

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES

Issue: *Ecological Economics Reviews*

The role of energy in economic growth

David I. Stern

Arndt-Corden Department of Economics, Crawford School of Economics and Government, Australian National University, Canberra, Australia

Address for correspondence: David I. Stern, Arndt-Corden Department of Economics, Crawford School of Economics and Government, Australian National University, Canberra, ACT 0200, Australia. sterndavid@yahoo.com

This paper reviews the mainstream, resource economics, and ecological economics models of growth. A possible synthesis of energy-based and mainstream models is presented. This shows that when energy is scarce it imposes a strong constraint on the growth of the economy; however, when energy is abundant, its effect on economic growth is much reduced. The industrial revolution released the constraints on economic growth by the development of new methods of using coal and the discovery of new fossil fuel resources. Time-series analysis shows that energy and GDP cointegrate, and energy use Granger causes GDP when capital and other production inputs are included in the vector autoregression model. However, various mechanisms can weaken the links between energy and growth. Energy used per unit of economic output has declined in developed and some developing countries, owing to both technological change and a shift from poorer quality fuels, such as coal, to the use of higher quality fuels, especially electricity. Substitution of other inputs for energy and sectoral shifts in economic activity play smaller roles.

Keywords: energy; economic growth; technological change; substitution; industrial revolution

Preferred citation: David I. Stern. 2011. The role of energy in economic growth in "Ecological Economics Reviews." Robert Costanza, Karin Limburg & Ida Kubiszewski, Eds. *Ann. N.Y. Acad. Sci.* 1219: 26–51.

Introduction

Toman and Jemelkova argue that most of the literature on energy and economic development discusses how development affects energy use rather than the reverse.¹ The principal mainstream growth models² do not include energy as a factor that could constrain or enable economic growth, though macroeconomics pays significant attention to the impact of oil prices on economic activity in the short run.³ Resource economists have developed models that incorporate the role of resources, including energy, in the growth process, but these ideas remain isolated in the resource economics field. Ecological economists, on the other hand, often ascribe to energy the central role in economic growth. So, is energy an important driver of economic growth and development and, if so, what factors affect the strength of the relationship between energy and growth? This paper attempts to answer these questions by surveying the literature and bringing some of the strands together in a synthesis. In the remain-

der of this introduction, I outline the plan of this review.

We cannot understand the role of energy in economic growth without first understanding the role of energy in production. Therefore, I first review basic physical principles and economic concepts that define the role of energy in economic production. Because thermodynamics implies that energy is essential to all economic production, criticism of mainstream economic growth models that ignore energy is legitimate. On the other hand, theories that try to explain growth entirely as a function of energy supply, while ignoring the roles of information, knowledge, and institutions, are also incomplete.

However, institutions also affect how these roles play out and, therefore, the mainstream theory of economic growth, which focuses on these considerations, is reviewed next. I discuss ecological economics views on energy and growth in the following section. A significant component of the ecological economics contribution is the critique of mainstream theory. These criticisms focus on

limits to substitution between energy and other inputs and limits to technological change.

As a step toward reconciling mainstream and ecological economics models of economic growth, I discuss theories that can explain economic growth in the long run including the industrial revolution that transitioned the preindustrial economy to the modern economic growth regime. I present a simple model that embeds the Solow growth model within a more general framework where energy and capital are poor substitutes. This shows that when energy is scarce it will strongly constrain economic growth, but when energy becomes more abundant it is much less of a limiting factor and the conventional mainstream model explains economic growth fairly well.

In recent decades, and in some cases over much longer periods, significant reductions in energy intensity have been achieved in many developed and some developing countries.^{4,5} I use a production function approach to examine the factors that could reduce or strengthen the linkage between energy use and economic activity over time. The key factors are as follows:

- substitution between energy and other inputs within an existing technology;
- technological change;
- shifts in the composition of the energy input; and
- shifts in the composition of economic output.

I dedicate a subsection to the discussion of each of these themes. Penultimately, I review the empirical evidence on the causal link between energy and growth focusing on reduced form time-series models that do not specify the structural linkages between energy and output. This literature shows that energy and gross domestic product (GDP) cointegrate and energy use Granger causes GDP when capital and other production inputs are included in the vector autoregression model. The final section of the paper presents conclusions.

Energy as a factor of production

The laws of thermodynamics and the conservation of matter describe the immutable constraints within which the economic system must operate.^{6,7} The mass-balance principle means that, in order to obtain a given material output, greater or equal quantities of matter must be used as inputs with the residual a pollutant or waste product.⁶ Therefore,

there are minimal material input requirements for any production process producing material outputs. The second law of thermodynamics (the efficiency law) implies that a minimum quantity of energy is required to carry out the transformation or movement of matter or, more generally, perform physical work. Carrying out transformations in finite time requires more energy than these minima.⁸ All production involves work. Therefore, all economic activities must require energy, and there must be limits to the substitution of other factors of production for energy so that energy is always an essential factor of production.⁹

Some inputs to production are nonreproducible, while others can be manufactured, at a cost, within the economic production system and are said to be reproducible. Capital, labor, and, in the longer term, even natural resources, are reproducible factors of production, while energy and matter are nonreproducible factors of production. Energy vectors—fuels and electricity—and raw materials such as minerals are in theory reproducible factors, though with the exception of agricultural and forest products they are usually harvested from nature and represent the accumulated work of the planet's biogeochemical cycles, which in turn are powered by energy from the sun and the Earth's internal heat.

As neither energy nor matter is reproducible they must be captured from the environment with implied environmental disruption. This is especially relevant for energy, as due to the entropy law, exergy—useful energy—cannot be recycled. While some forms of energy capture are possibly more hazardous to human health or damaging to environmental quality, all methods—nuclear energy, fossil fuels, hydropower, wind power, biomass, etc.—are environmentally disruptive. Solar energy is very diffuse compared to the concentrated stocks of fossil fuels and plants and animals are very inefficient converters of that energy into exergy and useful work, so that large quantities of energy must be invested to obtain a given amount of net energy.¹⁰ This is why the shift to fossil fuels in the industrial revolution released the prevalent constraints on energy supply, and therefore, on production and growth.¹¹

Some aspects of organized matter—that is information—might also be considered to be nonreproducible inputs in the same way as energy is.^{12–14} Energy is necessary to extract information from the environment while active use of energy

cannot be made without information and, possibly, accumulated knowledge. Obviously, energy can provide uncontrolled heating, lighting, etc., without any activity on the part of economic agents. But even nonintelligent organisms need to use (genetic) information to make controlled use of energy. Unlike energy, information and knowledge cannot be easily quantified. However, the fact that they must be incorporated into machines, workers, and materials in order to be made useful provides a biophysical justification for treating capital, labor, etc., as factors of production. Though capital and labor are easier to measure than information and knowledge, their measurement is, still, very imperfect compared to that of energy.

Primary factors of production are defined as inputs that exist at the beginning of the period under consideration and are not directly used up in production (though they can be degraded or accumulated from period to period), while intermediate inputs are those created during the production period under consideration and are used up entirely in production. Mainstream economists usually think of capital, labor, and land as the primary factors of production, and goods (such as fuels and materials) as intermediate inputs. The prices paid for the various intermediate inputs are seen as eventually being payments to the owners of the primary inputs for the services provided directly or embodied in the produced intermediate inputs. This approach has led to a focus in mainstream growth theory on the primary inputs, and in particular, capital and labor. The classical factor of land, including all natural resource inputs, gradually diminished in importance in economic theory as its value share of GDP fell steadily in the 20th century¹⁵ and today is usually subsumed as a subcategory of capital.

The mainstream theory of growth

Background: growth theory

The core mainstream growth models, including the Solow¹⁶ growth model and later developments,² do not include resources or energy. In the Solow model, the long-run rate of economic growth is given by the rate of technological progress while diminishing returns to capital mean that growth derived from capital accumulation is a transitional phase as the economy moves toward this long-run dynamic steady state. If the savings rate is increased, growth will occur for a while until a new equilibrium is reached,

though, the higher the savings rate, the lower the current standard of living of the population. Improvement in the state of technological knowledge raises the rate of return to capital, thereby offsetting the diminishing returns to capital that would otherwise apply a brake to growth. Without technological change, economic growth eventually halts but the economy can remain in a steady state indefinitely.

Early growth models did not explain how improvements in technology come about, so that these models are said to have exogenous technological change. More recent models attempt to endogenize technological change—explaining technological progress as the outcome of decisions taken by firms and individuals. Early endogenous growth models, such as Arrow's learning by doing model¹⁷ or Hicks' induced innovation model,¹⁸ allowed the state of technology to respond to changes in one of the variables in the model but do not explicitly model this as an optimizing process. More recent endogenous growth models are represented by the so-called AK models, which do not explicitly model research and development activities (R&D), and Schumpeterian growth models that do. In AK models, the relationship between capital and output can be written in the form $Y = AK$, where A is a constant and K is a composite of manufactured capital and disembodied technological knowledge is thought of as a form of capital. Technological knowledge is a nonrival good and thus is not depleted with use, and it generates positive externalities in production that benefit the economy as a whole. The economy can sustain a constant growth rate in which the diminishing returns to manufactured capital are exactly offset by the external effect of knowledge creation. In AK models, saving is directed to either manufactured capital accumulation or the increase of knowledge. The growth rate is permanently influenced by the savings rate; a higher savings rate increases the economy's growth rate, not merely its equilibrium level of income.¹⁹

In Schumpeterian growth models, there is imperfect competition in the capital goods industry and firms invest in R&D in order to receive monopoly profits.²⁰ Innovations appear stochastically and are embodied in new generations of capital or consumer goods. There are positive externalities to consumers or producers of final goods who benefit from innovation and to future researchers who benefit from past ideas. On the other hand, new innovations

make old vintages of capital obsolete. Both capital accumulation and innovation can determine the long-run growth rate. However, if there are diminishing returns in the innovation sector as technology becomes more complex, the economy could have a constant growth rate.²⁰

Growth models with resources and no technical change

Adding nonrenewable natural resources that are essential in production to the basic mainstream growth models means that capital also needs to be accumulated to compensate for resource depletion. Renewable but finite inputs, such as solar energy, also can pose problems. Nondeclining output is no longer guaranteed but instead depends on both the nature of technology and institutional arrangements.^{21,22} When there is more than one input—both capital and natural resources—there are many alternative paths that economic growth can take, determined by both the nature of technology and institutional arrangements. Technical conditions include the mix of renewable and nonrenewable resources, the initial endowments of capital and natural resources, and the ease of substitution among inputs. The institutional setting includes market structure (competition versus central planning), the system of property rights (private versus common property, etc.), and the system of values regarding the welfare of future generations. Models can attempt to either maximize the sum of discounted social welfare over some relevant time horizon or achieve sustainability (nondeclining social welfare).

Solow showed that sustainability is achievable in a model with a nonrenewable natural resource with no extraction costs and nondepreciating capital when the elasticity of substitution between the two inputs is unity, and when certain other technical conditions are met.²¹ Sustainability, and even indefinite growth in consumption, can occur when the utility of individuals is given equal weight without regard to when they happen to live. However, under competition the same model economy results in exhaustion of the resource and consumption and social welfare eventually fall to zero.²² Dasgupta and Heal show that with any constant discount rate the efficient growth path also leads to eventual depletion of the natural resource and the collapse of the economy.²³ The Hartwick rule shows that if sus-

tainability is technically feasible, a constant level of consumption can be achieved by reinvesting the resource rents in other forms of capital, which in turn can substitute for resources.²⁴ Dixit *et al.*²⁵ extended the rule to multiple capital stocks while Hartwick²⁶ extended the rule to open economies. It is difficult to apply this rule in practice, as the rents and capital must be valued at sustainability compatible prices, that is, the prices that would emerge if the sustainability constraint were imposed.^{27–29} In a practical sense, such prices are unknowable given that we have poor understanding of even the costs of current environmental damage and resource depletion or of the future development of the economy.

A common interpretation of this body of work is that substitution and technical change can effectively decouple economic growth from the use of energy and other resources. Depleted resources can be replaced by more abundant substitutes, or by “equivalent” forms of human-made capital (people, machines, factories, etc.). But this is a misinterpretation. Solow²¹ and others assume *a priori* that sustainability is technically feasible and explicitly dispose of cases where the elasticity of substitution between nonrenewable resources and capital is greater or less than unity. In the former case, substitution possibilities are large and, therefore, the possibility of nonsustainability is not an issue. In the latter case, sustainability is not feasible if an economy uses only nonrenewable resources, though sustainability is technically feasible when resources are renewable, at least in the absence of population growth. However, there is a tendency among mainstream economists to assume that sustainability is technically feasible unless proven otherwise.^{30–32}

Growth models with resources and technical change

In addition to substitution of capital for resources, technological change might permit growth or at least constant consumption in the face of a finite resource base. Stiglitz showed that, when the elasticity of substitution between capital and resources is unity, exogenous technical progress will allow consumption to grow over time if the rate of technological change divided by the discount rate is greater than the output elasticity of resources.³³ Technological change might enable sustainability even with an elasticity of substitution of less than

one. Once again, technical feasibility does not guarantee sustainability. Depending on preferences for current versus future consumption, technological change might instead result in faster depletion of the resource.³⁴

Recent work applies endogenous growth theories to analyzing capital-nonrenewable resource economies.^{20, 35–39} Smulders⁴⁰ provides a survey of earlier endogenous growth work, and Smulders and de Nooij⁴¹ and Di Maria and Valente³⁹ provide references to the more recent literature. An aim of much of this literature is to determine whether, and under what circumstances, technical progress is effective in ensuring sustained consumption.⁴² As noted earlier, due to externalities in knowledge production, there may be too little innovation in an endogenous growth world. As a result, depletion of a nonrenewable resource is nonoptimal, but this rate could be either too fast or too slow. A general finding is that the rate of resource augmenting progress must be strictly positive and at least equal to the discount rate to obtain nondeclining consumption in the long run.³⁹ Aghion and Howitt²⁰ found that an AK-type model with essential nonrenewable resources cannot allow for unbounded growth in consumption while a Schumpeterian-type model can. A Schumpeterian model with a renewable resource that affects utility directly can allow unlimited growth, but only under unlikely assumptions. But if the renewable resources do not affect utility, continued growth would be easier than in the nonrenewable case.

