

van Suntum, Ulrich; Rohwer, Anja; Hülsewig, Oliver; Heilemann, Ullrich;
Schmidt, Christoph M.

Article — Published Version

Rankings und Indikatoren: Hilfreiche Kennzahlen oder Verkürzungen mit fatalen Folgen?

Wirtschaftsdienst

Suggested Citation: van Suntum, Ulrich; Rohwer, Anja; Hülsewig, Oliver; Heilemann, Ullrich;
Schmidt, Christoph M. (2011) : Rankings und Indikatoren: Hilfreiche Kennzahlen oder
Verkürzungen mit fatalen Folgen?, Wirtschaftsdienst, ISSN 1613-978X, Springer, Heidelberg,
Vol. 91, Iss. 11, pp. 735-749,
<http://dx.doi.org/10.1007/s10273-011-1295-6>

This Version is available at:
<http://hdl.handle.net/10419/88825>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen
Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle
Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich
machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen
(insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten,
gelten abweichend von diesen Nutzungsbedingungen die in der dort
genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

*Documents in EconStor may be saved and copied for your
personal and scholarly purposes.*

*You are not to copy documents for public or commercial
purposes, to exhibit the documents publicly, to make them
publicly available on the internet, or to distribute or otherwise
use the documents in public.*

*If the documents have been made available under an Open
Content Licence (especially Creative Commons Licences), you
may exercise further usage rights as specified in the indicated
licence.*

Rankings und Indikatoren – hilfreiche Kennzahlen oder Verkürzungen mit fatalen Folgen?

Die Verdichtung von Informationen zu einer Kennzahl scheint praktisch: Sie erleichtert den globalen Vergleich verschiedener Länder. Aus der Positionierung innerhalb eines Rankings werden wirtschaftspolitische Empfehlungen abgeleitet. Der Politik wird signalisiert, genau an den Einzelkennziffern zu drehen, die im jeweiligen Ranking enthalten sind. Welche Vor- und Nachteile mit Indikatoren und Indikatorensystemen verbunden sind, diskutierten die Autoren im Rahmen der Panelsitzung, die der Wirtschaftsdienst anlässlich der Jahrestagung des Vereins für Socialpolitik 2011 organisiert hat.

Ulrich van Suntum

Keine bloßen Scheindebatten führen

Aus mehreren Einzelgrößen zusammengesetzte Gesamtindikatoren haben eine lange Tradition. In der Konjunkturbeobachtung und -prognose werden sogenannte Composite Indicators schon immer verwendet, angefangen von dem berühmten Harvard-Barometer in den 1930er Jahren bis hin zum aktuellen FAZ-Konjunkturindikator, der am Institut für Weltwirtschaft in Kiel berechnet wird. Ihr Zweck ist es hauptsächlich, die Vielzahl von Einzelinformationen zu bündeln, die täglich in Form von Daten über Auftragseingänge, von Börsenkursen, Stimmungsumfragen usw. auf die Öffentlichkeit einwirken. Viele dieser Teilgrößen sind selbst wiederum aus zahlreichen Einzeldaten zusammengesetzt. So bündelt etwa der ifo-Geschäftsklimaindex die Antworten von ca. 7000 befragten Unternehmen, und der ZEW-Konjunkturindikator fasst die Erwartungen von ca. 300 Finanzmarktexperten zusammen.

Von verschiedenen Organisationen werden regelmäßig auch Gesamtindikatoren publiziert, welche die Wettbewerbsfähigkeit oder die Standortqualität von Ländern messen sollen. Dazu gehören neben den bekannten Wettbewerbsindikatoren des World Economic Forum (WEF, Genf) und des Institute for Management Development (IMD, Lausanne) z.B. das Internationale Standortranking der Bertelsmann Stiftung (bis 2009) und der Ease of Doing Business Indicator der Weltbank. Auch in diese Länderrankings gehen zahlreiche Daten ein, die selbst wiederum hochaggregierte Indikatoren sind. Dazu gehören Korruptionsindizes, Regulierungsindizes und Umfragedaten zum Investitionsklima und zur Bürokratie

in den einzelnen Ländern. Internationale Organisationen wie die OECD, die Weltbank, der IMF und die EU publizieren regelmäßig zusammengesetzte Indikatoren zu verschiedenen Sachverhalten und bilden daraus vielfältige Ranglisten für Länder und Regionen.

Im Zuge der „Beyond GDP“-Diskussion sind zudem in den letzten Jahren zahlreiche zusammengesetzte Indikatoren zur Messung von Wohlstand und Nachhaltigkeit entwickelt worden. Sie reichen von relativ einfachen Maßen wie dem Human Development Index der Vereinten Nationen mit lediglich drei Einzelkomponenten bis hin zum Nachhaltigkeitsindikator der Kreditanstalt für Wiederaufbau, der nicht weniger als 37 Einzelkomponenten enthält.¹ Auch diesen Maßen ist gemeinsam, dass sie komplexe Sachverhalte durch die gewichtete Bündelung von dafür relevanten Einzeldaten abzubilden versuchen. Damit wollen sie nicht nur Politik und Öffentlichkeit Orientierungspunkte geben, sondern auch die empirische Untersuchung von komplexen Sachverhalten ermöglichen, für die es andernfalls keine quantitative Basis gäbe.

Komplexitätsreduktion ist unvermeidlich

Die reale Welt ist vielschichtig, erfordert aber dennoch oft eindeutige Entscheidungen. Das ist nicht nur in der

¹ Vgl. für eine Übersicht U. van Suntum, O. Lerbs: Theoretische Fundierung und Bewertung alternativer Methoden der Wohlfahrtsmessung, Studie im Auftrag der KfW Bankengruppe, Münster 2011.

Wirtschaftspolitik so, sondern auch im Alltag. Ob es der Kauf eines Autos ist, die Auswahl einer Universität für die eigene Ausbildung oder umgekehrt die Auswahl eines Bewerbers für das Masterstudium: Stets muss eine Vielzahl von Qualitätsfaktoren letztlich in eine eindeutige Rangfolge gebracht werden. Denn die Mittel sind begrenzt und viele Handlungsalternativen schließen sich gegenseitig aus. Die Frage kann also gar nicht sein, ob man komplexe Sachverhalte in eindeutige Rangordnungen bringen darf, denn das ist unausweichlich. Wer dies grundsätzlich in Frage stellt, führt deshalb eine bloße Scheindebatte. Worum es allein gehen kann, ist die Qualität entsprechender Rankings und Gesamtindikatoren.

Denn sie zu bilden, ist keine triviale Aufgabe. Nicht jeder Entscheidungsträger hat die Zeit und die Kompetenz, dies von Grund auf selbst zu leisten. Darum gibt es Autotests, Universitätsrankings und Zeugnisnoten. Sie leisten Entscheidungshilfe und reduzieren Komplexität, ohne damit die eigentliche Entscheidung vorwegzunehmen. Dies ist auch die Funktion von Länderrankings und hochaggregierten Kennzahlen, etwa zur Standortqualität oder zur Wettbewerbsfähigkeit eines Landes. Sie mögen im Einzelfall von unterschiedlicher Qualität und Aussagekraft sein – insoweit ist eine kritische Debatte natürlich angebracht. Wer solche Indikatoren aber in Bausch und Bogen als „unwissenschaftlich“, „pauschal“ oder „manipulativ“ abtut, setzt sich damit selbst dem Vorwurf aus, den er gegen die Rankings erhebt.² Dass solche Einwände oft von denen kommen, die im Vergleich schlecht abgeschnitten haben, liegt in der Natur der Sache.

Auch das BIP ist letztlich nichts anderes als ein hochaggregierter Gesamtindikator. Es soll, wenn schon nicht den Wohlstand, so doch die Wirtschaftsleistung eines Landes messen. Auch hier werden buchstäblich nicht nur Äpfel und Birnen, sondern unzählige weitere Güter zu einer einzigen Größe addiert. Und die Produktion jedes einzelnen Sektors besteht wiederum aus der Aggregation zahlloser Einzelgüter, die alle von ganz unterschiedlicher Qualität und Bedeutung sind. Das Gewichtungproblem, oft grundsätzlich gegen Gesamtindikatoren angeführt, besteht hier ganz genauso. Es wird bei der Berechnung des BIP bekanntlich dadurch gelöst, dass Marktpreise bzw. – im Fall öffentlicher Güter – Herstellungskosten verwendet werden: Die Gütermengen, jeweils bewertet mit ihren Marktpreisen bzw. Kosten, ergeben in der Summe die Bruttowertschöpfung, aus der das BIP berechnet wird.

2 Vgl. dazu im Einzelnen auch S. Gundel, U. van Suntum: Ist die Kritik an internationalen Standortrankings berechtigt?, in: Wirtschaftsdienst, 87. Jg. (2007), H. 7, S. 473-479.

Preise und Kosten sind weder im Zeitverlauf konstant noch als Gewichte unumstritten, da sie vielfältig verzerrt sein können und zudem verteilungsabhängig sind. Außerdem fehlen viele Leistungen im BIP, etwa die Arbeit im Haushalt, und andere, etwa Verkehrsunfälle, gehen (angeblich) sogar mit dem falschen Vorzeichen ein.³

Trotzdem denkt wohl niemand daran, auf die Berechnung des BIP zu verzichten. Ebenso wenig ist es sinnvoll, etwa über die Wettbewerbsfähigkeit Griechenlands zu diskutieren, ohne diese messen oder überhaupt definieren zu können. Eindimensionale Kenngrößen wie der Leistungsbilanzsaldo reichen dazu sicher nicht aus. Natürlich bestehen über die richtige Definition und Messung unterschiedliche Meinungen, das ist nun einmal Kennzeichen einer wissenschaftlichen Debatte. Aus den unvermeidlichen methodischen Problemen auf die grundsätzliche Unzulässigkeit der Bildung umfassender Indikatoren zu schließen, ist jedoch wenig überzeugend. Um beim Beispiel der Standortqualität eines Landes zu bleiben: Entweder man sagt, was darunter zu verstehen und wie sie zu messen ist, oder man streicht den Begriff aus dem wissenschaftlichen Wortschatz. In letzterem Fall würden allerdings wohl nicht viele ökonomische Begriffe übrig bleiben. Denn anders als in der Mathematik und der Physik ist die Ökonomie eine Gesellschaftswissenschaft. Es bleibt also nur übrig zu versuchen, die methodischen Probleme der Begriffsbildung bestmöglich zu lösen und stets offenzulegen, wie man im Einzelnen dabei vorgeht.

Es geht um das „Wie“ und nicht um das „Ob“

Die bloße Beschränkung auf Einzelindikatoren ist dazu keine Alternative, jedenfalls keine echte. Denn auch diese Einzelindikatoren, wie sie etwa von der Stiglitz-Sen-Fitoussi-Kommission und vom Sachverständigenrat für die Wohlstandsmessung präferiert werden,⁴ sind überwiegend selbst wiederum hochaggregiert. Das gilt für die Feinstaubbelastung der Bevölkerung ebenso wie für die Rohstoffproduktivität und natürlich auch für das Nettoneationaleinkommen sowie alle anderen aus der Volkswirtschaftlichen Gesamtrechnung abgeleiteten Größen. Es geht also gar nicht um ein Ja oder Nein zur Bildung

3 Bei genauerer Analyse erweisen sich viele dieser Einwände allerdings als nicht stichhaltig und das BIP als wesentlich aussagekräftiger, als es gemeinhin angenommen wird, vgl. U. van Suntum, O. Lerbs, a.a.O.

