

Jens Weidmann\*

## Hat sich die Prognosetreffsicherheit des Sachverständigenrates systematisch verändert?

Die häufigen Prognoserevisionen im Verlauf des Jahres 2001 und die deutliche Überschätzung der wirtschaftliche Aktivität durch sämtliche Prognoseinstitutionen hat die Öffentlichkeit irritiert und eine Diskussion über die Treffsicherheit und den Nutzen von Konjunkturprognosen ausgelöst<sup>1</sup>. Zwei Autoren, H. Dicke und H. H. Glismann<sup>2</sup>, gehen in diesem Heft sogar so weit, die Fehlprognosen des vergangenen Jahres auf eine Verschlechterung der prognostischen Arbeit im Zeitverlauf zurückzuführen, die ihrer Auffassung nach im Widerspruch zu den produktivitätssteigernden Effekten der neuen Technologien steht, die sich in einer erhöhten Treffsicherheit der Prognosen hätte niederschlagen müssen.

Im Folgenden soll untersucht werden, ob sich die Prognosequalität während der letzten vier Jahrzehnte systematisch verändert hat. Hierzu werden aus Gründen der Vergleichbarkeit mit der Studie von Dicke und Glismann die Prognosen des Sachverständigenrates zur Begutachtung der gesamtwirtschaftlichen Entwicklung für das Bruttoinlandsprodukt (BIP) herangezogen. Der Sachverständigenrat prognostiziert seit dem Jahre 1964 in seinen im November erscheinenden Jahresgutachten die wirtschaftliche Entwicklung des laufenden und des folgenden Jahres<sup>3</sup>.

### Systematische Einflüsse auf die Konjunkturprognose

Konjunkturprognosen sind Erwartungen über die zukünftige Entwicklung volkswirtschaftlicher Aggregate, die aus den zwischen den einzelnen relevanten Größen bestehenden definitorischen und in der Vergangenheit beobachteten Funktionalzusammenhängen sowie Annahmen in Bezug auf exogene Variablen abgeleitet werden. Sie sind ex post fast immer falsch. Das Ausmaß des Prognosefehlers hängt entscheidend von der Richtigkeit der gemachten Annahmen und Setzungen bezüglich der exogenen Größen, wie beispielsweise der Entwicklung der Rohstoffpreise und dem Kurs der Lohn- und Finanzpolitik, sowie der

Angemessenheit der zugrunde gelegten strukturellen Zusammenhänge ab.

Systematische Einflüsse auf die Prognosegüte – definiert als Funktion der Abweichungen der Prognosewerte von den Realisationen – gehen somit im Wesentlichen von der Fähigkeit des Prognostikers aus, die Struktur der Ökonomie abzubilden und die Entwicklung der exogenen Variablen adäquat zu setzen. Diese Fähigkeit wird ihrerseits determiniert (1) durch die Komplexität der strukturellen Zusammenhänge, (2) durch die Bedeutung der Zufallskomponente, d.h. die Größe der – per definitionem – nicht prognostizierbaren Schocks, (3) durch das Auftreten von Strukturbrüchen und die Geschwindigkeit des Strukturwandels, (4) durch die Zuverlässigkeit des statistischen Datenermaterials und potentiell (5) durch den technischen und wissenschaftlichen Fortschritt. Vor dem Hintergrund der hier betrachteten Fragestellung sind insbesondere die letzten drei Faktoren relevant.

### Strukturbrüche und Strukturwandel

Plötzliche Strukturbrüche oder kontinuierliche Veränderungen der ökonomischen Strukturen erschweren die Prognosetätigkeit, da die aus der Datenhistorie abgeleiteten und der Prognose zugrunde liegenden und als stabil angenommenen empirischen Regelmäßigkeiten nur noch begrenzte Aussagekraft besitzen. Eine der bedeutendsten strukturellen Veränderungen der letzten Jahre stellt die deutsche Vereinigung dar. Durch die Vereinigung änderte sich der institutionelle Rahmen und damit auch die Verhaltensmuster der Wirtschaftsakteure in den neuen Bundesländern gravierend. Der Transformationsprozess war in seinen Auswirkungen nur schwer abzuschätzen, was

\* Der Autor dankt Dr. Martin Albrecht, Sandra Eickmeier und Dr. Stephan Kohns für kritische Kommentare.

<sup>1</sup> Siehe z.B. W. Nierhaus: Konjunkturprognosen und Prognosereisiko, in: ifo Schnelldienst 16/2001, 2001, S. 17-21; oder M. Grömling: Konjunkturprognosen: Methoden, Risiken und Treffsicherheiten, in: iw-trends, 2/2002.

<sup>2</sup> H. Glicke, H. H. Glismann: Konjunkturprognosen und wissenschaftlich-technischer Fortschritt, in: WIRTSCHAFTSDIENST, 82. Jg. (2002), H. 3, S. 167-169; und H. Dicke, H. H. Glismann: Haben sich die Konjunkturprognosen des Sachverständigenrates verbessert?, in diesem Heft.

<sup>3</sup> Der Sachverständigenrat prognostiziert mehr als 20 Größen der Volkswirtschaftlichen Gesamtrechnungen, von denen das Bruttoinlandsprodukt regelmäßig das größte öffentliche Interesse hervorruft.

