Wageningen University – Department of Social Science

BSc Thesis - Chair Group Environmental Economics and Natural Resources

---

The Role of Energy in the Economic Process

June 2010

Student:	Rembrandt H.E.M. Koppelaar
Program:	BSc Economics and Policy (BEB)
Supervisor:	Dr. H. P. Weikard
Thesis code:	ENR - 80912

WAGENINGEN UNIVERSITY
WAGENINGEN UR

Student ID: 850321-467-060

ABSTRACT

In economics energy has never received sufficient attention as an input factor in the process of production, based on the assumption of unlimited substitutability between energy as an input and capital and labor. According to the laws of thermodynamics this assumption is incorrect as energy is a prerequisite for any type of economic activity. Therefore, without increasing high quality inputs of energy, economic growth would cease. This is a scenario of increasing probability as: 1) The depletion of fossil fuels poses limits to increasing production, 2) Non-fossil sources are of lower energy density, and cost more to retrieve and convert than fossil energy sources making them poor substitutes. To improve understanding of the effects of fossil fuel depletion on economic growth it is suggested to look at: A) how increased efficiency of transforming energy into work, measured by total energy inputs excluding energy conversion losses, has affected economic growth, B) how technological progress has enabled the usage of increasing high quality over time. Recommended empirical areas of study are energy inputs and economic output changes across countries, and energy to economic growth causality analyses including technological progress, prices, and energy delivered in production excluding energy conversion losses.

Keywords: economic growth, production inputs, thermodynamics, capital, labor, technological progress, energy efficiency, energy conversion losses.

Rembrandt Koppelaar, Department of Social Sciences, Environmental Economics and Natural Resources, P.O Box 9101, Wageningen University, 6700 HB, Wageningen, the Netherlands

ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr. Hans Peter Weikard. His guidance, trust, and flexibility helped me in writing this thesis. Special thanks go to Dr. Hannes Kunz for his invaluable comments in reviewing my thesis and my father Prof. Dr. Henk Koppelaar for his support. The author also conveys his thanks to Ajay Gupta and Prof. Dr. Charles Hall for providing sample data on energy efficiency. Last and not least, my gratitude goes to The Oil Drum: www.theoildrum.com, the largest blog on energy analysis on the internet, from whose contributors I learned most about energy which - hopefully - is reflected in this thesis.

TABLE OF CONTENTS

1. INTRODUCTION

In the past 200 years national income across the globe has grown enormously resulting in exceptionally high standards of living in comparison to pre-industrial civilizations. This transformation became possible by extraction and transformation of the earth's resources into usable materials for consumption, and as additional inputs in the production process. The chain of extraction and industrial processing has become more diverse and efficient with the continuous advancement of knowledge. Technological innovations resulted in an increased ability to transform raw resource inputs, by not only discovering processes themselves, but also through mastering the conditions under which such processes take place. For example by increasing control over temperature and pressure levels. Today this fountain of knowledge has made the continuous creation of an immense number of desired outputs in complex production chains a common phenomena. In economics this production process is usually explained by the increase of two input factors, capital and labor, both over time, combined with a general indicator for technological progress. This factor represents rising productivity of input factors which created the potential to produce far more per worker in the present than in the past (Krugman et al. 2008, Sørensen & Whitta-Jacobsen 2005). Resources in general are not incorporated in this view as a separate input factor, even though their input is a necessary condition to generate output. The explanatory reason lies in the expectation that technological advances will continue to result in an abundance of resource flows as has been the case in the past. Once an individual resource such as wood, or copper, becomes scarce, it is anticipated that rising prices will result in a pathway of technological progress that causes substitutes to be developed alleviating scarcity, as was the case with past scarcities in the industrial era. Hence, only capital and labor are seen as limiting factors to increasing production in this view.

This view has been questioned from a biophysical perspective on energy required to generate work. Each alteration of matter on earth requires an expenditure of usable energy from a higher to a lower quality state bound by the laws of thermodynamics. For every production process energy input is a necessary condition, and for an economy to grow in terms of increasing material output energy consumption has to grow accordingly (Georgescu-Roegen 1971; Hall et al. 1986). If this is the case, it is impossible to substitute between energy and other production input factors. The only manner to overcome scarcity of a non-renewable energy source, such as coal, is by substitution with an energy source that can provide the means to do a similar amount of work. In the past 200 years this replacement has occurred a few times in the transition from biomass to coal, and coal to the diverse fuel mix of natural gas, coal and oil. Each successive step meant an improvement from a lower quality to a higher quality fuel in terms of energy density. For example, oil and gas can generate more useful work per heat equivalent than coal (Hall et al. 1986). As fossil fuels are of a finite nature, substitution of these will be required when they become scarce. To date, however, no energy sources have been found of similar high quality as fossil fuels making substitution much more difficult as resource and hence monetary costs are much higher. All non-fossil energy sources such as nuclear, wind and solar, are of lower quality both in terms of concentration as well as the energy, labor and capital required to capture the energy from the sun, wind and growing biomass. For example, bio-ethanol from corn is the largest present day alternative to oil in terms of volume. But the total fossil fuel energy required to produce a volume unit of corn bio-ethanol is equivalent to at least forty percent of the energy eventually delivered from the fuel (Farrell et al. 2006). Furthermore, bio-ethanol has many drawbacks including large land requirements, high labor inputs, and a longer time span of production. Fossil fuels in comparison still require relatively low energy inputs to extract as most of the work to concentrate this energy source has been carried out over a very long time span in natural processes. Even though the required amount of energy inputs to produce fossil fuels has been increasing over time as easy to access and process fossil fuels are extracted first for economic reasons. Hence cheap fossil energy sources have already been depleted significantly (Gagnon et al. 2009). Another factor that hampers substitution from fossil to non-fossil energy sources is the sheer scale of energy usage nowadays. Total global energy usage has grown by a factor of 1400 since the year 1800 and is still growing rapidly. Because of both the quality of fossil fuels as well as the colossal size of usage, it is difficult to find a suitable replacement non-fossil energy source which can deliver similar work outputs and grow at a speed required for economic growth to continue unabated when fossil fuels become scarcer. These observations give credence to the thought of reconsidering the role of energy in the economic process, as energy may well be a limiting factor to the economy in the future. The largest source of global energy consumption is crude oil, which has been forecasted to reach an irreversible decline between 2004 and 2037, depending on assumptions for global reserves, demand growth and production growth (Hallock et al. 2004).

In this thesis a broad overview will be given on the role of energy in the economic process focusing on past attempts to investigate the importance of energy in explaining economic growth. In chapter 2 an overview will be given of historical thought on the inputs of production required for economic growth, starting from Adam Smith's *Wealth of Nations* up to present. In chapter 3 the role of energy efficiency developments in an economic context will be discussed, including a glance at some energy efficiency and economic growth data. In chapter 4 different aspects of the energy rebound effect are outlined including a literature review on the macroeconomic rebound effect. The rebound effect can be summarized as a decrease in marginal costs of an good resulting in more consumption, hence more energy use. In chapter 5 studies which attempt to quantify a causal relation between energy consumption and economic growth are summarized. Followed by chapter 6 where models that incorporating energy as an input factor in macroeconomic growth models are discussed. The thesis ends with a synthesis chapter summarizing key arguments and conclusions with a number of recommendations for further research.

2. DRIVERS OF ECONOMIC GROWTH

The origin of the analysis of economic growth can be traced back in a consistent form to Adam Smith's *Inquiry into the Nature and Causes of the Wealth of Nations* published in 1776. In his famous pin production example Smith explained that the division of labor into specialized tasks leads to increased production. People are able to perform tasks better when they train and exert them frequently. Most importantly Smith was aware that the division of labor made it possible to effectively employ machines that increase productivity.

> "This great increase of the quantity of work which, in consequence of the division of labor, the same number of people are capable of performing, is owing to three different circumstances; first to the increase of dexterity in every particular workman; secondly, to the saving of the time which is commonly lost in passing from one species of work to another; and lastly, to the invention of a great number of machines which facilitate and abridge labor, and enable one man to do the work of many." (Smith 1776; book I Ch. 1 section 5)

Smith also described productivity increases due to technological progress in a rudimentary form. When people become experts in their trade due to the division of labor, they are more likely to discover better methods of production, as they can devote all their attention to a single process. The machines constructed due to this change were seen as significant by Smith, and still are in economic theory, as they could lead to further accumulation of capital.

> "capital.. may be employed in the improvement of land, in the purchase of useful machines and instruments of trade." (Smith 1776; book II Ch 1. section 5)

Potential limiting factors in the area of resources to economic growth were only widely perceived to come from limited land availability in the 17th century and early 18th century. Most famous is Thomas Malthus' *Essay on the Principle of Population* first published in 1798. Although Malthus saw no absolute limits he thought that the speed at which population grows will eventually outweigh the speed of expansion of food production required to sustain the population. Because population grows exponentially, called geometrically by Malthus, and production grows in linear or arithmetic fashion, eventually the number of people can no longer be adequately fed and people will have to live on the brink of starvation in Malthus his argument.

> "No limits whatever are placed to the productions of the earth; they may increase for ever and be greater than any assignable quantity. yet still the power of population being a power of a superior order, the increase of the human species can only be kept commensurate to the increase of the means of subsistence by the constant operation of the strong law of necessity acting as a check upon the greater power." (Malthus 1798; book 1 Ch. 1 section 28)

The view of land as a limiting factor in growing national income was described out in greater detail by David Ricardo's *On the Principles of Political Economy and Taxation* published in 1817. Ricardo thought of land as being a requirement per additional unit of capital taken into use. As land is of a limited quantity and varying quality its cultivation will lead to surplus rent. Workers who work on high quality land will obtain similar wages as on lower quality land. The better productivity of high quality land results in a surplus which is usually incurred by the land owner, used to sustain his further capital expansion.

> "It is only, then, because land is not unlimited in quantity and uniform in quality, and because in the progress of population, land of an inferior quality, or less advantageously situated, is called into cultivation, that rent is ever paid for the use of it." (Ricardo 1817; Ch. 2 section 4)

In contrast to land neither Smith, Malthus, Ricardo or any other economist from the 17th and early 18th century mentioned energy in their publications as a potentially important factor affecting economic growth. The reason is probably twofold, due to a lack of conceptual knowledge of 'energy' and the sources from which energy was drawn at that time. Nearly all energy used was heat from wood constrained by available land. Large scale coal usage only took off in the 18th century, and the notion of limits to energy sources dragging on economic growth had no practical meaning. Also energy was seen as a constant property of materials, not changing over time, and because of this did not receive much attention. The realization that expending work requires an irreversible energy flow from higher to lower quality only came much later after Sadi Carnot in 1824 established this principle which formed the basis for the formulation of the second law of thermodynamics[1]. Hence, it is not surprising that the first large study on the impact of energy availability on the economy only came after the first version of the second law of thermodynamics was formulated in 1850 by Rudolf Clausius (Smil 2005). This was William Stanley Jevons' *The Coal Question* from 1865 which introduced reasoning about energy into economic theory. The first work linking energy usage to technological progress, recognizing that capital cannot operate without continuous energy inputs.

> "But the new applications of coal are of an unlimited character. In the command of force, molecular and mechanical, we have the key to all the infinite varieties of change in place or kind of which nature is capable. No chemical or mechanical operation, perhaps, is quite impossible to us, and invention consists in discovering those which are useful and commercially practicable… For once it would seem as if in fuel, as the source of universal power, we had found an unlimited means of multiplying our command over nature. But alas no! The coal is itself limited in quantity; not absolutely, as regards us, but so that each year we gain our supplies with some increase of difficulty." (Jevons 1865; Ch. 9 section 15)

In his argument Jevons used similar mathematical reasoning as Malthus, that coal consumption was increasing geometrically but supply eventually could not keep up with this rate. In contrast to Malthus, Jevons did not expect that humanity was on a path towards dismal existence under a constant population. Because of limits of coal a reversal of progress would happen leading towards a state of declining welfare and population.

> "The wave of population will break upon that shore, and roll back upon itself. And as settlers, unable to choose in the far inland new and virgin soil of unexceeded fertility, will fall back upon that which is next best, and will advance their tillage up the mountain side, so we, unable to

---

[1] In classical mechanics in the 17th and 18th century energy was perceived as kinetic and movement (kinetics) was assumed to be time-reversible. Heat was known as a by-product of kinetic energy (e.g. friction in movements). At that time it was assumed impossible that heat could be converted into persistent movement, therefore heat was not seen as energy and only as a by-product of movement. This view became outdated by the invention of steam engines which produce kinetic energy (movement) from heat. The eventual explanation that could reconcile the dichotomy between heat and classical mechanics was the recognition of probability. The chance being very high in a system consisting of heated gas – when unhampered – that a cooling down will take place. This because nature has a tendency to optimize effort and the inverse taking place being extremely improbable. The probability of system states has been translated into the term 'entropy'. The more common a state the higher its entropy and the lower the chance of being in a state, the higher is its economic 'quality' (Koppelaar 2010). The second law of thermodynamics then places a probabilistic restriction on processes occurring in a certain direction from low to high quality energy, asserting that energy has quantity as well as quality (i.e. likelihood of occurrence). The formula used to denote this is "quality = 1 – entropy". This probabilistic restriction has been described for refrigerators or heat pumps as the clausius statement: "It is impossible to construct a device that operates in a cycle and produces no effect other than the transfer of heat from a lower-temperature body to a higher-temperature body." (Çengel & Boles 2008; p. 291) and for heat engines as the Kelvin Planck statement: "It is impossible for any device that operates on a cycle to receive heat from a single reservoir and produce a net amount of work" (Çengel & Boles 2008; p. 296).

> discover new coal-fields as shallow as before, must deepen our mines with pain and cost." (Jevons 1865; Ch. 9 section 18)

2.1 Economic growth thinking in mainstream economics

Marshall popularized the concept of supply and demand as an adjustment process of marginal utility guided by market prices. He was the first to introduce an actual model of economic growth based on the savings rate and size and efficiency of the work force. Quite similar in its essence to Adam Smith's description of productivity increases of capital and labor as the main forces that drive output growth.

> "The gross real income of a country depends on (i) the number and average efficiency of the workers in it, (ii) the amount of its accumulated wealth, (iii) the extent, richness and convenience of situation of its natural resource, (iv) the state of the arts of production, (v) the state of public security." (Marshall 1881 cited in Whitaker 1975 cited in Rostow 1990; p.165)

Marshall did not foresee any stagnation as he saw no reason for productivity increases and investment to stop, although temporary shortcuts on a national scale, for England in his reasoning, could occur at any time. The creation of the supply and demand framework created during the marginalist revolution made it possible to work out aspects of economics in detail. Economists directed their efforts in one respect at creating tools for analysis like national income accounting and also on specific aspects of economic theory such as explaining price trends, in comparison to the much broader concepts laid out by earlier economists. Previous ideas were overhauled slowly, such as the classical theory of international trade based on fixed productive factors. According to John Williams in a paper written in the economic journal of 1929, international trade should be studied as a dynamic problem where trade, cost of transport and mobility of productive factors interact continuously. Studying these interactions in the context of a growing world would eventually lead to a better theory (Rostow 1990). Regarding technology one major new viewpoints in economic growth theory was laid out in Schumpeter's *The Theory of Economic Development* published in 1911. Seeing innovation as a spontaneous and discontinuous process, as opposed to the general view that technological progress occurred at an incremental pace. The most influential publication of that era was Keynes' *General Theory of Employment, Interest and Money* published in 1936. The short-run macroeconomics theory described in Keynes' work inspired two economists, Roy Harrod and Evsey Domar (Snowdon & Vane 2005) to independently create the first growth models. Their models initially focused on explaining unemployment and the effect of investment spending on demand, not necessarily economic growth. The output of the growth rate equation in their (nowadays combined) Harrod-Domar model depends on the proportion of savings by private and public bodies divided by the value of capital goods required in producing an additional unit of output. Assumptions of the model are an exogenous labor force growth, a technology with fixed factor proportions of capital and labor, and fixed capital to output ratios. Resource limits were seen from a relative framework. The old idea of absolute limits or development towards a continuous stationary state gave way to a more dynamic framework of supply and demand, combined with diminishing returns and relative price analyses for commodities. As long as investment in capital stock continued the economy could grow, only temporarily halted in business cycles with adverse effects like high unemployment in a downturn. In the economic view the main driver of growth became technological progress and capital accumulation, much more than in the traditional era.

A number of important observations on the growth model of Harrod-Domar slowly led to the current state of macroeconomic growth theory. Of central importance was the abandonment of the assumption that production takes place under fixed proportions of inputs, initiated by Solow and Swan in 1956.

