Institute of Energy Economics
at the University of Cologne

Institute of Energy Economics at the University of Cologne
Albertus-Magnus-Platz
50923 Cologne, Germany

EWI Working Paper, No 08.02

Cost Shares, Output Elasticities, and Substitutability Constraints

by

Reiner Kümmel, Jörg Schmid, Robert U. Ayres, and Dietmar Lindenberger

April 2008

Cost Shares, Output Elasticities and Substitutability Constraints

Reiner Kümmel,[a] Jörg Schmid,[a] Robert U. Ayres,[b] and Dietmar Lindenberger [c] [1]
[a] Institute for Theoretical Physics and Astrophysics, University of Würzburg,
D-97074 Würzburg, Germany,
[b] European School of Business Administration INSEAD, Fontainebleau 77305, France,
[c] Institute of Energy Economics, University of Cologne,
D-50923 Cologne, Germany.

April 16, 2008

Abstract:
The equilibrium conditions for an economic system that produces output with several factors of production and which is subject to technological constraints are derived. Optimization of either output minus cost or integrated utility yields the conditions that output elasticities must be equal to a modification of the usual factor cost shares, where shadow prices due to the constraints add to factor prices. In a model, where capital, labor and energy (exergy) are the factors of production, the technological constraints are identified as limits to capacity utilization and automation. The shadow prices depend on the output elasticities. These elasticities are determined for Germany, Japan and the USA by econometric estimations of energy-dependent production functions that are derived from the twice differentiability requirement and the law of diminishing returns.

JEL classifications: A12, C02, E13, Q43
Keywords: Optimization, technological constraints, shadow prices, output elasticities, energy, economic growth.

[1] Corresponding author; email: dietmar.lindenberger@uni-koeln.de

1. INTRODUCTION

Science and technology have contributed significantly to economic growth by providing the means for the exploitation of natural resources. Problems emerge, because industrial production is coupled to energy conversion and energy conversion to emissions. This has stimulated research into the economics of resources and the resources of economics [1]. The first and the second oil-price shock, and more recently climate change [2] and peak oil [3], indicate new challenges to welfare optimization faced by industrial economies. Constraints not taken into account before may become relevant. At the same time progress in information technology and automation has been reducing constraints on the technologically possible combinations of production factors, with significant consequences for the evolution of production and employment.

This paper does not attempt to analyze the impact of the new constraints on future growth. It is rather limited to a description of economic equilibrium under the hitherto existing technological constraints. Since aspects of the old constraints are related to the new ones, the analysis of the past may also benefit scenarios for the future.

The direct motivation of the paper is the observation that energy-dependent production functions that reproduce past economic growth without the Solow residual have output elasticities that deviate significantly from the factor cost shares [4] - [11]. Discussing this discrepancy previously, we have reasoned that technological constraints have prevented the economies from operating in the absolute profit maximum, where the partial derivatives of (output minus cost) vanish so that output elasticities and factor cost shares are equal. We now incorporate this qualitative view mathematically into the neoclassical vision, which involves economic agents optimizing subject to all relevant constraints.

The scope of the paper is the following. In Section 2 we briefly review some aspects of neoclassical and endogeneous growth theory. Section 3 presents the equilibrium conditions that result from profit maximization subject to technological constraints. These conditions are also derived in Appendix A by optimizing the integral of utility of consumption, starting from Ramsey's capital model [12] as presented by Samuelson and Solow [13] and modifying it by taking into account three factors of production, finite time horizons, and technological constraints. The equilibrium conditions are expressed as equalities between output elasticities and a modification of the usual factor cost shares, where shadow prices due to the constraints add to factor prices. Section 4 computes the Lagrange multipliers that enter the shadow prices for the case of two technological constraints. These are identified in Section 5 as constraints on capacity utilization and automation with appropriate slack variables. In Section 6 and Appendix B we calculate energy-dependent production functions and their output elasticities from the general partial differential equations that result from the requirement that second-order mixed derivatives be equal. Asymptotic boundary conditions that reflect the law of diminishing returns lead to the Linex production function, which reproduces past economic growth and the energy crises in Germany, Japan and the USA with small residuals and Durbin-Watson coefficients between 1.5 and 2. Section 7 indicates, how the shadow prices could be actually computed from the Linex function, its output elasticities and time series on factor prices, capacity utilization and two constraint parameters. Section 8 is Summary and Conclusions.

2. PERSPECTIVES OF NEOCLASSICAL AND ENDOGENEOUS GROWTH THEORY.

The oil price shocks 1973-1975 and 1979-1981 and the accompanying recessions known as the first and the second energy crisis prompted numerous investigations concerning the economic role of energy [14] - [20]. Jorgenson [19] summarizes his research: "My overall conclusion is that there was a dramatic impact of energy prices on economic growth during the energy crisis." Denison [21] disagrees and reasons: "Energy gets about 5 percent of the total input weight in the business sector ... If ... the weight of energy is 5 percent, a 1-percent reduction in energy consumption with no change in labor and capital would reduce output by 0.05 percent"; see also [22]. According to this argument, the decrease of energy input E in the US economy by 5.2 percent between 1973 and 1975 should have only caused a decrease of US output by 0.26 percent. The observed decrease of output, however, was 1.0 percent (see also Fig. 4 of Section 6.2). In most of the industrially advanced countries the share of energy cost in total factor cost is similar to that in the USA and is therefore almost negligible compared with the cost shares of capital K and labor L. From this perspective the recessions of the energy crises are hard to understand.[2]

Cost share weighting of production factors results from the *technologically unconstrained* equilibrium of neoclassical growth theory. In this equilibrium a factor's output elasticity and cost share in total factor cost are equal (see, e.g., eqs. (5) and (10) in Section 3).

In addition to the difficulties with explaining the energy crises, factor weighting by cost shares has another problem: the Solow residual, which accounts for that part of output growth that cannot be explained by the weighted input growth rates; attributing this difference formally to technological progress "has lead to a criticism of the neoclassical model: it is a theory of growth that leaves the main factor in economic growth unexplained" [23]. As a consequence considerable research efforts have been dedicated to breaking down the observed economic growth into components associated with changes in factor input and the interpretation of the Solow residual as a measure of technological change. Barro [24] summarizes the basics of standard growth accounting.

Endogeneous growth theory proposes a variety of approaches to specify technological progress and analyze its influence on growth. In addressing a few of them we concentrate on assumptions regarding production factors, production functions, and output elasticities in the growth models.

Barbier [25] combines Stiglitz' [26] neoclassical production function, which includes a natural resource, with Romer's [27] model of endogeneous technological change and investigates paths of optimal growth, using a production function of the Cobb-Douglas type with constant returns to scale, multiplied by the stock of knowledge. Welsch and Eisenack [28] start from Romer's model, too, and extend it "to examine the impact of secular changes in energy cost on technological progress and long run growth." Their production function is also Cobb-Douglas like with constant returns to scale and output elasticities of K, L, and E that result from the equilibrium equations between factor prices

[2] In the US economy, Fig. 4 in Section 6.2, the changes from 1973 to 1974 and from 1974 to 1975 were for capital +3.3% and +2.7%, for labor +0.6% and -2.8%, for energy -2.6% and -2.6%; if one multiplies these changes with the factor cost shares and adds the contributions from the two time intervals, the sum of the weighted changes of capital and labor cancel, and the two weighted energy changes add up to -0.26%. The total change of output was -0.6% -0.4% = -1%.

and marginal products. Differing from the approach of the present paper, all these models consider technologically unconstrained growth-in-equilibrium. This is also the case for the subsequently discussed growth models.

Bertola [29] studies distributional implications of growth-oriented policies, describing aggregate output by the product of a term representing disembodied productivity with a constant-returns-to-scale Cobb-Douglas production function of aggregate capital K and a factor L that "might refer to land or (uneducated) labor"; the output elasticities are given by the shares of K and L in aggregate income. The economy's rate of balanced growth is calculated for the case that the disembodied productivity term grows in such a way with capital that output becomes a linear function of K. Similarly, Howitt [30] constructs a multicountry endogeneous growth model starting from a set of Cobb-Douglas production functions with constant returns to scale, each one multiplied by a productivity parameter that grows as the result of innovations. In the model proposed below, which describes growth subject to technological constraints, innovations may show both in increases of productivity parameters (e.g. the energy efficiency of the capital stock) and in a relaxation of technological constraints.

Barro [24], after exploiting the duality of factor quantities and prices in growth accounting, considers firm-specific Cobb-Douglas production functions that depend on the firm's inputs of capital and labor (with constant returns to scale) and the economywide capital stock K. The economywide output, aggregated across firms, results as a Cobb-Douglas production function of the aggregated inputs capital K and labor L, with (the possibility of) increasing returns to scale. "The interpretation of K ... depends on the underlying model. Griliches [31] identifies K with knowledge-creating activities, such as R&D. Romer [32] stresses physical capital itself. Lucas [33] emphasizes human capital in the form of education." While these models aim at endogenizing the origins of productivity increases, employing various interpretations of capital in conjunction with increasing returns, we will stay with the concept of physical capital and constant returns, in combination with explicitly modeled technological constraints – which is new.

Labor is considered the only factor of production in studies that focus on endogeneous growth driven by either "specialization as a crucial ... aspect of human-capital accumulation" [34] or research and development, measured by the number of scientists and engineers engaged in R&D, expenditures by manufacturers for R&D, and patent grants [35]; see also [36]. On the other hand, Moser and Nicholas [37] look into patents and their citations for tracing the dynamics of technological progress and test whether electricity matches the criteria of general purpose technologies (GPTs), credited with generating the increasing returns that drive endogeneous growth. They find that inventions in other industries, such as chemicals (which require large amounts of process heat), fulfill the criteria for GPTs at least as well as those in electricity. Similarly, Jorgenson [18] notes that "much research remains to be done before we obtain a complete understanding of the role of energy utilization in productivity growth ... The data support the hypothesis that electrification and productivity growth are interrelated. Somewhat surprisingly, the data also show that the utilization of nonelectric energy and productivity growth are even more strongly related".

Like these studies we focus our attention on innovations, growth, technology, and energy. But unlike them, and any other work we are aware of, we proceed from economic equilibrium under technological constraints that affect the combinations of capital, labor and energy.

3. CONSTRAINED EQUILIBRIUM AND FACTOR SHADOW PRICES

We consider an economic systems with three factors of production X_1, X_2, X_3, which can vary independently within certain technological constraints. They form the input vector $\vec{X}$. X_1 is capital. The other two factors will be identified in Section 5 as (routine) labor and energy.[3] Non-routine human contributions to economic progress ("creativity") are taken into account by time-changing technology parameters of the production function[4] $Y(\vec{X})$ that will be used to describe output and growth in Section 6.

It is convenient to work with dimensionless, normalized variables. Normalized output is $y(\vec{x}) \equiv Y(\vec{X})/Y_0$ and the normalized inputs are $x_i \equiv X_i/X_{i0}$, $i = 1, 2, 3$, where Y_0 and the X_{i0} are output and inputs in the base year t_0.[5]

We assume, and show in Section 5, that the technological constraints, which limit the independent variations of the production factors at a given time t, can be written with the help of slack variables in the form

$$f_a(\vec{x}, t) = 0, \quad a = \alpha, \beta, \gamma, \ldots \tag{1}$$

The exogenously given prices per unit of factors X_1, X_2, X_3 are p_1, p_2, p_3. They form the price vector $\vec{p}(t)$. Thus, total factor cost is

$$FC \equiv \vec{p}(t) \cdot \vec{X}(t) = \vec{P}(t) \cdot \vec{x}(t) = \sum_{i=1}^{3} P_i(t) x_i(t), \quad \vec{P}(t) \equiv (p_1 X_{10}, p_2 X_{20}, p_3 X_{30}). \tag{2}$$

We derive the criteria for economic equilibrium under technological constraints via two different models of optimization by economic agents. The first, and somewhat more straightforward model used in this section, assumes that the decisions of all economic agents result in maximization of profit, which is the difference between the macroeconomic output Y and the total factor cost FC . The second model assumes that society maximizes the time integral of utility of consumption. It is outlined in Appendix A.

The necessary condition for a maximum of output minus cost subject to the constraint of fixed cost FC and the technological constraints (1) is that the gradient in factor space of $Y_0 y(\vec{x}) + \mu(FC - \vec{P} \cdot \vec{x}) + \sum_a \mu_a f_a(\vec{x}, t)$ vanishes. This yields the equilibrium conditions

$$Y_0 \frac{\partial y}{\partial x_i} - \mu \left[P_i - \sum_a \frac{\mu_a}{\mu} \frac{\partial f_a}{\partial x_i} \right] = 0, \quad i = 1, 2, 3 \quad , \tag{3}$$

where μ and μ_a are Lagrange multipliers.

The equilibrium values of x_1, x_2 and x_3 can be computed from these three equations, if one knows the production function, factor prices and constraints.

[3]By "energy" we mean an input that is converted into physical work, or which is physical work itself. In the engineering sciences exergy (with x) is the name for the maximum amount of physical work that could be obtained in principle from an energy source. Primary energy in the form of coal, oil, gas, nuclear fuels or solar radiation could be converted to nearly 100 percent into physical work under appropriate conditions. It is an input ultimately supplied by nature.