The ecological economics approach

Introduction

Ecological economists derive their view of the role of energy in economic growth from the biophysical foundations of the economy discussed earlier.^{10, 43–50} This view is shared by some geographers⁵¹ and economic historians^{11, 52} who also believe that energy plays a crucial role in economic growth, as well as being an important factor in explaining the industrial revolution. Ecological economists also argue that substitution between capital and resources and technological progress can only play limited roles in mitigating the scarcity of resources.⁹ Some ecological economists also downplay the role of technological change, arguing that either increased energy use accounts for most apparent productivity growth, or that technological change is real but innovations mainly increase pro-

ductivity by allowing the use of more energy.^{10, 45, 47} Therefore, increased energy use is the main or only cause of economic growth.

A prominent tradition in ecological economics is represented by biophysical models that consider energy to be a primary factor of production and the only such primary factor. In this view, all value is derived from the action of energy that is directed by capital and labor. The flow of energy in the economy is the service of the reservoirs of fossil fuels and the sun, which represent the primary input in our terminology. In some biophysical economic models, geological constraints fix the rate of energy extraction so that the flow rather than the stock can be considered a primary input.⁵³ On the other hand, capital and labor are treated as flows of capital consumption and labor services rather than as stocks, in other words, they are considered as intermediate inputs that are created and maintained by the primary input of energy and flows of matter. The level of the flows is computed in terms of the embodied energy use associated with them. Prices of commodities should then be determined by embodied energy cost⁵⁴—a normative energy theory of value—or are seen as actually correlated with energy cost⁴⁴—a positive energy theory of value.⁵⁵ Assuming that there are constant returns to scale, we can represent the production process of the economy as a whole by a Leontief input–output model with a single primary factor of production.⁵⁶ This theory—like the Marxian paradigm—must then explain how labor, capital, etc., end up receiving part of the surplus. Energy surplus must be appropriated by the owners of labor, capital, and land^{10, 44, 53, 57} with the actual distribution of the surplus depending on the relative bargaining power of the different social classes and foreign suppliers of fuel.⁵⁷

Because information is also an important factor of production it is difficult to argue for a model where energy is the sole factor of production. For example, the quality of resources such as oil reservoirs—as measured by depth, pressure, etc.—is critical in determining the energy required to extract and process fuels and other intermediate resource flows, which increases as the quality of resources declines over time with depletion. Odum's emergy approach⁵⁸ and the framework developed by Costanza⁴⁴ address this issue by including the solar and geological energy embodied in minerals and other natural resource inputs into indicators of total

embodied energy. Thus, changing resource quality is represented by changes in the embodied energy of the primary resources themselves. An alternative approach is to measure material and energy inputs on the common basis of their exergy.^{59,60} But both approaches seem too reductionist. Other services provided by nature, such as nutrient recycling, the provision of clean air and water, pollination, the climate system, and so on, should also then be accounted for. These ecosystem services provide the conditions that make economic production—and life itself—possible. Georgescu-Roegen formulated a more flexible approach with a variety of different types of factors of production.⁴³ The neo-Ricardian models⁶¹ developed by Perrings⁶² and O'Connor⁶³ also allow any number of inputs while complying with thermodynamic and mass-balance constraints.

Energy return on investment and economic output

Energy return on investment (EROI) is the ratio of useful energy produced by a method of energy supply to the amount of energy invested in extracting that energy. Lower quality energy resources have lower EROIs. Biophysical economists argue that the more energy that is required to extract energy, the less energy is available for other uses and the poorer an economy will be. In preindustrial societies, most workers were engaged in growing food and collecting fuel. Only a small fraction of society could use the small energy surplus generated to produce other products and services. In this view, the increase in EROI allowed by the switch from biomass to fossil fuels enabled the industrial revolution and the period of modern economic growth that followed it.¹⁰

Declining EROI would threaten not just growth but also the level of output of the economy and, therefore, sustainability. Murphy and Hall document EROI for many energy sources, arguing that it is declining over time.⁵⁰ Wind and direct solar energy have more favorable EROIs than biomass fuels but worse than most fossil fuels. However, unlike fossil fuels, the EROI of these energy sources tends to improve over time with innovation.⁶⁴ But current usage is very small, and Murphy and Hall argue that there is no prospect of them replacing a large part of fossil fuel usage in the near future.

Declining EROI could be mitigated by substituting other inputs for energy or by improving the efficiency with which energy is used. However, biophysical economics argues that both these processes have limits.

Limits to substitution

Substitution can occur *within* a category of similar production inputs—for example between different fuels—and *between* different categories of inputs—for example between energy and machines. There is also a distinction to be made between substitution at the microlevel—for example in a single engineering process or in a single firm—and at the macrolevel—in the economy as a whole. Finally, some types of substitution that are possible in a single country are not possible globally. We can classify inputs as good or poor substitutes as measured by the Hicks or direct elasticity of substitution.⁶⁵ This elasticity reflects movement along an isoquant of the production function holding all other inputs constant. It can take values from zero (Leontief function) to infinity (linear production). Good and poor substitutes have elasticities of substitution of greater than and less than unity, respectively.

Solow argues that the within-category substitution, and in particular the substitution of renewable for nonrenewable resources, is most important and seems to assume that new substitutes will always be found.³² It is possible that the elasticity of substitution for within-category types of substitution exceeds unity. The long-run pattern of energy use in industrial economies has been dominated by substitutions from wood and waterpower to coal, oil, natural gas and primary electricity (see Fig. 1).^{10,66} When fossil fuels are economically exhausted, the next stage of energy development may see a return to solar energy, albeit captured in a more sophisticated way, rather than a move to a new substitute.⁶⁷ Meta-analysis of existing studies of interfuel substitution suggests that the long-run elasticities of substitution at the level of the industrial sector as a whole for oil–coal, oil–gas, oil–electricity, and gas–electricity pairs are significantly greater than unity. The remaining elasticities—between coal and gas and coal and electricity—are insignificantly different from unity or zero due to their large standard errors. But there seems to be less substitutability at the macrolevel.⁶⁸

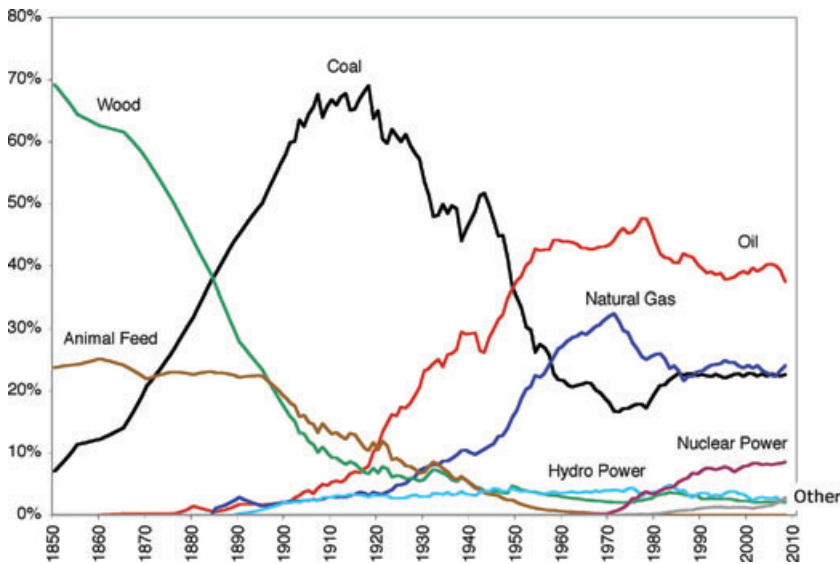


Figure 1. Composition of U.S. primary energy input, 1850–2008. Sources: Schurr and Netschert,¹⁴⁷ U.S. Energy Information Administration,¹⁸⁹ and U.S. Department of Commerce, Bureau of the Census.¹⁹⁰ (In color in *Annals online*.)

Ecological economists emphasize the importance of limits to the between-category type of substitution, and in particular, the substitution of manufactured capital for resources including energy.⁶⁹ A number of arguments for limited substitutability have been put forward, with the main ones that are relevant to the energy case described earlier. A meta-analysis of the existing empirical literature finds that the elasticity of substitution between capital and energy is less than unity.⁷⁰

Thermodynamic limits to substitution can be easily identified for individual processes but determining such limits for more complex technologies is difficult.^{71,72} It might be argued that standard production functions can account for mass balance and thermodynamic constraints if the elasticity of substitution between capital and resources is less than or equal to unity so that resources are essential. When the elasticity of substitution is unity (Cobb–Douglas production function), output is only zero when the energy input is zero (assuming positive nonenergy inputs), and strictly positive otherwise. This at least accounts for the fact that *some* amount of energy is required to produce goods and services but that this “essential” amount can be infinitesimal if sufficient manufactured capital is applied. Therefore, this condition does not satisfy thermodynamic considerations throughout the domain of the function. Thermodynamic limits can be approximated by a

production function with an elasticity of substitution significantly below unity.

Georgescu-Roegen’s fund-flow model describes production as a transformation process in which a flow of materials, energy, and information—the material cause—is transformed by two agents of transformation, human labor, and manufactured capital—the efficient cause.⁷³ Thus, Daly argues that adding to the stock of pulp mills (efficient cause) does not produce an increase in pulp unless there also is the wood fiber (material cause) to feed them.⁷⁴ From this perspective, capital should be a poor substitute for energy and other resources. Mainstream economists think about this question differently. First, they argue that though additional capital cannot conjure wood fibers out of thin air, more capital and “smarter” capital can be used with each amount of wood fibers to produce more sophisticated and valuable products. There is then no limit to the potential value of product created through sophisticated manipulation using larger amounts of capital.⁷⁵

In addition to these microeconomic limits to substitution, there may also be macroeconomic limits to substitution. The construction, operation, and maintenance of tools, machines, and factories require a flow of materials and energy. Similarly, the humans that direct manufactured capital consume energy and materials in the form of food, water,

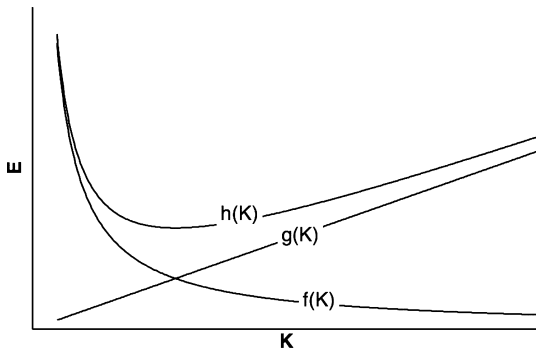


Figure 2. Macrolevel limits to substitution.

and other subsistence needs. Thus, producing more of the “substitute” for energy—manufactured capital—requires more of the thing that it is supposed to substitute for. This again limits potential substitutability.^{45,76–79}

Substitution does seem to be fundamentally more constrained at the economy-wide level of analysis than at the microlevel of analysis. In Figure 2, the curve $E = f(K)$ is an isoquant for a constant level of output, where E is energy, and K is capital.⁹ The indirect energy costs of materials are represented by $g(K)$ assuming, for simplicity, a constant energy cost per unit of capital. The diagram also unrealistically assumes that no capital is required in the extraction or capture of energy. Addition of direct and indirect energy costs results in the “net” isoquant $E = h(K)$, which has a positive slope for extreme capital/energy ratios. But the energy requirements of producing a unit of the capital input could also decrease if capital is also substituted for energy in the capital-producing sector of the economy. This would be the case if we were modeling the effects of an economy-wide tax on energy for example. However, a meta-analysis of the capital-energy elasticity of substitution shows a lower elasticity for more aggregated sectors than for less aggregated sectors,⁷⁰ which supports the idea that substitution is more restricted at the economy-wide level.

Kaufmann and Azary-Lee demonstrate the importance of accounting for this physical interdependency between manufactured and natural capital.⁸⁰ They found that from 1958 to 1984 the indirect energy costs of capital offset a significant fraction of the direct fuel savings in the U.S. forest products sector. In some years, the indirect energy costs of capital were greater than the direct fuel savings of using more capital.

