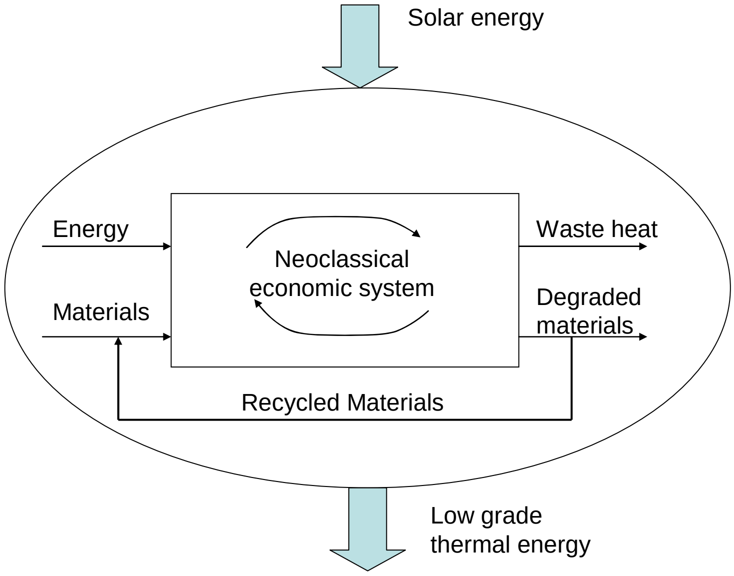


Source: Hall *et al* (1986)

In contrast, Figure 6.2 illustrates the conceptual model of ecological economics, in which the neoclassical economy is viewed as an open subsystem of the larger closed global ecosystem (Hall*, et al.*, 1986). Ecological economics views economic production as being wholly sustained by an irreversible, unidirectional flow of energy and materials from the environment, which travels through the economic system and returns back to the environment in the form of waste and low temperature heat (Cleveland and Ruth, 1997). In the terminology introduced in Section 2, the primary inputs to the economy are energy and materials with a 'high availablity', or *exergy* content, while the ultimate outputs are waste and low temperature heat, with low exergy content. The circular flow of exchange value, which forms the focus of neoclassical economics, is therefore but an intermediate step in a process powered by the unidirectional flow of energy and materials (Hall*, et al.*, 1986).

[81] Natural capital may be considered the planetary endowment of scarce matter and energy, along with the complex and biologically diverse ecosystems that provide goods and services necessary for human survival and well-being.

Figure 6.2 Ecological/biophysical model of economic production



Source: Hall *et al* (1986)

Industrial civilisation has developed through the exploitation of 'high quality' (i.e. high exergy) reserves of minerals and fossil fuels accumulated over millennia in the natural environment. As high-quality reserves are increasingly exploited, more energy is required for the extraction and processing of the relevant resources (Cleveland, 1992).[82] But given sufficient available energy, usable materials can in principle be extracted from even very low quality reserves. This suggests that the main limiting factor to economic development is likely to be the availability of (flow-limited) renewable energy, deriving ultimately from solar power (Ayres, 1998a). Under present conditions, however, it is the fragility of natural ecosystems and the services they provide that forms the greatest cause for concern.

The ecological perspective considers the standard neoclassical production function to be flawed. A key assumption of neoclassical production theory is that factors of production are substitutable, scarce, essential and *independent* inputs to economic production, implying that the availability of one input does not depend upon the use of other inputs. Prior to 1970, it was common to assume that labour services (L) and capital services (K) were the *only* independent factors of production, reflecting the fact that, in the standard presentation of national accounts, all income goes ultimately to labour (as wages) and capital (as interest, dividends, rents and royalties) (Ayres, 2001). Materials and energy inputs were not, until relatively recently, treated as independent factors of production because, in the national accounts, they appear as intermediate products. Daly is one of many to highlight the absudity of this approach:

> "….since the production function is often explained a technical recipe, we might say that Solow's recipe calls for making a cake with only the cook and his kitchen. We do not need

[82] Previous shortages of non-energy resources have generally been mitigated by new technologies that require more energy, both directly and indirectly. For example, water shortages have been mitigated by investment in energy intensive wells, pumps and pipeline networks. Similarly, the productivity of poor quality soil - or soil depleted by erosion - has been mitigated by the use of energy intensive fertilisers, pesticides and irrigation schemes (Hall*, et al.*, 1986).

> flour, eggs, sugar etc, or electricity or natural gas, or even firewood. If we want a bigger cake, a cook simply stirs faster in a bigger bowl and cooks the empty bowl in a bigger oven that somehow heats itself. Nor does the cook have any cleaning up to do, because the production recipe produces no wastes. There are no rinds, peelings, husks, shells, or residues, nor is there any waste heat from the oven to be vented. Furthermore, we can make not only a cake, but any kind of dish...without worrying about the qualitatively different ingredients, or even about the quantity of any ingredient at all! (Daly, 1997)

Since the 1970's energy crisis, energy (*E*) and materials (*M*) have routinely been incorporated into neoclassical production and cost functions - typically in value proportions that reflect the small share of the relevant intermediate goods in either the national accounts or the cost structure for individual sectors. But as Georgescu-Roegen (1971) pointed out, this does not necessarily solve the problem since such functions can, in principle, violate the second law of thermodynamics.[83] This is because many such functions implicitly embody the (physically impossible) assumption that output can be maintained with ever diminishing quantities of energy and material inputs, provided that capital and labour services can be increased sufficiently. However:

> "….. in actuality, the increase of capital implies an additional depletion of resources. And if $K \rightarrow$ infinity, then *R* will rapidly be exhausted by the production of capital. Solow and Stiglitz could not have come up with their conjuring trick had they borne in mind, first, that any material process consists in the transformation of some materials into others by some agents…..and that natural resources are the very sap of the economic process. They are not like any other production factor. A change in capital or labour can only diminish the amount of waste in the production of a commodity: no agent can create the material on which it works. Nor can capital create the stuff out of which it is made." (Georgescu-Roegen, 1971)

In contrast, ecological economists claim that *energy* (or more precisely, exergy) is the only primary factor of production, because it cannot be produced or recycled from any other factor and therefore must be supplied from outside the economic system (Hall*, et al.*, 1986). From this perspective, it is labour and capital that are the intermediate inputs, because they depend upon a net input of energy for their production and maintenance. In other words, the availability of energy is a necessary, but not sufficient, condition for the availability of labour and capital and hence for economic production and growth (Hall*, et al.*, 1986). So, far from being a secondary consideration, energy becomes the main focus of attention. Cleveland *et al.* (1984) set out a number of hypotheses that follow directly from this worldview and contrast these with conventional neoclassical assumptions. The first two are particularly relevant here:

- *Hypothesis 1*: A strong link between energy use and economic output exists and will continue to exist, both temporally and cross-sectionally. The correlation is strengthened when adjustments are made for energy quality and for the sector in which energy is used. (*Alternative hypothesis: This link can be and has been substantially decoupled, especially as the price of energy increases*.)
- *Hypothesis 2:* A large component of increased labour productivity over the past 70 years has resulted from increasing the ability of human labour to do physical work by

83 For a physical scientist such as Eddington, this is wholly unacceptable: "…..The law that entropy always increases, holds, I think, the supreme position among the laws of nature. If someone points out to you that your pet theory of the universe is in disagreement with Maxwell's equation's— then so much the worse for Maxwell's equations. If it is found to be contradicted by observation — well, these experimentalists do bungle things sometimes. But if your theory is found to be against the second law of thermodynamics I can give you no hope; there is nothing for it but to collapse in deepest humiliation." (Eddington, 1927)

empowering workers with increasing quantities of energy, both directly and indirectly as embodied in capital equipment and technology. (*Alternative hypothesis: Improvements in productivity have largely derived from exogenous technical change*).

6.3 Embodied energy and indirect rebound effects

6.3.1 Embodied energy – the limits to substitution

The differences between neoclassical and ecological perspectives centre in particular on the claimed scope for substitution between different factors of production and particularly on the scope for substitution between manufactured capital and 'natural capital'. These differences are most prominent in the debate between so-called 'weak' and 'strong' sustainability (Neumayer, 2004), but they are also relevant to the rebound effect.

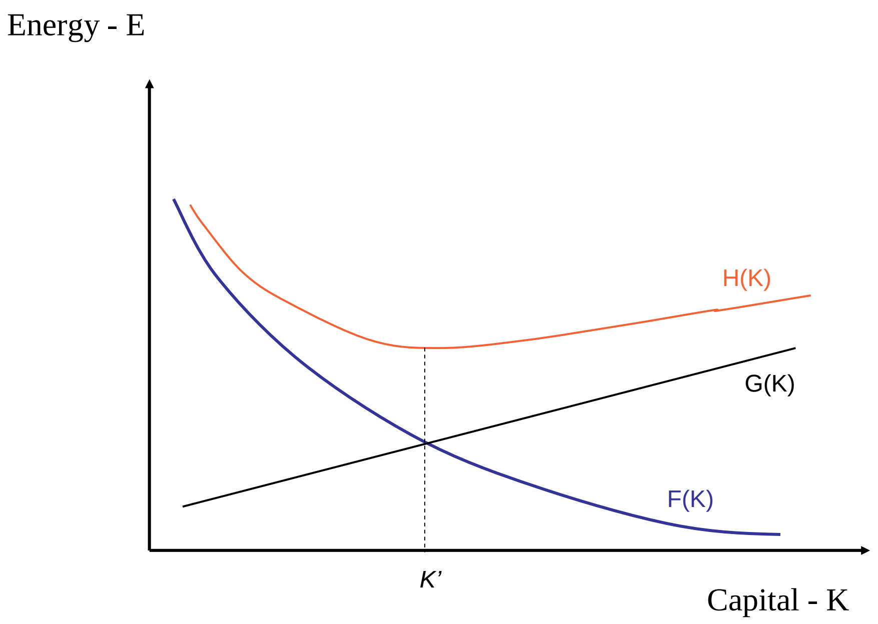
Ecological economists argue that the scope for substitution between manufactured and natural capital is very limited. One reason is that there are some essential, irreplaceable services that only natural capital can provide, such as regulation of the global climate and the maintenance of biodiversity (Cleveland and Ruth, 1997). Another is that the second law of thermodynamics imposes strict limits on the extent to which materials can be transformed into different states and the minimum amount of energy required to achieve that transformation (Stern and Cleveland, 2004). But for our purposes, the most important reason is that given by Georgescu-Roegen above, namely: *capital cannot create the stuff out of which it is made*. Natural capital is what is being transformed, while manufactured capital is what achieves the transformation. Hence, providing more of the substitute (manufactured capital) requires more of the thing that it is supposed to substitute for (natural capital). The same argument also applies to the scope for substitution between labour and natural capital, since increased labour inputs also require more natural resources (e.g. for food, shelter, housing, transport, warmth etc.).

Most improvements in energy efficiency can be understood as the substitution - within narrow system boundaries - of manufactured capital (e.g. thermal insulation) for a particular type of natural capital (energy from fossil fuels). It is these possibilities that are the main focus of energy efficiency programmes and which form the basis of estimates of the potential for improved energy efficiency in different sectors. These possibilities are also reflected in the econometric estimates of the elasticity of substitution between energy and capital that are reviewed in *Technical Report 3*. However, such estimates may not reflect the possibility for substitution between energy and capital within *wider* system boundaries, such as the economy as a whole. This is because they do not include the *indirect* energy consumption that is required to produce and maintain the relevant capital. For example, energy is required to produce and install home insulation materials and energy efficient motors. This is one reason why the net energy savings for the whole economy may be less than indicated by an analysis of individual energy efficiency opportunities - even when direct rebound effects are zero:

> "From an ecological perspective, substituting capital and/or labour for energy shifts energy use from the sector in which it is used to sectors of the economy that produce and support capital and/or labour. In other words, substituting capital and/or labour for energy increases energy use elsewhere in the economy" (Kaufmann and Azary-Lee, 1990)

Stern (1997) provides a useful graphical interpretation of this process (Figure 6.3). Here, *E=f(K)* is an isoquant of a neoclassical production function representing a different combinations of energy (*E*) and capital (*K*) that may be used to provide a given level of output for a particular firm or sector. But this capital also has indirect energy consumption associated with it (elsewhere in the economy), represented by the function *g(K)*. The summation of the two gives the 'net' isoquant *E=h(K)* for the economy as a whole. It can be seen that: first, the net energy savings from the substitution of capital for energy are less than implied by the neoclassical production function alone (*h(K)>f(K)*); and second, when capital inputs exceed a certain level (*K'*), the indirect energy consumption *exceeds* the direct energy savings – leading to backfire for the economy as a whole, even when the individual sector reduces energy consumption and when output from this sector is unchanged.

Figure 6.3 Indirect energy consumption and the limits to substitution



While this example relates to substitution between capital and energy in production, entirely analogous effects follow for substitution between capital and energy in consumption: for example, energy efficient refrigerators also require energy for their production.

In practice, backfire from this source alone (i.e. $G(K) > F(K)$) appears rather unlikely, since the cost of an energy efficient technology should reflect the cost of the embodied energy (Webb and Pearce, 1975). If the latter exceeds the saving in energy costs, it is unlikely that the investment would be cost-effective.[84] However, this assumes that the sole benefit of the investment is the reduced energy costs, which may not always be the case. Also, market imperfections may distort the relevant prices and costs.

In contrast to other sources of the economy-wide rebound effect, the contribution from this source may be expected to be smaller in the long-term than in the short-term. This is

[84] Note that this argument applies to measures of energy consumption weighted by the relative price of different energy carriers and not to energy consumption measured simply in terms of heat content.

because the embodied energy associated with capital equipment is analogous to a capital cost and hence diminishes in importance relative to ongoing energy savings as the lifetime of the investment increases.

Some authors argue that similar conclusions apply to the substitution of labour for energy, since energy is also required to feed and house workers and thereby keep them economically productive (Kaufmann, 1992). However, there is some dispute over whether and how to account for the 'energy cost of labour' (Costanza, 1980).[85] Similarly, while economists conventionally distinguish between substitution and technical change (Figure 3.1), the latter is also associated with indirect energy consumption since it is embodied in capital goods and skilled workers (Stern and Cleveland, 2004):

> "The arguments for technological change as a solution would be more convincing if technological change was really something different from substitution. The neoclassical approach assumes that an infinite number of efficient techniques coexist at any one point in time. Substitution occurs among these techniques. Changes in technology occur when new, more efficient techniques are developed. However, in a sense, these new techniques represent the substitution of knowledge for the other factors of production. The knowledge is embodied in improved capital goods and more skilled workers and managers, all of which require energy, materials and ecosystem services to produce and maintain. Thus, however sophisticated the workers and machinery become, there are still thermodynamic restrictions on the extent to which energy and material flows can be reduced" (Stern and Cleveland, 2004)

In principle, the implications of particular types of substitution should be assessed in a dynamic perspective (i.e. with an explicit time dimension) and taking into account the relevant alternatives - such as the indirect energy consumption associated with a non energy efficient refrigerator (thereby isolating the additional capital required for substitution). As an example, a mandatory requirement to replace existing refrigerators with more energy efficient models may either increase or decrease aggregate energy consumption over a particular period of time, depending upon the age of the existing stock, the lifetime of the new stock, and the direct and indirect energy consumption of different models of refrigerator. In practice, however, such estimates appear to be rare, with most analysts focusing instead upon the 'energy return on energy invested'[86] for different energy supply options (Cleveland, 1992).[87]

---

[85] Costanza (1980) estimates the energy cost of labour as the energy associated with all personal consumption expenditures. These energy costs are then assigned to individual goods and services in proportion to the labour required to produce them. Double counting is avoided by changing the boundaries of the traditional economic input-output analysis. The net result is to greatly increase the embodied energy estimated to be associated with labour-intensive goods and services. With this approach, the 'embodied energy intensity' of most sectors (excluding primary energy) is found to be broadly comparable. However, this conclusion depends entirely upon the particular methodology for calculating the energy cost of labour. This is quite different from conventional accounting approaches, which estimate much lower energy intensities for many sectors and greater variation between them. It also implicitly assumes that all personal consumption expenditures are necessary to support labour, which appears unjustified.

[86] The Energy Returned on Energy Invested (EroEI) is the ratio of the usable energy acquired from a particular energy resource to the amount of energy expended to obtain that energy resource. In principle, when the EROEI of a resource is equal to or lower than 1, that energy source can no longer be used as a primary source of energy. However, this measure neglects the relative economic productivity of different energy forms. When this is taken into account, resources with an EroEI of less than unity may still be economic to extract.

[87] Studies based upon embodied energy have fallen out of favour since the 1980s, when they were often associated with somewhat controversial 'energy theories of value' (Söllner, 1997). But there is no necessary link between these theories and use of 'embodied energy' estimates in empirical research. Consideration of rebound effects may provide a motivation for reviving this area of research.

The limits to substitution described by ecological economists form one component of the indirect rebound effect (Box 2.1 and Box 2.2). Indirect rebound effects derive from two sources: the energy required to produce and install the measures that improve energy efficiency, such as thermal insulation, and the indirect energy consumption that results from such improvements. The first of these relates to energy consumption that occurs prior to the energy efficiency improvement, while the second relates to energy consumption that follows the improvement.

However, the contribution of the first of these (i.e. the limits to substitution) to the economy-wide rebound effect is frequently overlooked. In the case of household consumption, for example, the indirect rebound effect is usually equated to energy content of the *other* goods and services that are purchased with the money saved from the lower energy bills following an energy efficiency improvement - the so-called *re-spending* effect. But a full accounting of indirect rebound effects also requires the embodied energy of the energy efficient equipment to be taken into account - relative, if appropriate, to that of inefficient equipment. For example, the lightweight materials required for fuel-efficient vehicles may be more energy intensive to produce than steel.

The following two sections examine the limited empirical evidence for indirect rebound effects, focusing in particular on studies that estimate the embodied energy associated with different categories of consumer goods and services. Despite the apparent potential of this approach, there appear to be very few applications to the rebound effect.

6.3.2 Evidence for limits to substitution

Some indication of the importance of embodied energy may be obtained from estimates of the own-price elasticity of *aggregate* primary, secondary or final energy demand. In principle, this measures the scope for substituting capital, labour and materials for energy, while holding output constant. Most energy price elasticities are estimated at the level of individual sectors and therefore do not reflect all the embodied energy associated with capital, labour and materials inputs. Since the own-price elasticity of aggregate energy reflects this indirect energy consumption, it should in principle be smaller than a weighted average of energy demand elasticities within each sector. However, the aggregate elasticity may also reflect price-induced changes in economic structure and product mix which in principle could make it larger than the average of sectoral elasticities (Sweeney, 1984). These two mechanisms therefore act in opposition.

Based in part upon modelling studies, Sweeney (1984) puts the long-run elasticity of demand for primary energy in the range -0.25 to -0.6. In contrast, Kaufmann (1992) uses econometric analysis to propose a range from -0.05 to -0.39, while Hong (1983) estimates a value of -0.05 for the US economy. A low value for this elasticity may indicate a limited scope for substitution and hence the potential for large indirect rebound effects.[88] But this interpretation is not straightforward, since direct rebound effects also contribute to the behaviour being measured. Also, measures of the quantity and price of 'aggregate energy' are sensitive to the methods chosen for aggregating the prices and quantities of individual energy carriers, while the price elasticity will also depend upon the particular composition of price changes (e.g. increases in oil prices relative to gas) (EMF 4 Working Group, 1981). In particular, when different energy types are weighted by their relative marginal productivity,

[88] This is in contrast to the own-price elasticity of energy demand for an individual energy service, where high values may indicate the potential for large direct rebound effects.

the estimated elasticities tend to be lower (Hong, 1983) As a result, the available estimates of aggregate price elasticities may be insufficiently precise to provide much indication of the magnitude of indirect rebound effects.

Relatively few empirical studies have investigated the embodied energy associated with specific energy efficiency improvements and those that have appear to focus disproportionately upon domestic buildings. In a rare study of energy efficiency improvements by producers, Kaufmann and Azary Lee (1990) estimate that, in the US forest products industry over the period 1954 to 1984, the embodied energy associated with capital equipment offset the direct energy savings from that equipment by as much as 83% (Box 4.1). But since their methodology is crude and the results specific to the US context, this study provides little indication of the magnitude of these effects more generally.

Box 6.1 Limits to substitution for producers

Kaufmann and Azary Lee (1990) examined the embodied energy associated with energy efficiency improvements in the US forest products industry over the period 1954 to 1984. First, they estimated a production function for the output of this industry and used this to derive the 'marginal rate of technical substitution' (MRTS) between capital and energy in a given year - in other words, the amount of gross fixed capital that was used to substitute for a thermal unit of energy in that year. Second, they approximated the embodied energy associated with that capital by means of the aggregate energy/GDP ratio for the US economy in that year - hence ignoring the particular type of capital used, as well as the difference between the energy intensity of the capital producing sectors and that of the economy as a whole. The product of these two variables gave an estimate of the indirect energy consumption associated with the gross capital stock used to substitute for a unit of energy. This was then multiplied by a depreciation rate to give the energy associated with the capital services used to substitute for a unit of energy.

Finally, they compared the estimated indirect energy consumption with the direct energy savings in the forest products sector in each year. Their results showed that the indirect energy consumption of capital offset the direct savings by between 18 and 83% over the period in question, with the net energy savings generally decreasing over time. The primary source of the variation was the increase in the MRTS over time, implying that an increasing amount of capital was being used to substitute for a unit of energy. However, the results were also influenced by the high energy/GDP ratio of the US economy, which is approximately twice that of many European countries. Overall, the calculations suggest that the substitution reduced aggregate US energy consumption, but by much less than a sector-based analysis would suggest. Also, their approach did not take into account any secondary effects resulting from the energy efficiency improvements.

The simplicity of this approach suggests the scope for further development and wider application. Accuracy could be considerably improved by the use of more flexible production functions and more precise estimates for the indirect energy consumption associated with specific types and vintages of capital goods. However, to date no other authors appear to have applied this approach to particular industrial sectors or to have related it to the broader debate on the rebound effect.

Estimates of the embodied energy of different categories of goods and services can be obtained from input-output analysis, life-cycle analysis (LCA) or a combination of the two (Chapman, 1974; Herendeen and Tanak, 1976; Kok*, et al.*, 2006). A full life-cycle analysis is time consuming to conduct and must address problems of 'truncation' (i.e. uncertainty over the appropriate system boundary)[89] and joint production (i.e. how to attribute energy consumption to two or more products from a single sector) (Leach, 1975; Lenzen and Dey, 2000). Hence, many studies combine standard economic input-output tables with additional information on the energy consumption of individual sectors, to give a comprehensive and reasonably accurate representation of the direct and indirect energy required to produce rather aggregate categories of goods and services. More detailed, LCA-based estimates are available for individual products such as building materials, but results vary widely from one context to another depending upon factors such as the fuel mix for primary energy supply (Sartori and Hestnes, 2007).

As an illustration, Sartori and Hestnes (2007) reviewed 60 case studies of buildings, and found that the share of embodied energy in life-cycle energy consumption ranged between 9 and 46% for low energy buildings and between 2 and 38% for conventional buildings – with the wide range reflecting different building types, material choices and climatic conditions. Two studies that controlled for these variables found that low energy designs could achieve substantial reductions in operating energy consumption with relatively small increases in embodied energy, leading to 'payback periods' for energy saving of as little as one year (Feist, 1996; Winther and Hestnes, 1999). Similar calculations were performed by the Royal Commission on Environmental Pollution (2007), who estimate a 15 year simple payback (in energy terms) for low energy new build houses in the UK. However, Casals (2006) shows how the embodied energy of such buildings could offset operational energy savings, even with an assumed 100 year lifetime. Such calculations typically neglect differences in energy quality and the results are sensitive to context, design and building type. However, the increasing availability of embodied energy coefficients at a national level (e.g. Alcorn and Baird (1996)) suggests the scope for greater use of such estimates in policy evaluation.

In the case of existing buildings, several studies suggest that retrofits of thermal insulation pay for themselves in terms of energy savings within a few months (compared to a useful life in excess of 25 years), while the corresponding period for double glazing is several years. In other cases, for example condensing boilers compared to conventional boilers, the variation of embodied energy within individual categories of boiler exceeds the difference between them. Hence, the contribution of embodied energy to the economy-wide rebound effect appears to vary widely from one situation to another and is inversely proportional to the lifetime of the energy saving measure. But the patchy nature of this evidence base, the lack of systematic comparisons of energy efficiency options and the dependence of the results on particular contexts all make it difficult to draw any general conclusions.

6.3.3 Evidence for secondary effects

By combining estimates of the embodied energy associated with different categories of goods and services with survey data on household consumption patterns, it is possible to estimate the total (direct plus indirect) energy consumption of different types of household; together with the indirect energy consumption associated with particular categories of expenditure (Kok, *et al.*, 2006). It is often found that the indirect energy consumption of households exceeds the direct consumption. Moreover, while indirect energy consumption

[89] For example, should the indirect energy costs of a building also include the energy used to make the structural steel and mine the iron ore used to make the girders? This is referred to as the truncation problem because there is no standard procedure for determining when energy costs become small enough to neglect

increases with income, direct energy consumption shows signs of saturation - suggesting that indirect energy consumption is becoming increasingly important over time.[90]

If this data is available at a sufficiently disaggregated level, it could also be used to estimate the secondary effects associated with energy efficiency improvements by households (Box 2.1 and Box 2.2) - provided that additional information is available on either the cross price elasticity between different product and service categories, or the marginal propensity to spend[91] of different income groups. By combining the estimates of embodied energy and secondary effects, an estimate of the total indirect rebound effect may be obtained. Such approaches are 'static in that they do not capture the full range of price and quantity adjustments, but could nevertheless be informative.[92] However, of the 19 studies in this area reviewed by Kok*, et al.* (2006), only three were considered to have sufficient detail to allow the investigation of such micro-level changes – largely because they combined input-output with LCA data (Bullard*, et al.*, 1978). Hence, estimation of secondary effects by this route appears to be in its infancy.