> "The characteristic and powerful conclusion of the Harrod-Domar line of thought is that even for the long run the economic system is at best balanced on a knife-edge of equilibrium growth…But this fundamental opposition of warranted and natural rates turns out in the end to flow from the crucial assumption that production takes place under conditions of fixed proportions. There is no possibility of substituting labor for capital in production. If this assumption is abandoned, the knife-edge notion of unstable balance seems to go with it." (Solow and Swan 1956 cited in Rostow 1990; p. 335)

The second important change was the concept of self-correcting stability caused by changes in income distribution and the ratio of savings to consumption. Introduced by Kaldor (1961 cited in Rostow 1990) this line of thought introduced the expectation that the economic system would return to a stable growth path after a deviation due to for instance an external shock. In case of higher output beyond the normal path, a shift occurs that inflates prices, diminishes savings and reducing demand, thereby pushing the system towards an equilibrium. These and other observations became possible by a shift towards quantified analyses in economics (Snowdon & Vane 2005). One of the major works in this respect was conducted by Simon Kuznets in the form of ten monographs under the title of *Quantitative Aspects of the Economic Growth of Nations* in the journal of *Economic Development and Cultural Change* from 1956 to 1964 (Rostow 1990). In this period the Harod-Domer model was extended and replaced by the Solow and Swan growth model published in 1956, which in its core form is still used in macroeconomic modeling today. It can be summarized as: economic growth in the form of GDP (Y) is derived from the impact of capital accumulation (K), labor supply growth (L) and technological progress (A). Formalized using a Cobb-Douglas production function:

$$1) \qquad Y = A_t K^a L^{1-a}$$

The Solow-Swan model and empirical testing of 'stylized' facts of growth as set out by Kaldor (1961 in Snowdon & Vane 2005) have led to widespread acceptance in the mainstream economics of what is called a balanced growth path of the economy[2]. Under such a path technological progress causing increasing capital and labor productivity will result in continued long-run growth of the global economy. Due to this world view, the focus in macroeconomics today lies in explaining technological progress. In the original model technological progress is taken as an exogenous process which gives little explanatory power on how technological development takes. To improve insights endogenous growth models have been introduced, measuring for instance the process of knowledge accumulation and its effect on productivity. A second focus of modern macroeconomics is the attempt to explain differences in growth trajectories as western OECD countries have grown significantly in the past two hundred years while many other countries have lagged behind.

2.2 An energy view on economic growth

The notion that most societies have arrived at a balanced growth path that can continue until an undetermined future has been criticized based on the role of energy in the economy. In mainstream economics energy is regarded as a resource in the form of a fuel like coal and gas or a flow such as wind or solar energy. It is assumed that energy inputs can be substituted by other inputs.

> "Our […] production functions assume that there are no limits to the technological possibilities of substituting capital and technologically augmented labour for scarce natural resources.

---

[2] A balanced growth path can formally be defined as a situation where consumption, investment, wages, and capital grow at a similar constant rate, g, the labour force grows at a constant rate, n, GDP, consumption and capital grow at the same rate g + n, and the capital to ouput ratio as well as the rate of return on capital remains constant (Sørensen & Whitta-Jacobsen 2005).

> Indeed, we implicitly assumed that production can continue to grow even as the input of natural resources becomes infinitely small relative to other inputs.” (Sørensen & Whitta-Jacobsen 2005; p. 209)

This notion is doubtful for energy based on the physics of the first and second law of thermodynamics. These laws make energy an inevitable prerequisite requirement for conducting work. The first law is the rule of closed systems, stating that the total amount of energy in the universe (or any other closed system) will always remain equal. An example of ‘closed’ systems are metals which are kept together by internal forces. Only when they are heated beyond a certain temperature threshold due to forces outside the system will metals melt and become ‘open’. Here meaning that these materials can be influenced by gravity and heat. The economic system itself can be conceived as an ‘open’ system as it interacts with its environment. This is not acknowledged in mainstream macroeconomics, however, as the process of energy flows that enters into the system is ignored.

The second law of thermodynamics relates to the optimization of a system, stating that the total entropy of the system is always increasing (i.e. drifting to the most probable state of lower quality energy). Human intervention attempts to achieve the opposite, however, by carrying out many industrial processes such as processing plastics, melting steel and heating food. This results in a state of lower entropy than if nature would have been left alone. However, processes making economic growth possible do cost intervention, i.e. expenditure of high quality energy. For an ‘open’ economic system to function, constant dissipation of energy (mechanically comparable to waste energy such as friction) is required, as otherwise an equilibrium with the environment via exchange cannot be obtained, let alone maintained.

The laws of thermodynamics therefore make it improbable to carry out work without a degradation from high to low quality energy. Quality being determined by the amount of work that can be performed without heat dissipation (i.e. energy losses). This implies that without continued increasing flows of high quality energy economic growth would cease, as substitution is impossible. The first major work to observe this was *Wealth, Virtual Wealth and Debt: The Solution of the Economic Paradox* written by the chemist Frederick Soddy in 1929. Soddy saw life as a struggle for energy. His views were in consequence similar to those of W. S. Jevons: without coal progress would stop causing humanity to fall back to a lower level of national income, but his focus on purely energy as a central tenet of creating national income was new.

> “A continuous stream of fresh energy is necessary for the continuous working of any working system […] Life is cyclic as regards the material substances consumed, and the same materials are used over and over again in metabolism. But as regards energy, it is unidirectional and no continuous cyclic use of energy is even conceivable. If we have available energy, we may maintain life and produce every material requisite necessary. That is why the flow of energy should be the primary concern of economics. Economics deals not with energy, but entirely with the flow of useful and available energy and its transformations into useless forms, and physical wealth as a product of the control and direction of this flow.” (Soddy 1929; p. 56, p. 104)

In Soddy’s view wealth is created by three factors, the discovery of new technologies and knowledge which takes place in a discontinuous manner, the continuous expenditure of energy from a useful form to a useless form as per the laws of thermodynamics, and human labor which Soddy describes as human ‘diligence’ as man in his view had become a tender of processes instead of a laborer due to scientific advances. Soddy does write about capital but sees it as an exponent of energy, in the form of long lasting materials made by embedding energy in them.

> “Iron embodies in itself a large part of the energy liberated in the combustion of the fuel used to smelt it from its ores, the possession of which causes it to rust.” (Soddy 1929; p. 117)

Capital, hence, forms long lasting wealth created by expending usable energy. Serving the purpose of multiplying the efficiency of time consumed in human labor. The main concern of society based on Soddy's view should be to secure increasing supplies of high quality energy to make future economic growth possible. The first thinker who described these concepts in a more formal manner was Georgescu-Roegen (1971). He criticized neoclassical economics for incorporating only a vector, or point estimate, in the production function as incorporated in the Solow-Swan model, instead of a process. In his view the component of time needs to be explicitly modeled to show the process of production. A fundamental distinction hence needs to be made between funds and flows. Funds being defined as services that can be used up only over a duration of time, as opposed to stocks which can be used in an instance of time. Flows being defined as a rate of change in a stock over a duration of time (Georgescu-Roegen 1971; p.223). Economic processes are in Georgescu-Roegen's view not well described by the relation between a set of numbers (inputs) and one number (output) but by the modeling of these inputs and outputs of these processes over time and the interaction between funds, stocks and flows. As an alternative Georgescu-Roegen proposed a flow-fund model to describe economic production consisting out of 8 elements, the economic funds Land (L), Capital (C) and Labor (H) and the flows of natural resources (N), input flows that generate products (I), input flows required to maintain capital equipment (M) and the output flow of products (Q) (Georgescu-Roegen 1971 p. 232). In this manner the most probable flow that thermodynamics has identified, from low to high entropy, embodied in the nature of every process, can be used to show the value of economic processes. As an example Georgescu-Roegen takes land, which is valuable as it is the only means by which humans can catch energy from the sun and its total surface cannot be changed. Land delivers usable energy to produce economic growth through, for example, food and wood production, by which energy becomes available which is eventually dispersed as heat. Scarcity is the result of a continuous attempt to keep our environment in a low entropy state by expending a given amount of high quality energy which can be used only once (Georgescu-Roegen 1971). Economic systems should hence be represented as a flow from high to low entropy by a throughput of energy converted from high quality to low quality unusable energy in the form of low temperature heat.

Building on the work of Georgescu-Roegen, Hall et al. (1986) describe a pure energy perspective for economic production. Starting with the observation that economic processes take place in the physical world and hence are constrained by thermodynamic laws. The system in which humanity operates can in their be conceptualized as an open system depending on a flow of low entropy energy (high quality energy) from outside the economic system defined in neoclassical economics. High quality energy must be continuously expended to keep the economy running. On that basis conceptually two levels of the economy can be distinguished. First, economic processes have a primary factor of production in the form of useful energy supplied from outside of the human economic system, either from solar inputs or mineral inputs. Second, economic processes require intermediate input factors in the form of labor, capital and technology which are dependent on useful energy for construction, operation and maintenance. Hall et al. (1986) denote every factor in the production function in terms of energy distinguishing between fuel, land, capital structures and labor. Fuel being the direct energy component providing heat upon combustion, land intercepting solar energy and transforming that into food and non-food biomass, capital structures representing energy expended in the past in their formation which is expended as the capital becomes worn down and eventually obsolete, and human labor a mechanism to transfer food and fossil energy into work. These components can be summed at an aggregate level or at the level of individual goods and services. In case of the latter, summing the energy costs of each input factor resulting in the embodied energy, or the total amount of energy used to produce a good or service. Regarding technology the energy perspective sees this as an enabler of using more energy, where productivity increases lead people to use energy directly in the form of fuel or indirectly through capital. Also technological advances are reliant partially on using higher quality fuels, as these increase the potential to do work. The implication of this is that there is a fundamental relation between productivity, technology and fuel resources which according to Hall et al. (1986) implies that energy consumption and decoupling of economic growth are a physical impossibility. Based on their reasoning of a causation between energy as a primary input and capital, labor and technology as an intermediate input a strong relationship can be expected between energy consumption and economic growth, citing a number of econometric studies that support their view. These, however, show a strong correlation between energy consumption and economic growth not necessarily causation.

Since the work of Georgescu-Roegen (1971) and Hall et al. (1986) some attempts have been made to measure the aforementioned concepts on the role of energy in the economic process. Probably the most wholesome and detailed work is Ayres and Warr's *The Economic Growth Engine: How Energy and Work Drive Material Prosperity* published in 2009. Their fundamental premise is that economic growth has been primarily driven by an increasing availability of energy. From the start of the industrial era monetary cost of energy decreased significantly over time, combined with an increasing abundance of energy, not only in absolute terms but also in terms of the productivity of energy (as the useful work possibly carried out from one unit of energy) has vastly improved over time due to technological progress. Technological progress is seen as the accumulation of a stock of knowledge, which results in one hand in continuous improvements in the efficiency of processes, and on the other hand in discontinuous change by the emergence of new technologies. In modeling technological progress Ayres & Warr (2009) replace the total factor productivity parameter from the Solow and Swan model with exergy conversion efficiency as a proxy for technological progress. Defined as "the ratio of actual work (output) to maximum work (exergy) input, for any given process." (Ayres & Warr 2009 p. 91) Exergy being a technical term denoting the maximum useful part of energy that is available to use and expended when work is performed, until energy is no longer usable by man as it is in equilibrium with the environment. For example in burning a pile of wood the energy embedded in the wood is transformed into heat that goes up in the air and cools down to air temperature. The relevance of using this ratio is that it shows the improvement resulting from technological progress in the work that can be conducted per amount of exergy exhausted in the process; i.e. knowledge accumulation has led man to utilize more energy next to more efficient usage of the exergy in fuels by reducing heat loss in industrial processes.

2.3 Summary

In this chapter the classical, neoclassical and energy view on the economic growth process have been briefly explored. Showing that the focus in neoclassical economics mainly lies in technology as the driver of growth, with the role of energy seen as minimal or neutral, as opposed to the energy critique which sees the increasing availability of energy as a prerequisite of economic growth. The importance of the disagreement cannot be understated as society today depends on a finite high-quality energy stock in the form of fossil fuels for 88% of its energy supply. If the role of energy in the economy is much more significant than assumed in neoclassical economics, the question of finding suitable fossil fuel alternatives is an economic imperative.

3. THE IMPORTANCE OF ENERGY EFFICIENCY

The energy required to create an economic output such as a ton of iron significantly decreased over time through efficiency increases in energy use. A distinction can be made at a physical level between two types of energy efficiency improvements. First, decreasing energy used to produce outputs in a known process. For example, decreasing fuel consumption by preheating air used in a blast furnace from heat produced in the process. This heat would previously have been wasted through dissipation in the environment outside of the blast furnace. Another example is more efficient isolation in a house to reduce dissipation of produced heat and hence less energy input requirements for the same heat production. Second, the invention of a new more efficient process to create the same type of output. For example, the introduction of catalytic cracking in the 1940s in the petroleum industry to split carbon atoms of oil into different types of products, which was more efficient than the previously used process of thermal cracking. Today the process of continuous catalytic cracking is still the common process for gasoline production. There are many similar examples showing that the large bulk of energy efficiency improvements have been due to new process inventions in industrial history, occurring mostly in the earlier part of the industrial revolution. Since roughly 1970 efficiency improvements at an economy level have been rising much more slowly as mainly incremental steps are being made to improve known processes, instead of the invention of new processes themselves (Ayres & Warr 2009).

For the economic process the importance of energy efficiency lies in cost saving to make societies wealthier as less energy expenditure is needed to maintain a similar production level. Another aspect of importance is the decrease in pollution when less input is required because of energy efficiency improvements. For these reasons, most countries across the globe have instituted policies to increase energy efficiency. The Commission of the European Communities (2006) has set a goal to save 20% of total primary energy usage in the European Union by means of energy efficiency versus a baseline reference scenario. In the accompanying policy paper it is stated that this would result in annual fuel savings in the order of 100 billion euro, supposedly sufficient to more than compensate for additional expenditures required to achieve efficiency targets. The commission expects the action plan to result in an absolute decrease in energy consumption from 1750 million tons of oil equivalent used per annum in 2005 to 1500 in 2020. If achieved this would imply a partial decoupling of economic growth from energy consumption. In physical terms full decoupling is impossible as all processes require energy consumption, but limited decoupling could in theory be achieved for a while when a sufficiently large increase of energy efficiency across the board offsets additional energy consumption required for economic growth. In recent history there are very few examples of countries where this has actually occurred. In the European Union the best example is Germany as shown in figure 1 where energy consumption has slowly declined since the 1970s while GDP has increased six fold.

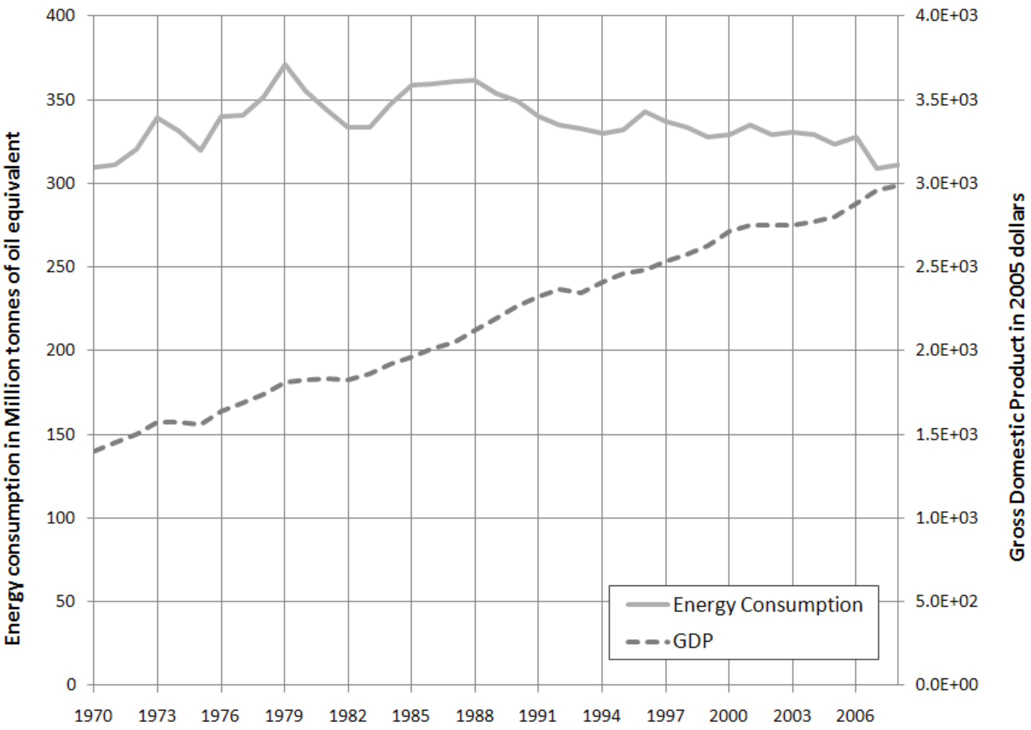


Figure 1 – Germany's primary energy consumption and GDP growth from 1965 to 2009. Energy consumption figures from the BP Statistical Review of World Energy 2008, and GDP figures are from the World Bank Development Indicators in real 2005 billions of dollars.