[4]The concept of the macroeconomic production function has been subject to criticism, summarized by Felipe and Fisher [38]. Lindenberger et al. [39] respond to this criticism elsewhere.

[5]If there were were more than three production factors, say n, one would have to replace $\sum_{i=1}^{3}$ by $\sum_{i=1}^{n}$ in the following equations.

Multiplication of eq. (3) with $\frac{x_i}{y}$ brings the equilibrium conditions into the form

$$\frac{x_i}{y}\frac{\partial y}{\partial x_i} = \frac{\mu}{Y_0 y} x_i \left[P_i - \sum_a \frac{\mu_a}{\mu}\frac{\partial f_a}{\partial x_i}\right], \qquad i = 1, 2, 3 \quad , \tag{4}$$

where

$$\frac{x_i}{y}\frac{\partial y}{\partial x_i} = \frac{X_i}{Y}\frac{\partial Y}{\partial X_i} \equiv \epsilon_i \tag{5}$$

also defines the output elasticity ϵ_i of the production factor x_i (X_i). Output elasticities will be calculated in Appendix B as approximate solutions of the set of partial differential equations that results from the requirement of twice differentiability of macroeconomic production functions.

We assume that the production function $y(\vec{x})$ is linearly homogeneous in (x_1, x_2, x_3) . Then we have constant returns to scale:[6]

$$\sum_{i=1}^{3} \epsilon_i = 1 \quad . \tag{6}$$

Combining the last three equations one obtains

$$Y_0 y = Y_0 y \sum_{i=1}^{3} \epsilon_i = \mu \sum_{i=1}^{3} x_i \left[P_i - \sum_a \frac{\mu_a}{\mu}\frac{\partial f_a}{\partial x_i}\right] = \mu \left[FC - \sum_{i=1}^{3} x_i \sum_a \frac{\mu_a}{\mu}\frac{\partial f_a}{\partial x_i}\right] \quad , \tag{7}$$

where the total factor cost FC is given by eq. (2). Equation (7) can also be written in the form

$$\mu = \frac{Y_0 y}{FC - \sum_{i=1}^{3} x_i \sum_a \frac{\mu_a}{\mu}\frac{\partial f_a}{\partial x_i}}. \tag{8}$$

We insert μ from eq. (8) into eq. (4) and observe eq. (5). This yields the equilibrium conditions in the form

$$\epsilon_i = \frac{x_i \left[P_i - \sum_a \frac{\mu_a}{\mu}\frac{\partial f_a}{\partial x_i}\right]}{FC - \sum_{i=1}^{3} x_i \sum_a \frac{\mu_a}{\mu}\frac{\partial f_a}{\partial x_i}}, \quad i = 1, 2, 3 \quad . \tag{9}$$

Without technological constraints all μ_a are absent in eq. (9). Then one obtains the non-constrained equilibrium conditions, which say that the output elasticities $\epsilon_{i,nc}$, are equal to the factor cost shares:

$$\epsilon_{i,nc} = \frac{x_i P_i}{FC} = \frac{X_i \cdot p_i}{FC} \quad . \tag{10}$$

For the general case with technological constraints we rewrite eq. (9):

$$\epsilon_i = \frac{x_i \left[P_i - \sum_a \frac{\mu_a}{\mu}\frac{\partial f_a}{\partial x_i}\right]}{\sum_{i=1}^{3} x_i \left[P_i - \sum_a \frac{\mu_a}{\mu}\frac{\partial f_a}{\partial x_i}\right]} = \frac{x_i X_{i0} \left[p_i - \frac{1}{X_{i0}} \sum_a \frac{\mu_a}{\mu}\frac{\partial f_a}{\partial x_i}\right]}{\sum_{i=1}^{3} x_i X_{i0} \left[p_i - \frac{1}{X_{i0}} \sum_a \frac{\mu_a}{\mu}\frac{\partial f_a}{\partial x_i}\right]} = \frac{X_i \left[p_i + s_i\right]}{\sum_{i=1}^{3} X_i \left[p_i + s_i\right]} \quad . \tag{11}$$

[6] At any fixed time t an increase of *all* inputs by the same factor λ must increase output by λ. Thus, the production function must be linearly homogeneous in (x_1, x_2, x_3), which means that $y(\lambda x_1, \lambda x_2, \lambda x_3) = \lambda \cdot y(x_1, x_2, x_3)$ for all $\lambda > 0$, all possible combinations (x_1, x_2, x_3) and all times t. Differentiating this equation with respect to λ according to the chain rule and then putting $\lambda = 1$ one obtains the Euler relation $(\partial y/\partial x_1) \cdot x_1 + (\partial y/\partial x_2) \cdot x_2 + (\partial y/\partial x_3) \cdot x_3 = y$. Dividing this by y yields $(x_1/y)(\partial y/\partial x_1) + (x_2/y)(\partial y/\partial x_2) + (x_3/y)(\partial y/\partial x_3) = 1$.

Here the shadow price of the production factor $X_i = X_{i0} x_i$ is defined as

$$s_i \equiv -\frac{1}{X_{i0}} \sum_a \frac{\mu_a}{\mu} \frac{\partial f_a}{\partial x_i} \quad . \tag{12}$$

It is due to the technological constraints $f_a(x_1, x_2, x_3, t) = 0, \quad a = \alpha, \beta, \gamma \ldots$.

The last equality in eq. (11) can also be written as

$$\epsilon_i = \frac{X_i(p_i + s_i)}{(FC + FC_S)}, \quad FC_S \equiv \sum_{i=1}^{3} X_i s_i \quad , \tag{13}$$

where FC_S is the total shadow cost of the input factors.

4. LAGRANGE MULTIPLIERS FOR THE CASE OF TWO CONSTRAINTS

In Section 5 we consider two specific technological constraints for the three factor model. Then the sum over subscripts a in the above equations contains only two terms, say α and β. It is convenient to abbreviate

$$\mu_A \equiv \frac{\mu_\alpha}{\mu}, \quad \mu_B \equiv \frac{\mu_\beta}{\mu}, \tag{14}$$

and define the partial derivatives of the two constraint equations

$$f_A(x_1(t), x_2(t), x_3(t), t) = 0, \quad f_B(x_1(t), x_2(t), x_3(t), t) = 0 \tag{15}$$

as

$$f_{Ai} \equiv \frac{\partial f_A}{\partial x_i}, \quad f_{Bi} \equiv \frac{\partial f_B}{\partial x_i}, \quad i = 1, 2, 3 \quad . \tag{16}$$

With that the equilibrum conditions (9) become

$$\epsilon_i = \frac{x_i \left[P_i - \mu_A f_{Ai} - \mu_B f_{Bi}\right]}{FC - \sum_{i=1}^{3} x_i \left[\mu_A f_{Ai} + \mu_B f_{Bi}\right]}, \quad i = 1, 2, 3 \quad . \tag{17}$$

We resolve the two independent ratios

$$\frac{\epsilon_1}{\epsilon_2} = \frac{x_1 \left[P_1 - \mu_A f_{A1} - \mu_B f_{B1}\right]}{x_2 \left[P_2 - \mu_A f_{A2} - \mu_B f_{B2}\right]} \tag{18}$$

and

$$\frac{\epsilon_1}{\epsilon_3} = \frac{x_1 \left[P_1 - \mu_A f_{A1} - \mu_B f_{B1}\right]}{x_3 \left[P_3 - \mu_A f_{A3} - \mu_B f_{B3}\right]}, \tag{19}$$

with respect to μ_A and μ_B. After some algebraic manipulations, and with the definitions

$$R_{21} \equiv \frac{x_2 \epsilon_1}{x_1 \epsilon_2}, \quad R_{31} \equiv \frac{x_3 \epsilon_1}{x_1 \epsilon_3}, \tag{20}$$

we obtain

$$\mu_A = \frac{(P_1 - P_2 R_{21})}{f_{A1} - f_{A2} R_{21}} + \frac{f_{B2} R_{21} - f_{B1}}{f_{A1} - f_{A2} R_{21}} \cdot \mu_B \quad , \tag{21}$$

$$\mu_B = \frac{(P_1 - P_3 R_{31})(f_{A1} - f_{A2} R_{21}) - (P_1 - P_2 R_{21})(f_{A1} - f_{A3} R_{31})}{(f_{B2} R_{21} - f_{B1})(f_{A1} - f_{A3} R_{31}) - (f_{B3} R_{31} - f_{B1})(f_{A1} - f_{A2} R_{21})} \quad . \tag{22}$$

5. CAPACITY UTILIZATION AND AUTOMATION.

We note from eqs. (10) and (11) that the standard equation of output elasticities with cost shares is valid only in the absence of constraints. The question arises: is that condition reasonable, or are there actually constraints on factor space that arise, for example, from technological relationships among the three factors of production? We postulate that such constraints do exist. Our premise is that capital $K(t)$, labor $L(t)$ and energy $E(t)$ are the factors of production that create value added Y at time t. The capital stock consists of all energy-conversion and information-processing devices together with all buildings and installations necessary for their protection and operation. Capital in the absence of energy (and supervision by humans) is functionally inert. Nothing happens. To be productive it must be activated by useful energy (exergy). Their is a minimum amount of exergy required in order that the machines of the capital stock can be productive.[7] To be economically productive it must also be allocated, organized and supervised by (human) labor.

Materials are passive partners of the production process. They do not contribute actively to the creation of value added. Capital, labor and energy arrange material atoms and molecules into the orderly patterns required for a useful product. Similarly, capital, labor and energy move electrons in materials as required for information processing. Land area matters mainly as site for production facilities of the industrial sector and for photosynthetic conversion of solar energy into the chemical energy of glucose in agriculture. Since usable land can be increased only a little if at all in industrial countries, and since land does not contribute actively to work performance and information processing, we disregard it when discussing industrial production and growth of the past. However, the limited emission-absorption capacity of the biosphere above the finite land area of earth is a constraint that will be felt more and more in the future. Since it depends on risk assessment by society and political weighting of the objectives "industrial growth" and "ecological stability" in welfare optimization, it is hard to express it mathematically by a constraint equation with slack variables. Instead, Kümmel [40] has proposed to model it by pollution functions that reduce output elasticities.

Economic activities of humans can be subdivided into two components: (1) routine labor, which (by definition) can be substituted by some combination of capital and energy and (2) a residual that cannot be replaced at any particular moment in time. The latter component could be given various names, but the one we prefer is creativity. Creativity, in this sense, is the specific human contribution to production (and growth) that cannot be provided by any machine, even a sophisticated computer capable of learning from experience. It includes ideas, inventions, valuations, and (especially) interactive decisions depending on human reactions and characteristics. It is important to recognize that the non-routine component of human labor may decline over time, but it is never zero.

The ultimate lower limit of labor inputs is probably un-knowable, because it depends to some degree on the limits of artificial intelligence. But we need not concern ourselves with the ultimate limit. At any given time, with a given technology, there is a limit to the extent that routine labor can increase output. In other words, we postulate a combination of capital and exergy such that adding one more unskilled worker adds nothing to gross

[7] Steam turbines, gas turbines, Diesel engines and petrol engines are the principal convertors of fossil and nuclear fuels into physical work. As a consequence of the Second Law of Thermodynamics their conversion efficiency can never be better than the Carnot efficiency.

economic output. (In some manufacturing sectors of industrialized countries this point does not seem to be far away.)

There is another fairly obvious technological constraint on the combinations of factors in factor space. In brief, machines are designed and built for specific exergy inputs. If the exergy available is less than the design point, production will be less than optimal. An example are power plants, whose conversion efficencies (generated electricity divided by fuel input) are reduced when operating in part load. In some cases (e.g. for some electric motors) there is a modest overload capability. Business buildings can be over-heated or over-cooled, to be sure, but it does not contribute to productivity. But on average the maximum exergy input is fixed by design. Both energy convertors and energy users have built-in limits. In other words, the ratio of exergy to capital must stay above a definite lower and below a definite upper limit.

Output Y and capital stock K, measured in deflated monetary units, are taken from the national accounts, routine labor L, measured in man hours worked per year, is given by the national labor statistics, and energy E, measured in petajoules (or tons of oil equivalents, or quads) per year, is obtained from the national energy balances. The prices p_K, p_L, p_E per units of capital, routine labor and energy can be calculated from data provided by economic research institutions and national statistics (see, e.g., [41]). (Aggregation of output and capital in physical units, defined by work performance and information processing, and the relation of these physical units to the monetary units has been discussed by Kümmel [42] and in [6, 7, 39].)

With these identifications the general variables of the preceding sections become:

$$\begin{aligned} y(x_1, x_2, x_3) &= y(k, l, e) \\ x_1 &= k = k(t) \equiv \frac{K(t)}{K_0}, \quad X_{10} = K_0 \\ x_2 &= l = l(t) \equiv \frac{L(t)}{L_0}, \quad X_{20} = L_0 \\ x_3 &= e = e(t) \equiv \frac{E(t)}{E_0}, \quad X_{30} = E_0 \quad . \end{aligned} \tag{23}$$

Here K_0, L_0 and E_0 are the production factors in the base year t_0.