4 Vgl. J. E. Stiglitz, A. Sen, J. P. Fitoussi: Report by the Commission on the Measurement of Economic Performance and Social Progress, 2009; Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung: Wirtschaftsleistung, Lebensqualität und Nachhaltigkeit: Ein umfassendes Indikatorensystem, Expertise im Auftrag des Deutsch-Französischen Ministerrates in Zusammenarbeit mit dem Conseil d'Analyse Economique, 2010.

von aggregierten Indikatoren, sondern nur um das Wie und um die Frage, bis zu welcher Stufe man dies treiben sollte.

Dass in der Messung unscharfer Begriffe auch viele Gefahren liegen, hat schon Simon Kuznets, der „Vater“ des BIP, zu Recht betont. Für die Beurteilung von entsprechenden Kennzahlen haben sich deswegen bestimmte Qualitätsanforderungen in der Indikatorendebatte durchgesetzt:⁵

- **Repräsentativität:** Hier geht es darum, dass ein Indikator möglichst genau das erfasst, was er erfassen soll. Dazu bedarf es zunächst einer klaren Definition des zu messenden Tatbestandes. Auf dieser Basis ist dann ein theoretisch konsistenter Indikator zu entwickeln, der einerseits umfassend genug, andererseits aber nicht redundant und schon gar nicht in sich widersprüchlich ist. Idealerweise können ökonomische Berechnungen dazu vorgelegt werden, wie eng der Indikator mit verwandten Referenzgrößen korreliert und welchen Erklärungsbeitrag die jeweiligen Einzelkomponenten leisten.⁶
- **Objektivität:** Werturteile lassen sich bei der Indikatorrekonstruktion nicht vermeiden, wie schon am Beispiel der Gewichtung des BIP gezeigt wurde. Jedoch sollten solche Werturteile und ihre Bedeutung, etwa für ein Länderranking, offengelegt werden. Alle verwendeten Einzeldaten sowie ihre Aggregation sollten gut dokumentiert und intersubjektiv nachvollziehbar sein. Idealerweise besteht für den externen Nutzer die Möglichkeit, die Sensitivität der Rangfolgen auf andere Gewichtungen oder Zusammensetzungen des Indikators selbst zu prüfen.⁷
- **Verlässlichkeit:** Hier geht es vor allem um die Robustheit der Ergebnisse und die Verfügbarkeit und Qualität der verwendeten Einzeldaten. Auch die Volatilität des Indikators kann eine wichtige Rolle spielen. So ist von einem langfristig orientierten Indikator der Standortqualität zu erwarten, dass er sich nicht sprunghaft von Jahr zu Jahr ändert.

5 Vgl. dazu A. Diekmann: Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen, Hamburg 2007; J. Bortz, N. Döring: Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler, Berlin 2006.

6 Dies ist z.B. für konjunkturelle Frühindikatoren möglich, die das BIP prognostizieren sollen, aber auch für Indikatoren, die z.B. die Wachstumsstärke einer Volkswirtschaft anzeigen sollen.

7 Diese Möglichkeit besteht z.B. bei den von der Bertelsmann Stiftung seit 2009 erhobenen Sustainable Governance Indicators, vgl. http://www.bertelsmann-stiftung.de/cps/rde/xchg/bst/hs.xsl/publikationen_92945.htm.

Unter pragmatischen Gesichtspunkten sind noch zwei weitere Kriterien hinzuzufügen, die vor allem wirtschaftspolitisch motiviert sind. So sollte ein Gesamtindikator zum einen übersichtlich und verständlich genug sein, um politisch überhaupt wahrgenommen und richtig interpretiert zu werden. Zum anderen sollte er unter vertretbarem Aufwand an veränderte Rahmenbedingungen und Datenlagen anpassbar sein, da „Eintagsfliegen“ der Politik keine langfristige Orientierung geben können.

Basis für die empirische Wirtschaftsforschung

Länderrankings erfüllen, wenn sie diesen Kriterien entsprechend gebildet werden, wichtige Aufgaben in der Politik und Politikberatung. Sie halten den Adressaten gewissermaßen den Spiegel vor und zwingen sie so, sich mit den Ergebnissen kritisch auseinanderzusetzen. Dies wird aber nur gelingen, wenn Öffentlichkeit hergestellt werden kann. Die Nachhaltigkeitsindikatoren des Sachverständigenrates sind eine gute Grundlage für weitere wissenschaftliche Arbeiten, aber eine breitere Öffentlichkeit wurde damit bisher nicht erreicht. Deshalb drohen sie politisch ebenso ignoriert zu werden wie vieles andere, was Wirtschaftswissenschaftler schreiben, aber nicht öffentlichkeitswirksam zu publizieren verstehen. Man mag dies nicht so schlimm finden. Damit aber überlassen die Ökonomen das Feld erst recht fragwürdigen Kennziffern wie etwa dem „ökologischen Fußabdruck“, der zwar äußerst simpel konstruiert, aber eben publikumswirksam vermarktet wurde.

Selbst methodisch angreifbare Indikatoren – und welcher wäre das nicht? – sind immer noch besser als der völlige Verzicht auf jeden Vergleich. Denn damit hätten es die politischen Adressaten zu leicht, sich unabhängig von ihren tatsächlichen Leistungen stets in günstigem Licht darzustellen. Das Ranking von Städten, Gesundheitssystemen, Universitäten oder anderen Organisationen erfüllt eine wichtige Funktion im Wettbewerb der Institutionen. Bei aller berechtigten Kritik etwa an der Bewertung von wirtschaftswissenschaftlichen Fakultäten und der – meist sehr einseitig gemessenen – Forschungsleistung ihrer Professoren: Wer wollte bestreiten, dass entsprechende Rankings einen Stachel in das Fleisch der deutschen Ökonomie getrieben und ihr Streben nach nationaler und internationaler Anerkennung wesentlich verstärkt haben? Wer ordnungspolitisch denkt und den Wettbewerb der Institutionen als ein wesentliches Element von Effizienz und Freiheit betrachtet, wird letztlich nicht gegen die Erstellung von entsprechenden Ranglisten argumentieren können.

Vorausgesetzt, die vorgenannten Qualitätsmerkmale werden erfüllt, sind hochaggregierte Gesamtindikator-

ren aber nicht nur eine Entscheidungshilfe für die Politik, sondern auch eine wichtige Basis für die empirische Untersuchung komplexer Sachverhalte. Wie sollte etwa der Einfluss von Korruption auf das Wirtschaftswachstum gemessen werden, wenn es gar keine quantitative Messgröße für Korruption gibt? Wie will man qualitative Institutionen wie Bürokratie, Freiheit, Rechtssicherheit und wirtschaftliche Freiheit einer empirischen Untersuchung zugänglich machen, wenn schon der Versuch ihrer Quantifizierung als „unwissenschaftlich“ abgelehnt wird? Es handelt sich hier nun einmal um komplexe Sachverhalte, die anders als etwa die Arbeitslosigkeit nicht in einer einzigen, offensichtlichen Kennziffer abgebildet werden können. Und selbst die Arbeitslosenquote ist, bei Lichte betrachtet, ein komplexer Gesamtindika-

tor, der Frauen und Männer, junge und ältere Arbeitslose, Langzeitarbeitslosigkeit und Kurzarbeitslosigkeit sowie auch alle regionalen Unterschiede zu einer einzigen Kennzahl zusammenfasst. Praktisch alle Einwände der Kritiker gegen Gesamtindikatoren lassen sich darum selbst gegen die gebräuchlichsten Kennzahlen der Ökonomie erheben.

Insoweit sind die – übrigens vorwiegend in Deutschland erhobenen – Vorbehalte gegenüber komplexen Indikatoren kaum zu verstehen. Wenn es richtig ist, dass „institutions matter“, wenn es weiterhin richtig ist, dass diese Institutionen komplexe Phänomene sind, dann führt an zusammengesetzten Indikatoren für eine gehaltvolle empirische Forschung kein Weg vorbei.

Anja Rohwer, Oliver Hülsewig

Länderrankings: Zu welchen Ergebnissen kommt die neuere Forschung?

„If we don't have good metrics, we don't have a good sense of how well we are doing.“ (Joseph E. Stiglitz)¹

Für die Regierung eines Landes stellen Länderrankings eine relevante Informationsquelle dar. Länderrankings ermöglichen es idealerweise, die strukturelle Lage der Wirtschaft international zu vergleichen. Sie weisen die Politik darauf hin, welcher Handlungsbedarf bei bestimmten wirtschaftspolitischen Feldern besteht. Da sich der Wettbewerb durch die Globalisierung der Weltwirtschaft zunehmend intensiviert, sind derartige Informationen besonders bedeutend.

Der Vorteil von Rankings kann darin gesehen werden, dass komplexe Zusammenhänge auf eine Zahl komprimiert werden. Diese ist der Öffentlichkeit leicht zu kommunizieren. Die Einfachheit von Rankings impliziert aber gleichzeitig den Nachteil, dass diese nur grundlegende Hinweise liefern können. So misst der „Corruption Perceptions Index“ von Transparency International beispielsweise lediglich den Grad der bei Beamten und Politikern wahrgenommenen Korruption. Dadurch ist es zwar möglich, ein Land mit einem anderen Land im Bereich der herrschenden Korruption zu vergleichen. Um aber detail-

liertere Informationen über die Korruption in einem Land zu erhalten, reicht der Index jedoch nicht aus, da dieser keine Auskunft darüber gibt, um welche Art von Korruption es sich handelt. Ein tieferer Einblick in die Details wäre wünschenswert. „Vielfach gelingt ökonomische Erkenntnis erst nach dem Studium rechtlicher Wirkungsmechanismen, die sich zwar durch Worte verständlich beschreiben lassen, sich aber einer rein numerischen Analyse entziehen.“² Zusätzlich ist zu erwähnen, dass Rankings nur Rangpositionen von Ländern angeben. Ein Beispiel ist der Index zur Messung der internationalen Wettbewerbsfähigkeit des World Economic Forum. Eine Vielzahl von Informationen (117 Indikatoren) werden hier zu einem Index gebündelt.³ In den Medien werden dann häufig nur die einzelnen Rangpositionen der Länder kommuniziert. Weiterreichende Informationen für den Bewertungsgrund der jeweiligen Wettbewerbsposition eines Landes sind jedoch nur durch eigene Recherchearbeit zu erhalten, beispielsweise dadurch, dass die im World Economic Forum publizierten Unterkategorien betrachtet werden.

Die Frage, ob der Entwicklungsstand von Ländern anhand von Rankings bzw. anderen Indikatoren, wie beispielswei-

1 12. Oktober 2011 in Paris auf der „Two years after the Stiglitz-Sen-Fitoussi report: What well-being and sustainability measures?“ Konferenz.

