Characterizing the development of sectoral gross domestic product composition

Raphael Lutz, Michael Spies, Dominik E. Reusser, Jürgen P. Kropp,* and Diego Rybski†
Potsdam Institute for Climate Impact Research (PIK), P. O. Box 60 12 03, 14412 Potsdam, Germany
(Received 31 January 2013; published 8 July 2013)

We consider the sectoral composition of a country's gross domestic product (GDP), i.e., the partitioning into agrarian, industrial, and service sectors. Exploring a simple system of differential equations, we characterize the transfer of GDP shares between the sectors in the course of economic development. The model fits for the majority of countries providing four country-specific parameters. Relating the agrarian with the industrial sector, a data collapse over all countries and all years supports the applicability of our approach. Depending on the parameter ranges, country development exhibits different transfer properties. Most countries follow three of eight characteristic paths. The types are not random but show distinct geographic and development patterns.

DOI: 10.1103/PhysRevE.88.012804 PACS number(s): 89.65.Gh, 45.30.+s

I. INTRODUCTION

During its development, humankind has passed through various stages of fundamentally different ecological and technological characteristics. In line with dramatic population growth, an increasing interaction with the biosphere and a domination of ecosystems took place. During the neolithic revolution, around 10 000 BCE (before common era), hunter-gatherer societies were progressively replaced by agrarian ones, with far-reaching consequences such as the formation of settlements. The industrial revolution is considered as the most significant development affecting all areas of human life, along with the systematic exploration of fossil energy sources.

From an economic point of view, the increasing significance of services is understood as an additional level of development. In fact, agrarian, industrial, and service sectors are commonly denoted primary, secondary, and tertiary, respectively. However, the production forms do not completely replace each other but are complements, and economies have more or less contributions from each sector.

Current theories and models on sectoral development are largely influenced by the work of Clark, Fisher, and Fourastie who developed the "three-sector hypothesis" in the first half of the 20th century, describing development as a process of shifting economic activities from the primary via the secondary to the tertiary sector [1–4]. Their research was mainly based on observed historical shifts of workforce between sectors in today's more developed countries. More recently, approaches have concentrated on describing the relationship between shifts in sectoral labor allocation or gross domestic product (GDP) shares and economic development, often focusing on specific countries or regions [5–7]. Yet, the universality of the three-sector hypothesis has been challenged, since it does not well represent labor allocation in today's developing countries [8,9]. Different from the historical pathways of industrialized countries, shifts of labor force from the primary to the secondary sector have been relatively low. Instead, advancement of the tertiary sector appears disproportionate early, which has been related to excessive urbanization and different structural conditions [8,9].

While existing work has mainly focused on modeling patterns observed in the United States or in Western Europe, applying a similar analysis in a universal model does not exist to the best of our knowledge. Furthermore, attention has mainly been given to sectoral resource allocation, e.g., labor input, rather than to economic output, e.g., the fractions of GDP. Thus, the objective is to develop a parsimonious description of a country's sectoral composition of GDP, which is also able to capture the early advancement of the tertiary sector observed in developing countries.

II. MODEL

We consider a country c and its sectoral GDP composition, where the fractions a, i, s, correspond to the agricultural, industrial, and service sector contributions, respectively. The fractions of the three sectors add up to unity, $a + i + s = 1$.

With economic development, i.e., increasing GDP per capita, the shares of the GDP shift between the sectors. We assume the transfer occurs according to a system of ordinary differential equations,

$$\frac{da}{dg} = -k_1 a, \tag{1}$$

$$\frac{di}{dg} = \alpha k_1 a - k_2 i, \tag{2}$$

$$\frac{ds}{dg} = (1-\alpha) k_1 a + k_2 i, \tag{3}$$

where g is the logarithm of GDP per capita (the natural logarithm is used in order to compensate for the broad distribution of GDP per capita values) and k_1, k_2, α are country-specific parameters.[1]

An additional parameter, g_0, emerges from the boundary conditions, i.e., from the value of g where $a = 1$, and has the character of a shift along the g axis. Figure 1 illustrates the model and shows schematic trajectories. Parameter k_1