*Dr. Jens Weidmann, 34, ist Generalsekretär des Sachverständigenrates zur Begutachtung der gesamtwirtschaftlichen Entwicklung, Wiesbaden. Der Autor vertritt hier seine persönliche Meinung.*

sich letztlich auch auf die Prognoseunsicherheit für Deutschland insgesamt auswirkte. Hinze<sup>4</sup> stellt hierzu fest, dass die Treffgenauigkeit von Konjunkturprognosen in der ersten Hälfte der neunziger Jahre merklich unter den Auswirkungen der deutschen und europäischen Einigung gelitten hat, allerdings weniger in Bezug auf die gesamtwirtschaftliche Wachstumsrate als in Bezug auf einzelne Verwendungskomponenten.

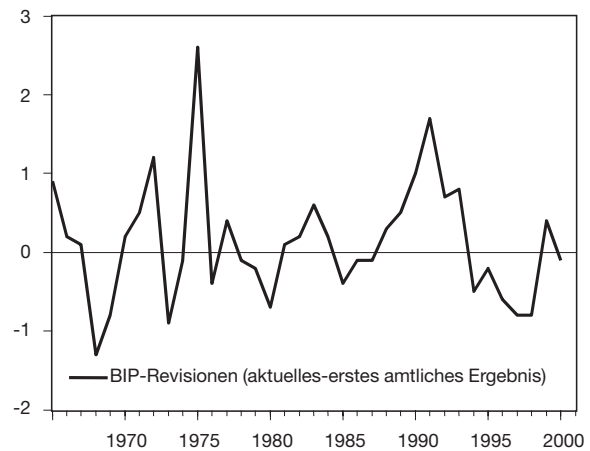
Während in den ersten Jahren nach der deutschen Vereinigung das Berichtssystem für die neuen Bundesländer noch größere Lücken aufwies, bestand später das prognoserelevante Problem vor allem in der regionalen Zuordnung wirtschaftlicher Transaktionen auf die beiden Gebietsstände. Der Sachverständigenrat bemängelte entsprechend in seinem Jahresgutachten 1994/95, dass die Schätzfehler in den einzelnen Verwendungsaggregaten der Volkswirtschaftlichen Gesamtrechnungen (VGR) ein „bedenklich“ großes Ausmaß erreicht hatten. Mit der Beendigung der gesonderten Volkswirtschaftlichen Gesamtrechnungen für die neuen Bundesländer Ende des Jahres 1994 konnten dann schließlich auch der wirtschaftliche Aufbauprozess und die ökonomische Entwicklung in Ostdeutschland nicht mehr in der gebotenen Breite analysiert werden.

Andere strukturelle Veränderungen, wie der Wandel von der Industrie- zur Dienstleistungsgesellschaft, technologieinduzierte Effizienzsteigerungen beispielsweise in der Lagerhaltung oder die zunehmende Bedeutung des internationalen Konjunkturzusammenhangs, laufen weniger sichtbar ab, wenngleich auch sie das Verlaufsmuster und die individuelle Ausprägung der Konjunkturzyklen beeinflussen. Empirisch sind diese Einflussfaktoren in Bezug auf die Prognosegüte beobachtungsäquivalent mit anderen Erklärungen einer veränderten Treffsicherheit – es kann also a priori nicht zwischen ihnen diskriminiert werden.

Dicke und Glismann ignorieren diese Sondereinflüsse in ihrer Arbeit gänzlich, obwohl die Gefahr besteht, dass der Strukturbruch der deutschen Vereinigung die empirischen Ergebnisse dominiert und bei der Interpretation des Vergleichs der Prognosegüte zwischen den neunziger Jahren und der Zeit davor nicht vernachlässigt werden sollte. Insofern stimmt bedenklich, dass die Schlussfolgerungen der beiden Autoren am Vergleich der Post-Vereinigungsperiode 1991-2000 mit der Zeit von 1968-1977 festgemacht werden. Die Auswahl der Teilzeiträume erscheint willkürlich, ist

<sup>4</sup> Jörg Hinze: Konjunkturprognosen unter veränderten Rahmenbedingungen – Auswirkungen der deutschen und europäischen Einigung, in: Hamburger Jahrbuch für Wirtschafts- und Gesellschaftspolitik, Bd. 41 (1996), S. 59-84.

**Schaubild 1**  
**BIP-Revisionen, 1965-2000**  
(in Prozentpunkten)



aber für die Ergebnisse von nicht unerheblicher Bedeutung.

**Zuverlässigkeit der statistischen Datenbasis**

In einem engen Zusammenhang mit strukturellen Veränderungen der Volkswirtschaft steht auch die Qualität des den Prognosen zugrunde liegenden statistischen Datenmaterials. Prognosen beruhen unter anderem auf den verfügbaren VGR-Daten aus den statistischen Ämtern, wobei es sich am aktuellen Rand um erste amtliche Ergebnisse handelt, die ihrerseits wiederum Schätzungen darstellen, die im Zeitverlauf noch mehrmals revidiert werden. Die Zuverlässigkeit dieser Information – gemessen am Ausmaß der nachfolgenden Revisionen – ist für sich genommen eine positive Funktion des zeitlichen Abstands zwischen Bezugszeitraum und Veröffentlichungszeitpunkt und der Qualitätssicherung in den statistischen Ämtern, sie hängt negativ ab vom Umfang des Strukturwandels in einer Ökonomie.

