

Ullrich Heilemann, Harald Lehmann, Joachim Ragnitz

## Länder-Rankings – Komplexitätsreduktion oder Zahlenalchemie?

Die Welt scheint vielen komplex und verflochten – Vereinfachung ist gefragt. Rankings und Ratings aller Art erfreuen sich großer, kaufkräftiger Nachfrage, und mit Laptop und Internet lässt sich alles in die richtige Ordnung bringen – Immobilien, Aktienfonds, Professoren, Hochschulen, Städte, Volkswirtschaften usw. Die festen Kosten der Prozeduren mögen hoch sein, die variablen sind es dank Standardisierung der Beurteilungsmuster und Automatisierung nicht. Dreimal „Hoch“ – eine neue Industrie ist geboren. Haftungs- oder Zertifizierungsfragen stellen sich in diesen Gewässern der Wissensgesellschaft bislang nicht – „anything goes“. Wer sich beschwert, ist meist schlecht weggekommen.

Vor diesem Hintergrund hatte das Bundesministerium der Finanzen nach einer Ausschreibung im Juni 2005 das Institut für Wirtschaftsforschung Halle (IWH) in Kooperation mit Ullrich Heilemann, Universität Leipzig, damit beauftragt, den Informationsgehalt des Rankings von Volkswirtschaften im Rahmen eines Gutachtens zu untersuchen. Dabei sollte vor allem der wirtschaftspolitisch relevante Aussagegehalt internationaler Ländervergleiche identifiziert und bewertet werden. Insbesondere sollte geprüft werden, ob das im Vergleich zu anderen führenden Wirtschaftsnationen relativ schlechte Abschneiden Deutschlands begründet ist, oder ob und in welchem Maße es lediglich Folge inhaltlicher Beschränkungen der einbezogenen Kriterien, der verwendeten Daten, spezifischer Methoden oder Gewichtungen des Rankings ist.<sup>1</sup> Das Gutachten wurde im April 2006 abgeschlossen und im November 2006 veröffentlicht, das Bundesministerium hatte eine Zusammenfassung im September 2006 publiziert.<sup>2</sup>

Sebastian Gundel und Ullrich van Suntum<sup>3</sup>, letzterer zusammen mit Klaus Hafemann Verfasser des

„Internationalen Standort Rankings“ der Bertelsmann-Stiftung, kritisieren unsere Vorgehensweise und namentlich das „Abschneiden“ des Bertelsmann-Standort-Rankings. Ehe wir darauf eingehen, ist es zweckmäßig, die Rankings selbst sowie Ansatzpunkte und Ergebnisse unserer Untersuchung vorzustellen.

### Drei Rankings

Zu untersuchen waren der „Global Competitiveness Report“ des Weltwirtschaftsforums (WEF), Genf, das „World Competitiveness Yearbook“ des International Institute for Management Development (IMD), Lausanne, und das „Internationale Standort-Ranking“ der Bertelsmann Stiftung, Gütersloh.<sup>4</sup> Mit Blick auf ihre Zielsetzungen sind sie als sehr anspruchsvoll anzusehen und erfreuen sich hoher öffentlicher Aufmerksamkeit. Ihre Methoden und ihre empirische Basis sind hinreichend dokumentiert, was eine externe Analyse und Bewertung – da sie bis in die 1990er Jahre zurückreichen – auch ihres prognostischen Gehaltes erlaubt. Die Ansätze unterscheiden sich zum Teil beträchtlich, so dass ihre Zusammenschau ein breit abgestütztes Urteil erlaubt.

Explizites oder implizites Ziel der Rankings ist die Bewertung der „Internationalen Wettbewerbsfähigkeit“ von Volkswirtschaften. Verständnis, Definition und Operationalisierung dieses Begriffes haben eine lange, und wenn man Krugmans und anderer generelle Kritik einbezieht, wechselvolle Geschichte.<sup>5</sup> Gegenwärtig wird „Internationale Wettbewerbsfähigkeit“ wahlweise als Absatzfähigkeit, Standortstärke, Anpassungsfä-

<sup>1</sup> Vgl. U. Heilemann, H. Lehmann, J. Ragnitz: Länder-Rankings und internationale Wettbewerbsfähigkeit. Eine kritische Analyse, Schriften des Instituts für Wirtschaftsforschung Halle, Bd. 24, Baden-Baden 2006.

<sup>2</sup> Vgl. Bundesministerium der Finanzen: Internationale Vergleiche von Volkswirtschaften im Rahmen von Ranking- und Benchmarking-Verfahren, in: Monatsbericht des BMF, September 2006, S. 41 ff.

<sup>3</sup> Vgl. S. Gundel, U. van Suntum: Ist die Kritik an internationalen Standort-Rankings berechtigt? in: WIRTSCHAFTSDIENST, 87. Jg. (2007), H. 7, S. 473-479.

<sup>4</sup> Vgl. Bertelsmann Stiftung (Hrsg.): Internationales Standort-Ranking 2004. Wachstum und Beschäftigung, Gütersloh 2004; Institute for Management Development (Hrsg.): The World Competitiveness Yearbook 2005, Lausanne 2005; und World Economic Forum (Hrsg.): The Global Competitiveness Report 2005-2006, Basingstoke u. a. 2005.

<sup>5</sup> Vgl. U. Heilemann, H. Lehmann, J. Ragnitz, a.a.O., S. 12 ff.

*Prof. Dr. Ullrich Heilemann, 62, ist Direktor des Instituts für Empirische Wirtschaftsforschung (IEW) der Universität Leipzig; Dr. Harald Lehmann, 36, ist wissenschaftlicher Mitarbeiter in der Abteilung Strukturökonomik des Instituts für Wirtschaftsforschung Halle (IWH); Dr. Joachim Ragnitz, 46, ist stellvertretender Leiter der Niederlassung Dresden des ifo Instituts.*

**Tabelle 1**  
**Themenbereiche und Daten der Rankings<sup>1</sup>**

	Bertelsmann 2004	Weltwirtschaftsforum (WEF) 2005/2006			Institute for Management Development (IMD) 2005
		Makro- ökono- mischer Index (GCI)	Mikro- ökono- mischer Index (BCI)	Mikro- und makroökono- mischer Index (GlobalCI)	
<i>Allgemeines</i>					
Entwicklungsniveau <sup>2</sup>	-	-	-	5/3	8/6
<i>Allgemeine Rahmenbedingungen</i>					
Institutionelle Rahmenbedin- gungen	-	0/7	0/10	0/9	6/30
Makroökonomi- sche Rahmen- bedingungen	4/0	5/2	-	4/0	8/1
Physische Infrastruktur	-	-	0/5	1/5	9/6
Technologie	-	7/11	0/2	4/11	29/12
Humankapital	-	-	0/3	3/5	6/7
Sozio-kulturelle Faktoren	-	-	-	-	3/13
<i>Akteursbezogene Kennzahlen</i>					
Staat	3/0	2/1	-	2/2	12/4
Unternehmen	-	-	0/17	0/12	2/13
Ausland	-	-	-	-	20/3
<i>Marktbezogene Kennzahlen</i>					
Arbeitsmarkt	9/0	-	-	0/7	16/9
Gütermarkt	-	-	0/1	5/7	-
Finanzmarkt	-	-	0/3	0/5	9/9
Insgesamt	16/0	14/21	0/41	24/66	128/113

<sup>1</sup> Der erste Wert bezeichnet harte/der zweite Wert subjektive Variablen.