2 H.-W. Sinn: Der richtige Dreiklang der VWL, in: Frankfurter Allgemeine Zeitung, Nr. 141, 22.6.2009.

3 Vgl. World Economic Forum: The Global Competitiveness Report 2011-2012, Genf 2011.

se dem Bruttoinlandsprodukt (BIP), angemessen erfasst werden kann, ist in jüngster Zeit auf eine breite Diskussionsbasis gestellt worden. Mit der Arbeit der Stiglitz-Sen-Fitoussi-Kommission (im Folgenden Stiglitz-Report genannt) im Jahr 2009 wurde eine Debatte darüber ausgelöst, wie die Glaubwürdigkeit der amtlichen Statistik erhöht werden kann und wie die Informationsbasis für wichtige Politikfelder zu ergänzen ist, um einen besseren Rahmen für Handlungsentscheidungen seitens der Politik zu ermöglichen. Diese Debatte richtete sich auf die Fragestellung, wie das Ziel eines nachhaltigen Wirtschaftens zu erreichen ist. Dafür untersuchte der Stiglitz-Report die Sinnhaftigkeit verschiedener Indikatoren zur Messung der Wirtschaftsleistung und des gesellschaftlichen Fortschritts.⁴ Anknüpfend daran erfolgte 2010 eine weiterführende Analyse durch den deutschen Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung in Zusammenarbeit mit dem französischen Conseil d'Analyse Économique (CAE).⁵ Auch die Europäische Kommission hat sich dieser Thematik mit ihrer Strategie Europa 2020 angenommen. Diese Strategie verfolgt ein intelligentes, nachhaltiges und integratives Wachstum.⁶ Anhand von fünf Kernzielen, die durch acht Indikatoren gemessen werden, sollen die Ziele der Strategie Europa 2020 erreicht werden.⁷ Galt bis vor einiger Zeit das BIP als „der“ Indikator für die Messung des Wohlstands einer Nation, wird heutzutage darüber diskutiert, ob das BIP wirklich ausreichend ist. Basierend auf dem Stiglitz-Report und seinen akademischen Vorläufern⁸ wird mit der Strategie Europa 2020 deutlich, dass dies nicht der Fall ist.⁹ Und auch hinsichtlich des Gebrauchs von Rankings wird in der einschlägigen Literatur eine ähnlich kritische Haltung eingenommen.

4 Vgl. J. E. Stiglitz, A. Sen, J. Fitoussi: Report by the Commission on the Measurement of Economic Performance and Social Progress, Paris 2009.

5 Vgl. Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung, Conseil d'Analyse Économique: Wirtschaftsleistung, Lebensqualität und Nachhaltigkeit: Ein umfassendes Indikatorensystem, Expertise im Auftrag des Deutsch-Französischen Ministerrats, Paris, Wiesbaden 2010.

6 Vgl. Europäische Kommission: Europa 2020 – Eine Strategie für intelligentes, nachhaltiges und integratives Wachstum, Brüssel 2010.

7 Eine genaue Beschreibung der acht Indikatoren findet sich bei Eurostat: http://epp.eurostat.ec.europa.eu/portal/page/portal/europe_2020_indicators/headline_indicators, letzter Zugriff: 30.9.2011.

8 Vgl. CESifo DICE Report 4/2010 (Winter).

9 In diesem Zusammenhang scheint es erwähnenswert, dass Bhutan als erstes Land das BIP durch ein „Bruttoglücksprodukt“ ersetzt hat (<http://www.ophi.org.uk/ophi-measure-used-in-creating-the-bhutan-gross-national-happiness-index/>). Immer mehr Regierungen ziehen nach. So hat z.B. der Premierminister von Großbritannien, David Cameron, das Nationale Amt für Statistik angewiesen, einen monatlichen Glücksindikator zu veröffentlichen. Und auch in Deutschland findet diese Diskussion immer mehr Anklang. So nahm Anfang 2011 die Enquetekommission „Wachstum, Wohlstand, Lebensqualität“ im Auftrag des Deutschen Bundestages ihre Arbeit auf (<http://www.bundestag.de/bundestag/ausschuesse17/gremien/enquete/wachstum/index.jsp>).

Da es in der Wissenschaft bislang keinen Konsens darüber gibt, welche Mindeststandards ein Ranking erfüllen muss, variieren die Methoden zur Berechnung von Rankings wie auch deren Ergebnisse erheblich. Im Nachfolgenden soll deshalb zuerst die Frage geklärt werden, was genau Rankings messen. Daran anknüpfend erfolgt eine Analyse, ob Rankings eine geeignete Grundlage für wirtschaftspolitische Handlungsempfehlungen sind. Zum Abschluss werden mögliche Alternativen vorgestellt.

Vier Schritte zur Rankingberechnung

Im Allgemeinen basiert ein Ranking auf mehreren Indikatoren, die auf eine bestimmte Art und Weise gewichtet zu einem Gesamtindex aggregiert werden. Vier Schritte sind zur Rankingberechnung notwendig:¹⁰

- Der erste Schritt der Rankingberechnung beruht auf der Festlegung einer Definition der Zielgröße. Diese sollte theoretisch fundiert sein. Probleme hierbei sind u.a., dass die Definition bzw. das Verständnis einer Zielgröße sehr facettenreich ausfallen kann. So existiert für den Begriff der internationalen Wettbewerbsfähigkeit beispielsweise keine allgemein akzeptierte Definition. Der Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung unterscheidet drei Konzepte der internationalen Wettbewerbsfähigkeit.¹¹ Die Auswirkungen der Entscheidung für ein Konzept auf das Endergebnis können bemerkenswert sein.¹²
 - Anhand der gewählten Definition werden im zweiten Schritt die relevanten Variablen identifiziert. Die Aussagekraft eines Rankings ist stark davon abhängig, welche Variablen einbezogen werden.¹³ Die Auswahl der Variablen sollte sich von dem theoretisch begründeten und empirisch belegten Zusammenhang zwischen Variable und Zielgröße und von der Messbar-
- 10 Eine gute Beschreibung der einzelnen Schritte zur Berechnung eines Rankings finden sich bei: OECD, Europäische Kommission, Joint Research Centre: Handbook on Constructing Composite Indicators: Methodology and User Guide, von M. Nardo, M. Saisana, A. Saltelli, S. Tarantola (EC/JRC), A. Hoffman, E. Giovannini (OECD), Paris 2008.
- 11 Vgl. Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung: Jahreshgutachten 2004/2005: Erfolge im Ausland – Herausforderungen im Inland, Stuttgart 2004.
- 12 Das Bundesministerium der Finanzen hat 2006 dem Institut für Wirtschaftsforschung Halle zusammen mit dem Institut für empirische Wirtschaftsforschung der Universität Leipzig einen Auftrag erteilt, drei international bedeutende Ranking-Studien zum Thema „Internationale Wettbewerbsfähigkeit“ zu vergleichen. Im Ergebnis kamen die drei Rankings zu drei unterschiedlichen Resultaten.
- 13 Vgl. W. Ochel, O. Röhn: Indikatorenbasierte Länderrankings, in: Perspektiven der Wirtschaftspolitik, 9 (2), 2008, S. 226-251.

keit der Variablen leiten lassen.¹⁴ Die Variablen sollten für die Fragestellung relevant sein und einer Korrelationsprüfung unterzogen werden.

- Da die ausgewählten Variablen in unterschiedlichen Dimensionen vorliegen, müssen sie im dritten Schritt normiert werden. Ziel der Normierung ist es, eine gemeinsame Basis zu finden, d.h. die Variablen dimensionslos zu machen. Es gibt hier unterschiedliche Herangehensweisen,¹⁵ die alle Vor- und Nachteile aufweisen.
- Der entscheidende Punkt bei der Rankingberechnung ist im vierten Schritt die Art und Weise, wie die ausgewählten Variablen gewichtet zu einem Gesamtindex aggregiert werden. Die ausgewählten Gewichte reflektieren die relative Bedeutung der einzelnen Variablen zum Gesamtindex. Von daher sollten Gewichte auf einer theoretischen und empirischen Grundlage beruhen. Da es an einer Theorie mangelt, wird oft die Gleichgewichtung bevorzugt. Die Gleichgewichtung impliziert allerdings, dass alle Variablen dieselbe relative Bedeutung haben. Dies heißt wiederum, dass alle Variablen als perfekte Substitute angesehen werden können.¹⁶ Ein Rückgang bei einem Variablenwert um einen bestimmten Betrag kann also vollständig durch eine ebenso starke Erhöhung bei einer anderen Variablen kompensiert werden. Um eine Gleichgewichtung oder eine andere willkürliche Gewichtung zu umgehen, sollte man daher auf statistische Verfahren (wie z.B. Regressionsanalysen, prinzipielle Komponentenanalyse oder Faktoranalyse) zurückgreifen. Der Vorteil dieser Verfahren ist, dass sie die den Daten innewohnenden Informationen nutzen. Das bedeutet, dass die Gewichte endogen bestimmt werden.¹⁷

Ein Verbesserungsvorschlag – der ifo Institutionenindex

Aufgrund der Kritik an der Berechnung von Rankings hat das ifo Institut 2007 mit der Entwicklung des ifo Institutionenindex für OECD-Länder versucht, einige Verbesserungsvorschläge umzusetzen. Der ifo Institutionenindex von Eicher und Röhn (2007)¹⁸ verfolgt das Ziel, Verände-

rungen des Pro-Kopf-Einkommens anhand von institutionellen Regelungen für 24 OECD-Länder zu erklären.¹⁹

Die Auswahl der einbezogenen Variablen beruht auf einer Kombination aus faktoranalytischen Verfahren und Regressionsanalysen. Im ersten Schritt werden zunächst institutionelle Variablen mit Hilfe der Faktoranalyse zu Subindizes verdichtet. Diese Subindizes werden im zweiten Schritt auf das Wirtschaftswachstum regressiert. Für die Bestimmung der Gewichtung jeder einzelnen institutionellen Komponente werden die Koeffizienten der Regressionsanalyse zusammen mit den Gewichten aus der Faktoranalyse verwendet.

Das Besondere an dieser Vorgehensweise ist, dass sowohl die Ableitung der relevanten Variablen als auch die Bestimmung der Gewichte ökonomisch erfolgen. Die Gefahr, dass die Auswahl der Variablen und Gewichte lediglich auf Plausibilitätsüberlegungen beruht, wird somit umgangen. Des Weiteren wird mit den 24 OECD-Ländern eine recht homogene Ländergruppe betrachtet, so dass die Vermutung nahe liegt, dass die einbezogenen Variablen in gleicher Weise auf die Zielgröße einwirken. Die Qualität des Index ließe sich aber weiter verbessern, wenn die Mängel in der Erfassung von Institutionen beseitigt würden.²⁰

Sind Länderrankings eine geeignete Grundlage für wirtschaftspolitische Handlungsempfehlungen?

Für wirtschaftspolitische Handlungsempfehlungen sollten Rankings nicht herangezogen werden. Dies hat mehrere Gründe. So wird unterstellt, dass Rankings eine genaue Vorstellung darüber geben, welche Faktoren die Zielgröße einer Volkswirtschaft bestimmen.²¹ Häufig beruht die Auswahl der Faktoren auf keiner systematischen Ableitung. Eine Ausnahme stellt der ifo Institutionenindex dar, da hier die Faktoren anhand von statistischen Methoden ausgewählt werden.

14 Vgl. M. Freudenberg: Composite Indicators of Country Performance: A Critical Assessment, in: OECD Science, Technology and Industry Working Papers, 2003/12, Paris 2003.

15 Im Handbook on Constructing Composite Indicators werden die jeweiligen Vor- und Nachteile der einzelnen Herangehensweisen zur Normierung aufgelistet.

16 Vgl. W. Ochel, O. Röhn, a.a.O.

17 Ebenda.

18 T. Eicher, O. Röhn: Institutional Determinants of Economic Performance in OECD Countries – An Institutions Climate Index, in: CESifo DICE Report 5 (1), München 2007, S. 38-49.