Three studies that use this general approach are summarised here. First, Brännlund *et al.* (2007) examine the effect of a 20% improvement in the energy efficiency of personal transport (all modes) and space heating in Sweden. They estimate an econometric model of aggregate household expenditure, in which the share of total expenditure for thirteen types of good or service is expressed as a function of the total budget, the price of each good or service and an overall price index. This allows the own-price, cross-price and income elasticities of each good or service to be estimated.[93] Energy efficiency improvements reduce the cost of transport and heating and lead to substitution and income effects that change overall demand patterns (e.g. improvements in transport efficiency are estimated to increase demand for clothes but to decrease demand for beverages). By combining these estimated changes in demand patterns with CO_2 emission coefficients for each category of good and service (based upon estimates of direct and indirect energy consumption) Brännlund *et al.* find that energy efficiency improvements in transport and heating lead to (direct + indirect) rebound effects (in carbon terms) of 120% and 170% respectively.

Brännlund *et al.*'s results are heavily dependent on the assumed carbon emission coefficients, but the source of these is not made explicit. The results also contradict the econometric evidence on direct rebound effects, since carbon emissions for heating and transport are estimated to increase. Furthermore, Brännlund *et al.* use an iterative

---

[90] Results vary widely with country, time period and methodology. For example, Herendeen (1978) found that indirect energy consumption in Norway accounted for one third of total energy consumption for a poor family and approximately two thirds for a rich family. Vringer and Blok (1995) found that 54% of total energy demand in Dutch households was indirect, while Lenzen (1998) found that 30% of total energy demand in Australian households was indirect.

[91] Defined as the change in expenditure on a particular product or service, divided by the change in total expenditure. The marginal propensity to spend on different goods and services varies with income and it is an empirical question as to whether the associated indirect energy consumption is larger or smaller at higher levels of income. However, the greater use of energy intensive travel options by high income groups (notably flying) could be significant in some cases.

[92] In technical terms, these provide a partial equilibrium analysis, as distinct from the general equilibrium analysis provided by CGE models.

[93] Brännlund *et al.* employ Almost Ideal Demand (AID) model, which has been shown to have a number of advantages over other models of consumer demand (Deaton and Mulbauer, 1980; Xiao*, et al.*, 2007). The model relies on the assumption of 'staged-budgeting': for example consumers are assumed to first decide on the proportion of their budget to spend on transport, and then to decide how to allocate their transport budget between different modes. While analytically convenient, this assumption is likely to be flawed.

estimation procedure, but only present the results from the first estimation step. This weakness is overcome by Mizobuchi (2007), who follows a very similar approach to Brännlund *et al*, but applied to Japanese households. Despite the differences in data sources and estimation procedures, the estimated rebound effects are broadly the same. However, Mizobuchi also examines the effect of the additional capital cost of energy efficient equipment and finds that these reduce the rebound effect significantly.

The third example adopts a different approach, using data on the marginal propensity to spend of different income groups in Sweden. Alfredsson (2004) calculates the direct and indirect energy consequences of 'greener' consumption patterns, which include both technical changes, such as buying a more fuel-efficient car, and behavioural changes such as car sharing. In the case of 'greener' food consumption (e.g. shifts towards a vegetarian diet), the total energy consumption associated with food items is reduced by around 5% and total expenditure on food items is reduced by 15%. But the re-spending of this money on a variety of items, notably travel and recreation, leads to indirect energy consumption that more than offsets the original energy savings (i.e. backfire). The results for a shift towards 'greener' travel patterns are less dramatic, but the secondary effects from re-spending reduce the overall energy savings by almost one third. A comprehensive switch to green consumption patterns in travel, food and housing is estimated to have a rebound effect of 35%.

Secondary effects are relatively large in this example because 'green' consumption reduces expenditure on more than energy alone. Also, the results from such studies depend upon the methodology and assumptions used, as well as the types of household analysed and the particular shifts in consumption patterns that are explored. For example, a more recent study (Kanyama*, et al.*, 2006) using a similar model and approach to Alfredsson, but employing Swedish rather than Dutch data on energy intensity, finds that a shift to 'green' food consumption could reduce overall energy consumption. Closer examination reveals that this result follows largely from the assumption that greener diets are more expensive (owing to the higher cost of locally produced organic food), thereby leading to a negative 're-spending' effect.

In sum, the potential of embodied energy approaches to estimating secondary effects has yet to be fully explored. While the studies reviewed here suggest that secondary effects may sometimes be larger than commonly assumed, the conclusions may change once methodological weaknesses are addressed or a different choice of independent variable is made. Hence, at present the available evidence is too small to permit any general conclusions to be drawn.

6.3.4 Evaluation

The evidence reviewed above is too small and diverse to allow any general conclusions to be drawn about the size of the indirect or economy-wide rebound effects. However, it is interesting, to note that several of the studies quoted above estimate the indirect rebound effect to be relatively large – for example, 18-83% in the case of Kaufmann and Azary Lee and 33% in the case of Alfredsson. This contrasts with the view expressed by many energy economists, that such effects should be relatively small. For example, Lovins (1998) states that:

> "Though [indirect] 'rebounds' can in principle make net savings slightly smaller than gross savings…..they cannot make net savings become less than zero, because nothing that can be bought with the money saved by saving energy contains more energy per £ than the direct energy purchase that was saved in the first place." (Lovins, 1998)

Closer examination reveals this to be a version of the cost share argument discussed in Section 4.5. Indirect rebound effects are assumed to be small because, first, energy typically makes up a small share of total consumer expenditure; and second, the energy content of other goods and services is typically small. For example, suppose energy efficiency improvements reduce natural gas consumption for space heating by 10%. If there is no direct rebound effect, consumers will reduce expenditure on natural gas by 10%. If natural gas for heating accounts for 5% of total consumer expenditure, consumers will experience a 0.5% increase in their real disposable income. If *all* of this were spent on gasoline for additional car travel, the net energy savings (in kWh thermal) will depend upon the ratio of natural gas prices to gasoline prices, and could in principle be more or less than one.[94] In practice, however, gasoline only accounts for a portion of the total cost of car travel and car travel only accounts for a portion of total consumer expenditure. If, for example, the 0.5% increase in real income were spent on DVDs, the indirect rebound effect would be much smaller. For the great majority of goods and services, life cycle analysis data suggests that the effective expenditure on energy should be less than 15% of the total expenditure. Hence, by this logic, the indirect rebound effect should be only around one tenth of the direct effect.

Greening and Greene (1998) give an analogous argument for producers, focusing on the effect of the lower price of output from one firm or sector on the input costs of other firms or sectors. For example, efficiency improvements in steel production should lower the cost of steel, lower the input costs to vehicle production, lower the cost of passenger cars and thereby increase the demand for car travel. In this example, the indirect impact of an improvement in the energy efficiency of steel production should, in principle, depend upon the size of the efficiency improvement, the share of energy in the total cost of steel production and the share of steel in the total cost of automobile production. Since energy forms a small share of total production costs for most firms and sectors (typically <3%) and raw materials form a small share of the total costs of most products, the product of these suggests an indirect effect that is much smaller than the direct effect.

But these arguments may be flawed for two reasons. First, they confine attention solely to the 're-spending' effect in the case of consumers, and the effect on output costs in the case

[94] Chalkley *et al* (2001) provide an example of the replacement of an inefficient (C rated) refrigerator with an efficient (A - rated) model. Lifetime carbon savings for the refrigerator are estimated at 1645 $kgCO_2$ and lifetime cost savings at £120.57. If these cost savings were spent wholly on petrol, the indirect CO_2 emissions would be 358 kg, giving an indirect rebound effect, in carbon terms, of 22%. However, spending all of the cost savings on petrol is unrealistic.

of producers, and therefore ignore the embodied energy associated with the capital (or labour) that is used to improve energy efficiency. In other words, they entirely overlook the 'limits to substitution' (illustrated in Figure 6.3) that ecological economists emphasise and which Kaufmann and Azary-Lee sought (rather crudely) to estimate.

Second, they assume that the only effect of the energy efficiency improvement is to reduce expenditure on energy consumption. But, as argued repeatedly in Section 4, many energy efficient technologies (e.g. electric arc furnaces) also improve the productivity of other factors of production and hence may lead to cost savings that exceed the savings in energy costs alone. This is also the source of Alfredsson's relatively large rebound effect for 'green' consumption patterns, where the cost savings available for re-spending turned out to be much greater than those associated with either the direct or indirect energy consumption associated with the relevant goods and services. It could be argued that this is an inappropriate example, since such shifts in consumption patterns should not be classified solely as an energy efficiency measure. But measures such as cycling and increased use of public transport are routinely advocated as energy efficiency measures and changes in consumption patterns to reduce indirect (rather than direct) energy requirements are being increasingly advocated as such (Benders*, et al.*, 2006). So, some of Alfredsson's examples may be valid. Again, the dispute over the magnitude of indirect rebound effects would appear to hinge in large part on the identification of the appropriate independent variable.

6.4 The importance of energy quality

In reviewing Brookes' arguments in favour of backfire, Section 4.3 discussed Schurr's pioneering work on energy and productivity and highlighted the importance of changes in *energy quality* (and particularly electrification) in explaining productivity growth. But while this work dates back to the 1960s, its influence on subsequent research appears to have been limited. Energy quality continues to be neglected in the majority of studies of the relationships between energy, productivity and economic growth and indices of aggregate energy consumption continue to be constructed on the basis of thermal content, rather than weighted by cost shares or some other method.

In sharp contrast to this widespread neglect of energy quality by conventional economists, ecological economists have repeatedly argued that improvements in energy quality are a crucial and neglected causal variable in explaining economic growth (Cleveland*, et al.*, 1984; Cleveland*, et al.*, 2000; Kaufmann, 2004; Stern and Cleveland, 2004). Kaufmann's (1992) econometric analysis of the determinants of changes in energy/GDP ratios provides a particularly good illustration of this work.

Kaufmann (1992) sought to quantify the factors that contributed to changes in the ratio of primary energy consumption (in kWh thermal) to real GDP in France, Germany, Japan and the UK during the period 1950-1990. The explanatory variables were the percentage share of different energy carriers in primary energy consumption; the fraction of GDP spent directly on energy by households; the proportion of the product mix that originated in energy intensive manufacturing sectors; and primary energy prices.

Despite the simplicity of this formulation, it was found to account for most of the variation in energy intensity for the four countries studied throughout the post-war period. Kaufmann argued that improvements in energy quality led to lower energy intensities by allowing more useful work to be obtained from each heat unit of energy input. The shift from coal to oil contributed greatly to declining energy/GDP ratios prior to 1973, while the rising contribution of primary electricity (hydro and nuclear) provided a significant contribution after 1973.

Since the energy intensity of household energy purchases is an order of magnitude greater than the energy intensity of other goods and services, falls in the former as a fraction of total expenditure should translate into falls in the energy/GDP ratio – and vice versa. The fraction of GDP spent directly on energy by households increased prior to 1973 and decreased thereafter and these trends were also found to be highly significant in explaining trends in the aggregate ratio.

In addition, changes in energy prices encouraged substitution between inputs, including the substitution of capital for energy, while shifts towards less energy intensive manufacturing sectors and towards the service sector reduced energy/GDP ratios. These mechanisms were found to be less important than those above, but when all four factors were taken into account, they were found to provide a more or less sufficient explanation for the observed trends in energy intensity.

By implication, Kaufmann's results suggest little role for energy-saving technical change - defined as advances in technology that allow the same type and quantity of output to be produced with less energy inputs (i.e. the economy-wide *AEEI*). Kaufmann tested this implication in three different ways,[95] but in each case failed to find statistically significant evidence for energy saving technical change. These results suggest that conventional econometric models that fail to take account changes in energy quality could be misleading, in that changes in energy intensity that derive from shifts to higher quality fuels could instead be attributed to energy-saving technical change. As he states:

> "[This] should not be interpreted as an argument that substitution or technical change cannot reduce the amount of energy used to produce a unit of output.....technical change has reduced the amount of energy (as measured in heat units) used to produce a unit of output. But characterising that technical change as "energy saving" is misleading. Over the last 40 years, technical change has reduced the amount of heat energy used to produce a unit of output by developing new techniques for using oil, natural gas, and primary electricity in place of coal. The technical innovations... take advantage of the physical characteristics of these energies that allow oil, natural gas and primary electricity to do more useful work per heat unit than coal. This interpretation implies that technical change is not something shaped solely by the mind of man... but rather technical change is shaped in part by the physical attributes of energies available from the environment. (Kaufmann, 1992)

As indicated above, Kaufmann also interprets the results as illustrating the limited scope, at the level of the macro-economy, for substituting capital and labour for energy. Estimated annually, the own-price elasticity of energy demand varies between -0.05 and -0.39, which is generally smaller than the elasticities estimated at the level of individual sectors. This arguably suggests that the indirect energy consumption associated with labour and capital inputs constitute a significant portion of the energy saved directly through energy efficiency improvements in each of those sectors.

The results also indicate that reducing the fraction of GDP spent directly on energy by households, may be the most effective way of reducing the energy/GDP ratio. This in turn suggests that rebound effects from energy efficiency improvements may be lower in the household sector than in producing sectors.

[95] Namely: a) seeking evidence for serial correlation and heteroscedasticity in the error term, which could be evidence of missing variable bias; b) including a time trend to represent energy-saving technical change; and c) using dummy variables to test for changes in the intercept or slope of individual regression coefficients during different time periods - such as may follow an increase in energy prices if this induces energy saving technical change.

Kaufmann's work provides some quantitative support for Schurr's 'energy quality' hypothesis (Section 4.3) and poses a challenge to conventional assumptions. Decreasing energy/GDP ratios appear to be largely explained by structural change, price-induced factor substitution and the shift towards higher quality fuels. Once these are accounted for, there seems to be little evidence for energy-saving technical change reducing the energy/GDP ratio. Hence, not only does the energy/GDP ratio reflect the influence of factors *other* than energy-saving technical change, but these other factors appear to be *sufficient* to explain the observed trends. The observed improvements in the thermodynamic efficiency of individual devices at the micro level do not appear to have significantly contributed to the observed reduction in energy intensity at the macro-level. Instead, the latter owes much more to the changing mix of energy sources and the technical opportunities they present.

Hence, whatever the limitations of Kaufmann's analysis, these results suggest that the conventional neglect of energy quality in energy-economic analysis has important consequences for the conclusions that are drawn. It also suggests that further insight into rebound effects would benefit from a close examination of the role of fuel mix, in addition to structural change, price induced factor substitution and technical change. Unfortunately, most neoclassical studies - including those reviewed in Section 4.4 - focus upon the latter, rather than the former. This is an important omission.

6.5 Causality between energy consumption and GDP

In contrast to neoclassical economists, ecological economists claim a *causal* relationship between energy consumption and economic growth. The suggestion is that increases in the availability of energy have driven economic growth in the past, and that the reduced availability of high quality energy may act as a limiting factor in the future (Cleveland*, et al.*, 1984). This has strong parallels with Schurr's conclusion, quoted by Brookes' (1984), that "….it is energy that drives modern economic systems rather than such systems creating a demand for energy." Both claims appear to contradict conventional neoclassical assumptions: "As no conventional economic growth model takes into account the effect of energy use on economic growth, all such models are misspecified, and economic theory would need to be changed." (Stern, 1993)

As discussed in Section 4.6, there is a strong correlation between GDP and energy consumption in both industrialised and industrialising countries. While there is some evidence for decoupling following the oil price shocks of the early 1970s, the strong correlation largely re-emerges when energy inputs are weighted by the quality of different energy types (Cleveland*, et al.*, 2000). As discussed in Section 4.6, the evidence for an Environmental Kuznets Curve for primary energy consumption is similarly weak (Stern, 2004b). But while these observations are consistent with the claims of ecological economists, they do not *prove* a causal relationship from energy consumption to GDP. The relationship may well very run in the opposite direction - from GDP to energy consumption – with the implication that future economic growth may be less dependent upon access to high quality energy. Conversely, the correlation may simply be the result of the two variables sharing a common time trend.

It is frequently claimed that if causality runs from energy consumption to economic growth, then reducing energy consumption could damage economic growth. Conversely, if causality runs from economic growth to energy consumption, policies for reducing energy consumption could be implemented with little adverse effect on economic growth (Yoo, 2005). Alternatively, the causality may run in both directions, implying a mutual interdependence of energy and the economy. However, such claims can be both oversimplified and misleading (Zachariadis, 2006).

Beginning with the work of the Nobel Prize winning economist, Clive Granger (1969), modern econometrics has developed a set of sophisticated techniques for exploring such 'causality' question more carefully. Following Kraft and Kraft (1978), an increasing number of authors have used these techniques to examine the relationship between energy consumption and GDP in a variety of contexts (Chontanawat, 2006; Lee, 2006; Yoo, 2006). But despite the methodological sophistication of these studies, the results are contradictory and the policy implications are poorly developed. This section briefly summarises the logic and approach of this type of study, the appropriate interpretation of the results and the possible implications for the rebound effect.

6.5.1 Granger causality and cointegration

Granger (1969) proposed a straightforward test for detecting the presence of a causal relationship between two variables. A time series (x_t) is said to 'Granger-cause' another time series (y_t) if the prediction of *y* is improved by the inclusion of past values of *x* in addition to past values of *y*. The test is designed to show whether one variable can meaningfully be described as dependent variable and the other as independent, or whether the relationship is bidirectional, or whether no relationship exists at all (Stern and Cleveland, 2004).[96] This is really a test of 'statistical precedence' rather than causality as normally understood, since the fact that *A* precedes *B* need not necessarily mean that *A* causes *B*. Hence, while a finding of 'Granger-causality' may sometimes indicate actual causality, this need not always be the case (Granger, 1980). For example, a met office prediction of rain can be shown to Granger cause rain!

Granger causality tests may be applied to time series data that is both stationary and non-stationary[97] although in the latter case it is usually necessary to apply the test with first differenced data.[98] However, the results will be invalid if *x* and *y* are 'cointegrated', meaning that a particular linear combination of *x* and *y* is stationary, even when *x* and *y* are each non-stationary. In this case, it is necessary to use an error correction model (ECM) (Engle and Granger, 1997) to explore the relationship between the variables.[99] The presence of cointegration between two or more variables implies that one variable cannot move 'too far'

---

[96] A simple representation of precedence between two time series can be described in a bivariate model: $x_t = \sum_{j=1}^{n} a_j x_{t-j} + \sum_{j=1}^{n} b_j y_{t-j} + e_t$, and $y_t = \sum_{j=1}^{n} c_j y_{t-j} + \sum_{j=1}^{n} d_j x_{t-j} + h_t$, where *x* and *y* are the two series, *j=1,n* is the number of lags and e and h are uncorrelated errors. Then *y* Granger causes *x* implies that in the first equation $b_j \neq 0$. Similarly, *x* Granger causes *y* if $c_j \neq 0$.

[97] A non-stationary time series is one that is strongly dependent upon its value in previous time periods. For example, the autoregressive model: $y_t = y_{t-1} + \varepsilon$ represents an integrated time-series of order 1 (*I(1)*) and will follow a 'random walk'.

[98] First differenced data is given by $\Delta y_t = (y_t - y_{t-1})$.

[99] The ECM model structure is given by: $\Delta x_t = \sum_{j=1}^{n} a_j \Delta x_{t-j} + \sum_{j=1}^{n} b_j \Delta y_{t-j} + \gamma_j (x_{t-1} - \beta y_{t-1}) + e_t$ and $\Delta y_t = \sum_{j=1}^{n} c_j \Delta y_{t-j} + \sum_{j=1}^{n} d_j \Delta x_{t-j} + \lambda_j (y_{t-1} - \beta x_{t-1}) + h_t$. Note that the variables are included in first differences, while coefficients γ and λ represent the speed of adjustment in long-run equilibrium..

away from another, because there is a long-term relationship between them. This may be because one variable Granger causes the other, or that they are both driven by a third, possibly omitted, variable. Cointegration analysis seeks to identify this relationship by detecting whether the stochastic (i.e. irregular) trends in a group of variables are shared by the series, so that the total number of unique trends is less than the number of variables.

Although many authors use error correction models for exploring the relationship between two variables, it is increasingly common to use Vector Auto Regressive (VAR) models to examine the relationship between several variables. A VAR model consists of a group of regression equations in which each dependent variable is regressed on lagged values of itself and of all the other variables in the system.

6.5.2 Empirical results

The testing of causality relationships between energy consumption and GDP appears to have become something of a mini industry over the last decade, but the results remain frustratingly ambiguous.[100] Some studies find that causality runs from GDP to energy consumption, some find the reverse, some find mutual causation and some fail to obtain statistically significant results at all. Worse still, several studies obtain different results for the same countries and time periods, suggesting a strong dependence on the particular estimation methods employed (Paul and Bhattacharya, 2004; Zachariadis, 2006).

Potential reasons for such discrepancies are not hard to find. For example, causality results are very sensitive to the number of lags used in the regressions. Although test statistics for the appropriate number of lags have been proposed in the literature, they are rarely implemented in practice. Similarly, tests for the presence of cointegration and unit roots (i.e. integrated time series) have low power in the small sample sizes commonly found in causality studies (eg 20-30 observations) (Zachariadis, 2006). Also, several studies select variables that are mismatched in terms of economic sector (e.g. total energy consumption and industrial output) or method of normalisation (e.g. total energy consumption and GDP per capita) (Zachariadis, 2006). But perhaps the most important reason is that many studies confine attention to two variables (i.e. energy consumption and GDP) and do not examine or control for the effect of other variables such as capital, labour and energy prices (Stern, 1993). The resulting missing variable bias may either lead to spurious correlations or hide actual correlations. To avoid this, a multivariate approach is preferred, using a VAR model or equivalent procedure (Stern, 1993). A VAR model incorporating capital, labour, energy and GDP allows the marginal effect of energy use on output to be observed, holding other factors of production constant, as well as allowing the investigation of indirect channels of causation.

Stern (1993)[101] adopts this approach to provide an informative and methodologically rigorous study of the relationship between US energy consumption and GDP over the period 1947 to 1989. When gross energy consumption was measured in standard thermal units, Stern found that causality ran from GDP to energy consumption - as the standard neoclassical model would predict. This contrasted with a simple bivariate model which showed no causality in either direction. However, when final energy consumption was

---

[100] For good overviews of this literature, including tabulated summaries of results, see Lee (2006), Chontanawat (2006) and Yoo (2006).

[101] This is one of the ten most cited papers published in the *Energy Economics* journal.

quality weighted (using a Divisia index) the direction of causation was reversed: i.e. energy was found to 'Granger cause' GDP (as was labour and capital). Very similar results were found in a later study of the same data, using a more sophisticated methodology (Stern, 2000). Both studies could be interpreted as lending empirical support to the claims of ecological economists as well as highlighting, once again, the importance of energy quality in economic growth. However, the use of such quality weighting appears to be exception in studies of this type.

6.5.3 Implications

Given the diversity and inconsistency of this literature, it is difficult to draw any firm conclusions. However, it is likely that the inconclusive results of the earlier causality studies were either due to the omission of necessary variables or the failure to account for the quality of energy and other inputs (Stern and Cleveland, 2004). A combination of multivariate models and quality weighting of energy inputs overcomes these difficulties to large extent, and when this is done the results suggest that energy may be statistically important in explaining economic growth (Stern, 1993; 2000). This is consistent with the claims of ecological economists, who point to the dependence of productivity improvements on high quality energy inputs, both embodied in capital equipment (indirect) and used by them (direct). However, the results from such studies also demonstrate that labour and capital 'Granger cause' GDP, so the relative importance of each variable remains to be established.

A finding that energy Granger causes GDP has been interpreted by several authors as implying that the decoupling of energy consumption from economic growth could be difficult and that reductions in energy consumption could damage economic growth.[102] But such interpretations may be oversimplified. The reason that energy is economically significant is that it is used to perform *useful work* - either in the form of mechanical work (including electricity generation) or in the production of heat. But more useful work can be obtained with the same, or less, energy consumption through improved (primary or secondary) conversion efficiency. Hence, while a reduction in the availability of useful work could potentially be damaging for an economy, a reduction in the availability of (high-quality) primary energy inputs need not be, if it is achieved by, or mitigated through, improvements in conversion efficiency (i.e. the substitution of capital for energy). Given the low second law efficiencies associated with many energy services (e.g. space heating), the scope for improved conversion efficiencies in many (if not all) sectors appears to be high (Hammond and Stapleton, 2001). However, as highlighted in the previous section, the indirect energy consumption associated with the relevant capital equipment also needs be taken into account. The point at which this becomes sufficiently large to offset the direct energy savings (Figure 6.3) can only be established empirically. But to the extent that aggregate savings in primary energy consumption are still feasible, it is incorrect to assume that 'energy conservation' will necessarily hinder economic growth.

A finding that quality weighted energy Granger causes GDP would appear to be consistent with Brookes' general argument that "….it is energy that drives modern economic systems

[102] For example, Stern (1993) observes that: "…. the policy implication of this research would be that raising taxes on energy or adopting other policies that reduce energy use... would reduce the rate of economic growth and, if severe enough, reduce the level of output.". Similarly, Lee and Chang (2005) observe that: "…. the empirical results shows unanimously in the long-run that energy acts as an engine of economic growth, and that energy conservation may harm economic growth."

rather than such systems creating a demand for energy" (Brookes, 1984). It also appears consistent with Schurr's electrification hypothesis (Section 4.3), the econometric evidence for energy using technical change (Section 4.4) the absence of evidence for an Environmental Kuznets Curve for quality weighted energy (Section 4.6); and Kaufmann's analysis of the factors determining energy/GDP ratios (Section 6.3). Hence, to the extent that each of these findings provides suggestive support for the Khazzoom-Brookes postulate, the causality literature adds a further source of evidence. However, the link is, at best, highly indirect. In particular, what is missing from the causality literature is the inclusion of data on conversion efficiencies and useful work (rather than energy consumption per se) as well as the quantification of the importance of energy relative to other variables. But this is precisely what is addressed by another stream of ecological economics literature that incorporates energy/useful work into alternative models of economic growth. Recent work in this tradition by Ayres and Warr (2005) appears to offer a potentially fruitful approach to exploring the macroeconomic rebound effect. This is discussed next.

