

# Wochenbericht

Königin-Luise-Straße 5  
14195 Berlin

Tel. +49-30-897 89-0  
Fax +49-30-897 89-200

www.diw.de  
postmaster@diw.de

**DIW** Berlin

Wirtschaft Politik Wissenschaft

## Warum Konjunkturprognosen?

Ulrich Fritsche  
ufritsche@diw.de

*Konjunkturprognostiker stehen im Kreuzfeuer der öffentlichen Kritik. Im vorliegenden Bericht wird folgenden Fragen nachgegangen: Was können Konjunkturprognosen leisten? Welchen Kriterien haben sie zu genügen? Wie sind die Leistungen von – zum Teil öffentlich geförderten – Institutionen, die Konjunkturprognosen liefern, zu beurteilen? Haben sich Konjunkturprognosen bewährt? Was sind die Gründe für Unterschiede bei der Prognosegüte?*

*Insgesamt zeigt sich, dass die untersuchten Prognosen im Großen und Ganzen den Kriterien der Unverzerrtheit und Rationalität genügen. Bezüglich Beschleunigungen und Verlangsamungen schneiden sie auf jeden Fall deutlich besser als ein Münzwurf ab. Allerdings sind die Fehlermargen relativ hoch und nehmen mit der Länge des Prognosezeitraums zu – Konjunkturprognosen sind unsicher. Die mit der Prognose verbundene Unsicherheit sollte angegeben werden. Prognoseinstitutionen sollten zudem ihre Prognoseinstrumente wissenschaftlich nachvollziehbar machen und die Anwendungsbedingungen ihrer Aussagen offen legen.*

### Konjunkturprognosen in der Kritik

Die Zeiten sind für Konjunkturforscher rau geworden. Seit geraumer Zeit nimmt die Kritik an Konjunkturprognosen – eine Domäne der öffentlich geförderten Wirtschaftsforschungsinstitute – zu. Vom „Blindflug der Forscher“ ist die Rede,<sup>1</sup> einzelne Journalisten fordern sogleich „Entlaßt die Experten“,<sup>2</sup> und selbst der Bundeskanzler klagt über „diese Art der Wissenschaft, die ihn an Meteorologie erinnere“.<sup>3</sup>

Solche Prognostikerschelte ist nicht neu<sup>4</sup> und auch nicht die – allzu verständliche – Verteidigungshaltung der Betroffenen, die auf die Bedingtheit ihrer Ergebnisse hinzuweisen nicht müde werden.<sup>5</sup> In der Öffentlichkeit entsteht daraus leicht der Eindruck, „getretene Hunde“ würden bellen. Umso wichtiger ist eine wissenschaftliche Auseinandersetzung mit Möglichkeiten und Grenzen von Konjunkturprognosen. Aus diesem Grund wurde hier ein Paneldatensatz von Prognosen für Deutschland, die von wichtigen Institutionen erstellt werden, untersucht. Für eine Bewertung ist es allerdings unumgänglich zu erörtern, was rationale Prognosen ausmacht und welche grundsätzlichen Probleme bei der Beurteilung von Prognosen bestehen.

<sup>1</sup> Vgl.: Blindflug der Forscher. In: Der Spiegel vom 2. Mai 2005, S. 94.

<sup>2</sup> Vgl.: Entlaßt die Experten. In: FAZ vom 19. April 2005, S. 33.

<sup>3</sup> Vgl.: Deutlich besser als ein Münzwurf. In: FAZ vom 6. Mai 2005, S. 12.

<sup>4</sup> Vgl. die zitierten Beispiele in M. Tietzel: Prognoselogik oder: Warum Prognostiker irren dürfen. In: Jahrbücher für Nationalökonomie und Statistik, Bd. 206, Nr. 6, 1989, S. 546–562.

<sup>5</sup> Vgl. J. Hinze: Konjunkturprognosen: Was sie leisten können und was nicht. In: Hamburger Jahrbuch für Wirtschafts- und Gesellschaftspolitik, Jg. 32, 1987, S. 37–47; E. Helmstädter: Die Ratgeber sind nicht ratlos. In: Die Zeit vom 18. September 1988, S. 26.

**Nr. 22/2005**

72. Jahrgang/1. Juni 2005

### 2. Bericht

**Warum Konjunkturprognosen?**  
Seite **361**

## Rationale Prognosen oder Kaffeesatzleserei?

Eine Vorhersage ist eine Aussage, die heute geäußert wird und das Eintreten eines bestimmten Ereignisses in der Zukunft beschreibt. Diese Aussage kann eintreten oder nicht, unabhängig davon, auf welchem Prognosemodell sie beruht. Was alle Wirtschaftsforschungsinstitute und seriöse Prognostiker von Propheten und deren Prophezeiungen unterscheidet, ist die wissenschaftliche Begründung der Prognose.<sup>6</sup> Eine Aussage über das künftige Auftreten eines Ereignisses wird deduktiv aus mindestens einem allgemeinen Anwendungsgesetz und mindestens einer singulären Anwendungsbedingung abgeleitet.

Beispielsweise seien die Auswirkungen einer Ölpreiserhöhung auf Wachstum und Beschäftigung betrachtet: So zeigt die Erfahrung, dass eine 10-prozentige Ölpreiserhöhung im Schnitt und über zwei Jahre gerechnet das Wachstum in Industrieländern um 0,2 bis 0,25 Prozentpunkte dämpft.<sup>7</sup> Unter der konkreten Anwendungsbedingung eines 20-prozentigen Ölpreisanstiegs wäre also eine konjunkturelle Abschwächung um 0,4 bis 0,5 Prozentpunkte über zwei Jahre zu erwarten.

Nur Gesetz und Anwendungsbedingung(en) zusammen sind also hinreichende Bedingung für das Prognoseargument. Vorhersagen, die nicht auf einem vollständigen oder sogar auf einem logisch defekten Prognoseargument beruhen, sind nicht wissenschaftlich oder rational, da sie sich der Überprüfbarkeit entziehen.

Im Allgemeinen erfüllen die meisten veröffentlichten Konjunkturprognosen diese Anforderungen.<sup>8</sup> Sie basieren auf einem ökonomisch fundierten Modell als dem allgemein gültigen Gesetz und enthalten mit der Analyse der Ist-Situation und der Darstellung der Annahmen der Prognose die konkreten Anwendungsbedingungen.