19 Eine Dokumentation der Datenbasis und der Ergebnisse des Institutionenindex findet sich in der DICE-Datenbank des ifo Instituts, http://www.cesifo-group.de/portal/page/portal/ifoHome/a-wininfo/d3iiv/_DICE_division?_id=6746666&_div=7209869, letzter Zugriff: 7.10.11. Eine Darstellung der Ergebnisse für 2011 findet sich bei A. Rohwer: Institutions Matter – Results of the Institutions Climate Index for OECD Countries, in: CESifo DICE Report 3/2011 (Herbst).

20 Qualitative Informationen über Regulierungen müssen häufig in quantitative Informationen überführt werden. Vgl. W. Ochel: Lernen von anderen Ländern: Zum internationalen Vergleich und Transfers von Arbeitsmarktinstitutionen, in: Perspektiven der Wirtschaftspolitik, 5 (2), 2004, S. 139-158.

21 Länder, die diesen Vorstellungen nicht entsprechen, werden im Ranking somit automatisch schlechter bewertet. Vgl. U. Heilemann, H. Lehmann, J. Ragnitz: Länder-Rankings und internationale Wettbewerbsfähigkeit: Eine kritische Analyse, in: Schriften des Instituts für Wirtschaftsforschung Halle, Bd. 24, Baden-Baden 2006.

Es stellt sich außerdem die Frage, ob die einbezogenen Variablen in allen Ländern in gleicher Weise auf die Zielgröße einwirken. Dies ist aber wahrscheinlich nicht der Fall, vor allem nicht bei der Betrachtung von heterogenen Ländern. Dieser Vorstellung unterliegen aber die meisten internationalen Rankings.

Außerdem vernachlässigen Rankings oftmals die Besonderheiten einzelner Länder. „Unterschiedliche institutionelle Rahmenbedingungen, das erreichte Niveau der wirtschaftlichen Entwicklung, wirtschaftsstrukturelle Besonderheiten, die Größe der Länder, die Ausstattung mit Ressourcen, Klima, geografische Gegebenheiten, kulturelle Faktoren, singuläre Ereignisse“²² (Beispiel Wiedervereinigung Deutschland) sind nur einige Größen, denen bei der Rankingberechnung Beachtung geschenkt werden muss, um einen wirklich vergleichenden Maßstab zu erzeugen. Hieran anschließend bleibt meistens auch die Frage unbeantwortet, inwieweit eine bestimmte Rangposition eines Landes konjunkturellen, strukturellen oder anderen Umständen geschuldet ist.²³

Auch ist ein Lernen aus der Vergangenheit mit Hilfe von Rankings nicht immer möglich, da es oftmals zu Änderungen bei der Berechnungsmethode sowie bei der Anzahl der einbezogenen Länder kommt. Nicht immer erfolgt dann eine Rückrechnung, die es erlaubt, Vergleiche über einen längeren Zeitraum hinweg anzustellen.²⁴

Allgemein kann gesagt werden, dass Rankings einen Handlungsbedarf signalisieren können, auch wenn die Methodik zu ihrer Berechnung Probleme aufwirft. Aber konkrete Handlungsempfehlungen sind nicht unbedingt ableitbar. Deshalb ist es sinnvoll, Rankings nur als Ausgangspunkt für weitere Analysen zu verwenden.²⁵

Alternativen – Benchmarking oder Dashboard?

Neben dem Versuch, bei der Berechnung von Rankings mehr auf ökonomische Verfahren hinsichtlich der Auswahl der Variablen und der Bestimmung der Gewichte zurückzugreifen, spielen in der heutigen Diskussion Verfahren wie Benchmarking oder Dashboards (Indikatorensystem) eine immer größer werdende Rolle. Länderrankings, wie wir sie kennen, rücken immer mehr ins Abseits. In der Wissenschaft herrscht mittlerweile ein Konsens darüber,

dass die Erfassung einer bestimmten Fragestellung in nur einem einzigen Indikator ihrer Komplexität nicht gerecht werden kann.²⁶

Das Benchmarking stellt einen Qualitätsvergleich dar. Dabei wird ein Land hinsichtlich einer bestimmten Fragestellung mit einem anderen Land, das als „best practice“ Land bei dem zu untersuchenden Thema gilt, verglichen. Genauer gesagt umfasst das Benchmarking die Analyse von Ursachen und inneren Zusammenhängen beobachteter Differenzen.²⁷ Mit dem Benchmarking wird versucht, Handlungsempfehlungen abzuleiten, indem relevante Maßnahmen des „best practice“ Landes identifiziert und evaluiert werden.²⁸ Für eine erfolgreiche Übertragung muss jedoch eine genaue Prüfung darüber erfolgen, welche Bedingungen erfüllt sein müssen, um wirksame wirtschaftspolitische Maßnahmen des „best practice“ Landes im eigenen Land zu implementieren.²⁹

Das Dashboard stützt sich hingegen auf eine Vielzahl verschiedener Einzelindikatoren. Aufgrund der Komplexität einiger Fragestellungen wird ein System relevanter Indikatoren eingeführt. Dies sollte zwar umfassend aber nicht zu detailliert sein.³⁰ Aufgrund der sich so ergebenden Übersichtlichkeit kann das Dashboard Informationen der Öffentlichkeit und auch der Politik auf einfachste Art überbringen. Ein erfolgreiches Beispiel für ein Dashboard stellt die „Better Life“³¹ Initiative im Zusammenhang mit der Studie „How’s Life“³² der OECD dar. Die OECD bietet elf Themenfelder mit unterschiedlichen Unterkategorien an, die die Frage beantworten sollen, wie die Lebensqualität sich in 40 verschiedenen Ländern darstellt. Trotz aller Vorteile des Dashboard birgt auch dieses Probleme in sich. Obwohl hier nicht nur ein Gesamtindikator betrachtet wird, besteht dennoch die Gefahr, dass nicht alle relevanten Informationen von dem Indikatorensystem erfasst werden. Zudem kann die Auswahl der Indikatoren wie auch deren Bedeutung auf subjektiven Empfindungen beruhen.

22 Vgl. W. Ochel, O. Röhn, a.a.O.

23 Vgl. U. Heilemann, H. Lehmann, J. Ragnitz, a.a.O.

24 Ein Beispiel wäre der „Corruption Perceptions Index“ von Transparency International, http://www.transparency.org/policy_research/surveys_indices/cpi, letzter Zugriff: 07.10.11. Die Berechnung erfolgt seit 2001. Ein Vergleich über die Zeit hinweg ist aufgrund von Änderungen verschiedenster Art bei der Erstellung des Index allerdings nicht möglich.

25 Vgl. U. Heilemann, H. Lehmann, J. Ragnitz, a.a.O.

26 Vgl. Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung, Conseil d’Analyse Économique: Wirtschaftsleistung, Lebensqualität und Nachhaltigkeit ..., a.a.O.

27 Vgl. U. Heilemann, H. Lehmann, J. Ragnitz, a.a.O.

28 Vgl. H. Schütz, S. Speckesser, G. Schmid: Benchmarking Labour Market Performance and Labour Market Policies: Theoretical Foundations and Applications, Wissenschaftszentrum Berlin für Sozialforschung, in: Discussion Paper, FS I 98-205, Berlin 1998.

29 Vgl. W. Ochel, a.a.O.

30 Vgl. J. E. Stiglitz, A. Sen, J. Fitoussi, a.a.O.; und Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung, Conseil d’Analyse Économique: Wirtschaftsleistung, Lebensqualität und Nachhaltigkeit ..., a.a.O.

31 OECD Better Life Initiative: <http://www.oecdbetterlifeindex.org/#/>, letzter Zugriff: 26.10.2011.

32 OECD: How’s Life, 2011, http://www.oecd.org/document/39/0,3746,en_21571361_44315115_48858599_1_1_1_1,00.html, letzter Zugriff: 26.10.2011.

Ullrich Heilemann

Die Krise 2008/2009: Daten oder Politik?

Nach der Krise wird nach Schuldigen gesucht. Wäre sie nicht zu vermeiden gewesen? Wenn wir akzeptieren, dass die Makroökonomie ihr Versprechen eines krisenfreien Wachstums nicht einlöste: Bekämpften wir die Krise wenigstens richtig? Auch dabei sieht es nicht gut aus.¹ Ein Schuldiger ist schnell bei der Hand: die gesamtwirtschaftlichen Prognosen. Sie sind nach wie vor mit beträchtlichen Fehlern behaftet und verhindern die rechtzeitige Krisenbekämpfung. Ihre Treffsicherheit ist das Produkt von „Theorie“, „Daten“ und „Methoden“, so dass es angesichts der methodischen Fortschritte der letzten Jahrzehnte nur an der Theorie und den Daten liegen kann. In der Tat erinnern sich nach der Krise 2008/2009 viele daran, dass unser Verständnis des Finanzsektors und dessen Beziehungen zur Realwirtschaft mangelhaft ist und unisono werden hier (und generell) „neue Theorien“ gefordert.² Ob sie für die Krisenbekämpfung mehr hergeben, bleibt abzuwarten – Krugmans³ vernichtendes Urteil über die Makroökonomie der letzten Dekaden stimmt wenig hoffnungsvoll.

Im Folgenden soll es jedoch um Konkretes gehen: „Haben die gängigen Indikatoren und Rankings (und Ratings) im Verlauf der Krise die richtigen Signale an die Politik gegeben?“ und „Wurden Sie von der Politik sinnvoll genutzt?“. Die erste Frage mutet paradox an: Auch wenn das „Informationszeitalter“ sicher nicht das „Zeitalter der Prognosen“ ist⁴ – das Angebot an Wirtschaftsdaten und Prognosen ist heute so groß wie niemals zuvor. Erst Recht gilt das hinsichtlich Aktualität, Verarbeitung und Dissemination der Informationen. Also sind es widersprüchliche oder fehlerbehaftete Informationen, die effizientes und effektives konjunkturpolitisches Handeln verhindern?

Krisen erkennen wir bekanntlich erst, wenn wir schon halbwegs durch sind, was auch damit zu tun hat, dass die Krisen

in Deutschland selbst bei breiter Definition kaum länger als vier Quartale dauern.⁵ Tempo und Ausmaß des Abschwungs (wie des Aufschwungs) werden dabei in der Regel deutlich unterschätzt. Die Krise 2008/2009 lieferte dazu offensichtlich ein besonders ärgerliches Beispiel (vgl. Abbildung 1).

Die Bedeutung der Prognosedefizite variiert – ein weiteres Paradoxon – mit der Tiefe der Krise: So war das späte Erkennen der Konsequenzen der Weltfinanz- und -wirtschaftskrise 2008/2009 für Deutschland angesichts ihrer Tiefe für die Effektivität der Bekämpfung daher weit weniger bedeutsam als in den vorausgegangenen, im Nachhinein „kleinen“ Krisen.