6.6 Ecological growth models

The starting point for this literature is a critique of the standard assumptions of neoclassical production theory and neoclassical growth theory - described earlier in Section 5. The most relevant assumption is that the primary factors of production are capital and labour services, with energy either being ignored or treated as an intermediate input. If output is only a function of capital and labour inputs, their marginal productivities should equal their corresponding payment shares in the national accounts. If energy and materials are included as inputs, their marginal productivities should be proportional to the share of the relevant intermediates in the national accounts Since this is small (<5%), the contribution of the growth in energy inputs to economic growth should also be small. Very similar conclusions follow for the contribution of energy to output growth at the level of individual sectors.

As discussed in earlier sections, techniques of 'growth accounting' have allowed the contribution of increases in factor inputs to increases in GDP to be estimated (Section 3), but these are generally found to be insufficient to explain economic growth (Fabricant, 1954). Instead, much of the increase in economic output is attributed to 'technical change' or improvements in total factor productivity (*TFP*) and represented by an exponential function of time that serves as an exogenous multiplier to the production function.

Work by Jorgenson and Griliches (1967) and others has reduced the size of this 'residual' by accounting for changes in the quality of labour and capital inputs, but has not substantially changed the conclusion that an important driver of growth is exogenous technical change (Easterly and Levine, 2001). Therefore, the origins, nature and determinants of economic growth remain partly unexplained by neoclassical growth theory. These difficulties have led to the development of endogenous growth theory, in which various mechanisms have been proposed to explain why observed returns to capital and rates of growth in the industrialised countries have not declined over time (Romer, 1986). But the explanatory variables of endogenous growth theory are difficult to measure, while measurable proxies such as educational expenditure appear insufficient to explain growth. Also, the role of energy and natural resources continues to be neglected (Ayres and Warr, 2002b).

Ecological economists have strongly criticised both the neoclassical and endogenous growth models and sought to develop alternative models that give a greater role to energy as a primary factor of production (Ayres, 2001; Hall*, et al.*, 2001). These models are of interest for a number of reasons, not least because they appear to fit historical data on GDP growth extremely well. Indeed, as a consequence of their modified assumptions and approach, these models considerably reduce the size of the 'Solow residual' or even eliminate it altogether. As such, they provide an alternative means of addressing a key weakness of neoclassical growth theory, as well as providing potentially useful insights into the conditions for future economic growth (Ayres, 2001; Ayres and Warr, 2002a; Ayres and van den Bergh, 2005). These 'alternative' growth models have one or more of the following features in common:

- The inclusion of energy (or some related measure such as exergy or useful work) as a factor within the aggregate production function.
- A departure from the traditional assumption that factor productivities are proportional to the share of that factor in the value of output. Energy inputs are instead measured in physical terms and marginal productivities are estimated directly from a production function, rather than indirectly from a cost function.
- A departure from traditional methods of estimating production functions, including the use of an unconventional, linear-exponential (LINEX) function in which factor productivities are not assumed to be constant.
- The use of improvements in thermodynamic conversion efficiency as a suitable proxy for technical change.

This ecological economics literature falls into three closely related groups:

- Research by the Canadian economist, Bernard Beaudreau (1995a; 1998; 2005).
- Research by a group of German researchers with a background in the physical scientists (Kummel, 1980; Kummel, 1982; Kummel, *et al*., 1985; Kummel, 1989; Kummel, *et al*., 2000; Kummel, *et al*., 2002).
- Research by Rob Ayres and Benjamin Warr at INSEAD (Ayres, 1998b; Ayres, 2001; Ayres and Warr, 2002a; Ayres, 2002; Ayres and Warr, 2002b; Ayres*, et al.*, 2003; Ayres and van den Bergh, 2005; Ayres and Warr, 2005)

These studies have much in common, but their theoretical models differ in important respects. Kummel and colleagues were the first to introduce the LINEX production function, which was subsequently borrowed by Ayres and Warr. Beaudreau, in contrast uses more conventional production functions and appears to have developed his ideas relatively independently. The work by Ayres and Warr is the most recent and accessible, as well as being the most clearly relevant to the rebound effect. But all these studies remain outside mainstream growth theory and all reach conclusions that run counter to it. Each will be reviewed below.

6.6.1 Beaudreau

Beaudreau (1998) argues that conventional growth theory is misleading because it ignores basic physical principles. He proposes an alternative production function where output is a function of 'useful work' and 'organisation'. The former represents the output of energy conversion devices and should in principle be given by the product of exergy inputs and the second law efficiency of the relevant conversion devices – an idea subsequently taken up by Ayres and Warr (see below). The latter is variously described as 'information' and 'supervision', but ultimately depends upon capital and labour.

Beaudreau (1995a) estimates this function empirically for value added in US manufacturing over the period 1950-84, while Beaudreau (2005) develops comparable estimates for Japanese and German manufacturing. However, given the difficulty in finding measurable proxies for the proposed factors of production, Beaudreau takes electric power as a proxy for useful work and conventional measures of labour and capital inputs as a proxy for organisation. This empirical work therefore goes only half way towards testing Beaudreau's full theoretical framework.

Traditionally, factor productivities are estimated from cost functions rather than production functions, because estimates of the latter are prone to bias. The conventional approach is based upon the theory of duality (Beattie and Taylor, 1993) and relies upon a number of assumptions, including perfectly competitive factor markets (Section 3.3). Beaudreau (1995a) considers this assumption to be invalid in the case of US electricity markets, which were heavily regulated over the period in question. He therefore estimates factor productivities directly from a production function, which is assumed to take a Cobb Douglas form. The output elasticity for electric power[103] is estimated to be 0.53, which implies that a 1% increase in electric power consumption results in a 0.53% increase in manufacturing value added. This is an order of magnitude larger than conventional estimates: for example, Berndt and Wood (1975) estimate aggregate energy to have an output elasticity of only 0.06 in US manufacturing. The marginal productivities of capital and labour are correspondingly estimated to be much smaller than conventionally assumed.

Beaudreau's results suggest that the growth in electricity consumption accounted for 79% of the growth in manufacturing value added over the period in question and that the substitution of capital and electricity for labour accounted for 97% of the improvement in labour productivity. Furthermore, if a Divisia index for aggregate inputs is formed by weighting by the estimated marginal productivities (rather than by cost shares), the growth in output is found to be fully explained by the growth in inputs – thereby eliminating the need for an exogenous multiplier for technical change. The implication is that the increased availability of low-cost electric power provides the main explanation of the growth in productivity in US manufacturing since 1950 – a result which echos Schurr's (1983) electrification hypothesis (Section 4.3) as well as Joregensen's (1984) findings of energy-using technical change (Section 4.4).

However, the reasons given by Beaudreau for the preferred use of a production function (notably lack of competition the US electricity markets) seem insufficient to account for the very large difference between his results and those of more conventional studies. In Beaudreau (1995a), the output elasticity for capital is found to be a statistically insignificant, which appears suspicious. Also, in Beaudreau (2005) the estimates for the output elasticities of capital and labour vary widely between the different countries, but no satisfactory explanation for this is given. Since Beaudreau's work has been overlooked by conventional economists, the strengths and weaknesses of his approach are difficult to gauge.

6.6.2 Kummel et al

Kummel and colleagues have published a series of studies on the role of energy in economic growth which take the same starting point as Beaudreau - namely, that the output elasticity of energy inputs should not be equated to the share of energy in total costs (Kummel, 1980;

[103] Namely: $\eta_E(Y) = \partial \ln Y / \partial \ln E$

Kummel, 1982; Kummel, *et al*., 1985; Kummel, 1989; Kummel, *et al*., 2000; Kummel, *et al.*, 2002). This also leads them to estimate a production function for economic output, rather than the more conventional cost function. However, instead of *assuming* a form for the production function and performing statistical fitting operations to estimate parameter values, Kummel (1982) begins with a functional form for the factor productivities and uses a partial integration to *derive* a production function. The starting point is an expression relating the marginal change of normalised output ($y = Y / Y_o$) to the marginal change of normalised capital (*k*), labour (*l*) and energy inputs (*e*), *without* any time trend to represent technical progress:

$$\frac{d \ln y}{dt} = \alpha \frac{dk}{dt} + \beta \frac{dl}{dt} + \gamma \frac{de}{dt} \qquad (6.1)$$

The parameters in Equation 6.1 represent the output elasticities, or marginal productivities of each factor (e.g. $\alpha = \partial \ln y / \partial \ln k$). Kummel (1982) adopt the conventional assumption that the production function exhibits constant returns to scale, which requires $\gamma = 1 - \alpha - \beta$. He then derives expressions for the marginal productivities of each factor from three differential equations "….that result from the requirement that the second order mixed derivatives of *y* with respect to *k,l,e* must be equal", together with additional assumptions about the asymtopic values of these marginal productivities (Kummel, 1982). Unfortunately, most of Kummel's published work fails to provide an adequate explanation of this important step and the proposed asymptopic conditions appear somewhat arbitrary. This leads to functional forms for the marginal productivities, which are further constrained to be non-negative. Partial integration then leads to a time dependent 'linear-exponential'[104] (LINEX) production function as follows:

$$y = e * \exp\left[a\left(2 - \frac{l+e}{k}\right) + ab\left(\frac{l}{e} - 1\right)\right] \qquad (6.2)$$

This procedure is difficult to interpret and difficult to compare with more conventional approaches. However, a useful feature of the LINEX function is that it imposes limits on the degree to which capital and labour can substitute for energy - as required by ecological theory. Another feature of this function is that it implies variable marginal productivities and variable elasticities of substitution throughout the period fitted. But the function also fails to satisfy the standard condition that production functions must be concave - implying that the marginal product of each input declines with increasing use of that input (Saunders, 2007). This feature may make it unacceptable to many conventional economists.

Kummel *et al* (2000) fit this model to time series data of industrial output in Germany, Japan and the US over the period 1960 to 1993. This involves the use of a non-linear estimation algorithm, but again this process is not adequately explained. The results are found to match the observed data extremely well, with R^2 as high as 0.999. The implication is that the growth in capital, labour and (primary) energy inputs can explain economic growth *without* the need for an exogenous technical progress term, provided that the marginal productivities are estimated in the unconventional manner indicated above. These productivities are not constant over time, but the average value for the marginal productivity of energy inputs is of the order of 0.5 (Kummel*, et al.*, 2000). As with Beaudreau, this exceeds conventional values (based upon the value share of energy) by a factor of 10. In contrast, the marginal productivity of labour is estimated to be in the range 0.05-0.2, which is much smaller than its value share.

[104] Since output depends linearly on energy and exponentially on factor ratios.

Kummel's published work lacks clarity in certain important respects and appears to have a number of statistical weaknesses. Because of the highly flexible nature of the LINEX function (e.g. a variable elasticity of substitution, without a priori restrictions on parameter values based on economic theory), it appears able to fit the data even when fluctuations in output are clearly due to exogenous factors, such as economic crises. Diagnostic statistics are poorly reported and where available suggest possible statistical problems. Notably, the traditional measures of goodness-of-fit (R^2) are suspiciously high (e.g. 0.999) while the standard errors for many of the coefficient estimates are small (e.g. 0.51 ± 0.02). Both of these are suggestive of serial correlation in the error terms[105], which is confirmed by the quoted values of the Durbin-Watson statistic.[106] Serial correlation, in turn is a sign that the equations are misspecified and the function has been 'overfitted' to the data (Hendry and Mizon, 1978). The authors do not comment on this and do not investigate alternative tests for serial correlation. Despite several publications of similar results, they also do not attempt to eliminate the misspecification through either the use of differenced data or through alternative estimation techniques. This is a major weakness.

Kummel's efforts have been overlooked by growth economists, although the basic approach and results were first published 20 years ago. However, his contributions have inspired more recent work by Ayres and Warr (2005), which takes energy efficiency directly into account.

6.6.3 Ayres and Warr

6.6.3.1 Approach

Of the three groups of 'alternative growth models', the work by Ayres and Warr (2005) is the most accessible and influential - as well as appearing the most relevant to the rebound effect. Their work is based upon a well articulated theoretical framework, in which economic growth is considered to be driven by one or more self-reinforcing feedback mechanisms - termed 'growth engines' (Ayres and van den Bergh, 2005). The proposed 'resource use (fossil fuel)' growth engine is closely related to the rebound effect (Ayres*, et al.*, 2003):

> "The generic energy-power feedback cycle works as follows: cheaper energy and power due to discoveries, economies of scale and technical progress (learning) in energy conversion, enable goods and services to be produced and delivered at lower cost…..Lower costs in competitive markets translate into lower prices for products and services. Through price elasticity effects, lower prices encourage demand. Since demand for final goods and services corresponds to the sum of factor payments, most of which go back to labour as wages and salaries, it follows that wages of labour tend to increase as output rises. This, in turn, stimulates further substitution of fossil energy and mechanical power for human (and animal) labour, resulting in further increases in scale, learning and still lower costs." (Ayres and Warr, 2006)

A key innovation by Ayres and Warr is the use of *exergy* as a generalised measure of both fuel and raw material inputs to the economy (see Section 3). Ayres *et al* (2003) have constructed a time series of exergy inputs to the US economy over the past century, incorporating fossil fuels, biomass, nuclear and renewable electricity, minerals and metals. Of these, fossil fuels provide by far the greatest contribution. Ayres*, et al* (2003) estimate

---

[105] Serial correlation means that the residuals from two time periods are correlated. Estimation by OLS will still be unbiased and consistent, but the estimated variances of the regression coefficients will be biased. If the serial correlation is positive and the independent variable is growing over time, then the standard errors will be underestimates of their true values.

[106] This test should give a value around 2 if there is no serial correlation. Postive serial correlation is indicated by values below a critical value, but instead of specifying a single critical value the DW test provides a range. Values falling below this range indicate positive serial correlation, while values falling within this range are inconclusive. Kummel's results indicate provide both positive and inconclusive DW results. However, there are number of drawbacks with this test and it is invalid if the time-series include lagged dependent variables.

that total exergy inputs to the US economy have grown by a factor of 8 over the course of the century, while the ratio of exergy inputs to GDP has fallen by two thirds.

Ayres *et al* (2003) have also painstakingly constructed a time series for the exergy (i.e. second law) conversion efficiencies of low, medium and high temperature heat, mechanical work and electricity production in the US economy. They estimate that overall exergy conversion efficiencies have improved five-fold over the course of the century, from 3% to 15%. The greatest improvement has been in electricity generation, where conversion efficiencies have improved from 4% to around 33%. These conversion efficiencies relate to particular types of technology and are not necessarily inconsistent with stable or even increasing energy intensities at higher levels of aggregation – such as found, for example, by Sue Wing and Eckaus (2006b).

The data on exergy inputs may be considered a quality-weighted time series of natural resource inputs (especially fossil fuels), while the data on exergy conversion efficiencies may be considered a particular measure of technical change. However, since the latter is based upon physical measures of thermodynamic conversion efficiency rather than economic measures of energy productivity, it is different from the conventional measures of 'energy-saving' technical change and the *AEEI*, defined in Section 3.

By combining their data on exergy inputs and exergy conversion efficiencies with estimates of the proportion of different fuels used in different applications, Ayres *et al* (2003) are able to derive a time series of the *useful work* inputs to the US economy (which they also term 'exergy services'). This is a measure of the productive inputs derived from both materials and energy and is divided into *mechanical* work by prime movers such as steam engines and gas turbines; *chemical* work to drive processes such as ore reduction; and *thermal* work to deliver low, medium or high temperature heat to a point of use.[107] By this means, Ayres *et al* are able to demonstrate that the useful work inputs to the US economy have grown by a factor of 18 over the past 100 years, implying that the useful work obtained from natural resources has grown much faster than the consumption of those resources themselves.

This empirical work is a valuable contribution in itself, since it is the first time that thermodynamic conversion efficiencies have been estimated on an economy-wide basis. But it also allows Ayres and Warr (2005) to develop a growth model in which useful work is included alongside capital and labour within an aggregate production function. This makes a great deal of sense, since it is the amount of exergy delivered in useful form that is likely to be economically productive, while the exergy inputs that are lost in conversion processes are effectively wasted (Ayres and Warr, 2006)

Ayres and Warr (2005) also construct time series of capital and labour inputs over the past century (although these are not quality weighted) and they show that US GDP has increased faster than either capital, labour or exergy inputs. Since production functions are conventionally assumed to exhibit constant returns to scale and since US GDP has increased faster than any homogeneous first-order combination of these three inputs, Ayres and Warr (2005) argue that standard production functions *cannot* represent output growth in the US without including a time-dependent factor representing technical progress. However, both Beaudreau and Kummel fitted production functions to historic GDP data without using a time

[107] Ayres *et al* also estimate the contribution of *muscle* work from farm animals, which was economically significant at the beginning of the century but is of negligible importance now

trend and both found very small residuals. So Ayres and Warr's argument appears to be at odds with Beaudreau and Kummel's results, although they cite them both.

Ayres and Warr borrow the general approach of Kummel*, et al.* (1985) as well as the specific form of the LINEX production function, but they replace primary energy inputs with their measure of useful work. As with Kummel*, et al.*, they find the fit to GDP trends be extremely good (R^2=0.99), thereby eliminating the need for exogenous *TFP* multiplier. They estimate that the marginal productivity of useful work exceeds that of capital and labour and has increased over time, while that of labour has steadily declined. This contrasts with declining resource prices and increasing wage rates over the same period (thereby suggesting a mismatch between economic returns and physical productivities). In 2000, the estimated marginal productivities were approximately 0.7 for useful work, 0.5 for capital and 0.05 for labour. Ayres and Warr (2005) conclude that the increasing productivity of physical work is by far the dominant driver of past growth and will continue to be for decades to come.[108]

6.6.3.2 Evaluation

The implication of Ayres and Warr's work is that improvements in exergy conversion efficiency is a plausible and quantifiable surrogate for all forms of technical change that contribute to economic growth:

> "Our core hypothesis is that the economy can be regarded as materials/exergy conversion system, when the technological knowledge is approximately proportional to the ratio of useful work output to the flow of primary natural resource exergy input. The five fold improvement in the flows of exergy services provided per unit of raw natural resource exergy.....is a rather good indicator of technological progress…."[109]

The enormous improvements in exergy conversion efficiencies have not reduced aggregate exergy consumption, but instead reduced the cost of exergy services and driven economic growth. Far from being a minor contributor to economic growth, improvements in exergy efficiency become the dominant driver – obviating the need for alternative measures of technological change, or improvements in 'human capital'.

Ayres and Warr's work implies that the rebound effect from improved energy efficiency is very large. To illustrate this, suppose that improved exergy efficiency accounted for one half of US economic growth over the course of the last century (Ayres and Warr do not provide a precise figure, but 50% is comparable to many estimates of the proportion of growth attributed to technical change). Since exergy efficiency improved five fold over the last 100 years, while economic output increased 20-fold, this implies that each 1% improvement in exergy efficiency accounted for a $2.0*\eta_Y(E)$ increase in exergy consumption, where $\eta_Y(E)$ represents the income elasticity of exergy demand. Assuming that the latter is broadly equivalent to the income elasticity of primary energy demand and taking a typical value for OECD countries of 0.5, this implies that each 1% improvement in exergy efficiency led to corresponding 1% increase in exergy demand – in other words, efficiency improvements led

---

108 The fact that GDP growth has exceeded the growth in factor inputs since 1975, leads Ayres and Warr to speculate that information and communication technologies may be providing an additional source of growth. Some statistical support for this hypothesis is provided in Ayres and Warr (2002a), who include the capital invested in information technology as a fourth factor of production. But more conventional accounting exercises place far more importance on ICT (Jorgenson, *et al*., 2005).

109 They further argue that some of the most dramatic and visible technological changes, such as medical progress and telecommunications, have not contributed significantly to economic growth but instead have improved quality of life.

to backfire.[110] The size of the rebound would be reduced if energy efficiency improvements were estimated to account for a smaller share of economic growth, but would be increased if the income elasticity of exergy consumption was estimated to be higher (as is likely to be the case in many developing countries).

However, Ayres and Warr's empirical work appears to have similar flaws to that of Kummel, *et al*. First, there are the difficulties with the LINEX function noted above, including its obscure nature, the apparent conflict with standard economic theory and the lack of adequate explanation. For example, Ayres and Warr refer to constrained non-linear optimisation using quasi-Newton methods, but do not cite a previous paper where this is used and do not provide the objective function that was optimised.

Second, the results from Ayres and Warr show the same signs of misspecification that were evident with Kummel. The authors note that standard R^2 measures are only valid in the absence of serial correlation, but then report strong evidence for serial correlation without further comment. This should have been picked up in peer review, since the presence of serial correlation renders statistical inference invalid.[111]. The results also suggest the presence of multicollinearity, but this possibility is not discussed.[112] The authors report modifications of the distributions of the *t* statistics (usually for non-constant variances in the residuals) but no methodology or source is cited for these modifications.

Third, the key message from the paper is that useful work provides a better fit to the data than raw exergy inputs and this forms the basis for Ayres and Warr's theoretical speculations. But although Ayres and Warr adopt Kummel's LINEX production function, they make no comment as to why he was able to reproduce economic growth without a time trend, while measuring energy inputs on a thermal basis (i.e. neither quality-weighting nor accounting for improvements in conversion efficiency). A comparison of the methods and result of these studies and an explanation for the differences is therefore still required.

6.6.4 Summary

The theoretical arguments that lay behind the work of Beaudreau, Kummel and Ayres and Warr are consistent with the perspective of ecological economics and seem very persuasive. If correct, their work would add further weight to the argument that the increased availability of energy is a primary driver of economic growth. Also, Ayres and Warr's work would add further weight to the argument that the economy wide rebound effect is large.

If the marginal productivity of energy inputs is as large as estimated here (i.e. ten times larger than the cost share), the argument that rebound effects must be small because the share of energy in total costs is small is undermined. Instead, improvements in the

---

[110] The efficiency improvements that are particularly relevant are those within upstream, process industries, including energy supply. Efficiency improvements in secondary conversion and consumer technologies are assumed to have played a smaller role. Nevertheless, although increases in energy consumption partly derive from the proliferation of downstream technologies (e.g. household appliances) made possible from the increased availability of low cost electricity.

[111] In a private communication, the authors reported that alternative estimates with a Cobb Douglas production function including a technology multiplier produced even stronger evidence for serial correlation. But that doesn't make the LINEX estimates valid.

[112] Also, there are typographical errors in Appendix B of Ayres and Warr (2005). The column of t values in Table B.1 has been inverted and the preceding text should read "…the correlations were significant, and the latter choice was by far the most significant, as indicated by the large t-value".

productivity of energy inputs, could have a dramatic effect on output growth and therefore on overall energy consumption. Such improvements may result both from changes in the quality of energy inputs and improvements in thermodynamic conversion efficiency, since both increase the amount of useful work obtained from a heat unit of energy inputs.

However, none of these authors have provided adequate empirical support for their claims and their published work contains a number of flaws. This may partly explain why their work has attracted such little attention from conventional growth economists, although a clash of 'world views' is another likely explanation. Nevertheless, the data collected by Ayres *et al* (2003) on exergy and useful work provides a valuable basis for further empirical investigation and the analysis of this data with more robust econometric methods (e.g. cointegration) could offer a promising way forward. We understand that Ayres and Warr are conducting such a study at present, but have yet to publish their results.

6.7 Summary and implications

Underlying Brookes' arguments in favour of the K-B postulate is an assumption about the contribution of energy to economic growth. While most economists assume that increases in energy inputs make a relatively small contribution to economic growth, Brookes endorses Sam Schurr's observation that: "…. it energy that drives modern economic systems rather than such systems creating a demand for energy" (Brookes, 1984). But this is precisely the claim made by contemporary ecological economists. Hence, while ecological economists have not directly investigated the rebound effect, much of their work may be relevant to it. This section has therefore reviewed the ecological perspective on energy and economic growth and has summarised and critiqued a number of associated empirical studies.

The theoretical and empirical studies reviewed here are consistent in many ways with those reviewed earlier in Sections 4 and 5. They may be used in support of the following:

- The increased availability of low-cost, high-quality energy sources has provided a necessary condition for the technological changes and economic growth experienced over the past century (Cleveland*, et al.*, 2000).
- Reductions in aggregate energy/GDP ratios are largely explained by structural change and improvements in the quality of energy inputs. Energy saving technical change at the micro-level has contributed much less to such reductions than is commonly assumed (Kaufmann, 1992).
- The observed reduction in aggregate energy/GDP ratios may therefore overstate the extent to which energy consumption has been decoupled from economic growth (Kaufmann, 2004).
- Future increases in the price of high-quality energy sources could have significant economic impacts (Stern, 1993).
- If the link between energy use and economic activity is stronger than commonly assumed, attempts to reduce energy consumption will be more expensive than is commonly assumed (Cleveland*, et al.*, 1984).

