The Need to Reintegrate the Natural Sciences with Economics

CHARLES HALL, DIETMAR LINDENBERGER, REINER KÜMMEL, TIMM KROEGER, AND WOLFGANG EICHHORN

ow long will researchers working in adjoining fields...abstain from expressing serious concern about the splendid isolation in which academic economics now finds itself?" Wassily Leontief, Nobel laureate in economics, asked almost two decades ago (Leontief 1982, p. 104). The question is extremely important, because economics is the foundation on which most decisions affecting agriculture, fisheries, the environment, and, indeed, most aspects of our daily lives are based. Natural scientists, including biological scientists, may have particular views on this or that economic policy, but few question the legitimacy of economics as a tool. We believe that, paraphrasing the great Prussian military historian Carl von Clausewitz, economics is too important to leave to the economists, and that natural scientists should not leave the procedures by which we undertake economics up to economists alone. Instead, natural scientists must contribute to a new discourse about the means, methods, and ends of economics.

"

This article is a response to Leontief's question. It is essential that economics be based on sound principles and that the policies generated from it have a solid foundation. Neoclassical economics, the form of economics derived in the mid-19th century that prevails today, focuses on problems related to value decisions, the behavior of economic actors, and the working of markets. These problems belong to the sphere of the social sciences (many of which, incidentally, have their own problems with neoclassical economic theory; see, for example, Marris 1992). But the wealth that is distributed in the markets must be produced in the hard sphere of the material world where all operations must obey the laws and principles of physics, chemistry, and biology. Our concern is that most NEOCLASSICAL ECONOMICS, THE DOMI-NANT FORM OF ECONOMICS TODAY, HAS AT LEAST THREE FUNDAMENTAL FLAWS FROM THE PERSPECTIVE OF THE NATURAL SCIENCES, BUT IT IS POSSIBLE TO DEVELOP A DIFFERENT, BIOPHYSICAL BASIS FOR ECONOMICS THAT CAN SERVE AS A SUP-PLEMENT TO, OR A REPLACEMENT FOR, NEOCLASSICAL ECONOMICS

production models of economics are not based on these biophysical laws and principles; in fact, they tend to ignore them (Georgescu-Roegen 1971, Daly 1973, 1977, Kümmel et al. 1985, Leontief 1982, Cleveland et al. 1984, Hall et al. 1986, Hall 1992, 2000).

This disregard of the biophysical aspects of production by economists was not the rule historically. Quesnay and other members of the 18th-century French physiocrat school focused on the use of solar radiation by biotic organisms and the role of land in generating wealth by capturing this energy through agricultural production. The classical economics of

Charles Hall (e-mail: chall@esf.edu) is a professor in the department of Environmental and Forest Biology and in the Program of Environmental Studies, College of Environmental Science and Forestry, State University of New York, Syracuse, NY 13210; also at SUNY, Timm Kroeger is a graduate student in the Graduate Program of Environmental Sciences at the College of Environmental Science and Forestry. Dietmar Lindenberger is senior researcher at the Institute of Energy Economics, University of Cologne, Albertus-Magnus-Platz, D-50923 Cologne, Germany. Reiner Kümmel is a professor of physics at the Institute of Theoretical Physics, University of Würzburg, D-97074 Würzburg, Germany. Wolfgang Eichhorn is a professor of economics at the Institute of Economic Theory and Operations Research, University of Karlsruhe, D-76128 Karlsruhe, Germany. © American Institute of Biological Sciences. Adam Smith, David Ricardo, and Karl Marx encompassed both the physical origin and the distribution of wealth (Ricardo 1891, Marx 1906, Smith 1937). Podolinsky, Geddes, Soddy, and Hogben were biological and physical scientists of the 19th and early 20th centuries who thought deeply about economic issues (Martinez-Alier 1987, Christensen 1989, Cleveland and Ruth 1997). Thus we find the degree to which neoclassical economics has displaced classical economics curious, almost a historical accident. The primary reason for this displacement was the superior mathematical rigor of neoclassical economics and the development of the marginal utility theory, which solved the "water vs. diamonds" paradox that classical economics could not. But the underlying biophysical perspective of Smith and Ricardo was not incorporated into the new mathematical elegance of the "marginal revolution."

Consequently, major decisions that affect millions of people and most of the world's ecosystems are based on neoclassical economic models that, although internally consistent and mathematically sophisticated, ignore or are not sufficiently consistent with the basic laws of nature. This leads to the failure of those economic policies that run counter to these laws and endanger sustainable development. In this article we examine this issue in more detail, making a case for including the laws of nature in economic theory and analysis, and in the policies derived from this theory, as carefully and explicitly as the assumptions on human preferences and choices. Both natural scientists and even many economists have been leveling severe criticisms at the basis of neoclassical economics for many years (Soddy 1926, Boulding 1966, Georgescu-Roegen 1966, 1971, Daly 1973, Binswanger and Ledergerber 1974, Cleveland et al. 1984, Hall et al. 1986, Ayres 1996, 1999). These criticisms, however, are largely ignored by neoclassical economists, and the rest of the scientific community seems to be largely unaware of them. We believe that it is time to exhume these criticisms and add to them more recent analytic work that gives them even greater validity.

Past criticisms of neoclassical economics from the perspective of natural scientists can be summarized as three fundamental arguments: (1) The structure of the basic conceptual neoclassical model is unrealistic because it is not based on the biophysical world and the laws governing it, especially thermodynamics (Figure 1a); (2) the boundaries of analysis are inappropriate because they do not include the real processes of the biosphere that provide the material and energy inputs, the waste sinks, and the necessary milieu for the economic process (Figure 2); and (3) the basic assumptions underlying the models used have not been put forth as testable hypotheses but rather as givens.

We substantiate these three criticisms below and present a new model of industrial production that we believe further supports our criticisms and our assessment of the importance of energy. In this new model, the output of the economic system and the maintenance of its components depend upon continuous input of energy into the system, as is true for all organisms and ecosystems.

Critique of neoclassical economics

"Anything as important in industrial life as power deserves more attention than it has yet received from economists.... A theory of production that will really explain how wealth is produced must analyze the contribution of the element energy" (Tryon 1927). "The decisive mistake of traditional economics...is the disregard of energy as a factor of production" (Binswanger and Ledergerber 1974).