<sup>2</sup> Lebenserwartung, Gesundheit, Umwelt.

Quelle: Eigene Darstellung nach Angaben des Weltwirtschaftsforums (WEF), des Institute for Management Development (IMD) und der Bertelsmann Stiftung.

higkeit, technologische Leistungsfähigkeit oder als Fähigkeit zur Wohlstandsmehrung verstanden.<sup>6</sup> Die Integration dieser partiellen Aspekte bereitet nicht nur bezüglich ihrer Gewichtung Probleme, sondern auch deshalb, weil sie auf unterschiedliche Fristen und entsprechend auch auf unterschiedliche Indikatoren abstellen. Eine geschlossene oder gar modellmäßige Erklärung der „Internationalen Wettbewerbsfähigkeit“ von Volkswirtschaften liegt jedenfalls bislang nicht vor. Die untersuchten Rankings zeigen sich davon freilich unbeeindruckt, wie die Wahl der von ihnen betrachteten Felder und der dazu verwendeten Indikatoren

<sup>6</sup> Vgl. dazu z.B. M. Stierle: Volkswirtschaften in der Globalisierung: Konzeptionelle Analyse der zentralen Indikatoren der internationalen Wettbewerbsfähigkeit von Volkswirtschaften. Speyerer Forschungsberichte, 182, Speyer 1998; H. Trabold: Die internationale Wettbewerbsfähigkeit einer Volkswirtschaft, in: DIW-Vierteljahreshefte zur Wirtschaftsforschung, 64, 2, 1995, S. 169 ff.

deutlich machen. Erörterungen dieser zentralen Frage fehlen weitgehend.

Im Einzelnen wurden der Growth Competitiveness Index (GCI), der Business Competitiveness Index (BCI) und der Global Competitiveness Index des Weltwirtschaftsforums (WEF 2005-2006), sodann der World Competitiveness Index des Institute for Management Development (IMD 2005) sowie der Erfolgs- und Aktivitätsindex 2004 der Bertelsmann Stiftung betrachtet.<sup>7</sup> Tabelle 1 gibt einen Überblick über die Analyseebenen, die einbezogenen Aktivitätsfelder und die Art der verwendeten Daten.<sup>8</sup> Mit der Verwendung von Strom- und Bestandsgrößen, Absolutzahlen und Verhältniszahlen, objektiven und subjektiven Daten weist die inhaltliche wie die empirische Basis der Rankings ein breites Spektrum auf.

Dem Ranking der Länder liegen Punkteberechnungen (Scores) zugrunde, die selbst in isolierter Betrachtung schon relativen Charakter tragen. Wie die Berechnungs- und vor allem die Standardisierungsvorschriften erkennen lassen, werden sie vom Kreis der einbezogenen Länder (dem Durchschnitt oder den jeweils „Besten“) und von der Zahl der Variablen sowie dem Betrachtungszeitraum bestimmt. Dies hat weitreichende Konsequenzen für die Interpretation der Rankingergebnisse und illustriert, wie schwierig offenbar eine objektive Messung von Wettbewerbsfähigkeit ist.<sup>9</sup> Das Vorgehen der Rankings widerspricht einem wesentlichen Aspekt des Indexgedankens, der den zeitlichen Vergleich zum Ziel hat und eine Referenzbasis erfordert. Insofern sind intertemporale Vergleiche der Ranking-Ergebnisse unzulässig, was weder die Studien noch deren mediale Darstellungen zu bekümmern scheint. Es gibt also immer „Verlierer“, selbst wenn sich alle Länder in allen Kennziffern verbessern und dort Ergebnisse erzielt werden, die eigentlich „ausreichend“ sind.

Die Breite der Rankings ist beeindruckend: Der Global Competitiveness Report 2005-2006 (WEF) untersucht 117 Länder (1996: 49) und stützt sich dabei auf 142 Kennzahlen (1996: 300). Konkret geht allerdings

<sup>7</sup> Vgl. zu den einzelnen Indizes U. Heilemann, H. Lehmann, J. Ragnitz, a.a.O., S. 28 ff. – Zu allen Rankings sind neue Ausgaben erschienen. Methodische Änderungen sind lediglich beim WEF zu verzeichnen – der ursprüngliche Growth Competitiveness Index (GCI) wurde eingestellt und durch den Global Competitiveness Index ersetzt. Letzterer hat nun unglücklicherweise das gleiche Kürzel erhalten (GCI), was beim Vergleich mit älteren Ausgaben zu beachten ist.

<sup>8</sup> Vgl. ebenda, S. 23 ff.

<sup>9</sup> Es ist wenig einsichtig, warum ein internationaler Vergleich nicht anhand einer zunächst autonomen Länderanalyse vorgenommen werden kann, vor allem angesichts der damit verbundenen Vorteile bei der Identifikation von intertemporalen Änderungen. Vgl. im Gegensatz dazu S. Gundel, U. van Suntum, a.a.O., S. 479.

**Tabelle 2**  
**Rankingergebnisse für 21 OECD-Länder<sup>1</sup>**

	Bertelsmann 2004	Institute for Management Development (IMD) 2005	Weltwirtschafts- forum (WEF) 2005			
	Erfolgs- index	Aktivi- täts- index	World Compe- titiveness Index (WCI)	Growth Compe- titiveness Index (GCI)	Business- Compe- titiveness Index (WCI)	Global Compe- titiveness Index (GobalCI)
Austra- lien	3	1	6	7	13	14
Belgien	15	18	16	19	14	15
Däne- mark	9	11	4	4	4	3
<b>Deutsch- land</b>	<b>21</b>	<b>20</b>	<b>15</b>	<b>12</b>	<b>3</b>	<b>5</b>
Finnland	17	17	3	1	2	2
Frank- reich	20	19	17	18	10	10
Grie- chenland	14	16	20	20	21	21
Großbri- tannien	8	8	14	10	5	7
Irland	1	3	7	16	16	16
Italien	19	21	21	21	20	20
Japan	13	9	13	9	7	8
Kanada	10	7	2	11	12	11
Neusee- land	5	2	11	13	15	17
Nieder- lande	6	13	8	8	8	9
Norwe- gen	4	6	10	6	17	13
Öster- reich	7	12	12	14	9	12
Portugal	16	10	19	15	19	19
Schwe- den	11	14	9	3	11	6
Schweiz	12	5	5	5	6	4
Spanien	18	15	18	17	18	18
USA	2	4	1	2	1	1

<sup>1</sup> Der Vergleich orientiert sich am Bertelsmann-Ranking. Die Rangziffern der WEF- bzw. IMD-Indizes wurden an dessen Länderkreis angepasst und weichen deshalb von deren ursprünglichen Werten ab.