Allerdings bestehen selbst bei hinreichend rational begründeten und vollständigen Prognosen theoretische Probleme:

- Das erste Problem bezieht sich auf den Geltungsbereich der der Prognose zugrunde liegenden allgemeinen Aussage. Ökonometrische Modelle und auch das Erfahrungswissen der Prognostiker sind vergangenheitsbezogen; die daraus abgeleiteten „Gesetzmäßigkeiten“ können streng genommen nur für die Vergangenheit Gültigkeit beanspruchen.<sup>9</sup> Denn die Struktur der ökonomischen Beziehungen kann sich verändern. Durch entsprechende Tests auf Strukturbrüche versucht man, dem Problem zu bege-

nen. Es ist jedoch sehr schwer, solche Veränderungen gerade am Ende des zurückliegenden Zeitraums zu entdecken. Sie für die Zukunft vorherzusagen, ist naturgemäß unmöglich.

- Das zweite Problem bezieht sich auf die Anwendungsbedingungen der Prognosen. Alle Konjunkturprognosen benutzen in irgendeiner Form Modelle, die die Komplexität reduzieren. Diese enthalten Anwendungsbedingungen oder exogen gesetzte Annahmen, die erfüllt sein müssen. Im genannten Beispiel ist es die Annahme über die Entwicklung des Ölpreises. Mit großen ökonometrischen Modellen kann man zwar immer mehr Annahmen „endogenisieren“; letztlich löst man damit aber das Problem nicht, sondern verschiebt es nur auf eine andere Ebene.
- Das dritte Problem bezieht sich darauf, dass Erwartungen und Prognosen selbst Bestandteil des zu analysierenden ökonomischen Systems sind.<sup>10</sup> Dieses Phänomen kann zu sich selbst erfüllenden Prognosen führen: Die Vorhersage einer Währungskrise löst eine Spekulationswelle aus, die die Währung ins Wanken bringt, oder die Prognose der Liquiditätskrise einer Geschäftsbank führt zum Abzug der Spareinlagen und damit genau zu einer Liquiditätskrise usw.

Was folgt daraus? Generell ist Zukunft nicht vorhersehbar. Eine Prognose ist deshalb noch nicht zwecklos. Die Abschätzung der Auswirkungen ökonomisch relevanter Entscheidungen kann – bei aller Beschränktheit – nur auf Basis fundierter Prognosen erfolgen. Solche Vorhersagen müssen allerdings wissenschaftlichen Kriterien genügen.

## Zur Beurteilung der Prognosen

Im Folgenden werden die Prognosen zahlreicher nationaler und internationaler Institutionen für das Wirtschaftswachstum und den Preisanstieg in Deutschland untersucht.<sup>11</sup> Die verwendeten Daten und Quellen sowie Details zu den verwendeten Prüfmaßen und Tests werden im Kasten beschrieben.

<sup>6</sup> Vgl. M. Tietzel, a. a. O., S. 548 f.

<sup>7</sup> Vgl. für einen Überblick D. Jones, P. Leiby und I. Paik: Oil Price Shocks and the Macroeconomy: What Has Been Learned Since 1996. In: Energy Journal, Vol. 25, No. 2, 2004, S. 1–32.

<sup>8</sup> Problematisch sind aus theoretischer Perspektive allerdings die häufig benutzten Trendextrapolationsmethoden. Solche Verfahren sind – wenn überhaupt – aus erkenntnistheoretischer Sicht nur auf streng geschlossene Systeme anwendbar, nicht jedoch auf offene und sich verändernde Systeme.

<sup>9</sup> Das Einstein-Jahr 2005 gemahnt daran, dass Denkgebäude ganzer Wissenschaftszweige in relativ kurzer Zeit umgeworfen werden können.

<sup>10</sup> Das Phänomen ist von Oskar Morgenstern und Karl Popper ausführlich diskutiert worden. Vgl. M. Tietzel, a. a. O., S. 554 ff.

<sup>11</sup> Vgl. U. Fritsche und J. Döpke: Growth and Inflation Forecasts in Germany – An Assessment of Accuracy and Dispersion. Diskussionspapier des DIW Berlin Nr. 399. Berlin 2004 ([www.diw.de/deutsch/produkte/publikationen/diskussionspapiere/docs/papers/dp399.pdf](http://www.diw.de/deutsch/produkte/publikationen/diskussionspapiere/docs/papers/dp399.pdf)).

Kasten

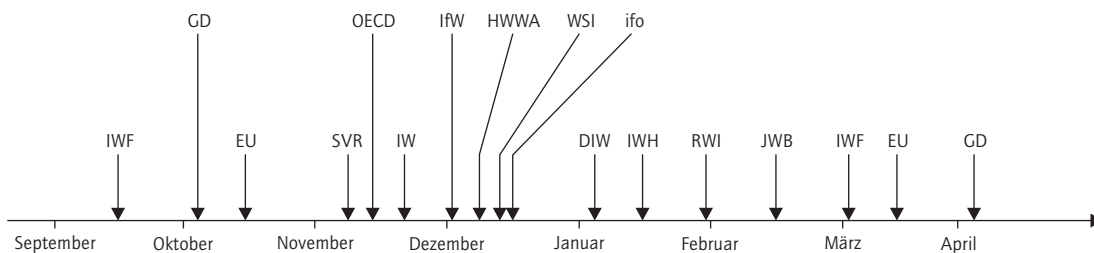
### Methoden der Prognoseevaluierung

Die untersuchten Prognosen sind alle im Winterhalbjahr für das kommende oder – wenn die Prognosen nach dem Dezember veröffentlicht sind – für das laufende Kalenderjahr erstellt.

Die Abbildung gibt auf einem Zeitstrahl die ungefähren – und historisch teilweise etwas variierenden – Veröffentlichungszeitpunkte der hier untersuchten Prognosen an: Das Panel beginnt mit der Prognose des Währungsfonds im Herbst und endet im April mit der Gemeinschaftsdiagnose vom Frühjahr. Die Quellen sind in Tabelle K1 angegeben.

Abbildung

### Institutionen und Prognosezeitpunkte



Legende:

IWF: Internationaler Währungsfonds; GD: Gemeinschaftsdiagnose; EU: Europäische Kommission; SVR: Sachverständigenrat; IW: Institut der deutschen Wirtschaft; OECD: Organisation für wirtschaftliche Zusammenarbeit und Entwicklung, Paris; IFW: Institut für Weltwirtschaft, Kiel; HWWA: Hamburgisches Weltwirtschaftsarchiv, Hamburg; WSI: Wirtschafts- und Sozialwissenschaftliches Institut bei der Hans-Böckler-Stiftung; ifo: ifo Institut, München; DIW: Deutsches Institut für Wirtschaftsforschung, Berlin; IWH: Institut für Wirtschaftsforschung, Halle; RWI: Rheinisch-Westfälisches Institut für Wirtschaftsforschung, Essen; JWB: Jahreswirtschaftsbericht der Bundesregierung.