Argument 1: Thermodynamics. Contemporary economics pays only marginal attention to the first and second laws of thermodynamics. This is a serious conceptual flaw and an obstacle to the design of economic policies that can successfully meet the challenges of pollution, resource scarcity, and unemployment. The two laws say that nothing happens in the world without energy conversion and entropy production, with the consequence that every process of biotic and industrial production requires the input of energy. Because

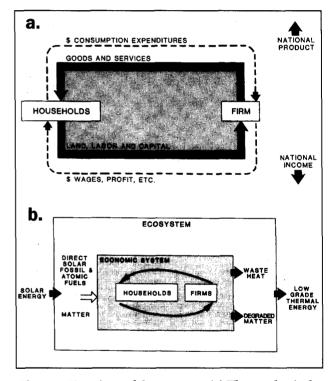


Figure 1. Two views of the economy. (a) The neoclassical view of how economies work. Households sell or rent land, natural resources, labor, and capital to firms in exchange for rent, wages, and profit (factor payments). Firms combine the factors of production and produce goods and services in return for consumption expenditures, investment, government expenditures, and net exports. This view represents, essentially, a perpetual motion machine. (b) Our perspective, based on a biophysical viewpoint, of the minimum changes required to make Figure 1a conform to reality. We have added the basic energy and material inputs and outputs that are essential if the economic processes represented in Figure 1a are to take place (redrawn from Daly 1977).

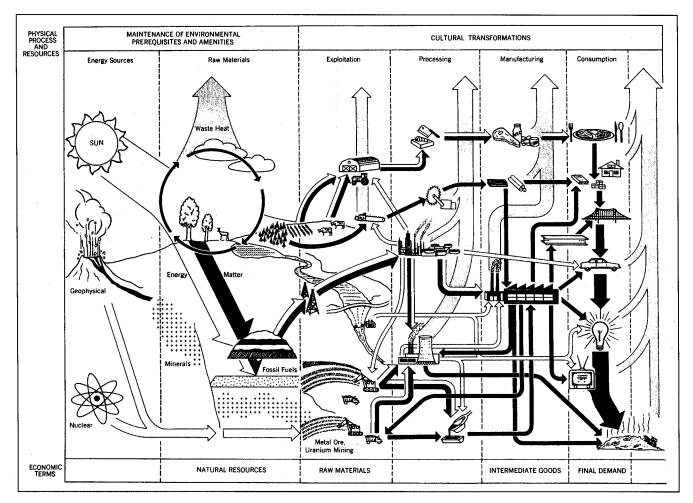


Figure 2. A more comprehensive and accurate model of how economies actually work. The second column of this diagram represents the entire global ecosystem milieu within which the rest of the global economy operates. Natural energies drive geological, biological, and chemical cycles that produce natural resources and public service functions and maintain the milieu essential for all other economic steps. Extractive sectors use economic energies to exploit natural resources and convert them to raw materials. Raw materials are used by manufacturing and other intermediate sectors to produce final goods and services. These final goods and services are distributed by the commercial sector to final demand. Eventually, nonrecycled materials and waste heat return to the environment as waste products. We believe this diagram to be the minimum model of how a real economy works.

of unavoidable entropy production, the valuable part of energy (*exergy*) is transformed into useless heat at the temperature of the environment (*anergy*), and usually matter is dissipated, too. This results in pollution and, eventually, exhaustion of the higher grade resources of fossil fuels and raw materials. Human labor, living on food, has been and continues to be replaced (at least in part) by energy-driven machines in the routine production of goods and services.

Although the first and second laws of thermodynamics are the most thoroughly tested and validated laws of nature, and they state explicitly that it is impossible to have a perpetual motion machine—that is, a machine that performs work without the input of exergy—the basic neoclassical economic model *is* a perpetual motion machine, with no required inputs or limits (Figure 1a). Most economists have accepted that incomplete model as the basis for their analyses and have relegated energy and other resources to unimportance in those analyses (e.g., Denison 1979, 1984). This attitude was fixed in the minds of most economists by the analysis of Barnett and Morse (1963), who found no indication of increasing scarcity of raw materials, as determined by their inflation-corrected price, for the first half of the 20th century (Smith 1989).

Their analysis, although cited by nearly all economists interested in the depletion issue, was nonetheless seriously incomplete. Cleveland showed that the only reason that decreasing concentrations and qualities of resources were not translated into higher prices for constant quality was because of the decreasing price of energy and its increasing use in the exploitation of increasingly lower-grade reserves in the United States and elsewhere (Cleveland 1991). Thus, although economists have argued that natural resources are not important to the economy, the truth is that it is only because of the abundant availability of many natural resources that economics can assign them low monetary value, despite their critical importance to economic production.

The perspective of Nobel laureate in economics Robert M. Solow is interesting. In 1974 he considered the possibility that "the world can, in effect, get along without natural resources," because of the technological options for the substitution of other factors for nonrenewable resources; he noted, however, that if "real output per unit of resources is effectively bounded-cannot exceed some upper limit of productivity which in turn is not too far from where we are now---then catastrophe is unavoidable" (Solow 1974, p. 11). More recently, Solow stated, "It is of the essence that production cannot take place without some use of natural resources" (Solow 1992, 1993). Clearly there is need for more analytical and empirical work on the relation between production and natural resources, especially energy but also all aspects of the supportive contributions of the biosphere. We believe that the attempt to simply put a monetary value on these services, although useful in some respects, is insufficient to resolve the issue, if only because such values are based on human perceptions that, in turn, are developed on the basis of imperfect information and-all too often-myopia.

Why does neoclassical economics assign a low value to natural resources?

The conventional neoclassical view of the low importance of energy and materials dates back to the first stages in the development of neoclassical economics. Initially the focus was not so much on the generation of wealth as on its distribution and the "efficiency of markets." Consequently, the early thinkers in economics started with a model of pure exchange of goods, without considering their production. With a set of mathematical assumptions on "rational consumer behavior," it was shown that, through the exchange of goods in markets, an equilibrium results in which all consumers maximize their utility, in the sense that it is not possible to improve the situation of a single consumer without worsening the situation of at least one other consumer (the so-called Pareto optimum). This benefit of (perfect) markets is generally considered as the foundation of free-market economics. It shows why markets, where "greedy" individuals meet, work at all. But later, when the model was extended to include production, the problem of the physical generation of wealth was coupled, inseparably, to the problem of the distribution of wealth as a consequence of the model structure: In the neoclassical equilibrium, with the assumption of profit-maximizing entrepreneurial behavior, factor productivities by definition had to equal factor prices. This means that in the resulting model the weights with which the production factors contribute to the physical generation of wealth are determined by the cost share of each factor. In other words, observations on contemporary social structure and entrepreneurial behaviors are used to draw inferences concerning the physical importance of production factors.