Ihr rechtzeitiges Erkennen und eine lehrbuchhafte Reaktion der Fiskalpolitik und dadurch das Vermeiden der Krise – „selbst-zerstörende“ wie selbst-erfüllende, reflexive Prognosen – bilden hier jedenfalls seltene Ausnahmen. Der in jüngerer Zeit offenbar einzige Fall einer sich völlig selbst-aufhebenden Prognose war im Zusammenhang mit dem weltweiten Börsenkrach 1987 zu beobachten. Eine partielle Selbstaufhebung der Krisen-Prognosen ist natürlich in allen modernen Wirtschaftskrisen festzustellen.⁶

Die Forderung nach mehr und besseren Daten über die Akteure und Aktivitäten, die helfen, dass sich der private Sektor auf die Krise einstellen und die Regierung wirtschaftspolitisch reagieren kann, zählt seit den ersten größeren Konjunkturkrisen im 19. Jahrhundert zu den routinemäßigen Appellen danach.⁷ Ungehört blieben die Forderungen nicht und ihr sind – nach der Weltwirtschaftskrise 1929 – u.a. der Aufbau der Volkswirtschaftlichen Gesamtrechnungen (VGR) und die Entwicklung der subjektiven Indikatoren zu danken.

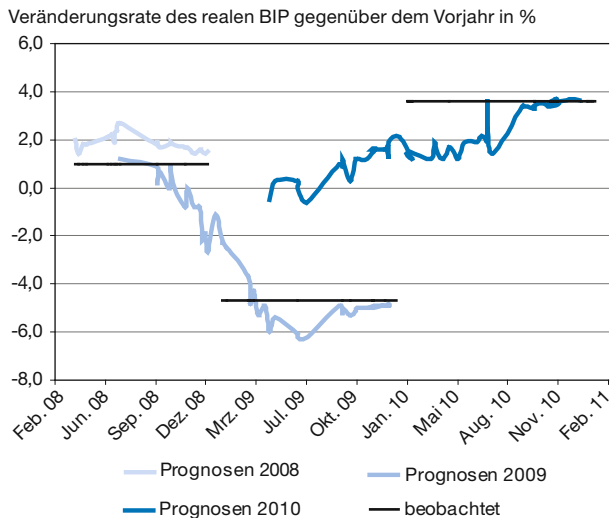
Konjunkturindikatoren und Krisenprognosen ...

Liegt es aber heute wirklich an den Informationen? Kann sich die Konjunkturanalyse hierzulande nicht auf das breiteste Informationsangebot seitens der amtlichen und nicht-amtlichen Statistik stützen – ungeachtet aller Kürzungen ih-

- 1 Vgl. dazu z.B. U. Heilemann, S. Wappler: „Haushalt geht vor Wirtschaft“ – Effizienz und Effektivität der fiskalischen Konjunkturpolitik 1967 bis 2010, erscheint demnächst.
- 2 Vgl. z.B. R. G. Rajan: Fault lines – How hidden fractures still threaten the world economy, Princeton 2010; J. Stiglitz: Im freien Fall – Vom Versagen der Märkte zur Neuordnung der Weltwirtschaft, Berlin 2010; G. A. Akerlof, R. J. Shiller: Animal spirits – How human psychology drives the economy and why it matters for global capitalism, Princeton 2009. Anders – der Titel spricht für sich – R. Skidelsky: Keynes – The return of the master, New York 2009. Zu einem wenig skeptischen Bild von den gegenwärtigen analytischen Möglichkeiten vgl. auch R. Fair: Analyzing macroeconomic forecastability. (Cowles Foundation and International Center for finance), New Haven 2009.
- 3 P. Krugman: A Dark Age of macroeconomics (wonkish), in: The New York Times vom 27. Januar 2009.
- 4 Vgl. dazu z.B. G. Minois: Geschichte der Zukunft – Orakel, Prophezeiungen, Utopien, Prognosen, aus dem Französischen von Eva Moldenhauer, Düsseldorf, Zürich 1998.

- 5 Vgl. dazu und dem Folgenden U. Heilemann: Déjà vu: Prognose in der Krise, in: Zeitgespräch „Welche Rolle spielen Prognosen?“, in: Wirtschaftsdienst, 89 Jg. (2009), H. 2, S. 90 ff.
- 6 Auch in der jüngsten Krise reagierten z.B. bereits allein die G7-Länder gemessen am Netto-Finanzierungssaldo 2008 mit -0,5% des BIP und je -1,6% 2009 und 2010; vgl. OECD: Economic Outlook Nr. 85, Paris 2009, S. 63.
- 7 Vgl. z.B. W. Roscher: System der Volkswirtschaft, Bd. 3, Nationalökonomie des Handels und des Gewerbetriebs, 3. Aufl., Stuttgart 1882, S. 770 ff. Teilweiser Wiederabdruck in: J. Burkhardt, B. P. Priddat: Geschichte der Ökonomie. Bibliothek der Geschichte und Politik, BDI, Bd. 21, Frankfurt a.M., S. 538 ff.

Abbildung 1
Ausgewählte¹ Wachstumsprognosen



¹ Im Einzelnen die Prognosen in diesem Zeitraum von Deutsche Bundesbank, DIW, EU-Kommission, „Gemeinschaftsdiagnose“, Ifo, IfW Kiel, IMF, IWH, Jahreswirtschaftsbericht der Bundesregierung und RWI.

Quelle: U. Heilemann, S. Wappler: „Haushalt geht vor Wirtschaft“ – Effizienz und Effektivität der fiskalischen Konjunkturpolitik 1967 bis 2010, erscheint demnächst, S.117.

res Programms („Verwaltungsvereinfachungen“)?⁸ Wie steht es um die Rolle und den prädikativen Gehalt von Indikatoren und Rankings mit Blick auf Krisen?

In der Tat lieferten z.B. die Gesamtindikatoren von Ifo und ZEW seit Herbst 2007 Hinweise auf eine Abschwächung, seit Frühjahr 2008 auf einen zunehmend dramatischen Rückgang der wirtschaftlichen Entwicklung. Die Konjunkturssignale der Gesamtindikatoren sind freilich qualitativer und gesamthafter Natur und haben eine prognostische Reichweite, die selten über drei Monate hinausreicht; sie deuten nicht immer alle in die gleiche Richtung und sind auch nicht frei von Fehlern erster wie zweiter Art.

Selbstverständlich finden im Rahmen der „Iterativen VGR“-Prognosen⁹ Indikatoren Verwendung. Wie bei eklektischen Vorgehensweisen dieser Art üblich, variieren Gewicht und Bedeutung der genutzten Informationen von Prognostiker zu Prognostiker und im Zeitablauf.

8 Vgl. z.B. U. Heilemann: Macroeconomic data, in: German Data Forum (RatSWD) (Hrsg.): Building on progress – expanding the research infrastructure for the social, economic, and behavioral sciences, Bd. 1, 2010, S. 289 ff.

9 Die iterativen VGR-Prognosen sind eine Kombination verschiedener Prognoseverfahren für die Volkswirtschaftliche Gesamtrechnung, die auch in Deutschland die am weitesten verbreitete Vorgehensweise ist, vgl. z.B. W. Nierhaus, J.-E. Sturm: Methoden der Konjunkturprognose, in: Ifo-Schnelldienst, 56. Jg. (2003), Nr. 4, S. 7 ff.

Prognostiker schauen selten zurück, auch die Analyse der Prognosen für die Krise hat bisher wenig Interesse gefunden. Waren die „Kosten“ einer definitiven „Rezessionsprognose“ – immerhin taucht sie im „Risikoszenario“ der „Gemeinschaftsdiagnose“ auf¹⁰ – den Prognostikern zu hoch?¹¹ Ob aus den jeweils verfügbaren Informationen die Krise zu erkennen war, müsste im Einzelnen untersucht werden, was hier nicht zu leisten ist. Da keiner die Krise sah, spricht dies dafür, dass sie einem neuen Muster folgte. In der Tat scheint die obere Wendepunktphase abrupt abgebrochen zu sein und mehr oder weniger unmittelbar in die Krisenphase übergegangen zu sein.¹²

Zum konjunkturdiagnostischen Informationsangebot sind bislang jedenfalls wenige Monita zu hören. Anders sieht es beim Finanzsektor aus, wobei vor allem auf die unzureichende statistische Erfassung des Sektors und die aus vielerlei Gründen instabilen Beziehungen innerhalb des Sektors hinzuweisen ist. An Indikatoren mangelt es alles in allem also offenbar nicht, aber ihre prognostische Reichweite und ihr Informationsgehalt sind beschränkt und wenig stabil.

... und Ratings und Rankings

Geben die in der Öffentlichkeit vielbeachteten Rankings von Volkswirtschaften, also die Bündelung von Indikatoren und die anschließende Kondensierung zu Kennzahlen Hinweise auf „Krisen“ oder mindestens Krisengefahren oder -dispositionen der betrachteten Volkswirtschaften? Bislang leider nicht.¹³ Unabhängig von ihrer nicht-konjunkturellen Orientierung fließen in die Rankings der „Internationalen Wettbewerbsfähigkeit“ von Volkswirtschaften weder Krisen und Krisendisposition ein, wie umgekehrt hohe „Internationale Wettbewerbsfähigkeit“ schwere Krisen nicht ausschließt oder – je nach Natur der Krisen – sogar dafür anfällig macht¹⁴ und damit in erheblichem Maße auch die gesamtwirtschaftliche Entwicklung bestimmt – zumindest dann nicht, wenn es sich um weltweite Krisen handelt wie

10 Projektgruppe Gemeinschaftsdiagnose: Deutschland am Rande einer Rezession – Gemeinschaftsdiagnose Herbst 2008, Essen 2008, S. 46.

11 Vgl. dazu und dem Folgenden D. Fintzen, H. O. Stekler: Why did forecasters fail to predict the 1990 recession?, in: International Journal of Forecasting, 15. Jg. (1999), S. 309 ff.

12 Vgl. R. Döhrn, G. Barabas, H. Gebhardt, T. Kitlinski, M. Micheli, T. Schmidt, S. Vosen, L. Zimmermann: Die wirtschaftliche Entwicklung im Inland: Feste Konjunktur in unsicherem Umfeld, in: RWI Konjunkturberichte, 62. Jg. (2010), S. 42 f., bzw. U. Heilemann, R. Schuhr, H. J. Münch: A perfect storm? – The present crisis and German crisis patterns, Universität Leipzig, Wirtschaftswissenschaftliche Fakultät, Working Paper, 93, Leipzig 2010.

13 Vgl. dazu und dem Folgenden U. Heilemann, H. Lehmann, J. Ragnitz: Länder-Rankings – Komplexitätsreduktion oder Zahlenalchemie?, in: Wirtschaftsdienst, 87. Jg. (2007), H. S. 480 ff.

14 Vgl. dazu z.B. E. Neuthinger: Staatsverschuldung und Wirtschaftspolitik, in: Wirtschaftsdienst, 91. Jg. (2011), H. 7, S. 486 ff.; oder Deutsche Bundesbank: Die außenwirtschaftliche Position Deutschlands vor dem Hintergrund zunehmender wirtschaftspolitischer Überwachung, in: Monatsberichte, 10, 63. Jg. (2011), S. 41 ff.

bei der Krise 2008/2009. Dies mussten nicht nur Irland oder die USA erfahren. Kurz: Der professionellen Analyse bieten diese Rankings nichts Neues, der Politik kaum Ansatzpunkte für konkretes Handeln – Komplexitätsreduktion hat ihren Preis.

