TWO PARADIGMS OF PRODUCTION AND GROWTH

Emeritus Professor Robert U. Ayres, INSEAD and Chalmers Institute of Technology[a]

Dr. Benjamin Warr, INSEAD

Center for the Management of Environmental Resources (CMER)
INSEAD
Boulevard de Constance
Fontainebleau
77300
mail to : Benjamin.Warr@free.fr

Abstract

This article contrasts two incompatible paradigms of economics and their implication for economic growth. The first paradigm is consistent with the micro-foundations of neoclassical theory, which assume that all goods and services are produced from other goods plus value added by some combination of capital and labor. The theory does not explain growth, but simply assumes that technological progress (or multi-factor productivity gains) will continue indefinitely along the supposedly `optimal’ path. Related endogenous growth theory, attempts to explain the so-called Solow residual in terms of spillovers and/or increasing knowledge embodied in `human capital’, but this theory is unquantifiable – lacking satisfactory metrics for knowledge or human capital – and it still neglects the role of energy and materials The second paradigm focuses on the economy as a material resource processor-convertor. It interprets economic growth as an evolutionary process driven mostly by technological innovations (not by capital accumulation), with a strong focus on materials processing and energy (exergy) conversion. We measure resource inputs and resource conversion efficiency in thermodynamic terms. Using a new

---

[a] This work was initiated when the first author was a part-time visiting professor at the UN University, Institute for Advanced Study in Tokyo, during 2000 and 2001. It has also been supported by the European Commission under project TERRA (FP5-IST) and by the Center for Management of Environmental Resources at INSEAD.

variable exergy services or '*useful work*' as a factor of production, historical economic growth in the US since 1900 is reproduced quite accurately. Much of the previously unexplained residual is the result of productive improvements in the efficiency with which useful work is delivered to the economy, the cumulative result of innovation, learning-by-doing and economies of scale.

1. **The neoclassical paradigm**

The neo-classical paradigm is a collection of assumptions and common understandings going back in some cases to the marginalist revolution in the 19th century. It has been increasingly formalized in the second half of the 20th century. The formal version is sometimes characterized by Solow's so-called `trinity': namely, `*greed*', *rationality* and *equilibrium*. `Greed' means selfish behavior; rationality means utility maximization and equilibrium refers to the Walrasian hypothesis. There are, of course, other features of the paradigm. Production and consumption are abstractions, linked only by money flows, payments for labor, payments for products and services, savings and investment. These abstract flows are governed only by equilibrium seeking market forces (the "invisible hand"). It assumes perfect competition, perfect information, Pareto-optimality and Walrasian equilibrium. The origins of physical production in this paradigm remain unexplained, since the only explanatory variables are abstract labor and capital *services.* In the closed economic system described by Walras (Walras 1954), Cassel, (Cassel 1932) von Neumann (von Neumann 1945) and Koopmans (Koopmans 1951), every product is produced from other products made within the system, plus capital and labor services.

The unrealistic neglect of materials (and energy) in the economic system was pointed out especially by Georgescu-Roegen (Georgescu-Roegen 1971), although his criticism did not immediately lead to a paradigm shift. Growth theory remained primitive because it lacked any empirical base until the 1950s when it was discovered by Fabricant (Fabricant 1954) and Abramovitz (Abramovitz 1956) that growth could not be explained by the accumulation of capital. But the key innovation in growth theory in the 1950s was the explicit introduction of an aggregate *production function* of labor services and capital services (Solow 1956; Solow 1957), (Swan 1956). Capital services are derived from an artifact called `capital stock'. This, in turn, is an accumulation based on investment and depreciation, although some have argued that aggregate capital cannot logically be measured independently of its rate of return, and – for this and other reasons – that the concept of production function itself is faulty (Robinson 1953-54; Pasinetti 1959; Sraffa 1960; Sylos Labini 1995). Labour services were considered to flow in proportion to the total number of hours worked.

While growth can theoretically be expressed in terms of the derivatives of the production function (marginal productivities) of the input variables, the simple two-factor version introduced by Solow does not explain growth. In fact, almost 90% of the observed growth in the US during the historical period Solow chose to investigate (1909-1949) remained unexplained by the increasing capital/labor ratio. Solow named this residual `technological progress' and the annual increments are called increases in "total factor productivity"(TFP). The annual increments tend to fluctuate around a long-term trend, and enormous effort has been expended on identifying these `business cycles' with various periodicities, and attempting to explain them. Productivity calculations and projections have become a mini-industry. The new `endogenous theory' offers qualitative explanations (spillovers from 'knowledge capital' increase TFP), but nothing quantitative. It is important to realize that in this paradigm the (presumed) long-term trend in TFP itself is assumed to have been exogenously determined. It is also assumed that it will continue, much as it has in the past.

It is both common practice and convenient, although somewhat inconsistent with atomistic competition, to assume constant returns to scale, which implies that the production function should be a homogeneous function of the variables, of the first order (the so-called Euler condition). In recent years (since the work or (Romer 1986) the possibility of non-constant (increasing) returns received a good deal of attention from theorists, as well as support from empirical studies focused on international comparisons (Easterly and Levine 2001). However, the work in question is overwhelmingly based on two-factor models, primarily of the Cobb-Douglas type, assuming a population of perfectly competitive producers of a single all-purpose good.

A simple model of income allocation (applicable, however, only to a single sector model) implies that the demand for capital services and labor services will be proportional to their respective marginal productivities. In equilibrium, it follows that, if output (GDP) is only a function of capital and labor service inputs, the marginal productivities (output elasticities) of the factors of production should be equal to the corresponding payment shares (factor costs) in the national accounts. The Cobb-Douglas production function with constant returns is particularly convenient because it provides an immediate economic interpretation for the parameters of the function as elasticities. However, when a third factor is introduced this interpretation falters (Ayres, 2001).

According to the neoclassical model only labor and capital are productive and the weights with which the production factors contribute to wealth creation are assumed constant and equal to the factor cost shares. However, this rather convenient assumption has not withstood empirical research. Easterly and Levine (2001, Table 1, p 183) present a wide selection of growth accounting results for individual countries. Only for capital does the empirically determined exponent approximate its cost share (typically in the range 0.2-0.3 in industrialised countries). The empircally determined elasticity of labor varies from –0.04 to 0.42, depending upon the time period, country and study (Easterly and Levine, 2001).

The underlying assumption of growth-in-equilibrium is also troubling. It is important to note that (1) the real economy is never actually in equilibrium and (2) if it were, there would be no opportunity or incentives to innovate. Furthermore, (3) the real economy is a complex non-linear system, and non-linear systems do not exhibit equilibrium states. Moreover, (4) even if the complex non-linear system could be optimized, a dynamic optimum is not the same as a static optimum. Finally, and most troubling, (5) the lack of any theory to explain physical production in physical terms (i.e. in terms of energy and materials.)

While technical progress is normally treated as an exogenous driving force, there is an endogenous mechanism that can explain some aggregate economic growth-in-equilibrium – beyond that which is accounted for by labor and capital accumulation -- *without* radical (structure changing) technological innovations. The mechanism in question is a simple positive feedback between increasing consumption, investment, increasing economies of scale and `learning-by-doing’. These result in declining costs and declining prices, stimulating further increases in demand and investment to increase supply (***Figure 1***) This simple feedback has been called the Salter cycle (Salter 1960) and it corresponds well to many aspects of the neoclassical model.

However, if this is the only type of technological change allowed by the model, there must be declining returns and an eventual limit to growth as the potential for incremental improvements in production technology is exhausted. The closed neo-classical economic system does not explain radical innovations that change the structure of the economy. Neither is there an essential role for energy or materials, except as a consequence, (not a cause) of economic growth. In its present two-factor form the Cobb-Douglas production function permits future physical economic growth even with no materials or energy consumption. This is significant, because if resource consumption is not needed to explain growth, then `decoupling' growth from resource consumption is conceptually easy: they were never coupled in the first place. Lacking any linkage between the efficiency of materials and energy use and productivity, there is no theoretical incentive to become more efficient. There are also no consequences from generating wastes and pollutants. In the closed Walrasian equilibrium system, where all products are abstractions, there is no such thing as material waste. The neo-classical conceptualization implies falsely that wastes and emissions – if they exist at all – do no economic harm and can be disposed of at no cost.

The evolutionary paradigm

In contrast, the disequilibrium evolutionary paradigm discussed hereafter characterizes the economy at the macro-level as an open multi-sector materials/energy processing system. The system is characterized by a sequence of value-added stages, beginning with extraction of crude resources and ending with consumption and disposal of material and energy wastes, which can do harm if not eliminated. Referring again to ***Figure 1***, above, if the system is open, then the causal link between materials and energy consumption and economic growth implied by this mechanism must be *mutual*. In other words, causality must be bi-directional, not uni-directional.

This means, *ceteris paribus,* that a two-factor production function involving only labor and capital services as inputs cannot reflect this mechanism. *A third factor representing resource flows (in some way) is minimally necessary to reflect the feedback between increasing resource consumption and declining production costs.* This is needed, for example, to explain the long-term declining resource prices documented by several scholars (Barnett and Morse 1962; Potter and Christy 1968; Barnett 1979).

However, the simple positive feedback (Salter cycle) mechanism sketched in the previous section allows for only one type of technological change: namely the combined effects of scale economies and experience or learning-by-doing *at the societal level*. These forces do not distinguish between sectors, hence they cannot explain structural change. But, in reality, there is not one single aggregate technology of production for a single composite universal product, nor even a single technology for each product as assumed by activity analysis, but multiple competing technologies for each product and in each sector[1].

The qualitative evolutionary change mechanism at the firm-level (assuming abstract products) has been described by Nelson and Winter (Nelson and Winter 1974; Nelson and Winter 1982). It applies in a multi-product, multi-sector system, although mechanisms to explain structural change at the sectoral level are not considered as such. As the rate of improvement of the existing dominant technology for one product slows down, the incentives to search for, and find, a new technology (or a new material or even a new product) grow in parallel. If the demand for continued improvement is sufficiently powerful, there will be enough R&D investment to achieve a `breakthrough' enabling some radically new innovations capable of displacing the older techniques (Ayres 1988). Schumpeter's evocative word for this process was `creative destruction' (Schumpeter 1934). Spillovers from radical innovations since the industrial revolution, especially in the field of energy conversion technology, have been among the most potent drivers of economic growth.

The disequilibrium evolutionary resource-conversion perspective elaborated below implies that long-term growth, and progress towards sustainability, will require more than the gradual efficiency gains resulting from economies of scale and social learning. Radical innovations, resulting in new products and services and structural change, are also necessary. Environmental constraints (arising from material extraction, processing and consumption) are becoming increasingly important. Continued economic growth, in the sense of welfare gains, will require multiple radical technological innovations, resulting in dramatic (`factor four'/`factor ten') reductions in raw materials and energy consumption, as well as more gradual improvements such as more recycling and end-of-pipe waste treatment. Concommittantly the 'productivity' of consumed materials and energy must increase dramatically. The rate of increase in productivity of materials and energy use must clearly offset the reductions in total consumption through dematerialisation (*ceteris paribus*) ; unless dramatic improvements in the quality of the other factors of production can substitute for the role of materials and energy in generating wealth.

The next section concerns terminology and measures, and can be omitted without serious loss of clarity, provided the reader is willing to accept that `exergy' is the correct all-purpose technical term for `energy' (as the latter word is normally used), while it is also applicable to minerals and non-fuel resources.

Thermodynamics and natural resources

The term `resources' is used in many ways in different disciplines. For purposes of this paper, a resource is an input to the economic process. Resources may be material or immaterial (e.g. information) and material resources may be of natural origin or man-made. Services provided by nature (e.g. climate, air, water, bio-diversity, `assimilative capacity') are often called resources. However, in this paper, the term natural resources is restricted hereafter to energy – actually exergy– carriers, products of photosynthesis (phytomass) and other industrial raw materials extracted from the natural environment by intentional human activity[2].

The word *energy* is widely misused, and for the sake of precision we will introduce a different term, *exergy* that is less familiar but more precise. Energy is a conserved quantity (the first law of thermodynamics), which means that it can only change form or quality (e.g. temperature) but can never be created or destroyed. Energy and mass are inter-convertible (by Einstein's formula $E = mc^2$), although nuclear reactions convert only infinitesimal amounts of mass into energy, while there are no practical processes for converting energy to mass. For all other processes of concern to humans, both the mass and the energy-content of materials and energy flows are independently conserved, which means the mass and energy of inputs (including water and air) are exactly the same as the mass and energy-content of the outputs, including waste products. What has changed is the *availability* of the energy in the inputs (solar insolation or fuel) for doing work. This availability is quantifiable. A number of terms have been used for it, including `available work', `availability', and `essergy', but by general agreement it is now denoted `*exergy*.'

The formal definition of exergy is the maximum amount of work that can be extracted from a material by reversible processes as it approaches thermodynamic equilibrium with its surroundings. Exergy is therefore a quantity that is not definable in absolute terms. It can only be defined in terms of a reference state, namely the environment. But exergy can be calculated for any material with reference to whatever environmental medium that material would be likely to reach thermodynamic equilibrium with, namely the atmosphere, the ocean or the surface layer of the earth's crust (topsoil or subsoil). Thus gases tend to equilibrate with the atmosphere, liquids or soluble solids with the oceans, insoluble solids with the land. (Detailed tabulations can be found for many materials in (Ayres 1999)).

The exergy content of most fuels, per unit mass, is very nearly the same as the measure usually tabulated, which is heat-of-combustion (or *enthalpy*, to be technically accurate) per unit mass. This means that when economists or engineers speak of energy, they usually mean exergy. However, even minerals and metal ores have characteristic exergy values, which are really measures of their `distance' from thermodynamic equilibrium as defined by the average mix of materials in the lithosphere. The higher the grade of ore, the more exergy would have to be expended to achieve that degree of concentration from the crustal average, hence the greater its intrinsic exergy value. It follows that the greater the exergy content of the ore itself, the less will be needed to refine it further.

In short, exergy is a very general way of keeping track of physical scarcity and the difficulty of separation and purification. Evidently different ores and minerals can be meaningfully compared in exergy terms. This leads to the possibility of measuring all kinds of resource reserves in common (i.e. exergy) terms, for purposes of both international and inter-temporal comparison (Wall 1977; Wall 1986). It is possible to measure copper reserves, iron ore reserves, coal reserves, petroleum or gas reserves and forest biomass in the same (energy) units e.g. kiloJoules (kJ) or petaJoules (pJ).

Just as resources can be measured in common physical (exergy) units, so can pollutants. Exergy analysis can also be used empirically as a measure of sustainability, to evaluate and compare wastes and emissions from period to period or country to country (Ayres 1998). The exergy content of wastes is not necessarily proportional to the potential environmental harm the wastes may cause, but the exergy content of a waste stream is certainly a rough measure of its reactivity in air or water, i.e. its tendency to initiate spontaneous uncontrolled chemical reactions in environmental media. In this regard, one can say that, although the exergy content of a waste stream is not a measure of human or ecotoxicity, it is certainly a better measure of its potential for causing harm, than is its total mass[3].

A word on value: the use of exergy as a quantity metric does not imply that it is a measure of value (although some have suggested the idea.) Some materials, such as diamonds, gold, platinum, palladium and rhenium, have enormous economic value per unit mass, because of their aesthetic or physical properties (e.g. as catalysts). Yet other elements have very little economic value because they have no especially useful properties or because they are extraordinarily difficult to work with. For instance, the light metals, beryllium, lithium, sodium, and magnesium are quite commonplace in the earth's crust, but rarely used in industry (as metals), at least in relation to their intrinsic availability. The 9th most common metal in the earth's crust, rubidium, has virtually no uses at all.

The Economy as materials processor

From an evolutionary perspective, as noted above, the economic system can be viewed as an open system that extracts and converts raw materials into products and useful services. The economy consists of a sequence of processing stages, starting with extraction, conversion, production of finished goods and services, final consumption (and disposal of wastes). Most of the non-structural materials are discarded in degraded form. These conversion processes correspond to exergy flows, subject to constraints (including the laws of thermodynamics). The objective of economic activity can be interpreted as a constrained value maximization problem (or its dual, an exergy minimization problem). Value is conventionally defined in terms of preferences for consumption goods, although other definitions are possible.

The simplest model representation consists of two sectors with a single intermediate product. The first sector would include extraction and primary processing, e.g. to finished materials. The second sector would include manufacturing and service activities. Three or more sectors would obviously add to the realism of the scheme. Of course, the more stages in the sequence, the more it is necessary to take into account feedbacks e.g. from finished goods to extraction of primary processing sectors. The N-sector version would be a Leontieff-type input-output model in which the sequential aspect tends to be obscured.

An adequate description of the materials processing system, in our view, must include materials and energy flows as well as money flows. These flows and conversion processes are governed by the laws of thermodynamics, as well as accounting balances. At each stage, until the last, mass flows are split by technological means into `useful' and waste categories. Value (and information) is added to the useful flows, reducing their entropy content and increasing their exergy content per unit mass (thanks to exogenous inputs of exergy), while the high entropy wastes are returned to the environment.

From a macroscopic perspective the output of the economic system – viewed as a materials/exergy convertor – can be decomposed into a product of successive conversion stages with corresponding efficiencies, viz.

$$\begin{aligned} GDP \quad &= R \text{ x } \frac{IO_1}{R} \text{ x } \frac{IO_2}{IO_1} \text{ x} \ldots \text{ x } \frac{GDP}{IO_n} \\ &= R \text{ x } f_1 \text{ x } f_2 \text{ x} \ldots g \end{aligned}$$

Equation 1

where f_1 is the conversion efficiency of the resource (exergys) inflow R into the first level intermediate product, f_2 is the conversion efficiency to the second level intermediate product, and so forth. The term g is just the ratio of output to the last intermediate product. The necessary feedbacks are implicitly taken into account in the efficiency definitions.