Quelle: Eigene Berechnungen nach Angaben der Bertelsmann Stiftung, des Weltwirtschaftsforums (WEF) und des Institute for Management Development (IMD).

nur ein Teil in ihre Rankings ein. Subjektiven Daten (Befragungsergebnisse von 11 000 Befragten gegenüber 2000 in 1996) kommt dabei ein großes Gewicht zu. Die Zahl der pro Land befragten Unternehmen ergibt sich offenbar eher zufällig (für Deutschland: 78). Im Einzelnen weist der Report 2005-2006 drei Indizes aus: Mit dem Business Competitiveness Index sollen die das Wohlfahrtsniveau bestimmenden mikroökonomischen Faktoren gemessen werden, wohingegen der Growth Competitiveness Index auf das Wirtschaftswachstum zielt und dazu die Faktoren „makroöko-

nomisches Umfeld“, „öffentliche Institutionen“ und „technologische Leistungsfähigkeit“ unterscheidet.<sup>10</sup>

Der World Competitiveness Index des IMD untersucht in der Ausgabe des Jahres 2005 60 Länder einschließlich neun subnationaler Regionen (1996: 46), dabei wurden 314 Kennzahlen herangezogen (1996: 224). Es wurden 4000 Experten befragt (1996: 3150), den so ermittelten „weichen“ Kennzahlen wird ein Drittel des Gesamtgewichts zugewiesen. Die Zahl der Befragten wird proportional zur wirtschaftlichen Bedeutung des jeweiligen Landes festgelegt, eine genaue Aufschlüsselung fehlt. Bei der Indexkonstruktion wird ausdrücklich die Vorgabe einer Zielgröße wie „Wohlstandsniveau“ abgelehnt, da dies Ausdruck vergangener Wettbewerbsfähigkeit oder Ergebnis natürlicher Ressourcenausstattung und -nutzung sein könne.

Das Internationale Standort-Ranking der Bertelsmann Stiftung vergleicht 21 westliche OECD-Staaten. Es stützt sich ausschließlich auf objektive Daten, fast ausnahmslos solche der OECD. Die „aktuelle Standortstärke“ soll im Erfolgsindex zum Ausdruck kommen, der sich gleichermaßen auf Arbeitsmarktsituation und Wirtschaftswachstum stützt. Daneben führt ein Aktivitätsindex Variablen zusammen, die in Hinblick auf den wachstums- und beschäftigungspolitischen Erfolg als relevant angesehen werden.

Die drei Rankings gelangen zu sehr unterschiedlichen Aussagen. So variiert beispielsweise Deutschland unter den 21 OECD-Staaten je nach Ranking zwischen Rang 3 und 21, für wenige Länder ergeben sich identische Positionen (vgl. Tabelle 2).

### Theoretische Fundierung

Angesichts der Breite und Offenheit des Konzepts der „Internationalen Wettbewerbsfähigkeit“ (oder der Standortqualität) überrascht es nicht, dass sich die Rankings auf vergleichsweise weiche Methoden stützen.<sup>11</sup> Viele strukturelle Beziehungen zwischen den Indikatoren bleiben im Dunkeln, es werden vielfach (Un-)Abhängigkeiten der Komponenten und Indikatoren

<sup>10</sup> Der BCI stützt sich auf Umfragevariablen, die anhand bivariater Regressionen mit dem Bruttoinlandsprodukts pro Kopf (in Kaufkraftparitäten) auf signifikante Korrelation geprüft werden. Vgl. World Economic Forum (Hrsg.), a.a.O., S. 76 f. Der GCI wird ökonomisch auf die mittlere Ex-Post-Wachstumsrate des Bruttoinlandsprodukts pro Kopf (in Kaufkraftparitäten) konditioniert. Vgl. J. McArthur, J. Sachs: The Growth Competitiveness Index: Measuring Technological Advancement and the Stage of Development, in: World Economic Forum (Hrsg.): The Global Competitiveness Report 2001-2002, Basingstoke u.a. 2001, S. 39. Der neue Global Competitiveness Index sollte ursprünglich den GCI und BCI ersetzen (in der Ausgabe 2007-2008 wurde aber lediglich der GCI aufgegeben) und stellt den Versuch dar, den statisch orientierten BCI (Wohlstandsniveau) und den dynamisch orientierten GCI (Wohlstandswachstum) zu verbinden.

<sup>11</sup> Vgl. U. Heilemann, H. Lehmann, J. Ragnitz, a.a.O., S. 46 ff.

unterstellt, von denen nicht ohne weiteres auszugehen ist. Auch hohe Korrelationen zwischen Indikandum und Indikator sind nur bei theoretisch fundierter Spezifikation aussagekräftig, will man sich nicht dem Vorwurf des *Measurement without theory* oder dem puren *Data mining* aussetzen. Es mag gute Gründe für die Verwendung empirischer Regelmäßigkeiten geben, aber diese sollten genannt werden.

Generell sind die Bezüge der Rankings zur ökonomischen Theorie eher schwach. Ganz überwiegend gründen die verwendeten Kennzahlen auf einfachen Plausibilitätsüberlegungen, gelegentlich gestützt auf Korrelationsanalysen, allenfalls implizit werden die verwendeten Hypothesen oder Konzepte deutlich. Vor allem bei den sehr breit angelegten Rankings von WEF und IMD entsteht häufig der Eindruck, als würden alle sich irgendwie anbietenden, potenziell relevanten Bezüge zum Untersuchungsgegenstand „Internationale Wettbewerbsfähigkeit“ einbezogen. Wo objektive Daten fehlen, müssen subjektive diese Lücke schließen und spielen entsprechend häufig eine große Rolle, im Falle des Business Competitiveness Index (WEF) bestimmen sie die Ergebnisse sogar ausschließlich.

Gänzlich offen bleibt in den Rankings die Frage der Stabilität der unterstellten Beziehungen und vor allem die Frage, ob sie für alle Länder gleichermaßen gelten. Besonders erklärungsbedürftig ist die Annahme monoton-linearer Beziehungen zwischen den Kennzahlen und der „Internationalen Wettbewerbsfähigkeit“. Vor allem bei Verhältniszahlen führt dies zu Ungereimtheiten. Eine Sparquote von 100% ist so wenig optimal wie Steuersätze von 0% – was den Autoren des IMD und WEF entgangen zu sein scheint.

### Berechnungsmethodik

Die Berechnung<sup>12</sup> der Ranking-Scores basiert auf der Punktsummennmethode<sup>13</sup> – verschiedene Kennzahlen ( $x_j$ ) werden standardisiert bzw. in Punkte transformiert und anschließend mit Gewichten  $b_j$  versehen summiert:

$$\text{Rankingscore} = b_1x_1 + b_2x_2 + \dots + b_mx_m.$$

Diese Addition vieler Merkmale eines Untersuchungsobjektes bedarf – anders als die der Ausprägungen eines Merkmals – besonderer Rechtfertigung, um eine zentrale Voraussetzung der Indexbildung

– Kommensurabilität (Messen mit dem gleichen Maß) – mindestens näherungsweise zu gewährleisten.<sup>14</sup>

Die Verwendung der Indexmethode könnte mit Hinweis auf die Messfehlerreduktion oder auf die sachliche Aggregation eines Ganzen aus seinen (additiven) Bestandteilen zu rechtfertigen versucht werden. Erstere kommt für die Rankings allerdings nicht in Frage, denn die Auswahl der Kennzahlen wird von den Autoren inhaltlich und nicht datentechnisch begründet. Bleibt also die sachliche Aggregation als Rechtfertigung; das heißt die Kennzahlen als „Teil des eigentlich zu messenden Konzepts“<sup>15</sup> zu interpretieren. In diesem Fall stellen sich freilich eine Reihe von Fragen, welche die Rankings unbeantwortet lassen. Wie viele und welche Kennzahlen werden benötigt? Wie soll die Korrelation der Kennzahlen beschaffen sein oder spielt sie keine Rolle? Welcher Zusammenhang besteht zwischen der Größe des Teileinflusses und dem Gewicht der Kennzahl?