Here lies the historical source of the economists' underestimation of the production factor energy, because in advanced industrial market economies the cost of energy, on the average, is only 5% to 6% of the total factor cost (Baron 1997). Therefore, economists tend to either neglect energy as a factor of production altogether or they argue that the contribution of a change of energy input to the change of output is equal only to energy's small cost share of 5% to 6% (Denison 1979, 1984). However, it can be argued that energy has a small share in total production costs not because it is relatively less important than capital or labor as a production factor, but rather because of the free work of the biosphere and the geosphere, it has been abundant and cheap; moreover, not all costs of its use are reflected in its market price (i.e., the problem of "externalities"). That energy actually has much more leverage was demonstrated by the impact of the two energy price explosions in the periods 1973-1975 and 1979-1981, which had significant impacts on economic growth (Cleveland et al. 1984, Jorgenson 1984, 1988).

Neoclassical models that do not include energy cannot explain the empirically observed growth of output by the growth of the factor inputs labor and capital. There always remains a large unexplained growth residual that formally is attributed to what economists call "technological progress." "This...has led to a criticism of the neoclassical model: it is a theory of growth that leaves the main factor in economic growth unexplained" (Solow 1994). As we argue below, weighting a factor by its cost share is an incorrect approach in growth theory. Likewise, the finite emission-absorption capacity of the biosphere is vastly more important to future economic production than its present (often zero) price indicates.

The human economy uses fossil and other fuels to support and empower labor and to produce and utilize capital, just as organisms and ecosystems use solar-derived energy to produce and maintain biomass and biotic functions. Labor productivity has been correlated highly with increasing energy use per worker. This has been especially critical in agriculture (Hall et al. 1986). Energy, capital, and labor are combined in human economies to upgrade natural resources (generated by natural energy flows) to useful goods and services. Therefore economic production, like biotic production, can be viewed as the process of upgrading matter into highly ordered (thermodynamically improbable) structures, both physical structures and information. Where one speaks of "adding value" at successive stages of production, one may also speak of "adding order" to matter through the use of free energy (exergy). The perspective of examining economics in the hard sphere of physical production, where energy and material stocks and flows are important, is called biophysical economics. It *must* complement the social sphere perspective.

Argument 2: Boundaries. Another problem with the basic model used in neoclassical economics (Figure 1a) is that it does not include boundaries that in any way indicate the physical requirements or effects of economic activities. We believe that at a minimum Figure 1a should be reconstructed as Figure 1b, to include the necessary resources, the generation of wastes, and the necessity for the economic process to occur within the larger system, the biosphere (Daly 1977, Cleveland et al. 1984, Dung 1992, Ayres 1996, Dasgupta et al. 2000). Taking this assessment one step further, we believe that something like Figure 2 is the diagram that should be used to represent the actual physical aspects of an economy's working. It shows the necessity of the biosphere for the first steps of economic production and as a milieu for all subsequent steps. Figure 2 further emphasizes the flow of energy and matter across the boundary separating the reservoirs of these gifts of nature from the realm of cultural transformation within which subboundaries indicate the different stages of their subsequent transformation into the goods and services of final demand. Some such diagram should be presented to every student in an introductory economics course so that the way in which the economic process operates in the real world is properly understood.

Argument 3: Validation. Natural scientists expect theoretical models to be tested before they are applied or developed further. Unfortunately, economic policy with farreaching consequences is often based on economic models that, although elegant and widely accepted, are not validated (Daly 1977, Cleveland et al. 1984, Dung 1992, Ayres 1996). Empirical tests to validate economic models are undertaken even less frequently in the developing countries, where these models are followed regularly (e.g., Kroeger and Montagne 2000). As Nobel laureate in economics Wassily Leontief noted, many economic models are unable "to advance, in any perceptible way, a systematic understanding of the structure and the operations of a real economic system"; instead, they are based on "sets of more or less plausible but entirely arbitrary assumptions" leading to "precisely stated but irrelevant theoretical conclusions" (Leontief 1982).

Most noneconomists do not appreciate the degree to which contemporary economics is laden with arbitrary assumptions. Nominally objective operations, such as determining the least cost for a project, evaluating costs and benefits, or calculating the total cost of a project, normally use explicit and supposedly objective economic criteria. In theory, all economists might come up with the same conclusions to a given problem. In fact, such "objective" analyses, based on arbitrary and convenient assumptions, produce logically and mathematically tractable—but not necessarily realistic—models.

Where there have been empirical analyses (of, for example, consumer choice), the results frequently have shown that the behavior of real people in experimental or laboratory situations was quite different from the assumptions of a given neoclassical model (Schoemaker 1982, Smith 1989, Hall 1991). On the one hand, this is not surprising, because social science models of human behavior sometimes apply, and sometimes they do not, depending upon which modeled subset of the infinite set of human behavioral patterns is matched by the actual group of people to which the model is applied. On the other hand, the authority economists often assign to their models is perplexing, because unavoidably fuzzy economic models do not become precise just because they emulate the mathematical rigor of physics.

For example, Hamiltonians are used in economics in analogy to the Hamiltonians in physics. In fact, in physics a Hamiltonian is an energy function representing the sum of kinetic and potential energy in a system from which one can derive the equations of motion of the particles of the system. In neoclassical production theory, the price vector is given by the gradient of the output in the space of the production factors. This corresponds formally to the vector of a conservative physical force, which is given by the gradient of potential energy in real space (Mirowski 1989). This formal analogy would result only in an appropriate description of economic situations if the economy evolved in a state of equilibrium characterized by a profit maximum that lies in the interiornot on the boundary-of the factor space accessible to the production system, according to its state of technology. However, as we show in the next section, this equilibrium has not been satisfied during three decades of industrial evolution in the United States, Japan, and Germany under the reign of low energy prices. Rather, the economies have been sliding downhill on the slope of the cost mountain inclined toward the cost minimum in the state of total automation. This state is characterized by minimum inputs of expensive labor and maximum inputs of cheap energy combined with highly automated capital. Because of technological constraints, this cost minimum has not yet been reached (Kümmel and Strassl 1985).

Validation also proves difficult or impossible because both classical and neoclassical theories were originally developed using concepts of production factors as they existed in agrarian societies. These theories have been transferred more or less unchanged to applications in the modern industrial world. Very often no provisions have been added to the basic theory for industrialization and its consequences.