The resume of the above considerations is that the use of capital, labor and energy in industrial production is subject to technological constraints that are the consequence of limits to capacity utilization and to the substitution of capital and energy for labor. We use the term “automation” to mean “degree of substitution of capital plus energy (exergy) for labor”. Entrepreneurial decisions, aiming at producing a certain quantity of output y within existing technology, result in the absolute magnitude of the total capital stock k, its degree of capacity utilization η and its degree of automation ρ.

The macroeconomic degree of capacity utilization η is defined as the appropriate average over the degrees of capacity utilization of the individual production units that make up the total capital stock.

The degree of automation is defined by

$$\rho = \eta \cdot \frac{k}{k_m(y)} \quad ; \tag{24}$$

$\rho = 1$, if at $\eta = 1$ the capital stock k is equal to the capital stock $k_m(y)$ that would be required for maximally automated production of output y; in this state the economic weight, i.e. the output elasticity, of routine labor would be vanishingly small [5] - [7].[8] Obviously, ρ and η are functions of capital k, labor l, and energy e. They are definitely constrained by $\rho(k,l,e) \leq 1$ and $\eta(k,l,e) \leq 1$, i.e. the maximum degree of automation (at a given time) cannot be exceeded, and a production system cannot operate above design capacity. Here it is important that (productive) energy input into machines and other capital equipment is always limited by their technical design.

However, there is a technical limit to automation that lies below 1. We call it $\rho_T(t)$. It depends on mass, volume and energy demand of machines, especially information processors, in the capital stock. Imagine the vacuum-tube computers of the 1960s, when the tiny transistor, invented in 1947 by Bardeen, Brattain and Shockley, had not yet diffused into the capital stock. A vacuum-tube computer with the computing power of a 2008 notebook computer would have had a volume of many thousands of cubic meters. In 1960 a degree of automation, as it is standard 40 years later in the highly industrialized countries[9], would have resulted in factories many orders of magnitude bigger than today, probably exceeding the available land area. The resources required to build such factories would have driven the cost of pushing automation to present-days standards far above the savings generated by the substitution of capital and energy for labor.

In the course of time, the technical limit to automation, $\rho_T(t)$, moves towards the theoretical limit, 1, as, e.g., the density of information processors on a microchip increases. According to "Moore's Law" this density has doubled every 18 months during the last decades. It may continue like that for a while, thanks to nano-technological progress. But there is a thermodynamic limit to transistor density, because the electricity required for information processing eventually ends up in heat; if this heat can no longer escape sufficiently rapidly out of the microchip, it will melt down the conducting elements of the integrated circuits and destroy them. We do not know exactly, how far the technical limit to automation can be pushed. For the purpose of the present paper, however, it is sufficient to know that this limit exists. Since the technical properties of the capital stock do not change with η, the corresponding constraint equation applies to the situation of maximum capacity utilization. It reads, with $\eta = 1$ in eq. (24) and the slack variable k_ρ,

$$f_A(k,l,e,t) \equiv \frac{k + k_\rho}{k_m(y)} - \rho_T(t) = 0 \quad ; \tag{25}$$

k_ρ is the capital stock that has to be added to k so that the total capital stock $k + k_\rho$, working at full capacity, exhausts the technologically possible automation potential $\rho_T(t)$.

With the slack variables l_η, e_η the constraint equation for capacity utilization becomes

$$f_B(k,l,e,t) \equiv \eta(k, l + l_\eta(t), e + e_\eta(t)) - 1 = 0 \quad ; \tag{26}$$

$l + l_\eta$ and $e + e_\eta$ are the quantities of labor and energy required for full capacity utilization of the capital stock k at time t. The technological state of the capital stock determines the relation between $e_\eta(t)$ and $l_\eta(t)$.

We need an explicit functional form for the degree of capacity utilization η. Since η does not change, if k, l and e all change by the same factor r, it is a homogeneous

[8] In prior publications the symbol k_t was used instead of k_m, and we (R.K. and D.L.) called "total automation" what we now call "maximum automation".

[9] Lindenberger [9] has computed an increase of automation in German industry by 50% between 1960 and 1990.

function of degree zero: $\eta = \eta(l/k, e/k)$. A trial form can be derived from a Taylor expansion of $\ln\eta[\ln(l/k), \ln(e/k)]$ around some point '0'$\equiv (\ln(l/k)_0, \ln(e/k)_0)$, up to first order in $\ln(l/k) - \ln(l/k)_0$ and $\ln(e/k) - \ln(e/k)_0$. This yields

$$\eta = \eta_0 \left(\frac{l}{k}\right)^{\lambda} \left(\frac{e}{k}\right)^{\nu}, \tag{27}$$

where λ and ν are the derivatives of $\ln\eta$ with respect to $\ln l/k$ and $\ln e/k$ in the point '0'. In analogy to Lindenberger [9] we propose to use eq. (27) as a phenomenological function for the degree of capacity utilization.

Knowledge of the production function $y(k, l, e)$ facilitates calculation of the capital stock $k_m(y)$ required for maximally automated production of a given quantity of output y that is actually produced by k, l and e. This quantity, which enters the constraint equation (25), will be calculated in Section 7. The technology parameters η_0, λ and ν can be determined from empirical data on capacity utilization.[10] One can also compute f_{Ak}, f_{Al}, f_{Ae} and f_{Bk}, f_{Bl}, f_{Be} from eq. (25) and the combination of eq. (26) with eq. (27). The output elasticities in eqs. (21) and (22) for μ_A and μ_B and in the equilibrium conditions (9) are also known, if one knows $y(k, l, e)$. Thus, calculation of the production function is all that is left to do, in order to solve the equilibrium problem under technological constraints in principle.

6. MACROECONOMIC PRODUCTION FUNCTIONS

6.1 Energy-dependent Cobb-Douglas and Linex functions.

In this section and the following ones we change the notation for output and production functions from $Y(K, L, E) = Y_0 y(k, l, e)$ to $Q(K, L, E) = Q_0 q(k, l, e)$.[11]

We calculate macroeconomic production functions $Q(K, L, E) = Q_0 q(k, l, e)$ from the capital-labor-energy-creativity (KLEC) model, outlined in Appendix B. This model is based on the law of diminishing returns and the conditions for twice differentiability of energy-dependent production functions that also satisfy the Euler relation, i.e. constant returns to scale. It consists of the growth equation (74), the set of partial differential equations (77) for the output elasticities of capital, $\alpha \equiv \epsilon_1$, and labor, $\beta \equiv \epsilon_2$, and the (asymptotic) boundary conditions (79) and (80) on α and β. According to the theory of partial differential equations the most general solutions of (77) are any differentiable functions of l/k and e/k; one could determine them uniquely – and thus compute the exact production function for any given economic system – if one knew β at all points on a certain boundary surface in k, l, e space and α on some boundary curve in that same space [40]. Then one would also know the exact output elasticity of energy, $\gamma = 1 - \alpha - \beta \equiv \epsilon_3$. However, such information is not, and never will be available. Therefore, one has to be content with approximations.

If one disregards the asymptotic boundary conditions (79) and (80), one may consider the trivial solutions of the differential equations (77). These are constant output elastici-

[10] Lindenberger [9] had actually performed the corresponding fitting procedure for a somewhat different phenomenlogical model of η.

[11] The reason is consistency with the notation used by three of us (R.K., J.S., D.L.) in previous publications. This notation is the one used in Paul A. Samuelson's "Volkswirtschaftslehre I, II (1975)", translated from the 1973 textbook "Economics" [43].

ties of capital, α_0, labor, β_0, and energy, $1-\alpha_0-\beta_0$. They result in the energy-dependent Cobb-Douglas function

$$q_{CDE}(k,l,e) = q_0 k^{\alpha_0} l^{\beta_0} e^{1-\alpha_0-\beta_0}. \tag{28}$$

If there were no technological constraints on factor combinations so that eq. (10) holds, and if one could approximate the factor cost shares by constants, then α_0, β_0, and $1-\alpha_0-\beta_0$ would be the (approximately) constant factor cost shares. In the presence of technological constraints, however, one cannot determine the output elasticities this way. Another procedure is required.

There are also technological reasons for looking beyond the Cobb-Douglas function. Most important is that the Cobb-Douglas function allows for practically all combinations of production factors, and thus asymptotically complete substitutability. However, machines don't run without energy conversion, and there are thermodynamic limits to energy conversion efficiency.[12] Furthermore, several billion of hard working people would be required in mayor industrial economies in order to deliver the amount of physical work that is just numerically equivalent to the physical work delivered by primary energy conversion in the capital stock.[13] Therefore, in modern economies there are absolute limits to the substitution of capital and labor for energy. The Cobb-Douglas function implies otherwise. Limits to substitution can be taken into account via the simplest, non-trivial, factor-dependent solutions of the partial differential equations (77) and their asymptotic boundary conditions (79) and (80). These solutions are the output elasticities

$$\alpha = a \cdot \frac{l+e}{k}, \quad \beta = a \cdot \left(c\frac{l}{e} - \frac{l}{k} \right), \quad \gamma = 1 - \alpha - \beta \quad . \tag{29}$$

The output elasticity of capital, α, approaches zero, if the capital stock increases more rapidly than labor and energy. This is the law of diminishing returns and describes the approach to the Schumpeter stationary state, where capital has a zero net productivity; see also p. 549 of [13]. The technology parameter a gives the weight, with which labor/capital and energy/capital combinations contribute to the output elasticity of capital. In this sense a can be interpreted as a measure of capital effectiveness. If innovations change this weight in time, a becomes time-dependent $a(t)$. It should increase, if work performance and information processing increase at unchanged l/k and e/k.

The negative term $-al/k$ in β results from eq. (78) as the mathematical consequence of the functional form of α in eq. (29). Thus, the law of diminshing returns, which is imprinted on α, leaves its mark on β, too. The first term in β, which must be positive and which can be only a function of l/e, has been chosen as the simplest function that yields decreasing β at increasing e. This describes the approach to the state of maximum automation, where $\beta = 0$, so that a marginal increase of labor would no longer contribute to the increase of output q. At a given output q the state of maximum automation would require the capital stock $k = k_m(q)$, introduced in Section 5, and an energy input $e = e_m \equiv ck_m(q)$. The technology parameter c may be interpreted as indicating the

[12] The limits are drawn by the First and Second Law of Thermodynamics. These are the most fundamental laws of nature. The First Law implies that work performance requires energy conversion. The Second Law states that entropy is produced by all processes that occur in finite times. Entropy production due to energy conversion depreciates energy quality and results in emissions of heat and particle currents.

[13] This can be seen by comparing the daily work calorie requirement of 2.9 kilowatthours for very heavy work load with average primary energy input per day.

energy demand of the fully employed, maximally automated capital stock. If innovations reduce/enhance the energy demand of fully employed capital, c becomes time-dependent $c(t)$ and decreases/increases.

If one inserts the output elasticities (29) into the growth equation (74) and integrates at fixed time t one obtains the (first) Linex production function[14]

$$q_{Lt} = q_0 e \exp\left[a(2 - \frac{l+e}{k}) + ac(\frac{l}{e} - 1)\right] , \tag{30}$$

which depends *lin*early on energy and *ex*ponentially on ratios of capital, labor and energy. The multiplier q_0 is a third integration constant, besides a and c.[15]

This function, proposed in [42], has been applied to the description of production in mayor industrial countries [5, 6, 7, 9, 44, 45, 11]. It reproduces economic growth and the first two energy crises satisfactorily. Recent results are shown in Section 6.2.

If the Linex function depends explicitly on time t via a time dependence of the technology parameters, one may also introduce an output elasticity of creativity, δ, by

$$\delta \equiv \frac{t-t_0}{q}\frac{\partial q}{\partial t} \quad = \quad \frac{(t-t_0)}{q_{Lt}}\left[\frac{\partial q_{Lt}}{\partial a}\frac{\mathrm{d}a}{\mathrm{d}t} + \frac{\partial q_{Lt}}{\partial c}\frac{\mathrm{d}c}{\mathrm{d}t} + \frac{\partial q_{Lt}}{\partial q_0}\frac{\mathrm{d}q_0}{\mathrm{d}t}\right]. \tag{31}$$

The output elasticities represent the weights, with which marginal relative changes of the production factors k, l, e and of time t contribute to the relative change of output q.[16] In this sense they measure the productive powers of capital, labor, energy, and creativity.

We assume that the economic actors choose only such factor combinations for which the marginal increase of an input will not cause a decrease of output. Thus, the output elasticities must be non-negative:

$$\alpha \geq 0, \quad \beta \geq 0, \quad \gamma = 1 - \alpha - \beta \geq 0 \quad . \tag{32}$$

Application of these conditions to the Linex output elasticities (29) yields the inequalities

$$\alpha = a\frac{l+e}{k} \geq 0, \quad \beta = al\left(\frac{c}{e} - \frac{1}{k}\right) \geq 0, \quad \gamma = 1 - a\frac{e}{k} - ac\frac{l}{e} \geq 0 \quad . \tag{33}$$

These relations allow certain predictions. For instance, imagine an economic system with decreasing energy inputs. The decrease may be due to resource, environmental or political obstacles to fossil fuel consumption, and insufficient development of technologies for the grand-scale exploitation of nuclear and solar potentials, including extraterrestrial options like solar power satellites. As a consequence, not all countries can afford any longer increasing or at least constant energy inputs. Let us assume that the engineers of the system have been so creative that the energy demand of the capital stock, c, has been driven down to its thermodynamic minimum. Then, within our model,

[14] Kümmel et al. [5] and Lindenberger [9] have related the Linex function to the energy-dependent Translog function. They also calculated more complicated output elasticities that solve the differential equations (77) and yield higher Linex functions. However, the simplest Linex function seems to be sufficient for the description of past economic growth.