Limits to technological change

The argument that resource augmenting technological change can overcome limited substitutability would be more convincing if technological change were really something different from substitution. This is not really the case. The neoclassical production function assumes that an infinite number of efficient techniques coexist at any one point in time. Substitution occurs among these techniques. Changes in technology occur when new more efficient techniques are developed. However, these new techniques really represent the substitution of knowledge for the other factors of production. The knowledge is embodied in improved capital goods and more skilled workers and managers. But there are still thermodynamic restrictions on the extent to which energy and material flows can be reduced in this way. The difference between knowledge and other inputs is that knowledge is nonrival in use—in other words, the same idea can be used simultaneously in different locations and production processes without any reduction in the productivity of the knowledge in the different locations and processes. This means that there can be constant returns to the application of knowledge in production while other inputs experience diminishing returns. But knowledge must be used in conjunction with the other inputs, such as energy, and the productivity of knowledge is still determined by the available quantities of those inputs.

Synthesis: unified models of energy and growth

The mainstream growth models discussed earlier either ignore energy or investigate whether limited energy or other resources could be a constraint on growth. By contrast, the ecological economics literature posits a central role for energy in driving growth but argues that limits to substitutability and/or technological change might limit or reverse growth in the future. But none of the models and theories reviewed so far really provides a satisfactory explanation of the long-run history of the economy. Until the industrial revolution, output per capita was generally low and economic growth was not sustained.⁸¹ Ecological economists and other researchers, discussed earlier in the paper, point to the invention of methods to use fossil fuels as the cause of the industrial revolution. But the mainstream growth

models that ignore energy resources can at least partly explain economic growth over the last half a century.⁸²

There are currently two principal mainstream theories that explain the growth regimes of both the preindustrial and modern economies and the cause of the industrial revolution, which formed the transition between them. The endogenous technical change approach, represented by Galor and Weil⁸³ and Lucas,⁸⁴ emphasizes the role of human capital and fertility decisions in the transition. The rate of technological change in Galor and Weil's model is a function of the size of the population and the level of education. Initially, there is a low technological change and education steady-state equilibrium. As population grows, a second high technological change and education equilibrium emerges, which eventually is the only equilibrium.⁸⁵

The second approach, represented by the work of Hansen and Prescott,⁸⁶ models the slow transition from a stagnating traditional economic sector—the Malthusian Sector—to a modern economic sector that experiences sustained economic growth—the Solow Sector. The “Malthusian Sector” depends on a land input and has decreasing returns to combined labor and capital. The modern “Solow Sector” does not use land and has constant returns to capital and labor combined. The transition starts when technical change first makes the operation of the modern technology profitable. Over time, the modern sector grows and the traditional sector declines as technological progress increasingly reduces the cost of production in the modern sector. Though Hansen and Prescott note that the Solow sector probably utilizes a technology that uses fossil fuels in place of land, they do not model energy explicitly.

Economic historians place a greater emphasis on the role of energy in the industrial revolution. Wrigley discussed the differential development of the Dutch and British economies.¹¹ Both countries had the necessary institutions for the industrial revolution to occur, but capital accumulation in the Netherlands faced a renewable energy resource constraint while in Britain domestic coal mines provided a way out from the constraint. This explanation emphasizes low substitutability between capital and energy with newly available energy sources relaxing the binding constraint on capital accumulation. With restricted energy supplies, capital accumulation faced rapidly diminishing returns.

Allen places energy innovation center stage in his theory of why the industrial revolution occurred in Britain.⁵² Like Wrigley, he compares Britain to the Netherlands as well as Belgium. These were the most developed economies in the world in the early modern age with much higher wages than elsewhere due to their dominant position in world commerce. In both countries the price of fuel wood was rising in the early modern period relative to the price of coal. But the price ratio of traditional fuel to coal was higher in London than in the Low Countries and even higher in the coalfield areas of northeastern England and western Britain. Coal was lower quality than wood as a heating and cooking fuel but became a “backstop technology”⁸⁷ once the relative price of wood to coal rose sufficiently, while in the Netherlands peat replaced coal. But innovations were required in order to use coal effectively in new applications from home heating and cooking to iron smelting. These induced innovations sparked the industrial revolution. Initially, the new innovations were only profitable in Britain where wages relative to energy prices were the highest in the world. Continued innovation eventually made coal-using technologies profitable in other countries too. Wilkinson⁸⁸ presented similar arguments, but also argued that no real welfare gains came from this Boserupian⁸⁹ response to scarcity. As a first step to integrating the different approaches, I propose to modify Solow's growth model¹⁶ by adding an energy input that has low substitutability with capital and labor, while allowing the elasticity of substitution between capital and labor to remain at unity. In this model, depending on the availability of energy and the nature of technological change, energy can be either a constraint on growth or an enabler of growth. Omitting time indexes for simplicity, the model consists of two equations:

$$Y = \left[(1 - \gamma) (A_L^\beta L^\beta K^{1-\beta})^\phi + \gamma (A_E E)^\phi \right]^{\frac{1}{\phi}}, \quad (1)$$

$$\Delta K = s(Y - p_E E) - \delta K. \quad (2)$$

Equation (1) embeds a Cobb–Douglas function of capital (K) and labor (L) in a constant elasticity of substitution (CES) function of value added and energy (E) that produces gross output Y ; $\phi = \frac{\sigma-1}{\sigma}$, where σ is the elasticity of substitution between energy and the value-added aggregate; p_E is the price

of energy; and γ is a parameter reflecting the relative importance of energy and value added. A_L and A_E are the augmentation indexes of labor and energy, which can be interpreted as reflecting both changes in technology that augment the effective supply of the factor in question and changes in the quality of the respective factors. Equation (2) is the equation of motion for capital that assumes like Solow¹⁶ that the proportion of gross output that is saved is fixed at s and that capital depreciates at a constant rate δ . Equation (1) explicitly ignores land. We could see it as a model of the industrial sector of the economy. By ignoring materials, we assume that they can be aggregated together with energy in the energy input.

As $\sigma \rightarrow 1$ and $\gamma \rightarrow 0$ we have the Solow model as a special case, where in the steady state, K and Y grow at the rate of labor augmentation. Additionally, depending on the scarcity of energy, the model displays either Solow-style or energy constrained behavior.⁹⁰ For a given elasticity of substitution, when energy is superabundant the steady-state level of the capital stock and output are determined (apart from a scaling factor) by the same functions of the same factors—the savings rate, the level of labor augmenting technology, and the rate of depreciation as the Solow model. But when energy is relatively scarce the steady state depends on the level of energy supply and the level of energy-augmenting technology. Therefore, in the preindustrial era when energy was scarce due to the limited supply of land, the steady-state level of output was determined by the supply of energy and the level of energy augmentation. After the industrial revolution, as energy became more and more abundant, the long-run behavior of the model economy becomes more and more like the Solow model where the growth of output is determined by the rate of labor augmentation. Additionally, if the cost share of energy falls over time more work effort can be directed to final output, boosting the growth rate of GDP per hour worked. Similarly, Ayres and Warr argue that historical declines in energy prices have been a major driver of economic growth.⁴⁸

Why assume that the elasticity of substitution between energy and capital-labor is less than unity? After all, if the energy supply is constant, then energy will act as a constraint on growth in the same way that a fixed land supply constrains growth in the Malthusian sector of Hansen and Prescott's model.

First, the cost share of energy in Sweden, a country for which we have data from 1800 till the present, fell over time (Fig. 3). This is not possible for a single sector Cobb–Douglas technology. The extent of the fall from a ratio of 90% of GDP (50% of gross output) in 1800 to close to 10% of GDP today appears to rule out structural change as the explanation either. Second, we cannot distinguish between labor augmenting innovations and energy augmenting innovations in a Cobb–Douglas technology. Using a CES production function allows us to distinguish between labor and energy augmenting technical change. Fröling develops a unified energy-growth model that allows the separate augmentation of energy in the production of energy services but it is assumed that total factor productivity (TFP) growth augments energy services, labor, and land in final production at the same rate.⁹¹ Furthermore, because final output is produced by a Cobb–Douglas function, the model cannot reproduce the decline in the energy cost share. Similarly, Tahvonen and Salo⁶⁷ develop a model of energy transitions between renewable and nonrenewable energy sources in the long run. They also assume that final output is produced by a Cobb–Douglas function and that there is technical progress in the extraction of fossil fuels (but not in the production of final output).

Third, the ecological economics literature and econometric results discussed earlier indicate that the elasticity of substitution between capital and energy is less than unity. The production function in Equation (1) has two limits to substitution.⁹ The “microeconomic” limit to substitution results from $\sigma < 1$ so that a minimum quantity of energy is required to produce any given level of output and energy is essential to production. The “macroeconomic” limit to substitution results from energy being required to produce capital and as long as $\delta > 0$, depreciation means that maintenance of the capital stock requires an ongoing energy input.

If this model is a reasonable representation of reality, then mainstream economists are not wrong to ignore the role of energy in economic growth in developed economies where energy is abundant but their models have limited applicability to other situations. However, in the long run, even if there is labor augmenting technical change, energy supply or energy efficiency must also be increasing or eventually energy will again begin to constrain

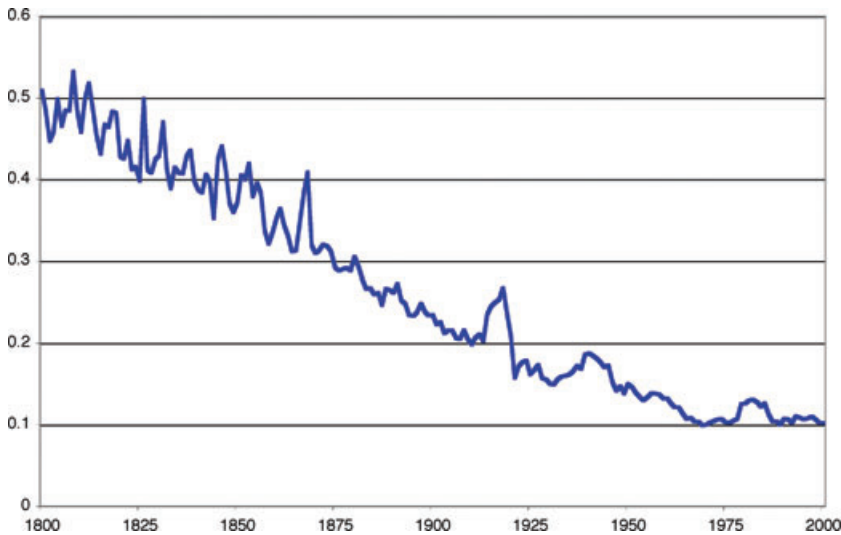


Figure 3. Cost of energy relative to the value of gross output: Sweden 1800–2000. Source: Kander.⁹⁸ (In color in *Annals online*.)

economic growth. Preliminary research results show that this model can simulate the observed features of the Swedish economy in the last two centuries reasonably well, including the fall in the cost share of energy and the decline in energy intensity over time.⁹⁰

Ayres and Warr also developed an approach that seeks to explain the role of energy in long-run economic growth while containing some elements of mainstream models.^{48,49} Capital, labor, and useful work performed by energy combine in a continuous “LINEX” production function. In the REXS growth model,⁹² capital accumulates as in the Solow growth model. But, in contrast to mainstream models, technological change only acts to raise the ratio of useful work performed per joule of exergy, as there is no labor augmentation or factor neutral technical change. Output elasticities are not constrained to equal cost shares and the LINEX production function assumes that as capital accumulates and the economy becomes “more automated” the output elasticity of labor falls.⁹³ Also labor and energy are assumed to be q -substitutes—increases in energy use reduce the marginal product of labor, which can become negative when energy is very abundant. As a result, the output elasticity of labor falls over time and the role of energy grows. This is in contrast to my model where, as energy use increases, the economy is better represented by the Solow growth model, output elasticities are assumed to be equal to cost shares, all inputs are q -complements—increases in

each input increase the marginal products of the other inputs—and technical change can augment both labor and energy.

Factors affecting the linkage between energy and growth

Introduction

So far, we have established that energy is an essential input, that in theory in the long-run energy availability could constrain economic growth, and that the current relative abundance of energy sources has alleviated these constraints. It is commonly asserted that there has been a decoupling of economic output and resources and it is often implied that there has been a permanent change in the relationship between energy and growth.^{94–96} Taking the example of the U.S. economy, energy consumption grew at a much slower pace after 1973 than before (Fig. 4). This was despite a significant increase in GDP. Energy consumption in several European countries was still below 1973 levels in 2007.⁹⁷ Evidence from the United States (Fig. 5), Sweden,⁹⁸ and other European countries⁴ shows that the energy/GDP ratio has declined for the last couple of centuries and not just over the last half century. However, if only modern energy—fossil fuels, nuclear, and hydropower—are included in the energy aggregate, energy intensity appears to first rise and then fall (Fig. 5). What were the reasons for the decline in energy intensity and more recent periods of stable or falling energy use in some countries?