As a first approximation, it is convenient to assume a two-stage system with a single intermediate product, U. We argue that this intermediate product can conveniently be identified as exergy services, or `useful work'. Then

$$Y = Rfg = Ug$$

Equation 2

where f is the overall technical efficiency of conversion of `raw' exergy inputs R to useful work output U, as shown in *Figure 2*. We consider the derivation of useful work in detail in the next section[4].

If we define $g = Y/U$ then *equation (1)* is an identity, not a theory or model. However, the right-hand side of *(equation 1)* might be interpreted as an aggregate production function provided g is a suitable homogeneous function of order zero, whose arguments are labor L, capital K, and resource flows R (or, as we propose, useful work, U). We consider production functions again subsequently.

Energy (exergy) conversion and useful work

Writers on energy commonly use the term `energy conversion' with reference to the use of energy to perform `useful work' (not to be confused with human `labor', as that term is understood in economics). The best explanation of useful (physical) work may be historical. Work was originally conceptualized in the 18th century in terms of a horse pulling a plow or a pump raising water against the force of gravity[5]. Since the discovery of the pendulum it has been realized that raising a bucket of water or going up a hill converts kinetic energy into gravitational potential energy, which can be re-converted to kinetic energy by reversing the process. Work is also performed when a force acting on a mass increases its velocity (accelerates it) and thus increases its kinetic energy. A piston compressing a gas does work by increasing the pressure of the gas, just as a gas expanding against a piston can do work by turning a wheel. Effectively a change in the pressure of a subsystem can generate a force capable of acting against resistance or accelerating a mass. Finally, work is performed when a process increases the chemical energy of a target substance, as when a metal ore is smelted.

Adding heat to a compressible fluid in a fixed volume (increasing its temperature) increases its pressure. This fact makes it possible to convert heat into work. However, it turns out that whereas kinetic and potential energy are inter-convertible without loss (in principle), this is not true of heat and pressure. The theory of heat engines, beginning with Sadi Carnot (1816) and subsequently extended to other engines (Rankine, Stirling, etc.) is all about converting `thermal energy’ in the form of heat into `kinetic energy’ capable of doing useful work.

It is convenient at this point to introduce the notion of `quasi-work’ not involving kinetic energy of motion. This refers to driving an endothermic chemical process or moving heat from one place to another across some thermal barrier. (Metal smelting is an example of the first; space heating or water heating are examples of the second.) Electricity can be regarded as `pure’ work, since it can perform either mechanical or chemical work with very high efficiency, i.e. with very small frictional or entropic losses. It is also convenient to distinguish primary and secondary work, where the latter is work done by electrical devices or machines. In all of these cases the physical units of work are the same as the units of energy or exergy. In physical terms, *power* is defined as work performed per unit time. Before the industrial revolution there were only four sources of mechanical power, of any economic significance. They were human labor, animal labor and water power (near flowing streams) and wind power. The advent of practical steam power in the mid-18th century, using coal as a fuel, triggered the industrial revolution.

Exergy allocation and conversion efficiency[6]

Inanimate sources of mechanical work (hydraulic turbines, steam engines and windmills) exceeded animal work in the US for the first time in 1870. However, only during the twentieth century has the primary exergy contribution from fossil fuels outstripped the contribution from biomass (agriculture and forests). In many developing countries the agricultural and forest contributions to exergy inputs are still dominant, since much of the useful work is still performed by animals and human muscles.

It is possible to estimate human (and animal) outputs of mechanical work crudely on the basis of food or feed intake, multiplied by a biological conversion efficiency. For humans this must be adjusted for the fraction of a lifetime spent doing physical (muscle) work. However, since human labor is treated independently in economic analysis – and since human muscle power is not an important component of human labor in the industrial world as compared to eye-hand coordination and brainwork – we can safely neglect it hereafter[7]. However, work done by animals, especially on farms, was still important in the US at the beginning of the 20th century and remained significant until trucks and tractors displaced horses and mules by mid-century[8].

According to Dewhurst 18.5 units of animal feed are needed to generate one unit of work (Dewhurst 1955) pp. 1113-1116, cited in (Schurr and Netschert 1960) footnote 19 p. 55. This implies an effective energy conversion efficiency of 5.4% for work animals. To confuse matters, however, more recent estimates by several authors converge on 4% efficiency or 25 units of feed per unit of work (e.g.(Grübler 1998), Box 7.1 p.321 and references cited therein). We choose the latter figure, right or wrong. Luckily, higher precision is probably unnecessary for the quantitative estimates in the US case because the magnitude of animal work is relatively small compared to inanimate power sources. For purposes of empirical estimation of other types of work, it is helpful to distinguish between two categories of fuel use. The first category is fuel used to generate heat *as such*, either for industry (process heat and chemical energy) or for space heat and other uses such as hot water for washing and cooking heat for residential and/or commercial users. The second category is fuel used to do mechanical work, via so-called `prime movers', including all kinds of internal and external combustion engines, from steam turbines to jet engines. (Electric motors are not prime movers because a prime mover, such as a steam turbine, is needed to generate the electricity in the first place).

Consumption by prime movers for the three major fossil fuels (coal, petroleum and natural gas) is plotted in *Figures 3-5*. (Fuelwood has never been used to a significant extent for driving prime movers, except in early 19th century railroads or Mississippi River steamboats, and there are no statistics on consumption.) The first of these graphs (***Figure 3***) shows the fraction of coal consumption fuel allocated to mechanical work, since 1900. During the first half of the century steam locomotives for railroads were the major users, with stationary steam engines in mines and factories also significant contributors. These uses are not distinguished in published US statistics prior to 1917. Industrial uses for heat and work were estimated by assuming that fuel consumption for each category is proportional to total horsepower in that category of prime movers, for which data have been estimated separately[9].

Figure 4 for petroleum, is based on published data for liquid fuels, by type[10]. At the beginning of the century only natural gasoline – a very small fraction of the petroleum consisting of hydrocarbons with 6 to 12 or so carbon atoms – was used for motor vehicles. The heavier, less volatile fractions had little value except for `illuminating oil' (kerosine) used for lamps in rural areas. The rapid increase in motor vehicle production and use created a correspondingly rapid growth in demand for gasoline, which led to a series of technological developments in `cracking' heavier petroleum fractions. Thermal cracking was later supplanted by catalytic cracking. Today roughly half of the mass of crude petroleum is converted into gasoline, with other liquid fuels (diesel oil, jet fuel, residual oil) accounting for much of the rest. Evidently the fraction of crude oil used to drive prime movers, rather than for heating, has been increasing for a long time.

Figure 5 for natural gas, shows the fraction of all gas consumption that is used to drive prime movers. There are two types, compressors in the gas pipelines themselves, and gas turbines used by electric utilities to generate electric power during peak periods when demand exceeds baseload supplied by other means [[USDOEEIA 1999g], Table 3].

Finally, ***Figure 6*** combining the other three, shows the fraction of all fossil fuel exergy used to drive prime movers (i.e. to perform mechanical work). This fraction has been increasing more or less continuously since the beginning of the 20th century, mostly because of the increasing fraction of fossil fuels that has been devoted to electric power generation. Transportation uses have remained roughly constant as a fraction. The other uses of fuel exergy, including industrial heating (direct or via steam), space heating, water heating, and cooking, have been declining in relative importance.

Figures 3 through *6* discussed above, reflect two different phenomena. One is structural change, including the substitution of machines for animals in transportation and agriculture, and for humans in factories and workshops. The other is increasing efficiency of conversion of heat into useful work. Needless to say, efficiency changes, reflected in prices for exergy or power, drove some of the structural changes, via the Salter cylce mechanism (*figure 1*). It is worth noting that the dramatic increases in demand for purposes of doing mechanical work have occurred despite – indeed, arguably because of – dramatic technological improvements in exergy conversion efficiency. In other words, increasing efficiency did not lead to reduced fuel consumption. Exactly the contrary occurred: prices fell sharply and demand rose even more sharply. (This phenomenon has been called the `rebound effect'.[11])

The fuel required to perform a unit of mechanical work (e.g. a horsepower-hour or kilowatt hour) from steam has decreased dramatically since 1900. In the case of electric power, the so-called `heat rate' has fallen from 90,000 Btu/kWh in 1900[12] to just about 10,000 Btu/kWh today. The heat rate is the inverse of conversion efficiency, which has increased by nearly a factor of ten, from 3.6% in 1900 or so to nearly 34% on average (including distribution losses) and 48% for the most advanced units ***Figure 7*** [Federal Power Commission, various years]. We have plotted the retail price of electricity (in constant dollars) to residential and commercial users. It will be noted that the average price fell continuously before 1972, but has risen slightly since then before stabilizing.

Estimated historical trends in the efficiency of exergy conversion (to primary work) since 1900 by category are shown in ***Figure 8***. The dominant contribution, referring to electric power generation, is measured directly and is available from the Federal Power Commission and other published sources. The other conversion efficiency trends have been estimated by the authors (Ayres and Warr 2003). Using the exergy flow and conversion efficiency data, the useful work (exergy services) performed by the US economy since 1900 can be calculated. The results are plotted for two cases – with (U_B) and without animal work (U_E) – in ***Figure 9*** along with the corresponding useful work / GDP ratios. Curiously, there is a very sudden change of slope in the useful work / GDP curve, with a peak in at 1973-74.

From exergy inputs and calculated work outputs, it is possible to derive the technical efficiency (useful work output divided by total exergy input, by year) The overall technical efficiency of exergy (resource) conversion to primary work in the US economy, since 1900, is shown for the same two cases depending on whether or not agricultural phytomass and animal work are included (***Figure 2***). In both cases the curves are monotonically increasing, as one would expect, given that technical efficiency reflects and accumulation of knowledge, even though the knowledge is largely embodied in equipment and systems.

Production and growth

The micro-foundations underlying the evolutionary theory have been discussed briefly already. While rejecting most of the neoclassical equilibrium and optimality assumptions, we retain the assumption that a production function of three factors (variables) is definable and meaningful. We also retain the assumption of constant returns to scale, meaning that it must be a homogeneous function of the first order (Euler condition).

The test of a theory is whether it can predict. For a theory of growth, if one does not want to wait twenty or thirty years for confirmation, the best hope is to explain past economic growth for a very long period, such as a century. This is what we attempt now. To do so we also need to specify a production function with as few independent parameters as possible that fits (i.e. 'explains') historical data.

There are two possible ways to determine a plausible mathematical form for a production function. One is to choose the simplest possible multi-parameter mathematical form that satisfies relevant constraints and model assumptions. The most straightforward way to incorporate resource consumption into a Cobb- Douglas (or any other) production function is to insert an additional term for resource inputs R in a Cobb-Douglas production function, which is, essentially, the simplest mathematical form that satisfies the constant returns to scale (Euler) condition and integrability conditions. While in the two-factor case the productivities (elasticities) can also be interpreted as factor shares in the national accounts, this last assumption is not plausible for the three (or four) factor case and not for a very long period of time (Sylos Labini 1995; Ayres 2001).

We have tested two versions of R, the first R_B, includes biomass (agricultural and forest products) plus non-fuel minerals. The other version R_E is limited to commercial fuels and energy sources (see *Appendix A*). Both versions are defined and measured in terms of exergy. In statistical tests (summarized in *Appendix B*) the more inclusive definition of resource inputs consistently provides a significantly better fit to the GDP data, regardless of choice of production function. While the mathematical development (equations 3-7) is valid for either definition, the quantitative results illustrated in ***Figures 11-16*** hereafter are based on the more inclusive definition of resources, which reflects the economic importance of animal labor in agriculture during the first half of the century. We note, however, that if the fitting period is shorter (say 1960-2000) the narrower definition fits equally well.

In the following we index variables to their values in the initial year (1900), viz. $y=Y/Y_0$, $k=K/K_0$, $l=L/L_0$, $r=R/R_0$ and (for later application) $u=U/U_0$. The standard form can be written

$$y = A(t)(C_K k)^a (C_L l)^b (C_R r)^{1-a-b} \qquad \textbf{Equation 3}$$

and, taking natural logarithms,

$$\ln y = \ln A(t) + a \ln(C_K k) + b \ln(C_L l) + (1-a-b)\ln(C_R r) \qquad \textbf{Equation 4}$$

Evidently a and b are still the output elasticities (marginal) productivities) of labor and capital, respectively. Assuming constant returns to scale, the term (1-a-b) is the output elasticity of resource inputs, γ. The term $A(t)$, is usually referred to as total factor productivity (TFP) while C_L, C_K, and C_R are quality multipliers of the respective variables. In the general case, all of the multipliers are also functions of time, reflecting quality improvements in each factor.

In practice, A and C_L, C_K and C_R are usually assumed to be exponential functions of time, with a general form $F = F_0 \exp(c_j t)$, whence the rates of increase are constants c_j determined by statistical analysis of past history[13]. The sum of the individual rates of increase yields the *overall* rate of increase in technical progress, i.e. the long-term multi-factor productivity growth trend mentioned earlier. *In effect, this approach incorporates the assumption that there is a basic (unexplained) productivity trend that will continue indefinitely, albeit subject to short-term fluctuations.*

As noted already, in the two-factor Cobb-Douglas scheme, a and b are the marginal productivities of labor and capital, respectively, and theorists equated those productivities with factor shares in the national accounts. Adding a third factor (resource inputs) with an assumed productivity negates the traditional identification of marginal productivities with factor payments shares. The apparent share of payments for raw materials (rents to resource owners) is clearly very small, as a fraction of GDP, which implies – under this interpretation – that the productivity of resource inputs is correspondingly small. However, the third factor is not truly independent of the other two. In particular, capital and resource flows are strongly synergistic. If the identification with factor payment shares is abandoned and the parameters a and b are determined econometrically, as in more recent models (e.g. (McKibben and Wilcoxen 1994) (McKibben and Wilcoxen 1995; Bagnoli, McKibben et al. 1996) the imputed productivity of resources is always much greater than the factor-payments share.

Once the factor-payments share identification is abandoned it makes sense to equate C_R in the Cobb-Douglas function with the exergy conversion efficiency function f defined in the previous section. Improving exergy conversion efficiency reflects tangible improvements in the quality of the fuels and materials consumed. In a first approximation (and to permit direct comparison with results from other empirical studies) the other multipliers were set equal to unity[14].

The product of resource (exergy) r inputs times efficiency f is equal to work, u. There are two cases, depending on whether phytomass is included among the resource inputs u_b or not u_e. The values of the elasticity parameters a and b are then determined indirectly by a constrained OLS fit. (See *Appendix B* for details of the statistical methodology.) The factor inputs are shown graphically in ***Figures 10a-c,*** while the best fit to GDP is displayed graphically in ***Figures 11 and 12***. It is obvious that the fit to historical GDP data is quite good, and that the remaining unexplained Solow residual (***Figure 13***) is correspondingly small.

One obvious conclusion is that long term US economic growth can be explained approximately within the Cobb-Douglas framework, with constant returns to scale, by introducing the exergy conversion efficiency function f, as a resource-enhancing multiplier. Much of the Solow residual can be attributed to productivity gains brought about by qualitative improvements to the functioning of the physical economy.

It is possible that an equally good or better fit could be obtained with the C-D model by introducing time-varying quality multipliers of labor and/or capital, as well as resource flows (certainly, the estimated parameter values a and b would differ from those identified here.) For instance, it is certainly plausible to test the idea that information technology has increased the productivity of labor, and thus that the multiplier C_L is some function of the capital invested in information technology. The difficulty, of course, is the selection of plausible measures. The other possibility is to assume that C_L and C_K are exponential functions of time (as discussed above) and to determine the best values of the corresponding rates of change.

Assuming the Cobb-Douglas fit is very good (which it is) we could claim to have produced a quantifiable endogenous theory of economic growth. However we do not like to assume either that C_L and C_K are equal to unity – which is arbitrary to say the least – or that they are exponential functions of time with constant coefficients. As noted, this choice of assumption is built into the usual approach to quantifying equation (2). To avoid making one or the other assumption it is necessary to abandon the Cobb-Douglas model.

Following an approach introduced by Kümmel et al (Kümmel, Strassl et al. 1985) we have chosen to specify factor productivities in functional forms that satisfy appropriate asymptotic conditions. It is then possible to perform a partial integration to obtain the corresponding production function. As it turns out, this procedure yields a function with just three constants of integration (one fewer than the Cobb-Douglas case), one of which can be fixed by normalization. The other two parameters are determined by a non-linear fit against the actual GDP data.

The following functional forms for factor productivities were first proposed by Kümmel et al (Kümmel, Strassl et al. 1985). There is one critical difference between our version and theirs. In our version, we substitute useful work (exergy service output) ***u*** rather than commercial energy inputs (***e* = *E*/*E*₀**) in their empirical work. The assumed marginal productivities are as follows:

$$\begin{aligned} a &= a\left[\frac{l+u}{k}\right] \\ b &= a\left[\frac{l}{u} - b\frac{l}{k}\right] \\ g &= 1 - a - b = 1 - a\left(\frac{l}{u}\right) - b\left(\frac{u}{k}\right) \end{aligned}$$

Equation 5

The third term reflects the constant returns condition, as before. Partial integration yields the following linear-exponential (LINEX) function:

$$y = u \exp\left[a\left(2 - \left(\frac{l+u}{k} \right) \right) + ab\left(\frac{l}{u} - 1 \right) \right]$$ **Equation 6**

Comparing (6) with (2), it is clear that the function ***g*** becomes

$$g = \exp\left[a\left(2 - \left(\frac{l+u}{k} \right) \right) + ab\left(\frac{l}{u} - 1 \right) \right]$$ **Equation 7**

which is a zero-th order homogeneous function of the variables, as required for constant returns to scale.

The choice of (6) as our production function, taking ***u***, ***l/u*** and ***(l + u)/k*** as the three independent explanatory, overcomes many of the problems of fitting the Cobb-Douglas function statistically (see Appendix B). The combination of useful work and ratios of labor, capital and work (plotted in *Figure 10d*) as variables explains economic growth from 1900 to 2000 even better than the Cobb-Douglas production function (*Figures 11, 12 and 13*).