[15] It may change in time with changing monetary valuation of the basket of goods and services that make up Q_0.

[16] The growth equation (74) in Appendix B shows this directly.

at fixed k the economic actors will adjust labor input l to values at or below the critical limit $l_C(e) = e(1 - ae/k)/ac$, above which γ would be negative.[17] This critical limit defines the range, in which employment of routine labor may vary as e decreases.

6.2 Computing output elasticities for Germany, Japan and the USA

Appendix C describes various methods that thave been used in order to determine the parameters a, c and q_0. Already the simplest case of fitting the Linex function with three constant parameters to the empirical time series of output reproduces the general trend of economic growth, and one obtains output elasticities for capital, labor and energy that are of the same order of magnitude as the ones in Table 2. However, the Durbin-Watson coefficients d_W of autocorrelation have been mostly below 1. The best d_W value, indicating the absence of autocorrelations, is 2. The closer one comes to $d_W = 2$, the more confident one may be that no important factor has been left out.

In order to see, whether a reduction of autocorrelation has a significant impact on the output elasticities of capital, labor and energy, we allow for time dependencies of the technology parameters and model them by logistic functions, which are typical for growth in complex systems and innovation diffusion. Let $p(t)$ represent either the capital-effectiveness parameter $a(t)$ or the energy-demand parameter $c(t)$. Its logistic differential equation

$$\frac{d}{dt}\left(p(t) - p_2\right) = p_3\left(p(t) - p_2\right)\left(1 - \frac{p(t) - p_2}{p_1 - p_2}\right) \tag{34}$$

has the solution

$$p(t) = \frac{p_1 - p_2}{1 + \exp\left[-p_3\left(t - t_0 - p_4\right)\right]} + p_2, \tag{35}$$

with the free (characteristic) coefficients $p_1, \ldots, p_4 \geq 0$. As an alternative to logistics we have also looked into Taylor expansions of $a(t)$ and $c(t)$ in terms of $t - t_0$ with a minimum of free coefficients.

The free coefficients of the logistic functions, or of the Taylor expansions, are determined by minimizing the sum of squared errors (SSE) computed from the empiricial time series of output, $q_{empirical}(t_i)$, and the Linex function $q_{Lt}(t_i)$ with the empirical time series of k, l and e as inputs:

$$\sum_{i=1}^{T}\left[q_{empirical}(t_i) - q_{Lt}(t_i)\right]^2 \quad , \tag{36}$$

subject to the constraints (33) of non-negative output elasticities. These constraints turn into the constraints on $a(t)$ and $c(t)$, or on k, l, e for given a and c:

$$0 \leq a(t) \leq a_{max}(t) \equiv \frac{k(t)}{l(t) + e(t)}, \quad \frac{e(t)}{k(t)} \equiv c_{min}(t) \leq c(t), \quad 0 \leq a(t)\left[\frac{e}{k} + c(t)\frac{l}{e}\right] \leq 1. \tag{37}$$

We use the Levenberg-Marquardt method of non-linear optimization [46] in combination with a new, self-consistent iteration procedure that helps avoid divergencies in the

[17] The maximum of $l_C(e)$ is at $e = e_C = k/2a$. Since 1960 e has been below $k/2a$ in Germany, Japan and the USA, according to the empirical values of e, k in Figs. 1- 4 and the magnitudes of a given in Fig. 1 and Table 1.

fitting procedure or convergence in a side minimum. More details are given in Appendix C.

German reunification on October 3, 1990, provides an interesting test of the KLEC model. The sudden merger of the planned economy of the former German Democratic Republic with the market economy of the Federal Republic of Germany (FRG) into what continues to be the Federal Republic of Germany was a result of political, social and economic decisions with far-reaching consequences. As it turns out, it is possible to model this working of "creativity" phenomenologically by just five free coefficients that enter the Taylor series expansion for $a(t)$ and the combination of step functions[18] for $c(t)$ in the model for Germany's total economy, see Fig. 1. For the other considered systems a and c are given by the logistics function (35), with $p_i \equiv a_i$ for $a(t)$ and $p_i \equiv c_i$ for $c(t)$, $i = 1, 2, 3, 4$. The free coefficients $\{a_i\}$ and $\{c_i\}$ are given in Table 1.

Table 1 The free coefficients of $a(t)$ and $c(t)$ in the logistics for the Federal Republic of Germany's industrial sector "Warenproduzierendes Gewerbe" (FRG I), the Japanese sector "Industries" (Japan I), and the total economy of the USA (USA TE).

System	a_1	a_2	a_3	a_4	c_1	c_2	c_3	c_4
FRG I	0.33	0.67	0.19	32	1	1.46	19.1	31
Japan I	0.16	0.2	1.87	20.1	2.75	0.45	0.86	14.61
USA TE	0.21	0.49	0.97	22.64	2.63	0.81	0.81	17.24

The reproduction of economic growth in Germany, Japan and the USA since 1960 and the time-averaged output elasticities of capital, labor, energy, and creativity are shown in Figures 1 – 4 and Table 2.[19] The asterisk at the value 0.12 of $\bar{\delta}$ for FRG I in Table 2 indicates that the very large derivative of the logistic function $c(t)$ in 1991 has been omitted when calculating the time average. This has been also done for FRG TE, because the derivative of the step function $c(t)$ does not exist in 1991. The uncertainty ranges in the output elasticities result from the law of error propagation and the variances of the free coefficients in eq. (35)– or the alternative Taylor series expansions – computed by the statistics program SAS 8.1. Since δ, defined in eq. (31), depends on the temporal variations of all free coefficients, the law of error propagation yields the largest uncertainties for $\bar{\delta}$.

The energy crises 1973-1975 and 1979-1981 are well reproduced and the residuals are small in Figs. 1 – 4. The time-averaged output elasticities shown in Table 2 are for labor much smaller and for energy much larger than the cost shares of these factors. On an OECD average the factor cost shares have been roughly 0.25 for capital, 0.70 for labor and 0.05 for energy during the considered time spans. Prior studies with one set of constant technology parameters before 1978 and another one after 1978 had yielded similar mean output elasticities [5, 6, 45]. The main effect of the more detailed modeling

[18]Step function results from the logistic (35) for $p_3 \to \infty$.

[19]These figures and the table refer to economic sectors whose output is essentially generated by work performance and information processing. Consequently, the residential sectors have not been included in the data on output and inputs. The system called "US Total Economy"consists of the sectors called "Private Industries" and "Government" in the Statistical Abstracts of the United States. For more details on data see [42] and the Internet Supplement to [7].

Table 2 Time-averaged Linex output elasticities of capital ($\bar{\alpha}$), labor ($\bar{\beta}$), energy ($\bar{\gamma}$), and creativity ($\bar{\delta}$), adjusted coefficient of determination R^2 and Durbin-Watson coefficient d_W for the Federal Republic of Germany's total economy (FRG TE) and its industrial sector "Warenproduzierendes Gewerbe" (FRG I),the Japanese sector "Industries" (Japan I), and the total economy of the USA (USA TE) during the indicated time spans.

System	FRG TE 1960-2000	FRG I 1960-99	Japan I 1965-92	USA TE 1960-96
$\bar{\alpha}$	0.38±0.09	0.37± 0.09	0.18± 0.07	0.51± 0.15
$\bar{\beta}$	0.15±0.05	0.11±0.07	0.09±0.09	0.14±0.14
$\bar{\gamma}$	0.47±0.1	0.52±0.09	0.73±0.16	0.35±0.11
$\bar{\delta}$	0.19±0.2	$0.12^* \pm 0.13$	0.14±0.19	0.10±0.17
R^2	1	0.996	0.999	0.999
d_W	1.64	1.9	1.71	1.46

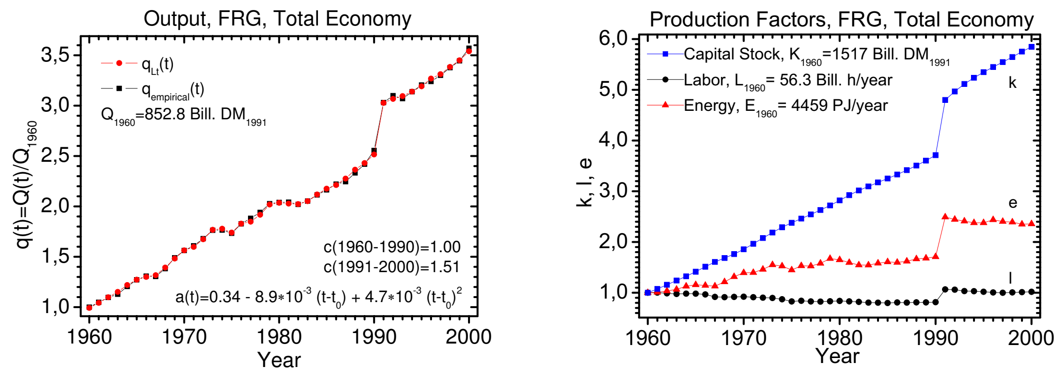

Figure 1: Left: Empirical growth (squares) and theoretical growth (circles) of the normalized output $q = Q/Q_{1960}$ of the total economy of the Federal Republic of Germany (FRG) between 1960 and 2000. Right: Empirical time series of the normalized factors capital $k = K/K_{1960}$, labor $l = L/L_{1960}$, and energy $e = E/E_{1960}$.

of the technology parameter's time dependence is the reduction of autocorrelation: the values of the Durbin-Watson coefficients d_W are closer to the best value 2 than in the prior studies. Furthermore, energy-dependent Cobb-Douglas functions with output elasticities close to the time-averaged Linex elasticities, reproduce observed economic growth rather satisfactorily, too, albeit with worse Durbin-Watson coefficients. In this sense the Cobb-Douglas function can also be used for an approximate description of past economic growth.

Ayres and Warr [11], using exergy inputs multiplied by appropriate conversion efficiencies (and physical work by animals) as energy variable, have fitted the Linex function with two constant technology parameters to the gross domestic product of the US economy between 1900 and 1998. The results are shown in Fig. 5. Their exergy data [47] already include most of the improvements in the efficiency of converting primary energy into useful work,[20] which have occurred in power plants, motors and other energy-consuming

[20] "Useful work", defined as the product of energy (exergy) inputs multiplied by a conversion efficiency, has been computed for some countries by Ayres et al. [47].

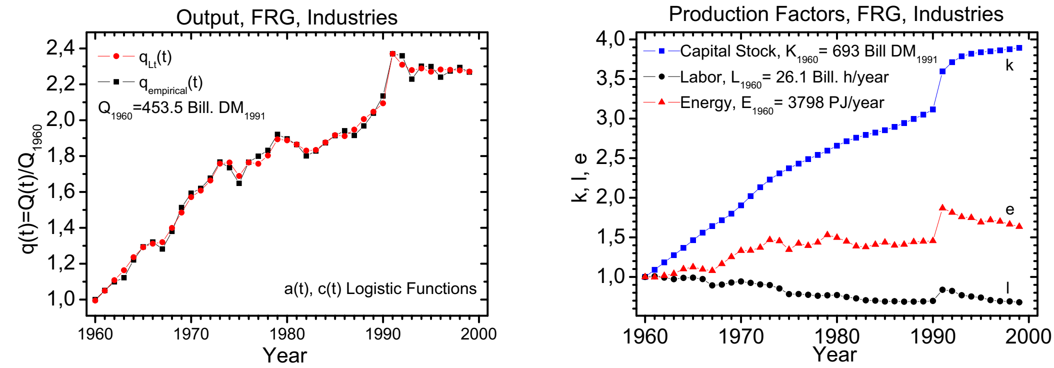

Figure 2: Left: Empirical growth (squares) and theoretical growth (circles) of the normalized output $q = Q/Q_{1960}$ of the German industrial sector "Warenproduzierendes Gewerbe" (GWG) between 1960 and 1999. Right: Empirical time series of the normalized factors capital $k = K/K_{1960}$, labor $l = L/L_{1960}$, and energy $e = E/E_{1960}$.