This apparent development of decoupling appears to confirm the probability that the energy action plan of the EU commission to reduce absolute energy usage is realistic, however, energy expenditure in producing imported materials and goods are not included in energy consumption figures for Germany. This could imply that the apparent decoupling is mostly due to the export of energy intensive activities, and to a lesser extent because of energy efficiency. Resulting in a considerable amount of energy consumption no longer taking directly place in Germany because these goods are imported, but trade and consumption still ending up on the GDP balance sheet. The GDP balance has also changed as over time Germany has had to import more and more of its energy from abroad. In 1965 Germany produced approximately 60% of its energy domestically which in 2008 had dropped to 20%. In conclusion, a large share of energy consumed to maintain the same economic output level is not incorporated in German energy statistics while GDP resulting from net energy imports is. The total effect could be of such significance that there is virtually no decreasing reliance on energy at all but only a shift taking place where the energy is consumed globally.

To shed more light on the role of energy efficiency in the economy, and whether energy efficiency can lead to much less energy usage this chapter examines three topics. First, methodologies employed to measure different types of energy efficiency. Second, an overview of GDP growth and energy use over time. Third, a discussion on the effect of structural economic change on energy intensity (or the energy usage per unit of GDP produced).

3.1 Measuring energy efficiency in economics

A common definition of energy efficiency is the ratio of useful outputs versus the energy inputs of a process. This description, however, lacks a precise definition of outputs and inputs on which further elaboration is required in order to quantify efficiency. In Patterson (1996) a thorough discussion can be found on four different ways in which energy efficiency can be measured:

1) Thermodynamic, expressed in either thermal efficiency or second law efficiency. Thermal efficiency is the ratio of usable energy output to energy inputs (in terms of heat or work). Second law efficiency is the ratio of actual thermal efficiency to the maximum possible thermal efficiency or ratio of useful work output to the maximum possible work output (Çengel & Boles 2008)

2) Physical-thermodynamic, the ratio between energy input in thermodynamic units of work or heat and output in some quantitative physical unit (e.g. mass).

3) Thermodynamic-economic, the ratio between energy input in thermodynamic and/or physical units and output in terms of market prices.

4) Economic, the ratio between energy input in and services delivered (output) in market prices.

In economics the usual method of measuring efficiency is thermodynamic-economic. This approach is chosen because using only thermodynamic metrics does not capture the valuation of consumers of end products, as consumption appreciation is not based on heat content or work potential (Patterson 1996). Usually in the thermodynamic-economic approach the amount of energy is compared with the amount of value generated in terms of GDP at a macroeconomic level, as in this manner the entire economic system can be analyzed using commonly used economic metrics. In this approach these are:

- The energy intensity of the economy, units of energy required per unit of GDP:

$$EI = \frac{E_t}{Y_t}$$

- Economic energy efficiency, units of GDP produced per unit of energy:

$$EE = \frac{Y_t}{E_t}$$

- Energy coefficients, growth rate of energy consumption to growth rate of national output:

$$EC = \frac{\frac{E_t}{E_0}^{1/n} - 1}{\frac{Y_t}{Y_0}^{1/n} - 1}$$

- Income elasticity of energy demand, percentage growth in energy consumption due to percentage growth in GDP:

$$\varepsilon_d = \frac{\partial E}{\partial Y}\frac{Y}{E}$$

Where E is energy consumption, Y is national output in GDP or GNP terms in respectively year 0 and year t, and ε_d is the income elasticity of energy demand.

There are four different problems with the economic-thermodynamic approach. The first is using GDP as an indicator for economic output because in GDP measurements non-market transactions and underground economic activities such as illegal trade are not incorporated (Karanfil 2008). In addition destructive or negative processes such as repairing vandalized property contribute to GDP growth but not to increasing useful output (van den Bergh 2009). Second, on the input side technical efficiency is not measured directly but indirectly because the aggregate level of the economy is measured. This can cause several factors to interfere with the thermodynamic-economic efficiency ratio that are unrelated to energy efficiency. These factors are 1) structural changes in the sector mix of the economy, i.e. different sectors use different amounts of energy per unit of GDP produced due to the type of processes employed, 2) substitution between energy and labor which are not caused by efficiency changes but by, for example, price shocks, 3) changes in the energy input mix as fuels have different heating values, making a fuel switch changing work output even when a similar amount of fuel is used. To overcome these problems the changes in energy intensity need to be decomposed in changes due to technical efficiency improvements, structural economic change and fuel switches. Third, incorporating all energy inputs and aggregating them in a sensible manner. The cost of energy needs to be subtracted from GDP before estimating the efficiency of output. Also, often only external energy sources are included in measuring thermodynamic-economic efficiency. However, also labor and capital use energy indirectly in the form of energy required for the food, housing, and luxury of laborers and maintenance of capital goods. These are often ignored in energy input calculations, which will be discussed in more detail in chapter 6. Fourth, in addition to methodological issues the effects of missing data needs to be taken into consideration. For example, Murishaw and Shipper (2001) note the United States manufacturing data to be incompletely reported (per 4 years with a time lag of 2 years), data on fuel economy of automobiles and trucks is incomplete, and that no information exists on energy use in agriculture, mining and construction.

3.2 A glance at Economic-Thermodynamic efficiency data

To obtain insights in the relation between energy consumption and GDP both statistics can be compared in a time-series for different geographical entities. Usually such comparisons are carried out at a country level and seldom at sector level. When comparing global change in energy consumption and GDP growth between 1970 and 2008 it can be observed that energy consumption per capita grew by a factor of 1.24 while GDP per capita in real terms using 2005 dollars grew by a factor 1.72, as shown in figure 2. These figures show a relatively small effect of energy efficiency improvements on a global scale. In the period after the second oil crisis between 1979 and 2001 a slight amount of decoupling appears to take place as GDP grows at a faster pace than energy consumption. Since then GDP and energy consumption have risen in an almost pair wise fashion, as shown in figure 3 with the annual percentage change in GDP and energy consumption. Overall no evidence can be found from these statistics on any form of decoupling between energy consumption and economic growth. These figures do not incorporate energy quality corrections, however, which need to be made due to a changing global fuel mix.

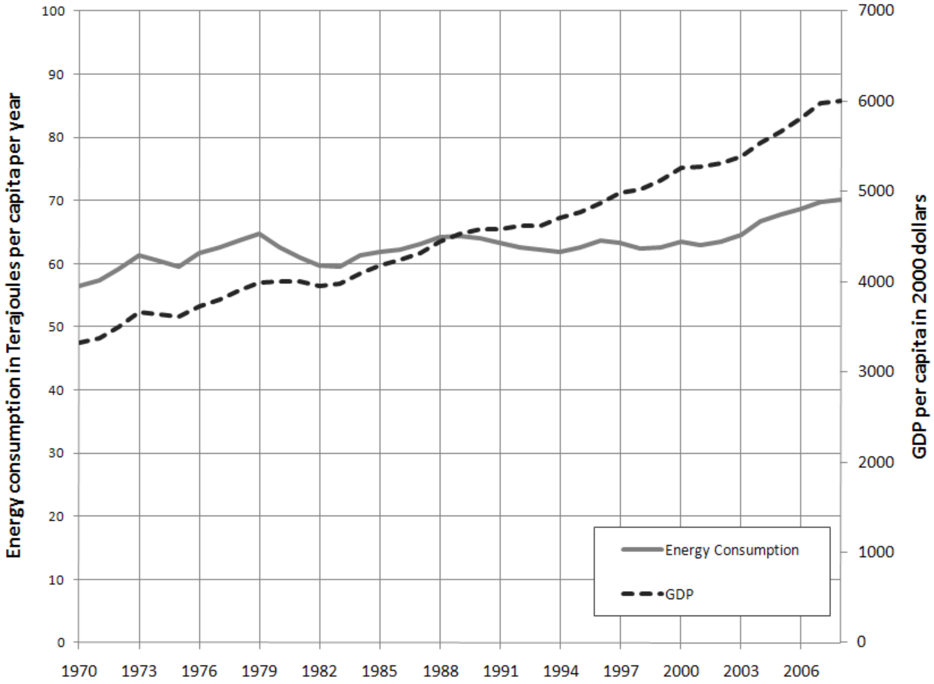


Figure 2 – Global energy consumption per capita and global GDP per capita between 1970 and 2008. GDP data from World Bank in 2005 dollars and energy consumption data from the BP Statistical Review on World Energy.

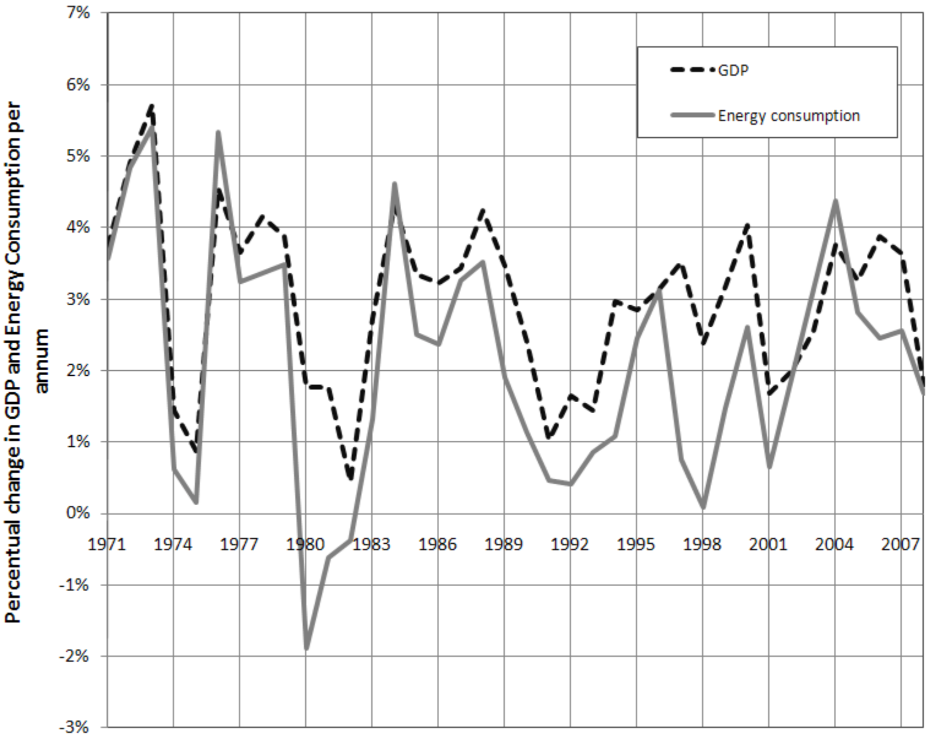


Figure 3 – YoY change in global energy consumption and global GDP between 1971 and 2008. GDP data from World Bank in 2005 dollars and energy consumption data from the BP Statistical Review on World Energy.

Studies on the effect of structural, fuel mix, and technical efficiency changes show a varied picture across countries. Patterson (1993) analyzed energy to GDP changes for New Zealand from 1971 to 1984 decomposed into sector changes, energy input mix changes, factor substitution effects, household lifestyle energy consumption and residual changes attributed to technical efficiency change. Total increase in efficiency was 15.45% over the total study period which could mainly be attributed to sector shifts because of the startup of an aluminum smelter, and increases in steel and petrochemical production. The effect of technical change was minor averaging 0.42% per year. As a probable case for low technical efficiency change Patterson (1993) notes low energy prices and lack of governmental policy support of energy saving in New Zealand during the period of analysis.

Murtishaw & Schipper (2001) analyzed energy efficiency in the United States economy from 1988 to 1998 in six sectors looking at energy intensity and structural change. They showed a significant decline in shares of fuel consumption of residential space heating, heavy manufacturing and services from 1988 to 1998 while energy consumption in freight, light manufacturing and air travel increased. For most of the total period declining energy intensity had mainly to do with energy efficiency changes, while in the latter period from 1994 onwards the effect of structural change in usage within sectors overtook this effect as activities shifted.

Wilson et al. (1994) analyzed energy efficiency changes in Australia between 1973 and 1990 finding that energy intensity (when not accounted for structural change) decreased by 6.8%. The decrease was much higher at 9.5% when structural change was accounted for because a large shift occurred towards more energy intensive sectors. When decomposing the intensity excluding structural change into fuel mix and technical energy efficiency changes, however, the biggest effect came from fuel mix changes at a decrease of 5.8% while technical energy efficiency accounted for a decrease of 3.7%. Therefore, Wilson et al. (1994) conclude that changes in energy efficiency were quite minor over such a long time span.

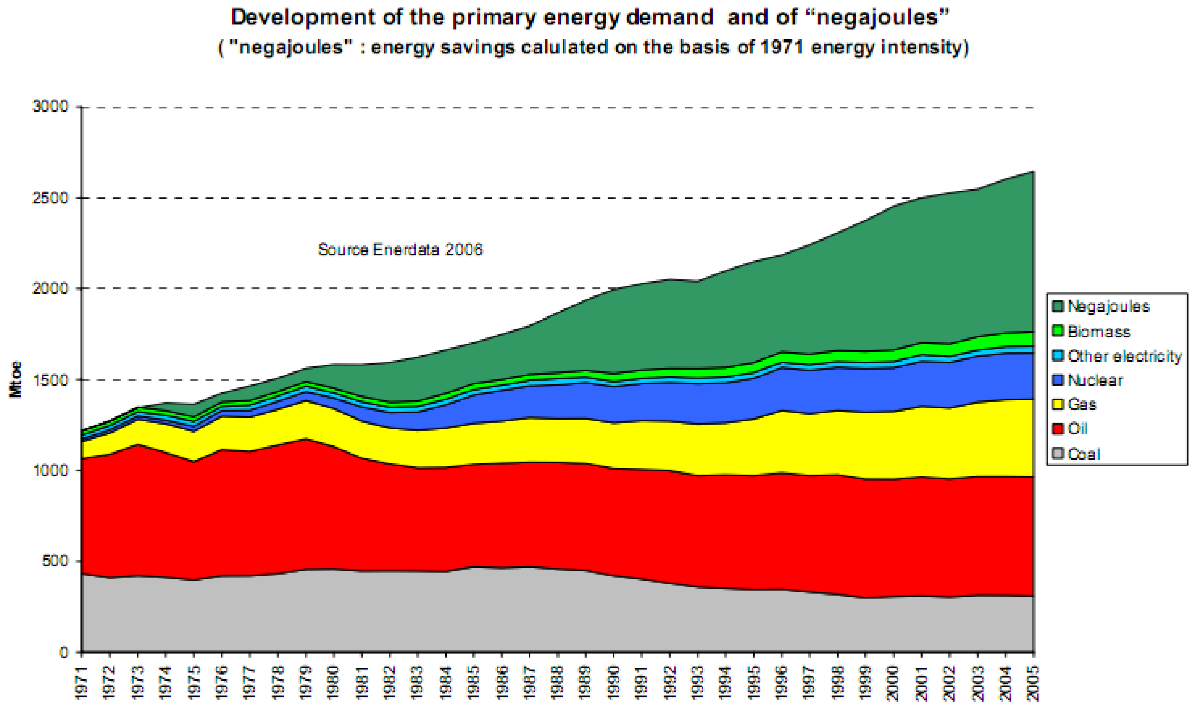
Geller et al. (2006) took a macro view of energy efficiency changes in OECD countries. Based on direct observation it was shown that in all OECD countries energy intensity fell by 2.2% per year between 1973 and 1983, 1.3% per year between 1983 and 1990, 0.6% between 1990 and 1997 and 1.8% between 1997 and 2000. To investigate the effect of structural change 30 subsectors for 11 countries were investigated. The change in the Energy to GDP ratio was decomposed into the effect of growth in energy services such as more travel and more electric household appliances, and energy intensity changes excluding this growth. The results show that for 11 studied OECD countries energy intensity across 30 subsectors declined by 1.6% per year on average, and energy services per unit of GDP declines by 0.3% on average between 1973 and 1998. Based on this finding Geller et al. (2006) conclude that energy services in general increase less than GDP except during recessions, and that the decline in the energy per GDP ratio between 1973 and 1994 are related mostly to sub-sector intensities and can, hence, be interpreted as technical energy efficiency improvements

Pen & Sévi (2009) analyzed whether a trend could be found in the development of energy intensity of 25 OECD countries between 1960 and 2004 and 73 non-OECD countries between 1971 and 2004. It was shown that for only 10 out of 25 OECD countries and 22 out of 73 non-OECD countries a significant deterministic trend could be found at 0.05 significance level. The authors found similar results when testing for oil intensity of GDP, concluding that a stochastic trend is more applicable.