Aufgrund der Disjunktheit der Teileffekte sollte man erwarten, dass die Indikatoren unabhängig voneinander sind, eine Voraussetzung, die verletzt ist, wie eine multivariate Korrelationsanalyse zeigt.<sup>16</sup> Entgegen der Interpretation der Indizes werden redundante und sich überschneidende Informationen aufgenommen („Doppelzählungen“), was einmal mehr am Informationsgehalt der Kennzahlen zweifeln lässt. Die Gewichtung der einzelnen Kennzahlen wird nicht begründet und ist zum Teil auch abhängig von der Anzahl der Kennzahlen, die ein Subindex enthält (GCI, Global Competitiveness Index, World Competitiveness Index). Dass dies aus Mangel an besserem Wissen z.B. als Erklärung für Gleichgewichtung geschieht, ist zwar nachvollziehbar, verträgt sich aber kaum mit den wissenschaftlichen Ansprüchen der Rankings. Vor allem sollten derlei Setzungen dem Leser auch mitgeteilt werden und gegebenenfalls ihre Bedeutung durch Sensitivitätsrechnungen illustriert werden, zumal sich der rechentechnische Aufwand dazu für die Autoren in Grenzen halten dürfte.<sup>17</sup> Insofern wird auch die Interpretation der Rankingindizes als Summation von Teilaspekten „Internationaler Wettbewerbsfähigkeit“ den zu stellenden methodischen Anforderungen nicht gerecht.

<sup>14</sup> Dagegen mag man den Preisindex ins Feld führen. Letztlich fußt dieser aber auf der Definition eines Warenkorbtes, der wenig mehr ist als eine Konvention von Statistikern. Vgl. R. Poliak: The Consumer Price Index: A Research Agenda and Three Proposals, in: Journal of Economic Perspectives, 1998, 12, S. 74.

<sup>15</sup> Vgl. P. von der Lippe, A. Kladraba, a.a.O., S. 118.

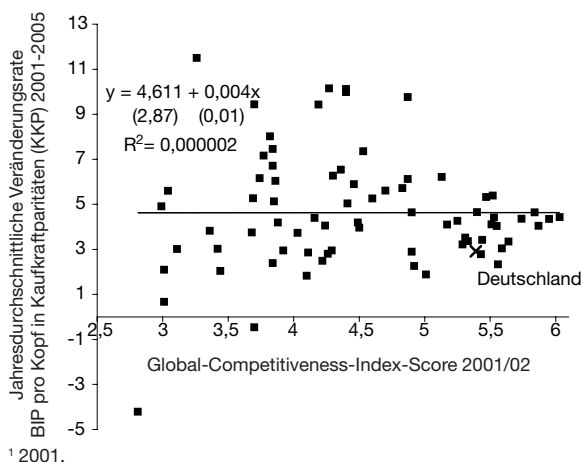
<sup>16</sup> Vgl. U. Heilemann, H. Lehmann, J. Ragnitz, a.a.O., S. 70.

<sup>17</sup> Vgl. ebenda, S. 89 ff.

<sup>12</sup> Vgl. ebenda, S. 65 ff.

<sup>13</sup> Vgl. P. von der Lippe, A. Kladraba: Messung komplexer Variablen als Summe von Punktvariablen. Eine beliebte Methode des measurement without theory, in: Jahrbücher für Nationalökonomie und Statistik, 224, 2004, S. 115 ff.

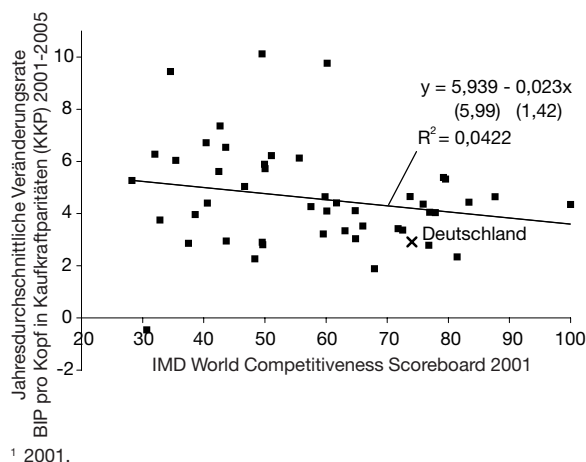
**Abbildung 1**  
**Zusammenhang von Global Competitiveness**  
**Index-Score<sup>1</sup> und Wirtschaftswachstum**  
 (t-Werte in Klammern)



<sup>1</sup> 2001.

Quelle: Eigene Darstellung und Berechnung nach Angaben des Weltwirtschaftsforums (WEF) und IMF.

**Abbildung 2**  
**Zusammenhang von World Competitiveness**  
**Index-Score<sup>1</sup> und Wirtschaftswachstum**  
 (t-Werte in Klammern)



<sup>1</sup> 2001.

Quelle: Eigene Darstellung und Berechnung nach Angaben des Institute for Management Development (IMD) und IMF.

### Empirische Basis

Die Aktualität<sup>18</sup> der Daten ist überwiegend hoch, ohne dass damit die suggerierte Zeitnähe zum Erscheinungsjahr des jeweiligen Rankings gegeben wäre bzw. beruht auf unterjährigen Schätzungen, wie im Falle des Internationalen Standort-Rankings. Beim IMD beeinträchtigen die intravariaten Aktualitätsunterschiede den Vergleich; zudem sind viele Kennzahlen des IMD „ländergrößenverzerrt“, und auch die Einbeziehung der Regionen, deren Angaben fast durchweg geschätzt wurden, schmälert den Informationsgehalt beträchtlich.