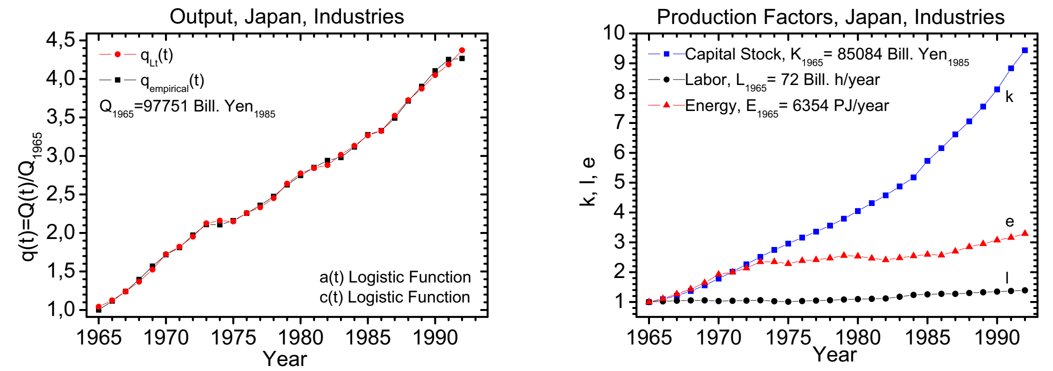

Figure 3: Left: Empirical growth (squares) and theoretical growth (circles) of the normalized output $q = Q/Q_{1965}$ of the Japanese sector "Industries" between 1965 and 1992. This sector produces about 90% of Japanese GDP. Right: Empirical time series of the normalized factors capital $k = K/K_{1965}$, labor $l = L/L_{1965}$, and energy $e = E/E_{1965}$.

devices during the 20th century. If, on the other hand, one uses primary energy input as energy variable, as it is done in the Linex functions that yield Figs. 1 – 4, one needs the time-dependent technology parameters, which model efficiency improvements in the use of primary energy and information processing. Most recent studies of Ayres and Warr [48] for the USA and Japan between 1900 and 2004, but excluding the years 1941-1948, show very good agreement between empirical and theoretical growth. Time-dependent Linex technology parameters have reduced residuals to much smaller values than in the best fit of Fig. 5.

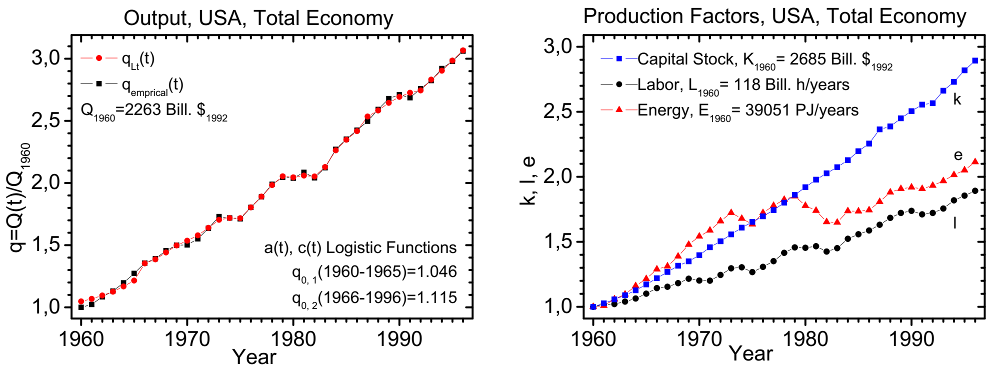

Figure 4: Left: Empirical growth (squares) and theoretical growth (circles) of the normalized output $q = Q/Q_{1960}$ of the total US economy between 1960 and 1996. Right: Empirical time series of the normalized factors capital $k = K/K_{1960}$, labor $l = L/L_{1960}$, and energy $e = E/E_{1960}$. Without recalibration of q_0 from q_{02} to q_{01} in 1965 agreement between the theoretical and the empirical growth curve would be somewhat worse during the first five years.

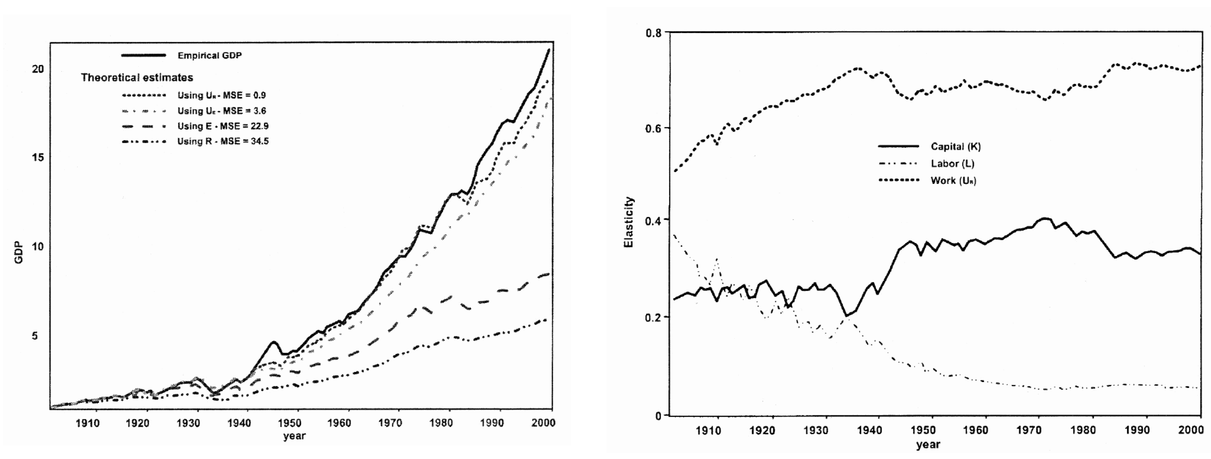

Figure 5: Left: Empirical growth (solid line) and theoretical growth (broken) lines of US GDP between 1900 and 1998. Theoretical growth has been calculated using the Linex production function with two constant technology parameters. The best fit, the dotted line, is obtained with physical work derived from *all* exergy inputs as energy variable. The other theoretical lines exclude some exergy sources, like, e.g. animal work. For more details see [11]. Right: Time-dependent Linex output elasticities of capital, solid line, labor, dashed-dotted line, and physical work, dotted line, for the US economy between 1900 and 1998. The time averaged output elasticities are for capital $\bar{\alpha} = 0.27$, labor $\bar{\beta} = 0.09$, physical work $\bar{\gamma} = 0.64$.

7. SHADOW PRICES OF CAPITAL, LABOR AND ENERGY.

The shadow prices of eq. (12) translate technological constraints into monetary terms. They are for capital, labor and energy

$$s_K = -\frac{1}{K_0}\left[\mu_A \frac{\partial f_A}{\partial k} + \mu_B \frac{\partial f_B}{\partial k}\right], \quad s_L = -\frac{1}{L_0}\left[\mu_A \frac{\partial f_A}{\partial l} + \mu_B \frac{\partial f_B}{\partial l}\right] \tag{38}$$

and

$$s_E = -\frac{1}{E_0}\left[\mu_A \frac{\partial f_A}{\partial e} + \mu_B \frac{\partial f_B}{\partial e}\right] \quad . \tag{39}$$

The (ratios of) Lagrange multipliers μ_A and μ_B are given by eqs. (21) and (22), where one has to identify $x_1 = k, x_2 = l, x_3 = e$, $\epsilon_1 = \alpha$, $\epsilon_2 = \beta$, $\epsilon_3 = 1 - \alpha - \beta$, and replace subscripts 1,2,3 by k, l, e. The output elasticities are those of eq. (29). (As a rough approximation one may also use the average output elasticities of Table 2, which correspond to energy-dependent Cobb-Douglas functions.) The functions f_A and f_B, which model the technological constraints, are given by eqs. (25) - (27). The capital stock for maximally automated production of output q, $k_m(q)$ in eq. (25) (with y changed to q), can be calculated from the Linex production function $q_{Lt}[k, l, e; t]$ of eq. (30) by demanding that

$$q_{Lt}[k, l, e; t] = q_{Lt}[k_m, l_m, ck_m; t], \tag{40}$$

where $c(t)k_m$ is the energy input in the state of maximum automation according to eq. (80). The routine labor l_m that remains in the state of maximum automation is certainly much smaller than k_m. If one neglects $l_m/k_m \ll 1$, eq. (40) becomes

$$q_0 e \exp\left[a(t)(2 - \frac{l+e}{k}) + a(t)c(t)(\frac{l}{e} - 1)\right] = q_0 c(t) k_m \exp\left[a(t)(2 - c(t)) - a(t)c(t)\right]. \tag{41}$$

This yields the capital stock for the maximally automated production of an output q that at time t is produced by the factors $k(t), l(t)$ and $e(t)$ as

$$k_m(q) = \frac{e}{c(t)} \exp\left[a(t)c(t)(1 + \frac{l}{e}) - a(t)\frac{l+e}{k}\right] \quad . \tag{42}$$

Thus, constraint equation (25), with the technical limit to automation $\rho_T(t)$ and the slack variable k_ρ, becomes

$$f_A(k, l, e, t) \equiv \frac{(k + k_\rho)}{k_m(q)} - \rho_T(t) = (k + k_\rho)\frac{c}{e} \exp\left[-ac(1 + \frac{l}{e}) + a\frac{l+e}{k}\right] - \rho_T(t) = 0 \quad . \tag{43}$$

The equation for the constraint on capacity utilization results from eqs. (26) and (27) as

$$f_B(k, l, e, t) \equiv \eta_0 \left(\frac{l + l_\eta(t)}{k}\right)^\lambda \left(\frac{e + e_\eta(t)}{k}\right)^\nu - 1 = 0 \quad . \tag{44}$$

According to eqs. (43) and (44) we have the slack-variable relations

$$k + k_\rho = k_m(q)\rho_T(t) \tag{45}$$

and

$$e + e_\eta = \frac{k}{\eta_0^{1/\nu} \left(\frac{l+l_\eta}{k}\right)^{\lambda/\nu}}. \tag{46}$$

The derivatives of f_A and f_B are obtained from eqs. (43)-(46) as

$$\frac{\partial f_A}{\partial k} = \frac{1}{k_m(q)} - a\frac{l+e}{k^2}\rho_T \tag{47}$$

$$\frac{\partial f_B}{\partial k} = -\frac{\lambda+\nu}{k} \tag{48}$$

$$\frac{\partial f_A}{\partial l} = -a\left(\frac{c}{e} - \frac{1}{k}\right)\rho_T \tag{49}$$

$$\frac{\partial f_B}{\partial l} = \frac{\lambda}{l+l_\eta} \tag{50}$$

$$\frac{\partial f_A}{\partial e} = \left(\frac{a}{k} + \frac{acl}{e^2} - \frac{1}{e}\right)\rho_T \tag{51}$$

$$\frac{\partial f_B}{\partial e} = \frac{\nu}{e+e_\eta} = \frac{\nu}{k}\eta_0^{1/\nu}\left(\frac{l+l_\eta}{k}\right)^{\lambda/\nu}. \tag{52}$$

The technology parameters a and c are given for Germany, Japan and the USA in Fig. 1 or by eq. (35). In the latter case $p_i \equiv a_i$ for $a(t)$ and $p_i \equiv c_i$ for $c(t)$, $i = 1, 2, 3, 4$, where the $\{a_i\}$ and $\{c_i\}$ are given in Table 1. In a rough approximation one may assume proportionality between the slack variables in the constraint on capacity utilization: $e_\eta(t) = d(t) \cdot l_\eta(t)$; $d(t)$ is the second constraint parameter besides $\rho_T(t)$. We call it the "labor-energy-coupling parameter at full capacity". Ideally, one should be able to determine it from measurements of the energy and labor increases required in order to go from any degree of capacity utilization to 1. Then l_η can be calculated from eq. (46). The multiplier η_0 and the exponents λ and ν may be obtained by fitting the phenomenological η of eq. (27) to empirical time series of η, which are available from economic research institutions.[21] The technical limit $\rho_T(t)$ to the degree of automation can be any number between 0 and 1. General business inquiries may give clues to it. Alternatively, one has to compute the time series of the shadow prices (38) and (39) for a number of scenarios for $\rho_T(t)$. Finally, one needs the deflated time series of factor prices p_K, p_L and p_E. Then the problem of economic equilibrium under technological constraints is solved.

8. SUMMARY AND CONCLUSIONS

Optimization of profit, or of integrated utility, yields new equilibrium conditions for economic systems subject to technological constraints. In a model, where capital, labor and energy (physical work) are the factors of production, the equilibrium values of the factors are determined by the conditions that the output elasticity of each factor is equal to its "shadowed" factor cost share. The shadowed cost share of a factor is the product of this factor with a two-component price term, divided by the sum over all factors times their respective price terms. Each price term consists of the factor (market) price plus a shadow price, which is due to the technological constraints. The constraints affect the degree of

[21] Lindenberger [9] used data from the "Sachverständigenrat für die Gesamtwirtschaft" when computing η_0, λ and ν for a somewhat different phenomenological model of η.

automation and the degree of capacity utilization. There are two constraint parameters: the technical limit to the degree of automation and the labor-energy-coupling parameter at full capacity. The latter enters a phenomenological function for capacity utilization. Exponents of factor ratios in this function can be obtained by fitting to empirical time series of capacity utilization.

The output elasticities have been determinded as functions of capital, labor and energy by econometric estimations of the Linex production function for Germany, Japan and the USA. Alternatively, as proposed by Lindenberger et al. [39], output elasticities of a macroeconomic production function may be obtained from equating this function to the sum of known microeconomic production functions. The latter depend on capital and labor coupled to energy-dependent productivity factors.

Cointegration analysis by Stresing et al. [49] has confirmed the order of magnitude of the Linex output elasticities in Table 2.