DIW Berlin 2005

Die Wachstumsprognosen beziehen sich auf das reale Bruttosozialprodukt für den Zeitraum 1983 bis 1989 und das reale Bruttoinlandsprodukt für alle anderen Jahre. Im Falle von Intervallprognosen wurde der Mittelwert verwendet. Bis 1992 beziehen sich die Größen auf Westdeutschland, danach auf Gesamtdeutschland. Für die Inflationsprognosen wurde der Deflator des privaten Verbrauchs herangezogen. Als Vergleichsreihe für die Realisierungen wurden die Erstveröffentlichungen der entsprechenden Reihen durch das Statistische Bundesamt benutzt.

Für die Untersuchung wurde eine Reihe von Prüfmaßen berechnet.

- Der mittlere Fehler  $ME = \frac{1}{T} \sum_{t=1}^T e_{t+1,t}$ , mit  $e_{t+1,t} = y_{t+1} - \hat{y}_{t+1,t}$  als Prognosefehler jeder Periode, definiert als tatsächlicher Wert (in  $t + 1$ ) minus dem prognostizierten Wert (in Periode  $t$  für Periode  $t + 1$ ). Ein positiver (negativer) Wert des mittleren Fehlers bedeutet Unter-(Über)schätzungen der Variablen.
- Der mittlere absolute Fehler  $MAE = \frac{1}{T} \sum_{t=1}^T |e_{t+1,t}|$ .
- Die Wurzel des mittleren quadratischen Fehlers  $RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T e_{t+1,t}^2}$ .

### Text auf Unverzerrtheit der Prognose

Hier bietet sich die sogenannte Mincer-Zarnowitz-Regression an, das heißt, es wird die Gleichung  $y_{t+1} = \beta_0 + \beta_1 \hat{y}_{t+1,t} + u_t$  mit  $y_{t+1}$  als realisiertem Wert und  $\hat{y}_{t+1,t}$  als dem entsprechenden Prognosewert in Periode  $t$  für die Periode  $t + 1$  geschätzt.<sup>1</sup> Die Hypothese der Unverzerrtheit ist gleichbedeutend mit einer Konstanten von null und einem Koeffizienten bezüglich der Prognose von 1. Die zu testende Nullhypothese lautet  $H_0: \beta_0 = 0 \wedge \beta_1 = 1$ . Die entsprechende Zeile in den Tabellen 1 und 2 enthält die Irrtumswahrscheinlichkeit der Ablehnung der Nullhypothese.

<sup>1</sup> Vgl. F. X. Diebold: Elements of Forecasting. Cincinnati 2001, Kap. 11.

Tabelle K1

**Übersicht über die Quellen der verwendeten Prognosen**

Institution	Quelle	Zeitraum	Durchschnittlicher Prognosehorizont in Monaten	Bemerkungen
Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung, Wiesbaden	Jahresgutachten, diverse Ausgaben	1970–2004	13	
Arbeitsgemeinschaft wirtschaftswissenschaftlicher Forschungsinstitute	Gemeinschaftsdiagnose Frühjahr, diverse Ausgaben	1970–2004	8	
	Gemeinschaftsdiagnose Herbst, diverse Ausgaben	1970–2004	14	
DIW Berlin	„Wochenbericht des DIW Berlin“, diverse Ausgaben	1970–2004	11	Zwei Prognosen pro Jahr, hier nur Januarprognose
Institut für Weltwirtschaft, Kiel	„Die Weltwirtschaft“, diverse Ausgaben	1970–2004	12	Keine Prognosen 1975 und 1976. Hier wurde die Prognose der GD benutzt, da das Kieler Institut daran beteiligt war und dieser Prognose zustimmte. Für Prognosen generell die Dezemberprognose
HWWA, Hamburg	„Wirtschaftsdienst“, diverse Ausgaben	1970–2004	12	Zwei Prognosen pro Jahr, hier nur Herbstprognose
ifo Institut, München	„ifo Schnelldienst“, diverse Ausgaben	1970–2004	12	
Institut für Wirtschaftsforschung, Halle	„Wirtschaft im Wandel“, diverse Ausgaben	1993–2004	11	Institut im Jahre 1992 gegründet
Rheinisch-Westfälisches Institut für Wirtschaftsforschung, Essen	„Economic Report“, diverse Ausgaben	1970–2004	11	
OECD, Paris	„Economic Outlook“, diverse Ausgaben	1970–2004	13	Zwei Prognosen pro Jahr, hier nur Herbstprognose
IWF, Washington, D.C.	„World Economic Outlook“, Frühjahrsprognose, diverse Ausgaben	1974–2004	9	Einige Prognosen nicht publiziert; Daten aus Artis (1987)
	„World Economic Outlook“, Herbstprognose, diverse Ausgaben	1974–2004	15	
Europäische Kommission, Brüssel	„European Economy“, Frühjahrsprognose, diverse Ausgaben	1974–2004	9	
	„European Economy“, Herbstprognose, diverse Ausgaben	1974–2004	14	
Wirtschafts- und Sozialwissenschaftliches Institut der Hans-Böckler-Stiftung, Düsseldorf	„WSI Mitteilungen“, diverse Ausgaben	1974–2004	12	2001 keine Wachstums- und 2001/02 keine Inflationsprognose
Institut der Deutschen Wirtschaft, Köln	„IW-Trends“, diverse Ausgaben	1970–2004	13	
Bundesregierung, Berlin/Bonn	Jahreswirtschaftsbericht, diverse Ausgaben	1970–2004	10	
				<b>DIW Berlin 2005</b>

**Tests auf Rationalität**

Ein Test auf Rationalität der Prognose kann direkt aus der Mincer-Zarnowitz-Regression abgeleitet werden, indem diese um exogene Variablen erweitert wird. Es wird also die Gleichung  $y_{t+1} = \beta_0 + \beta_1 \hat{y}_{t+1,t} + \beta_2 X_{t-1} + u_t$  geschätzt, wobei  $X_{t-1}$  für eine zum Zeitpunkt der Prognoseerstellung bekannte exogene Variable steht. Ein üblicher Wald-Test kann dann zur Überprüfung der Hypothese  $H_0: \beta_0 = 0 \wedge \beta_1 = 1 \wedge \beta_2 = 0$  verwendet werden.<sup>2</sup> Für eine optimale und rationale Prognose sollte es nicht möglich sein, eine Variable zu finden, die die Prognosefehler erklärt. Wenn diese Hypothese nicht abgelehnt werden kann, kann die Prognose als rational bezeichnet werden, da die exogene Variable keinen zusätzlichen Erklärungswert liefert. Es wird eine „schwache“ und eine „starke“ Version dieses Tests durchgeführt. Die „schwache“ Version des Tests bezieht sich auf einen Test auf Autokorrelation der Prognosefehler. Wenn keine Variable den Prognosefehler erklärt, sollten auch verzögerte Prognosefehler nicht informativ für die Erklärung der heuti-