*University of Potsdam, Dept. of Geo- & Environmental Sciences, 14476 Potsdam, Germany.
†ca-dr@rybski.de

[1]Equations (1)–(3) can also be expressed with a different set of parameters: $\frac{da}{dg} = -k_{ai} a - k_{as} a$, $\frac{di}{dg} = k_{ai} a - k_{is} i$, $\frac{ds}{dg} = k_{as} a + k_{is} i$, with $k_{ai} = \alpha k_1$, $k_{is} = k_2$, and $k_{as} = (1-\alpha) k_1$.

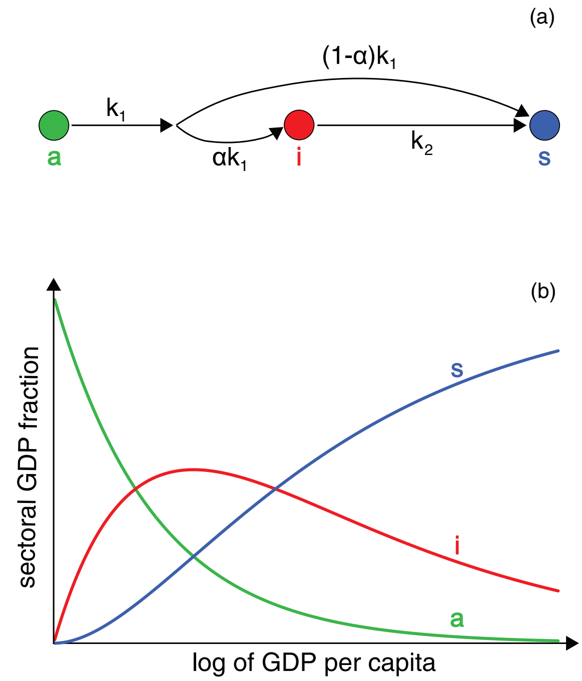

FIG. 1. (Color online) Illustration of the transfer model. (a) The three sectors, agrarian a, industrial i, and service s, are represented by colored circles. The arrows indicate possible transfer paths and their parameters as defined in Eqs. (1)–(3). (b) Typical evolution of sectoral composition as a function of the logarithm of per capita GDP (here $k_1 = 1$, $k_2 = 0.5$, $\alpha = 1$).

determines the transfer from the agrarian sector, which, depending on α, is split into contributions to the industrial or service sectors. Moreover, k_2 determines the transfer from industrial to service. For example, for $k_1 > 0$, $k_2 > 0$, and $\alpha = 1$ transfer takes place from a to i and continuously from i to s, leading to monotonically decreasing agrarian and increasing service, whereas the industrial sector exhibits a maximum [Fig. 1(b)]. Except for the trivial case, the model does not have any steady state.

III. RESULTS

We fit the model Eqs. (1)–(3) with a two step procedure, using global country-level data.[2] In the first step, the logarithmic form of Eq. (4), as introduced later, was used to identify k_1 and an initial value of g_0 by using a linear regression between $\ln(a)$ and g. In the second step, k_2, α, and g_0 were estimated by using the R implementation of the shuffled complex evolution algorithm [11] to minimize the sum of the mean squared errors between a, i, s from the model and the corresponding observed values. To obtain more reasonable fits, we restricted the parameter ranges as follows: $-5 < k_2 < 5, 0 < \alpha < 5$, and $1 < g_0 < 15$.

For 176 out of 246 countries the available data was sufficient, i.e., data on GDP per capita were available for at least four years, and for 137 countries the fitting worked reasonably well (we choose 0.1 as the threshold for the sum of the mean squared error between the data and the fits of all sectors). Due to an anomalous decline of g after dissolution of the Soviet Union, disbandment of the Warsaw Pact, and the breakup of Yugoslavia, respectively, the data before 1995 have been disregarded, in the case of the corresponding countries. For the same reason, data from Liberia and Mongolia prior to 1995 were omitted.