Es liegt nahe, in diesem Zusammenhang kurz auch auf das Rating einzugehen, genauer: das Sovereign Credit Rating, also der Bestimmung des Risikos, dass ein staatlicher Emittent seine Kredite nicht vereinbarungsgemäß zurückzahlt.¹⁵ Als Prognoseverfahren spielten die Ratings keine Rolle, wohl aber ist die gesamtwirtschaftliche Entwicklung eines Landes eine wichtige Determinante des Ratingergebnisses, wie dieses umgekehrt mindestens seine fiskalische und kreditwirtschaftliche Situation und damit indirekt auch seine gesamtwirtschaftliche Entwicklung beeinflusst. Im Unterschied zu den Länderrankings ist also das Ziel der Ratinganalyse sehr viel bestimmter. Angesichts seiner kredit- und volkswirtschaftlichen Bedeutung unterliegen die Analysekonzepte und deren Anwendung einer strengen staatlichen Aufsicht, die letztlich zu erheblichen fachlichen und berufsethischen Anforderungen an die Bearbeiter führt.¹⁶ Selbstverständlich kommen die Beurteilungen, wie jede Kreditbeurteilung, nicht ohne subjektive Elemente aus und sind auch nicht frei von Fehlern – nicht anders als gesamtwirtschaftliche Prognosen, auf denen sie fußen, oder die erwähnten Rankings. Im Unterschied zu letzteren lassen diese sich jedoch identifizieren.

Blickt man aus der Informations- oder Prognose-Perspektive auf die Krise 2008/2009 in Deutschland zurück, so lassen sich also wie bei den vorausgegangenen Krisen ohne Frage beträchtliche Defizite ausmachen.¹⁷ Besondere Anstrengungen des Fachs, die datenseitigen Mängel zu beheben, sind in Deutschland bislang nur vereinzelt zu registrieren.¹⁸ Ins Auge springende Erfolge sind dabei noch nicht zu erkennen und mehr als inkrementelle Verbesserungen wohl auch nicht zu erwarten. Selbst die beträchtlichen Verbesserungen des Angebots der amtlichen und nicht-

amtlichen Statistik der letzten Jahrzehnte z.B. was Periodizität und Aktualität der VGR angeht, erhöhten die Treffsicherheit der gesamtwirtschaftlichen Prognosen kaum. Das bedeutet nicht, dass die Datenbasis nicht verbessert werden kann. Ein Bruchteil der Ressourcen, die jetzt auf die partielle („Wachstumsorientierung“) und fundamentale Umorientierung („Wohlstandsmessung“) der VGR verwendet werden, dürfte etliches bessern.

Und die Politik?

Zur zweiten Frage: Hat die Politik die Signale sinnvoll genutzt? Nochmals, paradoxerweise sind 2007 ff. die Gründe für die an den Lehrbuch-Erwartungen gemessenen Effizienz- und Effektivitätsdefizite der Krisenbekämpfung nicht in den exorbitanten Prognosefehlern zu sehen. Sondern zum einen unmittelbar bei den Akteuren: Der Finanzminister zögerte Monate, da er durch fiskalpolitische Maßnahmen seine ehrgeizigen Konsolidierungspläne bedroht sah. Die bis dahin generellen Zweifel an der Effektivität fiskalpolitischer Maßnahmen, „crowding out“-Befürchtungen oder Sorgen um Lohnmäßigung, spielten dieses Mal zum Glück keine Rolle. Zögerlich waren auch die Maßnahmen selbst, und sie erfolgten auch dieses Mal in zwei Schritten, was ihre Wirkungen schmälerte. Natürlich kann dabei auch auf prognostische Unsicherheiten verwiesen werden, zumal ausgabenpolitische Maßnahmen weniger flexibel als steuer- oder geldpolitische sind.

Darüber hinaus litten Effizienz und damit auch Effektivität der Interventionen wiederum unter zahlreichen Implementierungs- und Wirkungsverzögerungen – vom Beschluss der Maßnahmen bis zu ihren ersten Wirkungen verging abermals mehr als ein halbes Jahr. Eine der Ursachen war ohne Frage, dass die im Stabilitätsgesetz vorgesehenen Schubladenprogramme nicht vorlagen. Das verzögerte die Implementierung der Maßnahmen und hatte auch zur Folge, dass – anders als im Stabilitätsgesetz vorgesehen – eben keine ohnehin vorgesehenen Ausgaben getätigt wurden, sondern „zusätzliche“, denen künftig keine Minderausgaben gegenüberstehen werden.

Fazit

Der Anteil der Daten- und Prognosedefizite an den Mängeln der Krisenbekämpfung 2008/9 hält sich in Grenzen. Auch wenn die fiskalischen Impulse von insgesamt knapp 5½% des BIP ein halbes Jahr früher gegeben worden wären – der Einbruch 2009 hätte noch immer bei 4% des realen BIP, also weit über allen Erfahrungen gelegen. Dies gilt erst recht, wenn berücksichtigt wird, dass das insgesamt kurzfristig zu mobilisierende Fiskalvolumen zur Krisenbekämpfung gegenwärtig kaum über 2½% des BIP hinausgehen dürfte. Mit Wirtschaftskrisen haben viele

15 Zu Entstehung, Struktur, Praxis, Bedeutung und Problemen des Ratinggeschäfts vgl. z.B. T. J. Sinclair: *The New Masters of capital – American bond rating agencies and the politics of creditworthiness*, Ithaca, London 2008. Zur Rolle von Rating Agenturen in der Finanzkrise seit 2007 in den USA; vgl. Financial Crisis Inquiry Commission: *The Financial Crisis Inquiry Report – Final report of the National Commission on the causes of the financial and economic crisis in the United States*, submitted by the Financial Crisis Inquiry Commission pursuant to Public Law 111-21, New York 2011.

16 Für Europa vgl. z.B. die Verordnung (EG) Nr. 1060/2009 des Europäischen Parlamentes und des Rates vom 16.9.2009 über Ratingagenturen und das am 19.9.2010 in Kraft getretene dazugehörige Ausführungsgesetz des Deutschen Bundestages.

17 Vgl. dazu im Einzelnen U. Heilemann, S. Wappler, a.a.O.

18 Vgl. z.B. K. A. Kholodilin, B. Siliverstovs: *Do forecasters inform or reassure? Evaluation of German real time data*, DIW-Discussion Papers, 888, Berlin 2008; R. Döhrn, C. M. Schmidt: *Information or institution? On the determinants of forecast accuracy*, in: *Jahrbücher für Nationalökonomie und Statistik*, Bd. 231, 2011, S. 9 ff.

eben nicht mehr rechnen wollen und dazu passte es, dass der Wissenschaftliche Beirat beim Bundesministerium für

Wirtschaft und Technologie noch 2008 die Aufgabe des Stabilitätsgesetzes forderte.

Christoph M. Schmidt*

Möglichkeiten und Grenzen umfassender Indikatorensysteme

Rezessionen und größere Krisen erinnern uns daran, dass wirtschaftliche Schwächephasen ein Teil des modernen Lebens sind. Die Finanz- und Wirtschaftskrise der vergangenen Jahre hat uns zudem eindrucksvoll vor Augen geführt, dass sich eine vordergründig günstig verlaufene Phase der wirtschaftlichen Entwicklung in retrospektiver Betrachtung ganz anders darstellen kann. So zeigten die Immobilienpreisblasen in den USA und manchen europäischen Ländern zunächst ein rapide ansteigendes Privatvermögen an, bevor ihr Platzen eine globale Rezession auslöste. Rasches materielles Wachstum mag genuine wirtschaftliche Potenziale widerspiegeln, kann aber auch auf Kosten der Umwelt oder der langfristigen wirtschaftlichen Stabilität erkaufte sein. Aus dieser Einsicht ergibt sich für Ökonomik und Statistik die Aufgabe, in der regelmäßigen statistischen Berichterstattung auch Hinweise auf Probleme der ökonomischen, sozialen oder ökologischen Nachhaltigkeit zu geben.

Um den tatsächlichen Stand von Wohlstand und Fortschritt möglichst gut zu erfassen, sollte das statistische Berichtswesen, das sich traditionell auf „harte“ Indikatoren der Wirtschaftsleistung, insbesondere das Bruttoinlandsprodukt (pro Kopf), konzentriert hat, in geeigneter Weise um nicht-materielle Aspekte der Lebensqualität und um Indikatoren der Nachhaltigkeit ergänzt werden. Diese Arbeit kann auf umfassende Vorarbeiten aufbauen. Sowohl die Sozialberichterstattung als auch die Berichterstattung zur Nachhaltigkeit können in Deutschland aus einer jahrelangen Erfahrung schöpfen; im internationalen Kontext haben die EU-Kommission und die OECD in den vergangenen Jahren umfangreiche Bemühungen zum Zwecke einer reichhaltigeren statistischen Berichterstattung angestrengt.

In der öffentlichen Diskussion hat die vom französischen Präsidenten im Jahr 2009 in Auftrag gegebene Analyse der hochkarätigen Kommission um das Führungstrio Stiglitz, Sen und Fitoussi (SSF) die größte Aufmerksamkeit auf sich gezogen.¹ Ihr Abschlussbericht skizziert „Wirtschaftsleistung“, „Lebensqualität“ und „ökologische Nachhaltigkeit“

als die drei Säulen einer umfassenden Wohlfahrtsberichterstattung und gibt erste Empfehlungen für ihre Ausgestaltung. Im Jahr 2010 haben der deutsche Sachverständigenrat für Wirtschaft (SVR) und sein französisches Pendant, der Conseil d'Analyse économique (CAE), diese Arbeiten in einer gemeinsamen Expertise vertieft und zum konkreten Vorschlag eines umfassenden Indikatorensystems für Wirtschaftsleistung, Lebensqualität und Nachhaltigkeit weiterentwickelt.² Diese Expertise bildet den Hintergrund für die weiteren Ausführungen.

Grundsätze für eine umfassende Berichterstattung

Eine wissenschaftlich fundierte Studie muss dem Anspruch genügen, die von ihr vorgeschlagenen Indikatoren zur umfassenden Wohlfahrtsberichterstattung prinzipiengestützt abzuleiten. Natürlich ist es nicht möglich, bei der Auswahl zu einer eindeutig optimalen Lösung zu kommen. Das von CAE und SVR vorgeschlagene Indikatorenbündel ist daher auch das Ergebnis von Abwägungen und Wahlentscheidungen, aber keinesfalls eine ad hoc entschiedene Zusammenstellung.

Am Anfang steht die ernüchternde Erkenntnis, dass nicht alles Wünschenswerte auch machbar ist. So erweist sich die Entscheidung, im Indikatorensatz nicht-materielle Aspekte der Lebensqualität prominent zu führen, gleichzeitig als eine Entscheidung gegen den Versuch, vollkommene internationale Vergleichbarkeit herzustellen. Denn zumindest ein gewisses Ausmaß an internationaler Vergleichbarkeit ist nur dann zu gewährleisten, wenn man sich auf den Kontrast zwischen objektiv messbaren Größen bezieht. Würden hingegen subjektive Indikatoren zur Lebensqualität einen hohen Stellenwert erhalten, dann wäre bestenfalls ein intertemporaler Kontrast innerhalb eines Landes aussagekräftig, nicht aber der Vergleich zwischen verschiedenen Ländern.

Ähnliche Überlegungen betreffen die Berichterstattung zur Nachhaltigkeit. Wenn es darum geht, Aussagen zur lang-

* Der Autor dankt Nils aus dem Moore herzlich für seine Kommentare.
1 J. E. Stiglitz, A. Sen, J. Fitoussi: Report by the Commission on the Measurement of Economic Performance and Social Progress, 2009, <http://www.stiglitz-sen-fitoussi.fr>.