Schaubild 1 suggeriert, dass der Umfang der Revisionen über die Zeit abgenommen hat, allerdings liefert eine Regression der absoluten Revisionen auf eine Konstante und einen Zeittrend keinen Hinweis auf eine signifikante trendmäßige Verringerung des Revisionsumfangs. Der durchschnittliche absolute Fehler (MAE) der ersten Hälfte des Untersuchungszeitraums unterscheidet sich nicht signifikant von dem der zwei-

<sup>5</sup> Der maximal verfügbare Beobachtungszeitraum wurde in zwei ungefähr gleichgroße Teilstichproben von 1965-1982 und 1983-2001 mit 18 beziehungsweise 19 Beobachtungen unterteilt. Alternativ wurden, auch um die Robustheit der Ergebnisse in Bezug auf die Verteilung exogener Schocks in den Teilperioden zu untersuchen, alle Berechnungen mit Teilstichproben von 1965-1978 und 1979-2001 durchgeführt – die qualitativen Ergebnisse bleiben jedoch auch bei dieser Unterteilung des Beobachtungszeitraums bestehen.

**Tabelle 1**  
**Stichprobenstatistiken der Revisionen**

Zeitraum	Mittelwert <sup>a</sup>	MAE <sup>a</sup>	Autokorr. <sup>b</sup>
1965-2000	0,13 (0,77)	0,57 (0,52)	0,17 (1,04)
1965-1982	0,11 (0,88)	0,61 (0,63)	-0,09 (-0,25)
1983-2000	0,14 (0,67)	0,54 (0,40)	0,58 (2,86)

<sup>a</sup> Standardabweichung in Klammern. MAE = durchschnittliche absolute Revision; <sup>b</sup> Autokorr.: Koeffizient und t-Wert der um eine Periode verzögerten Revision in einer der Regression  $R=a+bR(-1) + e$ .

ten Hälfte, gleichwohl ist er etwas größer (Tabelle 1)<sup>5</sup>. Allerdings bestätigt der signifikante Autokorrelationskoeffizient der zweiten Teilperiode den visuellen Eindruck, dass die Revisionen in der zweiten Hälfte der Stichprobe nicht mehr seriell unabhängig sind. Die serielle Korrelation der Revisionen ist vermutlich auf die deutsche Vereinigung zurückzuführen. Während der Anfangszeit der gesamtdeutschen Statistik lagen für Ostdeutschland bis zum Frühjahr des Jahres 1991 keine Volkswirtschaftliche Gesamtrechnungen nach den Vorschriften des System of National Accounts (SNA) vor<sup>6</sup>. Außerdem stimmt dies mit der Beobachtung von Zarnowitz<sup>7</sup> für die Vereinigten Staaten überein, dass „[i]n times of strong shocks and surprising developments measurement of short term changes in the economy is particularly difficult [and] systematic [measurement] errors are frequent“.

#### Technischer und wissenschaftlicher Fortschritt

Technischer Fortschritt könnte beispielsweise durch erhöhte Rechenkapazitäten sowohl die „Prognoseproduktion“ wie auch den wissenschaftlichen Fortschritt positiv beeinflussen. Wissenschaftlicher Fortschritt könnte in Form zuverlässigerer Statistiken, deren Auswertung im Rahmen der Diagnose der aktuellen konjunkturellen Lage die Basis jeder Prognose bildet, präziseren Prognosemodellen und neuen empirischen Erkenntnissen in die Arbeit des Prognostikers einfließen. Prinzipiell ist aber davon auszugehen, dass die Rolle des technischen und wissenschaftlichen Fortschritts in der Prognoseproduktion relativ komplex und vermutlich auch nicht linear ist.

Die Palette der Prognosemethoden lässt sich in „geschlossene“ ökonomische Verfahren, bei denen die Einschätzung des Prognostikers in der Regel nur

in die Spezifikation des Modells einfließt, und Ansätze, bei denen die Einschätzung des Prognostikers eine bedeutendere Rolle spielt, da lediglich Einzelzusammenhänge ökonometrisch erfasst werden, einteilen. Zu letzterer Gruppe zählt der iterativ-analytische Ansatz, der heute den meisten Prognosen, also auch denen der wirtschaftswissenschaftlichen Forschungsinstitute und des Sachverständigenrates, zugrunde liegt. Ist der Zusammenhang zwischen verfügbarer Rechenleistung und Output makroökonomischer Modelle recht direkt, wirkt sich der technische Fortschritt auf die Treffsicherheit iterativ-analytischen Methoden weniger direkt aus. Auch dies wird in dem Artikel von Dicke und Glismann nicht diskutiert.