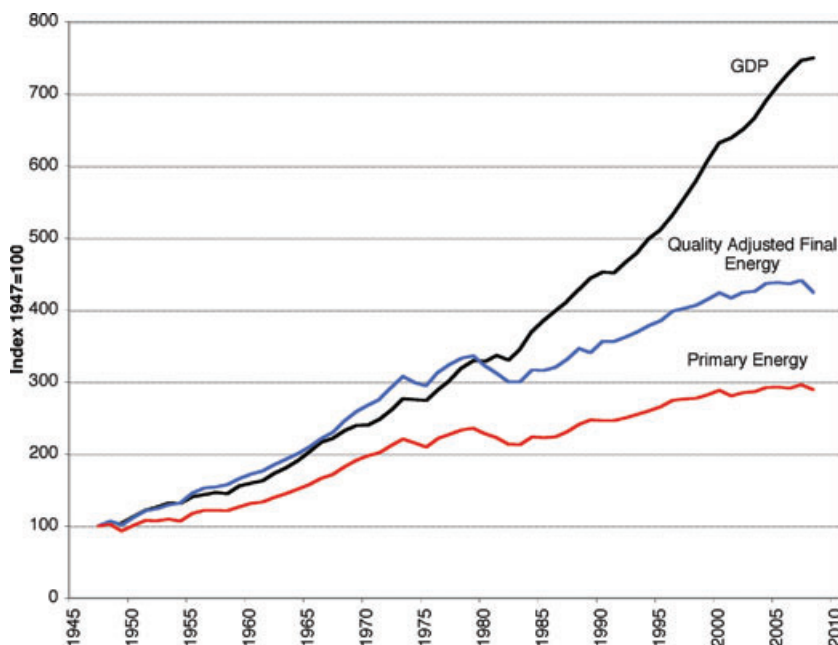


Figure 4. U.S. GDP and primary energy use and quality-adjusted final energy. Notes: GDP is in constant dollars, that is, adjusted for inflation. Primary energy use is the sum of primary energy joules. Quality-adjusted final energy use is a Divisia index of the principal final energy use categories—oil, natural gas, coal, primary electricity, wood, and other biofuels. The different fuels are weighted according to their average prices. Sources: U.S. Energy Information Administration (2010),¹⁸⁹ and U.S. Department of Commerce, Bureau of the Census (1975).¹⁹⁰ (In color in *Annals online*.)

We can use a production function approach to examine the factors that could weaken or strengthen the linkage between energy use and economic activity over time and summarize the empirical evidence on each of these mechanisms. A general production frontier, assuming for simplicity that there is separability between inputs and outputs, can be represented as

$$(Q_1, \dots, Q_m)' = f(A, X_1, \dots, X_n, E_1, \dots, E_p), \quad (3)$$

where the Q_i are various outputs, such as manufactured goods and services, the X_i are various nonenergy inputs, such as capital and labor, the E_i are different energy inputs, such as coal and oil, and A is the state of technology as defined by the TFP indicator. The relationship between energy and an aggregate of output, such as GDP, can then be affected by

- substitution between energy and other inputs;
- technological change—a change in A ;
- shifts in the composition of the energy input; and
- shifts in the composition of output.

Also, shifts in the mix of the other inputs—for example to a more capital-intensive economy from a more labor-intensive economy—can affect the relationship between energy and output but this issue has not been extensively discussed in the literature and so will not be pursued further here. It is also possible for the input variables X or E to affect TFP. This possibility is discussed in the subsection on technological change below.

Substitutability of energy and capital

There is a large empirical literature on the issue of whether capital and energy are substitutes or complements and on how substitutable they are.^{70,99–103}

We have already discussed the difficulty of substitution as measured by the Hicks or direct elasticity of substitution. Most of the empirical literature focuses, though, on the separate concept of price-substitutability versus price-complementarity. Two inputs are said to be p -substitutes (p -complements) if the quantity of one increases (decreases) when the price of the other increases. This discrimination is best made according to the sign of the cross-price elasticity (CPE).⁶⁵

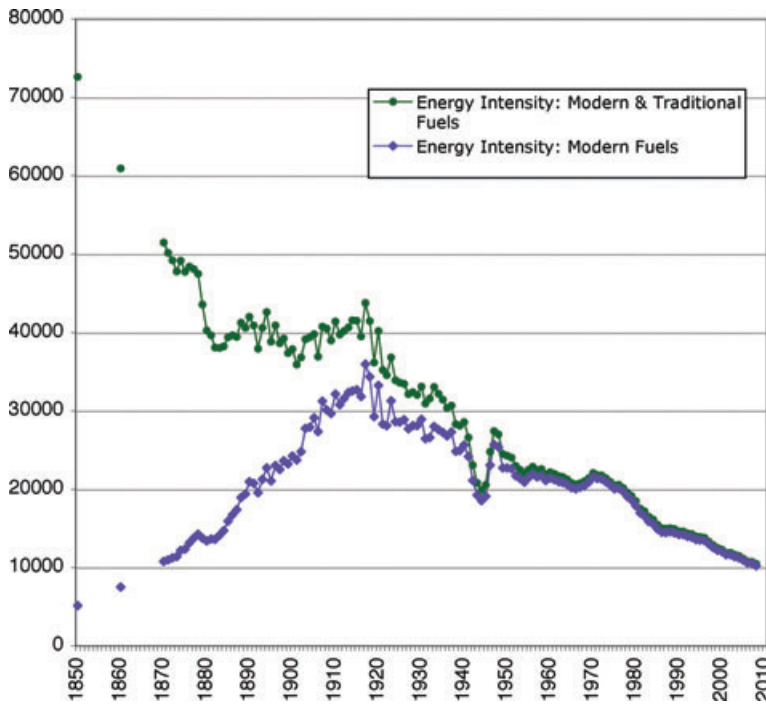


Figure 5. U.S. energy/GDP ratio, 1850–2008. Units are British Thermal Units per 1990 US\$. Sources: Maddison¹⁴⁸ (and data from <http://www.ggd.net/MADDISON/oriindex.htm>),⁸¹ Schurr and Netschert,¹⁴⁸ U.S. Energy Information Administration,¹⁸⁹ U.S. Department of Commerce, Bureau of the Census.¹⁹⁰

Econometric studies have come to varying conclusions regarding whether capital and energy are *p*-complements or *p*-substitutes.^{99,100} Based on the differences between time-series and cross-sectional results, Apostolakis concluded that capital and energy act more as substitutes in the long run and more as complements in the short run.¹⁰⁰ Frondel and Schmidt¹⁰² revisit the studies reviewed by Apostolakis and additional data from Germany and find that evidence of complementarity only occurs in cases where the cost share of energy is small. When materials are included, the cost shares of capital and energy are smaller and a finding of complementarity is more likely. Obviously the cost of materials should be included if possible and econometric results that exclude this variable are likely to be biased. Similarly, Berndt and Wood found that engineering studies and econometric studies that included capital, labor, and energy but not materials as inputs indicate substitution, while econometric cost functions that also include materials indicate complementarity.⁹⁹

The Morishima elasticity of substitution (MES) captures the concept of difficulty of substitution.¹⁰⁴

It is, however, asymmetric—the elasticity takes different values, depending on whether the price of energy or capital increases. The Hicks elasticity of substitution is a weighted average of the two MESs. Koetse *et al.* conduct a meta-analysis of the MES and CPE between capital and energy for an increase in the price of energy.⁷⁰ Their base case finds that energy and capital are complements and that the MES is 0.216. The MES estimated using panel and cross-section data is greater (0.592 and 0.848, respectively) while the CPE is positive in cross-sectional data. It is likely that these larger values reflect long-run elasticities and the lower values short-run elasticities.⁶⁸ Koetse *et al.* found that exclusion of materials had no significant effect.

It seems, in conclusion, that the microlevel Hicks elasticity of substitution between capital and energy is less than unity, especially in the short run. Capital and energy are likely *p*-complements in the short run and *p*-substitutes in the long run. In Atkeson and Kehoe's putty-clay model, energy is used in fixed proportion to capital but different types of capital have different energy requirements.¹⁰⁵ Thus

in the short run, energy and capital are poor substitutes but in the long run quite a lot of substitution is possible as the capital stock turns over. They find that this model matches changes in U.S. macroeconomic data much better than their alternative models.

Innovation and energy efficiency

Changes in the energy/GDP ratio that are not related to changes in the relative price of energy are called changes in the autonomous energy efficiency index or AEEL.¹⁰⁶ These could be due to any of the determinants of the relationship between energy and output listed at the beginning of this section and not just technological change. Even A in Equation (3) is just general TFP and, therefore, includes the effects of technological change on augmenting other inputs as well as energy. There are several ways of measuring the level of technology that control for these other factors. The distance function approach asks: "What is the minimum energy requirement to produce a given level of output holding all other inputs constant?" The level of energy efficiency in period t relative to period 0, B_t , is given by

$$B_t = \frac{E_0(\mathbf{y}_0, \mathbf{x}_0)}{E_t(\mathbf{y}_0, \mathbf{x}_0)}, \quad (4)$$

where \mathbf{y} is the vector of outputs and \mathbf{x} the vector of nonenergy inputs with subscripts indicating the periods and $E_i(\cdot)$ is a function indicating the minimum energy required in period i in order to achieve the given outputs given the level of inputs. The functions in Equation (4) can be estimated econometrically⁵ or nonparametrically.

An alternative approach is an index of energy augmenting technical change. This involves a reformulation of the production function in Equation (3)

$$Q = f(A_1 X_1, \dots, A_n X_n, A_E E), \quad (5)$$

so that each input is multiplied by its own technology factor A_i that converts crude units of the input into "effective units." A_E is the index of energy augmenting technical change, which holds the use of all other inputs and their augmentation indexes constant. In some, but not all, situations, $A_E = B$.

Bottom up, engineering-based measurements of energy efficiency represent a third approach. For example, Ayres *et al.*¹⁰⁷ and Warr *et al.*¹⁰⁸ estimate the useful work performed per joule of exergy by

various fuels and uses of energy. With some exceptions, the general trend over the 20th century in the United States, United Kingdom, Japan, and Austria has been to greater energy efficiency measured in this way. Over shorter periods energy efficiency has declined in some countries as it has in the long term for some fuels, especially food and feed.

Estimates of the trend in AEEL, energy efficiency, or the energy augmentation index are mixed.¹⁰⁹ The direction of change has not been constant and varies across different sectors of the economy. Berndt *et al.* use a model with linear time trends to estimate augmentation trends labor, electricity, fuels, machines, and structures in the U.S. manufacturing industry between 1965 and 1987.¹¹⁰ The rates of augmentation were 11.8% and -3.4% per annum for electricity and fuels, respectively. Patterns for Canada and France were, however, entirely different.

Judson *et al.* estimate separate models for energy consumption in each of a number of energy-consuming sectors for a large panel of data.¹¹¹ They estimate time effects that show rising energy consumption over time in the household and other sectors but flat to declining time effects in industry and construction. This suggests that technical innovations tend to introduce more energy-using appliances to households and energy-saving techniques to industry.

When there is endogenous technological change, changes in prices may induce technological changes. As a result, an increase in energy prices does tend to accelerate the development of energy-saving technologies, while periods of falling energy prices may result in energy-using technological change. There can also be an effect on the general rate of TFP growth.¹¹² Jorgenson found that technical change was biased and tended to be energy-using.¹¹³ If this is the case, lower energy prices tend to accelerate TFP growth and vice versa. Newell *et al.* provide some information on the degree to which energy price increases induce improvements in the energy efficiency of consumer products.¹¹⁴ They decompose the changes in cost and energy efficiency of various energy-using appliances using the concept of a transformation frontier of possible cost and efficiency combinations. For room air conditioners, large reductions in cost, holding efficiency, and cooling capacity constant occurred from 1960 to 1980 in the United States. Also the cost of high efficiency air conditioners relative to inefficient ones was

reduced. From 1980 to 1990, the former trend ended but the mix of air conditioners offered from those that were feasible to manufacture shifted sharply in favor of higher efficiency. Only about one-quarter of the gain in energy efficiency since 1973 was induced by higher energy prices. Another quarter was found to be due to raised government standards and labeling. For gas water heaters, the induced improvements were close to one-half of the total, although much less cost reducing technical change occurred. Popp similarly finds that increased energy prices have a significant though quantitatively small effect on the rate of patenting in the energy sector.¹¹⁵

New energy-using technologies initially diffuse slowly due to high costs of production that are typically lowered radically by a fairly predictable process of learning by doing.¹¹⁶ Diffusion tends to follow a logistic curve with the speed of diffusion depending on among other things how well the innovation fits into the existing infrastructure. Energy-saving innovations, such as LED light-bulbs, would be expected to diffuse rapidly once their price becomes competitive while more radical innovations that require new support infrastructures diffuse much more slowly due to “network effects.”

Recent research investigates the factors that affect the adoption of energy efficiency policies or energy efficiency technology.^{5,117–121} Differences in the adoption of energy efficiency technologies across countries and states, over time, and among individuals might be optimal due to differences in endowments, preferences, or the state of technology. But the rate of adoption may also be inefficient due to market failures and behavioral factors. Market failures include environmental externalities, information problems, liquidity constraints in capital markets, failures of innovation markets, and principal-agent problems, such as between landlords and tenants.^{119,120} Fredriksson *et al.* find that the greater the corruptibility of policy makers the less stringent is energy policy, and that the greater lobby group coordination costs are the more stringent energy policy is.¹¹⁸ Matisoff finds that the most significant variable affecting the adoption of energy efficiency programs across U.S. states is citizen ideology.¹¹⁷ A broad band of states from Florida to Idaho has not adopted any policies. The initial level of criteria air pollutants was also a significant determinant of the number of programs adopted and

the adoption of a renewable portfolio standard. Wei *et al.* compute an energy efficiency index based on the data envelopment analysis approach to examine energy efficiency in China.¹²¹ Using 1997–2006 panel data for 29 provinces, they find that energy efficiency is negatively associated with the secondary industry share in GDP, the state-owned economic share in GDP and the government expenditure share in GDP, and is positively associated with the level of general technology and noncoal share in energy consumption.