<sup>2</sup> Vgl. K. Holden und D.A. Peel: On Testing for Unbiasedness and Efficiency of Forecasts. In: The Manchester School, Vol. 18, No. 2, 1990, S. 120–127.

gen Prognosefehler sein.<sup>3</sup> Definiert man  $y_{t+1} - \hat{y}_{t+1,t} = e_{t+1}$  und schätzt die Gleichung  $y_{t+1} = \beta_0 + \beta_1 \hat{y}_{t+1,t} + \beta_2 e_t + u_t$ , lautet die Nullhypothese des Wald-Tests  $H_0: \beta_0 = 0 \wedge \beta_1 = 1 \wedge \beta_2 = 0$ . Die entsprechende Zeile in den Tabellen 1 und 2 enthält wiederum die Irrtumswahrscheinlichkeit der Ablehnung der Nullhypothese.

Verzögerte Prognosefehler stellen jedoch nur eine Teilmenge der dem Prognostiker zur Verfügung stehenden Informationsmenge dar. Eine „starke“ Version des Tests – auch als Tests auf „rationale Erwartungsbildung“ oder „starke“ Rationalität bezeichnet – prüft, ob Prognosefehler unkorreliert mit *irgendeiner* zum Zeitpunkt der Prognose bekannten Information sind. Bei Vorliegen einer solchen Information – informativ bezüglich der Prognosefehler – ist die Prognose als nicht rational zu bezeichnen. Einige Autoren weisen zu Recht darauf hin, dass es schwer zu bestimmen ist, welche Informationen dem Prognostiker genau zum Zeitpunkt der Prognosen bekannt sind.<sup>4</sup> Aus diesem Grund wurden Variablen aus dem zweiten Quartal des Jahres herangezogen, in dem das hier verwendete Panel der Prognostiker startet. Konkret wurden folgende exogene Variablen getestet: kurzfristiger Zinssatz (Dreimonatsgeld), der reale effektive Wechselkurs, der Ölpreis sowie die Industrieproduktion in der OECD. Alle nichtstationären Variablen wurden über Vorjahresvergleiche in stationäre Variablen überführt. Die entsprechende Zeile in den Tabellen 1 und 2 enthält ebenso die Irrtumswahrscheinlichkeit der Ablehnung der Nullhypothese.

### Tests auf Informationsgehalt der Richtungsprognosen

Um den Informationsgehalt der Richtungsprognosen zu untersuchen, wurden die Wachstumsprognosen nach „Beschleunigung“ und „Verlangsamung“ in einer Kontingenztabelle klassifiziert<sup>5</sup> (Tabelle K2).

Tabelle K2

#### Kontingenztabelle (Richtungsprognosen)

Prognose	Resultat		Summe
	Beschleunigung	Verlangsamung	
Beschleunigung	$O_{ii}$	$O_{ij}$	$O_{i.}$
Verlangsamung	$O_{ji}$	$O_{jj}$	$O_{.j}$
Summe	$O_{.i}$	$O_{.j}$	$O$

**DIW** Berlin 2005

Mit dieser Klassifikation kann der Informationsgehalt anhand der Formel  $\frac{O_{ii}}{O_{ii} + O_{ji}} + \frac{O_{jj}}{O_{jj} + O_{ij}}$  gemessen werden.

Im Falle eines normalverteilten Zufallsprozesses (Münzwurf) ergäben sich bei entsprechend hoher Beobachtungszahl  $O_{ii} \approx O_{ji}$  sowie  $O_{jj} \approx O_{ij}$  und damit tendenziell ein Informationsgehalt von 1. Im Falle einer jederzeit perfekten Prognose ergäbe sich ein Informationsgehalt von 2. Alle Werte zwischen 1 und 2 zeigen also einen positiven Informationsgehalt an.

Zusätzlich kann getestet werden, ob der Informationsgehalt statistisch signifikant von 1 verschieden ist. Dazu werden die erwarteten Zellbelegungen unter der Nullhypothese keines Informationsgehaltes geschätzt. Der Schätzer für die Zellbelegungen der Kontingenztabelle lautet  $\hat{E}_{ij} = O_{i.}O_{.j}/O$ . Damit kann eine

Maßzahl (Pearsons  $\chi^2$ ) nach der Formel  $\sum_{i,j=1}^2 (O_{ij} - \hat{E}_{ij})^2 / \hat{E}_{ij}$  konstruiert werden, die unter der Nullhypothese

keines Informationsgehaltes  $\chi^2$ -verteilt ist. Tabelle 3 enthält den Informationsgehalt sowie die Irrtumswahrscheinlichkeit der Ablehnung der erläuterten Nullhypothese.

<sup>3</sup> Vgl. G. Kirchgässner: Sind die Erwartungen der Wirtschaftssubjekte rational? Eine empirische Untersuchung für die Bundesrepublik Deutschland. In: Weltwirtschaftliches Archiv, Bd. 120, 1984, S. 279–300.

<sup>4</sup> Vgl. G. Tichy: Konjunktur: stilisierte Fakten, Theorie, Prognose. 2., völlig neu bearb. Auflage. Berlin 1994.

<sup>5</sup> Vgl. F. Diebold und J. Lopez: Forecast Evaluation and Combination. NBER Technical Working Paper 192. Cambridge, MA 1996.

Für die Beurteilung der Prognosen wurde hier insbesondere folgenden Fragen nachgegangen: Wie hoch sind die Prognosefehler der Institutionen, und gibt es signifikante Unterschiede? Wie sind Prognosen anhand der Kriterien Unverzerrtheit und Rationalität zu bewerten? Wie gut werden Richtungsänderungen der Konjunktur prognostiziert?