It is interesting to note that the residuals for both Cobb-Douglas and LINEX models (*Figure 13*) depart from zero in identical periods, corresponding to the two world wars, the Great Depression and the energy crisis of the mid-1970s. These departures are qualitatively unsurprising, given the large imbalances characteristic of those periods. In fact, it would be quite surprising if a simple model fit perfectly during such periods. However another conclusion can be drawn from the results, namely that there is no consistent long term unexplained residual. Based on the simple theory and the empirical results exhibited above, it seems clear that useful work *done* (or exergy services delivered) can be regarded as a quality-adjusted factor of production, at least in the same sense that labor or capital could be so regarded.

In effect, much of the observed TFP or `technical progress' observed in the US over the past century can now be interpreted rather well in terms of the increasing technical (thermodynamic) efficiency with which raw materials are converted into exergy services. However, the mechanisms underlying these dramatic improvements are still 'indirectly' attributed to improvements in 'knowledge capital' (and R&D.) The difference is that, in our model, we provide a physical measure that summarizes many of the end results of these innovations and incremental improvements in a single measure, the aggregate technical efficiency *f.*

The (non-constant) marginal productivities corresponding to the LINEX function are plotted in *Figure 14.* Evidently when useful work is included as a third factor of production the corresponding productivity dominates, while the productivity of `pure' labor exhibits a long-run decline to almost zero[15]. In the Cobb-Douglas case, using three factors and constrained OLS, the fitted productivity of labor actually vanishes, yielding (in effect) a two-factor (capital-work) model. To the extent that such a simple model can represent reality, the obvious interpretation would be that the gain in aggregate economic output resulting from an additional increment of unskilled labor, *ceteris paribus*, is now very small because there is now very little that an unskilled worker can do without either tools and machines (capital) or raw materials (exergy).

Conclusions and interpretation

We have carried out necessary *gedanken* experiments, statistical tests, as summarized in *Appendix B*. They confirm the (statistical) significance of useful work in explaining economic growth[16] (*Tables B1 to B4*). Our findings, while original do not contradict those of previous, albeit slightly different studies, which revel the exemplary productivity of public investment in core infrastructure[17] (Aschauer, 1989; Munnel, 1992; Lynne and Richmond, 1993; Otto an Voss, 1996, Easterly and Rebelo, 1993). De Long et al (1991) found a very strong association between economic growth and the level of investment in equipment and machinery[18]. They attribute a large part, (over three-quarters) of the decline in the US private sector TFP since 1970 to declines in investment in public infrastructure.

De Long and Summers (1991) argue that investment in core infrastructure increases TFP growth, i.e. it is a spillover benefit, which does not depreciate or suffer from decreasing marginal productivity. While they suggest that the underlying economic mechanism is one of ‘learning-by-doing’, they do not go so far as to specify the types of ‘technical progress’ responsible for the observed productivity gains. We have identified more precisely a significant component of the mechanism that they postulate (*Figure 1*). Clearly new equipment allows producers to develop new processes and products, which are produced more efficiently (using fewer resources at lower cost). The spillovers occur as others use the equipment (e.g. computers) or simply observe the benefits that accrue and copy or innovate along similar lines. Worker mobility, poaching of employees and ideas and reverse engineering are all plausible mechanisms for such spillovers to occur.

However, simply investing in new equipment alone cannot be the only mechanism by which productivity grows. Clearly there are interactions between investment levels, policies and institutions that together make possible the process of innovation, learning-by-doing and sharing expertise. The benefits of the approach we take come, as we are able to identify those investments and exergy requiring activities that are most productive and in turn those that are not.

Some additional general interpretive conclusions seem justified at this point. One is that, since most of the thermodynamic efficiency gains have been concentrated in the primary processing stages of the process-chain (especially the conversion of fossil fuel to electric power)[19] there seems to be little or no role for downstream efficiency improvements in explaining economic growth, even though improvements in the efficiency of various industrial processes have been documented in some detail, especially since the energy crisis in the 1970s (Ayres and Warr 2003).

However there is a plausible explanation for this apparent anomaly: briefly, cost reductions for capital equipment and useful work (delivered to a point of use) have also encouraged a proliferation of low efficiency downstream uses of electric power (and a diversity of associated power devices) that compensate for upstream efficiency gains. This phenomenon can be characterized as a ‘rebound effect’. It is this rebound effect that provides perhaps the most convincing evidence for the direction of causality. Where investment in equipment and machinery are high, growth rates also tend to be high and equipment and energy prices declining. This suggests that falling equipment prices stimulate supply driven growth. If growth itself induced demand for equipment and machinery, it would also tend to push prices higher (*demand driven)*.

It can be documented that low temperature electric heat – both in industry and for space heating – has sharply increased its share of electric power production, displacing more efficient sources of heat by direct combustion (Ayres and Ayres 2003; Ayres, Ayres et al. 2003). In the industrial case, the substitution of coreless induction (electric) heaters for process steam seems to have been driven by low capital costs. In the consumer domain, new uses of low temperature electric heat include laundromats, clothes driers, dishwashers, hair driers, irons, coffee-makers, toasters, electric blankets, water bed heaters, pool heaters and so on. Electric heat continues to displace direct combustion for water heating and cooking, as well as apace heating. Space heating use has been promoted by some utilities seeking markets for surplus nuclear power capacity (especially at night).

Another example of this delayed rebound phenomenon is in the enormous proliferation of devices based on small electric motors. Examples range from vacuum cleaners and refrigerators to air-conditioners, fans, washing machines, freezers, dishwashers, garbage grinders, mixers, power tools, hair driers, record players, lawn mowers, hedge clippers, and so forth. Automobiles are full of small electric motors, not only for ventilation and climatization, which bleed off a significant fraction of the power produced by the automobile engine. In recent years a large number of rechargeable battery-powered appliances have appeared using very small electric motors. Electric carving knives, electric toothbrushes, electric drills and screw-drivers are examples, but laptop PCs are by far the most important The overall thermodynamic efficiency of battery-powered devices (taking into account the discharge-recharge cycle) is very low, but these uses take up an increasing share of total output. The driver for all these uses seems to have been the development of inexpensive small motors and solid-state controls, plus relatively cheap rechargeable nickel-cadmium batteries.

A comparable proliferation of low efficiency uses is observable in other domains. For example, recreational power boats, motorcycles, ski-doos, jet-skis, small garden tractors, power mowers, snow-blowers and leaf blowers, chain saws and hedge clippers have exploited cost reductions from mass production of (mostly) small two-stroke internal combustion engines. These examples illustrate how gains in the efficiency (of primary power sources) as far back as the 1920s have driven mass-production and R&D in downstream industries with delayed substitution effects that have continued to increase downstream demand for products utilizing electric or ICE power many decades later. This new demand has had two effects: (1) it has slowed down the overall increase in thermodynamic efficiency of the economy, but (2) it has contributed significantly to continued GDP growth. Thus, the `energy intensity' measure (*E/GDP* or, in our terms *R/GDP*) has not decreased as fast as it would if the product mix had remained constant.

There is an additional factor to consider, namely that process efficiency gains in industry have been concentrated largely in the upstream sectors where raw materials and energy (exergy) constitute a large fraction of costs. On the other hand, for most final products and services, direct inputs of energy (exergy) account for insignificant fractions of total costs; labor and capital costs are far more important. Thus, the incentives facing managers often work the other way, i.e. to reduce labor costs by substituting machines (and computers) for human labor. Moreover, until the 1970s, fuel costs and electricity costs were declining, while labor costs were rising sharply, which accelerated this substitution.

Considering both the new demand for inefficient, energy-consuming appliances and the continuing substitution of machines for human labor, Even though primary conversion efficiency it is understandable that the demand for resource (exergy) inputs has continued to grow in absolute and per capita terms. Work performed is the product of exergy inputs R (still growing smartly) times conversion efficiency f, (slowing down) and it is the product of the two variables (work) that seems to drive economic growth.

We have not yet had an opportunity to test this theory on other industrial countries, still less on developing countries. The reason is simply that we do not have reliable historical time series on the uses of fuels or other material resources, from which the efficiency functions f can be calculated. Qualitatively, the differences in output per capita are fairly easy to understand: developing countries have less capital invested in machines of all kinds. A much larger fraction of exergy inputs is devoted to producing low temperature heat for domestic purposes. Another rather obvious point of difference is that developing countries still depend very largely on human and animal labor in agriculture. There has been no increase in the efficiency of muscle work, whence no contribution from that source to economic growth. Another point of difference is that primary conversion (e.g. electric power generation) in developing countries (such as China and India) has typically been considerably less efficiency than in the West, because of smaller scale and (until recently) lack of access to the most advanced turbine technology.

Nevertheless, high potential growth rates in developing countries (subject to other enabling conditions) are qualitatively easy to understand. The efficiency function f can be increased by any investment in mechanization that takes people away from agriculture, substitutes machines for animals and uses the energy of fossil fuels for purposes other than space heating and cooking. Electrification is, indeed, the primary tool of industrialization, with combustion of liquid fuels close behind. We would expect quantitative analysis to show a very strong correlation between electrification, consumption of liquid fuels, and economic growth.

Appendix A

Table A1: Fuel mass & exergy database for the US 1900-1998; Fuel total

	Economic data		Total fuel: MASS			Total fuel: EXERGY		
	GDP (billion)) 1992$	Population millions	Production MMT	Apparent Consump- MMT	Apparent Consumption. MMT	Production eJ	Apparent Consumption eJ	Apparent Consumption eJ
1998	7552.1	270.3	1870.336	2158.685	1277.287	63.213	77.140	59.210
1997	7269.8	267.7	1858.817	2148.759	1280.759	63.302	76.888	60.123
1996	6994.8	265.2	1837.189	2094.088	1270.574	62.938	75.013	60.513
1995	6761.7	262.8	1805.465	2037.113	1232.061	62.052	73.191	58.556
1994	6610.7	260.3	1803.096	2023.080	1229.167	62.062	73.016	57.534
1993	6389.6	257.7	1714.163	1996.321	1208.046	59.683	71.799	56.179
1992	6244.4	255.0	1764.822	1937.172	1180.548	61.183	70.276	55.120
1991	6079.4	252.1	1767.326	1911.237	1165.039	61.086	69.261	53.464
1990	6136.3	248.8	1789.624	1920.600	1177.468	61.546	69.631	53.549
1989	6062.0	246.8	1747.098	1913.252	1201.329	61.025	69.765	54.981
1988	5865.2	244.5	1745.023	1895.014	1184.618	61.402	68.988	54.241
1987	5649.5	242.3	1710.178	1832.135	1124.558	60.282	66.942	51.661
1986	5487.7	240.1	1681.125	1783.015	1105.953	59.380	65.315	49.876
1985	5323.5	237.9	1705.953	1776.434	1089.339	60.383	64.464	49.471
1984	5140.1	235.8	1728.915	1783.384	1111.162	61.299	65.480	49.903
1983	4803.7	233.8	1576.096	1682.076	1066.115	56.835	62.144	47.566
1982	4620.3	231.7	1658.910	1696.151	1082.813	59.399	63.315	47.722
1981	4720.7	229.5	1664.015	1784.437	1174.505	60.272	67.036	50.120
1980	4615.0	226.5	1675.408	1804.728	1229.933	60.733	68.528	51.858
1979	4630.6	224.6	1611.657	1833.445	1312.267	59.140	70.453	53.961
1978	4503.0	222.1	1493.828	1760.530	1316.147	56.269	68.521	53.754
1977	4273.6	219.8	1480.165	1740.364	1277.686	55.675	68.181	52.736
1976	4082.9	217.6	1452.131	1639.868	1224.895	55.075	64.648	51.456
1975	3873.9	215.5	1439.400	1558.010	1161.023	54.875	61.659	48.478
1974	3891.2	213.3	1452.538	1572.100	1218.216	56.748	62.817	50.717
1973	3916.3	211.4	1487.147	1606.187	1265.018	58.746	64.475	52.541
1972	3702.3	209.3	1501.208	1529.065	1203.690	59.468	61.995	50.983
1971	3510.0	206.8	1465.746	1481.433	1135.730	58.693	60.328	49.003
1970	3397.6	203.3	1516.310	1481.824	1112.419	60.088	60.006	48.146
1969	3393.6	201.3	1438.635	1437.223	1071.108	57.420	58.129	47.187
1968	3293.9	199.3	1393.728	1392.085	1016.692	55.532	56.108	45.348
1967	3147.2	197.4	1367.415	1331.281	966.782	53.911	53.387	43.060
1966	3069.2	195.5	1302.918	1297.471	948.799	51.282	51.720	42.074
1965	2881.1	193.5	1240.890	1229.541	896.729	48.697	48.958	39.980
1964	2708.4	191.1	1204.529	1186.083	863.431	47.405	47.405	38.413
1963	2559.4	188.4	1167.599	1144.294	831.104	45.969	45.632	37.052
1962	2454.8	185.7	1106.324	1098.746	806.707	43.720	43.783	35.969
1961	2314.3	183.0	1072.810	1065.852	789.196	42.379	42.350	34.575
1960	2262.9	179.3	1072.907	1055.671	781.037	42.098	41.698	34.187
1959	2210.2	177.1	1057.683	1029.408	763.873	41.219	40.531	32.936
1958	2057.5	174.1	1020.443	992.192	745.551	39.348	38.633	31.870
1957	2078.5	171.2	1119.353	1056.827	764.301	42.234	40.596	31.614
1956	2040.2	168.1	1120.507	1061.581	778.855	42.010	40.355	31.916
1955	2001.1	165.1	1055.371	1021.134	761.420	39.695	38.341	30.803
1954	1868.2	161.9	957.730	930.300	699.550	36.206	34.968	28.215
1953	1881.4	159.0	1019.445	983.504	748.999	37.758	36.439	28.822
1952	1798.7	156.4	1018.510	956.371	739.389	37.469	35.484	28.466
1951	1734.0	154.0	1069.080	987.730	771.329	38.270	35.681	28.738

Table A1: Fuel mass & exergy database for the US 1900-1998; Fuel total

	Economic data		Total fuel: MASS			Total fuel: EXERGY		
	GDP (billion)) 1992$	Population millions	Production MMT	Apparent Consump- MMT	Apparent Consumption. MMT	Production eJ	Apparent Consumption eJ	Apparent Consumption eJ
1950	1611.3	151.3	994.175	924.269	746.158	35.116	32.582	27.182
1949	1479.8	148.7	897.257	879.446	717.942	31.424	30.410	25.205
1948	1488.6	146.7	1070.135	960.263	794.101	36.402	32.726	24.841
1947	1426.2	144.1	1068.513	955.133	738.433	35.556	31.642	24.326
1946	1437.9	140.7	958.112	880.363	702.655	31.977	28.948	22.813
1945	1632.7	133.4	990.669	928.802	737.279	32.603	29.947	23.364
1944	1660.5	133.9	1024.175	946.117	741.805	33.273	30.204	23.271
1943	1551.6	135.1	955.553	915.756	719.896	30.617	28.698	22.072
1942	1370.4	134.6	924.135	847.904	667.232	29.256	26.492	20.401
1941	1213.3	133.7	869.305	824.349	656.252	27.422	25.707	20.060
1940	1044.7	132.5	807.503	750.389	604.699	25.438	23.350	18.396
1939	963.6	130.9	735.491	689.501	567.683	23.095	21.317	17.090
1938	887.5	129.8	679.177	639.744	540.082	21.243	19.658	16.101
1937	934.4	128.8	775.429	730.360	601.227	23.950	22.190	17.734
1936	888.7	128.1	743.516	700.055	582.790	22.382	20.950	16.869
1935	779.6	127.3	667.943	636.901	542.328	19.981	18.848	15.477
1934	709.6	126.4	651.710	617.335	526.236	19.182	17.911	14.666
1933	650.8	125.6	619.715	588.384	510.131	18.359	17.211	14.392
1932	663.7	124.8	580.960	572.632	501.423	17.075	16.711	14.075
1931	778.3	124.0	655.205	638.053	542.575	19.532	18.918	15.515
1930	843.5	123.1	741.285	730.055	608.543	22.146	21.834	17.561
1929	937.2	121.8	807.374	767.058	631.507	24.324	22.854	18.017
1928	878.3	120.5	755.338	732.134	611.767	22.636	21.675	17.371
1927	873.0	119.0	771.815	722.444	605.739	22.972	21.179	17.021
1926	873.2	117.4	802.141	747.114	625.506	23.538	21.906	17.602
1925	825.1	115.8	732.710	702.712	590.941	21.606	20.563	16.650
1924	760.9	114.1	718.580	696.858	597.006	21.087	20.219	16.661
1923	763.1	112.0	796.235	732.064	615.024	22.912	21.173	17.141
1922	680.3	110.1	610.495	608.639	525.072	17.434	17.206	14.281
1921	587.4	108.5	634.238	599.567	534.940	17.553	16.519	14.150
1920	643.6	106.5	766.811	706.050	600.469	21.086	19.389	15.765
1919	673.1	104.5	662.151	635.292	547.859	18.133	17.343	14.254
1918	698.7	103.2	775.976	743.896	641.390	21.078	20.105	16.498
1917	621.5	103.3	745.221	709.872	607.943	20.308	19.217	15.648
1916	617.1	102.0	681.922	649.693	554.473	18.616	17.578	14.262
1915	573.2	100.6	623.728	595.820	520.694	16.955	16.057	13.378
1914	578.3	99.1	604.371	579.401	515.684	16.386	15.567	13.226
1913	605.2	97.2	648.711	618.767	540.472	17.578	16.622	13.836
1912	598.1	95.4	612.067	586.325	512.668	16.548	15.699	13.076
1911	565.9	93.9	583.537	556.784	496.323	15.753	14.921	12.706
1910	552.3	92.4	582.620	562.480	494.594	15.746	15.064	12.636
1909	537.2	90.5	539.490	519.798	456.424	14.544	13.874	11.613
1908	461.6	88.7	499.435	481.343	437.427	13.425	12.809	11.142
1907	503.0	87.0	553.181	534.932	471.254	14.817	14.209	11.936
1906	494.4	85.5	488.977	473.889	417.112	13.033	12.510	10.490
1905	442.4	83.8	473.112	459.334	409.197	12.595	12.117	10.306
1904	413.1	82.2	435.924	423.697	385.930	11.540	11.118	9.695
1903	418.0	80.6	441.789	432.058	392.680	11.548	11.198	9.725
1902	399.0	79.2	392.577	383.927	344.964	10.203	9.877	8.469
1901	395.1	77.6	385.247	374.668	340.931	9.852	9.474	8.223
1900	354.0	76.1	365.776	355.194	323.484	9.290	8.920	7.754

Appendix B

This appendix summarises the results of empirical analyses, which provided much of the basis for the interpretations presented in the main text. Regression analyses and linear optimisation methods were used to parameterize both the natural resource augmented (three parameter) Cobb-Douglas production function and the LINEX production function using the data presented in Appendix A. Prior to fitting each function using the appropriate method, we performed an extensive regression analysis.