Strong correlations between the state of technology and the levels of other inputs can result in biased and inconsistent results of the trend in energy efficiency or energy augmentation indexes.⁵ Using a method intended to address the issue of biased estimation, Stern finds that energy efficiency shown in Equation (4) improved from 1971 to 2007 in most developed economies, former communist countries including China, and in India.⁵ But there was no improvement or instead a reduction in energy efficiency in many developing economies. Globally, such technological change resulted in a 40% reduction in energy use over the period than would otherwise have been the case. He finds that energy efficiency rises with increasing general TFP but is also higher in countries with more undervalued exchange rates in purchasing power parity adjusted terms. Higher fossil fuel reserves are associated with lower energy efficiency. Energy efficiency converges over time across countries with growing economies and technological change was the most important factor mitigating the global increase in energy use and carbon emissions due to economic growth.

Le Pen and Sévi applied a pair-wise cointegration test to convergence of energy intensities in 97 countries.¹²² They rejected the global convergence hypothesis. Previous work discussed by Le Pen and Sévi had mostly found convergence of energy intensity among developed economies but not in samples of both developed and developing countries.

The rebound effect

If energy-saving innovations induce an increase in energy consumption that offsets the technology-derived saving, there is said to be a rebound effect.¹²³ Rebound effects can be defined for energy-saving innovations in consumption and production. A consumer consumes energy services that are produced using energy itself and an appliance whose energy

use requirements are reduced by the innovation. Five rebound effects can be defined.

- (1) A substitution effect toward greater consumption of the now cheaper energy service and therefore of energy.¹²⁴
- (2) A direct income effect, which can be positive or negative depending on whether the energy service is a normal or inferior good.¹²⁵ Lovins argued that energy services were inferior goods in developed economies and, therefore, the negative income effect would outweigh the positive substitution effect.¹²⁵
- (3) Income effects on the consumption of other energy services by the consumer. Other energy services may be substitutes for or complements with the energy service that is now cheaper and, therefore, the effects are complicated.¹²³

Most empirical rebound studies are microeconomic studies that only include these first three effects. There are also two macroeconomic effects.

- (4) Increased real income also increases demand for all goods in the economy and, therefore, for the energy required to produce them.¹²³
- (5) There also may be economy-wide changes, such as adjustments in capital stocks, that result in a further increased long-run demand response for energy.¹²⁶

The production case is very similar, except that the income effect is replaced by an output effect. Consumers are constrained by a fixed nominal income, but producers' costs are not similarly constrained. Therefore, output effects can be large. For example, Darwin found that for wood-saving technological change in the U.S. Pacific Northwest output effects were sufficiently large as to increase the consumption of raw logs.¹²⁷

In partial equilibrium, the absolute value of the demand elasticity for energy should be an upper limit on the size of the rebound effect.¹²⁸ Brookes suggested that, due to long-run growth effects, the rebound effect could be larger than the initial saving resulting in higher, not lower, energy consumption, or "backfire" (also known as "Jevon's paradox").¹²⁹ Using a macro model with fixed energy prices, Saunders showed that this required that the elasticity of substitution between energy and other inputs is equal to or greater than unity, which as we have seen is unlikely.¹³⁰ Howarth, however, argues per-

suasively that even if the elasticity of substitution is one or greater than, when a distinction is made between energy services and energy use, the macrolevel energy rebound effect for a production innovation is less than the initial innovation-induced reduction in energy use, so improvements in energy efficiency do, in fact, reduce total energy demand.¹²⁶ However, backfire might be possible if the share of energy costs in total output was large and the share of energy costs in energy services was high. Both these conditions could have been true during the early industrial revolution⁹⁸ so that Jevons' paradox¹³¹ could have been valid then even if it no longer applies.

Extensive empirical studies have been conducted for both production and consumption. Surveying these existing studies, Greening *et al.* find that microlevel rebound effects for consumption are typically in the range of 10–30% and may typically be even smaller for industry.¹³² In subsequent studies, Bentzen¹³³ finds a 24% rebound in U.S. manufacturing, Haas and Biermayr¹³⁴ estimate a 20–30% rebound effect in Austrian space heating, and Berkhout *et al.*¹²³ find rebound effects of 15–27% for the Netherlands. Roy argues that because high-quality energy use is still small in households in India, demand is very elastic, and thus rebound effects in the household sector in India and other developing countries can be expected to be larger than in developed economies.¹³⁵ Sorrell *et al.* review the literature on the microlevel or partial equilibrium rebound effect for personal transport, household heating, and other household services, also finding that the effect appears to be less than 30%.¹³⁶ Schipper and Grubb survey energy-output changes for broad end-use categories in industrial nations and find sector-level rebounds of 5–15%.¹³⁷ They speculate that there are small macrolevel rebound effects.

Grepperud and Rasmussen use a general equilibrium model for the Norwegian economy with econometrically estimated parameters.¹³⁸ For an increase in the growth rate of the augmentation index of electricity, they find rebound effects greater than 100% in manufacturing industries where there are good substitution possibilities between electricity and other inputs, electricity dominates energy consumption, and the industries face perfectly elastic export demand which allows output to expand substantially. Allan *et al.* also use a CGE model and find a short-term rebound effect of 55% and long-run

effect of 30% for an increase in energy efficiency in production in the United Kingdom.¹³⁹ However, these results are sensitive to the assumed structure of the labor market, key production elasticities, the time period under consideration, and the mechanism through which increased government revenues are recycled back to the economy. Linares and Labandeira¹²⁰ summarize some additional empirical estimates.

Energy quality and shifts in composition of energy input

In the course of economic development, countries' fuel mix tends to evolve as they move up the "energy ladder."¹⁴⁰ Burke documents a similar progression for the power sources used in electricity generation.¹⁴¹ In the least developed economies, as in today's developed economies before the industrial revolution, the use of biomass and animate prime movers dominates. The evolution of the energy mix over the course of economic development and over history in the technologically leading countries depends on each country's endowments of fossil energy and potential for renewables, such as hydro-electricity, but some regularities apply. The share of electricity in total energy use tends to rise. Low-income countries tend to generate electricity from hydropower and oil, while high-income countries have more diverse power sources including nuclear power. Direct use of coal tends to rise and then fall over time and with income. Natural gas use has increased significantly in recent decades mostly in more developed economies. Finally electricity generated from solar and wind power are only now beginning to take off in more developed economies. Figure 1 illustrates this pattern for the United States.

Energy quality is the relative economic usefulness per heat equivalent unit of different fuels and electricity. Fuels have a number of physical attributes that will affect their relative qualities, including energy density; power density (rate of heat units produced per unit are per unit time); ease of distribution; the need for a transfer medium; controllability; amenability to storage; safety; and environmental impacts.^{142–144} Some fuels can be used for a larger number of activities and/or for more valuable activities. Some fuels, in particular electricity, require innovations that allowed their use that must be embodied in capital equipment can transform the workplace entirely and change

work processes, thus contributing to productivity gains.^{1, 145, 146}

Stern discusses alternative ways of measuring energy quality.¹⁴⁶ The most relevant approach to understanding the impact of relatively small changes in the composition of the energy input on economic output is the marginal product of the fuel. This is determined in part by the complex set of attributes described earlier that are unique to each fuel. It also varies according to what activities it is used in, how much and what form of capital, labor, and materials it is used in conjunction with, and how much energy is used in each application. More abundant fuels will be applied more widely and on the margin in less productive applications.¹⁴⁷ Therefore, energy qualities measured in this way are not fixed over time. However, it is generally believed that electricity is the highest quality type of energy followed by natural gas, oil, coal, and wood and biofuels in descending order of quality. This is supported by the typical prices of these fuels per unit of energy, which should be proportional to their marginal products. Under the assumption of optimizing behavior, marginal products should be approximated by prices, which are usually readily available. Other indicators of energy quality must be estimated. As energy vectors differ in quality, shifts in the energy mix should affect energy intensity.

Surprisingly, relatively few studies evaluate the role of the change in energy mix on energy intensity. Noting that the composition of energy use has changed significantly over time, Schurr and Netschert were among the first to argue that the general shift to higher quality fuels reduces the amount of energy required to produce a dollar's worth of GDP.¹⁴⁸ Berndt also noted the key role played by the shifting composition of energy use toward higher quality energy inputs.¹¹²

Cleveland *et al.*,⁴⁵ Kaufmann,^{78, 106} and the Office of Technology Assessment¹⁴⁹ presented analyses that explain much of the decline in the U.S. energy/GDP ratio in terms of structural shifts in the economy and shifts from lower to higher quality fuels. Kaufmann estimates a vector autoregressive (VAR) model of the energy/GDP ratio, household energy expenditures, energy mix variables, and energy price variables for the United States.¹⁰⁶ He finds that shifting away from coal use and in particular shifting toward the use of oil reduces energy intensity. This shift away from coal more than explains the decline energy intensity

over the entire 1929–1999 period. If decoupling is mainly due to the shift to higher quality fuels then there appear to be limits to that substitution. In particular, exhaustion of low-cost oil supplies could mean that economies have to revert to lower quality fuels, such as coal.⁷⁸

Figure 4 also includes a quality-adjusted index of final energy use that accounts for differences in the productivity of different fuels by weighting them by their prices.¹⁵⁰ There is less evidence of decoupling of energy use and GDP in these data than indicated by the primary energy series. The studies cited earlier (and those by Stern discussed later^{150,151}) used earlier GDP data that showed significantly less economic growth in the United States than more recently updated data do. Using these earlier data there was little decoupling of GDP from quality-adjusted energy use even after 1973. This change in the GDP data indicates that structural change and technological change must also contribute to lowering the energy/GDP ratio in the last three decades assuming that prices reflect the relative marginal products of the fuels.

Other studies find, however, a much larger role for technological change than for changes in the composition of energy in the reductions in energy intensity seen around the world. For example, Ma and Stern find that interfuel substitution has negligible effects on the decline in energy intensity in China between 1994 and 2003.¹⁵² Technological change reduced energy intensity by more than the actual reduction in energy intensity due to the intensity increasing effects of structural change. Stern finds that between 1971 and 2007, changes in fuel mix within individual countries increased world energy use by 4%, while global energy intensity declined by 40%. Shifts in the distribution of economic activity toward countries with lower quality energy mixes, such as China and India, contributed further to increasing energy intensity globally.⁵

Shifts in the composition of output

Output mix also typically changes over the course of economic development. In the earlier phases of development there is a shift away from agriculture toward heavy industry, while in the later stages of development there is a shift from the more resource-intensive extractive and heavy industrial sectors toward services and lighter manufacturing. Different

industries have different energy intensities. It is often argued that this will result in an increase in energy used per unit of output in the early stages of economic development and a reduction in energy used per unit output in the later stages of economic development.¹⁵³

However, there is reason to believe that the energy-saving effects of structural changes are overstated. When the indirect energy use, embodied in manufactured products and services, is taken into account, the U.S. service and household sectors are more energy intensive than they first appear.⁴⁴ Service industries still need large energy and resource inputs. The service being sold may be intangible but the office towers, shopping malls, warehouses, rental apartment complexes, etc., where the activity is conducted are very tangible and energy is used in their construction, operation, and maintenance. Furthermore, consumers use large amounts of energy and resources in commuting to work, shop, etc.

The effect of the Internet on the energy intensity of commerce has received increasing attention.¹⁵⁴ Obviously, individual technologies such as news websites versus newspapers can greatly reduce emissions¹⁵⁵ but the effects on other activities could outweigh the gains. Romm *et al.* argue that the environmental costs of the greater dispersal of population engendered by telecommuting would not outweigh the reduction in commuting costs suggesting a strong energy-conserving Internet effect.¹⁵⁶ But Matthews *et al.*¹⁵⁷ and Williams and Tagami¹⁵⁸ provide evidence that online book retailing uses more energy than traditional retail, while Herring and Roy show that electronic distance learning results in more energy use than traditional distance learning with printed material.¹⁵⁹

There may also be a tendency for consumers to use more energy directly over time as their consumption of the services provided by appliances, housing, transport, etc., increases. Judson *et al.* find that the consumer sector sees rising energy intensity over time, *ceteris paribus*, while the manufacturing sector sees decreasing energy intensity.¹¹¹

Furthermore, on a global scale there may be limits to the extent to which developing countries can replicate the structural shift that has occurred in the developed economies to the degree that this is due to outsourcing manufacturing overseas rather than simply from an expansion in service

activities. However, the evidence shows that trade does not result in reductions in energy use and pollution in developed countries through the offshoring of pollution-intensive industries.^{160–162} Additionally, if the service sector does require substantial material support, it is not clear whether the developed world can continue to shift in the direction of a growing service share of GDP indefinitely. In fact, as manufacturing prices have fallen relative to the prices of services, even the relative decline of manufacturing in developed countries is exaggerated when the relative sizes of the sectors are computed in current prices.¹⁶³

Kander⁹⁸ and Stern⁵ find a relatively small role for structural change in reducing energy intensity in Sweden (1800–2000) and the world (1971–2007), respectively. But, using a much finer disaggregation of industries, Sue Wing finds that structural change explained most of the decline in energy intensity in the United States (1958–2000), especially before 1980.¹⁶⁴

Empirical testing of the causal relationship between energy and growth

When two or more totally unrelated variables are trending over time, they will appear to be correlated simply because of the shared directionality. Therefore, traditional linear regression or correlation methods cannot be used to establish causal relations among a group of variables. Two methods for testing for causality among time-series variables are Granger causality tests¹⁶⁵ and cointegration analysis.¹⁶⁶ Hendry and Juselius¹⁶⁷ discuss the application of these methods to energy economics where they have been applied extensively to test for causality and cointegration between energy, GDP, and other variables from the late 1970s on.^{168,169}

Early studies relied on Granger causality tests on unrestricted VARs in levels of the variables, while more recent studies use cointegration methods. Another key characteristic that distinguishes between studies is whether a bivariate model of energy and output or a multivariate framework is used. A third way to differentiate among models is whether energy is measured in standard heat units or whether some type of indexing method is used to account for differences in quality among fuels.