The importance of energy to economic production

In industrial economies the capital stock consists of all energy conversion devices and the installations and buildings necessary for their operation and protection. Its fundamental components are heat engines and transistors (formerly mechanical switches, relays, and electronic valves), activated by energy and handled by labor. They provide the average citizen of the industrially developed countries with services that are energetically equivalent to those of 10 to 30 hard-laboring people-"energy slaves," if you will. These numbers would more than triple if one included energy for room and process heat. In 1995 primary energy consumption per capita per day was 133 kWh in Germany and 270 kWh in the United States. This would correspond numerically to about 44 and 90 energy slaves per capita in Germany and in the United States, respectively, each one delivering about 3 kWh per day. Huge armies of energy slaves create our wealth.

In order to demonstrate the economic importance of energy quantitatively, we present an econometric analysis of economic growth over three decades for the United States, Japan, and Germany (Kümmel 1980, 1982, 1989, Kümmel et al. 1985, Kümmel et al. 2000). This analysis shows how the proper inclusion of energy removes most of the unexplained residual encountered by neoclassical theory.

We make the fundamental assumption that wealth, as represented by the output Q of value added, is created by the cooperation of the production factors capital K, labor L, and energy E in conjunction with creativity Cr. Raw materials are the passive partners of the production process. They are critically important but do not contribute by themselves to the generation of value added. Their monetary value is not included in the national accounts' empirical time series on value added with which we compare our theoretical results. However, if materials become scarce in spite of recycling, growth of course will be constrained. In systems in which catalytic processes play a quantitatively important role, one might consider treating the catalytic materials as a factor distinct from the capital stock. Creativity is that specifically human contribution to economic evolution that cannot be made by any machine capable of learning and that cannot be realized by changing factor combinations. Creativity contributes ideas, inventions, value judgments, and decisions. Creativity's influence may be weak in the short run but important in the long run. In fact, creativity often has been about finding ways to increase energy subsidies for a task.

Q is of necessity measured in inflation-corrected monetary units, and so is K, whereas appropriate measures for E are Petajoules per year and for L man-hours worked per year. E and L are obtained from the national energy and labor statistics, and K and Q from the national accounts. Ideally, one would like to measure K by the amount of work performance and information processing that capital stock is able to deliver when totally activated by energy and labor. Likewise, the output Q might be measured by the work performance and information processing necessary for its generation. The detailed, quantitative technological definitions of K and Q are given by Kümmel (1980, 1982, Kümmel et al. 2000). However, these physical measurements of K and Q are not available. Therefore, we assume proportionality between them and the constant currency data. We normalize all variables to their values (Q_0, K_0, L_0, E_0) for a base year. For a quantitative analysis of growth, we employ production functions q = q[k(t), l(t), l(t)]e(t); t, which describe the evolution of the normalized output $q = Q/Q_0$ as the normalized inputs of capital, $k = K/K_0$, labor, $l = L/L_0$, and energy, $e = E/E_0$ change with time *t*; we allow for an explicit time-dependence of q in order to model the effects of creativity.1

We calculate production functions from the following growth equation that relates the (infinitesimal) relative change of the normalized output, dq/q, to the relative changes of the normalized inputs, dk/k, dl/l, de/e, and creativity's action: (1) $dq/q = \alpha(dk/k) + \beta(dl/l) + \gamma(de/e) + Cr$.

 α , β , and γ are called the *elasticities of production* of capital, labor, and energy in the language of economics. They measure the productive powers of the factors in the sense that

(roughly speaking) they give the percentage of output change when the corresponding inputs change by 1%. They, and *Cr*, involve the partial derivatives of q^2 . If one can approximately neglect the explicit time dependence of q, as we shall do for the moment, one has Cr = 0.

Our procedure for calculating the production function from equation (1) differs in one essential point from that of neoclassical economics: We do not set α , β , and γ equal to the cost shares of capital, labor, and energy in total factor cost. This stipulated equality of elasticities of production and cost shares is a result of the fundamental hypotheses underlying the neoclassical equilibrium model. Instead, we determine these coefficients differently using an econometric analysis and a set of three differential equations representing the integrability conditions of the production function.³

The simplest nonconstant solutions of these equations with technologically meaningful boundary conditions are $\alpha = a_0(l+e)/k$, $\beta = a_0(c_0(l/e)-l/k)$, and $\gamma = 1 - \alpha - \beta$, with technology parameters a_0 and c_0 .⁴ Here, a_0 gives the weight with which the ratios of labor to capital and energy to capital contribute to the productive power of capital, and c_0 indicates the energy demand $e_t = c_0 k_t(q_t)$ of the fully utilized capital stock $k_t(q_t)$ that would be required to generate the fraction q_t of output accessible to totally automated production with virtually no labor, while the production of $(q - q_t)$ is labor saturated; then β goes to zero as e and k approach e_t and k_t . If one inserts these elasticities of production into equation (1) and integrates, with Cr = 0, one obtains the (first) LINEX production function

(2) $q = q_0 e \exp[a_0(2 - (l+e)/k) + a_0 c_0(l/e - 1)],$

which depends *linearly* on energy and *exponentially* on quotients of capital, labor, and energy.

The integration constant q_0 is the third technology parameter of the theory. Its changes indicate changes in the monetary valuation of the original basket of goods and services making up the output unit Q_0 . Activities of creativity Cr that lead to an explicit time dependence of the production function can be modeled by allowing a_0 , c_0 , and q_0 to change in time. The elasticities of production, α , β , and γ , must be nonnegative to make sense economically. This poses important restrictions on the admissible factor quotients in α , β , and equation (2). (Integration of equation (1) with the constants α_0 , β_0 , and $\gamma_0 = 1 - \alpha_0 - \beta_0$, yields the energy dependent Cobb-Douglas production function $q = q_0 k^{\alpha_0} l^{\beta_0} e^{1-\alpha_0 - \beta_0}$. This function, however, violates the laws of thermodynamics because it allows for the almost complete substitution of energy by capital. Thus, it should be avoided in scenarios for the future.) Our model incorporates the limits to substitution, thanks to the restrictions on α , β , and γ . The LINEX function is of the type "variable elasticities of substitution." Its relation to the frequently used translog function has been discussed by Kümmel et al. (1985).