Empirical research into the causes of economic growth is fraught with problems, not least (i) choosing the appropriate (independent) explanatory variables (inputs to the production function) and (ii) specifiying the appropriate form of the production function. Of critical importance is (iii) the method with which the parameters of the production function are identified. The ‘open-endedness’ of the theory lies at the heart of the problem. Data error and model mispecification give rise to additional concerns when interpreting parameter estimates and predictions. To summarise the current situation, it is generally accepted that no single statistical test, considered in isolation, is sufficient to provide the evidence that can definitively validate one theory over another. The utility of statistical tests lies in their ability to provide, together, a ‘body of evidence’ that must be considered in light of alternative possible theoretical explanations and exchangeable models for observed trends, relationships and associations (Brock and Durlauf, 2001).

The majority of publised studies fitting (Cobb-Douglas) production functions have used ordinary least squares (OLS), either restricted to ensure constant returns to scale, or not. Necessary conditions for stable and unbiased parameter estimation, are often violated by long term economic time series. The OLS method minimizes the sum of the squared residuals. If certain assumptions can be met, OLS provides the so-called best linear unbiased estimators (BLUE). The assumptions include:

Assumption	***Consequence of failure***
(1) There is no correlation between explanatory variables and residuals (no simultaneity).	Biased estimates of the coefficients of explanatory variables.
(2) The expected or mean value of the residuals equals zero.	Biased estimate of the trend (i.e. the mean of a stationary time series (see 8 below.)
(3) Residuals are homoscedastic having	Inefficient estimates and biased tests of

constant variance (no heteroscedasticity).	hypotheses.
(4) Residuals are independently distributed (no serial correlation).	Inefficient estimates and biased tests of hypotheses.
(5) Explanatory variables are independent (no multicollinearity.)	Inefficient estimates and biased tests of hypotheses.
(6) Residuals are normally distributed with constant mean and variance.	Invalidates the use of the Student t-distribution in coefficient t –tests.
(7) Explanatory variables are measured without error (no errors in variables)	Biased estimates of the coefficients.
(8) Variables that are time series must be 2nd order stationary, i.e. have constant mean and variance.	Spurious regressions (large but meaningless r^2 values) may be observed.

While listing assumptions is easy, checking the data to ascertain whether any are contravened is not. It is not simply a matter of testing one after the other using simple statistical tests. Rather, the regression analysis must proceed in steps. The first stage in our analysis was to transform the data to a form suitable for the application of ordinary least squares. The second stage involved identification of the correct inputs, functional form and fitting procedure. The final step of the analysis involved fitting ‘optimal’ (constrained} versions of both models. The final results (using u_B) are presented in ***table 1*** (in the main text).

The constraints in the fitting procedure were required to ensure constant returns to scale and also that the marginal productivities of the factors were non-negative. The constraints were imposed on the Cobb-Douglas model by restricted regression and for the LINEX by using a non-linear optimisation algorithm. In the following two sections we outline how we arrived at these results in more detail.

Data Preparation

It is common practice prior to any econometric analysis involving ordinary least squares to standardise and (log) transform the observed data to linearise the production function (necessary conditions for assumptions 1 to 3). Errors from the OLS fitted model are assumed to be *independent* (from the predictions) and *identically* (Gaussian) *distributed* (*iid*), and therefore amenable to various statistical tests to quantify the reliability of the parameter estimates and to compare alternative models. Macroeconomic time series covering a hundred year time period are rarely if ever stationary. Moreover, it is not reasonable to assume that the error variance will be the same for predictions made in 1900 and those in 2000.

Stationarity and ergodicity are properties of the underlying stochastic process and not properties of a single realization of that process. Stationarity (2nd order, as mentioned in 8 above) is the property that the expected values of the 1st and 2nd moments (mean and variance) of the process are independent of the temporal index. Ergodicity is the property that the expected values of the moments of the process are equal to the time averages of the moments of the process. Since the expectation operator is the average over all realizations of the process, you can't say anything for sure with just a single realization of the process. If, however, you assume in advance that the process is ergodic, then you can draw conclusions about ensemble averages knowing only the time average, which you can estimate from a single realization. Thus, if you have only a single time series and you assume ergodicity, then you can test for stationarity by seeing whether the statistics (mean, variance, etc.) as estimated in one part of the time series are close enough to the statistics as estimated in other parts of the time series. Of course, any definition of stationarity depends entirely upon the scale of the period under consideration. As a rule of thumb, we should certainly not consider performing any type of staightforward statistics on ‘mixed populations’ where the trend is clearly different across regions[1]. There are no striaghtforward tests for ergodicity.

Often a natural log transformation of the data is enough to meet this condition is met. However, where the time series are particularly long, this may not be sufficient to remove the observed temporal trend in the mean (and hence the mean of the residuals) to render the time series *stationary*. *Figures 10a-*c show the raw data, natural log transforms and the increments of the log transforms. Only the latter approach stationarity. Hence, for the analyses involving OLS (fitting procedures to estimate the appropriate Cobb-Douglas parameters*)* we used increments of log-transforms[2]. It was not necessary to transform the data when fitting the

---

[1] The interested reader is referred to works by Matheron (1967), inventor of geostatistics, the study of continuous variables in space and time using statistical methods. In this field of study, the concept of stationarity and its implications are discussed in detail.

[2] For the purposes of comparison with other studies we also present the results using more commonly used log transforms. However, the standard error of the parameter estimates were considerably larger than when using the stationary increments of log transforms for prediction.

LINEX function, because optimisation methods make no prior assumptions about the statistical distribution of the variables[3], relying on brute (computing) force and robust constraints.

Prior transformation of the variables also overcomes problems with OLS caused by multicollinearity. Collinearity describes the presence of strong linear relationships between the explanatory variables and is common in economic data. The presence of collinearity can be assessed by looking at the correlations between the variables and plots of the series against time. However, collinearity is imprecisely defined. It is not possible simply by looking at sample statistics to determine whether or not collinearity will pose specific problems during fitting, or to estimate its direct contribution to uncertainty in the parameter estimates. It is important to be aware of the problems that may result and to identify robust models on the basis of the tests of alternative parameter estimates, between fits where trends from the data which may be responsible for the observed collinearity have been removed and from those where they have not.

The basic problem is that when two (or more) of the explanatory variables move together it is not possible to determine their separate influences on the variable of interest. In theory, this does not mean that the parameter estimates will necessarily be systematically biased, but rather that their estimated values are subject to considerable uncertainty, having large standed errors. Consequently, in the presence of strong multicollinearity it is not possible to draw robust conclusions from OLS parameter estimates.

In this study, the prior log-transformation and differencing (to form increments), to achieve stationarity removed the trend components of the time series that were likely to cause spurious correlation. In *Table B2* the figures in brackets are the coefficients between detrended variables (increments of logs). We conducted OLS fits for the Cobb-Douglas production function using (5yr smoothed) data, log-transforms (of the smoothed data) and finally year-to-year increments of the log transforms. These tests revealed the consistent reduction in the magnitude of the parameter estimate standard error and root mean squasre error of prediction, and increasing overall significance of the models as evidenced by the model F-statistics. The best fits were obtained using the time series of the increments of natural logarithms (i.e. the 2nd order stationary time series, as expected)

---

[3] It is also worth noting that it is also possible to use optimisation methods to estimate the parameters of the Cobb-Douglas production function, however any statistical interpretation of the paramters is lost.

While these tests alone are insufficient to prove the non-existence of multicollinearity it is important to note that the LINEX producution function is a linear function of one factor and an exponential function of ratios of (combinations) of the factors (*Figure 10d*). The correlations between these inputs suggest that there is little reason to consider that multi-collinearity should negatively affect the results of applying OLS to the factor inputs and ratios used in the LINEX function (*Table B2, Figure 10d*). Visual examination of the time series shows that the series do not grow in parallel, each desribing its own trend in time.

Prior to fitting the LINEX function, using an optimisation algorithm to minimise the sum of squared residuals, we applied OLS using the LINEX function inputs to predict output. The first model was an unrestricted regression for which three parameters were estimated (*TableB3*) corresponding to the weights for u_B, and the two factor ratios. The second model was a restricted form, wherein the weight for u_B was fixed equal to 1 (similar to the form of the LINEX function itself.) Tests to compare the overall significance of the two models indicated that two parameters (weighting just the ratios) suffice in the LINEX model. These results suggest that output is directly proportional to the amount of useful work services flowing through the economy, modified by the dynamics of the factor ratios in the model. Morevoer the estimated weights of the factor ratios were similar for both models. This type of consistency when modelling provides important insights into the stability of the model, an important tests to be completed prior to fitting the LINEX function by deterministic methods.

Choosing the appropriate inputs to the production function

Natural resource (energy) augmented production functions are a relatively new introduction to economic theory. There is no concensus about the appropriate definition for the natural resource representation. Among the various definitions employed include the mass of raw materials, electricity or commercial fuel energy consumption content and so on. The inputs to the production functions, from which we shall identify the optimal combination, are presented in Appendix A. They are independent variables in the sense that within technological limits entrepreneurs can vary each, the quantity of capital, labor or the degree of capital utilisation (and hence energy requirements) independently (Kümmel, 2001).

We suggest that useful work is the correct factor of production, not total natural resource exergy consumption. There are strong theroetical reasons to believe that it is the power (useful work/year) flowing through the economic system, as opposed to the total raw exergy required to provide this, which is productive in the economy. Indeed, a certain fraction of the primary exergy consumed, for example that part contributing to global climate change, may negatively impact output growth (particularly in the long run).

Tests revealed that past growth, in either u_B and u_E, was more strongly correlated with output growth than that of consumption of primary exergy (*Table B1*). All the tests that we describe subsequently were completed using each of the four variables. However, the test statistics of each model using either u_E or u_B as a predictor of output Y revealed the significant improvement in the predictive capacity of useful work over primary exergy. For the sake of clarity we present the results using u_B, because in each instance this variable provided the most statistically significant models, stable parameter estimates, and superior predictions. We can interpret this finding. Accounting for the substitution of useful work from powered 'tools' for animal and human muscle work improved the fit during the first part of the century, when this substitution was responsible for important productivity gains.

Also the results of using (unrestricted) OLS to predict output (using u_B) show clearly the improvement in fit provided by the inputs to the LINEX function over those used in the Cobb-Douglas production function (*Table B3*). *These results suggest that using ratios of the factors adds significant explanatory power to the model as opposed to using the factors directly*. These results confirm that in either model u_B is the most significant factor 'explaining' output growth[4]. Note that the sum of the exponents in the unrestricted Cobb-Douglas model is 1.3. However, after comparison of the unrestricted (F-statistic : 101) and the restricted model (F-statistic: 34), we accepted the hypothesis of constant returns to scale. Subsequently all Cobb-Douglas models were restricted to ensure constant returns to scale ($\alpha + \beta + \gamma = 1$). Similary the LINEX function was optimised under constraint that the elasticities (*Equation* 5) were non-negative and summed to 1[5].

Model Fitting

Following the tests described above, we fit the Cobb-Douglas model using restricted OLS and increments of logs, and the LINEX function to the raw data using a non-linear optimisation algorithm, under constraint of non-negativity of the marginal productivities (*Equation 5*). The results are presented in *Table B4*. The LINEX fit is far superior to the Cobb-Douglas

---

[4] Note that this is not a proof of causality.

[5] It is important to note that while the LINEX production function ensures that the assumption of constant returns to scale is met, that the elasticities are not necessarily constant over time (see Figure 15), reflecting the possibilities of dynamic substitution (Sylos Labini, 1995).

alternative as evidenced by the root mean square error and durbin-watson statistics. While the value of the latter was not sufficient to reject a hypothesis of serial correlation of the residuals, it is an acceptable value for a fit to a 100 year time series. *Figure 14* shows the temporally variable elasticities calculated using the optimal LINEX parameters (*a =0.12, b= 3.78, RMSE 0.355,* r^2 *0.99, DW 0.06,*). It is interesting to note that the elasticity for labor goes almost to zero by the end of the century. The reason for this is the constant 'technology parameters' used in the LINEX function. Kümmel et al (2001) use time dependent parameters, which vary according to diminishing returns, following a logistic-type trend over time. Their input to the production function was primary exergy consumed rather than useful work. Changes in technology are effectivley 'resumed' in the variable useful work, thereby removing the need for time variable parameters, and permitting a more parsimonious prediction of output using only a two-parameter function.

Table B1. Statistical measures of significance of regressions of work (*Ub or Ue*) or primary exergy (*B or E*) for output Y (*significant at 0.05 level).

Explanatory Variable	***RMSE***	***R^2***	***F-statistic****
B	1.061	0.967	2828
E	1.251	0.953	2007
Ub	0.722	0.984	6218
Ue	0.744	0.981	5851

Table B2. Correlation coefficients between variables (Figures in brackets for increments of logs).

	Y	***K***	***L***	***B***	***Ub***	***L/Ub***	***(L+Ub)/K***
Y	1						
K	0.99	1					
L	0.99 (0.72)	0.99	1				
B	0.98 (0.73)	0.97	0.97	1			
Ub	0.99 (0.74)	0.99	0.97	0.99	1		
L/Ub	-0.75 (-0.78)	-0.73	-0.78	-0.81	-0.77	1	
(L+Ub)/K	0.65 (0.43)	0.59	0.66	0.75	0.70	-0.85	1

Table B3. Comparison of unrestricted OLS fits for output (*Y*) using Cobb-Douglas and LINEX production functions. (Figures in brackets correspond to parameter standard errors and t-test statistics, respectively, all tested significant at 0.05 level)

	Cobb-Douglas	**LINEX**
Adjusted R^2	0.75	0.98
Standard Error	0.008	0.058
F statistic	101	7636

Cobb-Douglas (unrestricted OLS)

$$\Delta\ln(Y) = 0.11\Delta\ln(K) + 0.68\Delta\ln(L) + 0.53\Delta\ln(Ub)$$

(0.055; 2.01) (0.088 ; 7.7) (0.046 ; 11.6)

LINEX (unrestricted OLS)

$$\ln Y = 0.93\ln(Ub) + 0.14\ L/Ub - 0.11\ (L+Ub)/K$$

(0.014 ; 66) (0.031 ; 4.4) (0.012 ; -9.9)

Table B4. Summary of optimal model fitting results (* time average parameters.)

Production Function	***Fitting Method***	*a*	*b*	*g*	***Root Mean Square Error***	***Durbin-Watson Statistic***
Cobb-Douglas	*Restricted Regression*	0.01 (Standard Error: 0.06)	0.31 (Standard Error: 0.05)	0.66 (n/a)	1.86	0.006
LINEX	*Constrained Optimisation* a = 0.12 b = 3.78	(0.36)*	(0.08)*	(0.56)*	0.35	0.059

Endnotes

---

[1] Sectors are ultimately defined in terms of product families, which have gradually become increasingly differentiated over time. The sectoral structure of the economy has evolved as a consequence of a large number of micro-mutations (so to speak) at the product and process level.

[2] This is not to exclude, on a theoretical basis, the productive role of a more broadly defined set of ecosystem services, simply recognition that quantification of such services is impractical.

[3] The above remarks have direct relevance to the notion of using total (including indirect) mass flows associated with mining, agriculture and construction as a `measure' of sustainability. This idea has been popularized bySchmidt-Bleek (1992), Bringezu, S. (1993), Adriaanse, A., S. Bringezu et al. (1997)

[4] At this point it is important to note that we teseted two versions of U (and f), depending on whether phytomass exergy is included (subscripted with B) or not (subscripted with E). All statistical tests and fits were completed using both versions (Appendix B). In each test the inclusion of *useful work* from phytomass sources (i.e. fits using U_B and f_B) improved the quality and the significance of the fit.

[5] The first steam engines were used for pumping water from mines, an application where horses had previously been used. This enabled a direct comparison to be made. Ever since then power has been measured in terms of `horsepower' or a metric equivalent.

[6] The work described in the next section was performed in collaboration with Benjamin Warr. See Ayres, R. U. and B. Warr (2002, 2003).

[7] Human muscle work was already negligible by comparison at the beginning of the twentieth century. The US population in 1900 was 76 million, of which perhaps 50 million were of working age. but only 25 million were men (women worked too but their work was not monetized and hence did not contribute to GDP at the time). At least half of the male workers were doing things other than chopping wood, shoveling coal or lifting bales of cotton,

---

which depended more on eye-hand coordination or intelligence than muscles. The minimum metabolic requirement is of the order of 1500 Cal/day (for men), whereas the average food consumption for a working man was about 3000 Cal/day. Thus no more than 1500 Cal/day was available for doing physical (muscle) work. This comes to 18 billion Cal/day or about 0.16 EJ/year of food exergy inputs for work. Assuming muscles convert energy into work at about 15% efficiency, the overall food-to-work conversion efficiency for the human population as a whole would have been roughly 2.4%. In recent years, though most women now have jobs, given the changing nature of work, and the much greater life expectancy and retirement time, the average conversion efficiency has declined significantly.