Empirisch lässt sich ein eventuell bestehender Einfluss des wissenschaftlichen und technischen Fortschritts auf die Arbeit des Prognostikers außerdem nur dann nachweisen, wenn

- (1) die unbeobachtbaren Variablen wissenschaftlicher und technischer Fortschritt ausreichend gut approximiert werden können. Es ist davon auszugehen, dass beispielsweise die Diffusion neuer Technologien keinem linearen Prozess folgt und daher ein Zeitrend, wie implizit von Dicke und Glismann unterstellt, eine nur unbefriedigende Proxy-Variablen für diesen Einflussfaktor darstellt.
- (2) sich technischer Fortschritt wesentlich in einer Verbesserung der Prognosequalität äußert. Technischer Fortschritt kann aber auch dazu führen, dass entweder die Outputmenge erhöht wird oder eine gegebene Outputmenge in kürzerer Zeit produziert wird beziehungsweise die zur Prognoseerstellung eingesetzten Mengen der Produktionsfaktoren reduziert werden.
- (3) technischer Fortschritt mit der Prognoseunsicherheit unkorreliert ist. Gerade die Erfahrungen der „New Economy“ haben aber gezeigt, dass technologischer Wandel auch mit erhöhter Unsicherheit über den Wachstumspfad einer Volkswirtschaft einhergeht.

In der Summe gibt es also sowohl Argumente für eine Verbesserung wie für eine Verschlechterung der Prognosegüte über die Zeit. Der wissenschaftliche und technische Fortschritt jedenfalls ist lediglich ein Bestimmungsfaktor unter vielen. Solange empirisch lediglich die Veränderung der Prognosegüte über die Zeit untersucht wird und die Einzelfaktoren nicht als Regressoren in eine Schätzgleichung eingehen, kann der Einfluss verschiedener Determinanten nicht getrennt werden. Die von Dicke und Glismann gestellte Frage nach dem Zusammenhang zwischen

<sup>6</sup> In Arbeiten, die die Prognosen am endgültigen Ergebnis der VGR messen, schlägt sich dieser Messfehler natürlich auch in einer seriellen Korrelation der Prognosefehler nieder und kann zu einer Ablehnung der weiter unten diskutierten Informationseffizienz führen.

<sup>7</sup> V. Zarnowitz: Information, Measurement, and Prediction in Economics, NBER Working Paper 318, 1979.

Konjunkturprognosen und wissenschaftlichem sowie technischem Fortschritt oder der Qualität der prognostischen Arbeit kann mit dem von ihnen gewählten methodischen Vorgehen daher nicht beantwortet werden.

### Maße der Prognosegüte

Zur Beurteilung der Qualität einer Prognose können die interdependenten Kriterien Unverzerrtheit und Effizienz herangezogen werden. Zunächst einmal sollte eine gute Prognose nicht systematisch falsch sein: Prognosefehler sind zulässig, aber im historischen Durchschnitt sollten die Prognosewerte den Realisationen entsprechen. Die Bedingung der Unverzerrtheit kann mittels der so genannten Mincer-Zarnowitz-Regression der Prognose,  $y^p$ , auf die entsprechende Realisation,  $y$ , überprüft werden<sup>8</sup>.

$$y_t = \alpha + \beta y_t^p + \varepsilon_t \quad (1)$$

und entspricht der Nullhypothese ( $\alpha=0 \wedge \beta=1$ )<sup>9</sup>. Eine zweite Bedingung ist, dass die Treffsicherheit der Prognosen optimal ist, das heißt, dass die Abweichungen der Prognose von der Realisation minimal sein sollen. Ziel des Prognostikers ist es, den erwarteten „Verlust“ oder die erwarteten „Kosten“ aus einer Verlustfunktion,  $V(\cdot)$ , die im Allgemeinen die Summe der absoluten oder quadrierten Prognosefehler,  $u_t = y_t^p - y_t$ , als Argument enthält, zu minimieren. Diese Verlustfunktion sollte folgende Bedingungen erfüllen:  $V(0)=0$ ,  $V(U_1) > V(U_2)$  für  $U_1 > U_2 > 0$ , wobei  $U_i$  für das Argument der Verlustfunktion steht, beispielsweise die Summe der quadrierten Prognosefehler.

Der Vergleich der Prognosegüte zweier Prognoseverfahren oder eines Prognoseverfahrens über die Zeit kann mithilfe der Verlustfunktion durchgeführt werden, wobei zu beachten ist, dass Maße, die keine monotone Funktion von  $S_N^2 = 1/N \sum_{t=1}^N u_t^2$  sind, zu inkonsistenten Ergebnissen führen können<sup>10</sup>. Somit sind der durchschnittliche absolute Prognosefehler, der durchschnittliche quadrierte Prognosefehler sowie die entsprechenden standardisierten Maße<sup>11</sup> grundsätzlich geeignet, der „relative Prognosefehler“ in der Operationalisierung von Dicke und Glismann jedoch nicht.

Dies lässt sich wie folgt zeigen: Es seien  $\{u_t\}$  und  $\{v_t\}$  die Prognosefehler zweier unterschiedlicher Prognosen mit dem Erwartungswert Null und den Varianzen  $\sigma_a^2$  und  $\sigma_b^2$ , wobei  $\sigma_a^2 < \sigma_b^2$ . Mit anderen Worten  $b$  ist die „schlechtere“, weniger treffsichere Prognose.