2 CAE/SVR – Conseil d'Analyse économique und Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung: Wirtschaftsleistung, Lebensqualität und Nachhaltigkeit: Ein umfassendes Indikatorensystem. Expertise im Auftrag des deutsch-französischen Ministerrates, Wiesbaden 2010.

fristigen Stabilität eines ökonomischen, sozialen oder ökologischen Systems zu treffen, dann muss man aus dem vorhandenen Erfahrungsschatz heraus in geeigneter Weise extrapolieren. Dies stellt an die Daten ganz andere Anforderungen als die Anfertigung eines bloßen Schnappschusses. Das „Ist“ zu beschreiben, ist eine kleinere Herausforderung, als die Frage „Was wäre, wenn?“ zu beantworten. Je anspruchsvoller also das Berichtskonzept angelegt ist, desto eher mag es in der Praxis auf unüberwindbare Informationsdefizite stoßen.

Für eine praktikable Umsetzung einer umfassenden Wohlfahrtsmessung sind darüber hinaus zunächst die Möglichkeiten und Grenzen auszuloten, die durch die Vorarbeiten der statistischen Ämter eröffnet bzw. gesetzt werden. Letztendlich muss sich die Schar der vorgeschlagenen Indikatoren aus einem Reservoir speisen, das Informationen regelmäßig, zeitnah und in einer Art und Weise bereitstellt, die sich gegenüber kleinen Änderungen robust zeigt. Das schließt den Ruf nach neuen Datensätzen nicht aus, setzt aber deren dauerhafte Erhebung voraus, wenn ihre Information in die regelmäßige Berichterstattung einfließen soll.

In einem zweiten Schritt ist für jeden zusätzlichen Indikator zu prüfen, welchen Erkenntnisgewinn seine Berücksichtigung bringen würde. Dabei wird sich in vielen Fällen das folgende Muster herauschälen: Indikatoren, die gegenüber dem BIP einen hohen Informationsgehalt aufweisen, da sie nicht sehr stark mit dieser Größe korreliert sind, können häufig nicht sonderlich präzise erfasst werden. Solche Indikatoren hingegen, die man präzise erfassen kann, sind jedoch häufig so hoch mit dem BIP korreliert, dass ihr Informationsgewinn bescheiden bleibt.³ Die Kunst ist es daher, die beste Balance zwischen „Informationsgehalt“ und „statistischer Belastbarkeit“ zu finden.

Schließlich ist zu bedenken, dass die Erhebung von Statistiken meist erhebliche Ressourcen in Anspruch nimmt. Dort, wo sich ohne die zusätzliche Erhebung von Informationen kein Erkenntnisgewinn erzielen lässt, sind die Grenzerträge der erweiterten Erkenntnis und die Kosten dieser Erhebung gegeneinander abzuwägen und die Wünsche hinsichtlich ihrer Priorität zu ordnen.

Konzentration aufs Wesentliche

Wenngleich es dem Anspruch einer umfassenden Berichterstattung zu Wohlstand und Fortschritt entspricht, eine Fülle von Einsichten zu generieren, ist es dennoch wichtig, dass sie sich in angemessener Weise darauf konzentriert,

was sie wirklich zu leisten in der Lage ist. Insbesondere muss sie sich auf erkenntnistheoretisch positive Fragestellungen konzentrieren. Die erste dieser Fragen ist „Wie steht es um die Lebensqualität tatsächlich?“ Ganz offensichtlich ist bereits die Beantwortung dieser grundlegenden Frage angesichts der Breite des diskutierten Wohlfahrtsbegriffs alles andere als einfach. Ihr sollte im Rahmen der regelmäßigen Berichterstattung zu Wohlstand und Fortschritt die größte Aufmerksamkeit zukommen.

Die zweite Frage ist aus Sicht der statistischen Arbeit nochmals deutlich anspruchsvoller: „Welche Zusammenhänge gibt es zwischen materiellen und nicht-materiellen Aspekten der Lebensqualität?“ Es wird einer erweiterten Wohlfahrtsmessung wenig nützen, wenn die nicht-materiellen Indikatoren im Wesentlichen ein quantitativ leicht verändertes, doch im Grundsatz sehr ähnliches Bild zeichnen wie die Indikatoren der Wirtschaftsleistung. Nun wird es wohl kaum sinnvoll sein, im Rahmen der statistischen Berichterstattung für alle erhobenen Indikatoren neben den Mittelwerten auch die Kovarianzen zu berechnen. Aber zumindest bei der Ersteinführung von Indikatoren zur Lebensqualität sollte diese Frage ausführlich erörtert werden.

Mit der Erwartung, auf weitergehende Fragen – etwa nach den Auswirkungen bestimmter wirtschaftspolitischer Maßnahmen oder gar dem Hinweis darauf, was wirtschaftspolitisch zu tun ist – Auskünfte zu erhalten, wäre eine regelmäßige statistische Berichterstattung in der Form eines festen Indikatorensystems jedoch überfordert. Eine überzeugende Antwort auf die Frage nach der Wirkung wirtschaftspolitischer Maßnahmen muss von Fall zu Fall mit einem maßgeschneiderten Studiendesign gesucht werden.⁴ Nur durch sorgfältige Evaluationsstudien können die Wirksamkeit vergangener wirtschaftspolitischer Weichenstellungen beurteilt und belastbare Anhaltspunkte für erfolgversprechende neue Maßnahmen gewonnen werden.

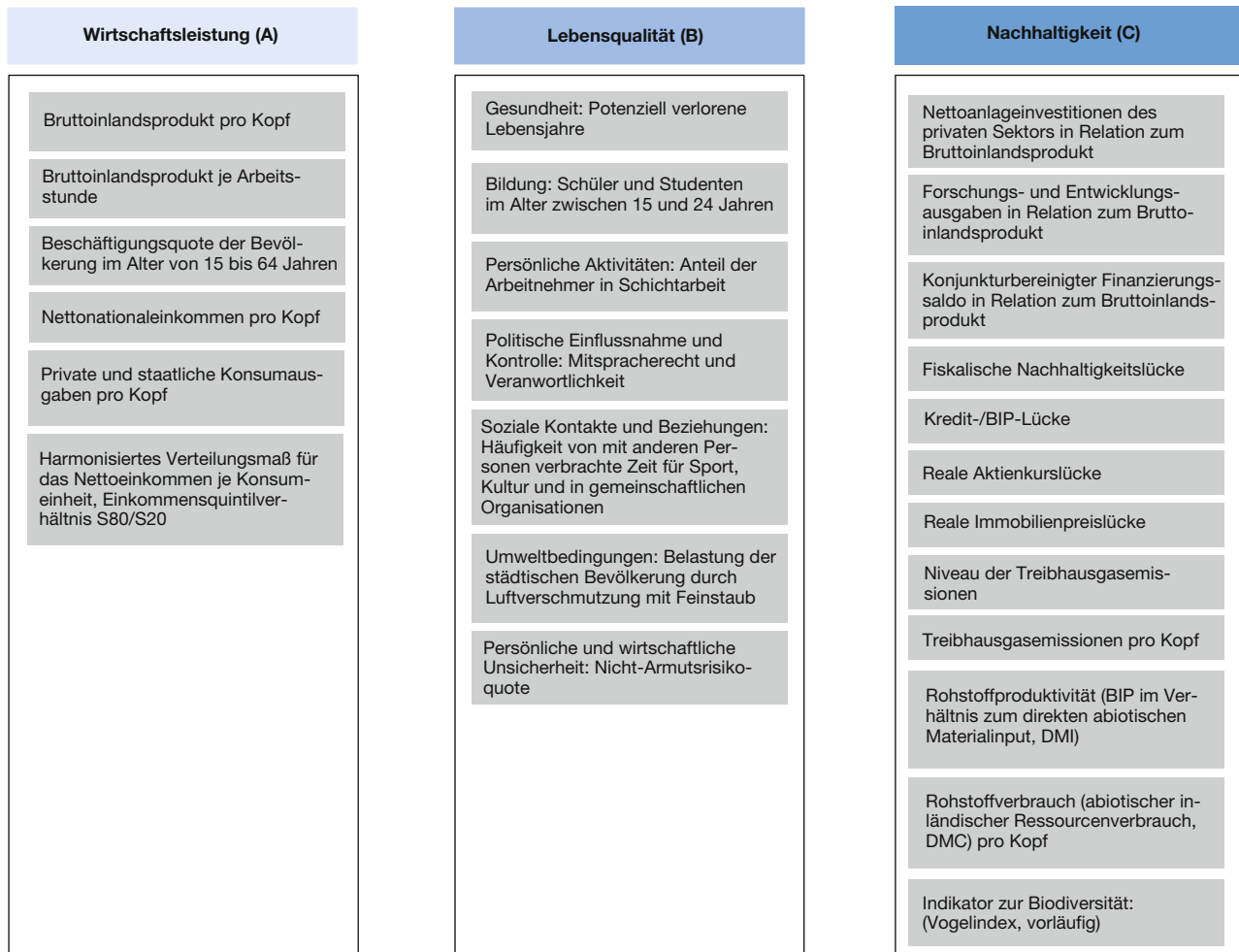
Schließlich gilt: Statistik ist Informationsreduktion, da viele einzelne Informationsbausteine in ein gemeinsames Bild zu verdichten sind. Diese Verdichtung erfordert immer eine gewisse Gewichtung der Einzelinformationen. Sie erfolgt über Personen, Unternehmen und andere Akteure hinweg, meist ohne dass dies größere konzeptionelle Probleme aufwirft. Dies gilt jedoch nur so lange, wie es sich um metrische Informationen handelt. Es wird jedoch zu einer gewissen Herausforderung, wenn die Verdichtung über eine Mischung von quantitativen und qualitativen Indikatoren hinweg erfolgen soll. Das Zusammenführen von Einzelaspekten der Lebensqualität oder der Nachhaltigkeit ist daher niemals eine harmlose Angelegenheit.

3 S. C. Kassenböhmer, Ch. M. Schmidt: Beyond GDP and Back: What is the Value-Added by Additional Components of Welfare Measurement?, Ruhr Economic Papers #239, Essen 2010.

4 Th. K. Bauer, M. Fertig, Ch. M. Schmidt: Empirische Wirtschaftsforschung. Eine Einführung, Berlin, Heidelberg 2009.

Abbildung 1

Indikatorensystem für Wirtschaftsleistung, Lebensqualität und Nachhaltigkeit



Quelle: CAE/SVR – Conseil d'Analyse économique und Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung: Wirtschaftsleistung, Lebensqualität und Nachhaltigkeit: Ein umfassendes Indikatorensystem. Expertise im Auftrag des deutsch-französischen Ministerrates, Wiesbaden 2010.

Bei Verbundindikatoren, die aus mehreren Variablen zusammengesetzt werden, öffnet das Gewichtungsproblem der Manipulierbarkeit im Prinzip Tür und Tor, wenn das Konzept einmal von der unabhängigen Wissenschaft aus der Hand gegeben wurde. Aus Sicht der Statistik muss daher alles daran gesetzt werden, dass das vorgeschlagene Indikatorensystem zur Wohlfahrtsberichterstattung gegen jegliche Versuche der Manipulation von vornherein gefeit ist. Nicht zuletzt aus dieser Einsicht heraus hat sich die CAE/SVR-Studie gegen eine übermäßige Verdichtung der Einzelinformationen ausgesprochen.