### Prognosefehler

In den Tabellen 1 und 2 sind die mittleren, die mittleren absoluten und die Wurzeln der mittleren quadrierten Fehler für die Wachstums- und die Inflationsprognosen berechnet. Die mittleren Fehler für das Wachstum sind in den meisten Fällen nega-

Tabelle 1

**Güte, Unverzerrtheit und Rationalität der Wachstumsprognosen**

	Mittlerer Fehler	Mittlerer absoluter Fehler	Wurzel des mittleren quadratischen Fehlers	Test auf Unverzerrtheit	Test auf „schwache“ Rationalität	Test auf „starke“ Rationalität unter Zuhilfenahme des kurzfristigen Zinssatzes	Test auf „starke“ Rationalität unter Zuhilfenahme des realen Außenwertes	Test auf „starke“ Rationalität unter Zuhilfenahme des Ölpreises	Test auf „starke“ Rationalität unter Zuhilfenahme der OECD-Industrieproduktion
	Prozentpunkte			Wahrscheinlichkeit <sup>1</sup> in %					
IWF, Herbst	-0,71	1,39	1,83	8	96	1	14	9	13
GD, Herbst	-0,39	1,21	1,65	33	86	6	29	28	37
Europäische Kommission, Herbst	-0,46	1,22	1,67	26	88	2	33	21	24
Sachverständigenrat	-0,36	1,02	1,45	35	95	9	31	54	34
OECD	-0,35	1,09	1,58	42	60	8	49	30	29
IW	-0,34	1,17	1,71	30	14	9	48	30	42
IfW	-0,26	1,08	1,55	35	95	13	30	32	36
HWWA	-0,22	0,98	1,38	67	57	29	61	56	63
WSI	-0,14	1,15	1,58	51	71	12	44	33	38
ifo	-0,13	0,96	1,33	86	44	40	86	50	64
DIW Berlin	0,02	1,03	1,51	50	62	9	46	18	50
IWH	-0,40	0,75	1,00	42	17	64	23	61	28
RWI	-0,30	1,02	1,34	40	65	15	51	55	56
Jahreswirtschaftsbericht	-0,26	0,95	1,40	56	38	29	58	35	66
IWF, Frühjahr	-0,19	0,88	1,22	40	12	20	49	21	65
Europäische Kommission, Frühjahr	-0,23	0,83	1,11	42	44	24	67	25	76
GD, Frühjahr	-0,24	0,84	1,18	48	91	20	66	24	58

<sup>1</sup> Angegeben ist die Irrtumswahrscheinlichkeit der Ablehnung der Nullhypothese.

Quelle: Berechnungen des DIW Berlin.

DIW Berlin 2005

Tabelle 2

**Güte, Unverzerrtheit und Rationalität der Inflationsprognosen**

	Mittlerer Fehler	Mittlerer absoluter Fehler	Wurzel des mittleren quadratischen Fehlers	Test auf Unverzerrtheit	Test auf „schwache“ Rationalität	Test auf „starke“ Rationalität unter Zuhilfenahme des kurzfristigen Zinssatzes	Test auf „starke“ Rationalität unter Zuhilfenahme des realen Außenwertes	Test auf „starke“ Rationalität unter Zuhilfenahme des Ölpreises	Test auf „starke“ Rationalität unter Zuhilfenahme der OECD-Industrieproduktion
	Prozentpunkte			Wahrscheinlichkeit <sup>1</sup> in %					
IWF, Herbst	-0,11	0,71	0,95	59	9	73	58	63	79
GD, Herbst	0,03	0,83	1,07	23	0	29	26	27	18
Europäische Kommission, Herbst	0,07	0,73	0,97	13	0	22	22	19	12
Sachverständigenrat	0,06	0,74	0,97	9	1	20	7	12	18
OECD	-0,05	0,58	0,79	85	30	89	74	95	95
IW	-0,03	0,55	0,77	27	26	27	29	44	44
IfW	0,19	0,81	1,20	22	0	16	21	20	18
HWWA	0,05	0,68	0,95	31	1	46	23	39	49
WSI	0,01	0,83	1,09	92	0	93	68	74	88
ifo	-0,02	0,63	0,87	66	18	77	50	73	76
DIW Berlin	0,21	0,72	0,95	17	2	33	29	27	31
IWH	0,04	0,53	0,74	22	42	0	22	42	30
RWI	0,13	0,62	0,86	3	0	8	2	2	7
Jahreswirtschaftsbericht	0,16	0,63	0,91	42	3	70	50	75	80
IWF, Frühjahr	-0,16	0,61	0,86	52	1	46	12	49	59
Europäische Kommission, Frühjahr	-0,05	0,38	0,48	62	15	78	80	66	59
GD, Frühjahr	0,05	0,45	0,60	52	19	61	33	86	55

<sup>1</sup> Angegeben ist die Irrtumswahrscheinlichkeit der Ablehnung der Nullhypothese.

Quelle: Berechnungen des DIW Berlin.

DIW Berlin 2005

tiv, was generell eine leichte Tendenz zur Wachstumsüberschätzung signalisiert. Statistische Prüfungen zeigen jedoch, dass diese Überschätzungs-

tendenz in keinem Fall signifikant ist, also auch als zufällig angesehen werden kann. Für die Inflationsprognosen liegt der mittlere Fehler in allen

Fällen nahe null. Die Wurzeln der mittleren quadratischen Fehler sind relativ groß – zwischen 1,1 und 1,8 Prozentpunkten für die Wachstumsprognosen sowie zwischen 0,5 und 1,2 Prozentpunkten für die Inflationsprognosen. Allerdings ist die Größe der Fehler über den betrachteten Zeitraum etwas zurückgegangen. Um eine Abschätzung über die Unsicherheit zumindest der Wachstumsprognosen zu geben, wurde der Median der Prognosen und darauf basierend die Wurzel des mittleren quadratischen Fehlers gegenüber dem Median aller verfügbaren Prognosen für ein rückwärts gerichtetes Zeitfenster von fünf Jahren berechnet.

In Abbildung 1 ist der Median der hier betrachteten Prognosen zusammen mit den Erstveröffentlichungen der Wachstumsergebnisse und einem aus dem Zweifachen der Wurzel des mittleren quadratischen Fehlers bestimmten 95-Prozent-Konfidenzintervall dargestellt. Das Konfidenzband um die Wachstumsprognosen ist weit. Die Unterschiede zwischen den Prognosefehlern sind vor allem auf den Erscheinungszeitpunkt der Vorhersagen zurückzuführen. Dies ist auch zu erwarten, da die Effizienz der Prognose mit der Länge des Prognosezeitraums negativ korreliert sein sollte. Abbildungen 2 und 3 zeigen für die einzelnen Institutionen die jeweiligen Wurzeln des mittleren quadratischen Fehlers, wobei die Prognosen nach ihrem Veröffentlichungszeitpunkt geordnet wurden. Die Regressionsgleichungen machen deutlich, dass die zeitliche Abfolge der Prognosen den Unterschied in den Prognosefehlern zu einem großen Teil erklärt.