Typical examples for which the model results were accepted are depicted in Fig. 2 together with the obtained fits. In all three examples the model agrees reasonably with the data. In the case of Pakistan, the fraction of industry overtakes the fraction of agriculture at $g_{a\times i} \approx 7.5$ ($1800 per capita). The industrial fraction of Finland reached its maximum at $g_{\max i} \approx 9.7$ ($16 300 per capita). The service sector is the largest and is still increasing for these examples. In the case of the USA, the agrarian sector has a very low contribution.

Different parameter ranges imply different behavior, e.g., $k_1 < 0$ means that the country transfers economic activity to agriculture with increasing GDP per capita. In total there are two different cases for each parameter leading to eight combinations. Table I gives an overview of the corresponding types together with the transfer behavior, i.e., economic transfer from which sector to which, and the frequency of each type. Almost half of the considered countries belong to type 1, the traditional path from the agrarian, via industrial, to the service sector. The second most frequent is type 3, which includes a transfer from the service to industrial sector. Another big group consists of type 2 countries, i.e., with transfer from agrarian to industry and flows between industry and services depending on the development. All other types are less populated; type 4 does not occur at all. The occurrence of types 5–8 might be due to noise in the data. Only types 1 and 2 and types 7 and 8 exhibit a maximum of the industrial sector at $g_{\max i} = \frac{\ln(k_1/k_2)}{k_1 - k_2} + g_0$. The examples from Fig. 2 are of type 3 (Pakistan), type 1 (Finland), and type 2 (USA).

On the world map, Fig. 3, one can see which country belongs to which type. Type 1 consists of big parts of Asia and Eastern Europe, some countries in Africa, Canada, and Mexico. The USA, Brazil, other South American countries, Western European countries, Japan, and Australia belong to type 2. Type 3 is mainly found in Africa, Middle East, Central Asia, Southeast Asia, and a few times in South America. A strong regionality can be observed and neighboring countries tend to belong to the same types. It is apparent that most developed countries belong to either type 1 or type 2. At this stage it is not clear what the decisive factor is, and further analysis including other economic data could help to pinpoint the most relevant influences of the countries economic paths. Methods from network theory have been applied to analyze the economic productions of countries, indicating that neighboring countries instead of diversifying tend to compete over the same markets [12].

[2]Country-level data for fitting the model were obtained from the World dataBank provided by the World Bank [10]. g: GDP per capita based on purchasing power parity (PPP) in constant 2005 international dollars (World dataBank Series Code NY.GDP.PCAP.PP.KD). a: Net output of the agricultural sector as percentage of GDP (World dataBank Series Code NV.AGR.TOTL.ZS). i: Net output of the industrial sector as percentage of GDP (World dataBank Series Code NV.IND.TOTL.ZS). s: Net output of the services sector as percentage of GDP (World dataBank Series Code NV.SRV.TETC.ZS). The data cover the period 1980–2005 in annual resolution.