The ecological perspective is well articulated and persuasive. However, the empirical evidence in support of this perspective remains patchy and in some cases flawed. In particular:

- Estimates of the indirect energy consumption required to achieve energy efficiency improvements are rare, while estimates of the indirect energy consumption associated with various categories of goods and services exhibit considerable diversity. Such studies are rarely detailed enough to allow either the embodied energy or the secondary effects associated with energy efficiency improvements to be estimated. As a result, they provide few empirical estimates of the magnitude of the indirect or economy-wide rebound effect and provide an insufficient basis on which to draw any general conclusions.
- The results of econometric investigations of causality relationships between energy and GDP remain ambiguous and the policy implications that are drawn are oversimplified. Methodologically rigorous studies that quality-weight energy inputs appear to suggest that energy Granger-causes GDP, but these results are far from definitive. Also, Granger causality is not necessarily the same as causality as conventionally understood.
- The different variants of 'ecological growth models' appear to suffer from misspecification, with insufficient attention being paid to the presence of serial correlation. As a result, claims that the marginal productivity of energy is in order of magnitude larger than its cost share, or that improvements in exergy conversion efficiency can act as a suitable proxy for technical change, must be treated with considerable caution.

The difficulty, at present, is that different assumptions of conventional and ecological perspectives seem to have prevented an objective comparison of their methods and conclusions. While ecological economists assume *a priori* that energy must play a dominant role in economic growth, neoclassical economists appear to assume *a priori* that it must play a minor role. The ecological perspective can only be reconciled with the neoclassical perspective if convincing evidence can be provided that the contribution of energy to economic growth substantially exceeds its share of total costs. The growth models reviewed above have so far failed to provide this. As Toman and Jemelkova (2003) note in their review of the literature on energy and economic development:

> "…..we do find some important illustrations of a disproportionate role for energy... (but) the amount of literature we found was very limited, and in many cases it was difficult to separate out various influences in the study to see how energy might be exerting a disproportionate role. This underscores our conclusion that while much is known about how the productivity of energy provision and use might be augmented at the micro-level, more work is needed to understand the magnitude of its importance for economic development at an economy wide level." (Toman and Jemelkova, 2003)

However, drawing upon ideas from endogenous growth theory, Toman and Jemolkova (2003) propose a useful way of understanding how energy could play such a disproportionate role. Their suggestion helps to bring together several of the observations of the preceding sections.

Toman and Jemolkova propose an aggregate production function in which economic output is produced from a combination of capital services (K), labour services (L) and 'energy services', where the latter may be understood as the output from energy conversion devices. This is closely analogous to Ayres and Warr's formulation, only the latter use data on useful work (*U*) to represent energy services. The key point is that increased inputs of energy services, or useful work, may *enhance* the productivity of capital and labour:

> "……increased energy use has augmentation effects on the productivity of other factors……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." (Toman and Jemelkova, 2003)

Mathematically, this may be represented by the use of 'augmentation multipliers' (τ_i) for capital and labour inputs, as introduced in Section 3.3.3. The key assumption is that the degree of factor augmentation depends upon the availability of useful work, as follows:

$$Y = f(\tau_K(U) * K, \tau_L(U) * L, U) \qquad (6.3)$$

The result is an aggregate production function that exhibits *increasing returns* to useful work (and hence energy) inputs – thereby departing from a core assumption of neoclassical theory. This approach is similar that adopted in endogenous growth theory, where factor augmentation is achieved through R&D, education and the provision of public goods. Toman and Jemelkova also suggest that the production of useful work may itself be subject to increasing returns, which provides another route by which increases in energy consumption could have a disproportionate impact on economic output.

Focusing in particular on households in developing countries, Toman and Jemolkova (2003) propose a number of ways in which the increased availability of useful work could improve capital and labour productivity and hence disproportionately affect economic output. For example, cheaper and better lighting could allow greater flexibility in time allocation throughout the day and evening and enhance the productivity of education efforts. The increased availability of electricity could promote access to safe drinking water (e.g. in deeper wells), allow the refrigeration of food and medicine and thereby improve both the health of workers and their economic productivity. Similarly, the increased availability of low-cost transport fuels could interact with investment in transport infrastructure to increase the geographic size, scale and efficiency of markets. Schurr's account of the impact of electricity on the organisation and productivity of US manufacturing (Section 4.3.1) provides an analogous example for developed countries, as does Lovins and Lovins (1997) micro-example of the higher labour productivity associated with energy efficient buildings.

It is an empirical question as to whether such benefits apply in practice and to what extent. It may be expected that the degree of factor augmentation will vary between different sectors, technologies and time periods, but the key question is the net effect on economic growth in the aggregate. Ecological economists appear to claim that such factor augmentation is the norm, with the result that the increased availability of energy (and hence useful work) has a greater impact on economic growth than is suggested by the share of energy in total costs. Ecological economists also appear to claim that the degree of factor augmentation is large, with the result that the increased availability of energy (and hence useful work) has historically been a primary driver of economic activity.

However, these arguments are normally couched in terms of the availability of high-quality energy sources (i.e. exergy), rather than useful work (Cleveland*, et al.*, 2000). The link between the two is thermodynamic conversion efficiency, which is the relevant variable for the rebound effect. But if the increased availability of exergy inputs has a disproportionate impact on productivity and economic growth, then improvements in energy efficiency should do the same. This is because the effect of both is to increase the availability of useful work. The large boost to productivity may in turn have a large impact on economic output and therefore on the demand for useful work. This in turn will increase demand for energy inputs, and in some cases this may be more than sufficient to offset the energy savings derived from the energy efficiency improvement. Schurr's historical work appears to provide evidence for this process in the US economy (Section 4.3), which is why Brookes cites it as evidence for backfire.

If it is useful work rather than raw energy (or exergy) inputs that drives economic activity, then improvements in thermodynamic conversion efficiency could mitigate the economic impact of increasing shortages of high-quality forms of energy. However, improvements in conversion efficiencies are necessarily associated with embodied energy and are ultimately constrained by thermodynamic limitations. Also, if these improvements have a disproportionate effect on economic output, they may also be associated with large rebound effects. Unfortunately, the jury is still out on whether the net effect will be to increase energy consumption in the aggregate. If so, the scope for decoupling economic growth from energy consumption may be limited. If not, a degree of decoupling may be achieved, although the process will ultimately be limited by the indirect energy consumption associated with both obtaining high-quality sources of energy and improving the thermodynamic efficiency of energy use. Despite the wide range of studies reviewed in this report, the available empirical evidence fails to provide a clear indication of which types of energy efficiency improvement are more likely to lead to backfire, or whether such an outcome is likely to be the norm.

7 Summary and Conclusions

7.1 In search of Nessie

In his introduction to the special edition of *Energy Policy* on the rebound effect, Lee Schipper (1998) likened the persistent debate on the issue to the repeated sightings of the Loch Ness monster. This is perhaps disingenuous since, unlike Nessie, there are strong theoretical grounds for 'believing' in rebound effects and good empirical evidence that they exist. The real issue is whether rebound effects routinely lead to backfire. Discovery of this 'monster' could have a profound effect on academia, government and civil society alike.

However, despite the range of arguments and evidence cited by both 'believers' and 'non-believers', the Khazzoom-Brookes 'monster' remains frustratingly elusive. This report has examined the theoretical arguments employed by the proponents of the K-B postulate, the empirical evidence that they cite and a wide range of other evidence that could potentially be used to either support or contest their case. But despite the number and range of studies covered, the report includes very few estimates of the economy-wide rebound effect.[113] The complexity of the economic system makes such estimates at best extremely difficult. As a consequence, authors such as Brookes rely upon 'suggestive' and indirect sources of evidence that embody a range of contested theoretical assumptions.

One reason why the debate on the rebound effect is so inconclusive is the lack of clarity over basic definitional issues, such as the identification of the appropriate independent variable. For example, there is a lack of clarity over the appropriate definition of energy efficiency, the system boundary to which it should apply, the appropriate method for aggregating different energy types and the extent to which energy efficiency improvements should be considered independently of any associated improvements in the productivity of labour and capital. As an illustration, the Bessemer process improved the energy efficiency of steel making, but also provided a revolutionary new process technology that manufactured a much improved product (mild steel) at lower cost. Historical improvements in energy efficiency are likely to have derived mainly from innovations such as these, rather than from investments designed solely to improve energy efficiency. This report has consequently paid a great deal of attention to definitional issues and to locating the rebound debate within the wider economic literature on productivity and economic growth. It has demonstrated how the particular definitions and theoretical assumptions employed can greatly influence the conclusions that are drawn. For example:

- The persistent neglect of changes in energy quality have led many studies to overstate the degree of improvement in aggregate energy intensity that has been achieved since the oil price shocks of the 1970s and to incorrectly attribute this to improvements in the thermodynamic efficiency with which individual fuels are used (Cleveland*, et al.*, 2000).
- Rebound effects may be hidden by the use of physical or economic measures of energy efficiency applicable to wide system boundaries, such as energy/GDP ratios. Aggregate indices of energy efficiency trends may also fail to distinguish the relative impact of price induced factor substitution and technical change. As a result,

[113] *Technical Report 4* reviews the estimates of economy-wide rebound effects that are available from CGE modelling. But even here, only a handful of studies are available, the results are very sensitive to the parameter values chosen and the methodologies rely upon questionable assumptions.

improvements in aggregate measures of energy efficiency may not necessarily be inconsistent with the existence of backfire following the introduction of individual, energy-efficient technologies (Saunders, 2000b).

- The results of econometric and modelling investigations of the rebound effect appear to be extremely sensitive to the choice of functional form for the relevant production function. Several commonly used production functions are effectively useless for investigating the rebound effect since they invariably predict backfire (Saunders, 2007).
- It is misleading to treat capital, labour and energy inputs as independent, since capital and labour require energy for their provision. While technical change and factor substitution may reduce energy consumption within one system boundary, they will increase energy consumption elsewhere in the economy.
- New technologies may often enhance the productivity of capital and labour inputs, as well as energy. From a consumer perspective, new technologies may reduce expenditure on more than energy consumption alone. This is one reason why the rebound effects from a particular technology may not be small simply because the share of energy in total costs is small.

7.2 Does Nessie exist?

One interpretation of the K-B postulate is that *all* cost-effective energy efficiency improvements increase energy consumption above where it would be without those improvements. Strong evidence is required to defend such a bold and counterintuitive claim. The main conclusion from this review is that such evidence does not exist. The theoretical and empirical evidence cited by authors such as Brookes is suggestive, only indirectly relevant to the rebound effect and flawed in a number of respects. For example:

- Brookes cites Sam Schurr's work in support of the postulate, but this applies primarily to the causal effect of shifts to higher quality fuels, rather than improvements in thermodynamic efficiency. Also, the patterns Schurr uncovered may not be as 'normal' as Brookes suggests.
- Brookes and Saunders also cite econometric evidence on energy-using technical change in support of the postulate. But these empirical results vary widely between different sectors, countries and time periods and are sensitive to minor changes in econometric specification. Moreover, the assumption of a fixed bias in technical change is flawed and the failure to check for the presence of cointegration or to account for changes in energy quality means that the estimates could be either biased or spurious. Moreover, even if energy-using technical change were to be consistently found, the relationship between this finding and the K-B postulate remains unclear.
- Brookes 'accommodation' argument is based upon a theoretical model that is unconventional in approach and difficult to interpret and calibrate. The model rests on the assumption that the income elasticity of 'useful' energy demand is always greater than unity, but the 1972 study on which this claim is based has not been updated. Contemporary research on Environmental Kuznets Curves has not tested this hypothesis, since useful energy consumption is not employed as the independent variable.

- Saunders work is entirely theoretical and therefore limited by the assumptions embedded in the relevant models, such as perfectly competitive markets, constant returns to scale and exogenous technical change. Many of these assumptions have been superseded by developments in endogenous growth theory and related fields, but these approaches have yet to be used to investigate the rebound effect. The results of neoclassical models are highly sensitive to both the choice of production function and the value of key parameters, which make them a questionable basis for empirical predictions.

Nevertheless, the arguments and evidence for the K-B postulate deserve far more serious attention than they have received to date. Many of the arguments and sources of evidence cited in favour of the postulate point to energy playing a significantly more important role in economic growth than is generally assumed. If these arguments are correct, they have pessimistic implications regarding the scope for decoupling economic growth from increased energy consumption. For example:

- Much technical change appears to be energy-using, in that it acts to increase aggregate measures of energy intensity, despite improvements in thermodynamic efficiency at the micro level. The conventional assumptions regarding the sign and magnitude of the *AEEI* parameter in energy-economic modelling may therefore be flawed.
- Reductions in aggregate energy/GDP ratios appear to be largely explained by structural change, changes in relative prices and improvements in the quality of energy inputs. Energy saving technical change at the micro-level may have contributed much less to such reductions than is commonly assumed.
- When changes in energy quality are taken into account, there appears to be no evidence for a turning point in the relationship between GDP and energy consumption. Hence, historical experience provides no support for the claim that economic growth can be maintained alongside absolute reductions in energy consumption.

The evidence for these claims is, again, far from definitive. But taken together, they provide further support for the argument that the economy-wide rebound effect from energy efficiency improvements is frequently large. Such a possibility has been dismissed by a number of leading energy economists (Howarth, 1997; Lovins, 1998; Laitner, 2000; Schipper and Grubb, 2000), but it becomes more plausible *if* it is accepted that energy efficiency improvements are routinely associated with improvements in capital, labour and hence total factor productivity. If this is the case, then rebound effects need not necessarily be small just because the share of energy in total costs is small.

Disagreement over this point divides proponents and opponents of the K-B postulate in much the same way that it divides ecological and neoclassical perspectives on the contribution of energy to economic growth. The neoclassical assumption appears to be that capital, labour and energy inputs have *independent* and *additive* effects on economic output, with any residual increase being attributed to exogenous technical change. Recent work on endogenous growth theory has modified these assumptions, but still attributes a relatively minor role to energy. In contrast, the ecological assumption appears to be that capital, labour and energy inputs have *synergistic* and *multiplicative* effects on economic output and that the increased availability of low-cost, high-quality energy sources provides a necessary condition for technical change. Moreover, improvements in thermodynamic conversion efficiency can be taken as a measurable proxy for technical change.

One approach to reconciling the neoclassical and ecological perspectives is therefore to investigate whether, how and to what extent different types of energy efficiency improvement at different levels of aggregation are associated with improvements in the productivity of other inputs and with improvements in total factor productivity. This question can be approached from an economy-wide perspective (Schurr*, et al.*, 1960), a sectoral perspective (Boyd and Pang, 2000) or a micro perspective (Pye and McKane, 1998), but at present the evidence base for all three remains weak.

7.3 Finding Nessie

In their insightful review of the literature on the rebound effect, Allan *et al* (2006) conclude that:

> "……the extent of rebound and backfire effects is always and everywhere an empirical issue. It is simply not possible to determine the degree of rebound and backfire from theoretical considerations alone, notwithstanding the claims of some contributors to the debate. In particular, theoretical analysis cannot rule out backfire. Nor, strictly, can theoretical considerations alone rule out the other limiting case, of zero rebound, that a narrow engineering approach would imply. However in an open economy such as the UK……….the zero rebound case seems extremely unlikely." (Allan*, et al.*, 2006)

The conclusion of this report is the same: since neither the theoretical arguments nor the available empirical evidence appear sufficient to claim that backfire is outcome of *all* energy efficiency improvements, the K-B postulate must be considered strictly incorrect. However, much of the evidence reviewed in this report points to economy-wide rebound effects being larger than is assumed by either policy makers or energy analysts (e.g. Schipper (1998)). To explore this question further, it is essential to move away from polarised debates over the validity of the K-B postulate and to investigate, both theoretically and empirically, the circumstances under which economy-wide rebound effects are more or less likely to be large. For example, on the basis of this review we may speculate that that rebound effects should be larger for energy efficiency improvements associated with:

- energy intensive production sectors compared to non-energy intensive sectors;
- energy supply industries compared to energy users;
- core process technologies compared to non-core technologies;
- technologies in the early stages of diffusion compared to those in the later stages; and
- technologies that improve capital and labour productivity, compared to those that do not.

Rebound effects may be particularly large for the energy efficiency improvements associated with 'general-purpose technologies', such as steam engines, railroads, automobiles and computers. General purpose technologies (GPTs) are defined by Lipsey*, et al* (2005) as technologies that:

- have a wide scope for improvement and elaboration;
- are applicable across a broad range of uses;
- have potential for use in a wide variety of products and processes; and
- have strong complementarities with existing or potential new technologies.

GPTs may begin as relative crude technologies with a limited number of uses, such as Savory's steam engine or the mainframe computers of the 1950s. But as they diffuse throughout the economy they evolve into much more complex forms with dramatic increases in their economic efficiency, their range of uses, the variety of outputs that they help produce and the range of new product and process technologies that they enable. Jevons steam engine example provides a paradigmatic illustration of the diffusion of GPTs in the 19th-century, as does the introduction of electric motors into manufacturing in the early 20th century and information and communication technologies in the late 20th century. Bresnahan and Trajtenberg (1995) observe that:

> "Most GPTs play the role of enabling technologies, opening up new opportunities rather than offering complete, final solutions. For example, the productivity gains associated with the introduction of electric motors in manufacturing were not limited to a reduction in energy costs. The new energy sources fostered the more efficient design of factories, taking advantage of the new-found flexibility of electric power." (Bresnahan and Trajtenberg, 1995)

Of course, this is precisely the observation made by Schurr and cited subsequently by Brookes as evidence for the K-B postulate. Hence, the empirical illustrations that best support the K-B postulate appear to those associated with GPTs. So perhaps the key to unpacking the K-B postulate is to distinguish the energy efficiency improvements associated with GPTs (the 'monsters') from other forms of energy efficiency improvement (mere 'waves'). As yet, however, the energy literature has not addressed the concept of GPTs, while the GPT literature has not addressed the specific importance of energy.

We may speculate that K-B postulate is more likely to hold for the energy efficiency improvements associated with GPTs such as steam turbines, electric motors and motor vehicles, particularly when these are used by producers and occur at an early stage of development and diffusion of the technology. The opportunities offered by these technologies have such long term and significant effects on innovation, productivity and economic growth that overall energy consumption is increased, rather than reduced. In contrast, the K-B postulate is less likely to hold for dedicated energy efficiency technologies such as improved thermal insulation, particularly when these are used by consumers or when they play a subsidiary role in economic production. These technologies have much smaller effects on productivity and economic growth, with the result that overall energy consumption is reduced.

A possible implication is that climate policy should focus on encouraging dedicated energy efficient technologies, rather than improving the energy efficiency of GPTs. However, these categories are poorly defined and the boundaries between them are blurred.[114] For example, lighting technology is not normally considered a GPT, but work by Fouquet and Pearson (2006) suggests that large rebound effects may also be present in this case, at least over the long term (Section 2.1). Similarly, the electric motor may have been a GPT in the early part of the 20th century, but given its widespread diffusion today, more energy efficient motors could still be effective in reducing aggregate energy consumption. Moreover, even if 'monsters' can be distinguished from 'mere waves', continued economic growth is likely to depend upon the diffusion and improvement of new types of GPT. If these account, directly or indirectly, for a significant proportion of overall energy consumption and if they depend upon the continued availability of low-cost, high-quality sources of energy, the gloomy conclusions cited above regarding the difficulty of decoupling energy consumption from continued economic growth may still apply.

In conclusion, while it is unlikely that all energy efficiency improvements lead to backfire, we still have much to learn about the factors that make backfire more or less likely. Put another way, we still can't distinguish 'monsters' from 'waves' and still know very little about the size and behaviour of the monster population!

[114] The innovation literature is itself ambiguous over the appropriate definition of a GPT and on the particular technologies which qualify. Also, the GPT concept overlaps with other related concepts such as technological regimes (Nelson and Winter, 1982).

References

Acemoglu, D., (2002), 'Directed Technical Change', *Review of Economic Studies*, **69**(4), 781-809.

Adams, F. G. and P. Miovic, (1968a), 'On relative fuel efficiency and the output elasticity of energy consumption in Western Europe', *The Journal of Industrial Economics*, **17**(1), 41-56.

Adams, F. G. and P. Miovic, (1968b), 'On relative fuel efficiency and the output elasticity of energy consumption in Western Europe', *Journal of Industrial Economics*, **17**, 41-76.

Aghion, P. and P. W. Howitt, (1998), *Endogenous growth theory*, MIT Press, Cambridge, Mass and London.

Ahern, J., (1980), *The exergy method of energy system analysis*, John Wiley, New York.

Alcorn, J. A. and G. Baird, (1996), *Embodied energy analysis of New Zealand building materials-methods and results*, Proceedings of the Embodied Energy: The Current State of Play Seminar, Deakin University, Geelong, Australia, 28-29 November

Alcott, B., (2005), 'Jevons' paradox', *Ecological Economics*, **54**(1), 9-21.

Alcott, B., (2006), 'Assessing Energy Policy: Should Rebound Count?' MPhil Thesis, Department of Land Economy, University of Cambridge, Cambridge.

Alfredsson, E. C., (2004), ''Green' consumption - no solution for climate change', *Energy*, **29**, 513-24.

Allan, G., N. Hanley, P. G. McGregor, J. Kim Swales, and K. Turner, (2006), 'The macroeconomic rebound effect and the UK economy', Final report to the Department Of Environment Food and Rural Affairs, Department Economics, University of Strathclyde, Strathclyde.

Ang, B. W., (1999), 'Decomposition methodology in energy demand and environmental analysis', in *Handbook of Environmental and Resource Economics*, J.C.J.M van der Bergh ed, Edward Elgar, Cheltenham, UK.

Ang, B. W., (2006), 'Monitoring changes in economy-wide energy efficiency: From energy-GDP ratio to composite efficiency index', *Energy Policy*, **34**(5), 574-82.

Arrow, K., Chenery, H., Minhas, B., and R., Solow, (1961), 'Capital-labour substitution and economic efficiency', *The Review of Economics and Statistics*, **43**(3), 225-50.

Arrow, K. J., (1962), 'The economic implications of learning by doing', *Review of Economic Studies*, **29**, 155-73.

Ayres, R., (1998a), 'Eco-thermodynamics: economic and the second law', *Ecological Economics*, **26**, 189-209.

Ayres, R., (1998b), 'Towards a disequilibrium theory of endogenous economic growth', *Environmental and Resource Economics*, **11**(3-4), 289-300.

Ayres, R., (1999), 'Products as service carriers: should we kill the messenger or send it back?' ZEF-EN-1999-1-D, United Nations University.

Ayres, R., L. W. Ayres, and V. Pokrovsky, (2005), 'On the efficiency of US electricity usage since 1900', *Energy*,

30(7), 1092-145.

Ayres, R. and B. Warr, (2002a), 'Useful work and information as drivers of growth', Working Paper, INSEAD, Fontainebleau.

Ayres, R. U., (2001), 'The minimum complexity of endogenous growth models: : the role of physical resource flows', *Energy*, **26**(9), 817-38.

Ayres, R. U., (2002), 'Resources, scarcity, technology and growth', INSEAD Working Paper No. 2002/118/EPS/CMER, Centre for the Management of Environmental Resources, INSEAD, Fontainebleau.

Ayres, R. U., L. W. Ayres, and B. Warr, (2003), 'Exergy, power and work in the US economy, 1900-1998', *Energy*, **28**(3), 219-73.

Ayres, R. U. and J. C. J. M. van den Bergh, (2005), 'A theory of economic growth with material/energy resources and dematerialization: Interaction of three growth mechanisms', *Ecological Economics*, **55**(1), 96-118.

Ayres, R. U. and B. Warr, (2002b), 'Two paradigms of production and growth', CMER Working Paper, INSEAD, Fontainbleau, France.

Ayres, R. U. and B. Warr, (2005), 'Accounting for growth: the role of physical work', *Structural Change and Economic Dynamics*, **16**(2), 181-209.

Ayres, R. U. and B. Warr, (2006), 'REXS: a forecasting model for assessing the impact of natural resource consumption and technological change from economic growth', *Structural Change in Economic Dynamics*, **17**(3), 329-78.

Azar, C., and H., Dowlatabadi, (1999), 'A review of technical change in assessment of climate policy', *Annual Review of Energy*, **24**, 513-44.

Azar, C. and H. Dowlatabadi, (1999), 'A review of technical change in assessment of climate policy', *Annual Review of Energy and the Environment*, **24**, 513-44.

Barro, R. J. and X. Sala-I-Martin, (1995), *Economic Growth*, McGraw-Hill, New York.

Beattie, B. R. and C. R. Taylor, (1993), *The Economics of Production*, John Wiley & Sons, New York, Chichester.

Beaudreau, B. C., (1995a), 'The impact of electric power on productivity : A study of US manufacturing 1950-1984', *Energy Economics*, **17**(3), 231-36.

Beaudreau, B. C., (1995b), 'The impact of electric power on productivity: a study of US manufacturing 1950-1984', *Energy Economics*, **17**(Three), 231-36.

Beaudreau, B. C., (1998), *Energy and Organisation: group and distribution re-examined*, Greenwood Press, Westport, CT.

Beaudreau, B. C., (2005), 'Engineering and economic growth', *Structural Change and Economic Dynamics*, **16**(2), 211-20.

Benders, R. M. J., R. Kok, H. C. Moll, G. Wiersma, and K. J. Noorman, (2006), 'New approaches for household energy conservation—In search of personal household energy budgets and energy reduction options', *Energy Policy*, **34**(18), 3612-22.

Bentley, R. W., (2002), 'Global oil & gas depletion: an overview', *Energy Policy*, **30**(3), 189-205.

Berndt, E. R., (1978), 'Aggregate energy, efficiency and productivity measurement', *Annual Review of Energy*, **3**, 225-73.

Berndt, E. R., (1990), 'Energy use, technical progress and productivity growth: a survey of economic issues', *The Journal of Productivity Analysis*, **2**, 67-83.

Berndt, E. R., and M., S., Khaled, (1979), 'Parametric Productivity Measurement and Choice among Flexibloe Functional Forms', *Journal of Political Economy*, **87**(6), 1220-45.