Die Verfügbarkeit/Transparenz der Daten ist im IMD durch die Vielzahl der Quellen eingeschränkt, bei den anderen Rankings ist sie tendenziell besser. Die Herkunft der Daten und Fallbesonderheiten in den Rankings selbst sind durchweg deutlich gemacht. Die zeitlich-räumliche Vergleichbarkeit der Daten ist indessen problematisch, da nationale Statistiken noch immer mehr oder minder unterschiedliche statistische Konzepte verfolgen – von den prinzipiellen Schwierigkeiten internationaler Vergleiche ganz abgesehen. Ungeachtet der beträchtlichen Bemühungen um einheitliche Systematiken, wenigstens der Angaben der Volkswirtschaftlichen Gesamtrechnung, sind internationale Vergleiche auch hier noch immer mit erheblichen Unsicherheiten behaftet.<sup>19</sup>

Die Qualität der Daten dürfte im Falle der Daten der amtlichen Statistik bzw. den darauf aufbauenden Da-

ten internationaler Organisationen gegeben sein, ohne dass auch in diesem Falle alle Zweifel ausgeräumt wären.<sup>20</sup> Die Qualität der großen Zahl nicht-amtlicher oder subjektiver Daten bleibt offen. Die Prüfung der Qualität der Survey-Daten (WEF, IMD) offenbarte angesichts hoher Antwortausfälle die Gefahr von Selektionseffekten. Bei der Expertenbefragung ergeben sich tendenziell Verzerrungen durch den „home bias“ (national geprägte Sicht) und den „perception bias“ (Abhängigkeit von Fremdstimmungen). Die Befragung richtet sich ausschließlich an Unternehmer bzw. Manager überwiegend größerer Unternehmen. Mehrheitlich sind Bereiche berührt, in denen die Befragten bestenfalls partielle Expertise besitzen. Die Antworten selbst stellen auf eine absolute Beurteilung der nationalen Verhältnisse ab und nicht wie die Rankings auf einen (internationalen) Vergleich. Eine auch nur ansatzweise Kontrolle der Verzerrungseffekte, wie dies z.B. im Rahmen der Delphi-Methode versucht wird, fehlt.

### Robustheit und prognostische Güte

Während die Reproduzierbarkeit der Ergebnisse – naturgemäß mit beträchtlichem Aufwand – gegeben ist,<sup>21</sup> überrascht allerdings, dass sich insbesondere beim WEF im Durchschnitt 1,2 (BCI) bis 2,2 (Global

<sup>18</sup> Vgl. ebenda, S. 71 ff.

<sup>19</sup> Vgl. ebenda, S. 76.

<sup>20</sup> Überschlägig entstammen ca. 60% der harten Kennzahlen des Weltwirtschaftsforums (70% Institute for Management Development, 100% Bertelsmann Stiftung) der amtlichen Statistik der einzelnen Länder oder von internationalen Organisationen. Der Trade-off zwischen Aktualität und Genauigkeit führt im Allgemeinen dazu, dass Angaben am aktuellen Rand kritischer zu sehen sind.

<sup>21</sup> Vgl. U. Heilemann, H. Lehmann, J. Ragnitz, a.a.O., S. 81 ff.

Competitiveness Index) Rangunterschiede je Land ergeben (beim GCI im Extremfall für Kanada sogar 8 Rangplätze Unterschied).

Ein zentrales Kriterium der Beurteilung von Indizes zur Bewertung der Wachstumsfähigkeit (Global Competitiveness Index des WEF, World Competitiveness Index des IMD) sollte ihre Prognoseleistung sein. Die diesbezüglichen Ergebnisse sind allerdings ernüchternd. Gemessen an der Veränderungsrate des realen BIP pro Kopf (in Kaufkraftparitäten) ist keinerlei Zusammenhang mit dem Ranking der Länder zu erkennen.<sup>22</sup> Mit anderen Worten: Die untersuchten Rankings haben keinen prädiktischen Gehalt (Abbildung 1 und Abbildung 2) und dem Leser bleibt rätselhaft, worauf sich dann ihre Ansprüche gründen, die Kommunikation über nationale wirtschaftspolitische Erfordernisse zu stimulieren und vor allem in die „richtige“ Richtung zu lenken.<sup>23</sup>

Bei den subjektiven Setzungen stellt sich, wie bereits erwähnt, die Frage, wie sich die Resultate bei Variation des Verdichtungsschemas ändern.<sup>24</sup> Im Einzelnen wurde dem Einfluss von weichen bzw. harten Kennzahlen, der Gewichtung und alternativer Aufbereitungsweisen der Kennzahlen nachgegangen. Es zeigte sich, dass die Ergebnisse in hohem Maße von der theoretisch wenig befriedigenden Zusammenstellung und Berechnungsweise abhängen und deren Variation zu beträchtlichen Abweichungen von den Original-Ergebnissen führt. In Einzelfällen ergaben sich Unterschiede von 30 und mehr Rängen, im Durchschnitt immer noch von 9 (GCI), 11 (Global Competitiveness Index) oder 7 (World Competitiveness Index), was einmal mehr die hohe Subjektivität – zurückhaltend ausgedrückt – der Rangfolgen illustriert, die dem Leser leider vorenthalten wird.

Für den Erfolgsindex des Internationalen Standort-Rankings wurde untersucht, wie sich eine Erweiterung des Zielgrößenkatalogs um Preisstabilität und Außenwirtschaftsgleichgewicht – also erklärte Ziele der deutschen Wirtschaftspolitik – auswirkt. Es ergaben sich wiederum gänzlich andere Bewertungen, wobei dann auch Deutschland – je nach Simulation, wie zu erwarten, deutlich besser abschnitt. Aus berechnungsmethodischer Sicht sind die Rankings insofern also instabil und bieten viel „Gestaltungsspielraum“.

<sup>22</sup> Vgl. ebenda, S. 85 ff.

<sup>23</sup> Vgl. z. B. R. Fendel, M. Frenkel: Wozu Studien zur Wettbewerbsfähigkeit von Volkswirtschaften?, in: WIRTSCHAFTSDIEST, 85. Jg. (2005), Heft 1, S. 26 ff.

<sup>24</sup> Vgl. hierzu und im Folgenden U. Heilemann, H. Lehmann, J. Ragnitz, a.a.O., S. 89 ff.

### Komplexitätsreduktion oder Zahlenalchemie?

Zusammenfassend: Die Länder-Rankings bündeln eine Vielzahl von objektiven und teilweise auch subjektiven Daten, um damit die relative Standortqualität oder die „Internationale Wettbewerbsfähigkeit“ zu bewerten. Auswahl und Gewichtung der einzelnen Indikatoren sind vielfach nicht unplausibel, ihre Bedeutung für den jeweiligen Ranking-Score bleibt jedoch im Dunkeln. Eine theoretische Fundierung des Konzepts, seiner Elemente und ihres Gewichts fehlt, ein Manko, welches auch dadurch nicht geheilt wird, dass hinter Konzeptionierung und Durchführung der Rankings ausgewiesene – wenn auch auf anderen Feldern – Autoren stehen. Die empirischen Grundlagen der Rankings sind gut dokumentiert, die Ergebnisse reproduzierbar, eine Ausnahme bilden die verwendeten, zum Teil sehr einflussreichen subjektiven Daten. Im Einzelnen sind jedoch erhebliche methodische Vorbehalte anzumelden. Als besonders kritisch erwies sich die geringe prognostische Leistungsfähigkeit der Rankings, würde man doch erwarten, dass dieses Kriterium bei der Indikatorenselektion und -gewichtung leitend gewesen wäre.