We emphasize that the suggested approach to modeling production takes into account explicitly that capital is utilized through energy and labor inputs. The conceptual basis is the notion that output is generated via work performance and information processing through the interaction of the factors capital, labor, and energy. Consequently, the utilization rate of capital is endogenous in the model (depending on the ratios of labor to capital and energy to capital), and the relevant output quantity is actual output – not potential output. This differs from alternative interpretations of production functions as production possibility frontiers, which consider potential output (output at full capacity), and thus employ – and have to construct for the purpose of numerical model application – utilization-adjusted factor inputs. The latter, of course, requires additional assumptions that are not needed in the present model, since capacity utilization is endogenous. Production functions that endogenize capacity utilization are the appropriate tool for estimating the economic impact of *actual* factor inputs.

Another feature of the proposed model, which differs from the concept of production possibility frontiers, is that it makes more explicit that component of technical progress which we call automation, i.e. the substitution of capital and energy for (routine) labor. At a given point in time at given state of technology, automation (depending on the ratios of labor and energy to capital) is limited by the automation potential $\rho_T(t)$. This represents a technological constraint that may be released in the course of technological progress.

If one wants to verify the evolution of past outputs and inputs according to the constrained equilibrium conditions, one needs time series for factor prices and constraint parameters. Business inquiries and technical analyses of the capital stock may provide the data for the technological constraints. Alternatively, one migth assume that in the past the economies have operated in constrained equilibria. Then one can use the empirical time series of inputs and factor prices in the equilibrium conditions and construct time series for the constraint parameters in the shadow prices. Extrapolation of these time series into the future, and guesses about the evolution of the technology parameters in the Linex output elasticities, will then allow predictions of economic growth within scenarios including market prices of capital, labor and energy.

Acknowledgement: The authors are grateful to Gerald Silverberg for stimulating discussions.

Appendix A: Intertemporal utility optimization subject to constraints

In their paper "A complete capital model involving heterogeneous capital goods", Samuelson and Solow [13] review Ramsey's problem [12] via their equation (1), on which they comment: "In words, society maximizes the (undiscounted) integral of all future utilities of consumption subject to the fact that the sum of current consumption and of current capital formation is limited by what the current capital stock can produce." Extending Ramsey's one-capital-good theory to their many-goods model they continue with maximizing undiscounted integrals of utility. One could argue that point. We don't, but rather, for the sake of simplicity, follow the optimization procedure of their Section I with the following modifications. 1. There is not one varying factor of production but three of different nature, which we label X_1, X_2, X_3. 2. There are contraints on magnitudes and combinations of these factors. 3. Utility optimization is done within finite time horizons, like the ones between the years 1960 and 2000. Thus, on our second way of deriving the equilibrium conditions (3) we assume that society has maximized the (undiscounted) integral W of utility U of consumption C between the times t_0 and t_1,[22] where the sum of consumption and capital formation at any time t between t_0 and t_1 is limited by what the capital stock $X_1(t)$ in combination with other production factors $X_2(t)$ and $X_3(t)$ can produce with due regard of the contraints on these factors. Thus, we have the optimization problem:
Maximize

$$W[s] = \int_{t_0}^{t_1} U[C]dt, \tag{53}$$

subject to constraints.

$W[s]$ is a functional of the curve $[s]$ along which the production factors evolve. This curve depends on the variables that enter C. In general, utility may depend on many variables. In the present case the utility function $U[C]$ depends on output minus capital formation.

Output (per unit time) is described by the macroeconomic production function $Y(\vec{X})$, where $\vec{X} \equiv (X_1, X_2, X_3)$. Part of Y goes into consumption C and the rest into new capital formation $\dot{X}_1 \equiv \frac{dX_1}{dt}$ plus replacement of depreciated capital. As usual we approximate the annual replacement rate by $\delta^d X_1$, where δ^d is the depreciation rate.[23] Then consumption (per unit time) is

$$C = Y(\vec{X}) - \dot{X}_1 - \delta^d X_1 \quad . \tag{54}$$

For the optimization of W we need the vector of the prices p_i per unit of factor X_i. This is $\vec{p} \equiv (p_1, p_2, p_3)$. Economic research institutions provide the price of capital utilization p_1 as the sum of net interest, depreciation and state influences. Furtheron we use this price. Since it already includes depreciation, we can omit explicit reference to the depreciation

[22] Modern notation for the integral to be maximized is W. Samuelson and Solow call it J. We define output Y as compared to their notation $f(S)$. Their term S seems to be what they call "abstract capital substance". The rate of change dS/dt must be interpreted as the sum of investment in new capital formation plus replacement of depreciated capital.

[23] The year is the natural time unit, because the annual cycle of seasons is decisive for agriculture, and important for construction. It also structures education, vacations (hence tourism and transportation) and some other industrial activities in the moderate climate zones. Thus, for practical purposes $Y(\vec{X})$ and $\dot{X}_1 + \delta^d X_1$ are annual output and annual capital formation, respectively.

rate, i.e. the term $\delta^d X_1$ in eq. (54), hereafter.[24]

As in Section 3 we work with dimensionless, normalized variables:

$$y(\vec{x}) \equiv \frac{Y(\vec{X})}{Y_0}, \quad x_i \equiv \frac{X_i}{X_{i0}}, \quad i = 1, 2, 3 \quad , \tag{55}$$

where Y_0 is the output and the X_{i0} are the inputs in the base year t_0. With that eq. (54) for consumption (without the term $\delta^d X_1$) becomes

$$C(\vec{x}, \dot{x_1}) = Y_0 y(\vec{x}) - X_{10}\dot{x_1} \quad . \tag{56}$$

The magnitude of the factors is constrained in the maximization of W by the requirement that total factor cost

$$\vec{p}(t) \cdot \vec{X}(t) = \vec{P}(t) \cdot \vec{x}(t) = \sum_{i=1}^{3} P_i(t) x_i(t), \quad \vec{P}(t) \equiv (p_1 X_{10}, p_2 X_{20}, p_3 X_{30}), \tag{57}$$

has finite magnitudes $FC(t)$.

Let there be other, technological constraints on the (normalized) facor inputs x_i that limit the technically accessible factor space. With the help of slack variables they can be written in the form

$$f_a(\vec{x}, t) = 0, \quad a = \alpha, \beta, \gamma, \ldots \tag{58}$$

Then, with the (generally time dependent) Lagrange multipliers μ, μ_a, the optimization problem becomes:
Maximize

$$W[s] = \int_{t_0}^{t_1} dt \left\{ U[C(\vec{x}, \dot{x_1})] + \mu(FC(t) - \vec{P} \cdot \vec{x}) + \sum_a \mu_a f_a(\vec{x}, t) \right\} \quad . \tag{59}$$

$W[s]$ is a functional of the curve $[s] = \{t, \vec{x} : \vec{x} = \vec{x}(t), \quad t_0 \le t \le t_1\}$. Consider another curve $[s, \vec{h}] = \{t, \vec{x} : \vec{x} = \vec{x}(t) + \vec{h}(t), \quad t_0 \le t \le t_1\}$ close to $[s]$, which goes through the same end points so that $\vec{h}(t_1) = 0$ and $\vec{h}(t_0) = 0$. Its functional is

$$W[s, \vec{h}] = \int_{t_0}^{t_1} dt \left\{ U[C(\vec{x} + \vec{h}, \dot{x_1} + \dot{h}_1)] + \mu \left(FC(t) - \vec{P} \cdot (\vec{x} + \vec{h}) \right) + \sum_a \mu_a f_a(\vec{x} + \vec{h}, t) \right\} \quad . \tag{60}$$

Since $\vec{h}$ is small, the integrand can be approximated by its Taylor expansion up to first order in $\vec{h}$ and $\dot{h}_1$. The necessary condition for a maximum of W is that the variation of W with respect to $\vec{h}$ vanishes:

$$\delta W \equiv W[s, \vec{h}] - W[s] = \int_{t_0}^{t_1} dt \left\{ \delta U - \mu \vec{P} \cdot \vec{h} + \sum_a \mu_a \sum_{i=1}^{3} \frac{\partial f_a}{\partial x_i} h_i \right\} = 0 \quad . \tag{61}$$

With the chain rule one obtains

$$\delta U \equiv U[C(\vec{x} + \vec{h}, \dot{x_1} + \dot{h}_1)] - U[C(\vec{x}, \dot{x_1})] = \frac{\partial U}{\partial C} dC = \frac{\partial U}{\partial C} \left[\sum_{i=1}^{3} \frac{\partial C}{\partial x_i} h_i + \frac{\partial C}{\partial \dot{x_1}} \dot{h}_1 \right] \quad . \tag{62}$$

[24] If we keep this term in the following optimization procedure, we would have a term proportional to $\delta^d X_1$ added to p_1 everywhere.

Partial integration yields

$$\begin{aligned}
\int_{t_0}^{t_1} dt \frac{dU}{dC}\frac{\partial C}{\partial \dot{x_1}}\dot{h}_1 &= \left[\frac{dU}{dC}\frac{\partial C}{\partial \dot{x_1}}h_1(t)\right]_{t_0}^{t_1} - \int_{t_0}^{t_1} dt \frac{d}{dt}\left(\frac{dU}{dC}\frac{\partial C}{\partial \dot{x_1}}\right)h_1(t) \\
&= -\int_{t_0}^{t_1} dt \frac{d}{dt}\left(\frac{dU}{dC}\frac{\partial C}{\partial \dot{x_1}}\right)h_1(t), \qquad (63)
\end{aligned}$$

because $h(t_1) = 0 = h(t_0)$. Combination of eqs. (61) - (63) results in

$$\begin{aligned}
\delta W &= \int_{t_0}^{t_1} dt \left\{\frac{dU}{dC}\sum_{i=1}^{3}\frac{\partial C}{\partial x_i}h_i - \frac{d}{dt}\left(\frac{dU}{dC}\frac{\partial C}{\partial \dot{x_1}}\right)h_1(t) - \mu\sum_{i=1}^{3}P_i h_i + \sum_a \mu_a \sum_{i=1}^{3}\frac{\partial f_a}{\partial x_i}h_i\right\} \\
&= 0 \quad . \qquad (64)
\end{aligned}$$

Since the small h_i are arbitrary for $t_0 < t < t_1$, the integral can only vanish, if the coefficients of the h_i vanish in the integrand. This yields the following conditions for $\delta W = 0$, i.e. for equilibrium of the economic system:

$$\frac{dU}{dC}\frac{\partial C}{\partial x_1} - \frac{d}{dt}\left(\frac{dU}{dC}\frac{\partial C}{\partial \dot{x_1}}\right) - \mu P_1 + \sum_a \mu_a \frac{\partial f_a}{\partial x_1} = 0 \qquad (65)$$

$$\frac{dU}{dC}\frac{\partial C}{\partial x_2} - \mu P_2 + \sum_a \mu_a \frac{\partial f_a}{\partial x_2} = 0 \qquad (66)$$

$$\frac{dU}{dC}\frac{\partial C}{\partial x_3} - \mu P_3 + \sum_a \mu_a \frac{\partial f_a}{\partial x_3} = 0 \quad . \qquad (67)$$

(Identifying $U[C(\vec{x}, \dot{x}_1)]$ with the Lagrangian $L(\vec{x}, \dot{x}_1)$, one notes the formal equivalence of these equations with the constrained Lagrange equations of motion in classical mechanics.)

With C from eq. (56) the last three equations turn into

$$\frac{dU}{dC}Y_0\frac{\partial y}{\partial x_1} - \mu P_1 + \sum_a \mu_a \frac{\partial f_a}{\partial x_1} = -X_{10}\frac{d}{dt}\left(\frac{dU}{dC}\right), \qquad (68)$$

$$\frac{dU}{dC}Y_0\frac{\partial y}{\partial x_2} - \mu P_2 + \sum_a \mu_a \frac{\partial f_a}{\partial x_2} = 0, \qquad (69)$$

$$\frac{dU}{dC}Y_0\frac{\partial y}{\partial x_3} - \mu P_3 + \sum_a \mu_a \frac{\partial f_a}{\partial x_3} = 0. \qquad (70)$$

Equations (68) – (70) are the the general equilibrium conditions for an economic system subject to cost limits and technological constraints. (For $\mu = 0 = \mu_a$ they correspond to eq. (2) of Samuelson and Solow [13]. If y does not depend explicitly on time t, they imply the conservation law $U + \frac{dU}{dC}\dot{x}_1 X_{10}$ = constant. The conserved Legendre transform of utility, $U + \frac{dU}{dC}\dot{x}_1 X_{10}$, corresponds to the Hamiltonian in classical mechanics.)

In general one assumes decreasing marginal utility. A special case is $U[C] = \ln C$. If in this case one approximates $\ln C$ by its Taylor expansion up to first order in $C - 1$ for sufficiently small $C - 1$,[25] one has

$$U[C] \approx C - 1, \quad \to \frac{dU}{dC} = 1, \quad \to \frac{d}{dt}\left(\frac{dU}{dC}\right) = 0 \quad . \qquad (71)$$

[25] A linear approximation of $\ln x$ is an acceptable approximation for $x < 4$.

With that the equilibrium conditions (68) – (70) can be summarized by

$$Y_0 \frac{\partial y}{\partial x_i} - \mu \left[P_i - \sum_a \frac{\mu_a}{\mu} \frac{\partial f_a}{\partial x_i} \right] = 0, \quad i = 1, 2, 3 \quad . \tag{72}$$

These conditions are the same ones as in eq. (3), derived from profit maximization.