The studies discussed above point towards technical energy efficiency being an important component of changing energy intensity. The size of this change varies significantly, however, across countries and time periods, indicating that this change cannot be measured in a straightforward manner from the aggregate energy consumed to GDP ratio of economies.

3.3 Examining Efficiency in a globalised context

The effect of structural change in the economy especially deserves attention in today's globalised context. In the past twenty years intensive energy processes have been outsourced from western countries to countries with cheap labor and cheap energy supplies. Many industrial processes such as fertilizer production through the Haber-Bosch process and steel production in oxygen furnaces now occur in countries where energy and labor prices are cheapest (IIER 2010). This globalization process has resulted in a restructuring of activities in the global economy with analogous shifts in energy usage. Hence, changes in energy intensity over time need to incorporate structural changes to enable correct comparisons. However, often energy intensity is studied from a perspective of what it would have been on a national scale in case that past efficiency improvements would not have been implemented. Actual developments in energy intensity are compared with a theoretical scenario, without compensating for economic shifts to sectors with less energy intensity. For example, Laitner (2000) looks at energy efficiency improvements within the US economy. Stating that in 1973 the United States consumed 78.4 exajoules of energy, due to for every dollar of GDP produced 20 megajoules were required in that year. In 1998 based on GDP and energy consumption in that year a total of 13.2 megajoules per dollar of GDP was required, as total consumption reached 100 exajoules. Based on these figures Laitner (2000) concluded that without the decline in energy intensity a total of 150 exajoules would have been consumed in 1998. As the factor that contributed most to this change Laitner (2000) mentions the shift to less energy intensive sectors, failing to note that for the production of GDP the United States has become more and more reliant on cheap energy and labor from abroad via outsourcing manufacturing activities. The same mistake is made at a policy level in case of the Commission of the European Communities (2006) action plan on energy were the energy intensity of the economy in 2005 is directly compared with the intensity in 1971 without adjusting for structural change, as shown in figure 4.


Development of the primary energy demand and of "negajoules"
("negajoules" : energy savings calulated on the basis of 1971 energy intensity)

Source Enerdata 2006

Figure 4 – development of energy demand based on energy intensity of 1971 compared with actual energy demand in 2005 after the Commission of the European Communities (2006).

By ignoring structural change across borders European policymakers likely overestimate the potential of future efficiency changes. Fuel mix changes play a less important role in case of Europe as the fuel mix remained quite similar since the 1970s. Incorporating the effect of structural change is difficult, however, as no study to date has looked in a comprehensive manner at how global energy consumption has shifted due to outsourcing of energy intensive and other industries from western countries to new upcoming economies. Studies in which this effect has been partially investigated show that it potentially can be very significant. Kahrl & Roland-Holst (2008) estimated the increase of Chinese energy consumption in relation to exports. They assessed that net-export growth caused 12 exajoules of energy consumption growth in China between 2002 and 2006, more than total energy consumption growth of 10 exajoules in EU and US from 1995 to 2005.

One methodology to measure this process in more accuracy would be to index individual technologies and their related production chain from extraction of the raw resources, to processing into the final good and the end point of consumption disposal or recycling, including the imports and exports of goods and their energy components (IIER 2010). If this chain has changed over time it then becomes possible to see to what extent consumed energy has shifted because of globalization and to what extent apparent decoupling or energy intensity improvements in Western countries is a result of increasing dependence on embedded energy in imports. This method can also be used to look at technological change over time and how efficiency develops from extraction to end disposal. For instance, the resulting higher efficiency of conversion of electricity to light when changing from an incandescent to a fluorescent light bulb, also will have positive or negative effect on changes in efficiency earlier in the production chain. Because materials required for these light bulbs differ, also the energy costs to produce the light bulbs may differ significantly. Few studies on such energy chain analyses are publicly available. In case of light bulbs a summary of a study by OSRAM (2009), a semiconductor company, can be found in which standard incandescent light bulbs have been compared with compact fluorescent lamps (CFL) and LED. In the study energy manufacturing costs for these three technologies were calculated. The study assumed optimal life conditions with LED lamps lasting 25.000 hours, CFL lights 10.000 hours, and incandescent lamps 1000 hours. The energy costs of manufacturing a 8 watt LED lamp was found to be 16 times higher than a standard incandescent light bulb at respectively 0.6 and 9.9 kWh, including transport to the sales location. Due to the long life span of the LED lamp, however, additional energy costs in manufacturing were more than compensated for in the entire life cycle, because 25 incandescent light bulbs require 15.3 kWh, in total 1.5 times as much as for the LED lamp. The study concludes that over the ‘entire life span’ of in total 25.000 hours of usage from manufacturing to end life the energy usage of a LED and CFL are similar at 667 kilowatt hours, and that a standard light bulb requires 3302 kilowatt hours. The study however ignores: 1) Energy costs in producing the raw materials required for the assembly of the lamps, 2) Disposal and recycling costs and the risk of early disposal (throwing away the lamp, for instance due to altering the building), and 3) Heat gain variations from using different light bulbs, which all could significantly influence the analysis. In addition also consumer adoption of more energy efficient light bulbs plays an important role which can remain limited because of the trade-off between higher purchasing costs and lower operating costs (Maria et al. 2010).

3.4 Summary

In this chapter it has been shown that energy efficiency has played a major role in increasing economic output during the industrial revolution. Especially in the earlier phase of the revolution due to the invention of new industrial processes the efficiency to transform energy into useful output improved vastly. In recent decades efficiency improved much more slowly. In comparing global energy consumption and GDP figures it can be shown that since 2000 there has been an almost fully coupled increase in energy consumption and growth in economic output, signaling a near lack of aggregate efficiency improvements. This brings the realism of policy plans to decrease energy consumption while maintaining economic growth into doubt. As an example for this the case

Commission of the European Communities (2006) action plan on energy is given. Studies investigating the cause of energy efficiency changes show that the apparent efficiency change is partially due to real technical efficiency change, but that also effects such as sector activity shifts and fuel mix changes play an important role. Making it difficult to straightforwardly interpret energy efficiency changes over time from a specific sector or country. To better understand the role of energy efficiency over time in today's globalised context, it is suggested to investigate the effect of structural sector changes across countries and continents over time. This because many energy intensive activities have been shifted over time towards countries where labor and energy is cheaper for economic reasons. Such research would shed more light on whether energy efficiency increases at country level actually did occur in recent decades, or whether these changes are to a large extent caused by the outsourcing of energy intensive production and importing these products from abroad.

4. THE REBOUND EFFECT

In the previous chapter it was shown that energy efficiency has played a significant role in making economies more productive. During the industrial era processes have become many times more efficient in terms of work output per unit of energy input. Despite a decrease in the required energy for industrial processes, however, energy consumption has continued to rise tremendously. This apparent paradox has raised doubts on whether energy efficiency leads to a large decrease in energy consumption, or whether it actually causes more energy consumption. The causal mechanisms are related to the saved energy which is reinvested in new forms of consumption inducing energy costs, and to lower prices which also result in more consumption as more goods can be purchased. The overall mechanism is popularly known under the name of 'Jevons's paradox' as it was first raised by William Stanley Jevons in his 1865 book *the Coal Question*, and in the economic literature under the 'rebound' effect. Jevons first wrote down the mechanism under which the rebound effect occurs in the context of coal usage in blast furnaces:

> "…if the quantity of coal used in a blast-furnace, for instance, be 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. And if such is not always the result within a single branch, it must be remembered that the progress of any branch of manufacture excites a new activity in most other branches, and leads indirectly, if not directly, to increased inroads upon our seams of coal." (Jevons 1965; ch. 7 p. 7)

The question is not so much whether the rebound effect occurs as there is a large amount of literature supporting its occurrence, but what the size of the rebound effect is. Especially the question is important whether a rebound occurs that results in more energy being consumed than was saved originally, referred to as 'backfire'. If this would be the case, then energy efficiency policy is detrimental to the economy from an energy savings perspective, even though it could make sense from an economic growth point of view. A second question of importance is what type of policies can counter the rebound effect to improve the effectiveness of energy policies. Answering these questions boils down to investigating which factors cause growth in energy consumption and to what extent these are at play. In the economy two types of factors can result in more energy consumption, one is an increase in productivity which spurs economic growth and increases income, and the other is an increase in population. The increase in productivity may occur for many inputs of production including capital, labor or energy. In case of the latter it is referred to as a 'rebound effect', even though a similar effect theoretically applies to capital and labor. The question of the energy rebound effect is whether an increase in energy efficiency plays a minor role in increasing energy consumption or whether it is one of the main causal factors.

In this chapter the rebound effect is investigated in more detail. First by looking at how the rebound effect can be defined. Second by briefly summarizing the direct rebound effect which has been well studied and documented. Third by a review of studies of the rebound effect on a macroeconomic level. Fourth by focusing on the substitution between production factors which is one of the core drivers of the rebound effect.

4.1 Defining the rebound effect

The discussion on the rebound effect was ignited by Leonard Brookes (1978) and Daniel Khazzoom (1980 in Polimeni et al. 2008). Both questioned the notion that efficiency standards imposed after the oil crisis would result in lower energy consumption at a national level due to a variety of mechanisms. After Brookes and Khazzoom many scholars have attempted to describe these mechanisms and quantify the rebound effect.

The most studied of these mechanisms is the direct rebound effect which relates to behavior changes by consumers when introducing energy saving techniques. One of the best studied examples is that an improvement in the mileage of a car leads to more miles driven. The explanation can be found in micro-economics by looking at the effect of cost reductions on a good. This leads to a shift in demand as marginal utility of consuming another good becomes higher than marginal cost of adding consumption, because of which more is consumed (IIER 2010). A much more difficult to quantify, and therefore less studied aspect, is the indirect rebound which occurs at higher levels of aggregation of the economy due to price and substitution effects. One of the mechanisms that occurs, for instance, is that at increasing cost reductions the marginal utility of consuming more units of the goods will decrease, resulting in substitution. The total of these indirect effects have also been referred to as the macroeconomic rebound, general equilibrium rebound, or economy-wide rebound effect. The divergence in terminology employed in the scientific literature signals the lack of agreement on the size of the effect, because of different theoretical approaches as well as a lack of comparability as the mechanisms of the rebound effect are often poorly described (Madlener & Alcott 2009). One of the consequences of the lack of a coherent framework is that little is known about the extent of the macroeconomic rebound effect. Greening et al. (2000) concluded based on a survey of 75 rebound effect studies that at the time of their analysis only speculation was possible on the macroeconomic rebound effect as no empirical evidence exists from sources studied. Since Greening et al. (2000) several macroeconomic studies have been conducted summarized in detail under section 3. Four types of indirect rebound related effects described in the literature are:

- *Re-investment of reduced spending*, as efficiency results in less consumption of energy per good consumed, the amount spent on the good drops resulting in an expansion of the consumption or production possibility frontier. Income is freed up and spent either on the same good which falls under the direct rebound effect, or a different good which falls under the indirect rebound effect which also consumes energy. This is also called the 're-spending' part of the rebound effect.

- *Production factor cost changes*, as efficiency increases the amount of energy required is diminished, resulting in a drop of energy costs relative to the costs of labor and capital used. Depending on the amount of substitution possible between these input factors, a shift will occur from labor and capital as an input of the production process to energy as an input. In mainstream economic theory it is often assumed that full substitution is possible, although this is not always true as discussed below in chapter 6.

- *Sector shifts*, the increase in efficiency of production in a certain sector can lead to a shift of companies activities towards that sector when it is more profitable to do so. This structural change effect of the economy would probably require a number of subsequent efficiency changes during a certain time period as switches in economic activity take time and need to be worthwhile given initial capital requirements. No distinction is made in the rebound literature, however, on the time required for structural effects to occur.

- *Economic growth*, the indirect effects described above cause increased economic growth because expansion of the production and consumption possibility frontier results in more output which is sold and leads to a secondary investment effect resulting in more energy consumption.

4.2 Direct rebound effect

In early studies of the direct rebound effect the ratio of inputs in the form of energy and economic output was either estimated or directly measured. In many cases potential efficiency was used to estimate the rebound effect instead of realized efficiency improvements (Greening et al. 2000) which historically led to overestimates of the direct rebound effect. As actual efficiency realizations began to be measured the size of the direct rebound effect has declined over time to more realistic values. The discrepancy between realized and potential efficiency is often rather large as efficiency changes in a product that could result in less energy consumption are in many cases not translated into efficiency but into more power output or pollution abatement. In industrial society a well known example lies in transportation. Potential energy savings due to more efficient motors could have triggered much larger fuel savings in the EU area since the 1970s than actually was the case. Ruzzenenti & Basosi (2008) showed - by comparing vans and trucks of similar technological capacities between 1978 and 2005 - that technological improvements in motors mainly led to increased power output. Power output of vans increased by 65% over the studied time period while efficiency increased by 30%, and power output for trucks increased by 22% while efficiency increased by 35%. Without improvements in power output, unnecessary to carry out the amount of work, the total efficiency changes could have been much higher. The reason behind the development of power output over efficiency is according to Ruzzenenti & Basosi (2008) caused by the low oil prices during the 1980s and 1990s.

In recent studies the direct rebound effect is usually quantified by estimating the own price elasticity of energy consumption or production. An overview of over 75 studies on the direct rebound has been given by Greening et al. (2000), showing that the total direct rebound effect is relatively small between 0% and 30% of energy savings. Including studies on space heating, space cooling, water heating, residential lighting, appliance, automotives transport, process uses, and lighting use in production. A similar survey by Sorrell et al. (2009) of 31 econometric studies on the direct rebound effect finds that in general the effect is less than 30%. An example of an individual study is Bentzen (2004) who estimated the own price elasticity for energy consumption in the US manufacturing sector using time series data between 1949 to 1999 of the US Bureau of Labor Statistics. Finding an elasticity of -0.24 at a 10% level of significance, indicating a direct rebound effect of 24%.

One problem with studies of the direct rebound is that they are usually conducted using data from Western societies where the level of affluence is already very high and additional investments do not go into basic necessities but into luxury goods. The difference is important because it is probable that in developing countries a higher share of money freed up through energy savings and price changes goes directly into more energy consumption making the effect bigger. Furthermore energy savings itself are often higher because used technologies often are at a low level of energy efficiency in developing countries, making the potential for the rebound effect higher. Roy (2000) conducted a study on an energy efficiency promotion program in India which intended to save on kerosene consumption in lighting. In the program villagers were supported to use lanterns with light derived from photovoltaic solar power instead of kerosene lamps. After the program was implemented villagers increased their lighting usage, from two hours initially to between four and six hours a day, resulting in a direct rebound of between 50% and 80% for kerosene usage as the actual savings were only 1.25 liters per week instead of 2.5. In addition the saved kerosene went into other types of usage, in this case mainly cooking, which when incorporated in the rebound calculations resulted in backfire. In the end the energy usage for the villagers doubled, and a rebound of 200% occurred. However, the study only looked at one village and did not control for other potential factors of growth that could have influenced the results. Nevertheless, distinguishing between countries development level because of technological differences seems to be prudent in doing rebound effect analyses.

4.3 Indirect rebound price effects literature review

The key issue with measuring a rebound effect on a macroeconomic level lies in the decomposition of causes of energy consumption increases. There are four types of economic changes which could result in increased energy consumption. Population increases, improvements in labor productivity, improvements in capital productivity, or improvements in energy efficiency. The rebound effect concerns itself with energy efficiency only and hence other factors should in some manner be separated in the analysis. For population this can easily be done as analyses can be conducted using energy consumption per capita and income per capita. For capital and labor this becomes much more difficult, however, as technological changes which result in less energy consumption almost always have an effect on other input factors. Labor, capital and energy savings are nearly always not mutually exclusive. For example, the usage of more efficient gas turbines resulted in not only less natural gas inputs per joule of electricity generated, but also in less labor requirements per unit of output in this part of the natural gas usage chain. Another example is the advancement of electric motors. In 1887 Tesla developed the first electric motors which already by 1900 had led to more than 100.000 motors used in American households, One of the first appliances in which such a motor was used was the electric vacuum cleaner introduced in 1905 by Chapman & Skinner (Smil 2005). Because of efficiency increases of the electric motor at the time, costs became sufficiently low for usage of the electric motor in all types of household appliances, which vastly diminished labor requirements and eventually resulted in a vast increase in household energy consumption. It is difficult, if not impossible at this moment in time to separate the productivity effects on different production inputs as this requires a level of detail in analysis on technological development for which data does not yet exist in an integral fashion.

To estimate the macroeconomic rebound effect the common approach taken so far is to create a theoretical model of an economy, in order to study the effect of changing variables related to energy input productivity in the production process. In this way the potential effects of energy efficiency at a macroeconomic level can be simulated. The key issue of discussion on the validity of this approach lies in the assumptions used to create the model. An overview of individual studies is given below, of which a summary can be found in table 1.