Berndt, E. R., Morrison, C., J., and G., C., Watkins, (1981), 'Dynamic models of energy demand: an assessment and comparison', in *Modeling and Measuring Natural Resource Substitution*, E. R. Berndt, and B., C., Field ed, MIT press, Cambridge , MA.

Berndt, E. R. and D. O. Wood, (1975), 'Technology, Prices, and the Derived Demand for Energy', *Review of Economics and Statistics*, **57**(3), 259-68.

Berndt, E. R. and D. O. Wood, (1979), 'Engineering and econometric interpretations of energy-capital complementarity', *American Economic Review*, **June**, 259-68.

Berndt, E. R. and D. O. Wood, (1982), 'The specification and measurement of technical change in US manufacturing', in *Advances in the Economics of Energy and Resources: volume for*, J.R. Moroney ed, JAI Press, Greenwich, CT.

Berndt, E. R. and D. O. Wood, (1986), 'Energy price shocks and productivity growth in US and UK manufacturing', *Oxford Review of Economic and policy*, **2**(3), 1-31.

Berndt, E. R. a. D., Hesse, (1986), 'Measuring and assessing capacity utilisation in the manufacturing sectors of nine OECD countries', *European Economic Review*, **30**(5), 961-89.

Besen, S. M. and L. L. Johnson, (1982), 'Comment on "Economic implications of mandated efficiency standards for household appliances"', *Energy Journal*, **3**(1), 110-16.

Binswanger, H. P., (1974), 'The measurement of technical change biases with many factors of production', *The American Economic Review*, **64**(6), 964-76.

Birol, F. and J. H. Keppler, (2000), 'Prices, technology development and the rebound effect', *Energy Policy*, **28**(6-7), 457-69.

Boyd, G. A. and J. X. Pang, (2000), 'Estimating the linkage between energy efficiency and productivity', *Energy Policy*, **28**, 289-96.

Brännlund, R., T. Ghalwash, and J. Nordstrom, (2007), 'Increased energy efficiency and the rebound effect: effects on consumption and emissions', *Energy Economics*, **29**(1), 1-17.

Bresnahan, T. and M. Trajtenberg, (1995), 'General-purpose technologies: engines of growth', *Journal of Econometrics*, **65**, 83-108.

Brookes, L., (1990a), 'Energy efficiency and the greenhouse effect', *Energy and the Environment*, **14**, 318–33.

Brookes, L., (2004), 'Energy efficiency fallacies—a postscript', *Energy Policy*, **32**(8), 945-47.

Brookes, L. G., (1972), 'More on the output elasticity of energy consumption', *The Journal of Industrial Economics*, **21**(1), 83-92.

Brookes, L. G., (1978), 'Energy policy, the energy price fallacy and the role of nuclear energy in the UK', *Energy Policy*, **6**(2), 94-106.

Brookes, L. G., (1984), 'Long-term equilibrium effects of constraints in energy supply', in *The Economics of Nuclear Energy*, Leonard Brookes and H. Motamen eds, Chapman and Hall, London.

Brookes, L. G., (1990b), 'The greenhouse effect: the fallacies in the energy efficiency solution.' *Energy Policy*, **18**(2), 199-201.

Brookes, L. G., (1992), 'Energy efficiency and economic fallacies - a reply', *Energy Policy*, **20**(5), 390-92.

Brookes, L. G., (1993), 'Energy efficiency fallacies - the debate concluded', *Energy Policy*, **21**(4), 346-47.

Brookes, L. G., (2000), 'Energy efficiency fallacies revisited', *Energy Policy*, **28**(6-7), 355-66.

Brown, M. A., M. D. Levine, S. W., and J. G. Koomey, (2001), 'Scenarios for a clean energy future', *Energy Policy*, **29**(14), 1179-96.

Bullard, C. W., P. S. Penner, and D. A. Pilati, (1978), 'Net energy analysis: handbook for combining process and input–output analysis', *Resources and Energy 1 (1978) (3), pp. 267–313*, **1**(3), 267–313.

Casals, X. G., (2006), 'Analysis of building energy regulation and certification in Europe: their role, limitations and differences', *Energy and Buildings*, **38**, 381-92.

Chalkley, A. M., E. Billett, and D. Harrison, (2001), 'An investigation of the possible extent of the respending rebound effect in the sphere of consumer products', *The Journal of Sustainable Product Design*, **1**(163-170).

Chambers, R. G., (1988), *Applied Production Analysis: A Dual Approach*, Cambridge University Press, Cambridge.

Chang, K. P., (1994), 'Capital-Energy Substitution and Multilevel Ces Production Function', *Energy Economics*, **16**(1), 22-26.

Chapman, P., (1974), 'Energy costs: a review of methods', *Energy Policy*, **2**(2), 91-103.

Chiang, A. C., (1984), *Fundamental methods of Mathematical Economics*, McGraw-Hill, London.

Chontanawat, J., L. Hunt, C.,, and R. Pierse, (2006), 'Causality between energy consumption and GDP: evidence from 30 OECD and 78 non-OECD countries', Working Paper, Department of Economics, University of Surrey, Guildford.

Chontanawat, J., Hunt, L., C., and R., Pierse, (2006), 'Causality between energy consumption and GDP: evidence from 30 OECD and 78 non-OECD countries', Department of Economics, University of Surrey, Guildford, UK.

Christensen, L. R., D. W. Jorgensen, and L. L. Lau, (1975), 'Transcendental Logarithmic Production Frontiers', *The Review of Economics and Statistics*, **55**(1), 28-45.

Christensen, L. R., D. W. Jorgenson, and L. J. Lau, (1969), 'Transcendental Logarithmic Production Frontiers', *Review of Economics and Statistics*, **55**(1), 28-45.

Cipolla, C. M., (1962), *The Economic History of World Population*, Pelican Books, London.

Clark, J. S. and C. E. Youngblood, (1992), 'Estimating duality models with biased technical change: a time series approach', *American Journal of Agricultural Economics Economic Review*, **74**, 353-60.

Cleveland, C. J., (1992), 'Energy quality and energy surplus in the extraction of fossil fuels in the U.S.' *Ecological Economics*, **6**(2), 139-62.

Cleveland, C. J., (1999), 'Biophysical Economics - from Physiocracy to Ecological Economics and Industrial Ecology', in *Bioeconomics and Sustainability: Essays in Honour of Nicholas Gerogescu-Roegen*, J. Gowdy and K. Mayumi eds, Edward Elgar, Cheltenham.

Cleveland, C. J., R. Costanza, C. A. S. Hall, and R. K. Kaufmann, (1984), 'Energy and the US economy: a biophysical perspective', *Science*, **225**, 890-97.

Cleveland, C. J., R. K. Kaufmann, and D. I. Stern, (2000), 'Aggregation and the role of energy in the economy', *Ecological Economics*, **32**, 301-17.

Cleveland, C. J. and M. Ruth, (1997), 'When, where and by how much do biophysical limits constrain the economic process?' *Ecological Economics*, **22**, 203-23.

Cobb, C., C., and P., H., Douglas, (1928), 'A theory of production', *American Economic Review*, **18**(1), 139-65.

Common, M. and S. Stagl, (2006), *Ecological economics: an introduction*, Cambridge University Press, Cambridge.

Costanza, R., (1980), 'Embodied energy and economic valuation', *Science*, **210**, 1219-24.

Dahmus, J. B. and T. G. Gutowski, (2005), *Efficiency and production: historical trends for seven industrial sectors*, Third biennial conference of the US Society for Ecological Economics, Tacoma, WA, July 20-23

Daly, H. E., (1997), 'Georgescu-Roegen versus Solow/Stiglitz', *Ecological Economics*, **22**(3), 261-66.

Daly, H. E. and J. B. Cobb, (1989), *For the Common Good*, Beacon Press, Boston.

Dargay, J. M., (1992), *Are price & income elasticities of demand constant? The UK experience*, Oxford Institute for Energy Studies, Oxford, U.K.

David, P., A., and T., Van de Klundert, (1965), 'Biased efficiency growth and capital-labour substitution in the US, 1899-1960', *American Economic Review*, **55**, 357-94.

David, P. A. and T. van der Klundert, (1965), 'Biased efficiency growth in capital-Labour substitution in the US: 1899-1960', *American Economic Review*, **55**(357-394).

Deaton, A. and J. Mulbauer, (1980), 'An almost ideal demand system', *The American Economic Review*, **70**, 312-26.

DeCanio, S., (1997), 'Economic Modeling and the False Tradeoff Between Environmental Protection and Economic Growth', *Contemporary Economic Policy*, **15**(4).

DeCanio, S., (2003), *Economic models of climate change: a critique*, Palgrave-Macmillan, New York.

Denison, E. F., (1985), *Trends in American Economic Growth 1929-1982*, The Bookings Institute, Washington, DC.

Diewert, E., (1971), 'An application of the Shephard duality theorem: a generalised Leontief production function', *Journal of Political Economy*, **79**, 481-507.

Diewert, W. E., (1976), 'Exact and superlative index numbers', *Journal of Econometrics*, **For**, 115-45.

Diewert, W. E. and T. J. Wales, (1987), 'Flexible functional forms and global curvature conditions', *Econometrica*, **55**(1), 43-68.

Dincer, I., (2002), 'The role of exergy in energy policy making', *Energy Policy*, **30**, 137-49.

Dinda, S., (2004), 'Environmental Kuznets Curve Hypothesis: A Survey', *Ecological Economics*, **49**(4), 431-55.

Easterly, W. and R. Levine, (2001), 'It's not factor accumulation: stylised facts and growth models', *The World Bank Economic Review*, **15**(2), 117-219.

Eddington, A. S., (1927), *The Nature of the Physical World*.

Einhorn, M., (1982), 'Economic implications of mandated efficiency standards for household appliances: an extension.' *Energy Journal*, **3**(1), 103-09.

Elliot, R. N., S. Laitner, and M. Pye, (1997), *Considerations in the estimation of costs and benefits of industrial energy efficiency projects*, 32nd Annual Intersociety Energy Conversion Engineering Congress, Honolulu,

EMF 4 Working Group, (1981), 'Aggregate elasticity of energy demand', *Energy Journal*, **2**(2), 37-75.

Engle, R. F. and C. W. J. Granger, (1997), 'Cointegration and error correction: representation, estimation and testing', *Econometricaa*, **55**, 251-76.

Fabricant, S., (1954), 'Economic progress and economic change', 34th Annual Report, National Bureau of Economic Research, New York.

Feather, J., (2005), 'The Khazzoom-Brookes' postulate', Unpublished manuscript.

Feist, W., (1996), *Life-cycle energy balances compared: low energy house, passive house, self-sufficient house*, Proceedings of the International Symposium of CIB W67, Vienna, Austria,

Fisher, F. M., (1993), *Aggregation: aggregate production functions and related topics*, MIT Press, Cambridge, MA.

Fouquet, R. and P. Pearson, (2006), 'Seven centuries of energy services: the price and use of light in the United Kingdom (1300-1700)', *The Energy Journal*, **27**(1), 139-77.

Frondel, M. and C. M. Schmidt, (2004), 'Facing the truth about separability: nothing works without energy', *Ecological Economics*, **51**, 217-23.

Geller, H., P. Harrington, A. H. Rosenfeld, S. Tanishima, and F. Unander, (2006a), 'Policies for increasing energy efficiency: thirty years of experience in OECD countries', *Energy Policy*, **34**, 556-73.

Geller, H., P. Harrington, A. H. Rosenfeld, S. Tanishima, and F. Unander, (2006b), 'Policy for increasing energy efficiency: 30 years of experience in OECD countries', *Energy Policy*, **34**, 556-73.

Georgescu-Roegen, N., (1971), *The Entropy Law and the Economic Process*, Harvard University Press, Cambridge, MA.

Giampietro, M., (2006), 'Comments on "The energetic metabolism of the European Union and the United States" by Haberl and colleagues: theoretical and practical considerations on the meaning and usefulness of traditional energy analysis', *Journal of Industrial Ecology*, **10**(74), 173-85.

Granger, C., (1980), 'Testing for causality: a personal viewpoint', *Journal of Economic Dynamics and Control*, **2**(329-342).

Granger, C. W. J., (1969), 'Investigate causal relation with econometric and cross-sectional methods', *Econometrica*, **37**, 424-38.

Granger, C. W. J. and P. Newbold, (1974), 'Spurious regressions in econometrics', *Journal of Econometrics*, **2**(111-120).

Greening, L. A. and D. L. Greene, (1998), 'Energy use, technical efficiency, and the rebound effect: a review of the literature', Report to the U.S. Department of Energy, Hagler Bailly and Co., Denver.

Grubb, M., J. Kohler, and D. Anderson, (2002), 'Induced technical change in energy and environmental modelling: analytic approaches and policy implications', *Annual Review of Energy and the Environment*, **27**, 271-308.

Grubb, M. J., (1990), 'Energy efficiency and economic fallacies', *Energy Policy*, **18**(8), 783-85.

Grubb, M. J., (1992), 'Reply to Brookes', *Energy Policy*, 783-85.

Gullickson, W. and M. J. Harper, (1987), 'Multi factor productivity in US manufacturing 1949-1983', *Monthly Labour Review*, (October 18-28).

Hall, C. A. S., C. J. Cleveland, and R. K. Kaufmann, (1986), *Energy and Resource Quality: The Ecology of the Economic Process.*, Wiley Interscience, New York.

Hall, C. A. S., D. Lindenberger, R. Kummel, T. Kroeger, and W. Eichhorn, (2001), 'The need to reintegrate the natural sciences with economics', *Bioscience*, **51**(8), 663-73.

Hammond, G. p., (2004), 'Towards sustainability: energy efficiency, thermodynamic analysis and the two cultures', *Energy Policy*, **32**, 1789-98.

Hammond, G. P. and A. J. Stapleton, (2001), 'Exergy analysis of the United Kingdom energy system', *Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy*, **215 (A2)**, 141-62.

Harvey, A. C., (1989), *Forecasting, structural time series models and the Kalman filter*, Cambridge University Press, Cambridge.

Hendry, D. F. and G. E. Mizon, (1978), 'Serial correlation as a convenient simplification, not a nuisance: a comment on the study of the demand for money by the Bank of England', *The Economic Journal*, **88**(549-563).

Henly, J., H. Ruderman, and M. D. Levine, (1988), 'Energy savings resulting from the adoption of more efficient appliances: a follow-up', *Energy Journal*, **9**(2), 163-70.

Herendeen, R., (1978), 'Total energy cost of household consumption in Norway, 1973', *Energy*, **3**(5), 615-30.

Herendeen, R. and J. Tanak, (1976), 'The energy cost of living', *Energy*, **1**(2), 165-78.

Herring, H. and M. Elliot, (1990), 'Letter to the editor', *Energy Policy*, **18**(8), 786.

Hicks, J. R., (1932), *The theory of wages*, London.

Hogan, W., and A., Manne, (1977), 'Energy-Economy Interactions: The Fable of the Elephant and the Rabbit?' *EMF Modeling Forum*, **1**(3).

Hogan, W., and Dale, Jorgenson, (1991), 'Productivity Trends and the cost of reducing CO2 emissions', *The Energy Journal*, **12**(1), 68-85.

Hogan, W. W. and D. W. Jorgenson, (1991), 'Productivity trends and the cost of reducing CO_2 emissions', *The Energy Journal*, **12**(1), 67-85.

Hogan, W. W. and A. S. Manne, (1970), 'Energy-the economy interactions: a fable of the elephant and the rabbit', Energy and the Economy: Report 1 of the Energy Modelling Forum, Stanford University.

Hogan, W. W. and A. S. Manne, (1977), 'Energy-the economy interactions: a fable of the elephant and the rabbit', Energy and the Economy: Report 1 of the Energy Modelling Forum, Stanford University.

Hong, N. V., (1983), 'Two measures of aggregate energy production elasticities', *The Energy Journal*, **4**(2), 172-77.

Howarth, R. B., (1997), 'Energy efficiency and economic growth', *Contemporary Economic Policy*, **15**(4), 1.

Hunt, L. C., G. Judge, and Y. Ninoyima, (1999), *Modelling technical progress: an application of the stochastic trend model to UK energy demand*, British Institute of Energy Economics Academic Conference, St. John's College, Oxford,

IBRD, (1992), *World Development Report: development and the environment*, Oxford University Press, New York.

Jevons, W. S., (1865), 'The Coal Question: Can Britain Survive?' in *The Coal Question: An Inquiry Concerning the Progress of the Nation, and the Probable Exhaustion of Our Coal-mines*, A.W. Flux ed, Augustus M. Kelley, New York.

Johansson, O. and L. Schipper, (1997), 'Measuring long-run automobile fuel demand: separate estimations of vehicle stock, mean fuel intensity, and mean annual driving distance', *Journal of Transport Economics and Policy*, **31**(3), 277-92.

Jones, C. J., (2001), *Introduction to Economic Growth*, W.W. Norton, New York, London.

Jones, H. G., (1975), *An introduction to modern theories of economic growth*, Van Nostrand Reinhold, London.

Jorgenson, D., (1998), 'Econometric Methods for Applied General Equilibrium Analysis', in *Growth v.2: Energy, the Environment and Economic Growth*, MIT Press, Cambridge MA.

Jorgenson, D., and B., Fraumeni, (1981), 'Relative prices and Technical Change', in *Modeling and Measuring Natural Resource Substitution*, E. R. Berndt, Field, B.C., ed, The MIT Press.

Jorgenson, D. and B. Fraumeni, (1981a), 'Relative prices and Technical Change', in *Modeling and Measuring Natural Resource Substitution.*, E.R. Berndt and B.C. Field eds, The MIT Press, Boston.

Jorgenson, D. and B. Fraumeni, (1981b), 'Relative prices and technical change', in *Modeling and Measuring Natural Resource Substitution*, E.R. Berndt and B.C. Field eds, The MIT Press.

Jorgenson, D. W., (1984), 'The role of energy in productivity growth.' *The Energy Journal*, **5**(3).

Jorgenson, D. W., (1996), *Growth Vol. 2 Energy, the Environment, and Economic Growth*, MIT Press, Boston.

Jorgenson, D. W. and Z. Griliches, (1967), 'The explanation of productivity change', *Review of Economic Studies*, **34**(3), 249-82.

Jorgenson, D. W., M. S. Ho, and K. Stiroh, (2005), *Productivity, Volume 3: Information Technology and the American Growth Resurgence*, The MIT press.

Kaldor, N., (1968), 'Capital Accumulation and Economic Growth', in *The Theory of Capital*, F A Lutz and D. C. Hague eds, Macmillan, London.

Kanyama, A. C., R. Engstrom, and R. Kok, (2006), 'Indirect and direct energy requirements of city households in Sweden: options for reduction, lessons from modelling', *Journal of Industrial Ecology*, **9**(1-2), 221-35.

Kaufmann, R. K., (1992), 'A biophysical analysis of the energy/real GDP ratio: implications for substitution and technical change', *Ecological Economics*, **6**(1), 35-56.

Kaufmann, R. K., (1994), 'The relation between marginal product and price in US energy markets: implications for climate change policy', *Energy Economics*, **16**(2), 145-58.

Kaufmann, R. K., (2004), 'The mechanisms of autonomous energy efficiency increases: a cointegration analysis of the US energy/GDP ratio', *The Energy Journal*, **25**(1), 63-86.

Kaufmann, R. K. and I. G. Azary-Lee, (1990), *A biophysical analysis of substitution: does substitution save energy in the US forest products industry?*, Ecological economics: its implications for forest management and research, Proceedings of a Workshop held in St Paul, Minnesota, April 2-6

Kemfert, C., (1998), 'Estimated substitution elasticities of a nested CES production function approach for Germany', *Energy Economics*, **20**(3), 249-64.

Kemfert, C. and H. Welsch, (2000), 'Energy-capital-labour substitution and the economic effects of CO_2 abatement: evidence for Germany', *Journal of Policy Modelling*, **22**(Six), 641-60.

Kemfert, C. a. H., Welsch, (2000), 'Energy-Capital-Labour substitution and the economic effects of CO2 abatement: Evidence for Germany', *Journal of Policy Modelling*, **22**(6), 641-60.

Khazzoom, J. D., (1980), 'Economic implications of mandated efficiency in standards for household appliances', *Energy Journal*, **1**(4), 21-40.

Klump, R., and H., Preissler, (2000), 'CES production functions and economic growth', *Scandinavian Journal of Economics*, **102**(1), 41-56.

Kohler, J., M. Grubb, D. Popp, and O. Edenhofer, (2006), 'The transition to endogenous technical change in climate-economy models: a technical overview to the innovation modelling comparison project', *The Energy Journal*, **Special Edition on Endogenous Technological Change and the Economics of Atmospheric Stabilisation**, 17-55.

Kok, R., R. M. J. Benders, and H. C. Moll, (2006), 'Measuring the environmental load of household consumption using some methods based on input–output energy analysis: A comparison of methods and a discussion of results', *Energy Policy*, **34**(17), 2744-61.

Koomey, J. G., R. Koomey, C. Richey, S. Laitner, R. J. Markel, and C. Marnay, (1998), 'Technology and greenhouse gas emissions: an integrated scenario analysis using the LBNL-NEMS model', LBNL-42054/EPA, 430-R-98-021, Lawrence Berkley National Laboratory, Berkley, California.

Kouris, G., (1983), 'Energy consumption and economic activity in industrialised economies - a note', *Energy Economics*, 207-12.

Kraft, J. and A. Kraft, (1978), 'On the relationship between energy and GNP', *Journal of Energy and Development*, **3**(401-3).

Kummel, R., (1980), *Growth dynamics of the energy dependent economy*, Oelgeschlager Gunn and Hain, Cambridge, MA.

Kummel, R., (1982), 'The impact of energy on industrial growth', *Energy*, **7**(2), 189-203.

Kummel, R., (1989), 'Energy as a factor of production and entropy as a pollution indicator in macroeconomic modelling', *Ecological Economics*, **1**(2), 161-80.

Kummel, R., J. Henn, and D. Lindenberger, (2002), 'Capital, labor, energy and creativity: modeling innovation diffusion', *Structural Change and Economic Dynamics*, **13**(4), 415-33.

Kummel, R., D. Lindenberger, and W. Eichhorn, (2000), 'The productive power of energy and economic evolution', *Indian Journal of Applied Economics*, **8**, 231-62.

Kummel, R., W. Strassl, A. Gossner, and W. Eichhorn, (1985), 'Technical progress and energy dependent production functions', *Nationalökonomie - Journal of Economics*, **3**, 285-311.

Kydes, A. S., (1997), 'Sensitivity of energy intensity in U.S. energy markets to technological change and adoption.' U.S. Department of Energy, Washington DC.

Laitner, J. A., (2000), 'Energy efficiency: rebounding to a sound analytical perspective', *Energy Policy*, **28**(6-7), 471-75.

Leach, G., (1975), 'Net energy analysis — is it any use?' *Energy Policy*, **3**(4), 332-44.

Lee, C. C., (2006), 'The causality relationship between energy consumption and GDP in G-11 countries revisited', *Energy Policy*, **34**(9), 1086-93.

Lee, C. C. and C. P. Chang, (2005), 'Structural bricks, energy consumption and economic growth revisited: evidence from Taiwan', *Energy Economics*, **27**, 857-72.

Lenzen, M., (1998), 'Energy and greenhouse gas cost of living for Australia during 1993/94', *Energy*, **23**(66), 497-516.

Lenzen, M. and C. Dey, (2000), 'Truncation error in embodied energy analyses of basic iron and steel products', *Energy*, **25**(6), 577-85.

Lim, H. and C. R. Shumway, (1997), 'Technical change and model specification: US agricultural production', *American Journal of Agricultural Economics*, **79**, 543-54.

Lipsey, R., G., K. I. Carlaw, and C. T. Bekar, (2005), *Economic Transformations: General Purpose Technologies and Long-term Economic Growth*, Oxford University Press, Oxford.

Loschel, A., (2002), 'Technological change in economic models of environmental policy: a survey', *Ecological Economics*, **43**, 105-26.

Löschel, A., (2002), 'Technological change in economic models of environmental policy: a survey', *Ecological Economics*, **43**(2-3), 105-26.

Lovins, A. B., (1998), 'Further comments on Red Herrings', Letter to the New Scientist, No. 2152, 18 September

Lovins, A. B., J. Henly, H. Ruderman, and M. D. Levine, (1988), 'Energy saving resulting from the adoption of more efficient appliances: another view; a follow-up', *The Energy Journal*, **9**(2), 155.

Lovins, A. B. and L. H. Lovins, (1997), 'Climate: making sense and making money', Rocky Mountain Institute, Old Snowmass, Colorado.

Mankiw, N. G., D. Romer, and D. N. Weil, (1992), 'A Contribution to the Empirics of Economic Growth', *The Quarterly Journal of Economics*, **107**(2), 407-37.

Manne, A., and R., Richels, (1990), 'CO2 emission limits: an economic cost analysis and probability assessment', *Energy Journal*, **11**(2), 51-74.

Manne, A., and R., Richels, (1991), 'Global CO2 Emission Reductions: the impacts of rising energy costs', *The Energy Journal*, **12**(1), 87-107.

Manne, A., and R., Richels, (1992), 'CO2 emission limits: an economic cost analysis for the USA', in *International Energy Economics*, Thomas Sterner ed, Chapman & Hall, London.

Manne, A. S. and R. G. Richels, (1990), 'CO2 emissions limits: an economic cost analysis for the USA', *The Energy Journal*, **11**(4), 51-74.

Mawson, P., K. Carlaw, and N. McLellan, (2003), 'Productivity measurement: alternative approaches and estimates', Working Paper 03/12, New Zealand Treasury, Auckland.

McFadden, D., (1963), 'Constant elasticity of substitution production functions', *Review of Economic Studies*, **30**(233-236).