Der „relative Prognosefehler“ in der Definition der beiden Autoren lässt sich vereinfacht schreiben als  $e_t = |u_t| / (\hat{\sigma}t)$ , wobei  $\hat{\sigma}t$  für den geschätzten Trend der zu prognostizierenden Größe, hier des Bruttoinlandsprodukts, steht<sup>12</sup>. Das korrespondierende Maß der Prognosegüte ist dann  $V_{DG} = 1/(\hat{\sigma} \cdot N) \sum_{t=1}^N |u_t|/t$ . Strebt  $N$  gegen unendlich, gilt  $\lim_{N \rightarrow \infty} V_{DG} \rightarrow 0$  und zwar unabhängig von der Varianz der Prognosefehler, also der Treffsicherheit der Prognose. Das von Dicke und Glismann verwendete Maß vermag also nicht zwischen guten und schlechten Prognosen zu unterscheiden. Zur Berücksichtigung einer sich potentiell verändernden prognostischen Unsicherheit sollte deshalb im gegebenen Kontext auf standardisierte Maße und nicht auf den „relativen Prognosefehler“ in der Operationalisierung von Dicke und Glismann rekuriert werden.

Die Untauglichkeit der gewählten Methode zeigt sich aber auch an ihrer mangelnden Robustheit in Bezug auf die Wahl des Beobachtungszeitraums. Wird ein anderer Startpunkt des Untersuchungszeitraums gewählt, nämlich beispielsweise 1975, 1976, 1977, 1978 oder 1979, dann lässt sich kein signifikanter Zeittrend in den Veränderungsraten des Bruttoinlandsprodukts nachweisen. Der relative Prognosefehler entspricht dann dem durch eine Konstante dividierten Prognosefehler im Falle durchschnittlich positiver Zuwachsraten oder er ist unendlich im Falle einer stagnierenden Ökonomie. Eine systematische Veränderung der Prognosequalität lässt sich dann auch empirisch nicht mehr zeigen: Für all diese Untersuchungszeiträume ist der Zeittrend des relativen Prognosefehlers entweder insignifikant oder nicht berechenbar. Gilt die Hypothese der beiden Autoren etwa nicht für andere als den gewählten Zeitraum?

Eine weitere Anforderung an eine Prognose ist die Bedingung der Informationseffizienz, die bedeutet, dass eine Prognose nicht durch zum Prognosezeitpunkt verfügbare Informationen verbessert werden kann, dass also alle relevanten Daten in sie eingegangen sind. Formal lässt sich die Bedingung der Informationseffizienz darstellen als

<sup>12</sup> Genau genommen müsste durch eine Konstante und den Zeittrends dividiert werden. Zur Vereinfachung der Darstellung wird hierauf jedoch im Folgenden verzichtet. Es sollte darauf hingewiesen werden, dass sich die Definition des relativen Prognosefehlers durch Dicke und Glismann von der üblichen Definition, die den Prognosefehler auf die tatsächliche Änderungsrate des Bruttoinlandsprodukts bezieht, unterscheidet.

<sup>8</sup> Siehe M. Clements, D. Hendry: *Forecasting Economic Time Series*, Cambridge University Press, 1998.

<sup>9</sup> Diese Nullhypothese entspricht der 45-Grad-Linie in einem Prognose-Realisationsdiagramm.

<sup>10</sup> Vgl. C. Granger, P. Newbold: *Forecasting Economic Time Series*, Academic Press, San Diego, 1986, S. 279.

<sup>11</sup> Die standardisierten Maße besitzen den Vorteil, dass die mit zunehmender Variabilität der zu prognostizierenden Zeitreihe steigende prognostische Schwierigkeit berücksichtigt wird.

$E\{V(S_t^2)|I_{t-j}\} = Min.$ , wobei  $I_{t-j}$  für die verfügbare Informationsmenge zum Zeitpunkt der Prognose,  $t-j$ , steht. Da sich die verfügbare Informationsmenge in empirischen Anwendungen jedoch nur schwer abbilden lässt, wird vielfach auf die schwache Form der Informationseffizienz zurückgegriffen, nach der eine Prognose informationseffizient ist, wenn  $E\{V(S_t^2)|L_{t-j}\} = Min.$ , wobei  $L_{t-j}$  für die vorangegangenen eigenen Prognosen steht<sup>13</sup>. Eine mögliche empirische Prüfung auf schwache Informationseffizienz besteht daher in dem Test der Nullhypothese  $\beta=0$  in einer Regression der Prognoserevisionen auf ihre Vergangenheit,

$$(y_{t,t-1}^p - y_{t,t-2}^p) = \alpha + \beta (y_{t,t-1}^p - y_{t,t-2}^p) + \varepsilon_t \quad (2)$$

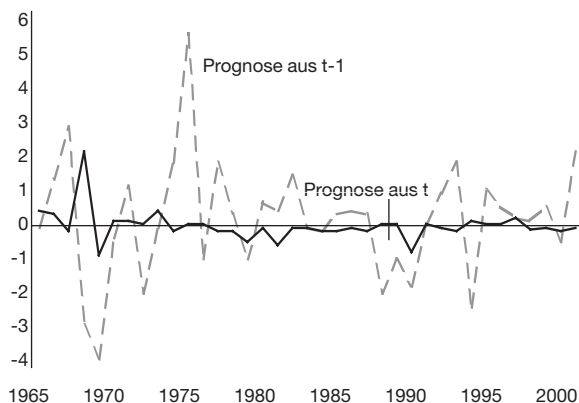
wobei  $y_{t,t-1}^p$  für die zum Zeitpunkt  $t-1$  erstellte Prognose von  $y_t$  steht<sup>14</sup>.