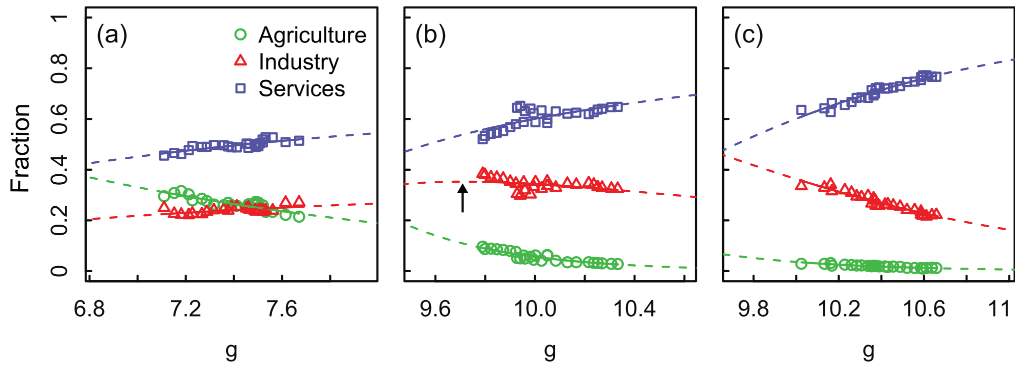

FIG. 2. (Color online) Examples of sectoral GDP fractions versus the logarithm of GDP per capita. (a) Pakistan, (b) Finland, and (c) the United States. Open symbols represent the observed values, solid lines the fitted functions according to Eqs. (1)–(3), and dashed lines extrapolations for illustration. The fitted parameters are (a) Pakistan ($k_1 = 0.56$, $k_2 = -0.01$, $\alpha = 0.32$, $g_0 = 8.12$), (b) Finland ($k_1 = 2.29$, $k_2 = 0.35$, $\alpha = 0.50$, $g_0 = 8.74$), and (c) the United States ($k_1 = 1.76$, $k_2 = 0.94$, $\alpha = 1.27$, $g_0 = 5.02$). For Finland, the maximum of the industrial sector, $g_{\max i} \approx 9.7$, is indicated by a black arrow in (b).

The inset of Fig. 3 shows the histograms of g (year 2005) for types 1–3. Type 1 and 3 countries are spread over a wide range of GDP per capita, whereas there is a tendency of type 1 countries toward higher GDP per capita ($g > 9$) compared to the type 3 case. Type 2 countries generally tend to larger GDP per capita. Surprisingly, many high GDP per capita countries belong to type 2 and not to type 1. Accordingly, their economic growth follows the traditional path $a \to i \to s$, but depending on the state of a and i, the flow between $i \to s$ may be increased (high i), decreased, or even reversed (high a) (see Table I). Type 3, which is the second most frequent one, also follows the traditional path from the agrarian to the industry and service sectors, but comes along with a transfer from the service sector to the industry sector, $a \to i \leftarrow s$, $a \to s$. This version seems to be characteristic for many developing economies, although not exclusively.

TABLE I. Types of sectoral GDP transfer as obtained from the fitting parameters of the model Eqs. (1)–(3). The types are defined according to the sign of the parameters k_1 and k_2 as well as if $0 < \alpha < 1$ or $\alpha > 1$. The parameters also specify between which sectors there is a transfer and in which direction. For $\alpha > 1$, instead of a transfer from the agriculture to the services sector, a transfer occurs between industry and services, depending on the values of k_1 and k_2 (indicated by a "–"). If k_1 and k_2 have opposite signs, the flow is in one direction for all possible values of a, i, and s. If k_1 and k_2 are positive, flow may occur from s to i for large a, while it is reversed for large i (indicated by $\leftrightarrows$). Vice versa for k_1 and k_2 negative. Only types 1 and 2 have a convergent asymptotic behavior, namely $a \to 0$, $i \to 0$, $s \to 1$ for $g \to \infty$. Most countries belong to the types 1–3. The remaining types are type 5 (Guinea-Bissau, Madagascar, Vanuatu), type 6 (Ivory Coast), type 7 (Cameroon), and type 8 (Burkina Faso, Morocco, Sudan, Venezuela, South Africa). The special cases $k_1 = 0$, $k_2 = 0$, $\alpha = 0$ or $\alpha = 1$ did not occur.

Type	k_1	k_2	α	Transfer behavior	Number of countries
1	>0	>0	<1	$a \to i, i \to s, a \to s$	59
2	>0	>0	>1	$a \to i, i \leftrightarrows s, a - s$	25
3	>0	<0	<1	$a \to i, i \leftarrow s, a \to s$	43
4	>0	<0	>1	$a \to i, i \leftarrow s, a - s$	0
5	<0	>0	<1	$a \leftarrow i, i \to s, a \leftarrow s$	3
6	<0	>0	>1	$a \leftarrow i, i \to s, a - s$	1
7	<0	<0	<1	$a \leftarrow i, i \leftarrow s, a \leftarrow s$	1
8	<0	<0	>1	$a \leftarrow i, i \rightleftarrows s, a - s$	5