It is now understood that in the absence of cointegration between the variables a Granger causality test on a VAR in levels is invalid. Ohanian¹⁷⁰

and Toda and Phillips¹⁷¹ showed that the distribution of the test statistic for Granger causality in a VAR with nonstationary variables is not the standard chi-square distribution. This means that the significance levels reported in the early studies of the Granger-causality relationship between energy and GDP may be incorrect, as both variables are generally integrated series. If there is no cointegration between the variables then the causality test should be carried out on a VAR in differenced data, while if there is cointegration, standard chi-square distributions apply when the cointegrating restrictions are imposed.¹⁷² Toda and Yamamoto developed a modification of the standard Granger causality test on the variables in levels that are robust to the presence of unit roots.¹⁷² But it is still, of course, subject to possible omitted variables bias. Cointegration tests can be used to test for omitted nonstationary variables. A lack of cointegration implies that variables essential to cointegration are omitted from the model. Therefore, testing for cointegration is still a necessary prerequisite to causality testing on data with potential unit roots.

The results of the early studies that tested for Granger causality using a bivariate model were generally inconclusive.¹⁵¹ Where nominally significant results were obtained, they mostly indicated that causality runs from output to energy. However, in many cases results differed depending on the samples used, the countries investigated, etc. Most economists believe that capital, labor, and technical change play a significant role in determining output, yet early studies used only energy as an independent variable. When relevant variables are omitted from the model, there will be no cointegration and a spurious regression will result. Results are frequently sample dependent in the face of omitted variables and noncointegration.¹⁷³ This may explain the very divergent nature of the early causality literature. Stern tested for Granger causality in a multivariate setting using a VAR model of GDP, capital and labor inputs, and a Divisia index of quality-adjusted energy use in place of gross energy use.¹⁵¹ When both the multivariate approach and a quality-adjusted energy index were employed, energy was found to Granger cause GDP.

These results were supported by studies that found that changes in oil prices Granger-cause changes in gross national product or industrial production and unemployment in VAR models whereas

oil prices are exogenous to the system.^{174–175} More recently, Blanchard and Gali used VAR models of GDP, oil prices, wages, and two other price indexes, to argue that the effect of oil price shocks has reduced over time.¹⁷⁶ Hamilton deconstructs their arguments to show that past recessions would have been mild or have merely been slowdowns if oil prices had not risen.³ Furthermore, he argues that the large increase in the price of oil that climaxed in 2008 was a major factor in causing the 2008–2009 recession. However, as the short-run elasticity of demand for oil and other forms of energy is low, the main short-run effects of oil prices are expected to be through reducing spending by consumers and firms on other goods, services, and inputs rather than through reducing the input of energy to production.^{3,177} Therefore, models using oil prices in place of energy quantities may not provide much evidence regarding the effects of energy use itself on economic growth.

Yu and Jin conducted the first cointegration study of the energy–GDP relationship.¹⁷⁸ Again, the results of this and subsequent studies differ according to the regions, time frames, and measures of inputs and outputs used. If a multivariate approach helps in uncovering the Granger causality relations between energy and GDP, a multivariate approach should be used to investigate the cointegration relations among the variables. When multivariate cointegration methods are used, a picture emerges of energy playing a central role in determining output in a diverse set of developed and developing nations. Stern estimated a dynamic cointegration model for GDP, quality-weighted energy, labor, and capital using the Johansen methodology.¹⁵⁰ The analysis showed that there is a cointegrating relation between the four variables and that energy Granger causes GDP either unidirectionally or possibly through a mutually causative relationship depending on which version of the model is used. Warr and Ayres replicate this model for the United States using their measures of exergy and useful work in place of Stern's Divisia index of energy use.¹⁷⁹ They find both short- and long-run causality from either exergy or useful work to GDP but not vice versa. Oh and Lee¹⁸⁰ and Ghali and El-Sakka¹⁸¹ apply Stern's methodology to Korea and Canada, respectively, and arrive at exactly the same conclusions, extending the validity of Stern's results beyond the United States. Lee and Chang¹⁸² and Lee *et al.*¹⁸³ use panel data

cointegration methods to examine the relationship between energy, GDP, and capital in 16 Asian and 22 Organisation for Economic Co-operation and Development (OECD) countries over a three- and four-decade period, respectively. Lee and Chang¹⁸² find a long-run causal relationship from energy to GDP in the group of Asian countries while Lee *et al.*¹⁸³ find a bidirectional relationship in the OECD sample. Taken together, this body of work suggests that the inconclusive results of earlier work are probably due to the omission of nonenergy inputs. However, using a panel vector error correction model of GDP, energy use, and energy prices for 26 OECD countries (1978–2005), Costantini and Martini find that in the short run energy prices cause GDP and energy use, and energy use and GDP are mutually causative.¹⁸⁴ However, in the long run they find that GDP growth drives energy use and energy prices. Other researchers who model a cointegrating relation between GDP, energy, and energy prices for individual countries produce mixed results. For example, Glasure¹⁸⁵ finds very similar results to Costantini and Martini for Korea, while Masih and Masih¹⁸⁶ and Hondroyannis *et al.*¹⁸⁷ find mutual causation in the long run for Korea and Taiwan and Greece, respectively. Using an idea from meta-analysis,¹⁸⁸ we should probably put most weight on the largest sample study—that is, Costantini and Martini¹⁸⁴—concluding that these models identify a demand function relationship where in the long run GDP growth drives energy use.

Conclusions

In conclusion, the theoretical and empirical evidence indicates that energy use and output are tightly coupled with energy availability playing a key role in enabling growth. However, the greater availability of energy, technical progress, and the employment of higher quality fuels have allowed less energy to be used per unit output and have reduced the constraint that energy resources place on the output of the economy and economic growth. Even so, energy remains important.

Energy is important for growth because production is a function of capital, labor, and energy, not just the former two or just the latter as mainstream growth models or some biophysical production models taken literally would indicate. Both theory and time-series results support these claims. Furthermore, the elasticity of substitution between

energy and capital is likely to be low, energy is also needed to produce the other inputs to production, and energy is available in finite quantities on the Earth's surface and is non recyclable.

However, the estimated output elasticity of energy and natural resources in general should be small in recent decades reflecting the market price determined output share. The current low price of energy reflects a low marginal productivity because of this heavy use. Resources have become increasingly abundant since the Industrial Revolution—evidence suggests that the energy cost share has declined continuously since then alongside the energy intensity of GDP.

Various factors have contributed to declining energy intensities, but research is less clear on the relative importance of these variables. Different energy qualities have differing productivities. In particular, modeling the effect of electricity on output is important. Part of the reduction in energy intensities in developed economies may be due to the shift to higher quality fuels. Some research indicates that most of the historical reductions in energy intensity in developed economies and China have been due to technical change, but other research finds a much larger role for structural change. Technological change tends to be offset to some degree by the rebound effect. Structural change toward more service-intensive economies tends to have less impact than is commonly thought because service industries in fact need energy-intensive infrastructures. In fact, energy-saving technical progress in manufacturing industry that reduces the apparent share of manufacturing in the economy may be more important.¹⁶³

As this survey shows, there is clearly much scope for further research to clarify the prospects for decoupling energy use and economic growth and for understanding the role of energy in growth.

Acknowledgments

I thank Jeroen van den Bergh, Cutler Cleveland, Astrid Kander, Cameron Murray, Catherine Norman, Jack Pezzey, Mark Diesendorf, Robert Ayres, and two anonymous referees for useful comments and discussions on various versions of this paper.

Conflicts of interest

The author declares no conflicts of interest.

References

1. Toman, M.A. & B. Jemelkova. 2003. Energy and economic development: an assessment of the state of knowledge. *Energy J.* **24**: 93–112.
2. Aghion, P. & P. Howitt. 2009. *The Economics of Growth*. MIT Press. Cambridge, MA.
3. Hamilton, J.D. 2009. Causes and consequences of the oil shock of 2007–08. *Brookings Papers Econ. Activity* **Spring**: 215–261.
4. Gales, B., A. Kander, P. Malanima & M. Rubio. 2007. North versus south: energy transition and energy intensity in Europe over 200 years. *Eur. Rev. Econ. History* **11**: 219–253.
5. Stern, D.I. 2010. Modeling international trends in energy efficiency and carbon emissions. Environmental Economics Research Hub Research Report 54. Crawford School of Economics and Government, Australian National University, Canberra.
6. Ayres, R.U. & A.V. Kneese. 1969. Production, consumption and externalities. *Am. Econ. Rev.* **59**: 282–297.
7. Boulding K. 1966. The economics of the coming spaceship Earth. In *Environmental Quality in a Growing Economy*. H. Jarett, Ed.: 3–14. Johns Hopkins University Press. Baltimore, MD.
8. Baumgärtner, S. 2004. Thermodynamic models. In *Modelling in Ecological Economics*. P. Safonov & J. Proops, Eds.: 102–129. Edward Elgar. Cheltenham.
9. Stern, D.I. 1997. Limits to substitution and irreversibility in production and consumption: a neoclassical interpretation of ecological economics. *Ecol. Econ.* **21**: 197–215.
10. Hall, C.A.S., C.J. Cleveland & R.K. Kaufmann. 1986. *Energy and Resource Quality: The Ecology of the Economic Process*. Wiley Interscience. New York.
11. Wrigley, E.A. 1988. *Continuity, Chance, and Change: The Character of the Industrial Revolution in England*. Cambridge University Press. Cambridge, UK.
12. Spreng, D. 1993. Possibilities for substitution between energy, time and information. *Energy Policy* **21**: 13–23.
13. Chen, X. 1994. Substitution of information for energy: conceptual background, realities and limits. *Energy Policy* **22**: 15–28.
14. Ruth, M. 1995. Information, order and knowledge in economic and ecological systems: implications for material and energy use. *Ecol. Econ.* **13**: 99–114.
15. Schultz, T.W. 1951. A framework for land economics: the long view. *J. Farm Econ.* **33**: 204–215.
16. Solow, R.M. 1956. A contribution to the theory of economic growth. *Q. J. Econ.* **70**: 65–94.
17. Arrow, K.J. 1962. The economic implications of learning-by-doing. *Rev. Econ. Stud.* **29**: 155–173.
18. Hicks, J.R. 1932. *The Theory of Wages*. Macmillan. London.
19. Perman, R. & D.I. Stern. 2001. Sustainable development, growth theory, environmental Kuznets curves and discounting. In *Encyclopedia of Life Support Systems*. UNESCO. Paris.
20. Aghion, P. & P. Howitt. 1998. *Endogenous Growth Theory*. MIT Press. Cambridge, MA.
21. Solow, R.M. 1974. Intergenerational equity and exhaustible resources. *Rev. Econ. Stud.*, Symposium on the Economics of Exhaustible Resources: 29–46.