Table 1 – summary of studies using a General Equilibrium model on the indirect rebound effect

Author(s)	**Country**	**Function**	**Elasticity of Substitution**	**Efficiency increase %**	**Rebound %**	**Notes**
Dufournaud et al. (1994)	Sudan	CES	0.2-0.4	100-200	35.9 - 77	Household utility level
Vikström (2004)	Sweden	CES	0.39-0.87	15	60	Production sectors, household & government
Grepperud & Rasmussen (2004)	Norway	CES	0 - 1	100	< 100	Elasticity of substitution values not specified
Washida (2004)	Japan	CES	0.3 - 0.7	1	35 - 70	Rebound calculated indirectly from CO2 reduced
Barker et al. (2007)	UK	CES	?	8.9	11	Elasticities of substitution not documented

Allan et al. (2007)	UK	CES	0.1 – 0.7	5	14.4 – 66.8	Figures noted for long run case (variable capital)
Hanley et al. (2009)	Scotland	CES	0.1 – 0.7	5	128 – 144	Figures noted for long run case (variable capital)
Anson & Turner (2009)	Scotland	CES	0.3	5	38.3	Efficiency effect only introduced in transport sector

Dufournaud et al. (1994) created a general equilibrium model for efficiency improvements in wood consumption for cooking purposes in Sudan. Using a framework with a nested CES utility function that contains two levels, one representing the substitution between work and leisure, the other the substitution between consuming wood or other goods. Two variants were employed, in number one the purchase of other goods was treated as fixed while leisure and wood use was variable, in number two the consumption of leisure was fixed while wood and other goods were variable. An elasticity of substitution between 0.2 and 0.4 and a doubling or tripling of the wood stove efficiency from 10% up to 30% was modeled. This gave an increase in wood energy consumption in case of model variant 1 of 44.6% for a doubling and 77.1% for a tripling of efficiency. In case of model variant 2 the result was an increase in energy consumption from wood of 35.9% for the doubling of efficiency and 61.9% when efficiency was tripled.

Vikström (2004) modeled the Swedish economy distinguishing 22 sectors such that each produces one good with five among them being energy sources. Using a CES framework that nests different production inputs starting with energy and capital, and subsequently labor and capital plus energy, into the final production function level. At a household and governmental level non-energy and energy goods demand was modeled. The model was calibrated using 1957 data on the Swedish economy. For the most important energy goods at that time in terms of shares in energy consumption, electricity, and oil and gas, an elasticity of substitution between 0.39 and 0.87 was modeled. At the household level an elasticity of substitution was assumed between aggregate non-energy and energy goods of 0.4 and 0 at a government level. Modeling an efficiency improvement of 15% during a period of five years for all sectors excluding energy, and 12% for the energy producing sectors resulted in a rebound effect of 60%. After this simulation 3% labor growth, 20% capital growth and 10% productivity growth was introduced based on Swedish accounts between 1957 and 1962. The final result was a GDP growth of 21.3% and an increase in total energy use of 15.3% in the five year period, showing that the growth in the economy counteracted the efficiency effects.

Grepperud & Rasmussen (2004) looked at the rebound effect in the Norwegian economy using the MSG-6 General Equilibrium Model which contains six economic sectors, paper manufacturing, metal manufacturing, chemical and mineral products, finance and insurance, fisheries, and road transport. The model assumes full price clearing on markets with flexible prices, high substitution possibilities between technologies and rational consumer behavior under a list of constraints. A nested CES framework with three levels was employed using substitution between capital and energy aggregate (electricity and fuels), capital/energy and labor, and capital/energy/labor and transport. Substitution factors were estimated from Norwegian national account data and ranged between 0 and 1, which unfortunately have not been further specified in this study. In the scenario run the productivity change over time from 2000 to 2050 for six sectors is doubled versus a baseline scenario using measured productivity increases of 1% per annum. The outcome is an absolute reduction in energy consumption in 5 out of 6 sectors pointing to mostly a weak to mild rebound, with the exception of metal manufacturing. In that sector the consumption of both electricity and oil backfires as consumption increases over time by respectively 17.8 and 87.5 percent versus the baseline scenario while gross production increases by 31.9 percent. Unfortunately the authors do not mention the overall rebound of the economy or provide figures in energy units, making it impossible to estimate the size of the rebound effect of the economy as a whole.

Washida (2004) created a general equilibrium model for the Japanese economy modeling 33 industrial sectors of which four (coal, oil, electricity and gas) are energy producers. In addition households optimizing between ordinary and energy goods based on income were included next to government expenditures and investments. To measure production a multi-level CES production function was used with at the first level substitution between aggregate energy and other inputs (capital and labor) and at the second level substitution between these three input factors and other intermediate inputs. For the elasticity of substitution between aggregate and other inputs a value between 0.3 and 0.7 was simulated using a 1% increase in efficiency for production and consumption. The rebound found was between 35% and 70%. This was not directly calculated based on energy, however, but using an indirect proxy by subtracting calculated CO_2 reductions in percentage from the efficiency improvements. As CO_2 emissions differ per type of fuels with respect to mass and the energy delivered they are not a correct proxy for energy consumption, making direct interpretation of results problematic because compensating for this effect can change the results significantly.

Barker et al. (2007) examined the rebound effect in the UK simulating the economy between 2000-2010 using the macroeconomic model MDM-E3. The model contains 51 sectors including individual energy sectors. The rebound effect is calculated by taking the projected energy saving and comparing this with net saving per sector after accounting for an assumed direct rebound effect of 15%. The uniqueness of this study is that the actual UK policy situation is simulated by taking energy efficiency policies/measures in place. These policies were estimated to reduce energy use from a baseline level by 8.9% from 2000 to 2010, excluding additional consumption by economic growth. Also an oil price increase was assumed from 25 dollars per barrel to 56 in 2005 and 40 in 2010, and a carbon price from 0 in 2002 to 18 in 2005 and 32 per ton of CO_2 in 2010. Results show a reduction in energy use of 8.1% between 2000 and 2010 which can be translated in a rebound effect of 11%. However, the elasticity of substitution between different input factors is not described in the model, inhibiting the interpretation of results in a meaningful manner.

Allan et al. (2007) computed a general equilibrium model called UKENVI with 25 sectors for the United of which 5 were energy producing sectors. Also households and government expenditures and investments were incorporated. They used a nested CES model with capital, labor, energy and materials (KLEM), aggregating energy sources for substitution between aggregate energy and materials, and substitution between intermediates (energy and non-energy composites) and labor/capital. The model differentiated between the short and long run with respectively fixed and variable capital. In the base case an elasticity of substitution of 0.3 was used for all sectors and a sudden efficiency increase of 5%. it was found that on the short run with fixed capital a macroeconomic rebound of 61.6% occurred for electricity and 54.6% for non-electricity energy, and on the long run with variable capital a 27% rebound for electricity and 30.8% for non-electricity energy demand, across all sectors. In case of the long run analysis several sensitivity analyses were conducted. When changing substitution elasticities between energy and non-energy composites from 0.1 to 0.7 this resulted in an electricity and other energy rebound of respectively 11.6% to 58.2% and 13.2% to 66.8%. In case of an elasticity change between 0.1 and 0.7 of value added (labor and capital) and intermediate inputs (energy and non-energy composites) the rebound for electricity and other energy rebound was respectively 14.4% to 52.6% and 22% to 49%. The contribution of increased economic output to the rebound effect was found to be relatively small as the largest change due to this effect was a 7% increase in non-electric energy production. It was concluded that the rebound was mainly caused by a redistribution of factor inputs to more energy inputs in production due to price changes.

Hanley et al. (2009) conducted a study using the model framework from Allan et al. (2007) and applying it to the Scottish economy as a subsection of the United Kingdom, renamed the AMOSENVI model. In their model they incorporated a similar sudden 5% increase in energy efficiency that led to a rebound effect of 63% for electricity and 54% for non-electricity energy-usage on the short run, and a rebound in the long run of 131.6% for electricity and 134.1% for non-electricity energy usage. The effect of varying the elasticity of substitution at different levels in the nested CES formulation between 0.1 and 0.7 had little effect, this only caused the electricity rebound effect to vary between 127.8% and 139.3% and non-electricity energy between 129.2% and 143.8%. Additional sensitivity analyses were carried out where the energy efficiency improvement of 5% was only applied to specific sectors. In case of simulating only the efficiency shock in the energy supply sector a rebound of 243.8% occurred, in case of non-energy supply sectors a 34.8% rebound, and in case of only heavier energy using sectors a rebound of 40.3%.

Anson & Turner (2009) use the AMOSENVI model to study the rebound in one of the 25 sectors of the model, the transport sector. As in earlier studies the impact of a sudden 5% increase in energy efficiency was analyzed using an elasticity of substitution of 0.3 in the base case. In addition to the model framework used by Hanley et al. (2009) a disinvestment effect was added, described as the effect of declining prices due to the energy efficiency increase on capital investments. It is assumed that due to lower demand the price of oil falls and firms therefore invest less and instead run down on their capital stocks. Because of less investment capacity contraction occurs until the point that energy prices begin to rise again due to lower supply reaching a new point of equilibrium, in the model framework at the level of capital costs. When including this effect, Anson & Turner (2009) find a macroeconomic rebound effect for oil consumed in the transport sector of 38.8% in the long run and 36.5% in the short-run.

The macroeconomic studies described above show that the total rebound effect is potentially much larger than assumed in direct rebound studies. Although the results are still inconclusive the majority of studies suggest that the macroeconomic rebound effect is larger than 50% and the possibility for backfire, a rebound higher than 100%, is not ruled out. The main variable that influences the outcome of these studies is the elasticity of substitution (assumed between production input factors). In practice substitution depends on the chosen timeframe and physical possibilities, while to a certain extent substitution between input factors always is possible, its limitation depends on the type of process. In addition the indirect energy costs of labor and the accounting of different energy inputs for different capital structures are not included in the macroeconomic model framework. If they were it could lead to a higher rebound. This issue of energy accounting for different production inputs will be discussed further in chapter 6. Other aspects that could significantly change the results are related to simplification of model specifications. These include ignoring the energy content of imported goods (Barker et al. 2007), omission of financial flows in general equilibrium models (Allan et al. 2007; Hanley et al. 2009), the capital costs of the energy efficiency improvement (Sorrell & Dimitropoulos 2008), and the disinvestment effect described by Anson & Turner (2009).

4.4 Coping with the rebound effect

The rebound effect poses a serious challenge to policy makers from the perspective of diminishing energy consumption in order to reduce CO_2 emissions and to decrease dependency on fossil fuels. The nature of the rebound effect is such that its occurrence is persistent due to price interaction between different production input factors. This necessitates, when creating energy policy, to understand the rebound effect better in its complexity, as many gaps still exist in the analysis as outlined in section 4. Often in designing energy policies at a governmental level the rebound effect is not included at all causing an overestimation of the potential to decrease energy consumption from a baseline level (Madlener & Alcott 2009). The main example given in this report is the lack of accounting for the rebound effect in the energy efficiency action plan of the Commission of the European Communities (2006). As a starting point at least the current preliminary conclusions of the macroeconomic rebound can be taken, on top of which further research on the size of the rebound effect is required to make better informed policy decisions. A simple framework based on marginal utility and substitution effects from micro-economics could be used to quantify the potential rebound effect of different policies based on estimated energy cost changes. Another aspect of importance is to design policies accounting for the trade-off between energy efficiency and rebound induced economic growth. The induced growth can be used to steer the economy in a desired direction. For example, in case of the desire to decrease dependency on fossil fuels, by re-investing the money from the efficiency gain in technologies and energy infrastructure further reducing fossil fuel dependency. In theory a possible mechanism to do this is taxing energy to counteract price decreases that are the result of improved energy efficiency, and use the new flow of funds to invest in certain technologies or energy infrastructure. This raises the problematic issue, however, on how specific causes for an energy price change can be distinguished, because many factors are involved which may at least on the short term have a much bigger influence than energy efficiency changes.

5. THE RELATION BETWEEN ENERGY AND ECONOMIC GROWTH

Past analyses on the relation between energy consumption and economic growth were carried out by looking at the correlation between fuel consumed over time in the economic process and economic growth. Two methods have been used. Direct correlation estimated through the coefficient between energy consumed and economic growth, and indirect correlation by adding energy as a variable in economic growth models and testing model outputs against actual economic growth figures. The underlying expectation of causality between energy and economic growth stems from Physics in that energy is required to create useful work. Based on the necessary condition of energy inputs to produce output it may be expected that there is a strong causality between energy consumption and economic growth. Broadening our view of causality between energy and economic growth four directions of causality are possible as distinguished by Hu & Lin (2008):

- Unidirectional causality from economic growth to energy consumption, implicating that policies which reduce energy consumption have no effect on economic growth, also referred to as the 'conservation hypothesis'.
- Unidirectional causality from energy consumption to economic growth, implicating that restricting energy consumption will negatively impact economic growth, also referred to as the 'growth hypothesis'.
- Bidirectional causality between energy consumption and economic growth, economic growth causes energy consumption and energy consumption drives economic growth, also called the 'feedback hypothesis'.
- No causality, there is no relation of energy consumption on economic growth or vice versa, this is also called the 'neutrality hypothesis'.

In order to test for these types of causality the time series methodology, first developed by Granger (1969), has become common practice with dozens of analyses conducted. This chapter reports the outcome of a number of these studies. First, by describing the methodology employed in performing causality analyses. Second, by giving an overview of the assumptions, methodology and outcomes for macroeconomic rebound. Third, by looking at studies on energy prices, economic growth and energy consumption.

5.1 Causality analyses between and energy and economic growth

Econometric analyses to uncover correlation and/or causality are highly dependent on the approach and testing procedure. In the early days of econometrics most analyses used simple linear models estimated by ordinary least squares (Asafu-Adjaye 2000, Keppler 2007). These are not able to take into account non-stationary aspects of time series, causing potentially spurious results. Also ordinary least squares (OLS) analyses may yield high correlations (high R^2 value) but do not explain underlying relations. For example, Hannon & Joyce (1980) analyzed the relation between energy and economic growth by using a production function with GDP as dependent variable and as independent variables capital, labor, technical progress and energy inputs. Nine different forms were produced with and without energy as an input factor, and these were subsequently tested by OLS, applying data from 1929 to 1969 for the United States. The results showed a high correlation (R^2 of 0.99) for all nine forms, despite the omission of energy or labor or capital, prohibiting any conclusions on causal relations from the data. Econometrists developed means to tackle this insensibility problem by improving correlation analyses via including time-series comparisons, starting with a seminal paper by Granger (1969) who introduced a new method to analyze causality between two variables. The method estimates inter temporal correlation for two models comparing how well these give explanatory power a dependent variable. For instance the dependent variable Y, in one model (2 below) only independent upon its past values, while the other model (1 below) also includes past values of another variable X. In case the fit of the second model is better than the first (due to a lower prediction error, (Granger, 1969)) causality is said to be established.

$$1)\quad Y_t = a + \sum_{i=1}^{T} \beta_i Y_{t-i} + \sum_{j=1}^{T} \theta_j X_{t-j} + \varepsilon_t$$

$$2)\quad Y_t = a + \sum_{i=1}^{T} \beta_i Y_{t-i}$$

This method became the most common in practice because of favorable Monte Carlo evidence testing (Keppler 2007). Nowadays the methodology in general obeys five steps:

1) The time-series is tested for non-stationarity, this is commonly done using the Philips-Perron 1988 test or the Augmented Dickey-Fuller test.
2) In case of non-stationarity the time-series is normalized by differencing[3] so that it becomes stationary.
3) The non-stationary time series is tested for cointegration, an idea first introduced by Engle and Granger (1987; cited in Dolado et al. 2003), tests whether a specific relation occurs between two non-stationary time series. In case the residuals of a simple OLS equation containing the two variable series is stationary, or in other terms contain a unit root[3], it is said that the two variables are co-integrated. An example, given in Keppler (2007), is $Y_t = a + \beta X_t + u_t$, in case $\hat{u}_t = Y_t - a - \beta X_t$ gives in its residual a random walk then the two variables X and Y are cointegrated. An alternative to test for cointegration is the Johansen & Juselius (1990) test.
4) In case of the presence of cointegration an error-correction procedure needs to be applied which corrects the model for shocks that drive it away from an equilibrium relation.
5) The model is tested using Granger causality either directly in case of a stationary time-series that is not co-integrated, or after the cointegration procedure. When no cointegration is present usually a vector autoregressive framework is used were the variable under question is explained by a number of lags, say p, of that variable, $Y_t = A_1 Y_{t-1} + \ldots + A_p T_{t-p} + U_t$, with A_i being coefficient matrices and U_t an independent white noise process[4] (Lütkepohl 2003). In case of cointegration the test is conducted with the error correction model as created under step 4.