Mizobuchi, K., (2006), *Estimation of the Rebound Effects caused by an improvement in energy efficiency: Japanese household consumption and emissions*, 3rd World Congress of Environmental and Resource Economics, Kyoto, Japan,

Mizobuchi, K. I., (2007), 'An empirical evidence of the rebound effect considering capital costs', Working Paper, Graduate School of Economics, Kobe University, Kobe, Japan.

Moroney, J., (1992a), 'Energy, Capital and Technological Change in the US', *Resources and Energy*, **14**, 363-80.

Moroney, J. R., (1992b), 'Energy, capital and technological change in the United States', *Resources and Energy*, **14**, 363-80.

Nelson, R. and S. G. Winter, (1982), *An Evolutionary Theory of Economic Change*, Harvard University Press, Cambridge, MA.

Neumayer, R., (2004), *Weak Versus Strong Sustainability: Exploring The Limits Of Two Opposing Paradigms*, Edward Elgar, Cheltenham.

Norsworthy, J., R., (1981), 'Energy prices, technical change and productivity growth', in *Energy, productivity and economic growth*, S. Schurr, Sonenblum, S., and D., Wood ed, Oelgeschlager, Gunn and Hain, Cambridge, MA.

Norsworthy, J. R., M. J. Harper, and K. Kunz, (1979), 'The slowdown in productivity growth: an analysis of some contributing factors', *Brookings Papers in Economic Activity*, **2**, 387-41.

Ohta, M., (1974), 'A note on the duality between production and cost functions: rates of return to scale and rate of technical progress', *Economic Studies Quarterly*, **25**, 63-65.

Olson, M., (1988), 'The productivity slowdown, the oil shocks and the real cycle', *Journal of Economic Perspectives*, **2**(4), 43-69.

Patterson, M. G., (1996), 'What is energy efficiency: concepts, indicators and methodological issues', *Energy Policy*, **24**(5), 377-90.

Paul, S. and R. N. Bhattacharya, (2004), 'Causality between energy consumption and economic growth in India: a note on conflicting results', *Energy Economics*, **26**(6), 977-83.

Perman, R. and D. I. Stern, (2003), 'Evidence from panel, unit root and cointegration tests that the environmental Kuznets curve does not exist', *Australian Journal of Agricultural and Resource Economics*, **47**, 325-47.

Pitchford, J. D., (1960), 'Growth and the Elasticity of Substitution', *Economic Record*, **36**, 491-503.

Pollak, R., A., Sickles, R., C., and T., J., Wales, (1984), 'The CES-Translog: specification and estimation of a new cost function', *Review of Economics and Statistics*, **66**, 602-07.

Pye, M. and A. McKane, (1998), *Enhancing shareholder value: making a more compelling energy efficiency case to industry by quantifying non-energy benefits*, Proceedings 1999 Summer Study on Energy Efficiency in Industry, Washington, DC,

Richmond, A. K. and R. K. Kaufmann, (2006), 'Energy prices and turning points: the relationship between income and energy use/carbon emissions', *The Energy Journal*, **27**(4), 157-79.

Romer, P. M., (1986), 'Increasing returns and long-run growth', *Journal of Political Economy*, **94**(5), 1002-37.

Rosen, M. A., (2004), 'Exergy analysis of energy systems', in *Encyclopaedia of Energy*, C.J. Cleveland ed, Elsevier Academic Press.

Rosenberg, N., (1983), 'The effects of energy supply characteristics on technology and economic growth', in *Energy, Productivity and Economic Growth*, Sam H. Schurr ed, Oulgeschlager, Gunn and Hain, Cambridge MA.

Rosenberg, N., (1989), *Energy Efficient Technologies: Past, Present and Future Perspectives*, How Far Can the World Get on Energy Efficiency Alone?,

Roy, J., J. Sathaye, A. H. Sanstad, P. Mongia, and K. Schumacher, (1999), 'Productivity trends in India's energy intensive industries', *The Energy Journal*, **20**(3), 33-61.

Royal Commission on Environmental Pollution, (2007), 'The Urban Environment', London.

Sanstad, A. H., J. Roy, and S. Sathaye, (2006), 'Estimating energy-augmenting technological change in developing country industries', *Energy Economics*, **28**, 720-29.

Sartori, I. and A. G. Hestnes, (2007), 'Energy use in the life-cycle of conventional and low-energy buildings: a review article', *Energy and Buildings*, **39**, 249-57.

Sato, K., (1966), 'On the Adjustment Time in Neo-Classical Growth Models', *The Review of Economic Studies*, **33**(3), 263-68.

Sato, K., (1967a), 'A two level elasticity of substitution production function', *Review of Economic Studies*, **34**, 201-18.

Sato, K., (1967b), 'A two-level constant-elasticity-of-substitution production function', *Review of Economic Studies*, **34**, 201-18.

Sato, K. and T. Koizumi, (1973), 'The production function and the theory of distributive shares', *The American Economic Review*, **63**(3), 484-89.

Saunders, H., (2006), 'Fuel Conserving (and Using) Production Functions', Working Paper, Decision Processes Incorporated, Danville, CA.

Saunders, H. D., (1992), 'The Khazzoom-Brookes postulate and neoclassical growth', *The Energy Journal*, **13**(4), 131.

Saunders, H. D., (2000a), 'Does predicted rebound depend on distinguishing between energy and energy services?' *Energy Policy*, **28**(6-7), 497-500.

Saunders, H. D., (2000b), 'A view from the macro side: rebound, backfire, and Khazzoom-Brookes', *Energy Policy*, **28**(6-7), 439-49.

Saunders, H. D., (2005), 'A calculator for energy consumption changes arising from new technologies', *Topics in Economic Analysis and Policy*, **5**(1), 1-31.

Saunders, H. D., (2007), 'Fuel conserving (and using) production function', Working Paper, Decision Processes Incorporated, Danville, CA.

Schipper, L., (1998), 'Editorial', *Energy Policy*, **28**, 6-7.

Schipper, L. and M. Grubb, (2000), 'On the rebound? Feedback between energy intensities and energy uses in IEA countries', *Energy Policy*, **28**(6-7), 367-88.

Schipper, L. and S. Meyers, (1992), *Energy efficiency and human activity: past trends, future prospects*, Cambridge University Press.

Schmalensee, R. T., T. M. Stocker, and R. J. Judson, (1998), 'World Carbon Dioxide Emissions: 1950-2050', *The Review of Economics and Statistics*, **80**(1), 15-27.

Schurr, S., (1982), 'Energy Efficiency and Productive Efficiency: Some Thoughts Based on the American Experience', *Energy Journal*, **3**(3), 3-14.

Schurr, S. H., (1983), 'Energy abundance and economic progress', *Annual Review of Energy*, **10**(3/4), 111-17.

Schurr, S. H., (1984), 'Energy, economic growth and human welfare', in *The Economics of Nuclear Energy.*, L. Brookes and H. Motamen eds, Chapman and Hall, London.

Schurr, S. H., (1985), 'Energy conservation and productivity growth : Can we have both?' *Energy Policy*, **13**(2), 126-32.

Schurr, S. H., B. C. Netschert, V. E. Eliasberg, J. Lerner, and H. H. Landsberg, (1960), 'Energy in the American Economy', Resources for the Future, John Hopkins University Press, Baltimore.

Schurr, S. H. and S. Sonenblum eds, (1986), *Electricity use, Productive Efficiency and Economic Growth*, Electric Power Research Institute, Palo Alto.

Sherwood, J. M., (1985), 'Engels, Marx, Malthus, and the Machine', *The American Historical Review*, **90**(4), 837-65.

Smulders, S. and M. de Nooj, (2003), 'The impact of energy conservation on technology and economic growth', *Resource and Energy economics*, **25**, 59-79.

Söllner, F., (1997), 'A reexamination of the role of thermodynamics for environmental economics', *Ecological Economics*, **22**(3), 175-201.

Solow, R., (1956), 'A contribution to the theory of economic growth', *The Quarterly Journal of Economics*, **70.**, 65-94.

Solow, R., (1957), 'Technical change and the aggregate production function', *The Review of Economics and Statistics*, **39**(3), 312-20.

Sorrell, S., E. O. Malley, S. J., and S. Scott, (2004), *The Economics of Energy Efficiency: Barriers to Cost-Effective Investment*, Edward Elgar, Cheltenham, UK.

Stern, D. I., (1993), 'Energy and economic growth in the USA : A multivariate approach', *Energy Economics*, **15**(2), 137-50.

Stern, D. I., (1997), 'Limits to substitution in the irreversibility in production and consumption: a neoclassical interpretation of ecological economics', *Ecological Economics*, **21**, 197-215.

Stern, D. I., (2000), 'A multivariate cointegration analysis of the role of energy in the US macroeconomy', *Energy Economics*, **22**, 267-83.

Stern, D. I., (2004a), 'Elasticities of substitution and complementarity', Working Paper in Economics, No. 0403, Rennsselaer Polytechnic Institute.

Stern, D. I., (2004b), 'The rise and fall of the Environmental Kuznets Curve', *World Development*, **32**(Eight), 1419-39.

Stern, D. I., (2004c), 'The rise and fall of the Environmental Kuznets Curve', *World Development*, **32**(8), 1419-39.

Stern, D. I. and C. J. Cleveland, (2004), 'Energy and economic growth', Rensselaer Working Paper in Economics, No. 0410, Rensselaer Polytechnic Institute, Troy, NY.

Sue Wing, I., (2006), 'Representing induced technological change in models for climate policy analysis', *Energy Economics*, **28**, 539-62.

Sue Wing, I., and R., Eckhaus, (2004), 'Explaining long-run changes in the energy intensity of the US economy', MIT, Cambridge MA.

Sue Wing, I. and R. S. Eckaus, (2006a), 'The Decline in U.S. Energy Intensity: Its Origins and Implications for Long-Run CO2 Emission Projections', Working Paper, Kennedy School of Government, Harvard University, Boston.

Sue Wing, I. and R. S. Eckaus, (2006b), 'The implications of the historical decline in US energy intensity for long-run CO2 emission projections', Working Paper, Join Programme on the Science and Policy of Global Change, MIT, Boston.

Sutherland, R. J., (1998), 'The impact of potential climate change commitments six industries in the USA', *Energy Policy*, **24**(4), 765-76.

Sweeney, J. L., (1984), 'The response of energy demand to higher prices: what have we learned?' *American Economic Review*, **74**(2), 31-37.

Temple, J., (2006), 'Aggregate production functions and growth economics', *International Review of Applied Economics*, **20**(3), 301-17.

Toke, D., (1990), 'Increasing energy supply not inevitable', *Energy Policy*, **18**(7), 671-73.

Toman, M. and B. Jemelkova, (2003), 'Energy and Economic Development: An Assessment of the State of Knowledge', *The Energy Journal*, **24**(4).

Tucker, M., (1995), 'Carbon dioxide emissions and global GDP', *Ecological Economics*, **15**, 215-23.

Van der Zwaan, B., C., Gerlagh, R., Klaassen, G., and L., Shrattenholzer, (2002), 'Endogenous technological change in climate change modelling', *Energy Economics*, **24**, 1-19.

Vringer, K. and K. Blok, (1995), 'The direct and indirect energy requirements of households in the Netherlands', *Energy Policy*, **23**(10), 893-910.

Wackernagel, M. and M. Rees, (1996), *Our Ecological Footprint: Reducing Human Impact on the Earth.*, New Society Publishers, Gabriola Island, BC.

Wall, G., (2004), 'Exergy', in *Encyclopaedia of Energy: Volume 2*, C.J. Cleveland ed, Elsevier Academic Publishers, Amsterdam, New York.

Wall, G. E., (1990), 'Exergy conversion in Japanese society', *Energy*, **15**(5), 435-44.

Waverman, L., (1984), 'Commentary on Chapters 4 and 5', in *Energy, productivity and economic growth*, S.H.. Schurr, S. Sonneblum and D.O. Wood eds, Oelgeschlager Gunn & Hain, Cambridge, Mass.

Webb, M. and D. Pearce, (1975), 'The economics of energy analysis', *Energy Policy*, **3**(4), 318-31.

Wei, T. Y., (2007), 'Impact of energy efficiency gains on output and energy use with Cobb–Douglas production function', *Energy Policy*, **35**, 2023-30.

Welsch, H. and C. Ochsen, (2005), 'The determinants of aggregate energy use in West Germany: factor substitution, technological change, and trade', *Energy Economics*, **27**, 93-111.

Winther, B. N. and A. G. Hestnes, (1999), 'Solar versus Green: the analysis of a Norwegian row house', *Solar Energy*, **66**(6), 387-93.

Worrell, J. A., M. Ruth, H. E. Finman, and J. A. Laitner, (2003), 'Productivity benefits of industrial energy efficiency measures', *Energy*, 1081-98.

Xiao, N., J. Zarnikau, and P. Damien, (2007), 'Testing functional forms in energy modeling: An application of the Bayesian approach to U.S. electricity demand', *Energy Economics*, **29**(2), 158-66.

Yoo, S., (2005), 'Electricity consumption and economic growth: evidence from Korea', *Energy Policy*, **33**, 1627-32.

Yoo, S. H., (2006), 'The causal relationship between electricity consumption and economic growth in the ASEAN countries', *Energy Policy*, **34**(18), 3573-82.

Zachariadis, T., (2006), 'On the exploration of causal relationships between energy and the economy', Discussion Paper 2006-05, Department of Economics, University of Cyprus.

Zarnikau, J., (1999), 'Will tomorrow's energy efficiency indices prove useful in economic studies?' *The Energy Journal*, **20**(3), 139-45.

Annex 1 Introduction to the Solow-Swan Model of Economic Growth

A1.1 The basic model

In 1956, Solow published a seminal paper on economic growth and development (Solow, 1956). For this work and his subsequent contributions to the understanding of economic growth, Solow was awarded the Nobel Prize in economics in 1987. The model provides one framework for exploring the determinants of economic growth and for explaining the differences in output levels and growth rates across countries and over time. This Annex introduces the basic framework of Solow's growth model, states the underlying assumptions and provides the conceptual basis for understanding the application of the model to the rebound effect. The notation and approach is based upon standard textbooks (Barro and Sala-I-Martin, 1995; Jones, 2001) and is simplified where possible.

The economy-wide production function may be written in general terms as:

$$Y = F(K, L) \qquad \text{(A.1.1)}$$

This function is assumed to be 'well-behaved' in that it satisfies the following standard conditions:

- The marginal products of all inputs are positive and diminishing ($\partial y / \partial i > 0; \partial^2 y / \partial i^2 < 0$). This assumption reflects the classical law of diminishing returns.
- The production function exhibits constant returns to scale: i.e. if all inputs are doubled, output will double (i.e. $F(aK, aL) = aF(K, L)$). A function of this type is said to be homogeneous of degree one.
- Inputs are scarce, meaning that their marginal products will approach infinity as input quantities approach zero (as $i \to 0$, $\partial y / \partial i \to \infty$).
- The economy is closed (no international trade) and consists of a sector producing a single homogeneous good that is consumed, invested or used for the accumulation of capital.
- The are a large number of firms in the product market and perfect competition ensures that individual firms are price takers..
- Agents have full and symmetric information.
- Production does not impose externalities, like pollution or waste by-products.

In this simple model, the labour force is assumed to be equal to the total population, with no unemployment. Firms are assumed to pay workers a wage (*w*) for each unit of labour and to pay a rental rate (*r*) for each unit of capital. Firms hire labour until the marginal product of labour is equal to the wage rate $(\partial Y / \partial L = w)$ and rent capital until the marginal product of capital is equal to the rental price $(\partial Y / \partial K = r)$. Payments to inputs completely exhaust the value of output produced so economic profits are zero.

The production function is conventionally written in terms of output per worker (*y=Y/L*) and capital per worker (*k=K/L*) as follows:

$$y = f(k) \qquad \text{(A.1.2)}$$

It is assumed that workers/consumers save a constant fraction (ς) of their combined wage and rental income, with the remainder spent on consumption (*c*). It is further assumed that economy is closed and these savings are used entirely for investment to accumulate capital.[115] It is also assumed that a constant fraction (δ) of capital stock depreciates every period, regardless of how much output is produced. This leads to an equation describing how capital accumulates over time. The change in capital stock is then determined *at any point in time* by the difference between investment (*I*) and capital depreciation (*D*):

$$\frac{dK}{dt} = \dot{K} = I - D = \varsigma Y - \delta K \qquad \text{(A.1.3)}$$

Population (and hence labour) is assumed to be growing at a constant rate, as an exponential function of time, that is:

$$L(t) = L_o e^{nt} \qquad \text{(A.1.4)}$$

So the *rate of growth* of population ($\dot{L}/L$) is given by n.[116]

Noting that $\ln k = \ln K - \ln L$ and differentiating this expression wrt time we obtain:

$$\frac{\dot{k}}{k} = \frac{\dot{K}}{K} - \frac{\dot{L}}{L} = \frac{\varsigma Y - \delta K}{K} - n$$
$$\frac{\dot{k}}{k} = \varsigma(Y/K) - (\delta + n) = \varsigma(y/k) - (\delta + n) \qquad \text{(A.1.5)}$$

Where $\dot{k} = \partial k / \partial t$. Rearranging:

$$\dot{k} = y\varsigma - (n + \delta)k \qquad \text{(A.1.6)}$$

This differential equation for the rate of capital accumulation per worker is the fundamental equation of the neoclassical growth model. The first term in the right hand side of this equation ($y\varsigma$) represents the *gross investment* per worker while the second term ($(n+\delta)k$) represents the *real* or *effective depreciation* per worker – taking into account both the rate of physical depreciation and the increase in population. The rate of change of capital per worker is therefore determined by the difference between gross investment and the real depreciation per worker.

The equations for output (A.1.2) and capital accumulation per worker (A.1.6) form the basis of the Solow model. With an assumed functional form for the production function, equation A.1.6 may be solved to express the endogenous variables (*y* and *k*) in terms of the exogenous variables (namely ς , δ , n and whatever terms are used to define the production function). If the production function meets certain assumptions, the dynamics of the equation leads to a *steady state* solution, where $\dot{k} = 0$ and the various quantities all grow at constant rates. At this point $k=k^*$ and:

---

[115] The corresponding absence of an explicit 'investment function' is one feature that distinguishes this model from earlier growth models.

[116] $\dot{L} = nL_o e^{nt} = nL$. The growth rate is given by $\dot{L}/L = n$.

$$f(k^*)\varsigma = (n+\delta)k^* \qquad \text{(A.1.7)}$$

As an illustration of this, assume that the production function takes the standard 'Cobb-Douglas' form $Y = K^{\alpha}L^{1-\alpha}$, where $o \leq \alpha \leq 1$ (this and other forms of production function are introduced in Annex 2). Rewriting in per worker terms, this becomes $y = k^{\alpha}$. If this is substituted into Equation A.1.6, the solution can be shown to take the following form:

$$k(t) = \left([k_0^{1-\alpha} - \varsigma / n]e^{-(1-\alpha)nt} + \varsigma / n\right)^{1/1-\alpha} \qquad \text{(A.1.8)}$$

Now as *n>o* and $o \leq \alpha \leq 1$, this means that $[k_0^{1-\alpha} - \varsigma / n]e^{-(1-\alpha)nt} \to 0$ as $t \to \infty$. Hence, over time the economy tends towards the steady state solution k^*:

$$k(t) \to k^* = (\varsigma / n)^{1/1-\alpha} \qquad \text{(A.1.9)}$$

Similar results follow for other 'well-behaved' production functions, provided the initial capital-labour ratio is greater than zero. In the steady state, output, labour and capital are growing, but output per worker, capital per worker and consumption per worker do not change*.* In these circumstances, the economy is said to be following a 'balanced growth' path, with all quantities growing at a constant exponential rate.

This behaviour of the basic Solow model can best be understood through a diagram (Figure A.1.1). Here, the curve *f(k)* shows output per worker (*y*) as a function of capital per worker (*k*). The gross investment per worker is represented by the curve $\varsigma f(k)$, which is proportional to the production curve *f(k)*. Consumption (*c*) is represented by the vertical distance between the production curve *f(k)* and the gross investment curve $\varsigma f(k)$ The capital-output ratio (*K/Y=k/y*) is given by the slope of a ray from the origin to the production function.

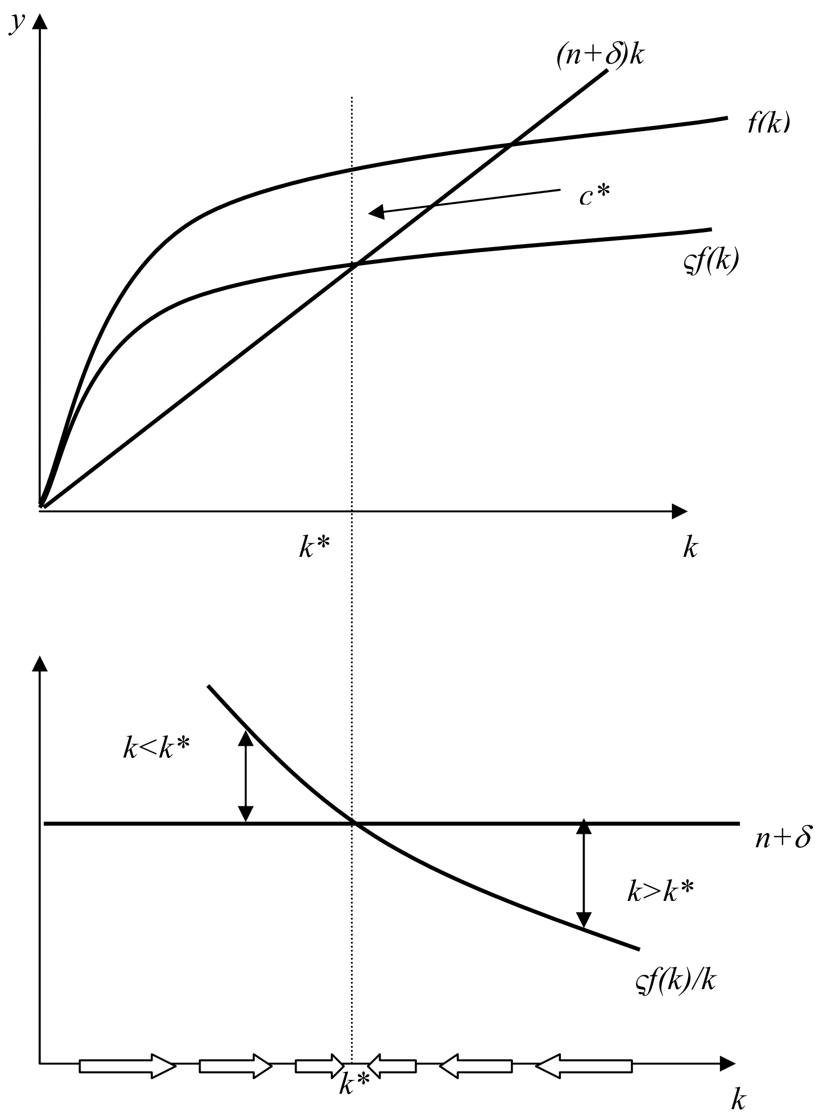
The effective depreciation line, $(n+\delta)k$, is a straight line, with a slope equal to the sum of constants *(n+δ)*. This line represents the amount of investment per worker required to keep the amount of capital per worker (*k*) constant. The difference between the gross investment curve ($\varsigma f(k)$) and the effective depreciation line ($(n+\delta)k$) represents the change in the amount of capital per worker for a particular level of capital per worker. When this change is positive and the economy is increasing the amount of capital per worker ($\dot{k} > 0$), *capital deepening* is said to occur. This occurs up to the steady state point *k*,* which satisfies (A.1.7). At this point, the amount of capital per worker is constant.

The lower part of Figure 1 illustrates the dynamics of the Solow-Swan model. Suppose an economy has an amount of capital that is *less* than k^* at a particular point in time. As the amount of investment per worker exceeds the amount needed to keep capital per worker constant, capital deepening occurs and the amount of capital increases. This will continue until the steady state point k^* is reached. Similarly, suppose that an economy has an amount of capital that is *greater* than k^* at a particular point in time. As the amount of investment per worker is less than the amount needed to keep the capital-labour ratio constant, the amount of capital per worker will decline. Again, this will continue until the steady state point k^* is reached. Hence, points to the right of the steady state (*k>k*)* imply a negative growth of capital per worker, and so *k* balances to the left towards *k**. Similarly, points to the left of the steady state *(k<k*)* imply a positive growth rate of capital per worker and so *k* balances to the right towards *k**. This dynamic balancing towards the steady state equilibrium is depicted by the arrows: the steady-state capital-labour ratio (k^*)

is stable in the sense that any other *k* will tend to approach it over time. The further the economy is from steady state, the faster *k* moves towards *k*.*

The long-run growth rate of the economy is equal to the long-run growth rate of population and is given by *n*. Notably, the long-run growth rate is entirely independent of the proportion of income saved for investment (ς). However, while an increase in this proportion will not increase the long-run growth rate of the economy it *will* increase the long-run level of output and consumption per worker - and hence overall economic welfare.

Figure A.1.1. The basic Solow diagram



The production function in the basic Solow model only includes capital and labour as inputs. However, the model is readily extendable to include energy in the production function as a third input. The production function then becomes:

$$Y = F(K, L, E) \qquad \text{(A.1.10)}$$

The function is assumed to satisfy the same assumptions as set out in the previous section. A possible objection is that 'energy' appears to be produced without any capital and labour inputs (Saunders, 1992). This important point is discussed in Section 5.3.2.

As in the basic model, it is assumed that there is no technical progress and prices are fixed in real terms. Under these conditions, it can be shown that in the steady-state equilibrium, economic output, capital inputs, labour inputs and energy inputs all grow at the same rate (Saunders, 1992). Real consumption per worker (the welfare measure in growth theory) stays fixed, as do capital per worker, energy per worker and output per worker. By implication the energy/GDP ratio (*E/Y*) also remains constant. Hence, under these simplifying assumptions, energy consumption grows in lock step with economic output (Saunders, 1992).