We tested our energy-dependent production function (eq. 2) with empirical data, examining the sectors "Industries" of the United States and Japan and the West German manufacturing sector (*Warenproduzierendes Gewerbe*). (The sectors "In-

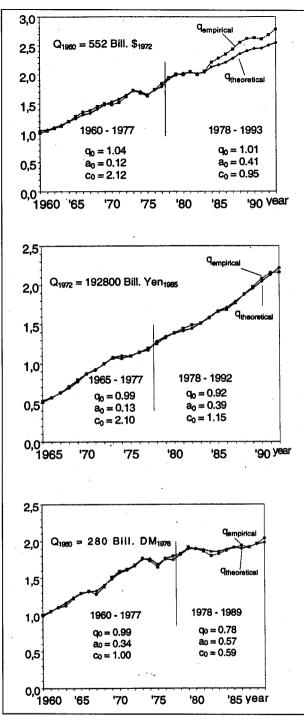


Figure 3. Theoretical (diamonds) and empirical (squares) growth of annual industrial production $q = Q/Q_0$ in the USA ($Q_0 = Q_{1960}$), top; Japan ($Q_0 = Q_{1972}$), middle; and West Germany ($Q_0 = Q_{1960}$), bottom. In all three systems the overall growth of the capital stock k is similar to the overall growth of the output q, and the ups and downs of energy inputs e and outputs q occur at the same times. Labor 1 rises in the United States, stays nearly constant in Japan, and decreases in West Germany. The empirical time series of k,l,e can be found at the Web site http:// theorie.physik.uni-wuerzburg.de/TP1/kuemmel/ profile.html.

dustries" are defined by the "system of national accounts" and include the services-producing sectors). We were able to obtain consistent sets of data for these sectors, which produce about 80%, 90%, and 50%, respectively, of gross domestic product. When we inserted the numerical values for the technology parameters given in Figure 3 and the annual empirical inputs of k, l, and e for the United States from 1960 to 1993, Japan from 1965 to 1992, and West Germany from 1960 to 1989 into the LINEX function, we obtained the theoretical outputs shown in Figure 3, together with the annual empirical outputs. For each country the numerical values of the three technology parameters have been determined by fitting the LINEX function to the empirical time series of output before and after 1977, using the Levenberg-Marquardt method (see Press et al. 1992). This results in the different sets of a_0 , c_0 , and q_0 shown in Figure 3, that is, a time dependence of the parameters between 1977 and 1978.

Results

The LINEX functions, which include the production factor energy, reproduce the output of all three production systems for all years considered with only minor residuals, including the recessions caused by the two major energy crises. The energy crises were triggered by the first and second oil-price explosions in 1973-1975 and 1979-1981 in the wake of the Yom Kippur war between Israel and its Arab neighbors and the war between Iraq and Iran, respectively. The influence of creativity in response to the oil price increase shows in the reduction of the energy demand of the capital stock, c_0 , and the enhancement of capital's productive power by the enhanced a_0 after 1977. These shifts of technology parameters are the result of the decisions of governments and entrepreneurs to invest in energy conservation technologies after the shock of the first oil-price explosion. Structural changes toward less energy-intensive economic activities played a role as well.

Of course, the limitation of the parametric time changes to one year is a consequence of our simple modeling of creativity's action as a single one-year pulse. If one goes a step further-that is, assumes that creativity is always active and models the transitions between the different values of a_0 and c_0 before and after the energy crises using continuous functions of time-the discrepancies between the theoretical and empirical USA curves after 1985 disappear and the results for Japan and Germany remain practically the same⁵ (Henn 2000). In any case, in the short run the changes caused by creativity are small compared with the changes caused by the changing combinations of capital, labor, and energy. Therefore, creativity's influence, and thus any explicit time dependence of the production function, can be neglected during time spans of at least a decade. Even without any parameter readjustments between 1977 and 1978, the evolution of production in Germany and Japan during three decades is reproduced by the LINEX function with residuals of less than 10% (Kümmel et al. 2000). Other energy-dependent production functions with mathematically simpler (i.e., constant) or more complicated elasticities of production yield quantitatively and qualitatively similar results⁶ (Lindenberger 2000).

The results of our analysis also demonstrate in all three cases that the productive power of energy is more important than that of capital or labor, and nearly an order of magnitude larger than the 5% share of energy cost in total factor cost. This follows from the time-averaged LINEX elasticities of production of capital, labor, and energy, ⁷ which are for the United States, $\overline{\alpha} = 0.36$, $\overline{\beta} = 0.10$, $\overline{\gamma} = 0.54$; for Japan, $\overline{\alpha} = 0.34$, $\overline{\beta} = 0.21$, $\overline{\gamma} = 0.45$; and for West Germany, $\overline{\alpha} = 0.45$, $\overline{\beta} = 0.05$, $\overline{\gamma} = 0.50$.

In addition, the time-averaged elasticity of production of labor, $\hat{\beta}$, is much smaller than labor's cost share (typically 0.70). In industrialized countries such as the United States, energy commands about 5%, labor about 70%, and capital about 25% of total factor cost. This means that one of the fundamental assumptions of neoclassical equilibrium economics, that is, the equality of elasticities of production and cost shares, has not been satisfied under the conditions of production prevailing in the United States, Japan, and Germany over the last three decades. Rather, under pressure to minimize costs, these economies have been driven into substituting cheap, powerful energy (in combination with increasingly automated capital stock) for expensive labor, which is weak economically in the sense that its elasticity of production is much smaller than that of energy. This substitution of energy for labor takes time because of technical constraints on the progress of automation, the demand for those products and services that cannot be produced in a totally automated fashion, and still-existing and respected laws and agreements. Therefore, the economies of the industrial countries have not yet reached the absolute cost minimum.

Some social implications of our analysis

If one accepts the importance of a biophysical basis for economics, then our analysis has some important implications for economics and for society.

The replacement of expensive labor in routine jobs with the combination of cheap energy and capital stock is likely to continue under the present incentive structure. This combination also reinforces the trend toward globalization, because goods and services produced in low-wage countries can be transported cheaply to high-wage countries. Thus, high unemployment (in most high-wage countries) will continue if the disparities between the productive powers and cost shares of labor and energy are not removed (for example, by adjusting fiscal policy). Certainly, the low price of fossil fuels relative to their productive power generates large profits. But, as is well known, it also prevents the market penetration of large-scale energy-conserving and nonfossil energy technologies, which could lower the demand for fossil fuels and relieve some of the burden of pollution. We therefore believe that the problems of unemployment, resource depletion, and pollution can be attacked successfully only if the pivotal role of energy as a factor of production is properly taken into account in economic and social policy.

Price does not always reflect scarcity and economic importance. Scarcity of a resource must be defined in terms of both shortand long-term resource availability. Price, the economist's usual metric of scarcity, reflects many important aspects of scarcity poorly because it is often based on short-term market values. Most important, as Norgaard (1990) and Reynolds (1999) show, is that uncertainty about the size of the base of a resource can obscure the actual trend in scarcity of that resource, with the result that "empirical data on cost and price...do not necessarily imply decreasing scarcity" (Reynolds 1999, p. 165). As an example of this phenomenon, in mid-1999 the real price of oil was at nearly its lowest level ever, despite the fact that most estimates of the time at which global oil production will peak range from 2000 to 2020 (Kerr 1998, Cleveland 1999).