In case of more than two variables a multivariate cointegration procedure applies. Also several variants of the causality procedure have been developed enabling cross-section analysis to be combined with time-series using panel cointegration tests. These include panel root test and panel-based error correction modeling (Karanfil 2009). Furthermore, an important factor that needs to be taken into account is structural breaks as these can lead to invalid inferences. In case of energy the most influential breaks are the oil crises of the '70s and 80s' and for economic growth the occurrence of recessions. To incorporate the effect of a structural break often a dummy variable is included in the model that accounts for shocks.

5.2 Results from studies on causality between energy consumption and economic growth

---

[3] A unit root test is conducted to discern whether the time series depends on its initial value or also on its previous moment(s), coined as respectively the first and second order of integration(i.e. second being a stationary time series after first differencing). First differencing is taking a replacement variable Z, such that Z(t) = Y(t) – Y(t-1) to transform the time series to the first type.

Several overviews have been made of causality studies. Payne (2010) summarizes literature on the causal relation between electricity consumption and economic growth per country, time frame, methodology, variables and causal relationship. He surveys 35 studies showing very mixed results: 31.35% gave evidence for the neutrality hypothesis, 27.87% for the conservation hypothesis, 22.95% for the growth hypothesis and 18.03% for the feedback hypothesis. The drawbacks mentioned on the validity of these studies are the likelihood of omitted variable bias because in 26 out of 35 only two variables are included, and that most of these studies did not investigate the coefficients in terms of the sign and magnitude of the relation. Ozturk (2010) summarized the literature on energy consumption and economic growth by country, time frame, methodology, and causal relationship. He surveys 26 studies including multiple countries, and 38 single country studies. The results again are mixed as all four hypotheses are shown as a regular outcome. For example, 15 different studies analyzed the causal relation between energy and economic growth in the USA, of which 2 gave as a result an unidirectional relation from GDP to energy consumption, 4 from energy consumption to GDP, 1 bidirectional causality, and 8 no causal relation.

This mix of results can be explained, according to Karanfil (2009), by the used methodology and the time period taken in consideration. One of the key issues is that a testing procedure with only two variables, energy consumption and economic growth, may prove to be too limited to be useful, making it difficult to rely on any policy implications assumed in these studies and requiring a need to develop new or improve better methods. Developments in econometrics need to be reviewed more frequently and substantially to know whether new methodologies can bear better results. In similar terms, Ozturk (2010) thinks that the mixed results are to a large extent caused by a too limited testing framework. These authors, however, do not provide any answers as to what type and number of variables need to be analyzed to improve insights in the energy versus economic growth connection. Although energy is a necessary condition for economic growth, it is not the only one. In the usage of energy capital is needed created through the advancement of knowledge. Hence, depending on the relative importance of energy to other factors that potentially drive economic growth in a certain time frame in the economy, the results of such analyses will differ. Three of the most important factors would be how to treat lags between energy consumption and economic growth assumed in the econometric analysis, in combination with the effect of energy prices, and technological change.

5.3 Results from causality analyses including price effects

Another relation between energy consumption and economic growth is by the effect of energy prices on economic growth. If energy has an important fundamental relation with economic growth, it can be expected that rising energy prices will have a significant influence on the economy as every single process is affected. The best studied effects are those involving the price rises of oil in the 20th and 21st century. It was shown by Hamilton (1983, 2005) that nine out of ten recessions in the U.S. after 1948 were preceded by oil price shocks, with a statistically significant correlation. He concluded that this supports the idea that oil shocks contribute to recessions. In a recent paper, looking at the 2007-2008 oil shock, Hamilton (2009) conducted a regression analysis comparing a situation with and without the oil price rise. Between January 2005 and July 2008 the average monthly price of oil rose from 33 dollars per barrel to a high of 133 dollars per barrel after which a decline set in. Based on the relation between consumption spending, oil prices and economic growth, Hamilton (2009) estimated that without the oil shock economic growth would have been 0.7% higher on average. From this he concludes that even though other factors were also of importance in the recent recession, especially the collapse of the housing market due to a lack of mortgage payments, these would not have led to such severe recession without oil price shock.

> “Eventually, the declines in income and house prices set mortgage delinquency rates beyond a threshold at which the overall solvency of the financial system itself came to be questioned, and

> the modest recession of 2007:Q4 – 2008:Q3 turned into a ferocious downturn in 2008:Q4. Whether we would have avoided those events had the economy not gone into recession, or instead would have merely postponed them, is a matter of conjecture. Regardless of how we answer that question, the evidence to me is persuasive that, had there been no oil shock, we would have described the U.S. economy in 2007:Q4 – 2008:Q3 as growing slowly, but not in a recession." (Hamilton 2009; p.40)

Similar results were found by Zhang (2008) who looked at the relation between oil price shocks and economic growth in the Japanese economy. Using a Granger causality testing framework he found a uni-directional causality running from oil shocks to economic growth. Jayaraman and Choong (2009) looked at the relation between economic growth and oil prices in fourteen small Pacific Island Countries, of which all except one fully depend on oil imports. Using time series data from 1980 to 2007 the relation between economic growth in real GDP as dependent variable was modeled with as independent variables the oil price, international currency reserves and a trend variable. The framework employed was an autoregressive distributed lag procedure with lags of dependent and independent variables in the model specification. The results show that a causal relation exists from oil prices to economic growth for all 14 countries at a 5% level of significance. The opposite hypothesis was also tested but yielded no significance even at a 10% level. This could have been expected given the small size of these countries in the world economy because of which economic growth which causes more oil consumption will not have a significant impact on oil markets.

In a first study of its kind Bessec & Méritet (2007) went beyond the relation between prices and economic growth by including technology and oil intensity developments, conducting a multivariate time-series analysis on real oil prices, technology and oil intensity of the economy. In total 15 OECD countries from 1960 to 2002 were studied using GDP in 1995 prices and oil prices in dollars per barrel. For technology a proxy was taken using the fuel rate which is obtained by taking total road vehicle fuel consumption and dividing it by the average mileage of road vehicles per individual country. The authors concluded that as road fuel consumption is about half of total oil consumption this measure is a sufficient estimator of oil consumption efficiency improvements and, hence, technological change. The model was set up using the multivariate cointegration framework proposed by Johansen (1988) including a vector error correction model. Also a dummy variable was introduced to account for the oil price shocks by adding a structural break in 1973. The results of the unit root test showed critical values to be exceeded for oil prices in all 15 countries critical, while for oil intensity 10 out of 15 showed significant results with inconclusive results obtained for Austria, Germany, Finland, Italy and the Netherlands. Based on these results the test was continued using a first differenced time series for all countries, cautioning the interpretation for the five countries with no significance for a unit root for oil intensity. Subsequent tests showed cointegration to be present between oil intensity, oil price and fuel rate in 12 countries, and no cointegration in Greece, Norway and New Zealand. For the latter three a VAR framework was specified while for the other 12 a VECM framework was used to test for Granger causality. The results of the Granger causality test are:

- "A unidirectional causality running from prices to oil intensity in 12 countries and a causality from the fuel rate to the oil intensity in 11 countries with a feedback effect from oil intensity to fuel rates in 7 countries.
- in 13 countries a relationship running from price to fuel rate and in 7 countries a causality running from oil intensity to fuel rate.
- Strongly exogenous oil prices, except in three countries (Germany, the United Kingdom and the United States); this means that this variable is generally unaffected by the changes in fuel rate

> and in oil intensity except in major countries like the United States.” (Bessec & Méritet 2007; p. 140)

The results show a strong link between energy prices and technological progress in terms of the fuel rate. In macroeconomic terms this process would lead to a shift in the demand curve as technological progress lowers costs making more consumption possible. Also a strong link from oil prices to oil consumption was found, indicating a downward sloping demand curve with rising prices resulting in lower consumption. Furthermore oil prices are according to the results only influenced by consumption in large consumer countries, Germany, the United Kingdom and United States, as smaller countries consume lower amounts which does not have an impact on the global market. The study therefore confirms that oil demand is sensitive to high prices, and that energy prices can have an impact on technological development in terms of energy efficiency.

5.4 Summary

The time series testing method developed by Granger looks at the inter temporal correlation for two models. Estimating how well both models can explain the development of a dependent variable Y over time, with one model encompassing only past lags of that variable Y, and the other also incorporating another variable X next to the past lags of Y. In case the fit of the second model with X is better than the first (due to a lower prediction error, causality is said to be established which can be in four directions (Granger 1969)). This includes unidirectional causality from X to Y, unidirectional causality from Y to X, bidirectional causality, or no causality found when both models perform similarly. In the literature this method has been often applied to measure causality between energy consumption and economic growth with inconclusive results. In more than 50 studies all four directions of causality have been established, often also for the same countries. The main reason given for this contradiction is that a testing procedure of only two variables is too limited (Karanfil 2009; Ozturk 2010). In the energy and economic growth literature little reflection is given as to what variables should be incorporated to overcome this issue. An area that appears promising in this respect is the relation between energy consumption, energy prices, technological development, and economic growth and energy consumption. Studies that look at the causal relation between energy prices and economic growth are less numerous but do show more firm results of causality running from oil prices to economic growth (Hamilton 1983, 2005, 2009; Zhang 2008; Jayaraman & Choong 2009).

6. MACROECONOMIC GROWTH MODELS INCLUDING ENERGY

In previous chapters three topics regarding the role of energy in economic growth were discussed, energy efficiency, the rebound effect, and the causal relation between energy and economic growth. The conducted literature surveys revealed many facets on the role of energy in the economic process, but often the results of the relations and size are disappointing and contradictory. Especially in literature on the causal relation between economic growth and energy, a continuance of the employed methodology does not yield better insights but only reproduces past inconclusive results. The probable cause is that only one or a few aspect of energy in the economic process have been investigated without placing the research in a greater context of the functioning of the economy. That leads to the question of how to formalize energy in an economic model framework to create coherence and a deeper understanding of the influence of energy on the processes that take place to create economic growth. On one end this would require better accounting of energy flows in the system as to make better measurements possible, and on the other end a theoretical production framework relating to these flows needs to be understood and modeled. In this chapter these issue will be discussed in reflecting upon attempts to incorporate energy in macroeconomic growth models.

6.1 Energy accounting in the economic system

The major problem of measuring energy in the economic system is the lack of incorporation of energy in the production function. A few attempts to explicitly model how energy affects the economic process have been made. Moon & Sonn (1996) created an endogenous growth model with capital and energy as inputs. In the model energy and economic growth interact first by an energy expenditure function wherein invested income is affected by energy costs, and second by an energy intensity function modeling the productivity change of energy that can increase economic growth. Smulders & de Nooij (2003) developed a growth model to study the effects of energy conservation policies on economic growth. They used a final goods production function with inputs in the form of labor services and energy services in a constant elasticity of substitution (CES) function specification, and underlying production functions that quantify these services through raw input and unspecified intermediaries as inputs. Zon & Yetkiner (2003) extended the Romer growth model by including energy consumption as an input step in providing intermediary capital services next to raw capital. Thereby combining possible endogenous energy saving through technical change as factor productivity of intermediary capital services also improves the amount of output generated with an amount of energy. In this manner economic growth can become affected by changes in energy prices if costs of intermediary goods change. Arbex & Perobelli (2010) created a macroeconomic model with capital, labor and renewable and nonrenewable energy inputs including a total factor productivity parameter that affects all inputs.

Three problems can be found with these approaches. First, energy is treated in a similar way to other input factors in assuming full substitution, implying that a process can be conducted with less energy by substituting for other input factors, or sometimes depending on the model structure, even without energy. This arises from the notion of a constant elasticity substitution between energy and other factor inputs in the production process, capital and labor, which is flawed from a Physics perspective, as energy is not substitutable for any other input except (conversion of) energy. Even though energy inputs can be diminished over time through efficiency improvements, energy can never be fully substituted making it doubtful whether the mainstream economic framework, that assumes differently, is a sufficient approximation of reality. For example, in case of assuming a constant elasticity of substitution in metal products manufacturing. When forging finished metal rods into a product, substitution between labor and a combination of energy and capital is possible to a certain extent, but labor itself is a transformation of energy embedded in food into mechanical work. Treating apparent non-energy production factors as having no energy component is incorrect from a thermodynamic point of view (Hall et al. 1986). The laborer needs to be fed, transported to work, and consumes normal and luxury goods, requiring energy. When expenditure in the form of direct energy inputs in a process is substituted for labor, this leads to higher labor requirements and hence more energy inputs to sustain the laborer. Depending on the structure of the economy labor will be taken away from other sectors and/or from unemployed citizens resulting in changing income and consumption patterns and hence energy consumption. Hence, when looking at labor and energy substitution what happens conceptually is a shift from energy consumption from external sources to labor energy consumption that needs to be accounted for. This is only possible to a very small extent, however, as the energy requirements for production processes globally by now vastly exceeds the labor capacity of human beings. Finally for some processes substitution between energy and labor is impossible due to the difference in quality of these inputs. For example in case of metal rods manufacturing, substitution between labor and energy is possible for molding the rods, but energy in form of heat transfer required for melting the steel cannot be substituted by labor as this is physically impossible due to the amount of heat required. The second problem in modeling energy as an input in the economic process in the models mentioned above is a lack of compensation for different heating values of fuels. When substituting between fuels work output can increase or decrease using similar quantities of fuel. The third problem is that in the energy models above energy is only investigated at an aggregated level prohibiting the investigation of the effect of structural changes inside the economy and shifts of production to other countries. Therefore the explanatory power of these models becomes limited as especially the process of globalization has resulted in many shifts across countries of energy intensive processes.

To increase the explanatory power of macroeconomic energy models a framework of accounting for energy changes of each production input at a sector and at chain level needs to be made. This would encompass energy accounting by: 1) calculating the indirect energy costs of labor for different processes across different societies and sectors, 2) the cases in which labor for energy substitution can happen as there is a limit to substitution, 3) the manner in which energy for energy substitution occurs and compensates for heating values, 4) accounting for the energy consumption changes that occur due to shifts between sectors within a country and between countries. The problem with this approach is determining the system boundaries with respect to the energy required to produce labor. Does this include only direct energy costs (in the form of food calories, and fuel required to drive to work) or also a part of the food chain and the energy required to consume other goods besides food and fuel? Another issue is the level of complexity required to look at energy from a sector and cross country perspective.

6.2 Modeling the role of energy in technological progress

The resulting picture from section 1 is static, however, as it only shows how energy ties in with the economic process over time, but does not explain how these changes come about. The changing combination of energy and capital through economic progress needs to be described in detail. The relation between capital and energy has for a long time been discussed, with different studies showing the two factors to be substitutes, while others show them as complementary. Apostolakis (1990) discusses 25 studies on capital and energy relation, finding in most cases time-series data studies to confirm the complimentary of capital. In contrast pooled cross section studies were found to support the hypothesis of substitutability. The different results probably can be explained by investigating the relation between capital and energy. From a theoretical point of view capital acts as a complement to energy as machines have as a necessary condition that they require energy to function, and contain embedded energy from when they were produced. The process of production always involves doing work, and capital is an intermediate agent in the process requiring energy inputs (Pokrovski 2003). The ratio between energy and capital used, however, does not need to be constant. Over time different machines may be developed to produce work in a more efficient manner, by increasing the transformation efficiency of energy resulting in less outputs needed for the same amount of work. Often this does lead to more sophisticated machinery, however, requiring more energy to construct and manufacture. The change in embedded energy in machinery needs to be taken into account to make a fair comparison possible of how the relation between capital and energy changes, often neglected in substitution studies. Concluding from this theoretical foundation it can be stated that capital and energy themselves are not substitutes, but that partial substitution between varying combinations of capital and energy over time is possible. In many cases this is measured as substitution because the production function employed in the studies described in Apostolakis (1990) is not able to measure dynamic changes. How to treat the relation in the theoretical economic framework is described in Kander & Schön (2007) who see the development between capital and energy as one of biased technical change. The difference being that substitution is a shift along the production curve of a selected factor of production, while biased technical change is a shift of the production curve itself caused by reducing one input while another input remains at the same level. This is shown graphically in figure 5. The driving process in case of energy being mainly improved efficiency of processes.