[8] The numbers of horses and mules, by year, in the US is given in United States Bureau of the Census (1975). Historical statistics of the United States, colonial times to 1970

[9] [[USCensus 1975], Table S 1-14, p. 818) Electric power generation gradually became the by far the dominant use of coal, as it is today [[USCensus 1975] Tables M-113,114, p.591 and S-100, p. 826.

[10] The basic sources of data are [[USCensus 1975] M-162-177 p. 596, and Energy Information Administration Office of Energy Markets and End Use (1999).

[11] The `rebound effect' has recently preoccupied energy conservation advocates Lovins, A. B. (1977, 197, 1998). The point is that in some circumstances energy efficiency gains translated into price reductions result in demand increases that over-compensate for the efficiency gains, thus undermining the case for attempting to achieve conservation through increasing efficiency.

[12] Btu refers to British thermal units, a measure still widely used in industry and government. For instance, the most common measure of energy in US statistics is the “quad” which is defined as 10^{15} Btus. The more usual

---

metric unit of energy is kiloJoules (kJ) and one Btu = 1.055 kJ. One kilowatt hour (kWh) of electric energy is equivalent to 3600 kJ or 3412 Btu. The conversion efficiency is the ratio of output to input, in consistent units. Thus 3412 Btu divided by 90,000 Btu corresponds to an efficiency of about 3.6 %.

[13] The statistical analysis involves several steps, to avoid problems of co-linearity. The first step is to transform from product (multiplicative) form to logarithmic (additive) form. However, the errors still show a (time) trend, whence it is usual to perform the OLS calculation on the annual increments. This does yield errors that approximate stationarity (no trend). We have done this both with and without the constraint of constant returns. Without the constraint, the sum of the three calculated output elasticities turns out to be of the order of 1.3, which is implausibly high.

[14] In Solow's original study (1957) he defined labour simply as hours worked and capital using standard definitions. We have done the same here and discuss the implictions and possible measures for C_K and C_L subsequently.

[15] We used constant parameters in the LINEX function. Kümmel et al (2001) use time dependent parameters, effecitively a year on year recalibration of the model. This is not a parsimonious solution, as a result of introducing time dependence, but does mean that the calculated marginal productivities do not reach asymptotic limits.

[16] Note in particular the highly significant correlation between either $\ln(u_B)$ and $\Delta\ln(u_B)$ in both regressions presented in Table B2.

[17] While public investment includes a braod mix of projects, including schools, hopsitals, but importantly including also power generation, transport and other communications facilities.

[18] Although some criticisms were raised in the 1992 Brookings papers, concerning the de Long and Summers' (1991) statistical analysis, the focus of concern was on causality and the presence of spurious correlations, to which de Long and Summer had several convincing rebuffs.

[19] This is logical, considering that it is in the primary processing sectors that raw material and fuel (exergy) inputs constitute a large fraction of manufacturing costs and consequently it is where reducing overall costs of production could best be accomplished by reducing raw material and energy costs.

References

Abramovitz, M. (1956). "Resources and output trends in the United States since 1870." American Economic Review **46**.

Adriaanse, A., S. Bringezu, et al. (1997). Resource flows: The material basis of industrial economies. Washington DC, World Resources Institute.

Ayres, R. U. (1988). "Barriers and breakthroughs: An expanding frontiers model of the technology industry life cycle." Technovation(7): 87-115.

Ayres, R. U. (1998). Technology, energy and materials. Rotterdam, The Netherlands, Research Center for Economic Policy, Erasmus University.

Ayres, R. U. (1999). "The 2nd law, the 4th law, recycling and limits to growth." Ecological Economics **29**(3): 473-384.

Ayres, R. U. (2001). "The minimum complexity of endogenous growth models: The role of physical resource flows." Energy – The International Journal, 26(9): 817-838.

Ayres, R. U. and L. W. Ayres (2003). Is the economy dematerializing? Main indicators and drivers. The economics of industrial ecology: methods and applications. M. Janssen and J. Van den Bergh. Cheltenham, UK, Edward Elgar (forthcoming).

Ayres, R. U., L. W. Ayres, et al. (2003). On the efficiency of electricity usage since 1900. Fontainebleau, France, Center for the Management of Environmental Resources, INSEAD**:** 149.

Ayres, R. U. and B. Warr (2002). Economic growth models and the role of physical resources. Unveiling wealth: On money, quality of life, and sustainability. P. Bartelmus. Dordrecht/Boston/London, Kluwer Academic Publishers**:** 171-188.

Ayres, R. U. and B. Warr (2003). "Exergy, power and work in the US economy 1900-1998." Energy-The International Journal **28**: 219-273.

Ayres, R. U. and B. Warr (2003). Useful work and information as drivers of growth. Fontainebleau, France, INSEAD.

Bagnoli, P., W. J. McKibben, et al. (1996). Future projections and structural change. Climate Change: Integrating Science, Economics and Policy, Laxenburg, Austria, International Institute for Applied Systems Analysis (IIASA).

Barnett, H. J. (1979). Scarcity and growth revisited. Scarcity and Growth Reconsidered. V. K. Smith. Baltimore MD, Johns Hopkins University Press**:** 163-217.

Barnett, H. J. and C. Morse (1962). Scarcity and growth: The economics of resource scarcity. Baltimore MD, Johns Hopkins University Press.

Bringezu, S. (1993). "Towards increasing resource productivity: How to measure the total material consumption of regional or national economies?" Fresenius Environmental Bulletin **2**: 437-442.

Cassel, G. (1932). The theory of social economy. New York, Harcourt-Brace.

Dewhurst, J. F. A. (1955). America's needs and resources: A new survey. New York, Twentieth Century Fund.

Easterly, W. and R. Levine (2001). "It's not factor accumulation: stylized facts and growth models." The World Bank Economic Review **15**(2): 177-219.

Energy Information Administration Office of Energy Markets and End Use (1999). Annual energy review 1998. Washington DC, United States Department of Energy, Energy Information Administration.

Energy Information Administration Office of Oil and Gas (1999). Historical natural gas annual 1930 through 1998. Washington DC, United States Department of Energy, Energy Information Administration.

Fabricant, S. (1954). Economic progress and economic change. 34th Annual Report. New York, National Bureau of Economic Research.

Georgescu-Roegen, N. (1971). The entropy law and the economic process. Cambridge MA, Harvard University Press.

Grübler, A. (1998). Technology and global change. Cambridge, UK, Cambridge University Press.

Koopmans, T. C., Ed. (1951). Activity analysis of production and allocation. New York, John Wiley and Sons.

Kümmel, R., W. Strassl, et al. (1985). "Technical progress and energy dependent production functions." Journal of Economics **45**(3): 285-311.

Kümmel, R., and D. Lindenbergyer, 2002. Energy-Dependent Production Functions and the Optimisation Model "PRISE" of Price Induced Sectoral Evolution. *Int. J. Applied Thermodynamics*. Vol. 5, No. 3, pp.101-107.

Lovins, A. (1998). "Energy efficiencies resulting from the adoption of more efficient appliances: another view." Energy Journal **9**(2): 155-162.

Lovins, A. B. (1977). Soft energy paths: Towards a durable peace. Cambridge MA, Ballinger Publishing Company.

Lovins, A. B. (1978). "Soft energy technologies." Annual Review of Energy **3**: 477-517.

Lovins, A. B., L. H. Lovins, et al. (1981). Least-cost energy: Solving the CO2 problem. Andover MA, Brickhouse Publication Co.

McKibben, W. J. and P. J. Wilcoxen (1994). The global costs of policies to reduce greenhouse gas emissions III,. Washington DC, The Brookings Institution.

McKibben, W. J. and P. J. Wilcoxen (1995). The theoretical and empirical structure of the G-cubed model. Brookings Discussion Papers on International Economics. Washington DC.

Nelson, R. R. and S. G. Winter (1974). "Neoclassical vs. evolutionary theories of economic growth: Critique and prospectus." Economic Journal **84**(336): 886-905.

Nelson, R. R. and S. G. Winter (1982). An evolutionary theory of economic change. Cambridge MA, Harvard University Press.

Pasinetti, L. L. (1959). "On concepts and measures of changes in productivity." Review of Economics and Statistics **XLI**(3): 270-286.

Potter, N. and F. T. J. Christy (1968). Trends in natural resource commodities. Baltimore MD, Johns Hopkins University Press.

Robinson, J. (1953-54). "The production function and the theory of capital." Review of Economic Studies **21**(1): 81-106.

Romer, P. M. (1986). "Increasing returns and long-run growth." Journal of Political Economy **94**(5): 1002-1037.

Salter, W. E. G. (1960). Productivity and technical change. New York, Cambridge University Press.

Schmidt-Bleek, F. B. (1992). "Ecorestructuring economies: Operationalizing the sustainability concept." Fresenius Environmental Bulletin **1**(46).

Schumpeter, J. A. (1934). Theory of economic development. Cambridge MA, Harvard University Press.

Schurr, S. H. and B. C. Netschert (1960). Energy in the American economy, 1850-1975. Baltimore MD, Johns Hopkins University Press.

Solow, R. M. (1956). "A contribution to the theory of economic growth." Quarterly Journal of Economics **70**: 65-94.

Solow, R. M. (1957). "Technical change and the aggregate production function." Review of Economics and Statistics **39**: 312-320.

Sraffa, P. (1960). Production of commodities by means of commodities: prelude to a critique of economic theory. Cambridge, UK, Cambridge University Press.

Swan, T. (1956). "Economic growth and capital accumulation." The Economic Record **32**(68): 334-361.

Sylos Labini, P. (1995). "Why the interpretation of the Cobb-Douglas production function must be radically changed." Sctructural Change and Economic Dynamics **6**: 485-504.

United States Bureau of the Census (1975). Historical statistics of the United States, colonial times to 1970. Washington DC, United States Government Printing Office.

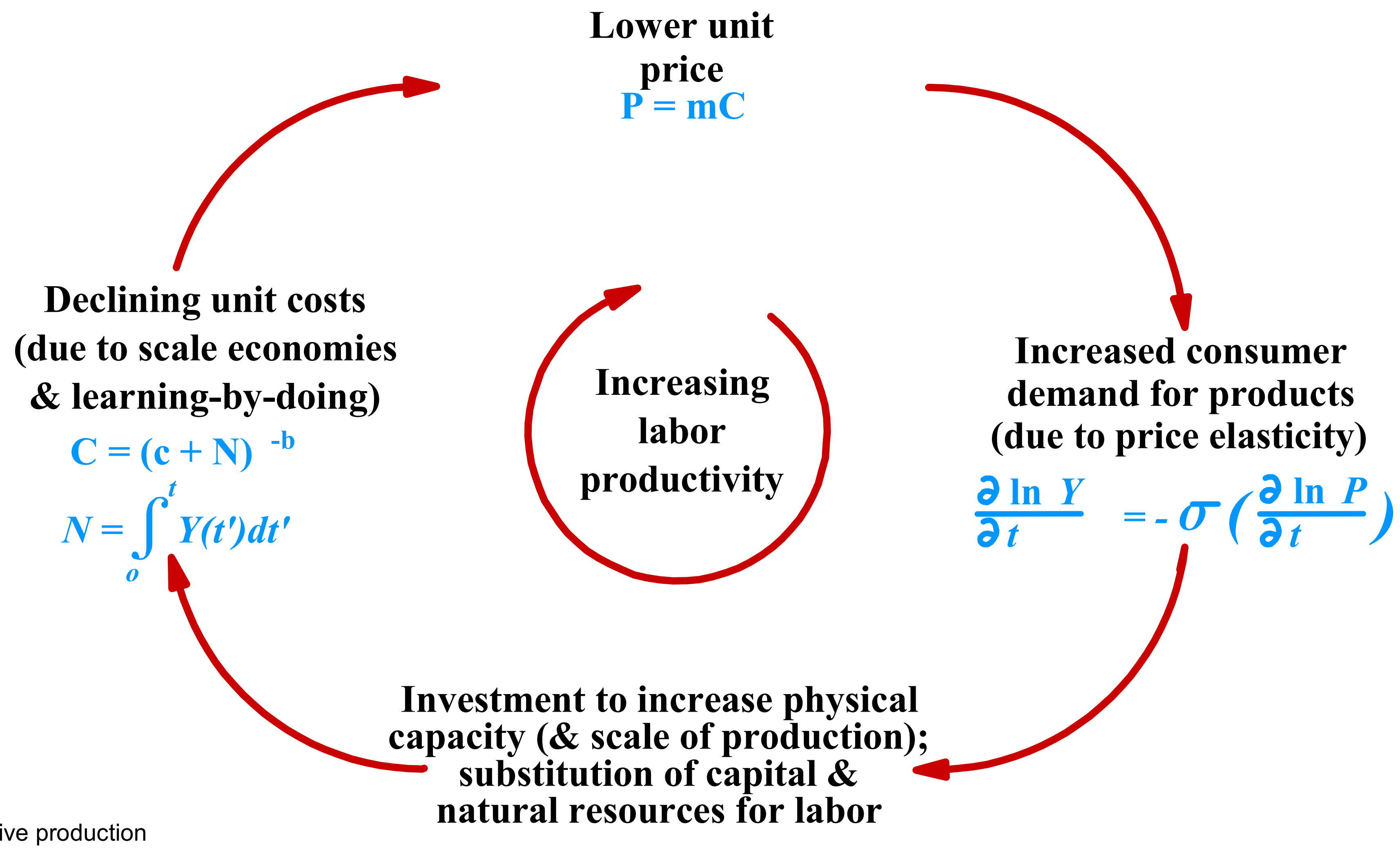
von Neumann, J. (1945). "A model of general economic equilibrium." Review of Economic Studies **13**: 1-9.

Wall, G. (1977). Exergy: A useful concept within resource accounting. Goteborg, Sweden, Institute of Theoretical Physics, Chalmers University of Technology and University of Goteborg.

Wall, G. (1986). "Exergy conversion in the Swedish society." Energy **11**: 435-444.

Walras, L. (1954). Elements of pure economics or the theory of social wealth. London, George Allen and Unwin.

Figure 1: Salter cycle growth engine



Lower unit price

$P = mC$

Increased consumer demand for products (due to price elasticity)

$$\frac{\partial \ln Y}{\partial t} = -\sigma\left(\frac{\partial \ln P}{\partial t}\right)$$

Investment to increase physical capacity (& scale of production); substitution of capital & natural resources for labor

Declining unit costs (due to scale economies & learning-by-doing)

$$C = (c + N)^{-b}$$

$$N = \int_o^t Y(t')dt'$$

Increasing labor productivity

C = cost
N = cumulative production
P = price
Y= economic output
sigma = price elasticity of demand
c,m parameters

Figure 2. Exergy conversion efficiency f, for two definitions of work and exergy, US 1900-1998