Appendix B: KLEC model and Linex function

We change notation for the reasons explained at the beginning of Section 6: the symbols Y, Y_0 and y for output are replaced by Q, Q_0 and q.

An infinitesimal change of (normalized) output, $\mathrm{d}q$, is related to the infinitesimal changes of (normalized) capital, $\mathrm{d}k$, labor, $\mathrm{d}l$, energy, $\mathrm{d}e$, and time, $\mathrm{d}t$, by the total differential of the production function $q[k, l, e; t]$:

$$\mathrm{d}q = \frac{\partial q}{\partial k} \cdot \mathrm{d}k + \frac{\partial q}{\partial l} \cdot \mathrm{d}l + \frac{\partial q}{\partial e} \cdot de + \frac{\partial q}{\partial t} \cdot \mathrm{d}t.$$

Dividing the total differential by q and using the abbreviations

$$\alpha \equiv \frac{k}{q} \frac{\partial q}{\partial k}, \quad \beta \equiv \frac{l}{q} \frac{\partial q}{\partial l}, \quad \gamma \equiv \frac{e}{q} \frac{\partial q}{\partial e}, \quad \delta \equiv \frac{t - t_0}{q} \frac{\partial q}{\partial t}, \tag{73}$$

one obtains the growth equation

$$\frac{\mathrm{d}q}{q} = \alpha \cdot \frac{\mathrm{d}k}{k} + \beta \cdot \frac{\mathrm{d}l}{l} + \gamma \cdot \frac{\mathrm{d}e}{e} + \delta \cdot \frac{\mathrm{d}t}{t - t_0}. \tag{74}$$

The quantities defined by eq. (73) are the output elasticities of capital, α, labor, β, energy, γ, and creativity, δ . They give the weights by which the marginal relative changes of the production factors and of time contribute to the marginal relative change of output. In this sense they measure the productive powers of capital, labor, energy, and creativity.

Since negative output elasticities do not make sense economically, the elasticities are subject to the constraints

$$\alpha \geq 0, \quad \beta \geq 0, \quad \gamma = 1 - \alpha - \beta \geq 0 \quad . \tag{75}$$

The usual assumption is made that the production function must be twice continuously differentiable with respect to the production factors. Then the conditions

$$\frac{\partial^2 q}{\partial k \partial l} = \frac{\partial^2 q}{\partial l \partial k}, \quad \frac{\partial^2 q}{\partial k \partial e} = \frac{\partial^2 q}{\partial e \partial k}, \quad \frac{\partial^2 q}{\partial l \partial e} = \frac{\partial^2 q}{\partial e \partial l} \tag{76}$$

must be satisfied. Combination of these conditions with eq. (73) and the condition for constant returns to scale, $\alpha + \beta + \gamma = 1$, eq. (6), yields three coupled differential equations for the output elasticities [40]:

$$k \frac{\partial \alpha}{\partial k} + l \frac{\partial \alpha}{\partial l} + e \frac{\partial \alpha}{\partial e} = 0, \qquad k \frac{\partial \beta}{\partial k} + l \frac{\partial \beta}{\partial l} + e \frac{\partial \beta}{\partial e} = 0, \qquad l \frac{\partial \alpha}{\partial l} = k \frac{\partial \beta}{\partial k}. \tag{77}$$

One can easily verify that the most general solutions of these equations are

$$\alpha = A\left(\frac{l}{k}, \frac{e}{k}\right), \qquad \beta = \int \frac{l}{k} \frac{\partial A}{\partial l} \mathrm{d}k + J\left(\frac{l}{e}\right), \tag{78}$$

where A and J are any continously differentiable functions of their arguments $l/k, e/k$, and $l/e = (l/k)/(e/k)$. The output elasticities of any mathematically reasonable constant-returns-to-scale production function must satisfy eqs. (77) and (78).

The number of solutions represented by the general form (78) is infinite. In order to obtain the unique solution that holds for a given economic system at a given time one needs the proper boundary conditions for the differential equations (77). The number and nature of the data required for these boundary conditions is found with the help of the "characteristic equations" and the "characteristic basis curves" related to (77). This has been analyzed in detail in [40]. The result of this analysis is: in order to determine the output elasticities uniquely one has to know i) β in all points on a certain boundary surface in K, L, E space[26] and ii) α on a boundary curve. Obviously, the required wealth of empirical data is not – and never will be – available, because we cannot do experiments with economic systems that put them successively in a huge number of points in K, L, E space. Therefore, there is no other way than guess and test output elasticities that are of the general form (78). Simplicity and plausibility are conventional-wisdom guidelines for such guess work.

The simplest elasticities are the trivial, constant solutions of eqs. (77): $\alpha = \alpha_0, \beta = \beta_0$. Inserting them into eq. (74), observing constant returns to scale, eq. (6), and integrating the equation of growth one obtains the energy-dependent Cobb-Douglas function (28).

The law of diminishing returns, "this famous technological-economic relation" [43], led Kümmel [42] to factor-dependent output elasticities. According to this law a small increase of a huge capital stock, which is operated and activated by relatively small quantities of labor and energy, will contribute practically nothing to output growth. Therefore, the simplest, factor-dependent output elasticity of capital should satisfy the asymptotic boundary condition

$$\alpha \rightarrow 0 \quad \text{for} \quad \frac{l+e}{k} \rightarrow 0 \quad . \tag{79}$$

Similarly, an additional unit of labor will practically contribute nothing to output growth, when the economy approaches the state of maximum automation, where the output q is produced by the fully employed, maximally automated capital stock $k_m(q)$ and the required energy $e_m = ck_m(q)$. Thus, the output elasticity of labor, β, should satisfy the asymptotic boundary condition

$$\beta \rightarrow 0 \quad \text{for } k \rightarrow k_m \text{ and } e \rightarrow e_m = ck_m \quad . \tag{80}$$

The simplest output elasticities that satisfy these asymptotic boundary conditions and the differential equations (77) are

$$\alpha = a \cdot \frac{l+e}{k} \quad \text{and} \quad \beta = a \cdot \left(c\frac{l}{e} - \frac{l}{k} \right) \quad . \tag{81}$$

We can understand a to incorporate a (possibly time-dependent) equivalence factor, which relates capital's monetary value to its technological value in terms of its capability of work performance and information processing [7].

[26]This surface can be built up by all so-called work functions $L_\eta(K, E)$, if one varies the degree of capacity utilization η between 0 and 1 and "puts one $L_\eta(K, E)$ besides the other". Each $L_\eta(K, E)$ is determined by a differential equation for *fixed* η and an appropriate boundary curve.

If one inserts these output elasticities into the growth equation (74) at fixed t, integrates its left-hand side from q_0 to q and the right-hand side along a convenient path in k, l, e space, such as $(k = 1, l = 1, e = 1) \to (k, 1, 1) \to (k, l, 1) \to (k, l, e)$, (incorporating in q_0 a possibly time-dependent equivalence factor, which relates output's monetary value to its technological value) one obtains the (first) Linex production function

$$q_{Lt}[k, l, e; t] = q_0 e \exp\left[a\left(2 - \frac{l+e}{k}\right) + ac\left(\frac{l}{e} - 1\right)\right], \tag{82}$$

which depends *lin*early on energy and *ex*ponentially on the quotients of capital, labor, and energy.

Appendix C: Determining the technology parameters: from point fitting to non-linear optimization.

Since the boundary surface for β and the boundary curve for α that would result from the method of the characteristics are unknown even in the vicinity of $\beta \to 0$ and $l/k \to 0$, $e/k \to 0$, the technology parameters a and c (and q_0 as well) must be determined by fitting the Linex function q_{Lt}, with k, l, e from empirical time series, to the empirical time series $q_{empirical}$ of output, subject to the constraints of non-negative output elasticities, eq. (75).

Given this situation, various methods of fitting have been used. The simplest one, fitting q_{Lt} to $q_{empirical}$ in three subsequent years, was done for the West German total economy and its industrial sector "Warenproduzierendes Gewerbe" from 1960 to 1978. The economic recession of the first energy crisis and the subsequent recovery were reproduced and residuals were small. This method did not work for the sector "Industries" of the USA, 1960-1978: because of the nearly parallel rise of k and l in the USA between 1960 and 1973, not observed in other countries like Germany and Japan, the fit equations for a and c involved quotients of small differences of large numbers, resulting in a wide spectrum for a and c within the error margins of k, l, e. This well known problem of collinearity in the USA makes fitting for the USA always difficult and leads some people to question fitting in principle. But, fortunately, collinearity is not such a problem in general. The problem was dealt with (not very elegantly, indeed) by playing around with the constant a and c until the residuals and the reproduction of the energy crisis were comparable with the results for Germany [42]. Still, in all systems considered there were systematic deviations of q_{Lt} from $q_{empirical}$.

Subsequently, the technology parameters were determined by minimizing the Sum of Squared Errors (SSE)

$$\sum_{i=1}^{T} \left[q_{empirical}(t_i) - q_{Lt}(t_i)\right]^2, \tag{83}$$

subject to the constraints (75) of non-negative output elasticities. The a and c obtained this way reduced the systematic deviations of the theoretical from the empirical growth curves, first for West Germany, 1960-1981, and the USA, 1960-1978, and then for West Germany, 1960-1989, and Japan, 1965-1992. Recalibrating the technology parameters in 1978, such as to increase the capital effectiveness parameter a and decrease the energy demand parameter c in 1978, improved things further for West Germany, 1960-1989, Japan, 1965-1992, and the USA, 1960-1993 without changing the time-averaged Linex output elasticities $\bar{\alpha}$, $\bar{\beta}$ and $\bar{\gamma}$ significantly: the elasticity $\bar{\beta}$ of labor does not exceed the order of 0.2, that of energy $\bar{\gamma}$ is above 0.4, and only capital's elasticity $\bar{\alpha}$, with magnitudes between 0.34 and 0.45, is roughly in equilibrium with capital's cost share [5, 6, 45]. From Linex-type service production functions Lindenberger [44] finds $\bar{\alpha} = 0.54$, $\bar{\beta} = 0.29$ and $\bar{\gamma} = 0.17$ for the German service sector "Marktbestimmte Dienstleistungen", 1960-1989. This shows that in services labor is (still) more important than energy. Energy-dependent Cobb-Douglas functions with output elasticities close to the time-averaged Linex elasticities reproduce observed economic growth not too badly, either.

The recalibrated a and c after 1978 are consistent with observed net efficiency improvements and energy demand reductions of the capital stock in response to the first and the second oil price shocks. The associated abrupt time change may be interpreted as the result of a "creativity pulse" between 1977 and 1978. However, such a one-year pulse is a crude approximation of the working of "creativity" at best. Therefore, modeling the technology parameters by functions that change continously in time tries to improve that, providing also the means of calculating creativity's output elasticity δ. This allows to see how much of the Solow residual is removed by energy and its non-cost-share weighting, and how much remains unexplained by the factors capital, labor and energy.

We have modeled $a(t)$ and $c(t)$ by logistic functions of the form (35) and by Taylor series expansions in terms of $t - t_0$. In the case of reunited Germany, Fig. 1, a combination of both with a total of only five free coefficients was used. In all cases the free coefficients $a_1 \dots a_s$ and $c_1 \dots c_r$ are determined by non-linear regression analysis observing the constraints (37). The Levenberg-Marquardt method [46] for the minimization of the sum of squared errors, SSE, (36) and the "Statistical Analysis System" (SAS) are used. The proper starting values for the numerical iterization (with up to 32 000 iteration steps) are crucial for convergence in the true minimum. Two methods are employed for obtaining them: a) The "brute force" method. Here a lattice is projected into the multidimensional space spanned by the free coefficients, and the SSE is computed for each lattice point. The "coordinates" of the lattice point with the smallest SSE are used as starting values. This method is employed for the total economy of reunited Germany, and as a control of the results for the other systems as well. b) A new iteration method developed by Julian Henn and employed first in [7]. It is indicated in Fig. 6 and involves the definitions: $x_i \equiv 2 - (l_i + e_i)/k_i, \quad y_i \equiv l_i/e_i - 1$, so that the Linex function (82) at time t_i can be written as

$$q_{Lt} = q_0 \cdot e_i \cdot \exp[a(t_i)(x_i + c(t_i)y_i)] \equiv q_i. \tag{84}$$

In the iteration scheme of Fig. 6 $c_{min}(t)$ and $a_{max}(t)$ are given by eq. (37). The search for the proper starting values according to this scheme proceeds as follows:
i) Make an educated guess for q_0 and $a(t)$. One option is $q_0 = 1$ and for $a(t)$ the choice of a step function that corresponds to a Linex fit with piecewise constant a and c and recalibration, e.g. in 1978. ii) Compute $c(t_i)$ for each point in time t_i from $c(t_i) = \frac{\ln(q_i/q_0 e_i)}{a(t_i)y_i} - x_i/y_i$; this is eq. (84), resolved with respect to $c(t)$. Insert $a(t_i)$ and q_0 from step i) into it. iii) If many of the $\{c(t_i)\}$ are smaller than the $\{c_{min}(t_i)\}$ from eq. (37), decrease q_0 and repeat again step i). iv) If most of the $\{c(t_i)\}$ are larger than $\{c_{min}(t_i)\}$, estimate the free coefficients $c_1 \dots c_r$ of the logistic function or Taylor expansion for $c(t)$, $f(c_1 \dots c_r)$, so that this $f(c_1 \dots c_r)$ fits the $\{c(t_i)\}$ satisfactorily. v) Compute $a(t_i)$ for

each point in time t_i from $a(t_i) = \frac{\ln(q_i/q_0 e_i)}{x_i + c(t_i) y_i}$; this is eq. (84), resolved with respect to $a(t)$. Insert $c(t) = f(c_1 \ldots c_r)$ and q_0 from step iv) into it. vi) If many of the $\{a(t_i)\}$ are outside the range of a values allowed by the first of eqs. (37), decrease q_0 and repeat again step v). vii) If most of the $\{a(t_i)\}$ are within the allowed range, estimate the free coefficients $a_1 \ldots a_s$ of the logistic function or Taylor expansion for $a(t)$, $g(a_1 \ldots a_s)$, so that this $g(a_1 \ldots a_s)$ fits the $\{a(t_i)\}$ satisfactorily. viii) Repeat steps ii) to vii) until the $q_0, a(t)$, and $c(t)$ don't change any more. The corresponding $a_1 \ldots a_s$, $c_1 \ldots c_r$, and q_0 are the starting values for SSE minimization.