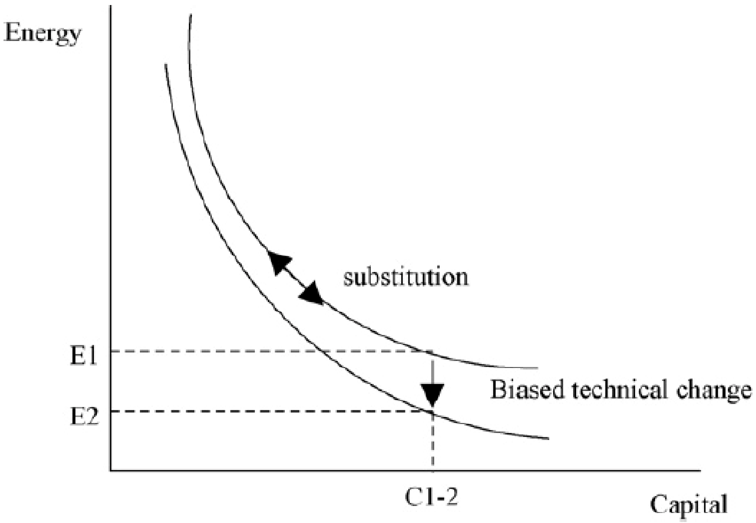


Figure 5 – Graphical overview of substitution and biased technical change after Kander & Schön (2007)

The discussion outlined above only considers the changing relation over time of capital and energy. A description of the explanatory power of why and how this relation changes is lacking. In that area most work has been conducted by Ayres & Warr (2009) who have looked at the effect of improving energy efficiency in creating economic output. As described in chapter 1 the approach taken is to replace the factor productivity parameter from the Solow and Swan model by exergy conversion efficiency[5] as a proxy for technological progress. As a model framework two production functions were created. The first Y1 describes the production of a single intermediate product of useful work through exergy inputs and technical efficiency changes. The second Y2 involves the transformation of useful work into GDP from inputs of capital, labor and useful work. The term useful work refers to the total exergy input for different types of end-use (work) times the average exergy conversion efficiency by which each type of work is produced. From this total inputs in terms of energy are obtained that are actually put to use in production. These include muscle work, mechanical work from prime movers such as heat engines, electricity, and heat used at point of use. Both sectors are described similarly with production (Y), Capital (K), Labor (L), and useful work (U).

---

[5] As mentioned in chapter 3 exergy conversion efficiency is defined as “the ratio of actual work (output) to maximum work (exergy) input, for any given process (Ayres and Warr 2009; p. 91).” Exergy being a technical term denoting the maximum potentially available useful part of energy to be expended by performing work, until exhaustion (e.q. is in equilibrium with the environment).

1) $Y_1(K^*, L^*, U^*) = (1 - \lambda)\, Y(K, L, U)$

2) $Y_2(K - K^*, L - L^*, U - U^*) = \lambda Y(K, L, U)$

3) $Y_1 + Y_2 = Y$

In this form the total production is separated in two stages. For application purposes, however, the traditional Cobb-Douglas production function was used $Y = A(t)K^{\alpha}L^{b}E^{\gamma}$ with the exception of the additional variable E denoting exergy input. Constant returns were assumed for capital and labor of the form $\gamma = 1 - \alpha - \beta$. The function was, however, transformed assuming marginal productivities after Kummel (1980 and 1985 in Ayres & Warr (2009) under the conditions:

4) $\alpha = a\frac{L+E}{K}$

5) $\beta = a(b\frac{L}{E} - \frac{L}{K})$

These conditions describe that for capital (as an input factor) to work labor or exergy will always be required and labor will be substituted over time by a combination of capital and exergy when capital intensity increases. The transformation resulted in the (linear-exponential) production function (6) through partial integration and setting the factor A(t) = 1 in the function $A(t)K^{\alpha}L^{b}E^{\gamma}$ as productivity is incorporated by calculating the useful work of a process and not by a general factor productivity term.

6) $Y = E * \exp(a(t)\left(2 - \left(\frac{L+E}{K}\right)\right) + a(t)b(t)\left(\frac{L}{E} - 1\right))$

To test their ideas the total amount of aggregate useful work (exergy services) over time for each sector was calculated in the Japanese and US economies from 1900. Showing that, for example, the overall exergy efficiency of electricity power generation and distribution in the US improved from 3.8% in 1900 to 17.3% in 1930 and 31.3% in 1960 after which it hardly increases as in 1990 the exergy efficiency was estimated to be 33.3%. By multiplying this metric with total inputs in terms of exergy consumed the total amount of useful work produced can be measured. To estimate the LINEX production function the exergy service units were used and various econometric methods were applied to test the validity of the approach. It was shown that a very good fit could be obtained with an R^2 of 0.99, better than the fits for a standard Cobb-Douglas production function. Further tests showed cointegration between the variables and Granger causality tests showed evidence of long-run causality running from exergy as well as useful work to GDP.

Kander & Schön (2007) expanded insights from Ayres & Warr (2009) studying the development of energy and capital in the Swedish economy between 1870 and 2000. Distinguishing between the ratio of capital to total energy consumption and the ratio between capital and the part of energy that is actually utilized in the production process. The latter calculated via the thermal efficiency of a process, which is the ratio of usable energy output in comparison to energy inputs in terms of heat or work. The results show that the capital stock grew much faster over the entire period than energy use, with the ratio of capital to energy rising by 1.7% annually on average over the 130 year period. Although during the end of the period around the 1980s the ratio started to stabilize, which reflects the finding of Ayres & Warr (2009) that exergy efficiency remained nearly stable since the 1960s. Relative price of energy in relation to machinery rose significantly in the same period by a factor 3 between 1890 and 2000. As a probable underlying explanation for the long term development of capital over energy the changes in the price of energy versus machinery is given as this at large precedes capital to energy ratio changes.

> "The general pictures is thus that trend breaks in the price ratio (pE/pM) preceded changes in the volume relation (K/E), that in turn were reinforced by the subsequent crises on the energy market." (Kander & Schön 2007; p. 298)

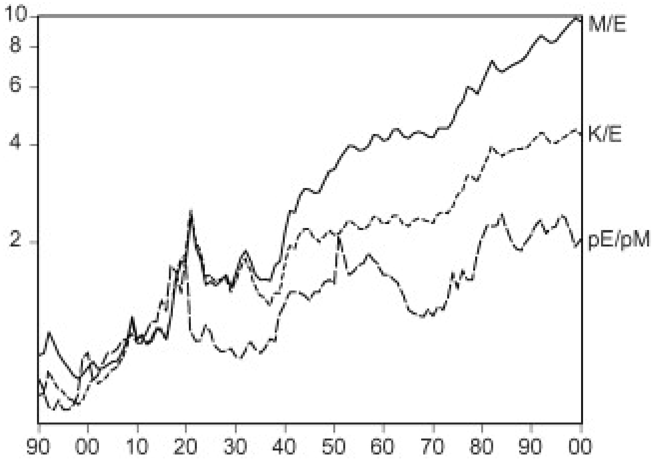


Figure 6 – Changes in the relative price of energy to machinery (pE/pM), capital to energy (K/E), and machinery to energy (M/E) for the Swedish economy between 1890 and 2000 after Kander & Schön (2007)

To test the hypothesis of prices inducing changing capital to energy ratios a model (about Sweden between 1890 and 2000) was estimated with relative price changes of energy to machinery (pE/pM) resulting in lagged change in the capital to energy ratio (K/E). The results show statistically significant coefficients. In case of looking at the ratio between capital and energy services, or the part of energy that is actually utilized in the production process, an hitherto entirely different picture emerged. Kander & Schön (2007) show the ratio between capital and energy services to be quite stable over the entire test period from 1900 to 1998 (short term fluctuations excluded). They concluded that capital and energy services are complementary and provided evidence that the increase in the capital to energy ratio is due to biased technical change, as less energy is required to provide the same quantity of capital, even though the capital itself may be different.

6.3 Summary

A few attempts have been conducted in the past to incorporate energy in macroeconomic growth models (Moon & Sonn 1996; Smulders & de Nooij 2003; Zon & Yetkiner 2003; Arbex & Perobelli 2010). These add energy as an input factor in production incorporating a partial aspect of energy to economic interaction. Including energy cost saving effects on production or the effect of (energy) factor productivity increases on output. The limited framework of these models prohibits a full analysis of the effect of energy on economic growth. Three problems are found with the approaches taken among which the assumption of full substitution between different input factors, the lack of compensation for different heating values of fuels, and the aggregate country level of measuring changes in energy consumption. The problem with assuming substitution between different input factors are physical limitations as energy is not substitutable for any other input except energy. Although changes in input factors are possible over time, these are likely not due to substitution but because of biased technical change with less energy being required per unit of capital as efficiency improves (Kander & Schön 2007). How the relationship between energy inputs and capital has developed over time is described in detail in Ayres & Warr (2009) who estimated the effect of improved energy efficiency over time for the United States and Japanese economy. In their model the total factor productivity parameter from the Solow and Swan A(t) model, $Y = A(t)K^{\alpha}L^{b}E^{\gamma}$, is replaced by exergy conversion efficiency as a proxy for technological progress. Also a variable E is added to represent energy inputs in production. Their transformed model shows a good fit in explaining economic growth, better than the standard Solow and Swan Cobb-Douglas model. In an extension Kander & Schön (2007) add an analysis of relative energy to capital price changes, showing that these precede the capital to energy ratio changes. The question that follows from these analyses is the order of events taking place. Does technological development result in increased energy efficiency, which causes reduced prices and hence an improved capital to energy ratio, or is the price reduction an artifact that does not influence the changing capital to energy ratio.

7. CONCLUSIONS AND RECOMMENDATIONS

In our industrial society energy inputs in the form of fossil fuel consumption play an essential role in providing transport, heat, and electricity. Every production process requires energy inputs, and in the present these inputs are derived for 88% from a finite high-quality fossil fuel energy stock. In practical terms “finite” means that fossil fuels can no longer meet energy constantly growing energy demand at some point in time, because increasing production becomes too costly. Eventually production will increase more slowly, reach a peak, and turn into irreversible decline. Economically speaking this will lead to an increasing price of fossil fuels and a shrinkage of demand to match stagnating or diminishing supplies. The first fossil fuel expected to reach such a decline is crude oil, with forecasts on the onset of decline ranging from now up to 2040. This range is so large due to uncertainty of estimates for oil stocks, demand growth and political and economic constraints to production growth. For instance, for oil production to grow significantly beyond 2015-2020 a key assumption is political stability in Iraq as this is one of the few remaining countries with large cheap to produce oil reserves. Despite these uncertainties it can be stated that in the coming decades a transition from oil to other energy sources in transport and petrochemical production will need to take place for economic growth to continue. In addition, because of the present scale of oil usage this poses a massive economic undertaking.

In mainstream economics this issue is, however, not taken into consideration in studies of economic growth. Energy has never been widely acknowledged as an input factor requiring explicit attention in economics. This is based on a view of the economic process driven by inputs of labor and capital and increasing the productivity of those factors through technological progress. Resulting in the assumption that there are no limits to substituting energy inputs for labor and capital (Sørensen & Whitta-Jacobsen 2005). In the few cases where energy has been treated as an explicit input factor, a further assumption is often made that due to technological progress easy substitution of fossil for non-fossil energy sources will occur. Based on the laws of thermodynamics it can be argued that assuming substitution of energy for non-energy sources is incorrect. Especially the second law, which in summarized form states that performing work always leads to a degradation from high to low quality energy, makes energy a prerequisite for any type of economic activity and hence an indispensable input in the production function. This implies that without continued increasing inputs of high quality energy, economic growth would cease, as substitution of energy for non-energy input factors is impossible. Substitution between fossil energy and labor is possible as labor is conversion of food produced with solar and fossil energy inputs into mechanical work. Thereby this substitution does not violate the second law of thermodynamics. This is only possible to a very small extent, however, as the energy requirements for production processes globally by now vastly exceeds the labor capacity of human beings. Technological progress has made it possible to increase the amount of work that can be performed per hour of labor, but this has come at the price of vastly increasing fossil fuel consumption. Because of this physical reality it becomes necessary to recognize energy as an explicit input, especially in light of expected scarcity, which is not unlikely in case of a limit to increasing oil production.

Regarding the assumption of easy substitution from fossil to non-fossil energy sources, it is true that in the past because of technological progress a number of transitions successfully occurred from using wood to coal, and from coal to a diverse fuel mix of oil, gas and coal. These transitions were a change, however, from lower to higher quality fuels in terms of energy density and cost of retrieval and conversion to their most useful state. Coal can generate more useful work per weight equivalent then wood, is more easily transported, and more versatile in use. The same holds true for oil and gas over coal. No energy sources have been found to date with similar qualities as fossil fuels, which is why substitution from fossil to non-fossil energy is an extremely difficult process. Energy sources such as wind, solar, biomass and nuclear are all more expensive in physical and monetary terms. The costs of non-fossil sources results from the technical complexity of nuclear power stations, and the low energy density that can be obtained from wind, solar and bio-energy relative to required inputs because these all are flow sources. Without a cheap form of energy storage, flow sources will remain too costly to implement at a large scale, contrary to fossil fuels which are nearly freely available stocks that only have to be extracted from the earth. As the substitution from fossil to non-fossil energy sources is difficult and costly and will remain so for at least the foreseeable future, it would be unwise to ignore the potential economic implications of a diminishing availability of (cheap) fossil fuels. As shown by Hamilton (1983, 2005, 2009) rising oil prices have had a significant recession inducing effect in the past as well as in the recent economic crisis. The chance for a repetition of energy constraints contributing to economic recessions is quite conceivable due to fossil fuel depletion. However, the negative effects on economic growth can be expected to be more severe as the chance is high that declining fossil fuel production will lead to long term scarcity, instead of the short term politically induced price rises of the past. The risks of such a turning point for societies hence makes it prudent to reconsider the role of energy in mainstream economics. Fortunately in past decades considerable insights have already been gained in this area, of which a large share has been summarized in this thesis focusing on efficiency, the rebound effect, economic analyses on the relation between energy and economic growth, and modeling energy in macroeconomics.

Although many aspects have been studied in this thesis two are found to be of key importance in understanding the role of energy in economics: A) how increased efficiency of transforming energy into useful work affects economic growth. Useful work being defined as the total work generated by energy inputs minus the energy lost in terms of heat dissipation or other non useful energy from the conversion process, and B) how technological progress has enabled the usage of increasing high quality energy over time for nearly all economic processes. Increasing the amount of empirical research on these aspects can lead to a more complete understanding of the role of energy in the economy. In addition knowledge on whether and in what manner these aspects will continue can lead to an improved understanding of the long term future development of the global economy in case of energy constraints.

The groundwork regarding energy conversion efficiency (A) has been laid out by Ayres & Warr (2009). Their work provides preliminary proof of a close relation between increased efficiency of transforming energy into useful work and economic growth. The process itself has been identified by showing vast increases of the amount of capital relative to required energy inputs over time, while the amount of capital to useful work performed remained fairly stable (Kander & Schön 2007). This approach appears to be much more promising than causality studies in which the relation between total energy consumption and economic growth is analyzed without taking conversion efficiency changes into account. More than fifty studies on the causal relation between total energy consumption and economic growth have led to inconclusive results as to whether an unidirectional, bidirectional or no causal relation can be found (Karanfil 2009; Ozturk 2010). Energy consumption itself is a prerequisite for economic growth but may not be a cause which is the likely reason for the inconclusive results. Taking useful work instead of total energy inputs may yield better results because the improvement of energy conversion efficiency is a good approximation for technological process and hence closely related to economic growth. This can be described in economic terms as technological progress resulting in biased technical change with less energy being required over time in relation to capital to maintain output production levels, caused by increasing the efficiency of energy conversion in industrial processes. In this manner a larger share of energy has become available over time making more economic output possible. Also, this has likely has led to enabling of energy use (B) as fossil fuel sources have become applicable to an increasing number of economic processes, making even more economic growth possible. From this it follows that to better anticipate for the future the extent to which conversion efficiency increases can continue needs to be studied. Based on the laws of thermodynamics there is a limit to the extent to which efficiency improvements can take place. This is not only because there is an upper bound efficiency improvements in creating work by transforming high to low quality energy, but also because there is a trade-off between efficiency versus power output in nature. As the rate of a process or power output increases, a larger portion of input energy will be lost as heat and vice versa (Hall et al. 1986). A maximally efficient process is hence a very slow process as the power output is very low. Hence, beyond a certain optimum of power output and efficiency, a more efficient process will need more time to perform a similar amount of work. Usually this optimum lies around 50% efficiency. This means that there is a limit for processes to become more efficient beyond a certain point, because power output would be too low for many processes in industrial society too still be economic when assuming a certain discount rate. The question is whether these limitations are already occurring, implying that efficiency is improving at a lower rate than in the past. Some evidence for this has been provided by Ayres and Warr (2009) by showing that the largest share of energy conversion improvements took place from 1900 to approximately 1960 in the US and Japanese economies due to the invention of new industrial processes. This shows that at least in western economies, the possibilities for further efficiency conversion improvements may prove to be limited. It is difficult to find firm proof for this, however, due to the globalized state of the world. An increasing number of goods is no longer produced in western countries but imported from abroad, the amount of energy used in their production no longer shows up in energy statistics, while GDP gains from trading and consuming these goods do. Studies on energy consumption and GDP do not take these transfers into account as they have been conducted on a country level. In addition, these studies are also confounded by apparent energy efficiency changes within countries not being caused by efficiency increases but by sector shifts within a nation itself and fuel mix changes (Patterson 1993; Wilson et al. 1994; Murtishaw & Schipper 2001; Geller et al. 2006). To improve knowledge more data is required on conversion efficiency changes on a global level. Preferably per type of technology as this can then be combined with knowledge from the domain of physics on the extent to which conversion efficiency improvements from total energy inputs to useful work are possible.