A1.2 The neoclassical growth model with technology

The central conclusions of this simple neoclassical growth model are that:

- The long-run rate of growth of national income is the rate of growth of population which is assumed to be an exogenous constant *n*.
- The economy invariably tends to a balanced growth path whatever the initial capital-worker ratio.
- Output per worker, capital per worker and consumption per worker are all constant in the long-run.
- Increases in the propensity to save increase the *levels* of output per worker and capital per worker but do not change the long-run *rate* of economic growth.
- Energy consumption grows in lock step with economic output and the energy/GDP ratio remains unchanged

Although it rests upon a number of extremely simplifying assumptions, this model *can* fit some of the 'stylised facts' of economic growth. However, the model does not predict the key empirical fact of economic growth, namely sustained growth in per capita income. The basic Solow model allows economies to grow for while, but not for ever. To generate sustained growth in per capita income, it is necessary to introduce technological progress into the model. Following Solow, this is traditionally achieved through the introduction of an exogenous, time-dependent multiplier *A(t)*.

Starting with the most general case, technology might contribute to the saving of all factor inputs by the same amount. This kind of technological progress is known as *neutral* (or *Hicks neutral*) and is represented as follows:

$$Y = A(t)F(K, L, E) \qquad \text{(A.1.11)}$$

Where *A(t)* is unknown function of time, which is assumed to be positive. It is commonly assumed that technology is improving at a constant rate (*g*):

$$\frac{\dot{A}}{A} = g \Leftrightarrow A = A_0 e^{gt} \qquad \text{(A.1.12)}$$

In this production function, technical progress improves the productivity of all factor inputs. However, technical progress may affect only one of the factor inputs. For example, the production function:

$$Y = F[A(t)K, L, E] \qquad \text{(A.1.13)}$$

only improves the productivity of capital inputs and so is known as *capital augmenting* (or *Solow neutral*). If the production function is of the form:

$$Y = F[K, A(t)L, E] \qquad \text{(A.1.14)}$$

then it is called *labour augmenting* (or *Harrod neutral*). If technology improves only the productivity of energy inputs the production function becomes:

$$Y = F[K, L, A(t)E] \qquad \text{(A.1.15)}$$

and so it is called *energy augmenting*. This type of technical progress may be interpreted as 'pure' energy efficiency improvements. However, most real-world technologies represent a mix of the above categories of technical change.

Most introductory textbooks explore the implications of labour augmenting technical change in the growth model (Jones, 2001). With this, the capital accumulation equation is the same as in the basic model:

$$\frac{\dot{K}}{K} - \frac{\dot{L}}{L} = \varsigma y - (n + \delta) \qquad \text{(A.1.16)}$$

However, with technical progress, capital per worker (*k*) is no longer constant in the long run. Hence, rather than setting up a differential equation for *k*, a new variable is introduced ($\tilde{k}$) representing the amount of capital per 'effective worker': $\tilde{k} = K / AL$. Since $\ln k = \ln K - \ln A - \ln L$, the growth rate of this new variable is given by:

$$\frac{\dot{\tilde{k}}}{\tilde{k}} = \frac{\dot{K}}{K} - \frac{\dot{A}}{A} - \frac{\dot{L}}{L} \qquad \text{(A.1.17)}$$

Combining this with Equation A.1.16, and noting that $\dot{A}/A = g$, we obtain:

$$\dot{\tilde{k}} = y\varsigma - (n + g + \delta)\tilde{k} \qquad \text{(A.1.18)}$$

This is a very similar equation to that for the basic model. The difference is that, first, the independent variable is 'capital per effective worker' ($\tilde{k}$) rather than capital per worker (*k*) and second, there is an additional term representing the growth rate of technology (*g*). As before, the model can be solved for a steady-state. In the steady-state, the ratios capital per effective worker ($\tilde{k}$), energy consumption per effective worker ($\tilde{e}$), output per effective worker ($\tilde{y}$), and consumption per effective worker ($\tilde{c}$) are all constant. But this is not informative of the welfare of the economy. It is *workers*, not effective workers, that receive the income and consume. Looked at on a per worker basis, in the steady-state output per worker (*y*), capital per worker (*k*), energy consumption per worker (*e*) and consumption per worker (*c*) are all growing at a the steady rate *g,* the rate of technological progress. Thus, although the steady-state growth has *effective* ratios constant, *actual* ratios are increasing. This compares with the basic model in which actual ratios were constant (e.g. *k*) and actual values were increasing (e.g. *K*). The model therefore predicts increasing income per capita – as required to meet one of the stylised facts of growth theory.

With technical progress introduced into the production function, the growth rates of output, capital and energy will be equal to each other and fixed, but at the same time *higher* than the 'natural' rate by which population grows. In other words, in the neoclassical growth model with exogenous technology we will always have:

$$\frac{\dot{Y}}{Y} = \frac{\dot{K}}{K} = \frac{\dot{E}}{E} > n \qquad \text{(A.1.19)}$$

Similar conclusions follow when other types of technical progress are introduced into the production function. However different technologies will affect the production function in different ways. For this reason, the choice of the exact functional form for a production function is crucial for the outcome of the neoclassical growth model with technology. Functional forms are the subject of the Annex 2.

Annex 2 Functional Forms for Production Functions

Economists traditionally work with a number of standard functional forms for production functions. These are briefly introduced below. Further details may be found in an standard textbooks on production economics (e.g. Beattie and Taylor (1993)).

The functional form describes the 'shape' of a production surface, representing the different combinations of factor inputs that may be used to produce a given level of output using existing technology. The form places restrictions on this shape, while the particular parameter values determined it precisely. Functional forms are chosen for their analytical tractability and may be considered as merely a convenient approximation to reality.

A2.1 The Hicks elasticity of substitution

A key difference between different functional forms is how they represent the ease with which one varying factor of production can be substituted for another, while still maintaining the same level of output. The technical measure for this is the *elasticity of substitution* (σ). The magnitude of the elasticity of substitution relative to zero is commonly used to define whether factors may be considered 'substitutes' or 'complements'. However, there are several competing definitions of the elasticity of substitution, incorporating different assumptions about which factor quantities are held fixed and which factor prices held fixed (for most measures, production output is assumed to be fixed). The value and sign of the estimated elasticity of substitution therefore depends upon which definition is used.

The most basic definition was introduced by Hicks (1932) and is termed the Hicks elasticity of substitution (σ). This measures the ease with which a decrease in one input can be compensated by an increase in another input while both output and other inputs are held constant. The definition relates to movement along an isoquant on the production surface and is a scale free measure of the curvature of the isoquant (the smaller the value of σ , the greater the curvature).

The original Hicks definition applies to a production function with only two inputs. The subsequent generalisation to multi-input production functions is sometimes termed the Direct elasticity of substitution (Chambers, 1988). In this case, the elasticity refers to a situation in which two inputs (*i* and *j*) vary, other inputs are fixed and output is fixed. For a multi-input production function that satisfies certain conditions,[117] σ_{ij} is then defined as:

$$\sigma_{ij} = -\left.\frac{\partial \ln(x_i / x_j)}{\partial \ln(f_j / f_i)}\right|_{Y\, and\, x_k\, constant\, for\, k \neq i,j} \quad \text{(A.2.1)}$$

Where $f_{x_i} = \partial f / \partial x_i$ represents the marginal productivity of input factor *i*. The ratio f_j / f_i in the denominator is termed the 'marginal rate of technical substitution' (*MRTS*).

[117] A continuous function with positive first order partial derivatives and continuous second-order partial derivatives that is quasi concave.

If other inputs are held fixed, output can only be held constant if a decrease in one input (*i*) is compensated by an increase in a second (*j*) - in other words, one factor must 'substitute' for another. Therefore, the Hicks elasticity of substitution classifies all inputs as 'substitutes'. This is different from other measures of the elasticity of substitution (e.g. the 'Allen Urzwa') where other inputs are not held fixed and where changes in the ratio of two inputs are measured with respect to changes in relative prices. In these cases, two inputs can be 'complements' in that demand for both can increase (decrease) when the price of one input falls (rises). The relationship between the various measures of the elasticity of substitution can be a source of much confusion. It is described in detail in *Technical Report 3*.

The Hicks elasticity of substitution can provide information on the effect of a change in usage of an input on the share of that input in the value of output ($s_i = x_i p_i / P_Y$).[118] It may be shown that (Sato and Koizumi, 1973):

$$\frac{\partial(s_i / s_j)}{\partial(x_i / x_j)} \overset{\geq}{\underset{\leq}{-}} 0 \text{ according to } \sigma_{ij} \overset{\geq}{\underset{\leq}{-}} 1$$

Hence, if $\sigma_{ij} > 1$ (< 1,) the share of input *i* in total costs becomes larger (smaller) relative to *j* as the usage of *i* becomes larger (smaller) relative to *j.* This observation is sometimes used as a basis to classify two inputs as either weak substitutes ($\sigma_{ij} < 1$) or strong substitutes ($\sigma_{ij} > 1$) - although the terminology here (as elsewhere) is not always consistent.

From Equation A.2.1, if the *MRTS* does not change at all with changes in the ratio x_i/x_j, it indicates that substitution is easy, because the ratio of marginal productivities of the two inputs does not change as the input mix changes. Alternatively, if *MRTS* changes rapidly for small changes in the ratio x_i/x_j, it indicates that substitution is difficult because minor variations in the input mix will have a substantial effect on the relative productivities of the two inputs.

Taking the two factor case, if σ_{ij} is large, *r* will not change much relative to the input ratio, x_i/x_j, and the isoquant will be relatively flat. On the other hand, if σ_{ij} is small, the isoquant will be sharply curved. The extremes are: a linear production function, where $\sigma_{ij} = \infty$, and a 'Leontief' (fixed proportions) production function, where $\sigma_{ij} = 0$. For a 'Cobb Douglas' production function $\sigma_{ij} = 1$, while for a 'Constant Elasticity of Substitution' (CES) production function, σ_{ij} is constant (as the name suggests) between 0 and infinity. The CES may be generalised to the multifactor case, but this places restrictive conditions on the elasticity values (McFadden, 1963). These restrictions may be avoided if the 'nested' CES form is used instead (see below), but this also has limitations because it relies on assumptions about the 'seperability' of different inputs which may not be supported by empirical evidence (Frondel and Schmidt, 2004). The 'translog' production function allows for multiple substitution possibilities between pairs of factors, so σ_{ij} can vary.

Since the extension of the Hicks definition to multi factor functions requires the assumption that other factor inputs are fixed, the practical value of this definition is limited. In practice,

[118] Under competitive market conditions, the latter is equal to the share of a factor in total input costs ($s_i = x_i p_i / C$).

it is likely that in practice any change in the ratio of two inputs will also be accompanied by changes in the levels of other inputs. Some of these inputs may be complementary with the ones being changed, whereas others may be substitutes, and to hold them constant creates a rather artificial restriction.

The econometric estimation of production functions is also prone to bias, and it is more common to estimate *cost functions* since the relevant independent variables (factor prices) can usually be assumed to be exogenous. For this reason, alternative definitions of the elasticity of substitution, based upon the cost function, are more relevant to empirical studies (McFadden, 1963; Stern, 2004a). However, this means that the elasticity of estimated by empirical studies are different from the Hicks elasticity and therefore cannot be used as a basis for determining the appropriate magnitude of the Hicks elasticity for use in theoretical and modelling studies.

A2.2 The Cobb-Douglas production function

A widely used functional form was originally suggested by the Swedish economist Knut Wicksell in 1901 and rediscovered independently by Cobb and Douglas (1928). This form, known ever since as Cobb-Douglas, is given by:

$$Y = K^{\beta} L^{\gamma} E^{1-\beta-\gamma} \qquad \text{(A.2.2)}$$

where β and γ are positive constants and $\beta + \gamma = 1$. This functional form satisfies the assumption of constant returns to scale (i.e. it is homogeneous of degree 1) and assumes constant and unitary elasticity of substitution between inputs ($\sigma = 1.0$). Hence, a percentage reduction in one input can be fully compensated by a percentage increase in another input.

A2.3 The CES production function

The limitations of Cobb Douglas functional form led Solow (1956) and Arrow *et al*. (1961) to seek an alternative that allowed flexibility in factor substitution but also allowed the elasticity of substitution to differ from unity. This led to the development of a new functional form, where the elasticity of substitution is still constant but not constrained to be unity. This functional form is known as Constant Elasticity of Substitution (CES) production function. For two inputs this functional form is given by:

$$Y = (aK^{\rho} + bL^{\rho})^{\frac{1}{\rho}} \qquad \text{(A.2.3)}$$

with *a>0, ,b*>0, *a+b=1* and $\rho = \frac{\sigma - 1}{\sigma} \Leftrightarrow \sigma = \frac{1}{1-\rho}$. With $\rho = 1$ ($\sigma \to +\infty$), the CES reduces to a simple linear production function. If $\rho \to -\infty$ ($\sigma = 0$), the CES reduces to a Leontief style production function, with zero substitutability between factors (see below). A popular case is when $\rho \to 0$ ($\sigma \to 1$), where the functional form becomes identical to the Cobb-Douglas.[119]

---

[119] This is not immediately obvious since as $\rho \to 0$, the exponent $1/\rho \to \infty$. But the equivalence may be demonstrated by the application of L/Hopital's rule: $\lim_{x \to a}[m(x)/n(x)] = \lim_{x \to a}[\dot{m}(x)/\dot{n}(x)]$ (Chiang, 1984).

The basic CES production function does not include energy as a third input. There is no universally accepted way of including energy in a CES function. In the most general case and assuming constant, non-unity elasticity of substitution between energy and other inputs, a potential form of this function can be:

$$Y = [a(bK^{-a} + (1-b)E^{-a})^{\frac{\rho}{a}} + (1-a)L^{-\rho}]^{-\frac{1}{\rho}} \qquad \text{(A.2.4)}$$

with all symbols as defined above. This more general form is called two-level CES because it contains a CES production function embedded (or 'nested') within another CES function. By analogy, energy can be included any type of nesting schemes like *(KL)E* or even *(EL)K.* (Kemfert, 1998). If more inputs are introduced (e.g. materials), or a distinction is made between different fuels (e.g. electricity and non-electricity), the CES can be transformed into a *multilevel* CES, with more than one function nested within the original one (Chang, 1994). Therefore, multilevel CES functions can be used for unlimited numbers of inputs. In principle, this form of nesting rests upon the assumption that the nests are *separable:* i.e. the marginal rate of technical substitution[120] between the inputs in one nest is independent of the level of inputs in the other nest. In practice, this assumption is not always tested and may not always hold (Frondel and Schmidt, 2004).

A special case of a nested CES is the combination of a Cobb-Douglas and a CES function. Such a functional form has been used by Hogan (1977) and Manne and Richels (1990) and takes the form:

$$Y = [a(K^{\gamma}L^{1-\gamma})^{\rho} + bE^{\rho}]^{\frac{1}{\rho}} \qquad \text{(A.2.5)}$$

with all symbols as defined above and with γ representing the share of capital in the *KL* nest. This functional form assumes unitary elasticity of substitution between labour and capital and σ elasticity of substitution between energy and non-energy inputs. This Cobb-Douglas CES can also be extended to include more inputs. In the case of Manne & Richels (1990) we have the form:

$$Y = [a(K^{\gamma}L^{1-\gamma})^{\rho} + b(E^{\delta}N^{1-\delta})^{\rho}]^{\frac{1}{\rho}} \qquad \text{(A.2.6)}$$

where *E* is electricity, *N* is non-electric energy and δ is the share of electricity in the energy nest.

A2.4 The Leontief production function

Leontief proposed a production function that imposes fixed proportions of inputs. This means that substitution possibilities between inputs, especially in the short-run are zero $(\sigma = 0.0)$. In the case of three inputs, Leontief's production function can be represented as:

120 The marginal rate of technical substitution (r) between two factor inputs (xi) provides a measure of how much of one factor is required to substitute for another – with other inputs and output fixed. For a general production fuction $Y = f(x_i, x_j, \ldots)$, it is defined as: $r = -\partial x_i / \partial x_j = f_{x_j} / f_{x_i}$ - where: $f_{x_j} = \partial f / \partial x_j$.

$$Y = \min(\frac{K}{a_K}, \frac{L}{a_L}, \frac{E}{a_E}) \quad \text{(A.2.7)}$$

Where the coefficients $a_K = \frac{K}{Y}, a_L = \frac{L}{Y}, a_E = \frac{E}{Y}$ represent constant Input-Output coefficients that transform a unit of input into a unit of output. Since these coefficients are constant, the elasticity of substitution between each pair of inputs will be zero, by definition.

A2.5 The Generalised Leontief production function

The assumption of constant elasticity of substitution, common to all the functional forms above, can be too restrictive in cases where empirical estimation is needed. This and other theoretical limitations of the CES family of functional forms have led to further generalizations of the production function.

An interesting generalization of the Leontief function was proposed by Diewert (Diewert, 1971) and allows for unlimited substitution possibilities among inputs, without explicitly imposing *a priori* conditions on the properties of the function. Diewert proposed a production function for unlimited numbers of inputs that was consistent with duality theory and could therefore be easily estimated by means of a cost function. Such a functional form for three inputs is:

$$Y = a_{KK}K + a_{LL}L + a_{EE}E + 2a_{KL}\sqrt{K}\sqrt{L} + 2a_{LE}\sqrt{L}\sqrt{E} + 2a_{EK}\sqrt{E}\sqrt{L} \quad \text{(A.2.8)}$$

where $a_{ij} = a_{ji}, a_{ij} \geq 0$ (i.e. all the *a* parameters are positive and are elements of a symmetric input matrix) is the only restriction imposed on this function. Diewert has shown that this function is flexible enough to account for non-constant returns to scale by re-writing the function as:

$$Y = H(a_{KK}K + a_{LL}L + a_{EE}E + 2a_{KL}\sqrt{K}\sqrt{L} + 2a_{LE}\sqrt{L}\sqrt{E} + 2a_{EK}\sqrt{E}\sqrt{L}) \quad \text{(A.2.9)}$$

where *H()* is a continuous, monotonically increasing function. It can be seen that in the special case $a_{ij} = 0, i \neq j$ with constant returns to scale, the Generalised Leontief can be reduced to the fixed input proportion function of Leontief. Several extensions of the Generalised Leontief function have been proposed in the literature.

A2.6 The Translog production function

Another important flexible form, is the Transcendental Logarithmic ('Translog'), production function, originally introduced by Christensen *et al* (1975). The term 'transcendental' derives from the fact that no factor can be solved for explicitly without being a function of itself (i.e. a 'transcendental' equation). Like the Generalised Leontief, the Translog does not impose any restrictions on the substitutability between different factor inputs. For three inputs, the Translog is given by:

$$\ln Y = a_K \ln K + a_L \ln L + a_E \ln E + \tfrac{1}{2}[\gamma_{KK}(\ln K)^2 + \gamma_{LL}(\ln L)^2 + \gamma_{EE}(\ln E)^2 + \gamma_{KL}(\ln K \ln L) + \gamma_{LK}(\ln L \ln K) + \gamma_{KE}(\ln K \ln E) + \gamma_{EK}(\ln E \ln K) + \gamma_{LE}(\ln L \ln E) + \gamma_{EL}(\ln E \ln L)] \quad \text{(A.2.10)}$$

With $a_K + a_L + a_E = 1$, for constant returns to scale, $\gamma_{ij} = \gamma_{ji}$ and $\sum \gamma_{ij} = \sum \gamma_{ji} = 0$, for *i,j=K,L,E*. If all the γ coefficients are zero, then it is easy to see that the above form is exactly equivalent to a Cobb-Douglas. Pollak *et al* (1984) have shown that this form can also be extended to combine a general CES with a translog, yielding a CES-translog[121]. Various other extensions of this flexible form have been proposed in the literature.

The translog is very popular in empirical work in energy economics. However, practitioners more often make use of the Translog cost function, where the factor volumes in the above equation are replaced with factor prices.

[121] A special case of the CES-Translog is the general CES, just as a special case of the Translog is the Cobb-Douglas.

Annex 3 Modelling of Technical Progress with CES production functions

While econometric studies tend to work with flexible translog production or cost functions, neoclassical growth and CGE models frequently use the nested, multifactor CES production function. This is despite the fact that the parameters of such functions are difficult to estimate empirically. This Annex explores the behaviour of these functions in neoclassical growth models, particularly with regard to their representation of technical change. Specifically, this Annex:

- Develops a proof for the condition for backfire with a nested CES production function of the Manne-Richels form.
- Clarifies the similarities and differences between variants of the CES function and compares two different approaches to representing improvements in 'energy efficiency'.
- Establishes the conditions under which a nested CES function of the Manne Richels form can model declining energy intensity (i.e. a positive *AEEI*).

A3.1 Proof of the condition for rebound with a nested CES production function

We analyse here a Cobb Douglas *KL* production function nested within a CES production function (*(KL)E)* incorporating three types of factor augmenting technical change:

$$Y = [a((\upsilon_K K)^{\beta}(\upsilon_L L)^{1-\beta})^{\rho} + b(\upsilon_E E)^{\rho}]^{1/\rho} \quad \text{(A.3.1)}$$

The *KL* nest is commonly referred to as 'value added'. Using the chain rule, the partial derivative of output with respect to energy is given by:

$$\frac{\partial Y}{\partial E} = b\upsilon_E^{\rho}E^{\rho-1}[a((\upsilon_K K)^{\beta}(\upsilon_L L)^{1-\beta})^{\rho} + b(\upsilon_E E)^{\rho}]^{\frac{1}{\rho}-1} \quad \text{(A.3.2)}$$

The term in brackets is actually output (*Y*) in the $(1-\rho)$ power:

$$\frac{\partial Y}{\partial E} = b\upsilon_E^{\rho}E^{\rho-1}Y^{1-\rho}$$
$$\frac{\partial Y}{\partial E} = b\upsilon_E^{\rho}(\frac{Y}{E})^{1-\rho} \quad \text{(A.3.3)}$$

According to standard neoclassical assumptions, in equilibrium, the marginal product is equal to the price of energy (p_E):

$$\frac{\partial Y}{\partial E} = p_E = b\tau_E^{\rho}(\frac{Y}{E})^{1-\rho} \quad \text{(A.3.4)}$$

Solving for energy, in terms of output and energy price, we have:

$$E = (\frac{p_E}{b})^{\frac{1}{\rho-1}} \upsilon_E^{\frac{\rho}{1-\rho}} Y \quad \text{(A.3.5)}$$

The demand for energy therefore depend upon output, energy prices, the energy augmenting multiplier the elasticity of substitution ($\sigma = 1/1-\rho$) and the parameter *b* which represents the share of income going to energy ($\upsilon_E E$). Now we can substitute for energy in the production function in (A.3.1) to obtain an expression for output in terms of capital and labour inputs and energy prices:

$$Y = [a((\upsilon_K K)^\beta (\upsilon_L L)^{1-\beta})^\rho + b(\upsilon_E (\frac{p_E}{b})^{\frac{1}{\rho-1}} \upsilon_E^{\frac{\rho}{1-\rho}} Y)^\rho]^{1/\rho}$$

$$Y = [a((\upsilon_K K)^\beta (\upsilon_L L)^{1-\beta})^\rho + b(\upsilon_E^{\frac{1}{1-\rho}} (\frac{p_E}{b})^{\frac{1}{\rho-1}} Y)^\rho]^{1/\rho}$$

Solve for *Y*:

$$Y^\rho - b(\upsilon_E^{\frac{1}{1-\rho}} (\frac{p_E}{b})^{\frac{1}{\rho-1}} Y)^\rho = a((\upsilon_K K)^\beta (\upsilon_L L)^{1-\beta})^\rho$$

Collecting terms:

$$Y^\rho [1 - b\upsilon_E^{\frac{\rho}{1-\rho}} (\frac{p_E}{b})^{\frac{\rho}{\rho-1}}] = a((\upsilon_K K)^\beta (\upsilon_L L)^{1-\beta})^\rho$$

Doing some simplifications within the brackets and rearranging:

$$Y = \frac{(\upsilon_K K)^\beta (\upsilon_L L)^{1-\beta} a^{\frac{1}{\rho}}}{\left[1 - b(\frac{p_E}{b})^{\frac{\rho}{\rho-1}} (\upsilon_E)^{\frac{\rho}{1-\rho}}\right]^{1/\rho}} \quad \text{(A.3.6)}$$

With more simplifications in the denominator, we finally obtain:

$$Y = \frac{(\upsilon_K K)^\beta (\upsilon_L L)^{1-\beta} a^{\frac{1}{\rho}}}{\left[1 - (\frac{\upsilon_E}{p_E})^{\frac{\rho}{1-\rho}} b^{\frac{1}{1-\rho}}\right]^{1/\rho}} \quad \text{(A.3.7)}$$

Equations. (A.3.5) and (A.3.7) are the two equations needed to conduct spreadsheet simulations of the effect of factor augmenting technical progress in the neoclassical growth model, in the spirit of Saunders (1992).