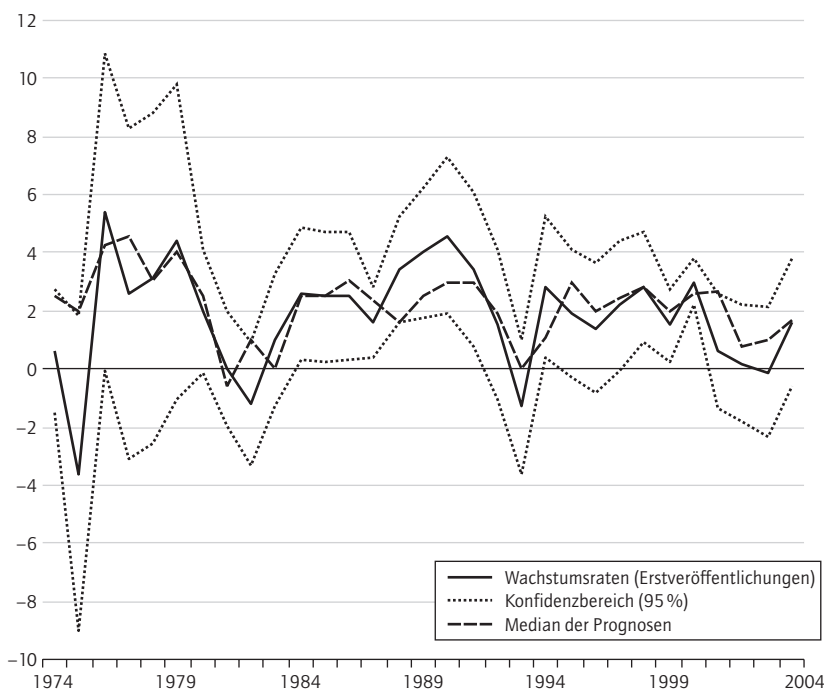
**Rationalitätstest**

Es wurden Tests auf Unverzerrtheit und Rationalität der Prognosen durchgeführt. Die Einzelheiten zu den verwendeten Tests und Hypothesen sind im Kasten detailliert erläutert. Die Ergebnisse zeigen, dass die Wachstumsprognosen im Allgemeinen den beschriebenen Kriterien von Unverzerrtheit und „schwacher“ Rationalität genügen. Bei den Inflationsprognosen wird das Kriterium der „schwachen“ Rationalität vielfach abgelehnt – was aber ein Reflex bestimmter Zeitreiheneigenschaften der Inflationsraten sein kann.<sup>12</sup> Tests auf „starke“ Rationalität ergeben, dass in einigen Fällen exogene Variablen einen Informationsgehalt bezüglich der Prognosefehler aufweisen. Insbesondere monetäre Variablen scheinen nicht von allen Prognostikern

<sup>12</sup> Dies liegt vor allem daran, dass die Inflationsrate die Zeitreiheneigenschaft einer ausgeprägt hohen Persistenz aufweist, d. h. Inflationsraten zweier aufeinander folgender Zeiträume sehr stark miteinander korreliert sind. Tests zeigen, dass die Inflationsrate – zumindest für die hier betrachteten Jahreswerte – jedoch die Eigenschaft der Stationarität aufweist. Dies ist Bedingung für die hier verwendete Spezifikation der Tests.

Abbildung 1

**Konfidenzbereich der Wachstumsprognosen**  
Veränderung in %

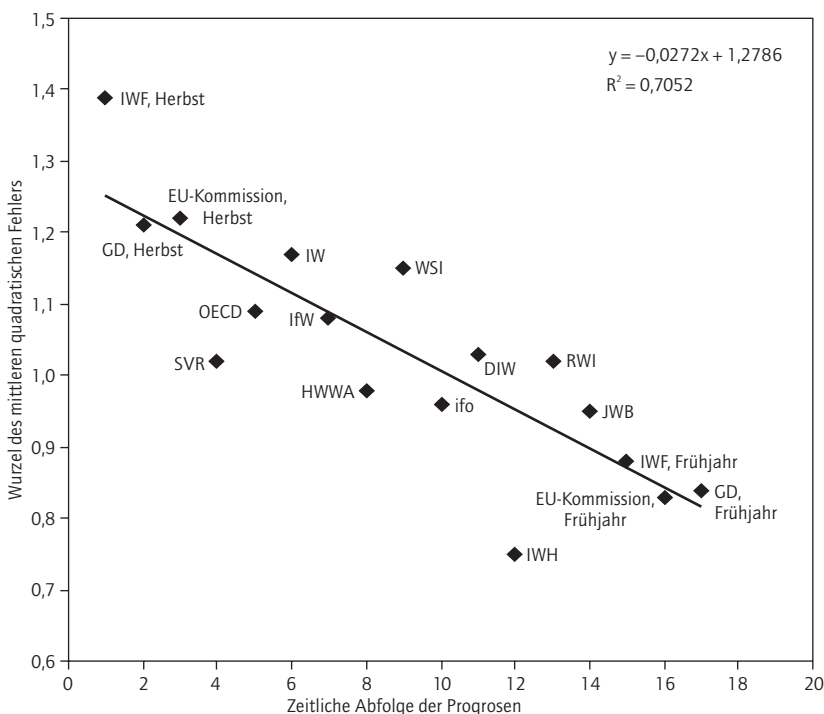


Quelle: Berechnungen des DIW Berlin.