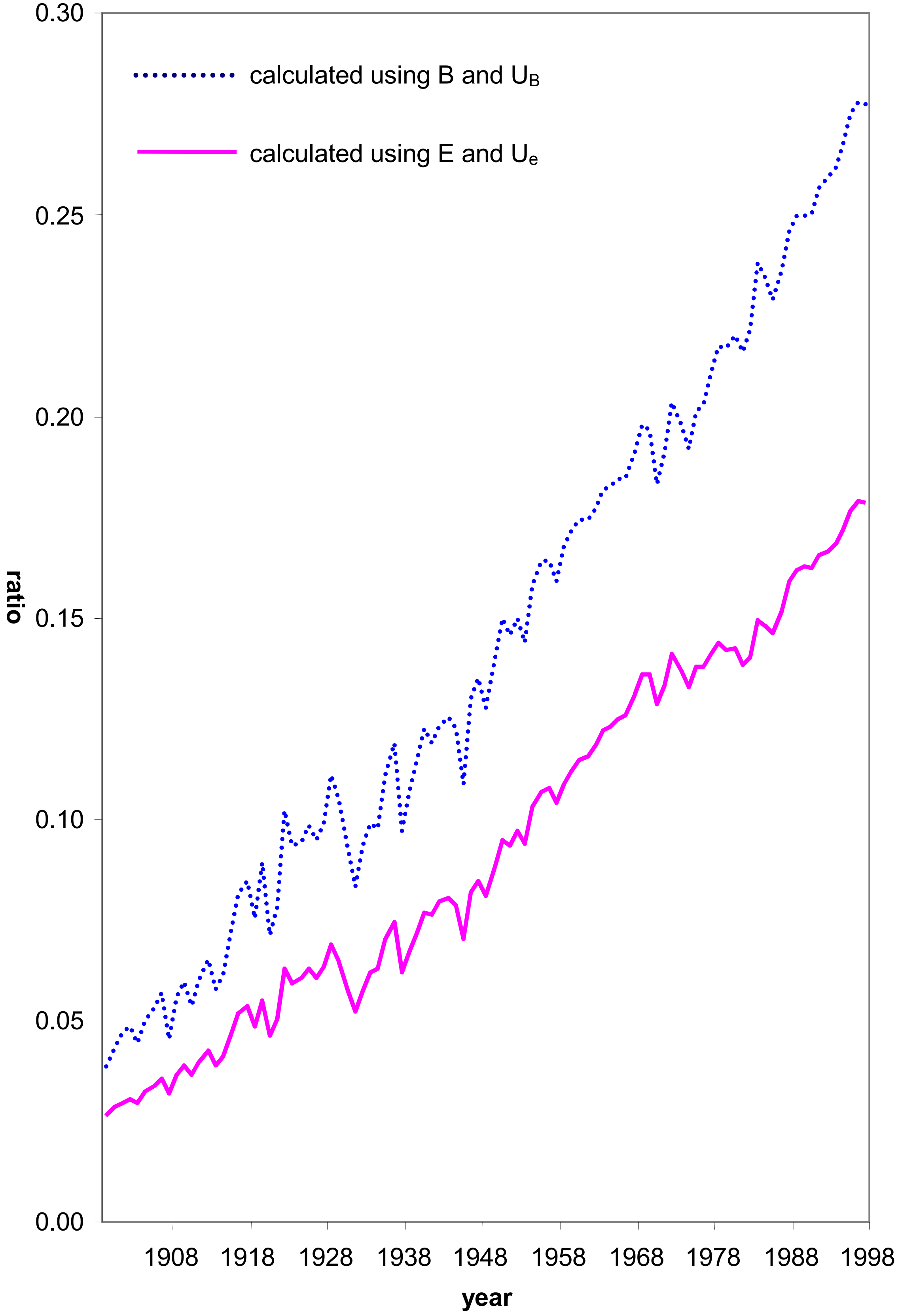


Figure 3: Coal consumption; Exergy allocation among types of work, USA 1900-1998

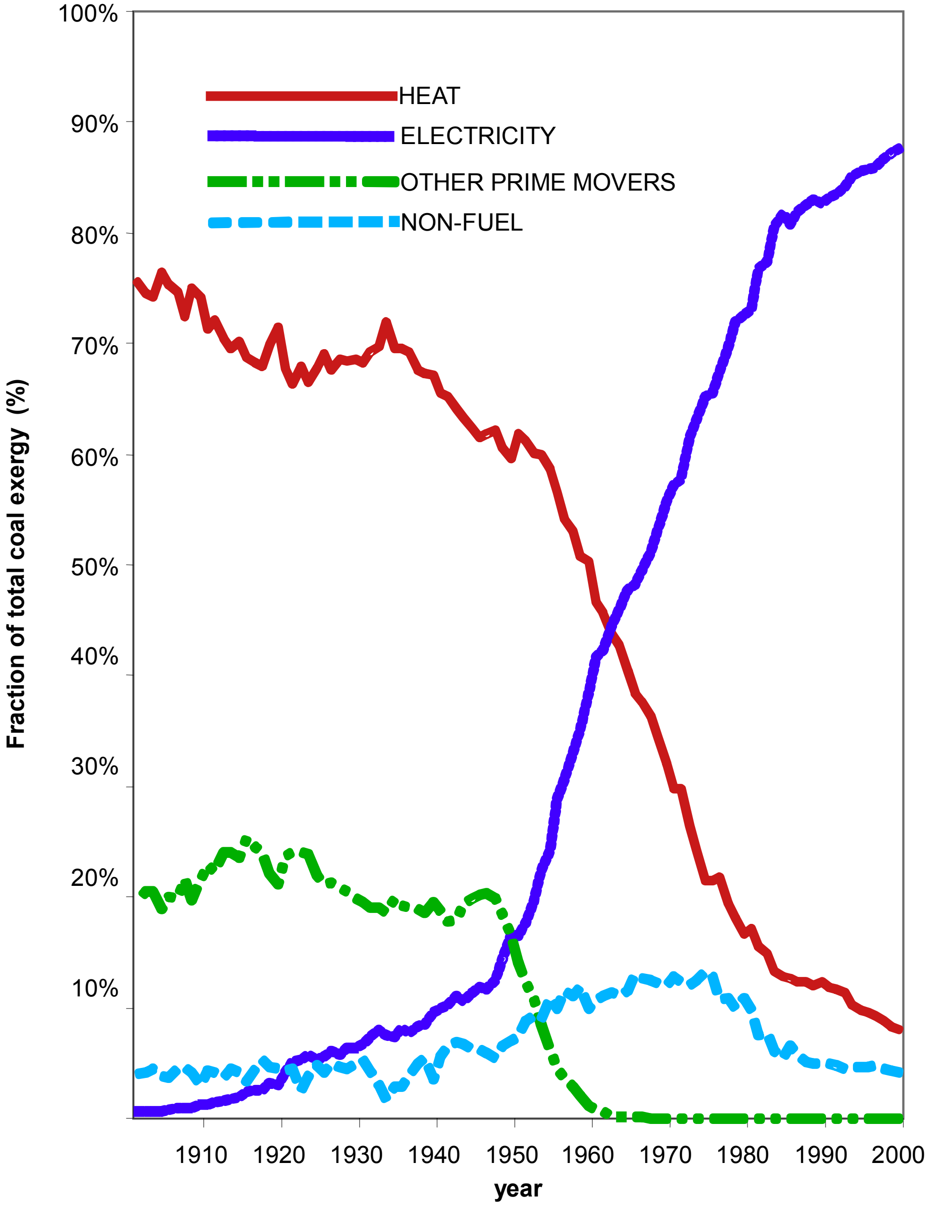


Figure 4: Petroleum consumption; Exergy allocation among types of work, USA 1900-1998

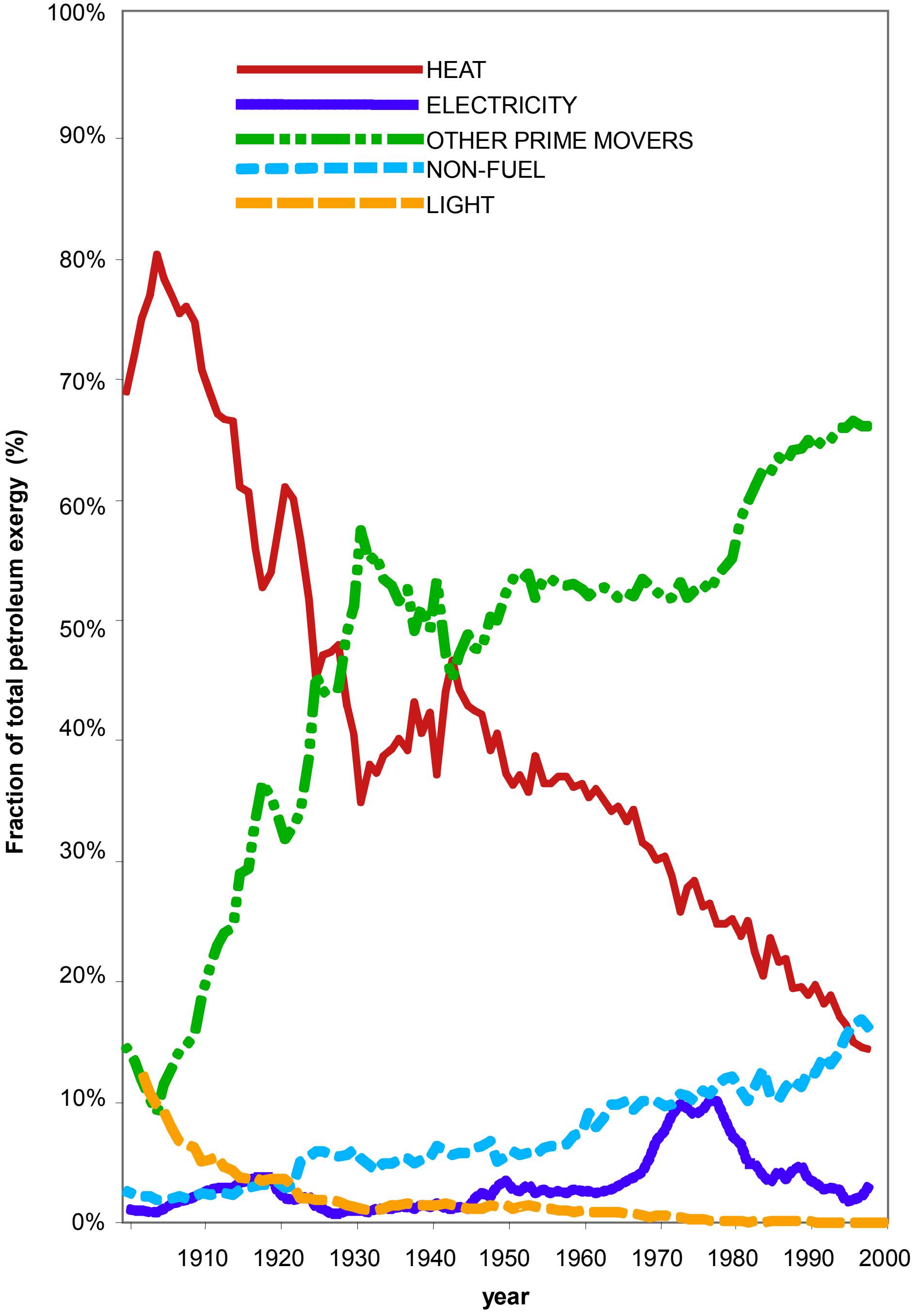


Figure 5: Gas consumption; Exergy allocation among types of work, USA 1900-1998

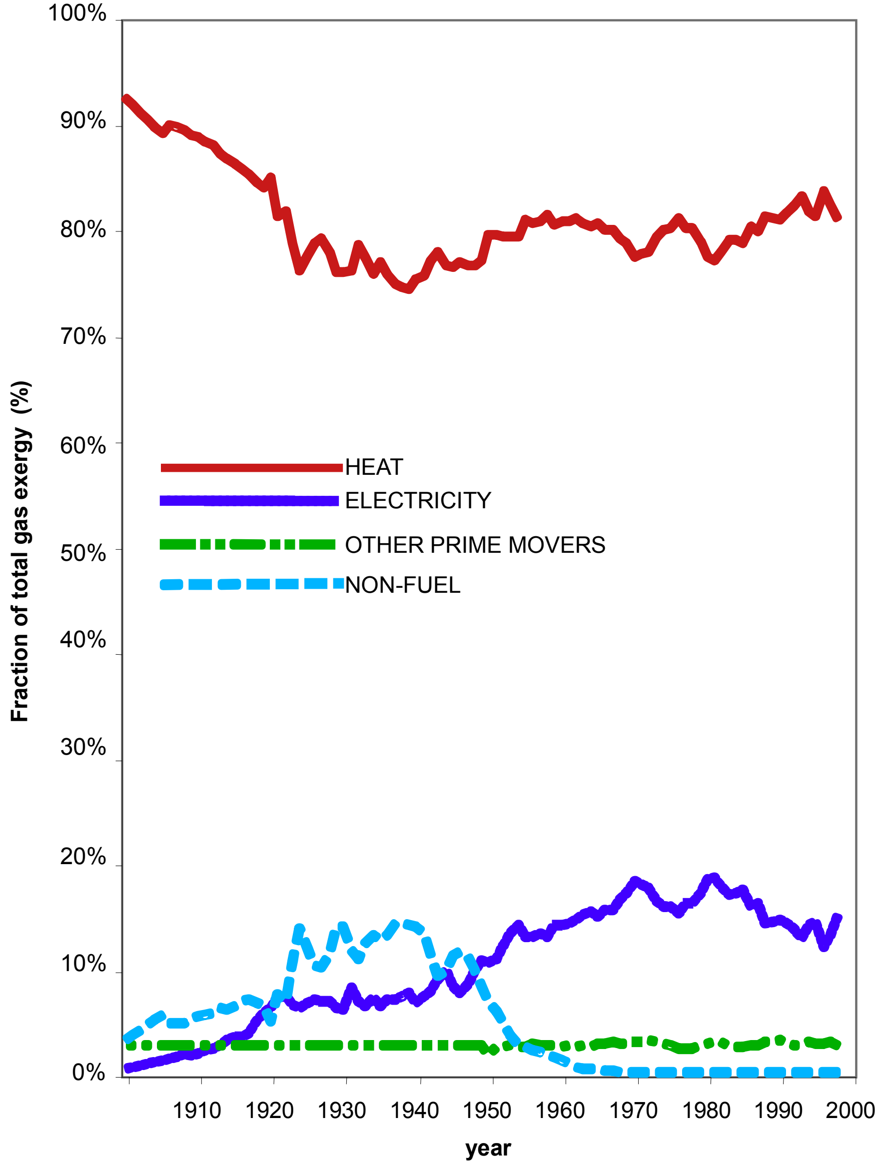


Figure 6: Fossil fuel consumption; Exergy allocation among types of work, USA 1900-1998

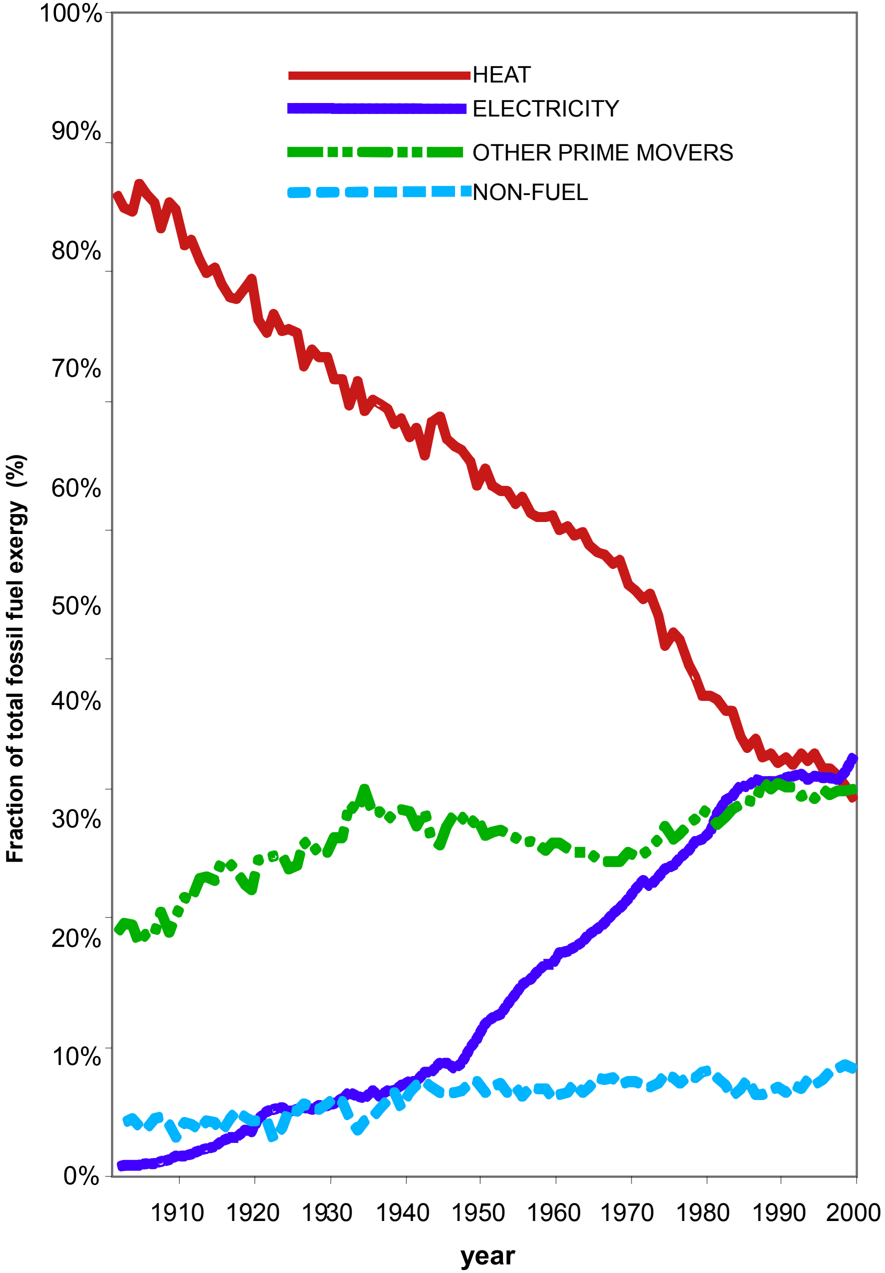
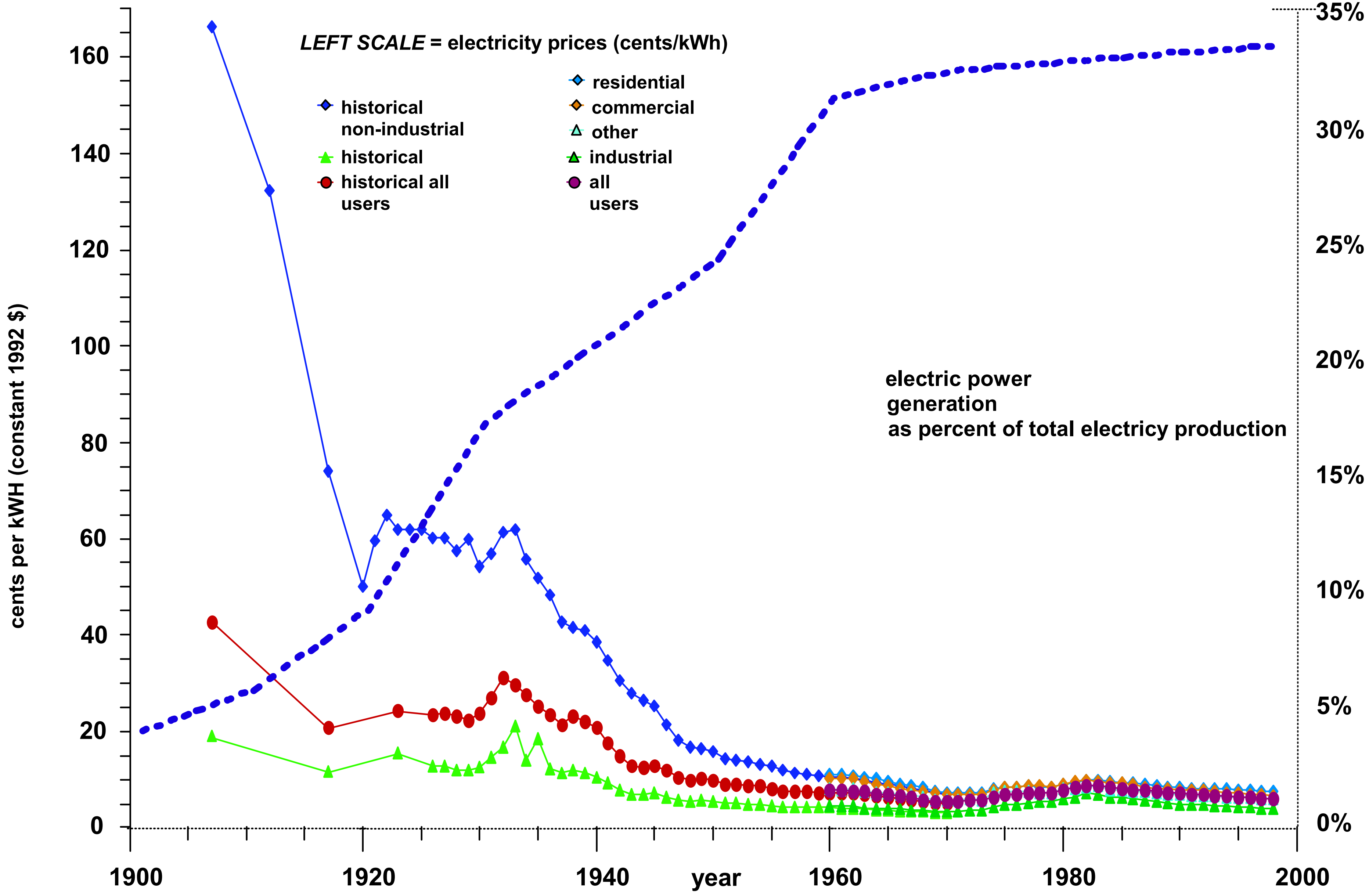