From (A.3.5), the economy-wide energy/output ratio is given by:

$$\frac{E}{Y} = (\frac{p_E}{b})^{\frac{1}{\rho-1}} (\upsilon_E)^{\frac{\rho}{1-\rho}}$$

(A.3.8)

By taking the partial derivative of this with respect to the energy augmenting technology multiplier (υ_E), we can derive an expression for how energy augmenting technical change affects the overall energy/output ratio:

$$\frac{\partial(\frac{E}{Y})}{\partial \upsilon_E} = \frac{\rho}{1-\rho} (\frac{p_E}{b})^{\frac{1}{\rho-1}} \frac{\upsilon_E^{\frac{\rho}{1-\rho}}}{\upsilon_E}$$

$$\frac{\partial(\frac{E}{Y})}{\partial \upsilon_E} = \frac{\rho}{1-\rho} \frac{E}{Y} \frac{1}{\upsilon_E}$$

(A.3.9)

Expressing this in elasticity terms gives:

$$\eta_{\upsilon_E}(\frac{E}{Y}) = \frac{\partial(\frac{E}{Y})}{\partial \upsilon_E} \frac{\upsilon_E}{(\frac{E}{Y})} = \frac{\rho}{1-\rho}$$

(A.3.10)

Or:

$$\eta_{\upsilon_E}(\frac{E}{Y}) = \sigma - 1$$

(A.3.11)

In other words, for this form of CES production function, the elasticity of the energy/output ratio with respect to the energy augmenting technical change multiplier is equal to the Hicks elasticity of substitution between energy and 'value added', minus one. This equation is consistent with Saunders's claim regarding the importance of the elasticity of substitution in determining the propensity to backfire with a CES. If $\sigma > 1$ then the energy augmenting technical change will increase overall energy intensity (backfire). Conversely, if $\sigma < 1$, energy augmenting technical change will reduce overall energy intensity.

Saunders considered this proof sufficient to investigate the propensity for CES functions to backfire. However, it does not fully account for the substitution possibilities between energy and other inputs in a cost minimising framework. As discussed below, with suitably modified assumptions the model can show also energy savings when the elasticity of substitution is greater than unity.

It should also be noted that the above proof uses a version of the CES function that *excludes* the multiplier for neutral technical progress (υ_N) that Saunders (1992) included in his general formulation of the CES, namely:

$$Y = \upsilon_N [a((\upsilon_K K)^{\beta} (\upsilon_L L)^{1-\beta})^{\rho} + b(\upsilon_E E)^{\rho}]^{1/\rho}$$

Following the same procedure as above, it is possible to derive an demand function for energy that *includes* the neutral weight:

$$E = (\frac{p_E}{b})^{\frac{1}{\rho-1}}(\upsilon_E\upsilon_N)^{\frac{\rho}{1-\rho}}Y \qquad \text{(A.3.12)}$$

Similarly, it is possible to derive an equation for output:

$$Y = \frac{\upsilon_N(\upsilon_K K)^{\beta}(\upsilon_L L)^{1-\beta}a^{\frac{1}{\rho}}}{\left[1-(\frac{\upsilon_E\upsilon_N}{p_E})^{\frac{\rho}{1-\rho}}b^{\frac{1}{1-\rho}}\right]^{1/\rho}} \qquad \text{(A.3.13)}$$

These two formulae are the correct ones for the form of the CES that Saunders's quotes in his 1992 paper. However, his simulations did not use this, but instead used Equations A.3.5 and A.3.7. The properties of the CES are such that the function can be equivalently written either with or without the neutral technology multiplier and the former is easier to apply. Equations A.3.5 and A.3.7 are used for the simulations described in Annex 3.3.

A3.2 Variants of the CES function

In the 50 years since Solow's (1956) introduction of the CES production function, there have been several variants of its form (Table A.3.1) and different interpretations of its components from different authors. These differences can make it difficult to interpret different modelling studies and to compare their results. This section discusses the similarities and differences between these variants and highlights the implications for the representation of technical change in CES production functions

Table A.3.1 Variants of the CES production function

Author	**Function**
Basic CES: two inputs only:	
Solow (1956)	$Y = (aK^{\rho} + L^{\rho})^{\frac{1}{\rho}}$
Pitchford (1960)	$Y = (aK^{\rho} + bL^{\rho})^{\frac{1}{\rho}}$
Arrow *et al* (1961)	$Y = C[aK^{\rho} + (1-a)L^{\rho}]^{\frac{1}{\rho}}$
David and van der Klundert (1965)	$Y = [(AK)^{\rho} + (BL)^{\rho}]^{\frac{1}{\rho}}$
Nested CES: more that two inputs	
Sato (1967b)	$Y = [a(bK^{\theta} + (1-b)L^{\theta})^{\frac{\rho}{\theta}} + (1-a)E^{\rho}]^{\frac{1}{\rho}}$
Hogan & Manne (1977)	$Y = [a(cK^{\gamma}L^{1-\gamma})^{\rho} + bE^{\rho} +]^{\frac{1}{\rho}}$
Manne & Richels (1992)	$Y = [a(K^{\gamma}L^{1-\gamma})^{\rho} + b(E^{\delta}N^{1-\delta})^{\rho}]^{\frac{1}{\rho}}$
Saunders (1992)	$Y = \upsilon_N[a((\upsilon_K K)^{\beta}(\upsilon_L L)^{1-\beta})^{\rho} + b(\upsilon_E E)^{\rho}]^{1/\rho}$

Source: adapted from Klump (2000)

In Solow's original formulation (Table A.3.1), the interpretation of the parameter *a* is unclear. Klump and Preissler (2000) comment that *a* could be interpreted as an 'acceleration factor', similar to the one used in a Leontief production function, or as 'a measure of the relative share of each factor in total income, as in Cobb-Douglas production function (implying 0<*a*<1). In this, as in all other variations of the CES, the parameter ρ is a transformation of the Hicks elasticity of substitution ($\rho = \sigma / 1 - \sigma$) - introduced in Annex 2.

Pitchford (1960) provided a more general formulation of the CES (see Table A.3.1), referring to *a* as a 'constant attached to capital' and introducing a comparable parameter *b* as a multiplier of labour inputs – and therefore implying that the parameters refer to relative factor shares. According to Pitchford, both terms depend upon the substitution parameter ρ.

Arrow *et al.* (1961) introduced the standard formulation of the CES function, which includes the substitution parameter (ρ) and an "efficiency" parameter *C* which represents neutral technical progress. In Arrow *et al*'s formulation, the parameter *a* is termed the 'distribution parameter' and restricted to be less than unity (as in a Cobb Douglas), implying that this parameter reflects relative value shares. Arrow *et al.* show that, *if technical change is neutral*, the ratio of income going to labour as wages (*w*) to that going to capital as rent (*r*) is given by:

$$\frac{wL}{rK} = \frac{1-a}{a}\left(\frac{K}{L}\right)^{\rho}$$

While this expression is independent of the efficiency parameter (*C*), it would not hold over time if technical change was non-neutral. Hence, equating the distribution parameters with the 'value share' of each factor is only valid in the context of neutral technical change.

In order to incorporate non-neutral technical change, David and van der Klundert (1965) proposed another version of the CES (see Table A.3.1). In this version, the coefficients *A* and *B* represent the factor augmentation multipliers (or 'technology multipliers', or 'efficiency levels') for capital and labour. The products *AK* and *BL* may then be interpreted as 'effective capital' and 'effective labour' inputs respectively (see Section 3.3.3). These two multipliers are assumed to be positive, but their growth rates can be either positive or negative.

David and van der Klundert (1965) relate these multipliers to the original Hicks definition of biased technical change (introduced in Section 3.3.3).[122] Under this definition, technical change is said to labour saving when the growth rate of the marginal product of labour is less than that the growth rate of the marginal product of capital ($[\partial \ln(\partial Y / \partial L /) / \partial t] - [(\partial \ln(\partial Y / \partial K /) / \partial t] < 0$ (i.e. new technology lowers the marginal product of labour relative to that of capital). For labour saving technical change, David and van der Klundert (1965) derive the following relationship between the growth rate of marginal products and the technology multipliers:

$$\frac{\partial \ln(\partial Y / \partial L)}{\partial t} - \frac{\partial \ln(\partial Y / \partial K)}{\partial t} = \frac{\sigma - 1}{\sigma}\left[\frac{(\partial \ln B)}{\partial t} - \frac{(\partial \ln A)}{\partial t}\right] < 0 \qquad \text{(A.3.16)}$$

[122] See also Acemoglu (2002).

With labour augmenting technical change, the growth rate of B exceeds that of A ($\partial \ln B / \partial t \geq \partial \ln A / \partial t$), so the expression in brackets on the right-hand side of this equation will be positive. Hence, the inequality will only be satisfied if the elasticity of substitution (σ) is less than unity (so $(\sigma - 1/\sigma) \leq 0$). In these circumstances, the two factors are sometimes termed 'weak substitutes' (see Annex 2). Hence, factor augmenting technical change will *only* be the same as the Hicks definition of factor saving technical change if the Hicks elasticity of substitution is less than unity (i.e. capital and labour are weak substitutes). Put another way, the bias of technical change (using the Hicks definition) depends upon *both* the relative magnitude of the factor augmentation multipliers *and* the elasticity of substitution between the two factors.

All the above versions of the CES function are for two inputs - capital and labour. Sato (1967a) proposed an alternative to the two input CES for multiple inputs (see Table A.3.1), termed a two (or multi)-level CES. This production function can accommodate unlimited inputs through the use of different 'nesting' schemes, although the behaviour will be sensitive to the particular scheme chosen. It also relies upon the notion of 'seperability' between the factors in separate nests (Frondel and Schmidt, 2004). Separability is commonly used within production theory to justify the *nesting*, of different inputs. The assumption is that producers in engage in a two-stage decision process: first optimising the combination of inputs within each nest, and then optimising the combination of nests required to produce the final output. Two factors may only be legitimately grouped within a nest if they are separable from factors outside of the nest. For example, a *(KL)E* nesting structure requires that capital and labour are 'separable' from energy. This means that the marginal rate of technical substitution between capital and labour is unaffected by the price of energy. In practice, the assumption of seperability is frequently not supported by empirical data.

Sato did not comment explicitly on how technical progress can be included in this framework. His idea, however, was picked up by Hogan and Manne (1977), who defined a two level CES (see Table A.3.1) incorporating energy inputs and assuming that the elasticity of substitution between capital and labour is unity (i.e. a Cobb-Douglas sub-function for 'value added'). This functional form contains an 'efficiency parameter' (*c)* for the subfunction together with the distribution parameters *a* and *b*, following the version of Pitchford.

A revised version of the CD-CES of Hogan & Manne was included in the ETA-MACRO model, whilst yet another version was used by Manne & Richels (1990) to predict the costs of a carbon emissions limits in their 2100 model. This version (see Table A.3.1) included *two* Cobb-Douglas sub-functions - one for 'value added' (capital and labour) and one for energy (electricity and non-electricity) - nested within a CES function. In this version, the parameters *a* and *b* denote the relative value shares of energy and non-energy inputs.

Manne & Richels (1990) use a *negative* growth rate of parameter *b* to represent energy efficiency improvements – which they refer to as the *AEEI*.[123] As discussed in Section 3.3.4, this definition of the *AEEI* is not the same as the one that adopted in this report (namely: $AEEI = \partial \ln(Y/E) / \partial t$). With the Manne & Richels approach, energy efficiency improvements are modelled as a gradual *reduction* in amount of energy required to provide a unit of output – as represented by the parameter *b*. Manne and Richel's increased energy efficiency scenario uses a growth rate of –1% per year for the parameter *b*.

[123] See also Hogan and Jorgenson (1991).

Finally, we arrive at the version of the CES use by Saunders (1992). This is a revised version of the Manne-Richels nested CES that includes *both* 'distribution parameters' to weight the relative share of the two inputs in the spirit of Arrow *et al.* (1961) *and* 'efficiency parameters' to represent factor augmentation in the spirit of David and van der Klundert (1965). As a result, Saunders formulation potentially combines two different approaches to modelling energy efficiency improvements, namely:

- The Manne and Richels approach of reducing the distribution parameter for energy (a negative growth rate for parameter *b*)
- The David and van der Klundert approach of increasing the technology multiplier for energy (a positive growth rate for parameter υ_E)

aunders (1992) uses a positive growth rate for the energy technology multiplier (υ_E) to model improvements in energy efficiency. As shown below, this approach is only identical to Manne and Richels assumption of a negative growth rate for parameter *b* if the elasticity of substitution is less than unity. Saunders' also refers to the *AEEI* parameter in the Manne and Richels study as positive, although it is implemented through a negative growth rate for the distribution parameter *b* ($\partial \ln b / \partial t \leq 0$). The declining energy intensity that results could also be simulated through a positive growth rate for the energy technology multiplier *provided* $\sigma \leq 1$. It also corresponds to a positive value for our definition of the *AEEI* ($AEEI = \partial \ln(Y/E)/\partial t$ - i.e. declining energy intensity) which, in this framework, may be influenced by changes in either *b* or υ_E . This illustrates how the lack of consistency in both the definition and the sign of the *AEEI* parameter can be a source of confusion.

David and van der Klundert's (1965) requirement that the elasticity of substitution be less than unity for a CES to depict input augmenting technical progress is still valid for nested CES functions. This is shown by Kemfert and Welsch (2000), who demonstrate that a nested CES can only model input saving technical progress if the elasticity of substitution between the nesting schemes is less than unity (i.e. the two nests are 'weak substitutes'). This condition can also be demonstrated in our case. Taking Equation A.3.3, we have for the marginal product of the CES in Saunders (1992):

$$\frac{\partial Y}{\partial E} = b\upsilon_E^{\rho}(\frac{Y}{E})^{1-\rho} \qquad \text{(A.3.17)}$$

Under competitive equilibrium, the value of the marginal product must equal its price so that the demand function for energy can be written as:

$$E = b^{\sigma}(\frac{p_Y}{p_E})^{\sigma}(\upsilon_E)^{\sigma-1}Y \qquad \text{(A.3.18)}$$

As mentioned above, Manne-Richels assumed a negative growth rate for *b* in the above equation while Saunders assumed a positive growth rate for υ_E . Both rates are exponential $b = e^{-\beta t}$ and $\upsilon_E = e^{\lambda t}$ (where $\beta \geq 0$ and $\lambda \geq 0$). Since CES functions from both authors can be traced back to the same basic form, as in (A.5.1), there has to be a relationship between the two different ways of representing technology.

Transforming (2.8.3) into logs yields:

$$\ln E = \sigma \ln b + \sigma \ln(\frac{p_Y}{p_E}) + (\sigma - 1)\ln \upsilon_E + \ln Y$$

Or:

$$\ln(\frac{E}{Y}) = \sigma \ln b + \sigma \ln(\frac{p_Y}{p_E}) + (\sigma - 1)\ln \upsilon_E \qquad \text{(A.3.19)}$$

If we differentiate this expression with respect to the energy saving weight we obtain another proof for equation A.3.11. But in this case we are interested in the long-run growth rates so we differentiate with respect to time to obtain:[124]

$$\frac{\partial \ln(\frac{E}{Y})}{\partial t} = \sigma \frac{\partial \ln b}{\partial t} + (\sigma - 1)\frac{\partial \ln \upsilon_E}{\partial t} \qquad \text{(A.3.20)}$$

Or:

$$-AEEI = (1-\sigma)\frac{\partial \ln \upsilon_E}{\partial t} - \sigma \frac{\partial \ln b}{\partial t} \qquad \text{(A.3.21)}$$

A similar expression is found in Kemfert and Welsch (2000). The growth rate of the technical weight ($\partial \ln \upsilon_E / \partial t$) is the method used for simulating the *AEEI* index in Saunders (1992), whereas the growth rate of the energy value share ($\partial \ln b / \partial t$) is the method used for simulating the *AEEI* in Manne & Richels (1992). Substituting for growth rates we have:

$$AEEI = (\sigma - 1)\lambda + \beta\sigma \qquad \text{(A.3.22)}$$

For declining energy intensity (positive *AEEI* under our definition), the term on the right-hand side must be positive.

The Manne Richels approach has $\lambda = 0$. Since $\sigma \geq 0$ and $\beta \geq 0$, then $AEEI = \sigma\beta \geq 0$. Hence, the assumption of a negative growth rate for the parameter *b* ($b = e^{-\beta t}$) will always lead to a positive *AEEI* (declining energy intensity) in this formulation.

The Saunders approach has $\beta = 0$. Since $\lambda \geq 0$, $AEEI = (\sigma - 1)\lambda$. Hence, the assumption of a positive growth rate for the parameter υ_E ($\upsilon_E = e^{\lambda t}$) will only lead to a positive *AEEI* (declining energy intensity) when $\sigma \leq 1$ in this formulation.

The value taken by the elasticity of substitution will for therefore determine the conditions under which the Manne-Richels representation of energy efficiency improvements (negative growth rate for *b*) is equivalent to that used by Saunders (positive growth rate for υ_E). There are 3 cases:

---

[124] The price ratio remains unchanged with time so its time derivative is zero.

- $\sigma < 1$*:* A negative growth rate for *b* is equivalent to a positive growth rate of the same magnitude for υ_E. In this case, energy is a 'weak substitute' to value added and technical change lowers the marginal product of energy relative to non-energy inputs. With $\sigma < 1$, the closer to unity is σ, the slower the decline in energy intensity.
- $\sigma \approx 1$: The CES reduces to a Cobb-Douglas and the growth rate of energy intensity is fixed, so neither energy augmenting technical progress nor a declining value for the distribution parameter will alter the proportion of energy in total output.
- $\sigma > 1, -\beta = -\lambda$: A negative growth rate for *b* is equivalent to a negative growth rate of the same magnitude for υ_E. In this case, energy is a 'strong substitute' to value added and technical change increases the marginal product of energy relative to non-energy inputs. Hence, when energy is a strong substitute, 'energy efficiency' has to be growing at a diminishing rate to describe declining energy intensity.

An assumption of a negative growth rate for the factor augmentation multiplier appears odd, since it seems to imply technical 'regress' rather than progress (i.e. factor productivity declining over time). However, the key point is the growth rate of the energy multiplier (υ_E) *relative to the others* (υ_L and υ_K). In practice, all the factor multipliers are likely to have positive growth rates, but if the growth rate of energy is *less* than the others, the relative productivity of energy is declining.

Consistent with the above, we find that models that depict energy saving technical progress invariably assume a value for the Hicks elasticity of substitution between energy and other inputs that is less than unity (Manne, 1991; 1992; Azar, 1999; Kemfert, 2000; Loschel, 2002; Löschel, 2002; Van der Zwaan, 2002). The robustness of these insights is further tested below using a neo-classical growth model comparable to that used by Saunders (1992).

A3.3 Sensitivity analysis with a nested CES function

This section uses simulations with a neoclassical growth model to establish the conditions under which the nested CES function can model declining energy intensity (i.e. a positive *AEEI*). The simulations adopt a similar approach to Saunders (1992) and focus upon varying the factor augmenting parameters in the production function.[125]

The previous section demonstrated that our definition of the *AEEI* depends upon both the growth rate of the energy augmenting multiplier (λ_E) and the elasticity of substitution (σ). It is important to stress that, although the nominal value of the efficiency parameter is strictly positive, its growth rate can be either positive or negative.

Saunders' general formulation of the CES aggregate production function:

$$Y = \upsilon_N [a((\upsilon_K K)^{\beta} (\upsilon_L L)^{1-\beta})^{\rho} + b(\upsilon_E E)^{\rho}]^{1/\rho} \quad \text{(A.3.23)}$$

[125] Throughout this section we have benefited greatly from personal communication with Harry Saunders.

can be re-written with exponential terms for technical weights:

$$Y = e^{|\lambda_N|t}[a((e^{|\lambda_K|t}K)^{\beta}(e^{|\lambda_L|t}L)^{1-\beta})^{\rho} + b(e^{|\lambda_E|t}E)^{\rho}]^{1/\rho} \quad \text{(A.3.24)}$$

where $|\lambda_i|$ is the absolute value of the growth rate for effective input *i*. The growth rate is given here in absolute terms because we assume that it may be increasing or decreasing, corresponding to either an increase or decrease in the ratio of effective to real factor inputs. An assumption of a negative growth rate for the factor augmentation multiplier appears odd, since it seems to imply technical 'regress' rather than progress (i.e. factor productivity declining over time). However, the aim is to isolate the growth rate of one factor augmentation multiplier *relative to the others*. In practice, all the factor multipliers are likely to have positive growth rates, but if the growth rate of one factor multiplier is less than the others, the relative productivity of that factor is declining. In the simulations of energy augmenting technical progress undertaken here, the growth rates of the technology multipliers for capital and labour are set to zero. Energy augmenting technical change is then represented by a positive growth rate for the multiplier for energy inputs ($\lambda_E > 0$), while what may be called 'energy diminishing' technical change is represented by a negative growth rate ($\lambda_E < 0$).

The basis for the simulations is Equations (A.3.15) and (A.3.16). Each simulation corresponds to a period of 100 years and makes different assumption about either the growth rate of the factor augmenting multipliers (λ_i), and/or the elasticity of substitution (σ). In all cases, the growth rate of the factor augmenting multipliers is set to either 1.2% per year ($|\lambda_i|$=0.012) or zero. The nominal (constant) price for energy is assumed to be 0.05, with capital accounting for 30% of value added. Other prices and initial quantities are normalised. Labour is assumed to grow at a "natural" rate of 3% per year.

Simulations are conducted for: a) neutral technology ($\lambda_N = +0.012$); b) energy augmenting technology ($\lambda_E = +0.012$); and c) energy 'diminishing' technology ($\lambda_E = -0.012$). In each case, six different values of the elasticity of substitution were examined, ranging from 0.2 to 1.8. Table A.3.2 summarises the results

Table A.3.2 Simulation Results (% annual growth rates)

Technical Weight	Annual Growth Rates (%)										
	Y	K	L	E σ=0.2	E σ=0.5	E σ=0.8	E σ=1.2	E σ=1.5	E σ=1.8	E σ≈1	E/Y
λ_N = 1.2%											
Neutral	4.8	4.8	3	4.8	0	0	0	0	0	0	0
	4.8	4.8	3	0	4.8	0	0	0	0	0	0
	4.8	4.8	3	0	0	4.8	0	0	0	0	0
	4.9	4.9	3	0	0	0	4.9	0	0	0	0
	4.9	4.9	3	0	0	0	0	4.9	0	0	0
	4.9	4.9	3	0	0	0	0	0	4.9	0	0
	4.9	4.9	3	0	0	0	0	0	0	4.9	0
λ_E = 1.2%											
Energy Augmenting	3	3	3	2	0	0	0	0	0	0	-0.9
	3	3	3	0	2.4	0	0	0	0	0	-0.6
	3.1	3.1	3	0	0	2.8	0	0	0	0	-0.2
	3.1	3.1	3	0	0	0	3.4	0	0	0	0.2
	3.3	3.3	3	0	0	0	0	3.9	0	0	0.6
	3.4	3.4	3	0	0	0	0	0	4.4	0	0.9
	3.1	3.1	3	0	0	0	0	0	0	3.1	0
λ_E = -1.2%											
Energy Diminishing	2.9	2.9	3	3.9	0	0	0	0	0	0	0.9
	2.9	2.9	3	0	3.5	0	0	0	0	0	0.6
	2.9	2.9	3	0	0	3.2	0	0	0	0	0.2
	2.9	2.9	3	0	0	0	2.7	0	0	0	-0.2
	2.9	2.9	3	0	0	0	0	2.3	0	0	-0.6
	2.9	2.9	3	0	0	0	0	0	1.9	0	-0.9
	2.9	2.9	3	0	0	0	0	0	0	2.9	0

The upper part of the table shows the results from neutral technical progress. This is found to increase the growth rates of output, capital and energy by 4.8% per year when energy is weak substitute to value added ($\sigma < 1$), and 4.9% per year when energy is a strong substitute ($\sigma > 1$). Changes in the technology neutral weight leave aggregate energy intensity (*E/Y*) unchanged (i.e. *AEEI*=0).

The middle part of the table shows the results from energy augmenting technical progress. When energy is weak substitute to value added ($\sigma < 1$), aggregate energy intensity is found to decrease (*AEEI*>0). Therefore, the model simulates energy savings compared to a scenario without technical progress. The size of the decline in aggregate energy intensity (or the magnitude of energy savings) depends on the elasticity of substitution. The larger the elasticity of substitution, the smaller the decrease in aggregate energy intensity. Results from this case are consistent with Saunders (1992) and Kemfert & Welsch (2000).

When energy is a strong substitute to value added ($\sigma > 1$), energy augmenting technical progress induces substitution *towards* energy and therefore increases overall energy intensity (backfire). In this case, a larger elasticity of substitution leads to a greater increase in energy intensity. Therefore we can say that the further the elasticity of substitution diverts from unity, the greater the impact on overall energy consumption.

When the elasticity of substitution is equal to unity (a Cobb-Douglas production function), energy augmenting technical progress has no effect on aggregate energy intensity. Hence, the Cobb-Douglas cannot depict either 'energy-saving' or 'energy-using' technical progress. Output, energy and capital all grow at the same rate (3.1%), which is slightly greater than without technical progress (3%).

The third part of the table shows the results for 'energy diminishing' technical progress. Here, the technical multiplier for energy grows at a negative rate of –1.2% per year. In all cases, output (and capital) grow slower (2.9%/year) than without technical 'progress' (3%/year). In this case, the cost of effective energy increases and energy becomes relatively unattractive as compared with other inputs. When energy is a weak substitute to value added ($\sigma < 1$), energy consumption increases relative to output and we observe increasing energy intensity. But when energy is a strong substitute to value added ($\sigma > 1$), the negative bias makes it less substitutable for other inputs, leading to a reduction in energy intensity.

Hence, in the last three lines of Table A.3.2, we observe 'energy savings' (i.e. declining energy intensity) as a result of 'deteriorating' energy efficiency ($\lambda_E < 0$). This is conditional upon energy being a strong substitute for value added ($\sigma > 1$). Of course, 'deteriorating' energy efficiency is not desirable in itself since output is reduced. But if all technology multipliers were growing at a positive rate, and if the technology multiplier for energy was growing slower than that for capital and labour, the result would be increasing output combined with declining energy intensity. The implication is that *less* efficient use of energy use could lead to actual energy savings – the inverse of the rebound effect.

The results demonstrate that, in this framework, declining energy intensity may result from either: a) energy augmenting technical progress when energy is a weak substitute to value added; or b) energy 'diminishing' technical progress when energy is a strong substitute to value added.