22. Stiglitz, J.E. 1974. Growth with exhaustible natural resources: the competitive economy. *Rev. Econ. Stud.*, Symposium on the Economics of Exhaustible Resources: 139–152.
23. Dasgupta, P.S. & G.M. Heal. 1979. *Economic Theory and Exhaustible Resources*. Cambridge University Press. Cambridge, UK.
24. Hartwick, J.M. 1977. Intergenerational equity and the investing of rents from exhaustible resources. *Am. Econ. Rev.* **66**: 972–974.
25. Dixit, A., P. Hammond, and M. Hoel. 1980. On Hartwick's rule for regular maximin paths of capital accumulation and resource depletion. *Rev. Econ. Stud.* **47**: 551–556.
26. Hartwick, J.M. 1995. Constant consumption paths in open economies with exhaustible resources. *Rev. Int. Econ.* **3**: 275–283.
27. Asheim, G.B. 1994. Net national product as an indicator of sustainability. *Scand. J. Econ.* **96**: 257–265.
28. Asheim, G.B., W. Buchholz, & C. Withagen. 2003. The Hartwick rule: myths and facts. *Environ. Resour. Econ.* **25**: 129–150.
29. Stern, D.I. 1997. The capital theory approach to sustainability: a critical appraisal. *J. Econ. Issues* **31**: 145–173.
30. Solow, R.M. 1978. Resources and economic growth. *Am. Econ.* **22**: 5–11.
31. Solow, R.M. 1993. An almost practical step toward sustainability. *Resour. Pol.* **19**: 162–172.
32. Solow, R.M. 1997. Reply: Georgescu-Roegen versus Solow/Stiglitz. *Ecol. Econ.* **22**: 267–268.
33. Stiglitz, J.E. 1974. Growth with exhaustible natural resources: efficient and optimal growth paths. *Rev. Econ. Stud.*, Symposium on the Economics of Exhaustible Resources: 123–138.
34. Smulders, S. 2005. Endogenous technical change, natural resources and growth. In *Scarcity and Growth in the New Millennium*. R. Ayres, D. Simpson & M. Toman, Eds.: Chapter 8. Resources for the Future. Washington, DC.
35. Barbier, E.B. 1999. Endogenous growth and natural resource scarcity. *Environ. Resour. Econ.* **14**: 51–74.
36. Scholz, C. & G. Ziemes. 1999. Exhaustible resources, monopolistic competition, and endogenous growth. *Environ. Resour. Econ.* **13**: 169–185.
37. Groth, C. & P. Schou. 2002. Can non-renewable resources alleviate the knife-edge character of endogenous growth? *Oxf. Econ. Pap.* **54**: 386–411.
38. Grimaud, A. & L. Rougé. 2003. Non-renewable resources and growth with vertical innovations: optimum, equilibrium and economic policy. *J. Environ. Econ. Manage.* **45**: 433–453.
39. Di Maria, C. & S. Valente. 2008. Hicks meets Hotelling: the direction of technical change in capital–resource economies. *Environ. Dev. Econ.* **13**: 691–717.
40. Smulders, S. 1999. Endogenous growth theory and the environment. In *Handbook of Environmental and Resource Economics*. J.C.J.M. van den Bergh, Ed.: 89–108. Edward Elgar. Cheltenham.
41. Smulders, S. & M. de Nooij. 2003. The impact of energy conservation on technology and economic growth. *Resour. Energy Econ.* **25**: 59–79.
42. Bretschger, L. 2005. Economics of technological change and the natural environment: how effective are innovations as a remedy for resource scarcity? *Ecol. Econ.* **54**: 148–163.
43. Georgescu-Roegen, N. 1971. *The Entropy Law and the Economic Process*. Harvard University Press. Cambridge, MA.
44. Costanza, R. 1980. Embodied energy and economic valuation. *Science* **210**: 1219–1224.
45. Cleveland, C.J., R. Costanza, C.A.S. Hall & R.K. Kaufmann. 1984. Energy and the U.S. economy: a biophysical perspective. *Science* **225**: 890–897.
46. Hall, C.A.S., D. Lindenberger, R. Kümmel, *et al.* 2001. The need to reintegrate the natural sciences and economics. *BioScience* **51**: 663–673.
47. Hall, C.A.S., P. Tharakan, J. Hallock, *et al.* 2003. Hydrocarbons and the evolution of human culture. *Nature* **426**: 318–322.
48. Ayres, R.U. & B. Warr. 2005. Accounting for growth: the role of physical work. *Struct. Change Econ. Dyn.* **16**: 181–209.
49. Ayres, R.U. & B. Warr. 2009. *The Economic Growth Engine: How Energy and Work Drive Material Prosperity*. Edward Elgar. Cheltenham.
50. Murphy, D.J. & C.A.S. Hall. 2010. Year in review: EROI or energy return on (energy) invested. *Ann. N.Y. Acad. Sci.* **1185**: 102–118.
51. Smil, V. 1994. *Energy in World History*. Westview Press. Boulder, CO.
52. Allen, R.C. 2009. *The British Industrial Revolution in Global Perspective*. Cambridge University Press. Cambridge, UK.
53. Gever, J., R.K. Kaufmann, D. Skole & C. Vörösmarty. 1986. *Beyond Oil: The Threat to Food and Fuel in the Coming Decades*. Ballinger. Cambridge, MA.
54. Hannon, B. 1973. An energy standard of value. *Ann. Am. Acad.* **410**: 139–153.
55. Common, M.S. 1995. *Sustainability and Policy: Limits to Economics*. Cambridge University Press. Melbourne.
56. Hannon, B. 1973. The structure of the ecosystem. *J. Theor. Biol.* **41**: 535–546.
57. Kaufmann, R.K. 1987. Biophysical and Marxist economics: learning from each other. *Ecol. Model.* **38**: 91–105.
58. Brown, M.T. & R.A. Herendeen. 1996. Embodied energy analysis and energy analysis: a comparative view. *Ecol. Econ.* **19**: 219–236.
59. Ayres, R.U., L.W. Ayres & K. Martinás. 1998. Exergy, waste accounting, and life-cycle analysis. *Energy* **23**: 355–363.
60. Ukidwe, N.U. & B.R. Bakshi. 2007. Industrial and ecological cumulative exergy consumption of the United States via the 1997 input–output benchmark model. *Energy* **32**: 1560–1592.
61. Sraffa, P. 1960. *Production of Commodities by Means of Commodities: Prelude to a Critique of Economic Theory*. Cambridge University Press. Cambridge, UK.
62. Perrings, C.A. 1987. *Economy and Environment: A Theoretical Essay on the Interdependence of Economic and Environmental Systems*. Cambridge University Press. Cambridge, UK.
63. O'Connor, M.P. 1993. Entropic irreversibility and uncontrolled technological change in the economy and environment. *J. Evol. Econ.* **34**: 285–315.

64. Kubiszewski, I., C.J. Cleveland & P.K. Endres. 2010. Meta-analysis of net energy return for wind power systems. *Renewable Energy* **35**: 218–225.
65. Stern, D.I. 2007. The elasticity of substitution, the capital-energy controversy, and sustainability. In *Frontiers in Ecological Economic Theory And Application*. J.D. Erickson & J.M. Gowdy, Eds.: 331–352. Edward Elgar. Cheltenham.
66. Smil, V. 1991. *General Energetics Energy in the Biosphere and Civilization*. John Wiley. New York.
67. Tahvonen, O. & S. Salo. 2001. Economic growth and transitions between renewable and nonrenewable energy resources. *Eur. Econ. Rev.* **45**: 1379–1398.
68. Stern, D.I. 2010. Interfuel substitution: a meta-analysis. *J. Econ. Surv.* In press.
69. Costanza, R. & H.E. Daly. 1992. Natural capital and sustainable development. *Conservation Biol.* **6**: 37–46.
70. Koetse, M.J., H.L.F. de Groot & R.J.G.M. Florax. 2008. Capital-energy substitution and shifts in factor demand: a meta-analysis. *Energy Econ.* **30**: 2236–2251.
71. Ruth, M. 1993. *Integrating Economics, Ecology, and Thermodynamics*. Kluwer Academic. Dordrecht.
72. Islam, S. 1985. Effects of an essential input on isoquants and substitution elasticities. *Energy Econ.* **7**: 194–196.
73. Georgescu-Roegen, N. 1976. *Energy and Economic Myths*. Pergamon. New York.
74. Daly, H.E. 1991. Elements of an environmental macroeconomics. In *Ecological Economics*. R. Costanza, Ed.: 32–46. Oxford University Press. New York.
75. van den Bergh, J.C.J.M. 1999. Materials, capital, direct/indirect substitution, and mass balance production functions. *Land Econ.* **75**: 547–561.
76. Georgescu-Roegen, N. 1979. Energy and matter in mankind's technological circuit. *J. Bus. Adm.* **10**: 107–127.
77. Ayres, R. & I. Nair. 1984. Thermodynamics and economics. *Phys. Today* **35**: 62–71.
78. Kaufmann, R.K. 1992. A biophysical analysis of the energy/real GDP ratio: implications for substitution and technical change. *Ecol. Econ.* **6**: 35–56.
79. Daly, H.E. 1997. Georgescu-Roegen versus Solow/Stiglitz. *Ecol. Econ.* **22**: 261–266.
80. Kaufmann, R.K. & I.G. Azary-Lee. 1991. A biophysical analysis of substitution: does substitution save energy in the U.S. forest products industry? In *Ecological Economics: Implications for Forest Management and Practice*. D.P. Bradley & P.O. Nilsson, Eds.: 111–123. The Swedish University of Agricultural Sciences. Garpenberg.
81. Maddison, A. 2001. *The World Economy: A Millennial Perspective*. OECD. Paris.
82. Durlauf, S.N., P.A. Johnson & J.R.W. Temple. 2005. Growth econometrics. In *Handbook of Economic Growth*. P. Aghion & S.N. Durlauf, Eds.: Vol. 1A, 555–677. North Holland. Amsterdam.
83. Galor, O. & D.N. Weil. 2000. Population, technology and growth: from Malthusian regime to the demographic transition. *Am. Econ. Rev.* **90**: 806–828.
84. Lucas, R.E. 2002. The industrial revolution: past and future. In *Lectures on Economic Growth*. R.E. Lucas, Ed.: 109–188. Harvard University Press. Cambridge, MA.
85. Galor, O. 2005. From stagnation to growth: unified growth theory. In *Handbook of Economic Growth*. P. Aghion & S.N. Durlauf, Eds.: Vol. 1A, 171–294. North Holland. Amsterdam.
86. Hansen, G.D. & E.C. Prescott. 2002. Malthus to Solow. *Am. Econ. Rev.* **92**: 1205–1217.
87. Nordhaus, W.D. 1973. The allocation of energy resources. *Brookings Pap. Econ. Activity* **3**: 529–570.
88. Wilkinson, R.G. 1973. *Poverty and Progress: An Ecological Model of Economic Development*. Methuen. London.
89. Boserup, E. 1965. *The Conditions of Agricultural Growth*. Allen and Unwin. London.
90. Stern, D.I. & A. Kander. 2010. The Role of Energy in Long-Run Economic Growth. Manuscript, Arndt-Corden Department of Economics, Crawford School of Economics and Government, Australian National University. Canberra.
91. Fröling, M. 2010. Energy use, population and growth, 1800–1970. *J. Popul. Econ.* In press.
92. Warr, B. & R. Ayres. 2006. REXS: a forecasting model for assessing the impact of natural resource consumption and technological change on economic growth. *Struct. Change Econ. Dyn.* **17**: 329–378.
93. Kümmel, R., R.U. Ayres & D. Lindenberger. 2010. Thermodynamic laws, economic methods and productive power of energy. *J. Non-Equilib. Thermodyn.* **35**: 145–180.
94. Bohi, D. 1989. *Energy Price Shocks and Macroeconomic Performance*. Resources for the Future. Washington, DC.
95. de Bruyn, S.M. & J.B. Opschoor. 1997. Developments in the throughput-income relationship: theoretical and empirical observations. *Ecol. Econ.* **20**: 255–268.
96. IBRD. 1992. *World Development Report 1992: Development and the Environment*. Oxford University Press. New York.
97. International Energy Agency. 2010. *Energy Balances of OECD Countries—2010 Edition*. IEA. Paris.
98. Kander, A. 2002. Economic Growth, Energy Consumption and CO₂ Emissions in Sweden 1800–2000, Lund Studies in Economic History No. 19. Lund.
99. Berndt, E.R. & D.O. Wood. 1979. Engineering and econometric interpretations of energy-capital complementarity. *Am. Econ. Rev.* **69**: 342–354.
100. Apostolakis, B.E. 1990. Energy-capital substitutability/complementarity: the dichotomy. *Energy Econ.* **12**: 48–58.
101. Thompson, P. & T.G. Taylor. 1995. The capital-energy substitutability debate: a new look. *Rev. Econ. Stat.* **77**: 565–569.
102. Frondel, M. & C.M. Schmidt. 2002. The capital-energy controversy: an artifact of cost shares? *Energy J.* **23**: 53–79.
103. Thompson, H. 1997. Substitution elasticities with many inputs. *Appl. Math. Lett.* **10**: 123–127.
104. Blackorby, C. & R.R. Russell. 1989. Will the real elasticity of substitution please stand up? A comparison of the Allen/Uzawa and Morishima elasticities. *Am. Econ. Rev.* **79**: 882–888.
105. Atkeson, A. & P.J. Kehoe. 1999. Models of energy use: putty-putty versus putty-clay. *Am. Econ. Rev.* **89**: 1028–1043.
106. Kaufmann, R.K. 2004. The mechanisms for autonomous energy efficiency increases: a cointegration analysis of the US energy/GDP ratio. *Energy J.* **25**: 63–86.