Teilbereiche einer umfassenden Wohlfahrtsberichterstattung

Sowohl der Report von Stiglitz, Sen und Fitoussi als auch die CAE/SVR-Expertise haben jeweils drei Teilbereiche als

Säulen einer umfassenden Wohlfahrtsberichterstattung diskutiert: Wirtschaftsleistung, Lebensqualität und Nachhaltigkeit. CAE und SVR haben in dieser Logik ein konkretes Indikatorensystem entwickelt (vgl. Abbildung 1).

Die erste Gruppe von Indikatoren betrifft Maße der Wirtschaftsleistung und des materiellen Wohlstands. Hier geht es vor allem um die Verbesserung traditioneller „harter“ Maße der regelmäßigen statistischen Berichterstattung, insbesondere des BIP. Als Maß der Wirtschaftsleistung und nicht als Wohlfahrtsmaß konzipiert, leistet das BIP zwar gute Dienste für die Vorbereitung zeitnaher stabilisierungspolitischer Entscheidungen und stabiler finanzwissenschaftlicher Planung. Es hat jedoch gewisse Probleme, wirtschaftliche Aktivitäten vollständig zu erfassen. Die Haushaltsproduktion oder Aktivitäten der Schattenwirtschaft entziehen sich der direkten statistischen Berichterstattung. Darüber hin-

aus ist es immer dann besonders schwer, den Wert einer Dienstleistung zu erfassen, wenn sie nicht über den Markt erbracht wird.

Diese und andere Mess- und Erfassungsprobleme werden von der amtlichen Statistik sukzessive adressiert und stellen sie bei der Aufgabe der zeitnahen Erfassung der Wirtschaftsleistung nicht vor prohibitiv große Probleme. Allerdings gilt es, wie es bereits seit fast fünf Jahrzehnten bei den Gutachten des SVR üblich ist, neben dem BIP weitere Maßzahlen in den Blick zu nehmen, um ein vollständiges Bild der aktuellen wirtschaftlichen Lage zu erhalten. Dazu gehören etwa die Arbeitslosenquote oder die Inflationsrate. Will man die Möglichkeiten zu internationalen Vergleichen wahren, dann müssen Maße wie das BIP auf die Bevölkerung eines Landes bezogen sein: Nicht ohne Grund ist das BIP pro Kopf in der makroökonomischen Fachliteratur der zentrale Diskussions- und Erkenntnisgegenstand.

Schließlich sollte die Verteilung ökonomisch relevanter Größen stärkere Beachtung finden. Zu einem vollständigen Bild über die Einkommenssituation einer Gesellschaft gehört daher als gleichwertiger Informationsbaustein auch ein Maß der Verteilung. Es bietet sich an, ein Ungleichheitsmaß wie den Gini-Koeffizienten oder – besser, da für die breite Öffentlichkeit leichter zu interpretieren – den Quotienten der Einkommen unterschiedlicher Dezile zu verwenden, etwa den 80/20-Quotienten. Dabei sollte man sich allerdings der Einschränkung bewusst sein, dass die Einkommen am unteren und am oberen Ende der Verteilung in der Regel nur unzureichend erfasst werden.

Die Indikatoren in der zweiten Säule dienen der Erweiterung der Berichterstattung in Richtung ganzheitlicher Lebensqualität. Ausgehend von der Einsicht, dass materielle Aspekte von essentieller Bedeutung für die menschliche Existenz sind, geht es nicht darum, das BIP oder die anderen Indikatoren der ersten Säule zu ersetzen, sondern um die zusätzliche Berücksichtigung nicht-materieller Aspekte der Lebensqualität. Überlegungen zu den Grenzen der Messbarkeit von Emotionen und Neigungen legen nahe, diese Ergänzung nicht in der Auswertung einer übergeordneten Frage nach der subjektiven Einschätzung der aktuellen Lebensqualität zu suchen, sondern in detaillierter Information zu einzelnen Facetten der nicht-materiellen Lebensqualität. Die dabei angesprochenen Dimensionen wie Gesundheit, Bildung oder Freiheit lassen sich allerdings schwerlich in ein übergreifendes Bild verdichten. Denn die individuellen Abwägungen, die Menschen zwischen diesen einzelnen Facetten treffen, sind empirisch kaum zu erfassen.

Letztendlich gilt auch für diesen Bereich des vorgeschlagenen Indikatorensetzes, dass robust und regelmäßig

erhobene statistische Informationen, beispielsweise zu Aspekten des Gesundheitszustands der Bevölkerung, genutzt werden sollten, bevor man sich darauf einlässt, subjektive Befragungsdaten in den Mittelpunkt der Berichterstattung zu stellen. Zu sensitiv reagieren erfahrungsgemäß die Antworten auf die konkrete Formulierung der Fragen, zu schwankend erscheinen häufig die subjektiven Einschätzungen im Zeitablauf, und zu sehr mag die Versuchung obsiegen, auf derartige Fragen mit strategischem Motiv zu antworten.

Die letzte Gruppe von Indikatoren nimmt in der dritten Säule schließlich zukunftsbezogene Aspekte der Nachhaltigkeit in den Blick. Die CAE/SVR-Studie geht dabei deutlich über die Arbeiten des Reports von Stiglitz, Sen und Fitoussi hinaus, indem sie neben der ökologischen Nachhaltigkeit explizit auch ökonomische Aspekte der Nachhaltigkeit diskutiert. Dabei werden sowohl die Möglichkeiten der Volkswirtschaft ausgelotet, zukünftig in angemessener Weise zu wachsen, reflektiert in den laufenden Investitionen in Kapital sowie Forschung und Entwicklung, als auch die langfristige Tragfähigkeit der öffentlichen Finanzen. Zudem wird nach Alarmsignalen gefahndet, die eine Blasenbildung auf Vermögens- oder Finanzmärkten abbilden können.

Zwei Aspekte erschweren jedoch die statistische Berichterstattung zur Nachhaltigkeit. Zum einen bedarf es hier einer Abkehr von der üblichen Vorgehensweise der amtlichen Statistik, die sich auf eine solide Beschreibung des Ist-Zustandes konzentriert. Denn bei der Untersuchung der Stabilität ökonomischer, sozialer oder ökologischer Systeme sind Fragen mit „Was wäre, wenn“-Charakter zu beantworten. Dafür sind Projektionen notwendig, die gegenwärtige Aktivitäten und Entscheidungen in die Zukunft fortschreiben und so die möglichen Konsequenzen eines bereits eingeschlagenen Aktionspfades beschreiben. Es handelt sich hierbei wohlgerne nicht um Prognosen, die gerade keine Handlungsstabilität unterstellen, sondern das künftige Entscheidungsverhalten der relevanten Akteure aus umfassenden Modellen ableiten. Das kann eine regelmäßige statistische Berichterstattung natürlich nicht leisten. Nachhaltigkeitsberichte zielen also darauf ab, durch ihr Mahnen künftige Weichenstellungen zu verändern, nicht künftige Entwicklungen vorherzusagen.

Zum anderen sind Aussagen zur Nachhaltigkeit, die sich auf nationale Umstände konzentrieren, weitgehend nutzlos, denn Wirtschaftskrisen und ökologische Probleme sind nur im internationalen Kontext zu verstehen: Weder die Auswirkungen einer Instabilität des Finanzsystems noch ökologische Herausforderungen wie die Emission von Klimagasen machen an den Grenzen der eigenen Volkswirtschaft halt. Die statistische Berichterstattung zur

Nachhaltigkeit muss daher grundsätzlich die internationale Dimension prominent berücksichtigen.

Vorläufiges Fazit

Bei aller Sehnsucht nach besseren, allumfassenderen, ganzheitlichen Maßen für Stand und Perspektiven des menschlichen Wohlstands haben sich aus meiner Sicht drei Erkenntnisse herausgeschält:

- Erstens sollten nach wie vor materielle Aspekte im Zentrum der Diskussion stehen. Sie werden vergleichsweise verlässlich erfasst, sind einer klaren Interpretation zugänglich und – was am wichtigsten ist – enthalten wertvolle Informationsbausteine zum Erkenntnisgegenstand.
- Zweitens sollte die Verdichtung der Einzelinformationen nicht überzogen werden: Eine aufgeklärte Gesellschaft muss mit einem Minimum an Komplexität umgehen können, wenn sie umfassend und transparent informiert werden will. „Ganzheitlich“ bedeutet somit in der Berichterstattung keineswegs „eine einzelne Zahl“.
- Drittens lohnt es sich zweifellos, Weiterungen der traditionellen Berichterstattung im Sinne des Reports von Stiglitz, Sen und Fitoussi und der CAE-SVR-Expertise zu explorieren, aber es gibt bei dieser Unternehmung wichtige Vorbehalte. So sind Verteilungsmaße bereits bei der Analyse des materiellen Wohlstands aus statistischer Sicht recht problematisch, denn sie stoßen an den Rändern der Verteilung oft an die Grenzen der Repräsentativität. Daher sind Erweiterungen des Berichtsspektrums in Richtung von Verteilungsmaßen zu subjektiven Indikatoren wohl mit äußerster Vorsicht zu genießen.

Darüber hinaus zeigt die aktuelle Literatur zum Thema „Happiness“ aus meiner Sicht durchaus, dass subjektive Befragungsdaten interessante Informationen zu Tage fördern können. Es dürfte ihnen jedoch nicht möglich sein, das (durchschnittliche) Niveau des Wohlbefindens in einer transparenten und nachvollziehbaren Art und Weise abzubilden.

Schließlich sollten bei der statistischen Berichterstattung zum Thema Nachhaltigkeit keine eindeutigen Antworten erwartet werden: Denn die Extrapolation aus dem bisherigen Erfahrungsschatz im Sinne eines „Was wäre, wenn“ muss immer in Abhängigkeit von Annahmen zum Verhalten aller relevanten Akteure geschehen. Trotzdem haben die auf diese Weise gewonnenen Projektionen einen hohen Wert: Sie zeigen uns, wohin der Weg führen kann, wenn wir die Richtung nicht ändern.

Die Autoren des Zeitgesprächs



Prof. Dr. Ulrich van Suntum ist Geschäftsführender Direktor des Zentrums für angewandte Wirtschaftsforschung der Universität Münster.



Anja Rohwer, Dipl.-Volkswirtin, ist Vorstandsreferentin am ifo Institut – Leibniz-Institut für Wirtschaftsforschung an der Universität München.



Prof. Dr. Oliver Hülsewig lehrt Wirtschaftspolitik an der Fakultät für Betriebswirtschaft der Hochschule für angewandte Wissenschaften in München und ist Research Associate am ifo Institut – Leibniz-Institut für Wirtschaftsforschung an der Universität München.



Prof. Dr. Ullrich Heilemann war langjähriger Vizepräsident des Rheinisch-Westfälischen Instituts für Wirtschaftsforschung (RWI) in Essen und zuletzt Direktor des Instituts für Empirische Wirtschaftsforschung (IEW) der Universität Leipzig.



Prof. Dr. Christoph M. Schmidt ist Präsident des Rheinisch-Westfälischen Instituts für Wirtschaftsforschung (RWI) in Essen, lehrt Wirtschaftspolitik und angewandte Ökonomie an der Ruhr-Universität Bochum und ist Mitglied des Sachverständigenrates zur Begutachtung der gesamtwirtschaftlichen Entwicklung.