IV. DATA COLLAPSE

In order to test the universality and applicability of the model, we derive a collapsed representation of all data. We start from solutions of Eqs. (1)–(3),

$$a(g) = e^{-k_1(g-g_0)}, \tag{4}$$

$$i(g) = \frac{\alpha k_1}{k_2 - k_1}\left(e^{-k_1(g-g_0)} - e^{-k_2(g-g_0)}\right), \tag{5}$$

$$s(g) = 1 - a(g) - i(g). \tag{6}$$

Eliminating $(g - g_0)$ in Eqs. (4) and (5), one obtains a relation between $a(g)$ and $i(g)$,

$$\frac{k_2 - k_1}{\alpha k_1} i(g) = \left(a(g) - a(g)^{k_2/k_1}\right), \tag{7}$$

allowing one to collapse the data of all countries and all years, i.e., independent of g. In Figs. 4(a)–4(c) we plot the transformed data and separate into panels the three most frequent types for better visibility. The data generally collapse onto the unity diagonal, despite a few countries where deviations are partly due to the fact that values before 1995 have been excluded from the fitting (see above) but the values are still displayed for completeness. The collapse suggests universality [13,14] and supports the applicability of the proposed model.

For the set of countries with reasonable fitting, some of the parameters are correlated nonlinearly, i.e., g_0 and k_1 as well as k_2 and α. Thus, by introducing global parameters, the number of country-specific ones could be reduced. Moreover, k_1 is weakly correlated with g, suggesting that—from the ensemble point of view—Eq. (4) has rather a log-normal shape, which indicates that the system is not ergodic.

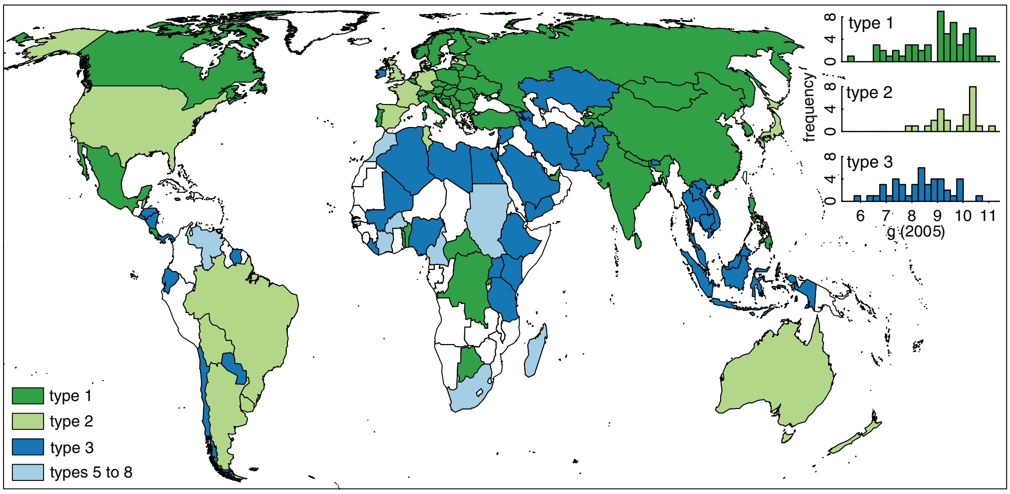

FIG. 3. (Color online) Types of sectoral GDP transfer, geographic location, and frequency. The map resolves the three most common types, 1–3 (Table I), and the remaining types 5–8. Countries with insufficient data or bad fitting are not colored. Countries of the same type tend to be neighbors. The inset in the upper right corner shows the frequency of countries according to their g (logarithm of GDP per capita) in 2005 for the types 1–3.