Vor diesem Hintergrund werden in dieser Arbeit zur Beurteilung der Prognosegüte folgende Maße herangezogen: (1) Unverzerrtheit (UV), (2) Schwache Informationseffizienz (SI) und (3) zwei Verlustfunktionen, nämlich der durchschnittliche absolute Fehler (MAE) sowie der durchschnittliche quadrierte Fehler (MSE), jeweils auch in ihrer standardisierten Form (SMAE und SMSE).

### Empirische Umsetzung

Bei der empirischen Umsetzung ist zunächst zu entscheiden, an welchen Realisationen die Prognosen gemessen werden. Von den Daten aus den Volkswirtschaftlichen Gesamtrechnungen, die im Zeitverlauf revidiert werden, können grundsätzlich das erste amtliche Ergebnis oder das aktuelle amtliche Ergebnis, das für weiter zurückliegende Datenpunkte gleichbedeutend mit dem amtlichen Endergebnis ist und alle zwischenzeitlich vorgenommenen Revisionen enthält, verwendet werden. In der Literatur finden sich zwar prinzipiell beide Vorgehensweisen, allerdings besteht Konsens, dass das erste amtliche Ergebnis besser den zum Zeitpunkt der Prognoseerstellung verfügbaren Informationsstand widerspiegelt<sup>15</sup>. Es versteht sich von selbst, dass die Konjunkturprognostiker mit dem offiziellen Datenkranz arbeiten und zukünftige, unbe-

**Schaubild 2**  
**Prognosefehler des Sachverständigenrates,**  
**1965-2001**  
 (in Prozentpunkten, Prognose-Realisation)



kannte Datenrevisionen damit nicht in die Prognose einfließen können.

Vergleiche mit den endgültigen Daten,  $y^f$ , führen in der Regel zu größeren Fehlern, die sich aus dem Prognosefehler im Vergleich zum ersten amtlichen Ergebnis und den nachfolgenden Revisionen zusammensetzen:  $y_t^p - y^f = (y_t^p - y_t) + (y_t - y^f) = \eta_t + \varepsilon_t$ , wobei  $\eta_t$  und  $\varepsilon_t$  positiv miteinander korreliert sind. Außerdem kann die Verwendung des aktuellen amtlichen Ergebnisses insofern zu verzerrten Ergebnissen führen, als für ältere Datenpunkte bereits das amtliche Endergebnis vorliegt, während für jüngere Beobachtungen mit vorläufigen Daten gearbeitet werden muss, die in der Zukunft weiteren Revisionen unterliegen. Der Prognosefehler der jüngsten drei bis fünf Jahre kann somit im historischen Vergleich unterschätzt werden. Diesen Überlegungen und der Mehrzahl der empirischen Studien folgend wird hier deshalb das erste amtliche Ergebnis der Volkswirtschaftlichen Gesamtrechnungen als Referenzreihe zur Messung des Prognosefehlers herangezogen<sup>16</sup>. Die Studie von Dicke und Glismann verwendet endgültige VGR-Daten, die nicht dem Informationsstand der jeweiligen Prognose, sondern dem Zeitpunkt ihrer eigenen Untersuchung entsprechen.

<sup>13</sup> Siehe hierzu W. Nordhaus: Forecasting Efficiency, in: The Review of Economics and Statistics (4), 1987, S. 667-674. Allerdings fassen einige Autoren den Begriff der schwachen Informationseffizienz auch etwas weiter und beziehen in die relevante Informationsmenge nicht nur die eigenen Prognosen, sondern auch solche anderer Institutionen ein.

<sup>14</sup> Wie bereits beschrieben prognostiziert der Sachverständigenrat jeweils im November die VGR-Größen des laufenden Jahres und des kommenden Jahres. Zu diesem Zeitpunkt liegen Ergebnisse der VGR bis einschließlich dem zweiten Quartal des laufenden Jahres vor.

<sup>15</sup> Vgl. V. Zarnowitz, P. Braun: Twenty-Two Years of the NBER-ASA Quarterly Economic Outlook Surveys: Aspects and Comparisons of Forecasting Performance, NBER Working Paper 3965, 1992.

<sup>16</sup> Beispiele für ein ähnliches Vorgehen sind J. Hinze, a.a.O.; H. Rinne: Bemerkungen zur Qualität der Konjunkturprognosen des Sachverständigenrates zur gesamtwirtschaftlichen Entwicklung, in: P. M. von der Lippe (Hrsg.), Wirtschafts- und Sozialstatistik heute: Theorie und Praxis, Festschrift für Walter Krug, 1997, S. 319-340; oder J. Döpke: Haben Konjunkturprognosen einen politischen Bias, in: Schmollers Jahrbuch (120), 2000, S. 587-620.