The concept and implementation of sustainable development as interpreted and advocated by most economists must be thought through much more carefully, given the requirement for energy and materials for all economic activity (see Hall 2000 for a detailed analysis of Costa Rica). Energy is in fact disproportionately more important in terms of its impact on the economy than its monetary value suggests, as evidenced by the events of the 1970s (i.e., inflation, stock market declines, reduced economic output, and so on), which appear to be reoccurring to some degree in 2000 partly in response to a similar proportional increase in the price of oil. Fundamentally, current societal infrastructure has been built and maintained on the basis of abundant, cheap supplies of highquality energy-that is, energy characterized by the large amount of energy delivered to society per unit of energy invested in this delivery (through exploration and development or through trade of goods for imported energy [Hall et al. 1986]).

In developing nations, investment policies based on neoclassical economic analyses encourage borrowing from developed countries and hence growing indebtedness. Pressure to service the debt encourages the quick extraction of resources to generate a cash flow so that payments of interest and repayment of principal can be maintained. In the meantime, the longterm productivity of the region may be destroyed. But those assessments are not included in neoclassical analyses; in the rare cases where resources are included in the analysis, their value is heavily discounted. For example, many tropical countries sell their forest products at a price far below their worth (Repetto 1988, Hall 2000), and the Russian government has been talked into abolishing its export tax on fossil fuels, which was the last source of secure revenues for highly indebted Russia. Developing countries and nations in transition to market economies should attribute more importance to their natural resources than they presently do under the influence of the reigning economic theory.

Humans tend to seek political explanations for events that in fact may have been precipitated by biophysical causes. For example, Reynolds (2000) shows how the sharp decline in the former Soviet Union's oil production may have precipitated the economic crises that led to the collapse of the Soviet Union.

Some biological implications of our analysis

Economies, just like ecosystems—or indeed any system can be represented as stocks and flows of materials and energy, with human material welfare largely a function of the per capita availability of these stocks and flows.

Present agricultural technologies, most wildlife management and conservation programs, and perhaps biomedical technologies are as dependent on the availability of cheap energy as anything else. For example, most increases in agricultural productivity have not come from genetics alone. In fact, for many crops there appears to be essentially no increase in gross photosynthesis but rather only an increase in the proportion of photosynthate that goes to the parts we eat, often seeds, while the organs and functions of a wild plant (e.g., growing roots to take up more nutrients and water, generating secondary compounds for insect defense) are increasingly supplied by industrially derived inputs from outside the plant (Smil 2000). In addition, the efficiency of agriculture tends to be inversely related to the intensity of use of land area or fertilizer (Hall et al. 1998, Hall 2000, chap. 12).

Human material well-being is derived essentially by redirecting energy stocks and flows from what natural selection and the accidents of geology dictated to ends determined by human needs and, increasingly, desires. Now some 40% to 60% of global primary production is exploited, in one way or another, by the human economy (Vitousek et al. 1986, 1997).

Outlook: The challenge to construct a model that includes the biophysical basis of the economy

Existing "economic" models cannot effectively represent a total economy, because none has a biophysical basis; some attempts to produce such a model have meen made, however. First, there are very detailed and comprehensive models of the flow of energy through each sector of the US economy (Hannon 1982). But these do not include the flows of nature (such as the energy associated with the hydrological cycle, flows of rivers, solar energy, photosynthesis, and other important components of the economic system). Another approach, one that garners considerable controversy, does include the energy flows of nature and the human economy: This is emergy (with an "m") analysis, which also attempts to give each energy flow a weighting, according to its quality (Odum 1996). This approach has been applied at an aggregated level to national economies and used as the basis for policy recommendations (Brown et al. 1995).

Finally, evolutionary economics looks for ways to model the economic process by combining nature's principle of selforganization with the growth of human knowledge and innovations⁸ (Witt 1997, Faber and Proops 1998).

We must conclude, however, that a truly useful and acceptable model that includes the biophysical basis of the economy is probably still far in the future. What then is the utility of bringing a biophysical perspective into economics? We believe that it is overwhelmingly heuristic. By thinking about economies as they actually are (i.e., Figure 1b or 2) instead of how we might conceptualize them for analytic ease and tractability (i.e., Figure 1a), we can teach a new generation of economists about the real operations of human economies and the various links to the "economies" of the natural world. We believe that doing so is especially important because science gives us to understand that there are at least constraints, and possibly even limits, to growth. Future generations of economists probably will not be able to treat such issues as overpopulation, oil and groundwater depletion, and changes in the composition of the atmosphere and the biosphere simply as "externalities" to be given a price and rolled into the larger analysis; these will have to be treated as fundamental components of the total economic model. We do not understand how that can be done without starting from a biophysical basis. We challenge a new generation of economists and natural scientists to think from this perspective.

Acknowledgments

We thank Rick Beal, Bart Daly, Jae-Young Ko, and Paul Christensen for helpful discussions; four anonymous reviewers for critical comments; Julian Henn for analytic help; the Roosevelt Wildlife Collection (Roosevelt Wildlife Station contribution number 2001–03); and the Denman Foundation for logistic inputs.

References cited

- Ayres RU. 1996. Limits to the growth paradigm. Ecological Economics 19: 117–134.
- Barnett HJ, Morse C. 1963. Scarcity and Growth: The Economics of Natural Resources Availability. Baltimore (MD): Johns Hopkins University Press.
- Baron R. 1997. Competitive Issues Related to Carbon/Energy Taxation. Paris: Expert Group on the United Framework Convention on Climate Change. Working paper no. 14.
- Binswanger HC, Ledergerber E. 1974. Pages 107–125 in Wolff J, ed. Wirtschaftspolitik in der Umweltkrise. Stuttgart (Germany): DVA.
- Boulding, KE. 1966. The economics of the coming spaceship Earth. In Jarrett H, ed. Environmental Quality in a Growing Economy. Washington (DC): Resources for the Future.
- Brown MT, Odum HT, Murphy RC, Christianson RA, Doherty SJ, Mc-Clanahan TR, Tennenbaum SE. 1995. Rediscovery of the world: Developing an interface of ecology and economics. Pages 216–250 in Hall CAS, ed. Maximum Power: The Ideas and Applications of H. T. Odum. Niwot (CO): University of Colorado Press.
- Christensen P. 1989. Historical roots for ecological economics-bio-physical versus allocative approaches. Ecological Economics 1: 17–36.
- Cleveland CJ. 1991. Natural Resource Scarcity and Economic Growth Revisited: Economic and Biophysical Perspectives. Pages 289–317 in Costanza R, ed. Ecological Economics: The Science and Management of Sustainability. New York: Columbia University Press.