Das reduktionistische Bild von einer sich ständig wandelnden komplexen ökonomischen Wirklichkeit weist mithin zahlreiche blinde Flecken auf. Die Schwächen von Daten und Methoden lassen eine Bildung von Ähnlichkeitsklassen angemessener erscheinen als eine scharfe dezimale Ordnung. Wer immer der Adressat der Rankings sein soll – der Wirtschaftspolitik helfen sie wenig. Spezifische Handlungsfelder oder Ansatzpunkte für die Politik werden nicht identifiziert, die reklamierte beanspruchte Signalfunktion ist wegen der Vernachlässigung definitorischer und funktionaler Beziehungen stark eingeschränkt. Fast durchweg bleibt die Rolle konjunktureller, struktureller oder akzidenteller Faktoren offen. Dass deren Herausarbeitung angesichts der Breite des Erklärungsziels und der gesuchten Aktualität nur mit beträchtlichem Aufwand und mit Blick auf die Vielzahl subjektiver Daten nur partiell zu leisten ist, liegt auf der Hand. Als entscheidende Mängel erweisen sich letztlich die nur schwer zu rechtfertigende ordinale Abstufung der Länder, vor allem aber die geringe Prognoseleistung der untersuchten Rankings.

### Zur Kritik von Gundel/van Suntum

In ihrem Beitrag kritisieren Gundel/van Suntum vor allem – systematisiert ausgedrückt – Methoden, Ergebnisse und Darstellungsweise unserer Untersuchung.<sup>25</sup> Die für uns zentrale Frage der Bewertung der

<sup>25</sup> Vgl. S. Gundel, U. van Suntum, a.a.O.

Rankings und ihrer Ergebnisse wird dagegen allenfalls beiläufig behandelt. Ein Teil der Kritik sollte sich für den Leser nach den vorangegangenen Abschnitten erledigt haben, so dass wir nur auf die noch klärungsbedürftigen Teile eingehen.

Generell bemängeln die beiden Verfasser, unsere Analyse sei unsystematisch und voller Fehler, insbesondere auch die „Vorzüge“ des Internationalen Standort-Rankings würden nicht deutlich oder gar absichtsvoll unterschlagen, und es werde pauschale Rankingschelte betrieben. Die Rankings würden im Gegenteil eine hohe Übereinstimmung der Befunde aufweisen, gerade auch für Deutschland.<sup>26</sup>

Dass unsere Analyse keineswegs erschöpfend ist, liegt angesichts des vorgegebenen Rahmens und der Literatur zum Thema auf der Hand, zumal die bisherigen Auseinandersetzungen eher grundsätzlicher und weniger empirischer Art sind.<sup>27</sup> Von einer Unterschlagung der „Vorzüge“ des Bertelsmann-Rankings kann keine Rede sein, letztlich aber zählt nur, ob es gelungen ist, die Komplexität hinreichend sachgerecht zu reduzieren. Rankings mögen mediale Nachfrage befriedigen, aber diesem Erfordernis könnte fraglos überzeugender und zum Teil auch einfacher entsprochen werden. Gerne bescheinigen wir Gundel/van Suntum auch explizit, dass sich die Schwächen im Detail vor allem bei ihren beiden Konkurrenten häufen. Gleichwohl ist auch das Vorgehen des Bertelsmann-Rankings in wichtigen Punkten zu kritisieren (siehe unten). In der zentralen Frage des Informationsgehalts oder dem Neuigkeitswert der Rankings wird angesichts des diffusen Verständnisses von „Internationaler Wettbewerbsfähigkeit“ und „Standortstärke“ im Grunde nichts Neues bzw. Überzeugendes geboten. Von Objektivität der Rankings wird durch das zum Teil große Gewicht subjektiver Daten (bei WEF und IMD) und subjektiver Gewichtungen, die sich zudem nur nach aufwändiger Lektüre erschließen, kaum die Rede sein. Weder die Grundkonzeptionen noch die theoretische, methodische oder empirische Auffüllung der untersuchten Rankings ist überzeugend, auch wenn im Einzelnen die Vorbehalte unterschiedlich ausfallen.

Vielleicht tröstet es die beiden Kritiker, wenn wir daran erinnern, dass es nicht um ein Ranking der Rankings ging. Ob aus unserer Untersuchung das „Bertelsmann-

Ranking als klarer Sieger des Rankingvergleichs von der Pauschalkritik [hätte] ausgenommen werden müssen“, wie die beiden Autoren an mehreren Stellen postulieren, bleibt dem Leser der Rankings oder unseres Gutachtens überlassen. Es ging uns auch nicht um die Präsentation von Alternativen. Wer das Gutachten liest, wird durchaus Anhaltspunkte für sachgerechtes Ranking finden – Objektivität, Reliabilität und Validität wären, wenig überraschend, zentrale Kriterien. Ob das unter den gegebenen Umständen zu leisten ist, müssen sich die Produzenten fragen. An mehreren Stellen im Gutachten haben wir deutlich gemacht, dass sich auch nach Beseitigung einzelner von uns diagnostizierter Mängel der untersuchten Rankings an der Position Deutschlands nicht viel ändern würde, was nach Lage der Dinge auch nicht zu erwarten war.

### Die Kritikpunkte im Einzelnen

Die Aussagekraft eines Rankings könne nur daran gemessen werden, was es eigentlich untersuchen will.<sup>28</sup> – Genau dieser Anspruch bleibt für den Leser ebenso diffus wie die Konzepte „Internationale Wettbewerbsfähigkeit“ oder „Standortqualität“. Das Interesse, das die Rankings in der Öffentlichkeit finden, und die Bedeutung, die ihnen zugeschrieben wird, lässt ihnen einen für viele fast divinen und von den Rankings auch nicht relativierten Anspruch zuwachsen, das Präsentationen des gleichen Informationsgehaltes in weniger sportlichem Gewand nie zuteil würde.<sup>29</sup> Die Autoren des Internationalen Standort-Rankings lehnen nach eigener Aussage den problematischen Begriff der Wettbewerbsfähigkeit ab.<sup>30</sup> Wenn es aber im ersten Satz auf der Rückseite des Bucheinbandes ihrer Veröffentlichung heißt – „Wie wettbewerbsfähig ist der Wirtschaftsstandort Deutschland im internationalen Vergleich?“<sup>31</sup> –, so belegt dies einen anderen Anspruch, der so auch in der Öffentlichkeit wahrgenommen wird.<sup>32</sup>

<sup>26</sup> Vgl. S. Gundel, U. van Suntum, a.a.O., S. 474.

<sup>29</sup> In Entgegnungen auf Kritik an reduzierten Ländervergleichen findet sich, anders als in den Rankings selber, mitunter der relativierende Hinweis, dass sie vor allem politische und öffentliche Aufmerksamkeit auf wirtschaftliche und gesellschaftliche Defizite lenken sollen. Warum dazu der mühsame Umweg über die Rankings mit allen hier aufgeführten Vorbehalten gewählt werden muss, bleibt dabei aber offen. Vgl. dazu R. Fendel, M. Frenkel, a.a.O., offenbar als Replik auf U. Heilemann: Deutschland im internationalen Vergleich – einige Fragezeichen, in: WIRTSCHAFTSDIENST, 84. Jg. (2004), Heft 12, S. 761 ff.

<sup>30</sup> Vgl. S. Gundel, U. van Suntum, a.a.O., S. 473 f.

<sup>31</sup> Vgl. Bertelsmann Stiftung (Hrsg.), a.a.O.