<sup>17</sup> Zusätzlich wurden in der vorliegenden Studie auch die Prognosefehler bei Verwendung der endgültigen bzw. aktuellen Ergebnisse der amtlichen Statistik berechnet. Wie erwartet erhöht sich der absolute Prognosefehler, die im Folgenden abgeleiteten qualitativen Ergebnisse sind aber robust bezüglich der Wahl der Referenzdatenreihe.

**Tabelle 2**  
**Güte der BIP-Prognosen des**  
**Sachverständigenrates**

Zeitraum	UV <sup>a</sup>	MAE	SMAE <sup>b</sup>	MSE	SMS <sup>b</sup>	IE <sup>c</sup>
Schätzung (laufendes Jahr)						
1965-2001	3,18 (0,05)	0,23	0,10	0,15	0,07	/
1965-1982	2,20 (0,14)	0,36	0,12	0,25	0,08	/
1983-2001	0,10 (0,91)	0,12	0,09	0,02	0,01	/
Prognose (nächstes Jahr)						
1965-2001	0,76 (0,47)	1,22	0,53	1,37	0,60	0,27 (0,76)
1965-1982	0,15 (0,34)	1,59	0,54	1,98	0,67	0,12 (0,89)
1983-2001	0,20 (0,82)	0,87	0,62	0,55	0,39	0,28 (0,76)

<sup>a</sup> F-Test und Signifikanzniveau der Nullhypothese ( $\alpha=0$ ,  $\beta=1$ ) aus Gleichung (1). <sup>b</sup> Standardisiert mit der Standardabweichung des BIP in der entsprechenden Stichprobe. <sup>c</sup> F-Test und Signifikanzniveau der Nullhypothese ( $\alpha=\beta=0$ ) in der Regression der Prognoserevision auf ihre eigene Vergangenheit in Gleichung (2).

Die Prognosefehler des Sachverständigenrates in Bezug auf die Veränderungsrate des realen Bruttoinlandsprodukts für das laufende und das jeweils nächste Jahr sind in Schaubild 2 dargestellt<sup>17,18</sup>.

Es zeigt sich, dass – wie erwartet – der absolute Prognosefehler mit zunehmendem Prognosehorizont, also zunehmender Unsicherheit, zunimmt und dass besonders große Prognosefehler in zeitlicher Nähe zur stürmischen Hochkonjunktur Ende der sechziger Jahre und den „Sonderereignissen“ erste Ölkrise und deutsche Vereinigung sowie im letzten Jahr auftraten.

Die verschiedenen Kriterien zur Messung der Prognosequalität sind in Tabelle 2 wiedergegeben und belegen, dass:

(1) die wirtschaftliche Entwicklung vom Sachverständigenrat im Mittel richtig eingeschätzt wurde. Die Prognosen des Sachverständigenrates sind erwartungstreu, auch wenn einzelne Teilzeiträume betrachtet werden. Dies gilt jedoch nur eingeschränkt für die Schätzung des laufenden Jahres: Hier ist die Steigung im Realisations-Prognose-Diagramm mit

einem Wert von 0,93 zwar recht nahe an, jedoch statistisch signifikant unterschiedlich von eins. Die empirische Verwerfung der Unverzerrtheit lässt sich auch auf das hohe korrigierte Bestimmtheitsmaß  $R^2$  in Gleichung (1) zurückführen; es beträgt 0,97. Anhaltspunkte für eine systematisch zu optimistische Sichtweise des Sachverständigenrates während der Neunziger Jahre lassen sich im Übrigen auch nicht aus den in Tabelle 2 von Dicke und Glismann wiedergegeben Mittelwerten ableiten. Entscheidend ist, ob der Mittelwert statistisch signifikant von Null differiert und ob der Unterschied zwischen den Teilstichproben statistisch signifikant ist. Eigene Berechnungen legen nahe, dass dies nicht der Fall ist. Es ist befremdlich, dass die beiden Autoren ihre Behauptung nicht durch empirische Tests belegen.

(2) die absoluten Prognosefehler über die Zeit nicht zugenommen haben, sondern tendenziell kleiner geworden sind. Eine Regression des absoluten Prognosefehlers auf einen linearen Zeittrend bestätigt diesen Eindruck. Der Koeffizient des Zeittrends ist bei 10-prozentiger Irrtumswahrscheinlichkeit nicht wie bei Dicke und Glismann positiv, sondern negativ. Um abzuschätzen, ob dies auf herausragende exogene Schocks und Strukturbrüche zurückzuführen ist, wurde die Gleichung zusätzlich mit Dummy-Variablen für die erste Ölkrise und die deutsche Vereinigung geschätzt. Auch in diesen Varianten<sup>19</sup> kann keine Zunahme des absoluten Prognosefehlers im Zeitverlauf festgestellt werden, das Signifikanzniveau des negativen Trends nimmt teilweise sogar zu. Für das Maß des mittleren quadrierten Prognosefehlers zeigt sich sogar eine ausgeprägtere Verbesserung der Treffsicherheit in der zweiten Teilstichprobe, was sich aus der höheren Gewichtung der großen Prognosefehler Ende der sechziger Jahre und Mitte der siebziger Jahre ergibt.