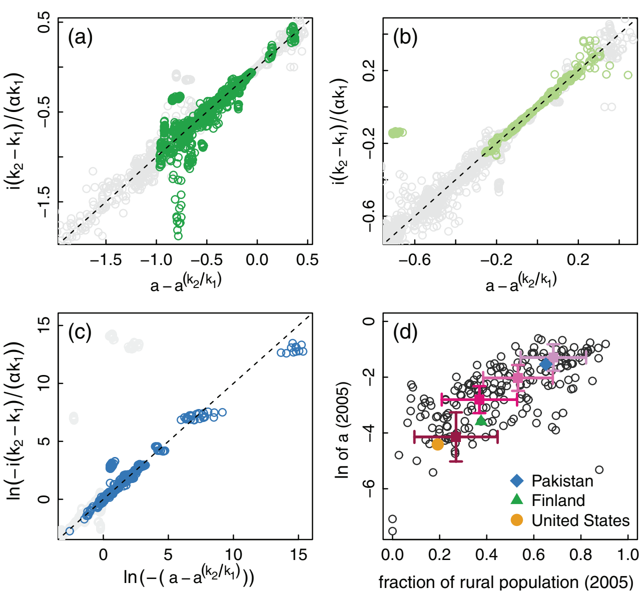

FIG. 4. (Color online) Data collapse according to Eq. (7) and correlations between agrarian sector and rural population. Panels (a)–(c) display the transformed values of all countries and all years. (a) type 1, (b) type 2, and (c) type 3; see Table I. For clarity, in (c) the signs have been changed and the natural logarithm taken. The underlying grey data points show the values of all other types. The dashed lines have slope 1 and intercept 0. For some countries the fitting does not perform well, in particular (a) Singapore (at approx. −0.8,−0.35) and Bulgaria (approx. between −0.8,−1.2 and −0.8,−1.8), (b) Denmark (at approx. −0.7,−0.14), and (c) Saudi Arabia (at approx. 0.6,2.9) and the Maldives (at approx. 14.6,12.9). Panel (d) shows for the year 2005 $\ln a$ versus the fraction of rural population, based on data from the World dataBank. The open circles represent the values of all countries with available data and exhibit a correlation coefficient of 0.69. The examples from Fig. 2 are highlighted with filled symbols, i.e., Pakistan (diamond), Finland (triangle), and USA (circle). The filled squares with error bars represent averages and standard deviations of GDP per capita quartiles, i.e. lowest (light, top right) to highest (dark, lower left).

We would like to note that since g is the logarithm of the GDP per capita, a decreases in a power-law manner with the GDP per capita. Studying the asymptotic behavior of the types as defined in Table I, it turns out that types 1 and 2 converge to $s = 1$ for $g \to \infty$. The model is confined to specific ranges of g, e.g., $g \leqslant g_0$ for negative parameter k_1. Thus, it is important to keep in mind that fitting the model only characterizes the transfer as it is included in the data. This means that the obtained parameters only capture the behavior of the past.

V. DISCUSSION

Finally, $\ln a$ is plotted versus the fraction of urban population for the year 2005 in Fig. 4(d). The two quantities are correlated with a correlation coefficient of 0.69. Despite not being completely linear, the correlations are considerable, implying that low agrarian contribution to the economy's GDP comes along with less rural population. In order to visualize the relation to overall economic output, Fig. 4(d) also includes averages and standard deviations of those countries belonging to GDP per capita quartiles, i.e. the quarter of all countries with highest GDP per capita, the second quarter of countries, etc. As one can see, with increasing GDP per capita, rurality and agrarian GDP share decrease. In other words, a high degree of urbanization comes along with economic development or vice versa. This can be related to the finding that per capita socioeconomic quantities such as wages, GDP, number of patents applied, and number of educational and research institutions increase by an approximate factor of 1.15 with increasing city size [15].

However, as Timberlake [8] has pointed out, in the case of developing countries an "overurbanization" with fast growing urban populations and excessive employment in the service sector can also hinder economic growth. Not without reason, most developing countries in our model belong to type 3, where economic growth is associated with a sectoral transfer from service to industry.

In summary, we propose a system of ordinary differential equations to characterize the development of the sectoral GDP composition. Despite being very simple and involving only four country-specific parameters (g_0 has only the character of shift along g), the model fits for the majority of the countries in the world. Relating agrarian and industrial fractions, we collapse the data of all countries and all years onto a straight line. This could be used as an alternative approach to fit the parameters by means of nonlinear techniques.