Cleveland CJ, Ruth M. 1997. When, where, and by how much do biophysical limits constrain the economic process? A survey of Nicholas Georgescu-

^{-------.} Presentation to Bank of America Securities Energy Conference; 29 June 1999; Houston, TX.

Roegen's contribution to ecological economics. Ecological Economics 22: 203–223.

- Cleveland CJ, Costanza R, Hall CAS, Kaufmann RK. 1984. Energy and the US economy: A biophysical perspective. Science 225: 890–897.
- Daly HE, ed. 1973. Toward a Steady-State Economy. San Francisco: W. H. Freeman.
- Daly HE. 1977. Steady-State Economics. San Francisco: W. H. Freeman.
- Dasgupta P, Levin S, Lubchenco J. 2000. Economic pathways to ecological sustainability. BioScience 50: 339–345.
- Denison EF. 1979. Explanations of declining productivity growth. Survey of Current Business 59 (Part 2): 1–24.
 - ——. 1984. Accounting for slower economic growth. Pages 1–45 in Kendrick JW, ed. International Comparisons of Productivity and Causes of the Slowdown. Cambridge (MA): Ballinger.
- Deutsche Bundesbank. 1996. Makro-ökonometrisches Mehr-Länder-Modell. Frankfurt (Germany): Deutsche Bundesbank.
- Dung TH. 1992. Consumption, production and technological progress: A unified entropic approach. Ecological Economics 6: 195–210.
- Faber M, Proops JLR. 1998. Evolution, Time, Production and the Environment. 3rd ed. Berlin: Springer-Verlag.
- Georgescu-Roegen N. 1966. Some orientation issues in economics. In Georgescu-Roegen N, ed. Analytical Economic Issues and Problems. Cambridge (MA): Harvard University Press.
- Hall CAS. 1991. An idiosyncratic assessment of the role of mathematical mod-
- —, ed. 2000. Quantifying Sustainable Development: The Future of Tropical Economies. San Diego (CA): Academic Press.
- Hall CAS, Cleveland CJ, Kaufmann RK. 1986. Energy and Resource Quality: The Ecology of the Economic Process. New York: Wiley-Interscience.
- Hall CAS, Ko J-Y, Lee C-L, Wang HQ. 1998. Ricardo lives: The inverse relation of resource exploitation intensity and efficiency in Costa Rican agriculture and its relation to sustainable development. Pages 355–370 in Ulgiati S, Brown M, eds. Advances in Energy Studies, Energy Flows in Ecology and Economy. Rome: Musis.
- Hannon B. 1981. Analysis of the Energy Cost of Economic Activities: 1963–2000. Urbana (IL): University of Illinois. Energy Research Group Doc. No. 316.
- Henn J. 2000. Die Produktionsmächtigkeit von Energie und Kreativität: Eine Zeitreihenanalyse für Deutschland, Japan und die USA. Diplomarbeit, Universität Würzburg, Germany.
- Jorgenson DW. 1984. The role of energy in productivity growth. American Economic Review 74: 26–30.
- ———. 1988. Productivity and economic growth in Japan and the United States. American Economic Review 78: 217–222.
- Kerr RA. 1998. The next oil crisis looms large—and perhaps close. Science 281: 1128–1131.
- Kroeger T, Montagne D. 2000. An assessment of the effects of structural adjustment policies in Costa Rica. Pages 665–694 in Hall C, ed. Quantifying Sustainable Development: The Future of Tropical Economies. San Diego (CA): Academic Press.
- Kümmel R. 1980. Growth Dynamics of the Energy Dependent Economy. Cambridge (MA): Oelgeschlager, Gunn and Hain.
 - ------. 1982. The impact of energy on industrial growth. Energy: The International Journal 7: 189–203.
- Kümmel R, Strassl W. 1985. Changing energy prices, information technology, and industrial growth. Pages 175–194 in Van Gool W, Bruggink JJC, eds. Energy and Time in the Economic and Physical Sciences. Amsterdam: North-Holland.

- Kümmel R, Strassl W, Gossner A, Eichhorn W. 1985. Technical progress and energy dependent production functions. Z. Nationalökonomie Journal of Economics 45: 285–311.
- Kümmel R, Lindenberger D, Eichhorn W. 2000. The productive power of energy and economic evolution. Indian Journal of Applied Economics 8 (Special Issue in Honour of Paul A. Samuelson): 231–262.

Leontief W. 1982. Academic economics. Science 217: 104-107.

- Lindenberger D. 2000. Wachstumsdynamik industrieller Volkswirtschaften: Energieabhängige Produktionsfunktionen und ein faktorpreisgesteuertes Optimierungsmodell. Marburg (Germany): Metropolis.
- Marris R. 1992. Implications for economics. In Egidi M, Marris R, eds. Economics, Bounded Rationality and the Cognitive Revolution. Brookfield (VT): Edward Elgar.
- Martinez-Alier J. 1987. Ecological Economics: Energy, Environment and Society. Oxford (UK): Blackwell.
- Marx K. 1906. Capital. New York: Modern Library.
- Mirowski P. 1989. More Heat than Light: Economics as Social Physics, Physics as Nature's Economics. Cambridge (UK): Cambridge University Press.
- Norgaard RB. 1990. Economic indicators of resource scarcity: A critical essay. Journal of Environmental Economics and Management 19: 19–25.
- Odum HT. 1996. Environmental Accounting: Energy and Environmental Decision Making. New York: Wiley and Sons.
- Press WH, Flannery BP, Teukolsky SA. 1992. Numerical Recipes in C. Cambridge (UK): Cambridge University Press.
- Repetto R. 1988. The Forest for the Trees? Government Policies and the Misuse of Forest Resources. Washington (DC): World Resources Institute.
- Reynolds DB. 1999. The mineral economy: How prices and costs can falsely signal decreasing scarcity. Ecological Economics 31: 155–166.
- ———. 2000. Soviet economic decline: Did an oil crisis cause the transition in the Soviet Union? Journal of Energy and Development 24: 65–82.
- Ricardo D. 1891. The Principles of Political Economy and Taxation. 3rd ed. London: G. Bell and Sons.
- Roberts PC. 1982. Energy and value. Energy Policy 10: 171-180.

Schoemaker PJH. 1982. The expected utility model: Its variants, purposes, evidence and limitations. Journal of Economic Literature 20: 529–563.