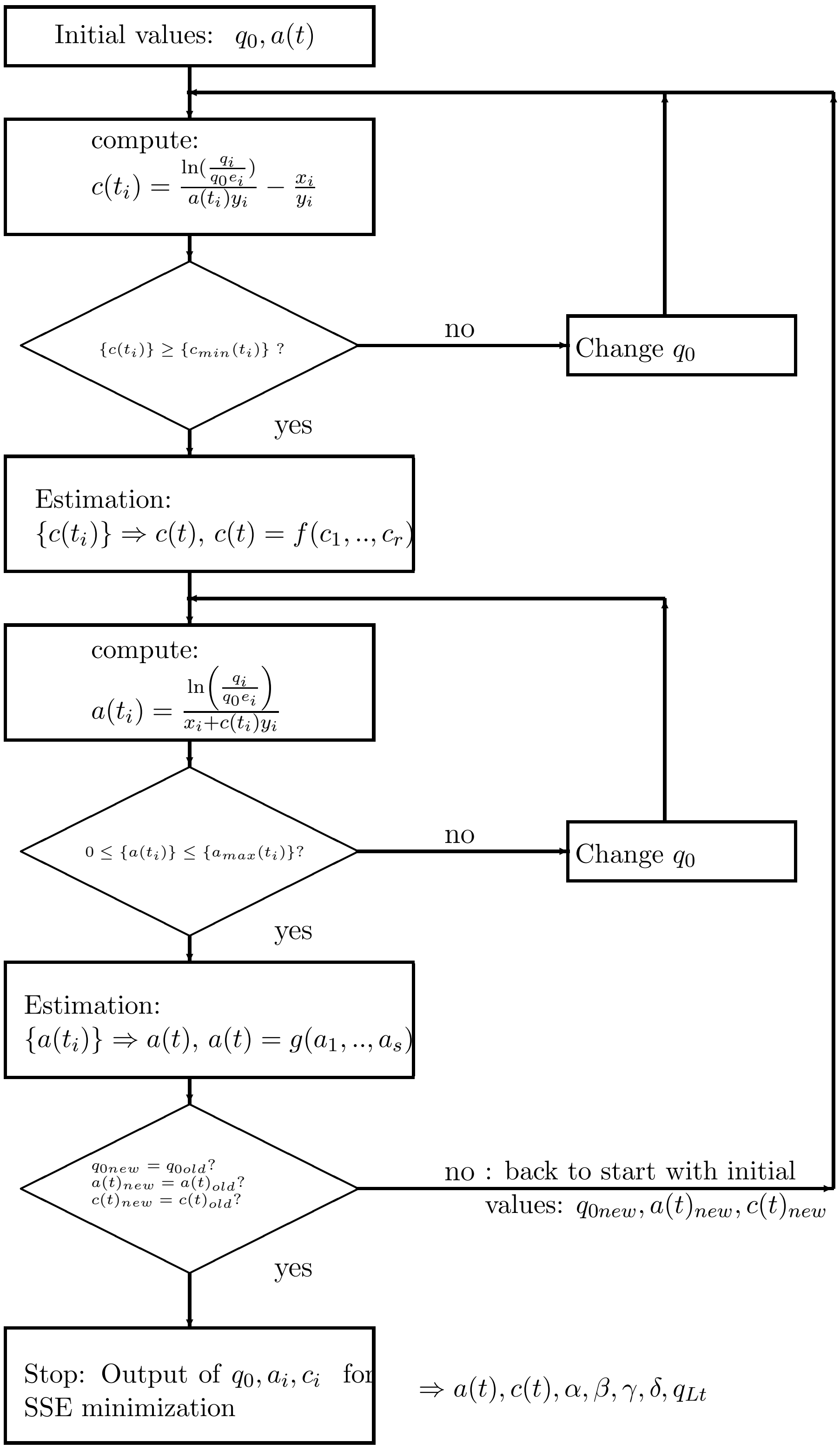

Figure 6: Iteration scheme for the determination of the starting values of the free coefficients in $a(t)$ and $c(t)$ that are to be used for SSE minimization.

References

[1] R. M. Solow, The Economics of Resources and the Resources of Economics, The American Economic Review 64 (1974) 1-14.

[2] D. Stern, Stern Review Report on the Economics of Climate Change, Cambridge University Press, Cambridge, 2007; (http://www.cambridge.org/9780521700801)

[3] D. Strahan, The Last Oil Shock, John Murray, London, 2007.

[4] G. Tintner, E. Deutsch, R. Rieder, A Production Function for Austria Emphasizing Energy, in: F.L. Altman, O. Kýn, H.-J. Wagner (Eds.), On the Measurement of Factor Productivities, Vandenhoek & Ruprecht, Göttingen, 1974, pp 151-164.

[5] R. Kümmel, W. Strassl, A. Gossner, W. Eichhorn, Technical progress and energy-dependent production functions, Z. Nationalökonomie – Journal of Economics 45 (1985) 285-311 .

[6] R. Kümmel, D. Lindenberger, W. Eichhorn, The Productive Power of Energy and Economic Evolution, Indian Journal of Applied Economics 8 (2000) 231-262 . Special Issue Essays in Honor of Professor Paul A. Samuelson.

[7] R. Kümmel, J. Henn, and D. Lindenberger, Capital, labor, energy, and creativity: modeling innovation diffusion, Structural Change and Economic Dynamics 13 (2002) 415-433, and Internet Supplement.

[8] B.C. Beaudreau, Energy and Organization: Growth and Distribution Reexamined, Greenwood Press, Westport CT, 1998

[9] D. Lindenberger, Wachstumsdynamik industrieller Volkswirtschaften: Energieabhängige Produktionsfunktionen und ein faktorpreis-gesteuertes Optimierungsmodell, Metropolis, Marburg, 2000.

[10] D. Lindenberger, R. Kümmel, Energy-Dependent Production Functions and the Optimization Model "PRISE" of Price-Induced Sectoral Evolution, International Journal of Applied Thermodynamics 5 (2002) 101-107.

[11] R. U. Ayres and B. Warr, Accounting for growth: the role of physical work, Structural Change and Economic Dynamics 16 (2005) 181-209.

[12] F.P. Ramsey, A Mathematical Theory of Saving, Economic Journal XXXVIII (1928) 543-559.

[13] P. A. Samuelson, R. M. Solow, A complete capital model involving heterogeneous capital goods, The Quarterly Journal of Economics, 70 (1956) 537-562 .

[14] E.H. Hudson, D.W. Jorgenson, U.S. Energy Policy and Economic Growth, 1975-2000, The Bell Journal of Economics and Management Science 5 (1974) pp. 461-514.

[15] E.R. Berndt, D.W. Jorgenson, How Energy and its Cost enter the Productivity Equation, IEEE Spectrum 15 (1978) 50-52.

[16] E.R. Berndt, D.O. Wood, Engineering and Econometric Interpretations of Energy-Capital Complementarity, The American Economic Review 69 (1979) 342-354.

[17] D. W. Jorgenson, The Role of Energy in the U.S. Economy, National Tax Journal 31 (1978) 209-220.

[18] D.W. Jorgenson, The Role of Energy in Productivity Growth, The American Economic Review 74, No. 2 (1984) 26-30.

[19] D.W. Jorgenson, Productivity and Economic Growth in Japan and the United States, The American Economic Review 78, No. 2 (1988) 217-222.

[20] E.F. Renshaw, Energy Efficiency and the Slump in Labor Productivity in the U.S.A., Energy Economics 3 (1981) 36-42.

[21] E. F. Denison, Explanation of Declining Productivity Growth, Survey of Current Business 59, No. 8, Part II (1979) 1-24 .

[22] E. F. Denison, Accounting for Slower Economic Growth: An Update, in: International Comparisons of Productivity and Causes of Slowdown, J.W. Kendrick (Ed.) Ballinger, Cambridge 1984, pp. 1-45.

[23] R.M. Solow, Perspectives on Growth Theory, Journal of Economic Perspectives 8 (1994) 45-54 .

[24] R. J. Barro, Notes on Growth Accounting, Journal of Economic Growth 4 (1999) 119-137.

[25] E.B. Barbier, Endogeneous Growth and Natural Resource Scarcity, Environmental and Resource Economics 14 (1999) 51-74.

[26] J.E. Stiglitz, Growth with Exhaustible Natural Resources: Efficient and Optimal Growth Paths, Review of Economic Studies; Symposium on the Economics of Exhaustible Resources (1974) 123-138.

[27] P. Romer, Endogeneous Technological Change, Journal of Political Economy 95 (1990) 71-102.

[28] H. Welsch, K. Eisenack, Energy Costs, Endogeneous Innovation, and Long-run Growth, Jahrbücher für Nationalökonomie u. Statistik/Journal of Economics and Statistics 222/4 (2002) 490-499.

[29] G. Bertola, Factor Shares and Savings in Endogeneous Growth, The American Economic Review 83, No. 5 (1993) 1184-1198.

[30] P. Howitt, Endogeneous Growth and Cross-Country Income Differences, The American Economic Review 90, No. 4 (2000) 829-846.

[31] Z. Griliches, Issues in Assessing the Contribution of Research and Development to Productivity Growth, Bell Journal of Economics 10 (1979) 92-116.

[32] P. Romer, Increasing Returns and Long-Run Growth, Journal of Political Economy 94 (1986) 1002-1037.

[33] R. E. Lucas, On the Mechanics of Economic Development, Journal of Monetary Economics 22 (1988) 3-42.

[34] S. Kim, H. Mohtadi, Labor Specialization and Endogeneous Growth, The American Economic Review 82, No. 2 (1992) 404-408.

[35] P.S. Segerstrom, Endogeneous Growth without Scale Effects, The American Economic Review 88, No. 2 (1998) 1290-1310.

[36] S.S. Kortum, Research, Patenting, and Technological Change, Econometrica 65 (1997) 1389-1419

[37] P. Moser, T. Nicholas, Was Electricity a General Purpose Technology? Evidence from Historical Patent Citations, The American Economic Review 94, No. 2 (2004) 388-394.

[38] J. Felipe, F.M. Fisher, Aggregation in production functions: what applied economists should know, Metroeconomica 54 (2003) 208-262.

[39] D. Lindenberger, R.U. Ayres, R. Kümmel, Macro- and Microeconomic Production Functions, Discussion Paper (2008).

[40] R. Kümmel, Growth Dynamics of the Energy-Dependent Economy, Oelgeschlager, Gunn & Hain, Cambridge, Mass., 1980.

[41] R. Kümmel, W. Strassl, Changing energy prices, information technology, and industrial growth, in: W. van Gool and J.J. Bruggink (Eds.), Energy and Time in Economic and Physical Sciences, North Holland, Amsterdam, 1985, pp. 175-194.

[42] R. Kümmel, The impact of energy on industrial growth, Energy –The International Journal 7 (1982) 189-203.

[43] P. A. Samuelson, Economics, McGraw-Hill, New York, 1973.

[44] D. Lindenberger, Service Production Functions, Journal of Economics – Z. Nationalökonomie 80 (2003) 127-142.

[45] Ch. Hall, D. Lindenberger, R. Kümmel, T. Kroeger, W. Eichhorn, The need to reintegrate the natural sciences with economics, Bioscience 51 (2001) 663-673.

[46] W. H. Press, S. A. Teukolsky, W. T. Vetterlin, B. P. Flannery, Numerical Recipes in C, Cambridge University Press, Cambridge, 1992

[47] R. U. Ayres, L. W. Ayres, B. Warr, Exergy, power and work in the US economy, 1900- 1998, Energy 28 (2003) 219-273.

[48] R. U. Ayres, B. Warr, On the importance of Exergy and Useful Work in the Economy: US and Japan 1900-2004, Discussion paper (2007).

[49] R. Stresing, D. Lindenberger, R. Kümmel, Cointegration of Output, Capital, Labor, and Energy, Discussion paper (2008).