Figure 7: Electricity; conversion efficiency and retail price in the USA 1900-1998



Sources: 1960-1998: Annual Energy Review Table 8.13 1907-1970: Historical Statistics: Vol 1, Series S116,118, 119

Figure 8: Energy (exergy) conversion efficiencies, USA 1900-1998

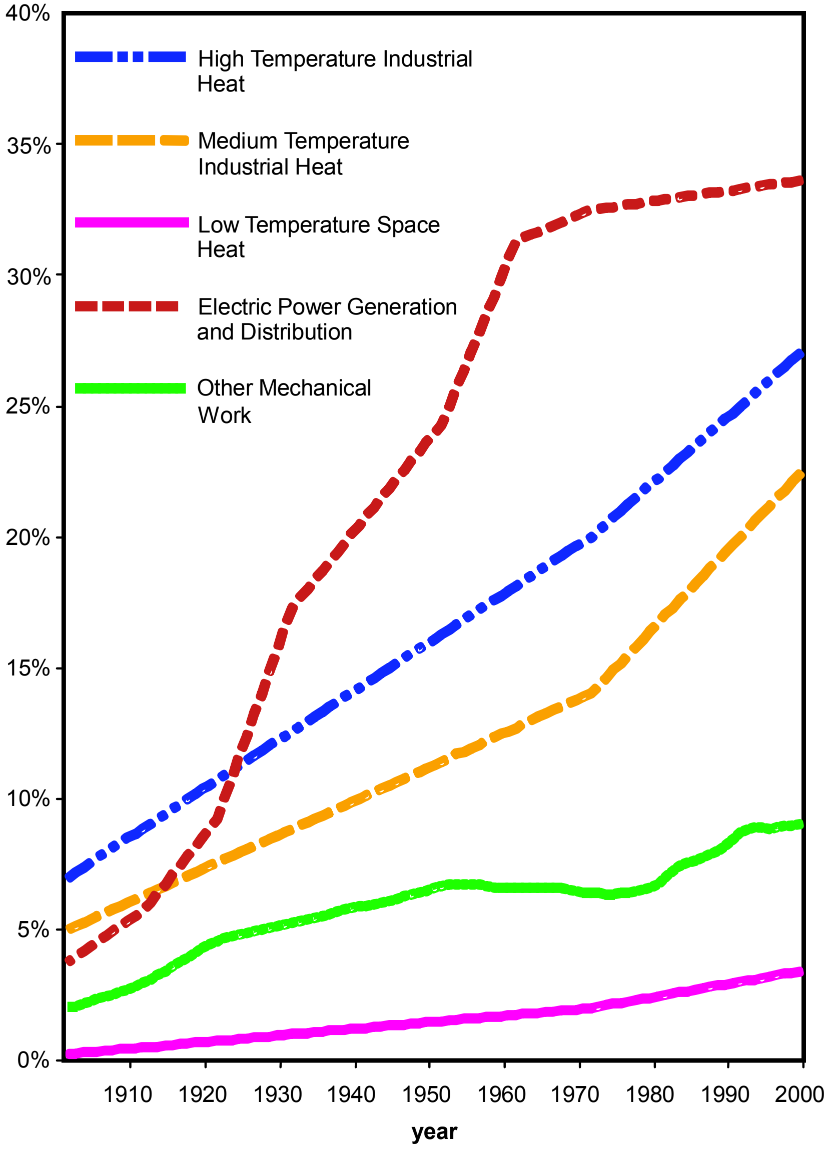


Figure 9. Useful work and the useful work GDP ratio, USA 1900-1998

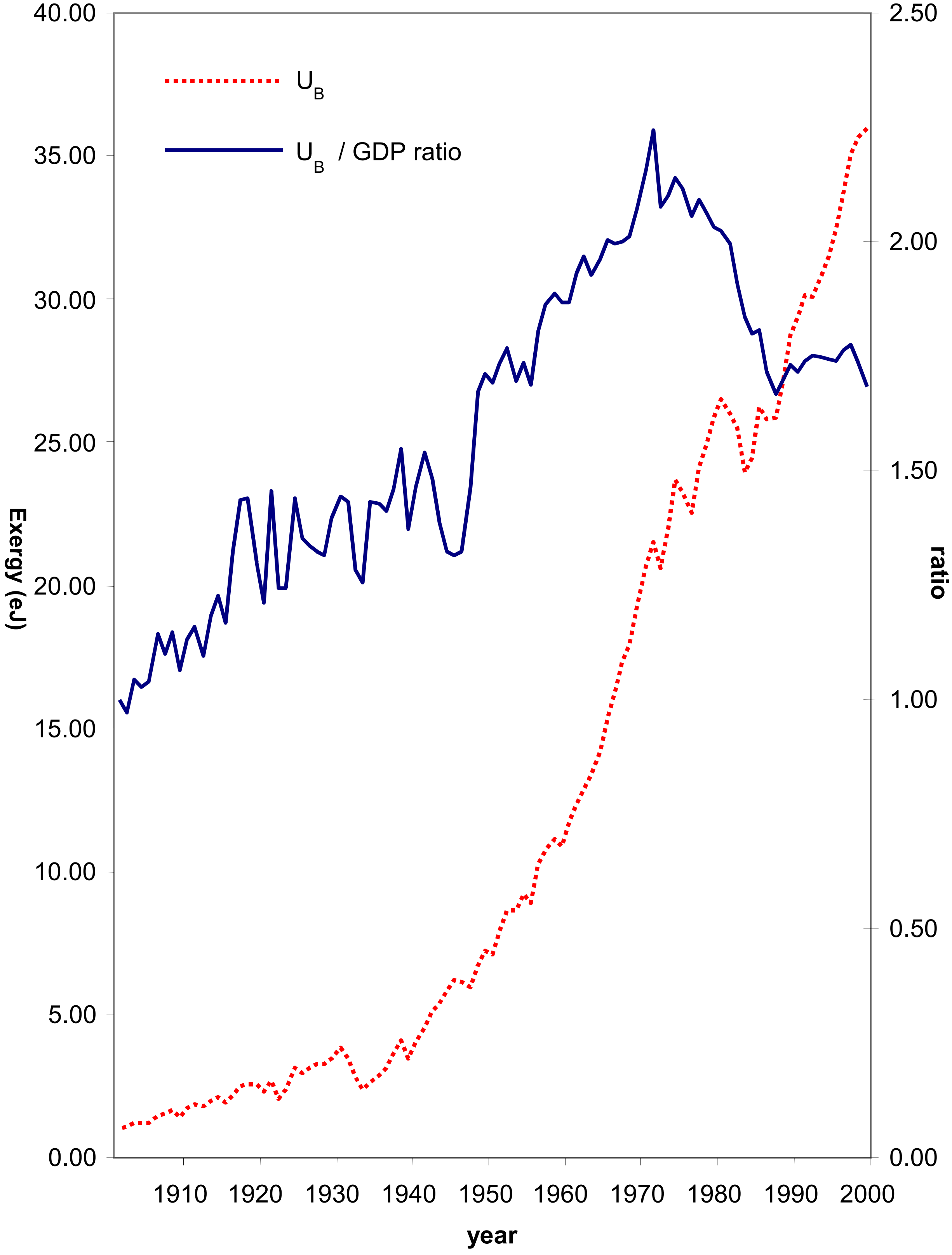


Figure 10a. Output and factors of production for two alternative definitions of natural resource inputs, USA 1900-2000

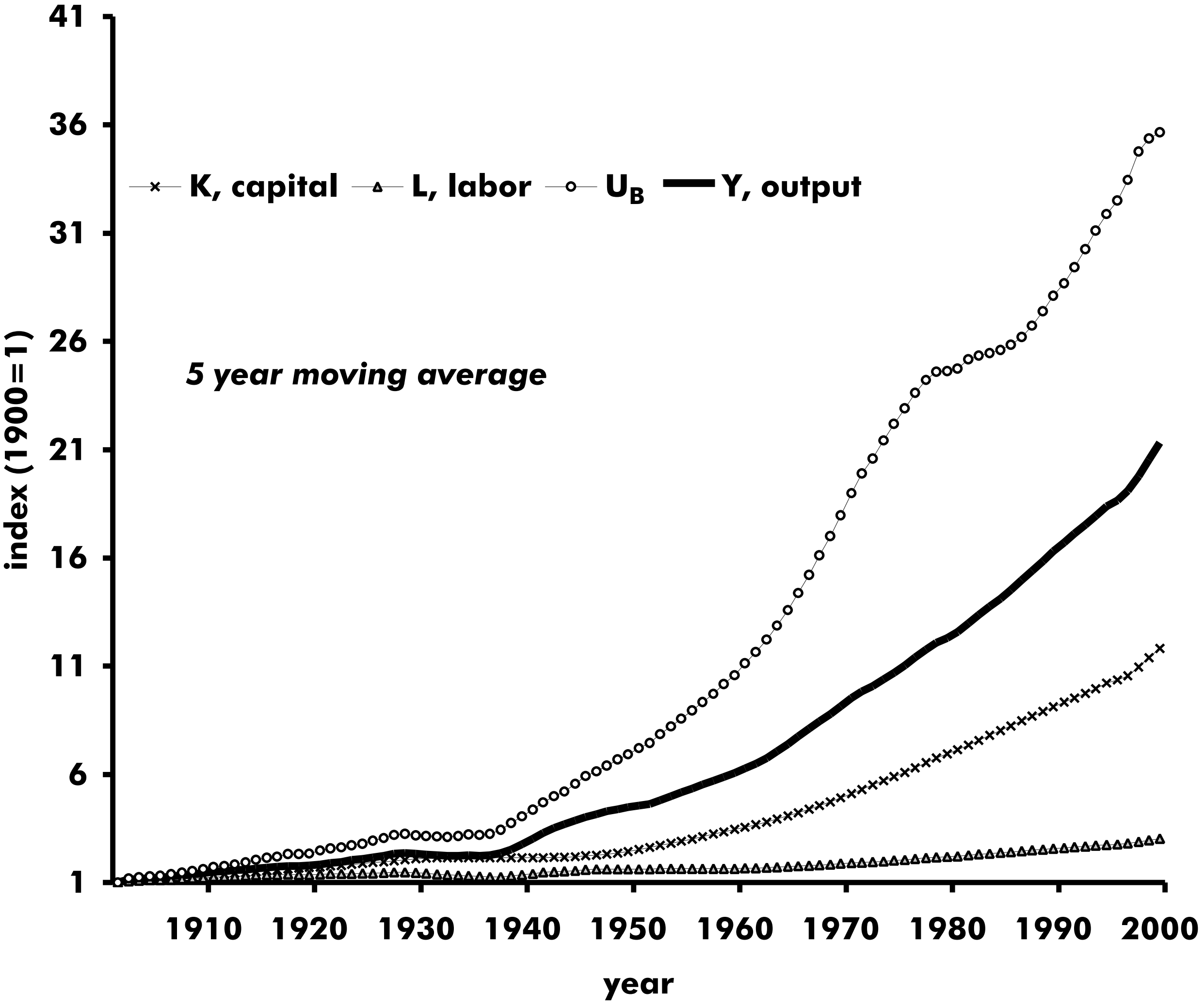


Figure 10b. Output and factors of production: log transformed variables, USA 1900-2000

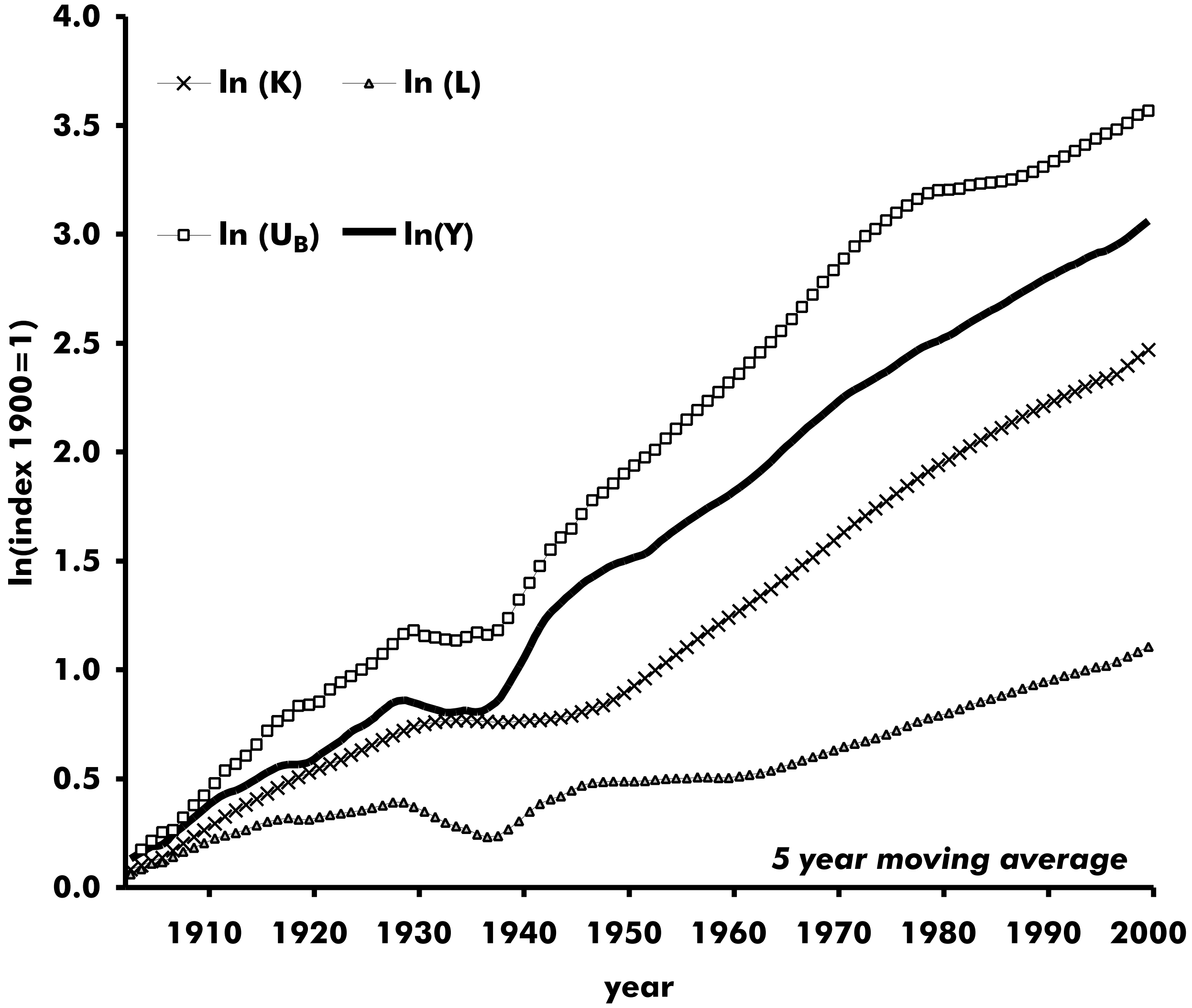
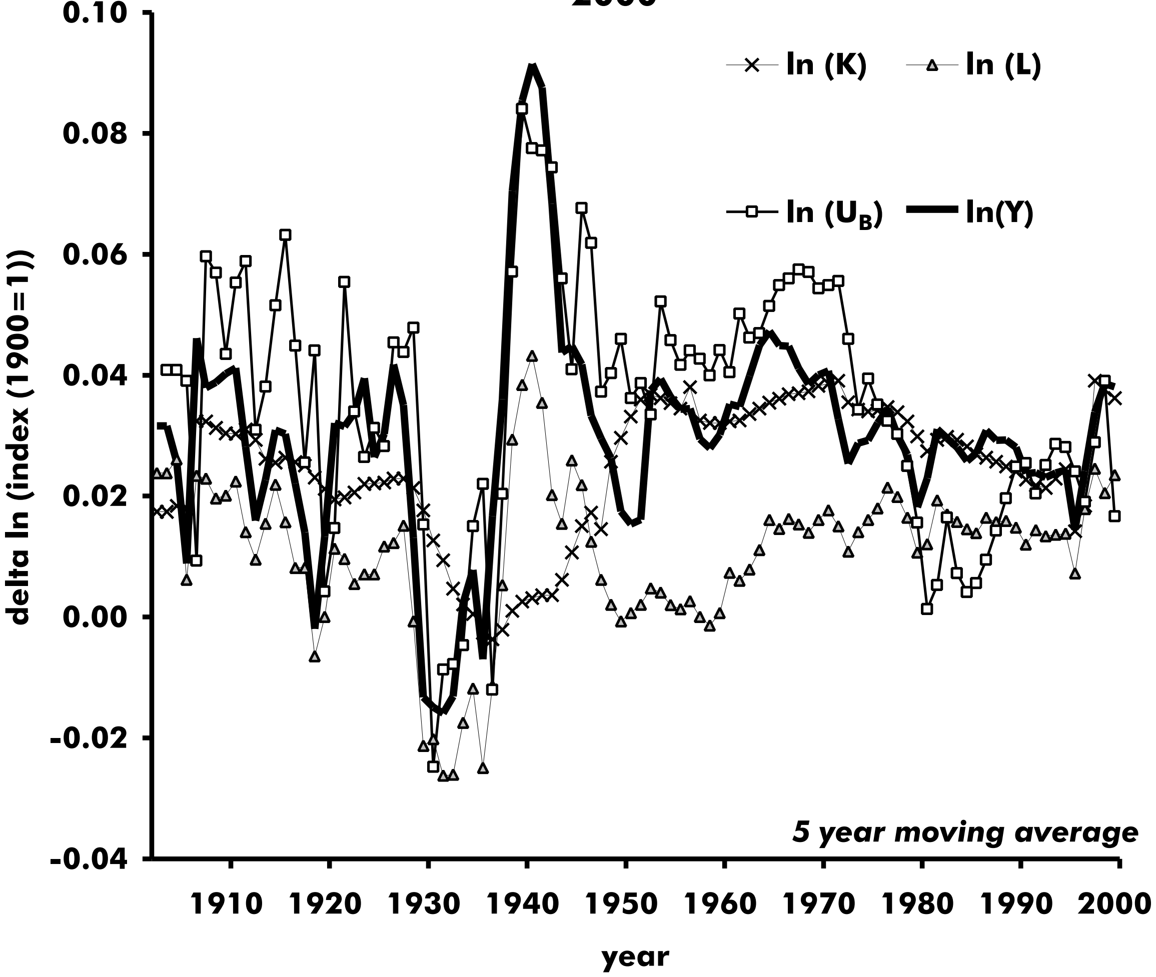