Next to the process itself also an understanding of driving mechanisms is required. Regarding the enablement of energy usage due to technological process (B), the rebound effect comes into play. This can be summarized as a shift in demand that occurs because of cost reductions that go together with efficiency increases in processes and/or in energy consumption. The reason is that the marginal utility of consuming a good becomes higher than the marginal cost of more consumption, hence the economy grows as more consumption can take place of the good which becomes more efficient to produce or consume. Once demand is satisfied for one good, freed up income can be used for other goods using the same amount of energy. In this manner saving energy through improving efficiency has a limited effect on overall energy consumption. This has been quantified in a comprehensive manner in the economic literature for the direct rebound effect, defined as an increase in consumption of the same good due to energy savings. It has been found that overall energy consumption rebounds at less than 30% for rich countries (Greening et al. 2004; Sorrell et al. 2009). In the literature developing countries have been ignored, this could influence the results as these have a different initial level of technology and this often enhances their potential for efficiency improvements. Also the potential for substitution could be higher in developing countries because more unsatisfied needs exist, although this is not of concern in direct rebound analyses. The effects of substitution, as well as other aspects including sector shifts, have been quantified in literature on the macroeconomic rebound effect. The literature on this is far from conclusive, however, as few analyses have been conducted, and the theoretical framework still lacks a number of important aspects. These aspects include the accounting for the energy transformation of labor, the energy content of imported goods (Barker et al. 2007), omission of financial flows in general equilibrium models (Allan et al. 2007; Hanley et al. 2009), the capital costs of energy efficiency improvements (Sorrell & Dimitropoulos 2008), and the disinvestment effect (Anson & Turner 2009). The testing approach for the macroeconomic rebound is to change energy input productivity values in the model, while keeping other factors constant, either by means of an instant efficiency shock or during a longer period of time. The majority of the studies show that the total rebound effect is larger than 50% with the variability of results being highly dependent on the assumed elasticity of substitution between input factors. This is one of the major points of study as in practice substitution depends on the chosen timeframe and physical possibilities which in case of energy are limited.

7.1 Results and recommendations

From the argumentation above, which summarizes this thesis, a number of conclusions are drawn:

- The assumption of unlimited substitutability between energy and non-energy inputs in mainstream economics is incorrect because of the laws of thermodynamics.
- The assumption of a smooth path of substitution from fossil to non-fossil energy sources is improbable and perhaps even impossible.
- Technological progress resulting in improved efficiency of energy conversion to useful work is an important driver of economic growth.
- There are important limits to improving energy efficiency due to physical boundaries in the form of maximum efficiency and the trade-off between power and efficiency.
- The limits to improving energy efficiency may already be in sight regarding many key processes in western economies.
- The rebound effect is one of the key mechanisms that enables more energy usage through technological progress resulting in cost reductions.

The consequence of these conclusions are significant, because a continuously increasing high-quality energy supply, needed in order to sustain the economic growth path (as experienced during most of the industrial revolution), will no longer be feasible.

To further investigate these conclusions two follow-up areas of research are suggested:

First, to investigate the effect of structural sector changes across countries and continents. This because many energy intensive activities have been shifted over time towards countries where labor and energy is cheaper. Such research would shed more light on whether country-level energy efficiency increases did actually occur in recent decades, or whether country-level energy efficiency changes are actually caused by or coupled with the geographical outsourcing of energy intensive production and importing these products from abroad.

Second, to bolster causality studies on energy consumption and economic growth. In one way by adding more variables such as technological progress and prices, and also by using the energy delivered in production, so excluding energy conversion losses, instead of using total energy inputs.

8. REFERENCES

Asafu-Adjaye, J., 2000. The relationship between energy consumption, energy prices and economic growth: time series evidence from Asian developing countries. *Energy Economics*, 22, p. 615-625.

Allan, G., Hanley, N., McGregor, P., Swales, K., Turner, K., 2007. The impact of increased efficiency in the industrial use of energy: A computable general equilibrium analysis for the United Kingdom. *Energy Economics*, 29, p. 779-798.

Anson, S., Turner, K., 2009. Rebound and disinvestment effects in refined oil consumption and supply resulting from an increase in energy efficiency in the Scottish commercial transport sector. *Energy Policy,* 37, p. 3608-3620.

Apostolakis, B.E., 1990. Energy-capital substitutability/complementarity: the dichotomy. *Energy Economics*, 12 (1), p. 48-57

Arbex, M., Perobelli, F.S., 2010. Solow meets Leontief: Economic growth and energy consumption. *Energy Economics*, 32, p. 43-53.

Ayres, R.U., Warr, B., 2009. *The Economic Growth Engine: How Energy and Work Drive Material Prosperity*. Cheltenham: Edward Elgar.

Barker, T., Ekins, P., Foxon, T., 2007. The macro-economic rebound effect and the UK economy, *Energy Policy*, 35, p. 4935-4946

Bentzen, J., 2004. Estimating the rebound effect in US manufacturing energy consumption. *Energy Economics*, 26, p. 123-134.

Bergh, J.C.J.M. van den, 2009. The GDP Paradox. *Journal of Economic Psychology*, 30, p. 117-135.

Bessec, M., Méritet, S., 2007. The Causality Link between Energy Prices, Technology and Energy Intensity. In Keppler, J.H., Bourbonnais, R., Girod, J., eds. *The Econometrics of Energy Systems*. Hampshire: Palgrave MacMillan. Ch. 6.

Brookes, L., 1978. Energy policy, the energy price fallacy and the role of nuclear energy in the UK. *Energy Policy*, 6 (2), p. 94- 106.

Çengel, Y.A., Boles, M.A., 2008. *Thermodynamics: An Engineering Approach*, 6th edition. New York: McGraw-Hill.

Commission of the European Communities, 2006. *Action Plan for Energy Efficiency: Realising the Potential* [online]. (Updated 19 Oktober 2006) Available at: http://ec.europa.eu/energy/action_plan_energy_efficiency/doc/com_2006_0545_en.pdf [accessed 26 April 2010].

Dolado, J., Gonzalo, J., Marmol, F., 2003. Cointegration. In Baltagi, B.H., ed. *A Companion to Theoretical Econometrics*. Malden: Blackwell Publishing. Ch. 30.

Dufournaud, M.C., Quinn, J.T., Harrington, J.J., 1994. An applied General Equilibrium (AGE) analysis of a policy designed to reduce the household consumption of wood in The Sudan. *Resource and Energy Economics*, 16, p. 67-90.

Farrel, A.E. et al., 2006. Ethanol Can Contribute to Energy and Environmental Goals. *Science*, 311, p. 506-508.

Gagnon, N., Hall, C.A.S., Brinker, L., 2009. A Preliminary Investigation of Energy Return on Energy Investment for Global Oil and Gas Production. *Energies*, 2(3), p. 490-503.

Georgescu-Roegen, N., 1971. *The Entropy Law and the Economic Process*. London: Harvard University Press.

Geller, H., Harrington, P., Rosenfeld, A.H., Tanishima, S., Unander, F., 2006. Policies for increasing energy efficiency: Thirty years of experience in OECD countries. *Energy Policy*, 54, p. 556-573.

Granger, C.W.J., 1969. Investigation Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica*. 37, p. 424-438.

Greening, L.A., Greene, D.L., Difiglio, C., 2000. Energy efficiency and Consumption – the Rebound Effect – a Survey. *Energy Policy*, 28, p. 389-401.

Grepperud, S. and Rasmussen, I, 2004. A General Equilibrium Assessment of Rebound Effects. *Energy Economics,* 26, p. 261-282.

Hall, C.A.S., Cleveland, C.J., Kaufmann, R., 1986. *Energy and Resource Quality: The Ecology of the Economic Process*. New York: John Wiley & Sons.

Hallock, J.L., Tharakan, P.J., Hall, C.A.S., Jefferson, Wu, W., 2004. Forecasting the limits to the availability and diversity of global conventional oil supply. *Energy*, 29, p. 1673-1696.

Hamilton, J.D., 1983. Oil and the Macroeconomy since World War II. *The Journal of Political Economy*, 91 (2), p. 228-248.

Hamilton, J.D., 2005. Oil and the Macroeconomy. [online]. (Updated August 24, 2005) Available at: dss.ucsd.edu/~jhamilto/JDH_palgrave_oil.pdf [accessed May 6 2010].

Hamilton, J.D., 2009. Causes and Consequences of the Oil Shock of 2007-08. *Brookings Papers on Economic Activity,* 1, p. 215-261.

Hanley, N. D., McGregor, P.G., Swales, J.K., Turner, K., 2009. Do increases in energy efficiency improve environmental quality and sustainability?. *Ecological Economics*. 68, p. 692-709.

Hannon, B., Joyce, J., 1981. Energy and Technical Progress. *Energy*. 6, p. 187-195.

Hu J., Lin, C., 2008. Disaggregated energy consumption and GDP in Taiwan: A threshold co-integration analysis, *Energy Economics*, 30, p. 2342-2358.

IIER, 2010. Personal communication with Hannes Kunz President of the Integrated Institute of Economic Research. Meilen, Zwitserland, Held Monday 12 April 2010.

Jayaraman, T.K., Choong, C., 2009. Growth and oil price: A study of causal relationships in small Pacific Island countries. *Energy Policy*, 37, p. 2182-2189.

Jevons, W.S., 1865. *The Coal Question: An Inquiry concerning the Progress of the Nation, and the Probably Exhaustion of our Coal-Mines* [e-book]. London: Macmillan and co. Available at: http://www.econlib.org/library/YPDBooks/Jevons/jvnCQ.html [accessed 8 March 2010].

Johansen, S., 1988. Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control,* 12, p. 231-254.

Johansen, S., Juselius, K., 1990. Maximum likelihood estimation and inference on cointegration – with applications to the demand for money. *Oxford Bulletin of Economics*, 51, p. 169-210.

Kahrl, F., Roland-Holst, D., 2008. Energy and Exports in China. *China Economic Review*, 19. p. 649-658.

Kander, A., Schön, L., 2007. The energy-capital relation – Sweden 1870-2000. *Structural Change and Economic Dynamics*, 18, p. 291-305.

Karanfil, F., 2008. Energy consumption and economic growth revisited: Does the size of the unrecorded economy matter?. *Energy Policy*, 36, p. 3029-3035.

Karanfil, F., 2009. How many times again will we examine the energy-income nexus using a limited range of traditional econometric tools?, *Energy Policy,* 37, p. 1191-1194.

Keppler, J.H., 2007.Causality and Cointegration between Energy consumption and Economic Growth in Developing Countries, J.H., Bourbonnais, R., Girod, J., ed. *The Econometrics of Energy Systems.* Hampshire: Palgrave MacMillan. Ch.4.

Koppelaar, 2010. Personal communication with Prof. Dr. Henk of TU/Delft. Breda, the Netherlands, Held Saturday 5 June 2010.

Krugman, P., Wells, R., Graddy, K., 2008. *Economics.* Worth Publishers, New York, United States.

Laitner, J.A., 2000. Energy efficiency: rebounding to a sound analytical perspective. *Energy Policy*, 28, p. 471-475.

Lütkepohl, H., 2003. Vector Autoregressions. In Baltagi, B.H., ed. *A Companion to Theoretical Econometrics*. Malden: Blackwell Publishing. Ch. 32.

Madlener, R., Alcott, B., 2009. Energy rebound and economic growth: A review of the main issues and research needs. *Energy,* 34, p. 370-376.

Malthus, T.R., 1798. *An Essay on the Principle of Population* [e-book]. London: John Murray. Available at: http://www.econlib.org/library/Malthus/malPlong.html [accessed 7 March 2010]

Maria, C.D., Ferreira, S., Lazarova, E., 2010. Shedding light on the light bulb puzzle: the role of attitudes and perceptions in de adoption of energy efficient light bulbs. *Scottish Journal of Political Economy*, 57 (1), p. 48-67.

Moon, Y., Sonn, Y., 1996. Productive energy consumption and economic growth: An endogenous growth model and its empirical application. *Resource and Energy Economics*, 18, p. 189-200.

Murtishaw, S., Schipper, L., 2001. Disaggregated analysis of US energy consumption in the 1990s: evidence of the effects of the internet and rapid economic growth. *Energy Policy*, 29, p. 1335-1356.

OSRAM, 2009. *Life Cycle Assessment of Illuminants: A Comparison of Light Bulbs, Compact Fluorescent Lamps and LED Lamps, Executive Summary*. [online]. Available at http://www.osram-os.com/osram_os/EN/About_Us/We_shape_the_future_of_light/Our_obligation/LED_life-cycle_assessment/OSRAM_LED_LCA_Summary_November_2009.pdf [accessed 4 april 2010]

Ozturk, I, 2010. A literature survey on energy-growth nexus. *Energy Policy*, 38, p. 340-349.

Patterson, M.G., 1993. An accounting framework for decomposing the energy-to-GDP ratio into its structural components of change. *Energy*, 18(7), p. 741-761.

Patterson, M.G., 1996. What is energy efficiency? Concepts, indicators and methodological issues. *Energy Policy*, 24(5), p. 377-390.

Payne, J.E., 2010. A survey of the electricity consumption-growth literature. *Applied Energy*. 87, p. 723-731.

Pen, Y.L., Sévi, B., 2010. What trends in energy efficiencies? Evidence from a robust test. *Energy Economics*, 32, p. 702-708.

Pokrovski, V. N., 2003. Energy in the theory of production. *Energy*, 28, p. 769-788.

Polimeni, J.M., Mayumi, K., Giampietro, M., Alcott, B., 2008. *The Jevons Paradox and the Myth of Resource Efficiency Improvements*. London: Earthscan.

Ricardo, D., 1817. *On the Principles of Political Economy and Taxation* [e-book]. London: John Murray. Available at: http://www.econlib.org/library/Ricardo/ricP.html [accessed 7 March 2010].

Rostow, W.W., 1990. *Theorists of Economic Growth from David Hume to the Present: With a Perspective on the Next Century.* New York: Oxford University Press.

Roy, J., 2000. The rebound effect: some empirical evidence from India. *Energy Policy*. 28, p. 433-438.

Ruzzenenti, F., Basosi, R., 2008. The role of the power/efficiency misconception in the rebound effect's size debate: Does efficiency actually lead to a power enhancement?. *Energy Policy*, 36, p. 3626-3632.

Smil, V., 2005. *Creating the Twentieth Century: Technical Innovations of 1867-1914 and their lasting impact*. 2005 printing. New York: Oxford University Press.

Smith, A., 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations* [e-book]. London: Methuen & Co. Available at: http://www.econlib.org/library/Smith/smWN.html [accessed 7 March 2010].

Smulders, S., de Nooij, M., 2003. The impact of energy conservation on technology and economic growth. *Resource and Energy Economics*. 25, p. 59-79.

Snowdon, B., Vane, H.R., 2005. *Modern Macroeconomics: Its Origins, Development and Current State*. Cheltenham: Edward Elgar Publishing.

Soddy, F., 1929. *Wealth, Virtual Wealth and Debt: The Solution of the Economic Paradox*. 1983 printing. London: George Allen & Unwin.

Sorrell, S., Dimitropoulos, J., 2008. The rebound effect: Microeconomic definitions, limitations and extensions. *Ecological Economics*. 65, p. 636-649.

Sorrell, S., Dimitropoulos, J., Sommerville, M., 2009. Empirical estimates of the direct rebound effect: A review. *Energy Policy*. 37, p. 1356-1371.

Sørensen, P.B., Whitta-Jacobsen, H.J., 2005. *Introducing Advanced Macroeconomics: Growth & Business Cycles.* McGraw-Hill Education, Berkshire, England.

Vikström, P., 2004. Energy efficiency and Energy Demand: A Historical CGE Investigation on the Rebound effect in the Swedish Economy 1957. *Working Paper*, Umea University, Sweden.

Washida, T., 2004. Economy-wide Model of Rebound Effect for Environmental Efficiency. *International Workshop on Sustainable Consumption*, University of Leeds.

Wilson, B., Trieu, L.H., Bowen, B., 1994. Energy Efficiency Trends in Australia. *Energy Policy*, 22(4), p. 287-295.

Zhang, D., 2008. Oil shock and economic growth in Japan: A nonlinear approach. *Energy Economics*, 30, p. 2374-2390.

Zon, A van., Yetkiner, I.H., 2003. An endogenous growth model with embodied energy-saving technical change. *Resource and Energy Economics*, 25, p. 81-103.

WAGENINGEN UNIVERSITY

WAGENINGEN UR