107. Ayres, R.U., L.W. Ayres & B. Warr. 2003. Exergy, power and work in the US economy, 1900–1998. *Energy* **28**: 219–273.
108. Warr, B., R.U. Ayres, N. Eisenmenger, *et al.* 2010. Energy use and economic development: a comparative analysis of useful work supply in Austria, Japan, the United Kingdom and the US during 100 years of economic growth. *Ecol. Econ.* **69**: 1904–1917.
109. Jorgenson, D.W. & P.J. Wilcoxon. 1993. Reducing US carbon emissions: an econometric general equilibrium assessment. *Resour. Energy Econ.* **15**: 7–25.
110. Berndt, E.R., C. Kolstad & J.-K. Lee. 1993. Measuring the energy efficiency and productivity impacts of embodied technical change. *Energy J.* **14**: 33–55.
111. Judson, R.A., R. Schmalensee & T.M. Stoker. 1999. Economic development and the structure of demand for commercial energy. *Energy J.* **20**: 29–57.
112. Berndt, E.R. 1990. Energy use, technical progress and productivity growth: a survey of economic issues. *J. Productivity Anal.* **2**: 67–83.
113. Jorgenson, D.W. 1984. The role of energy in productivity growth. *Energy J.* **5**: 11–26.
114. Newell, R.G., A.B. Jaffe & R.N. Stavins. 1999. The induced innovation hypothesis and energy-saving technological change. *Q. J. Econ.* **114**: 941–975.
115. Popp, D. 2002. Induced innovation and energy prices. *Am. Econ. Rev.* **92**: 160–180.
116. Grubler, A., N. Nakicenovic & D.G. Victor. 1999. Dynamics of energy technologies and global change. *Energy Policy* **27**: 247–280.
117. Matisoff, D.C. 2008. The adoption of state climate change policies and renewable portfolio standards: regional diffusion or internal determinants? *Rev. Policy Res.* **25**: 527–546.
118. Fredriksson, P.G., H.R.J. Vollebergh & E. Dijkgraaf. 2004. Corruption and energy efficiency in OECD countries: theory and evidence. *J. Environ. Econ. Manage.* **47**: 207–231.
119. Gillingham, K., R.G. Newell & K. Palmer. 2009. Energy efficiency economics and policy. *Annu. Rev. Resour. Econ.* **1**: 597–620.
120. Linares, P. & X. Labandeira. 2010. Energy efficiency: economics and policy. *J. Econ. Surv.* **24**: 583–592.
121. Wei, C., J. Ni & M. Shen. 2009. Empirical analysis of provincial energy efficiency in China. *China World Economy* **17**: 88–103.
122. Le Pen, Y. & B. Sévi. 2010. On the non-convergence of energy intensities: evidence from a pair-wise econometric approach. *Ecol. Econ.* **69**: 641–650.
123. Berkhout, P.H.G., J.C. Muskens & J. W. Velthuisen. 2000. Defining the rebound effect. *Energy Policy* **28**: 425–432.
124. Khazzoom, D.J. 1980. Economic implications of mandated efficiency standards for household appliances. *Energy J.* **1**: 21–39.
125. Lovins, A.B. 1988. Energy saving from more efficient appliances: another view. *Energy J.* **10**: 157–166.
126. Howarth, R.B. 1997. Energy efficiency and economic growth. *Contemp. Econ. Policy* **25**: 1–9.
127. Darwin, R.F. 1992. Natural resources and the Marshallian effects of input-reducing technological changes. *J. Environ. Econ. Environ. Manage.* **23**: 201–215.
128. Sorrell, S. & J. Dimitropoulos. 2008. The rebound effect: microeconomic definitions, limitations and extensions. *Ecol. Econ.* **65**: 636–649.
129. Brookes, L. 1990. The greenhouse effect: the fallacies in the energy efficiency solution. *Energy Policy* **18**: 199–201.
130. Saunders, H.D. 1992. The Khazzoom-Brookes postulate and neoclassical growth. *Energy J.* **13**: 131–148.
131. Madlener, R. & B. Alcott. 2009. Energy rebound and economic growth: a review of the main issues and research needs. *Energy* **34**: 370–376.
132. Greening, L.A., D.L. Greene & C. Difiglio. 2000. Energy efficiency and consumption: the rebound effect—a survey. *Energy Policy* **28**: 389–401.
133. Bentzen, J. 2004. Estimating the rebound effect in US manufacturing energy consumption. *Energy Econ.* **26**: 123–134.
134. Haas, R. & P. Biermayr. 2000. The rebound effect for space heating: empirical evidence from Austria. *Energy Policy* **28**: 403–410.
135. Roy, J. 2000. The rebound effect: some empirical evidence from India. *Energy Policy* **28**: 433–438.
136. Sorrell, S., J. Dimitropoulos & M. Sommerville. 2009. Empirical estimates of the direct rebound effect: a review. *Energy Policy* **37**: 1356–1371.
137. Schipper, L. & M. Grubb. 2000. On the rebound? Feedback between energy intensities and energy uses in IEA countries. *Energy Policy* **28**: 367–388.
138. Grepperud, S. & I. Rasmussen. 2004. A general equilibrium assessment of rebound effects. *Energy Econ.* **26**: 261–282.
139. Allan, G., N. Hanley, P. McGregor, *et al.* 2007. The impact of increased efficiency in the industrial use of energy: a computable general equilibrium analysis for the United Kingdom. *Energy Econ.* **29**: 779–798.
140. Hosier, R.H. 2004. Energy ladder in developing countries. *Encyclopedia of Energy, Elsevier* **2**: 423–435.
141. Burke, P.J. 2010. Income, resources, and electricity mix. *Energy Econ.* **32**: 616–626.
142. Berndt, E.R. 1978. Aggregate energy, efficiency, and productivity measurement. *Annu. Rev. Energy* **3**: 225–273.
143. Schurr, S. 1982. Energy efficiency and productive efficiency: some thoughts based on American experience. *Energy J.* **3**: 3–14.
144. Cleveland, C.J., R.K. Kaufmann & D.I. Stern. 2000. Aggregation and the role of energy in the economy. *Ecol. Econ.* **32**: 301–318.
145. Enflo, K., A. Kander & L. Schön. 2009. Electrification and energy productivity. *Ecol. Econ.* **68**: 2808–2817.
146. Stern, D.I. 2010. Energy quality. *Ecol. Econ.* **69**: 1471–1478.
147. Kaufmann, R.K. 1994. The relation between marginal product and price in US energy markets: implications for climate change policy. *Energy Econ.* **16**: 145–158.
148. Schurr, S. & B. Netschert. 1960. *Energy and the American Economy, 1850–1975*. Johns Hopkins University Press. Baltimore.
149. US Congress, Office of Technology Assessment. 1990. *Energy Use and the U.S. Economy*. OTA-BP-E-57, U.S. Government Printing Office. Washington, DC.

150. Stern, D.I. 2000. A multivariate cointegration analysis of the role of energy in the U.S. macroeconomy. *Energy Econ.* **22**: 267–283.
151. Stern, D.I. 1993. Energy use and economic growth in the USA, a multivariate approach. *Energy Econ.* **15**: 137–150.
152. Ma, C. & D.I. Stern. 2008. China's changing energy intensity trend: a decomposition analysis. *Energy Econ.* **30**: 1037–1053.
153. Panayotou, T. 1993. Empirical Tests and Policy Analysis of Environmental Degradation at Different Stages of Economic Development. Working Paper WP238, Technology and Employment Programme, International Labour Office, Geneva.
154. Yi, L. & H.R. Thomas. 2007. A review of research on the environmental impact of e-business and ICT. *Environ. Int.* **33**: 841–849.
155. Toffel, M.W. & A. Horvath. 2004. Environmental implications of wireless technologies: news delivery and business meetings. *Environ. Sci. Technol.* **38**: 2961–2970.
156. Romm, J., A. Rosenfeld & S. Herrmann. 1999. The Internet Economy and Global Warming: A Scenario of the Impact of E-Commerce on Energy and the Environment. The Center for Energy and Climate Solutions, The Global Environment and Technology Foundation, Arlington, VA.
157. Matthews, H.S., E. Williams, T. Tagami, and C.T. Hendrickson. 2002. Energy implications of online book retailing in the United States and Japan. *Environ. Impact Assess. Rev.* **22**: 493–507.
158. Williams, E. & T. Tagami. 2008. Energy use in sales and distribution via e-commerce and conventional retail: a case study of the Japanese book sector. *J. Ind. Ecol.* **6**: 99–114.
159. Herring, H. & R. Roy. 2002. Sustainable services, electronic education and the rebound effect. *Environ. Impact Assess. Rev.* **22**: 525–542.
160. Levinson, A. 2010. Offshoring pollution: is the United States increasingly importing polluting goods? *Rev. Environ. Econ. Policy* **4**: 63–83.
161. Aguayo, F. & K.P. Gallagher. 2005. Economic reform, energy, and development: the case of Mexican manufacturing. *Energy Policy* **33**: 829–837.
162. Kander, A. & M. Lindmark. 2006. Foreign trade and declining pollution in Sweden: a decomposition analysis of long-term structural and technological effects. *Energy Policy* **34**: 1590–1599.
163. Kander, A. 2005. Baumol's disease and dematerialization of the economy. *Ecol. Econ.* **55**: 119–130.
164. Sue Wing, I. 2008. Explaining the declining energy intensity of the U.S. economy. *Resour. Energy Econ.* **30**: 21–49.
165. Granger, C.W.J. 1969. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* **37**: 424–438.
166. Engle, R.E. & C.W.J. Granger. 1987. Cointegration and error-correction: representation, estimation, and testing. *Econometrica* **55**: 251–276.
167. Hendry, D.F. & K. Juselius. 2000. Explaining cointegration analysis: part I. *Energy J.* **21**: 1–42.
168. Kraft, J. & A. Kraft. 1978. On the relationship between energy and GNP. *J. Energy Dev.* **3**: 401–403.
169. Ozturk, I. 2010. A literature survey on energy-growth nexus. *Energy Policy* **38**: 340–349.
170. Ohanian, L.E. 1988. The spurious effects of unit roots on vector autoregressions: a Monte Carlo study. *J. Econometrics* **39**: 251–266.
171. Toda, H.Y. & P.C.B. Phillips. 1993. The spurious effect of unit roots on vector autoregressions: an analytical study. *J. Econometrics* **59**: 229–255.
172. Toda, H., & T. Yamamoto. 1995. Statistical inference in vector autoregressions with possibly integrated processes. *J. Econometrics* **66**: 225–250.
173. Stern, D.I. & M.S. Common. 2001. Is there an environmental Kuznets curve for sulfur? *J. Environ. Econ. Manage.* **41**: 162–178.
174. Hamilton, J.D. 1983. Oil and the macroeconomy since World War II. *J. Pol. Econ.* **91**: 228–248.
175. Burbridge, J. & A. Harrison. 1984. Testing for the effects of oil price rises using vector autoregressions. *Int. Econ. Rev.* **25**: 459–484.
176. Blanchard, O.J. & J. Galí. 2008. The macroeconomic effects of oil price shocks: why are the 2000s so different from the 1970s? In *International Dimensions of Monetary Policy*. J. Galí & M.J. Gertler, Eds.: 373–428. University of Chicago Press, Chicago.
177. Edelstein, P. & L. Kilian. 2009. How sensitive are consumer expenditures to retail energy prices? *J. Monet. Econ.* **56**: 766–779.
178. Yu, E.S.H. & J.C. Jin. 1992. Cointegration tests of energy consumption, income, and employment. *Resour. Energy* **14**: 259–266.
179. Warr, B. & R.U. Ayres. 2010. Evidence of causality between the quantity and quality of energy consumption and economic growth. *Energy* **35**: 1688–1693.
180. Oh, W. & K. Lee. 2004. Causal relationship between energy consumption and GDP revisited: the case of Korea 1970–1999. *Energy Econ.* **26**: 51–59.
181. Ghali, K.H. & M.I.T. El-Sakka. 2004. Energy use and output growth in Canada: a multivariate cointegration analysis. *Energy Econ.* **26**: 225–238.
182. Lee, C.-C. & C.-P. Chang. 2008. Energy consumption and economic growth in Asian economies: a more comprehensive analysis using panel data. *Resour. Energy Econ.* **30**: 50–65.
183. Lee, C.-C., C.-P. Chang & P.-F. Chen. 2008. Energy-income causality in OECD countries revisited: the key role of capital stock. *Energy Econ.* **30**: 2359–2373.
184. Costantini, V. & C. Martini. 2010. The causality between energy consumption and economic growth: a multi-sectoral analysis using non-stationary cointegrated panel data. *Energy Econ.* **32**: 591–603.
185. Glasure, Y.U. 2002. Energy and national income in Korea: further evidence on the role of omitted variables. *Energy Econ.* **24**: 355–365.
186. Masih, A.M.M. and R. Masih. 1997. On the temporal causal relationship between energy consumption, real income, and prices: some new evidence from Asian-energy dependent NICs based on a multivariate cointegration/vector error-correction approach. *J. Pol. Model.* **19**: 417–440.

187. Hondroyannis, G., S. Lolos & E. Papapetrou. 2002. Energy consumption and economic growth: assessing the evidence from Greece. *Energy Econ.* **24**: 319–336.
188. Stanley, T.D. & H. Doucouliagos. 2010. Picture this: a simple graph that reveals much ado about research. *J. Econ. Surv.* **24**: 170–191.
189. U.S. Energy Information Administration. 2010. *Annual Energy Review 2009*. United States Government Printing Office. Washington, DC.
190. U.S. Department of Commerce, Bureau of the Census. 1975. *Historical Statistics of the United States: Colonial Times to 1970*. United States Government Printing Office. Washington, DC.