(3) eine Berücksichtigung der in der zweiten Hälfte des Beobachtungszeitraums geringeren prognostischen Unsicherheit (SMAE und SMSE) eine ungefähr gleich große Treffsicherheit in den beiden Teilstichproben nahe legt. Mit anderen Worten: Die, anhand des mittleren absoluten und des mittleren quadratischen Prognosefehlers gemessene, verbesserte Prognose-

<sup>18</sup> Bis zum Jahresgutachten 1994/95 wurde die Prognose auf ½-Prozentpunkte, anschließend bis zum Jahresgutachten 1998/99 auf ¼-Prozentpunkte und seitdem auf Zehntelprozentpunkte genau ausgewiesen. Für das Jahr 1980 wurde ein Band von 2,5-3,0% angegeben. Hier wird dessen ungeachtet, wie auch in anderen Studien, zur Berechnung des Prognosefehlers jeweils der Prognosewert mit einer Nachkommastelle verwendet, beziehungsweise für das Jahr 1980 die Mitte des Prognosebandes. Durch die großzügigere Rundung in früheren Jahren dürfte dies zum Ausweis tendenziell größerer Prognosefehler bis Mitte der Neunziger Jahre führen.

<sup>19</sup> Die genaue zeitliche Abgrenzung der „exogenen Schocks“ und Strukturbrüche ist schwierig und sicherlich auch teilweise subjektiv. Aus diesem Grund wurden insgesamt drei Varianten überprüft: Erstens, eine Spezifikation, in der die Dummy-Variable im Jahr 1975 und 1990 den Wert eins annimmt, zweitens eine Spezifikation, in der eine Dummy-Variable in den Jahren 1975 und 1979 den Wert eins annimmt sowie eine weitere im Jahr 1990 den Wert gleich eins ist, und drittens eine Spezifikation, bei der der Ölkrisedummy der zweiten Variante entspricht, die Wiedervereinigungsdummy aber während der gesamten ersten Hälfte der neunziger Jahre den Wert eins annimmt.

güte in der zweiten Teilstichprobe steht in einem engen Zusammenhang zur geringeren durchschnittlichen prognostischen Unsicherheit in diesem Zeitraum.

(4) die Prognosen des Sachverständigenrates schwach informationseffizient sind.

### Fazit

Die Hypothese einer systematischen Veränderung der Treffsicherheit von BIP-Prognosen des Sachverständigenrates kann nicht bestätigt werden. Diese Ergebnisse sind konsistent mit denen anderer Studien<sup>20</sup>, die den Prognosen des Sachverständigenrates – auch im Vergleich zu den wirtschaftswissenschaftlichen Forschungsinstituten – eine hohe Treffgenauigkeit attestieren<sup>21</sup>. Aufgrund der Vielzahl von Einzelfaktoren, die die Prognosegüte über die Zeit beeinflussten, ist eine systematische Veränderung auch nicht zwangsläufig zu erwarten. Umgekehrt lassen sich aus einer unveränderten Prognosegüte alleine natürlich auch keine Rückschlüsse auf die Wirkung einzelner Einflussfak-

toren, beispielsweise die Qualität der prognostischen Arbeit, ziehen, wie dies Dicke und Glismann tun.

Die Ergebnisse von Dicke und Glismann erklären sich vor allem durch die Verwendung einer speziellen Operationalisierung des relativen Prognosefehlers, die als Maß der Prognosegüte nicht geeignet ist. Zudem werden die von den Autoren aufgestellten Behauptungen durch die Tabellen im Text lediglich illustriert, sie werden aber nicht formal getestet. Die Mittelwerte der Prognosefehler, auf die die Autoren ihr Fazit stützen, sind jedoch statistisch nicht signifikant, und angesichts der Schwere der Vorwürfe gegen den Sachverständigenrat erstaunt die Nachlässigkeit der empirischen Beweisführung, die einer näheren Überprüfung nicht standhält.

Letztlich kann die Fragestellung von Dicke und Glismann nach dem Zusammenhang zwischen Prognosegüte und wissenschaftlichem und technischem Fortschritt mit der von den Autoren gewählten Methode, die sich lediglich auf die Beschreibung der zeitlichen Entwicklung des Prognosefehlers stützt, nicht beantwortet werden, und so handelt es sich auch bei den Schlussfolgerungen in Bezug auf die Arbeit des Rates um Spekulationen, die jeglicher empirischen Basis entbehren.

---

<sup>20</sup> Vgl. beispielsweise J. Hinze, a.a.O.; und H. Rinne, a.a.O.

<sup>21</sup> Zur Treffsicherheit der Prognosen verschiedener Institutionen im Vergleich siehe z. B. auch D. Radowski: Wie treffsicher sind Prognosen? in: ZEW Konjunkturreport, März 2002, S. 10-12.