DIW Berlin 2005

Abbildung 2

**Wachstumsprognosen**  
Prognosefehler in Abhängigkeit von der zeitlichen Abfolge der Prognosen



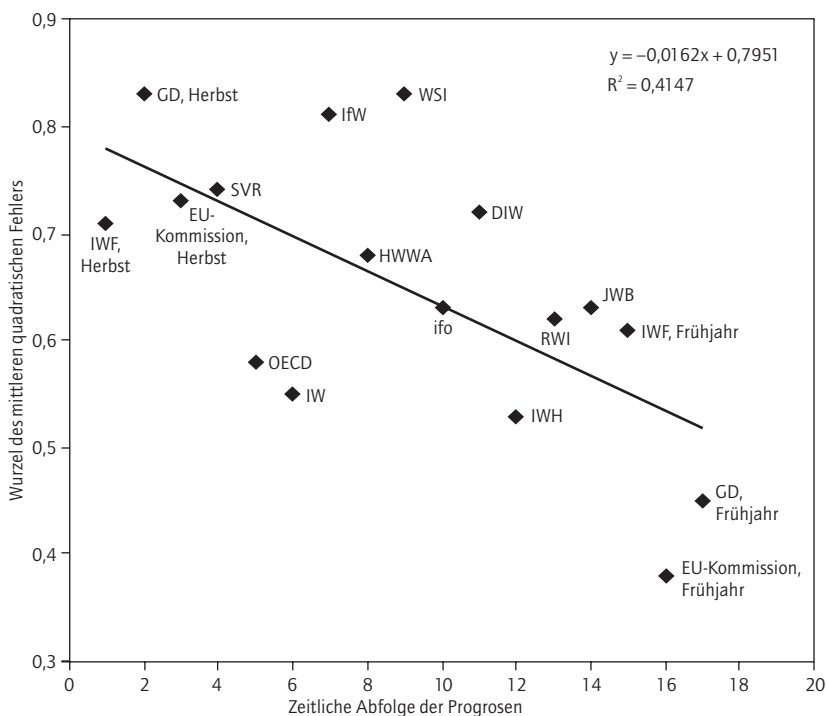
Quelle: Berechnungen des DIW Berlin.

DIW Berlin 2005

Abbildung 3

**Inflationsprognosen**

Prognosefehler in Abhängigkeit von der zeitlichen Abfolge der Prognosen



Quelle: Berechnungen des DIW Berlin.

DIW Berlin 2005

in ausreichendem Maße berücksichtigt worden zu sein.<sup>13</sup> Auch Tests im Rahmen eines Panel-Ansatz

Tabelle 3

**Informationsgehalt der Richtungsprognosen**

	Wachstumsprognosen		Inflationsprognosen	
	Informationsgehalt <sup>1</sup>	Unabhängigkeitstest <sup>2</sup>	Informationsgehalt <sup>1</sup>	Unabhängigkeitstest <sup>2</sup>
IWF, Herbst	1,47	1	1,62	0
GD, Herbst	1,34	2	1,46	0
Europäische Kommission, Herbst	1,39	2	1,58	0
Sachverständigenrat	1,31	5	1,51	0
OECD	1,65	0	1,72	0
IW	1,78	0	1,33	1
IfW	1,41	1	1,51	0
HWWA	1,36	3	1,51	0
WSI	1,41	1	1,45	0
ifo	1,47	0	1,40	1
DIW Berlin	1,47	0	1,50	0
IWH	1,83	0	1,43	3
RWI	1,36	3	1,39	0
Jahreswirtschaftsbericht	1,41	1	1,52	0
IWF, Frühjahr	1,39	1	1,47	0
Europäische Kommission, Frühjahr	1,53	0	1,76	0
GD, Frühjahr	1,52	0	1,76	0

<sup>1</sup> Im Falle eines normalverteilten Zufallsprozesses ergibt sich ein Informationsgehalt von 1, bei einer perfekten Prognose von 2. Alle Werte zwischen 1 und 2 zeigen also einen positiven Informationsgehalt an.

<sup>2</sup> Angegeben ist die Irrtumswahrscheinlichkeit der Ablehnung der Nullhypothese.

Quelle: Zusammenstellung des DIW Berlin.

DIW Berlin 2005

zes mit einer Korrektur des Einflusses der zeitlichen Abfolge auf die Prognosen bestätigen dies.

**Prognose von Richtungsänderungen**

Eine im Zusammenhang mit Konjunkturprognosen häufig gestellte Frage lautet: Liefern uns die Prognosen Informationen bezüglich der Richtungsänderung des Wachstums bzw. der Inflation? Um dies festzustellen, wurden die konjunkturelle Entwicklung und Prognosen nach den Merkmalen „Beschleunigung“ und „Verlangsamung“ klassifiziert, und es wurde überprüft, wie oft die Prognostiker hier richtig lagen. Dazu wurde eine Kontingenztafel aufgestellt und der Informationsgehalt ermittelt (nähere Erläuterungen siehe Kasten). Alle Prognosen haben einen deutlich höheren Informationsgehalt als auf Zufall basierende Prognosen („Münzwurf“); die Ergebnisse der durchgeführten  $\chi^2$ -Tests bestätigen die statistische Signifikanz dieser Aussage (Tabelle 3).

**Fazit**

Alles in allem zeigt sich, dass eine Bewertung von Prognosen ein komplexer Vorgang ist. Allein auf das Kriterium der Prognosegüte zu schauen, greift zu kurz, da prinzipiell auch unwissenschaftliche Prognosen Ereignisse zufällig richtig vorhersagen können. Mittelfristig können sich Prognosen allerdings „bewähren“ und zur Orientierung dienen. Für die Bewertung von Prognosen sind folgende Argumente entscheidend:

1. Prognosen müssen vollständig sein, das heißt, sie sollten eine Darstellung des zugrunde liegenden Modells wie auch der Anwendungsbedingungen (vor allem exogene Annahmen) enthalten. Dies kann formal – im Rahmen eines für die Prognose verwendeten Modells – oder verbal – wie bei den Wirtschaftsforschungsinstituten – erfolgen. So wird sichergestellt, dass Prognosen dem Kriterium der Nachprüfbarkeit genügen.
2. Prognosen sollten im längerfristigen Vergleich unverzerrt sein und dem definierten Kriterium von Rationalität genügen, das heißt, es sollte keine zum Zeitpunkt der Prognoseerstellung all-

<sup>13</sup> H. M. Hagen und G. Kirchgässner: Interest Rate Based Forecasts vs. Professional Forecast for German Economic Growth: A Comparison. Discussion Paper No. 9714. Universität St. Gallen 1997; G. Kirchgässner und M. Savioz: Monetary Policy and Forecasts for Real GDP Growth: An Empirical Investigation for the Federal Republic of Germany. In: German Economic Review, Vol. 2, No. 4, 2001, S. 339–365; C. R. Harvey: Interest Rate Based Forecasts of German Economic Growth. In: Weltwirtschaftliches Archiv, Bd. 127, 1991, S. 701–718. Sie kommen alle zu ähnlichen Ergebnissen.

gemein zugängliche Information geben, die ungenügend berücksichtigt wurde und so für den Prognosefehler maßgeblich ist. Die für Deutschland vorliegenden Prognosen erfüllen im Allgemeinen diese Anforderungen. Allerdings deutet der nachweisbare Informationsgehalt monetärer Variablen bezüglich der Prognosefehler bei Wachstumsprognosen einzelner Institutionen darauf hin, dass hier Verbesserungsbedarf besteht.