- Smil V. 2000. Feeding the World: A Challenge for the Twenty-First Century. Cambridge (MA): MIT Press.
- Smith A. 1937. An Inquiry into the Nature and Causes of the Wealth of Nations. 5th ed. New York: Modern Library.
- Smith VL. 1989. Theory, experiment and economics. Journal of Economic Perspectives 3: 151–169.
- Soddy F. 1926. Wealth, Virtual Wealth and Debt. New York: E. P. Dutton.
- Solow RM. 1974. The economics of resources or the resources of economics. American Economic Review 64: 1–14.
- ------. 1992. An Almost Practical Step toward Sustainability. Washington (DC): Resources for the Future.
- ———. 1993. Policies for Economic Growth. Pages 127–140 in Knoester A, Wellink AHEM, eds. Tinbergen Lectures on Economic Policy. Amsterdam: North-Holland.
- ——. 1994. Perspectives on growth theory. Journal of Economic Perspectives 8: 45–54.
- Tryon FG. 1927. An index of consumption of fuels and water power. Journal of the American Statistical Association 22: 271–282.
- Vitousek PM, Ehrlich PR, Ehrlich AH, Matson PA. 1986. Human appropriation of the products of photosynthesis. BioScience 36: 368–373.
- Vitousek PM, Mooney HA, Lubchenco J, Melillo JM. 1997. Human domination of Earth's ecosystems. Science 277: 494–499.
- Witt U. 1997. Self-organization and economics: What is new? Structural Change and Economic Dynamics 8: 489–507.

Endnotes

1. The constraints on economic growth attributable to entropy production (Kümmel 1980, 1989) will not be considered in this analysis of the past.

2. Equation (1) results from the total differential of the production function. The elasticities of production are $\alpha(k,l,e) \equiv (k/q)(\partial q/\partial k)$, $\beta(k,l,e) \equiv (l/q)(\partial q/\partial l)$, $\gamma(k,l,e) \equiv (e/q)(\partial q/\partial e)$, and the term due to the creativity-induced explicit time dependence of the production function is $Cr = (t/q)(\partial q/\partial t)(dt/t)$.

3. The differential equations result from the requirement that the second-order mixed derivatives of the production function with respect to the production factors are equal. With the assumption of constant returns to scale, i.e., $\gamma = 1 - \alpha - \beta$, the differential equation for α is $k(\partial\alpha/\partial k) + l(\partial\alpha/\partial l) + e(\partial\alpha/\partial e) = 0$, the equation for β has identical structure, and the coupling equation reads $l(\partial\alpha/\partial l) = k(\partial\beta/\partial k)$. The most general solutions of the first two equations are $\alpha = f(l/k, e/k)$ and $\beta = g(l/k, e/k)$, with arbitrary differentiable functions *f* and *g*. The boundary conditions that would unequivocally determine the solutions of this system of partial differential equations would require knowledge of β on a surface and of α on a curve in *k*, *l*, *e* space. It is practically impossible to obtain such knowledge. Therefore, one has to choose approximate or asymptotic boundary conditions.

4. These solutions take into account the possible approach to the state of total automation, as described in the paragraph preceeding equation (2), and the condition that α must vanish if (l + e)/k goes to zero: With zero labor and energy, i.e., zero capacity utilization of capital, capital growth cannot contribute to output growth. These "asymmetric" boundary conditions lead to the "asymmetric" solutions of the symmetric set of differential equations. When we tested other boundary conditions and more sophisticated elasticities of production with the corresponding "higher" LINEX functions, the quantitative results did not change significantly (Kümmel et al. 1985, Lindenberger 2000).

5. Yet another modeling of creativity's action is possible for West Germany, where we know the time series of the share of electricity El(t) in end-energy consumption: If one replaces e by [1+El(t)]e in the LINEX production function and determines the three technology parameters by only one fitting procedure for the period 1960 to 1989, one obtains a theoretical output that is barely discernible from the one in Figure 3, bottom (Kümmel et al. 2000). This is consistent with the observation that efficiency improvements nor-

mally require more electrical devices and confirms the view that electrification and technological progress are closely interrelated (Jorgenson 1984).

6. Like the Deutsche Bundesbank (Federal Reserve Bank of Germany; 1996) in its macroeconometric multicountry model, we present here the standard econometric quality measures, namely, the coefficient of determination, R^2 (the "best" possible value is 1.0), and the Durbin-Watson coefficient of autocorrelation, dw (the "best" possible value is 2.0). The R^2 and dw pertinent to the LINEX functions in Figure 3 are, for West Germany, 0.991 and 1.23 during 1960-1977, 0.782 and 0.96 during 1978-1989; for Japan, 0.995 and 1.22 during 1965-1977, 0.992 and 1.15 during 1978-1992; and for the United States, 0.983 and 0.65 during 1960-1977, and very small values during 1978-1993. In Julian Henn's (2000) innovation-diffusion model with continuously decreasing $c_o(t)$ and increasing $a_o(t)$ —not shown in Figure 3—one finds, for the United States, $R^2 = 0.997$ and dw =0.95 for the period 1960–1993; for Japan and Germany, the R² and dw are better than 0.993 and 1.57 for the full length of the observation times. The technology parameters have been determined with the help of the Levenberg-Marquardt method in nonlinear optimization, subject to the constraints of nonnegative elasticities of production (see Press et al. 1992).

The positive autocorrelations are due to the unavoidable approximations for the boundary conditions on the elasticities of production³ and, as a consequence, the necessarily approximate character of the production functions. When estimating the gross domestic product of the United States, Japan, and Germany between 1974 and 1995, using a translog-type production function of capital and labor with cost-share weighting and exponential time dependence, the econometricians of the Deutsche Bundesbank (1996) obtained 0.997, 0.995, and 0.97 for R^2 and 0.42, 0.32, and 0.24 for *dw*, respectively.

7. The time-averaged LINEX elasticities are approximately equal to the constant elasticities of production of the energy-dependent Cobb-Douglas production function $q = q_0 k^{\alpha_0} \beta_0 e^{1-\alpha_0-\beta_0}$ that also fits reasonably well to the empirical data. Thus, energy-augmented Cobb-Douglas functions approximate the LINEX functions on past growth-paths in factor space that, of course, did not violate the physical limits to substitution.

8. An opportunity to start this process was offered by the seminar "Economic Growth: Driving Forces and Constraints in the Perspective of Economics and the Sciences" of the WE-Heraeus Foundation (WE-Heraeus-Stiftung, PO Box 1553, D-63405 Hanau, Germany) from 22–25 October 2000, in Bad Honnef, Germany.