We find that, according to the parameter ranges, the countries belong to eight different types. Most countries are found in three of them; the members are distinct in geography and state of economic development. This suggests that countries with low current GDP per capita follow a different path from early developed countries. Our results could indicate a relation between transfer patterns and economic development. Further analysis of additional socioeconomic data could shed light on reasons of economic failure or success.

As with any model, our approach is a strong simplification of reality. Also, we assume that parameters are fixed over time and countries follow a given development pathway. This may be justified by structural, cultural, and bioclimatic conditions, which have been consistent over the period of observation. On the other hand, a transition between characteristic pathways is possible. For workforce distribution of 22 countries from the former Soviet Union and Central as well as Eastern Europe, such a transition analysis has been performed using another simple model based on the three-sector hypothesis [5]. A similar transition analysis could be an extension of the work presented.

Since it has been found that countries tend to develop goods which are similar to those they currently produce [16] and that economically successful countries are extremely diversified [17], it could be also of interest to extend the analysis to the level of products, in order to enable a more detailed analysis. Furthermore, the inclusion of a "quaternary" sector [18,19] in our model might provide additional insights, but sufficient data is not (yet) available. In this context we would also like to note that many developing countries exhibit an informal service sector which is not included in the official figures. Similarly, in developed countries the products can be very complex, so that the separation between industrial and service sectors might be fuzzy. Accordingly, already the data analyzed in this study are likely to be affected by inaccuracies.

ACKNOWLEDGMENTS

We thank Torsten Wolpert, Flavio Pinto Siabatto, Xavier Gabaix, Boris Prahl, and Lynn Kaack for useful discussions and comments. The authors acknowledge the financial support from the Federal Ministry for the Environment, Nature Conservation and Nuclear Safety of Germany who support this work within the International Climate Protection Initiative, and the Federal Ministry for Education and Research of Germany who provided support under the framework of the PROGRESS Initiative (Grant No. 03IS2191B).

[1] C. Clark, *The Conditions of Economic Progress* (Macmillan, London, 1940).
[2] A. G. B. Fisher, Econ. Rec. **15**, 24 (1939).
[3] J. Fourastie, *Le Grand Espoir du XXe Siècle* (Gallimard, Paris, 1949).
[4] J. J. Krüger, J. Econ. Surv. **22**, 330 (2008).
[5] M. Raiser, M. Schaffer, and J. Schuchhardt, Struct. Change Econ. Dynam. **15**, 47 (2004).
[6] C. Echevarria, Int. Econ. Rev. **38**, 431 (1997).
[7] P. Kongsamut, S. Rebelo, and D. Xie, Rev. Econ. Stud. **68**, 869 (2001).
[8] M. Timberlake and J. Kentor, Sociol. Quart. **4**, 489 (1983).
[9] K. Pandit and E. Casetti, Ann. Ass. Am. Geogr. **79**, 329 (1989).
[10] The World Bank, data retrieved December 19th, 2011, from World dataBank database, http://data.worldbank.org (2011).
[11] Q. Y. Duan, V. K. Gupta, and S. Sorooshian, J. Optimiz. Theory Appl. **76**, 501 (1993).
[12] G. Caldarelli, M. Cristelli, A. Gabrielli, L. Pietronero, A. Scala, and A. Tacchella, PLoS ONE **7**, e47278 (2012).
[13] H. E. Stanley, Rev. Mod. Phys. **71**, S358 (1999).
[14] R. D. Malmgren, D. B. Stouffer, A. S. L. O. Campanharo, and L. A. N. Amaral, Science **325**, 1696 (2009).
[15] L. Bettencourt and G. West, Nature (London) **467**, 912 (2010).
[16] C. A. Hidalgo, B. Klinger, A.-L. Barabási, and R. Hausmann, Science **317**, 482 (2007).
[17] A. Tacchella, M. Cristelli, G. Caldarelli, A. Gabrielli, and L. Pietronero, Sci. Rep. **2**, 723 (2012).
[18] Z. Kenessey, Rev. Income Wealth **3**, 359 (1987).
[19] T. Selstad, Nor. Geogr. Tidsskr. **44**, 21 (1990).