<sup>32</sup> Vgl. A. Struve: Die Wettbewerbsfähigkeit ganzer Volkswirtschaften ist nur schwer zu messen, in: Die Welt, Jg. 49, 14.10.2004, S. 12; K. Horn: Zwischen Genf und Gütersloh. Zur Aussagefähigkeit von Standort-Rankings, in: Frankfurter Allgemeine Zeitung, 18.10.2004, S. 13; D. Hess: Rote Laterne bleibt in Deutschland, in: Handelsblatt, 20.4.2006, S. 3.

<sup>26</sup> Vgl. ebenda, S. 479.

<sup>27</sup> Vgl. S. La11: Competitiveness Indices and Developing Countries: An Economic Evaluation of the Global Competitiveness Report, in: World Development, 29, 2001, S. 1501 ff.; J. Matthes: Die Position Deutschlands in Rankings zur internationalen Wettbewerbsfähigkeit, in: IW-Trends: Vierteljahresschrift zur empirischen Wirtschaftsforschung aus dem Institut der deutschen Wirtschaft Köln, H. 4, 2005, S. 77 ff.; oder W. Oche1, O. Röhn: Ranking of Countries – The WEF, IMD, Fraser and Heritage Indices, in: CESifo DICE Report, 4 (2), 2006, S. 48 ff.

Es sei wegen der unterschiedlichen Ausrichtung der Rankings gar nicht zu erwarten, dass die einzelnen Rankings zu vergleichbaren Ergebnissen kommen.<sup>33</sup> – Am Ende ihres Beitrags stellen Gundel/van Suntum allerdings fest, die verschiedenen Rankings zeigen „in Wahrheit trotz ganz unterschiedlicher Methodik ein hohes Maß an Übereinstimmung in den Ergebnissen, gerade auch in Bezug auf das schlechte Abschneiden Deutschlands“<sup>34</sup>. Was immer den Sinneswandel der Autoren auslöste und was mit „Übereinstimmung“ gemeint ist (siehe oben), der Nachweis der unterschiedlichen Gesamtbewertung der Länder ist wichtig (vgl. Tabelle 2). Er zeigt einmal mehr, dass es kein allgemein akzeptiertes oder verwendetes Konzept von Wettbewerbsfähigkeit gibt, auch wenn die Autoren der Rankings anderes glauben machen wollen.<sup>35</sup> Dies ist die zentrale Botschaft und nicht die triviale Feststellung, dass in Bereichen wie Staatsverschuldungsquote, Arbeitslosenquote etc. Deutschland in den letzten Jahren nicht zu den „besten“ Ländern gehört hat. Dazu hätte ein Blick in das Statistische Jahrbuch genügt.

Die Gegenüberstellung der Rankings anhand der Rangkorrelation sei unvollständig und die berechnete Deckungsgleichheit sei zudem recht hoch. Gundel/van Suntum stellen dazu eigene Berechnungen an,<sup>36</sup> die sie sich allerdings mit einem Blick in unser Gutachten hätten sparen können.<sup>37</sup> Es bleibt dem Leser überlassen, in Korrelationen zwischen dem Aktivitätsindex der Bertelsmann Stiftung und dem Global Competitiveness Index (WEF) von 0,41 bzw. dem World Competitiveness Index (IMD) von 0,57 ähnliche Befunde zu sehen.

Die Bewertung der finanzpolitischen Leistung sei nicht ausreichend fundiert und führt in allen Rankings zu einer ungünstigen Bewertung der deutschen Position.<sup>38</sup> – Eine solche Sicht ignoriert – wie oben ausgeführt –, dass die Qualität der Rankings grundsätzlich nicht anhand der Befunde nur eines Landes (Deutschland) diskutiert werden sollte. Dass Deutschland bei Kennziffern wie der Staatsverschuldung oder der Staatsquote in verschiedenen Rankings ähnlich

ungünstige Werte erzielt, sollte allein mit Blick auf die Kosten der deutschen Einheit und die schleppende gesamtwirtschaftliche Entwicklung nicht erstaunen. Das entworfen Bild bleibt zudem unvollständig, wenn es die Ausgabenseite (z.B. eben die Kosten der Eini-gung in Höhe von 4-5% des BIP) und die Effizienz der Aufgabenerfüllung der öffentlichen Hand außer Acht lässt.<sup>39</sup> Dem Befund von Gundel/van Suntum, dass der Bereich fiskalischer Leistungsvergleiche sehr viel mehr Aufmerksamkeit verdient, als es üblicherweise der Fall ist, mag man zwar zustimmen, aber dies ist kein Manko unserer Untersuchung, die sich auf vorlie-gende Rankings stützt.

Ausführlich legen Gundel/van Suntum dar, wes-halb das Gutachten aus ihrer Sicht auch in der meth-odischen Vorgehensweise nicht überzeugt.<sup>40</sup> – Bei näherer Betrachtung lösen sich die Vorwürfe freilich rasch auf. Der Vorwurf, beim Bertelsmann-Ranking sei unverständlicherweise eine intertemporale Stabilitäts-betrachtung seiner Ergebnisse unterlassen werden, ist ungerechtfertigt. Die intertemporale Korrelation der Rankings ist aufgrund der oben erwähnten sich jähr-lich ändernden Referenzpunkte im Grunde unzulässig. Leider sind allerdings auch wir in unserer Unters-uchung der Versuchung erlegen, die Ergebnisse inter-temporal zu interpretieren.<sup>41</sup>

Gundel/van Suntum monieren ferner, die ökonomet-rische Fundierung des Aktivitätsindex würde nicht ausreichend gewürdigt. – Die ökonometrische Aus-wahl der Kennzahlen im Aktivitätsindex ist – worauf wir hingewiesen haben<sup>42</sup> – keineswegs ein Spezifikum des Bertelsmann-Rankings. Die behaupteten hoch-signifikanten Zusammenhänge der zwölf Kennzahlen des Aktivitätsindex mit jeweils einer Kennzahl des Erfolgsindex<sup>43</sup> treffen selektiv zu.<sup>44</sup> Zudem scheinen die beiden Kritiker zu übersehen, dass zwischen Signi-fikanz und Einflussstärke eines Regressors zu trennen ist (standardisierte Schätzkoeffizienten sind nicht aus-gewiesen). Zudem sind die Regressionsbeziehungen in der gewählten Form nicht ohne weiteres kausal zu interpretieren. Sie mögen hilfreich bei der Exploration theoretischer Fundierung sein, aber sie ersetzen sie

<sup>33</sup> Vgl. S. Gundel, U. van Suntum, a.a.O., S. 474 f. Die Aussage, wonach allein der Aktivitätsindex den Vorzug hätte ökonomisch fundiert zu sein, ist dabei falsch. Auch der GCI und BCI – siehe Fußnote 10 – werden explizit ökonomisch hergeleitet.

<sup>34</sup> Vgl. ebenda, S. 479.

<sup>35</sup> Man lese die entsprechenden Vorworte der Rankings von WEF und IMD verschiedener Jahrgänge.

<sup>36</sup> Vgl. S. Gundel, U. van Suntum, a.a.O., S. 474 f.