Figure 10c. Output and factors of production: increments of log transformed variables, USA 1900-2000



5 year moving average

Figure 10d. Output and factors of production for log-linear LINEX function, USA 1900-2000

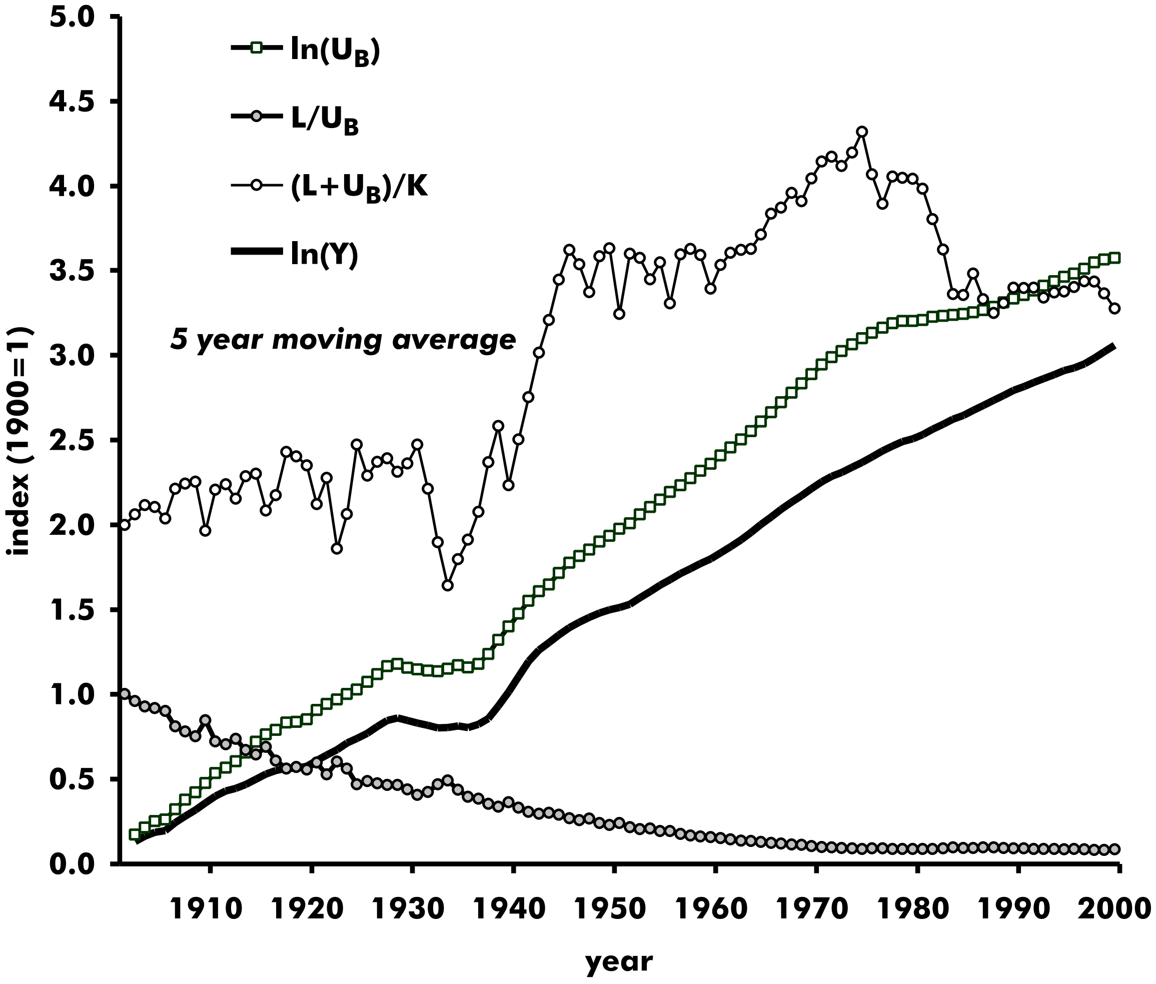


Figure 11. Predicted rate of change of output growth delta $\ln(Y/Y_0)$, USA 1900-2000.

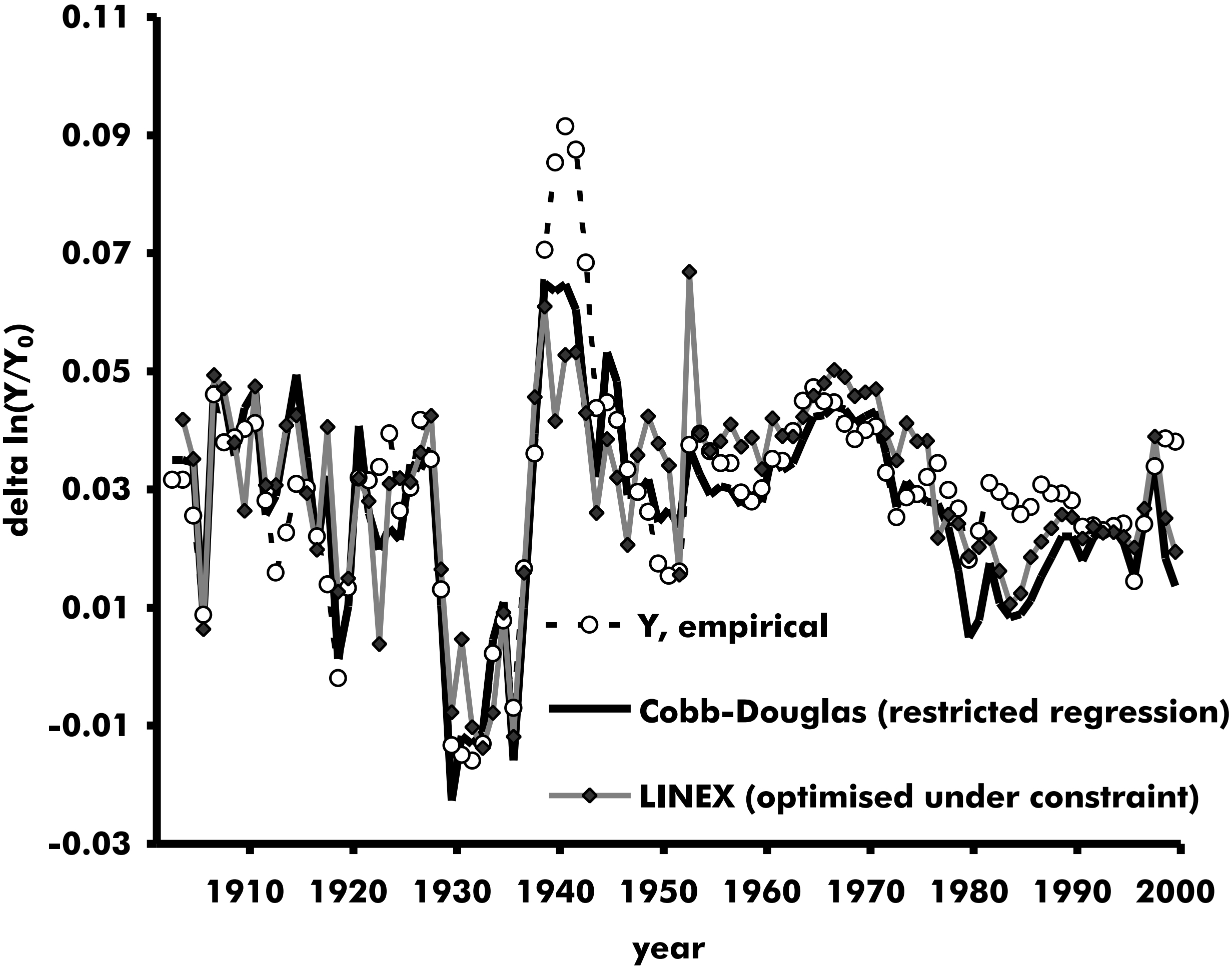


Figure 12. Predicted output (Y/Y_0) from alternative methods (assuming constant returns to scale), US 1900-2000.

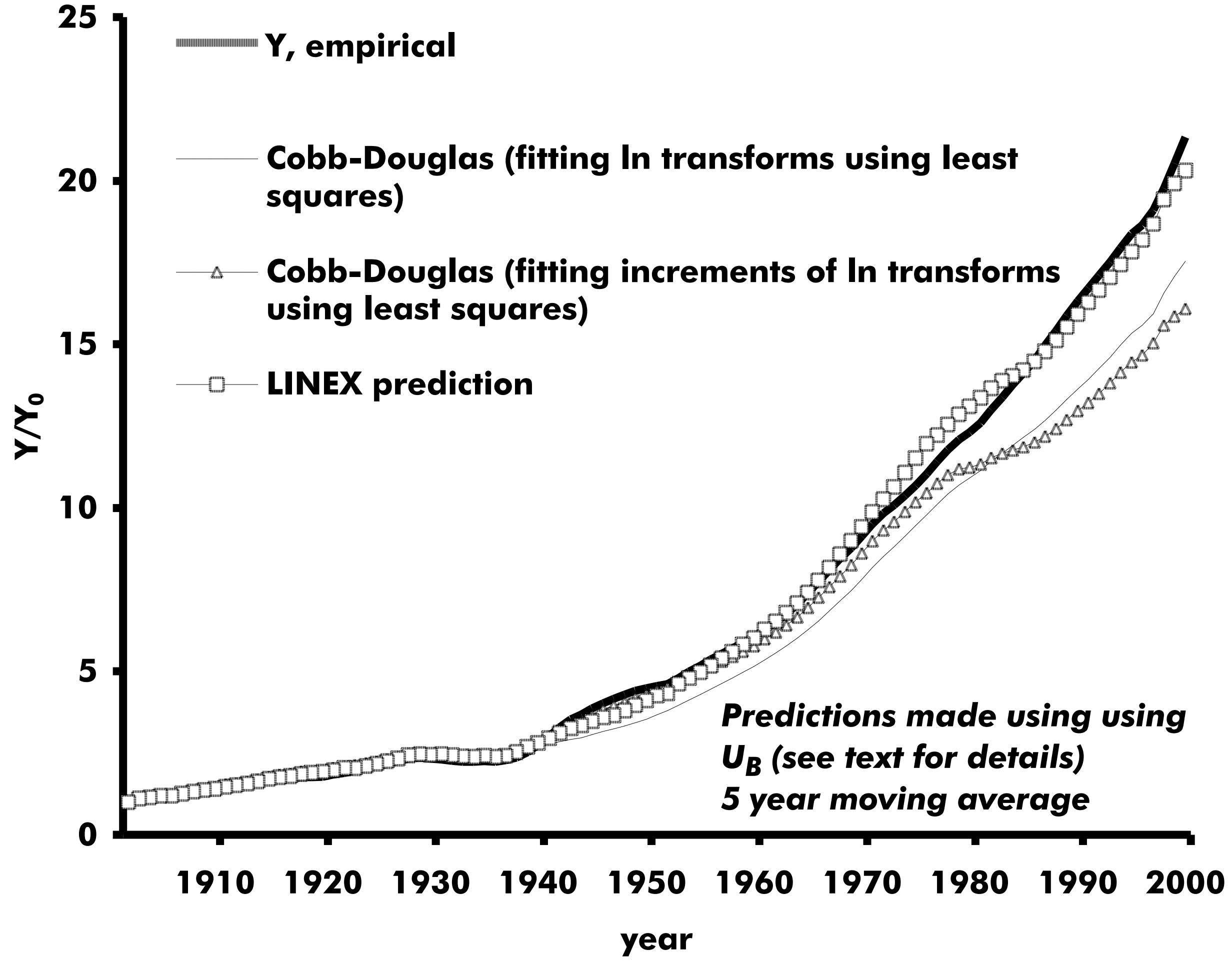
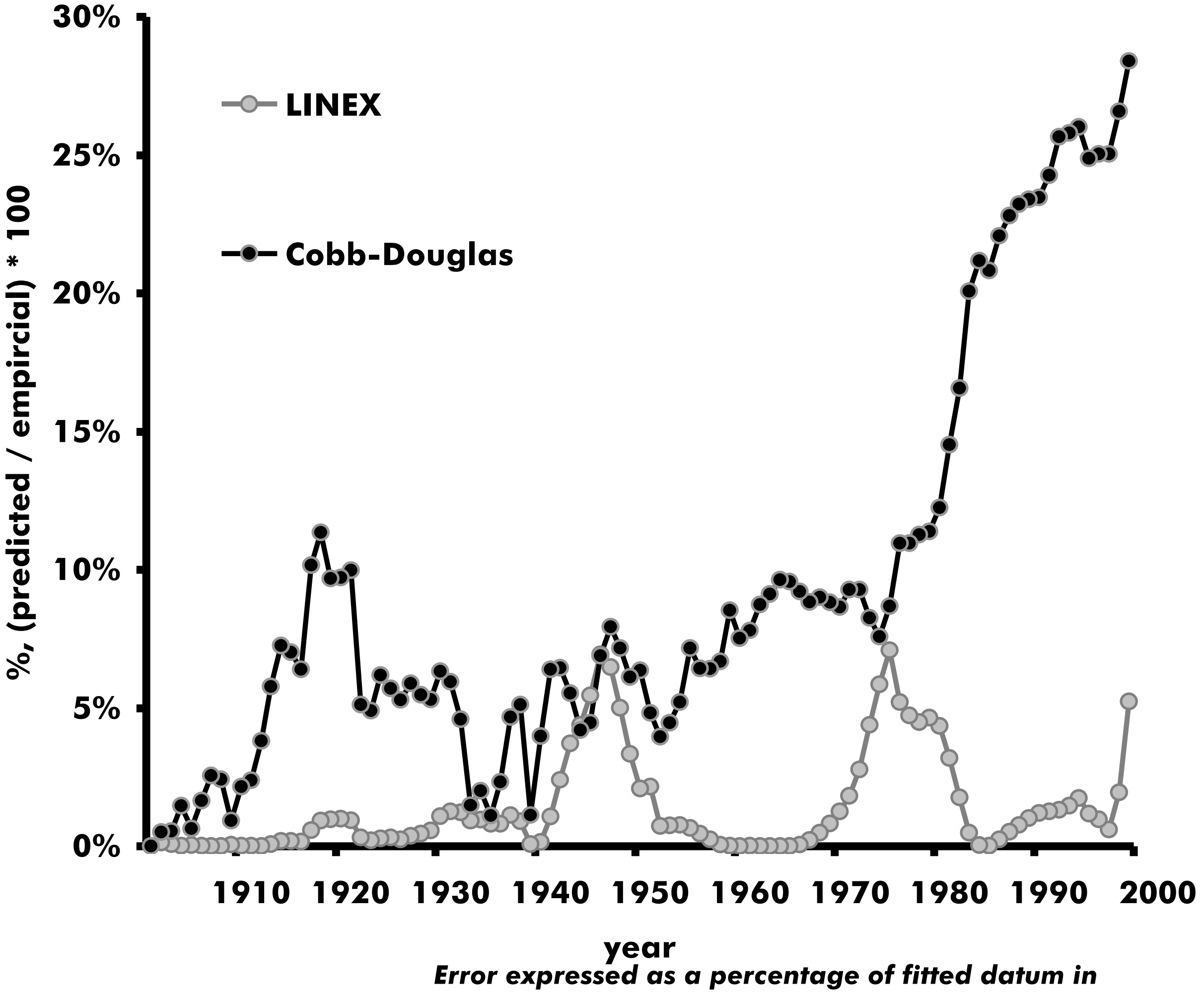


Figure 13. Residual from alternative predictions of output measures, US 1900-2000



Error expressed as a percentage of fitted datum in each year

Figure 14. Marginal productivities from the LINEX function, USA 1900-2000.