3. Prognosen können sich „bewähren“. Auf der Grundlage der historischen Prognoseabweichungen kann – solange die Prognosefehler normalverteilt sind – ein näherungsweise berechnetes Konfidenzband ermittelt werden. Dieses beträgt für ein 95-Prozent-Konfidenzintervall derzeit etwa  $\pm 2$  Prozentpunkte, das heißt, eine Wachstumsprognose von 1 % liegt mit 95-prozentiger Sicherheit im Intervall von  $-1$  bis  $+3$  %. Das 66-Prozent-Konfidenzintervall läge bei  $\pm 1$  Prozentpunkt.
4. Die absolute Größe der Prognosefehler ist in den vergangenen Dekaden zurückgegangen, wegen

der insgesamt niedrigeren Wachstums- und Inflationsraten jedoch nicht die relative Größe der Prognosefehler.<sup>14</sup> Im internationalen Vergleich sind die Prognosefehler für Deutschland nicht ungewöhnlich hoch.<sup>15</sup>

5. Für die Bewertung der Prognosen ist zu berücksichtigen, dass es eine Vielzahl von Faktoren gibt, die von Prognostikern nur schwer und mit hoher Unsicherheit einzuschätzen sind. Dazu gehören Annahmen über die exogenen Variablen wie Ölpreis und Wechselkurs, die sowohl die Mehrzahl der Marktteilnehmer wie auch viele Konjunkturforscher in den letzten Jahren vielfach falsch eingeschätzt haben. Externe Schocks – wie die konjunkturelle Abkühlung nach dem 11. September 2001 – sind ebenfalls beispielhaft für solche schwer absehbaren Faktoren.

<sup>14</sup> Vgl. U. Heilemann und H. O. Stekler: Has the Accuracy of German Macroeconomic Forecasts Improved? Discussion Paper of the German Research Council's Research Centre 475, 31/03. Dortmund 2003.

<sup>15</sup> Vgl. P. Loungani: How Accurate are Private Sector Forecasts? Cross-country Evidence from Consensus Forecasts of Output Growth. In: International Journal of Forecasting, Vol. 17, No. 3, 2001, S. 419–432.



## Aus den Veröffentlichungen des DIW Berlin

Marco Caliendo and Sabine Kopeinig

### Some Practical Guidance for the Implementation of Propensity Score Matching

Propensity Score Matching (PSM) has become a popular approach to estimate causal treatment effects. It is widely applied when evaluating labour market policies, but empirical examples can be found in very diverse fields of study. Once the researcher has decided to use PSM, he is confronted with a lot of questions regarding its implementation. To begin with, a first decision has to be made concerning the estimation of the propensity score. Following that one has to decide which matching algorithm to choose and determine the region of common support. Subsequently, the matching quality has to be assessed and treatment effects and their standard errors have to be estimated. Furthermore, questions like "what to do if there is choice-based sampling?" or "when to measure effects?" can be important in empirical studies. Finally, one might also want to test the sensitivity of estimated treatment effects with respect to unobserved heterogeneity or failure of the common support condition. Each implementation step involves a lot of decisions and different approaches can be thought of. The aim of this paper is to discuss these implementation issues and give some guidance to researchers who want to use PSM for evaluation purposes.

Diskussionspapier Nr. 485  
April 2005

Richard V. Burkhauser and Dean R. Lillard

### The Contribution and Potential of Data Harmonization for Cross-National Comparative Research

The promise of empirical evidence to inform policy makers about their population's health, wealth, employment and economic well being has propelled governments to invest in the harmonization of country specific micro data over the last 25 years. We review the major data harmonization projects launched over this period. These projects include the Luxembourg Income Study (LIS), the Cross-National Equivalent File (CNEF), the Consortium of Household Panels for European Socio-Economic Research (CHER), the European Community Household Panel (ECHP), the European Union Statistics on Income and Living Conditions (EU-SILC), and the Survey of Health, Aging and Retirement in Europe (SHARE). We discuss their success in providing reliable data for policy analysis and how they are being used to answer policy questions. While there have been some notable failures, on the whole these harmonization efforts have proven to be of major value to the research community and to policy makers.

Diskussionspapier Nr. 486  
April 2005

Die Volltextversionen der Diskussionspapiere liegen von 1998 an komplett als pdf-Dateien vor und können von der entsprechenden Website des DIW Berlin heruntergeladen werden ([www.diw.de/deutsch/produkte/publikationen/diskussionspapiere](http://www.diw.de/deutsch/produkte/publikationen/diskussionspapiere)).

#### Impressum

DIW Berlin  
Königin-Luise-Str. 5  
14195 Berlin

#### Herausgeber

Prof. Dr. Klaus F. Zimmermann (Präsident)  
Prof. Dr. Georg Meran (Vizepräsident)  
Dr. Tilman Brück  
Dörte Höppner  
Prof. Dr. Claudia Kemfert  
Dr. Bernhard Seidel  
Prof. Dr. Viktor Steiner  
Prof. Dr. Alfred Steinherr  
Prof. Dr. Gert G. Wagner  
Prof. Dr. Axel Werwatz, Ph.D.  
Prof. Dr. Christian Wey

#### Redaktion

Dr. habil. Björn Frank  
Dr. Elke Holst  
Jochen Schmidt  
Dr. Mechthild Schrooten

#### Pressestelle

Renate Bogdanovic  
Tel. +49 - 30 - 897 89-249  
[presse@diw.de](mailto:presse@diw.de)

#### Vertrieb

DIW Berlin Leserservice  
Postfach 74  
77649 Offenburg  
[leserservice@diw.de](mailto:leserservice@diw.de)  
Tel. 01805 - 19 88 88 \*dtms/12 Cent./min.

Abo-Betreuung durch  
Abonnenten Service Center GmbH  
Geschäftsführer: Heinz-Jürgen Koch  
Marlener Str. 4  
77656 Offenburg

#### Bezugspreis

Jahrgang Euro 180,-  
Einzelheft Euro 7,-  
(jeweils inkl. MwSt. und Versandkosten)  
Abbestellungen von Abonnements  
spätestens 6 Wochen vor Jahresende

ISSN 0012-1304

Bestellung unter [leserservice@diw.de](mailto:leserservice@diw.de)

#### Konzept und Gestaltung

kognito, Berlin

#### Satz

Wissenschaftlicher Text-Dienst (WTD), Berlin

#### Druck

Druckerei Conrad GmbH  
Oranienburger Str. 172  
13437 Berlin

Nachdruck und sonstige Verbreitung – auch auszugsweise – nur mit Quellenangabe und unter Zusendung eines Belegexemplars an die Abteilung Information und Organisation zulässig.