<sup>37</sup> Vgl. U. Heilemann, H. Lehmann, J. Ragnitz, a.a.O., S. 42, S. 143.

<sup>38</sup> Vgl. S. Gundel, U. van Suntum, a.a.O., S. 476.

<sup>39</sup> Darüber hinaus müsste man die Bedeutung temporärer Sonderlasten für beide Seiten der fiskalischen Aktivität – z.B. für Deutschland aus der Wiedervereinigung – angemessen berücksichtigen.

<sup>40</sup> Vgl. S. Gundel, U. van Suntum, a.a.O., S. 476-479.

<sup>41</sup> Vgl. U. Heilemann, H. Lehmann, J. Ragnitz, a.a.O., S. 69.

<sup>42</sup> Vgl. ebenda, S. 41, 57.

<sup>43</sup> Arbeitslosenquote, Erwerbstätigenzuwachs, Bruttoinlandsprodukt pro Kopf, Potentialwachstum.

<sup>44</sup> Vgl. Bertelsmann Stiftung (Hrsg.), a.a.O. S. 138 f.



nicht.<sup>45</sup> Der Interdependenz der Erfolgsgrößen und vieler Aktivitätsgrößen werden Einzelgleichungsregressionen jedenfalls nicht gerecht.

Des Weiteren wird kritisiert, die Simulationsrechnungen zur Robustheit des „Erfolgsindex“ seien willkürlich, durch unsere „beliebige“ Erweiterung des Erfolgsindex um die Ziele Preisstabilität und Leistungsbilanzüberschuss könne man jegliche Rangfolge erzeugen. – Das Argument wendet sich indessen gegen die beiden Kritiker: illustriert es doch – anhand durchaus diskussionswürdiger Kriterien – die hohe Sensibilität der Rankings und unterstreicht einmal mehr die Notwendigkeit, die verwendeten Kriterien überzeugend zu begründen. Wie ausgeführt ging es uns keineswegs um eine „Verbesserung“ des Erfolgsindex, obwohl auch in materieller Sicht über die Einbeziehung der beiden Kriterien nachzudenken ist.

Ein weiterer Kritikpunkt betrifft den Verzicht auf die Untersuchung des prognostischen Gehalts des Aktivitätsindex. – In der Tat wurde die Untersuchung der Ex-post-Prognosequalität auf den Growth Competitiveness Index, der explizit auf die Wachstumsrate des Bruttoinlandsprodukts/Kopf konditioniert ist, und den World Competitiveness Index (IMD), der explizit auf eine Erklärung von Wohlstandsniveau und -wachstum zielt,<sup>46</sup> beschränkt. Die Analyse von Gundel/van Suntum für den Erfolgsindex 2004 und den Aktivitätsindex 2001<sup>47</sup> kann ebenfalls nicht überzeugen. Durch die Identität der zur Hypothesengewinnung und zu ihrem Test verwendeten Daten wird, erstens, ohne Not die Macht des Tests sehr eingeschränkt und, zweitens, sind die Ergebnisse nicht mit denen für die beiden anderen Indizes vergleichbar. Von dem problematischen intertemporalen Vergleich der Rankings war bereits mehrfach die Rede.

Schließlich vermissen Gundel/van Suntum das Anbieten von Alternativen, die Kritik im Gutachten sei auf jegliche empirische Forschung anwendbar.<sup>48</sup> – Das Anbieten von „Alternativen“ war nicht Gegenstand unserer Untersuchung, aber alle Elemente unserer Kritik sind mindestens implizit Hinweise auf Verbesserungsmöglichkeiten. Was den zweiten Teil des Satzes angeht, so steht in der Tat jegliche – kognitiv wie handlungsorientierte – empirische Forschung in einem Spannungsfeld von Theorie, Daten und Methoden. Für etablierte Felder und Beziehungen wie die

makroökonomische Konsumfunktion oder die Preisbildung homogener Güter, trifft er nach vermutlich weit verbreiteter Einschätzung weniger zu, für ein noch immer terminologisch und theoretisch so unsicheres Terrain, wie „Internationale Wettbewerbsfähigkeit“ oder „Standortstärke“, eben mehr. Auch wenn sich die Möglichkeiten für den internationalen Vergleich von Volkswirtschaften seit mehr als 300 Jahren fraglos verbessert haben, sind sie eben noch immer begrenzt. Dies gilt erst recht, wenn nach einer objektiven Verdichtung etlicher Dutzend Aspekte zu einer einzigen „Zahl“ gesucht wird. Reduktionismus hat seinen Preis – auch für ihre Adepten.

### Zusammenfassung

Es fällt uns schwer, in der Kritik von Gundel/van Suntum Überzeugendes gegen unsere Untersuchung und ihre Ergebnisse auszumachen. Mit Recht können Gundel/van Suntum darauf hinweisen, dass einige Kritik an den beiden anderen Rankings auf das Bertelsmann-Ranking weniger oder gar nicht zutrifft. Dies wurde in der Arbeit auch vermerkt, freilich ohne damit in ein Ranking der Rankings einzutreten.

Mit vergleichsweise wenig Aufwand ließen sich einige der methodischen Mängel und Schwächen der Datenbasis beseitigen. Sehr viel mehr Mühe dürfte eine überzeugende Operationalisierung der „Internationalen Wettbewerbsfähigkeit“ und der „Standortstärke“ bereiten. Das Gleiche trifft auf die nicht zu objektivierenden Gewichtungen der einzelnen Faktoren und Indikatoren zu. Auf jeden Fall sollte der Leser über den Einfluss von Setzungen und Gewichtungen z.B. anhand von Sensitivitäts- oder Robustheitsrechnungen ins Bild gesetzt werden. Letzteres gilt a fortiori für die prognostische Leistungsfähigkeit der Rankings.

Der reduktionistische Ansatz der Rankings schmälert zwangsläufig ihren Nutzen für die Wirtschaftspolitik<sup>49</sup> – selbst ihrer „Alarmfunktion“ kann sie leicht entraten. Aber auch wenn die Ranking-Industrie nur eine öffentliche oder mediale Nachfrage mit wissenschaftlichem Anspruch bedienen will, sollten Objektivität, Reliabilität und Validität sehr viel stärker Beachtung finden, als dies bislang der Fall ist. Freilich würden die Rankings dann viel vom Charme der simplen Zahl verlieren, und ob die Ergebnisse dann noch Aufmerksamkeit wie bisher finden, ist ungewiss. Einfachheit konkurriert nun einmal mit Genauigkeit – nicht nur in der empirischen Wirtschaftsforschung.

<sup>45</sup> Vgl. S. Gundel, U. van Suntum, a.a.O., S. 478.

<sup>46</sup> Vgl. U. Heilemann, H. Lehmann, J. Ragnitz, a.a.O., S. 35.

<sup>47</sup> Vgl. S. Gundel, U. van Suntum, a.a.O., S. 478.

<sup>48</sup> Vgl. ebenda, S. 479.

<sup>49</sup> Der Benchmarking-Ansatz ist in dieser Hinsicht tendenziell vielversprechender. Vgl. U. Heilemann, H. Lehmann, J. Ragnitz, S.103 ff.