

# Treffgenauigkeit, Rationalität und Streuung von Konjunkturprognosen für Deutschland\*

Von Ulrich Fritsche und Jörg Döpke\*\*

## 1 Einleitung

Konjunkturprognostiker und deren prognostische Fähigkeiten standen schon oft im Kreuzfeuer der öffentlichen Kritik. Seit einiger Zeit mehren sich die Fragen über die Tauglichkeit von Konjunkturprognosen. Da war vom „Blindflug der Forscher“ zu lesen (vgl. Der Spiegel 2005), Feuilletonisten forderten „Entlasst die Experten“ (vgl. FAZ 2005a) und auch Altbundeskanzler Gerhard Schröder hegte wenig Sympathie für „diese Art der Wissenschaft, die ihn an Meteorologie erinnere“ (vgl. FAZ 2005b). Dass viele Wirtschaftsforschungsinstitute mit öffentlichen Geldern gefördert werden, steigert den Legitimationsdruck zusätzlich.

Umso wichtiger ist es, die Debatte zu versachlichen und sich Grenzen, Möglichkeiten und Chancen der Konjunkturprognosen unter Verwendung *nachprüfbarer* Kriterien vor Augen zu führen.<sup>1</sup> Der folgende Beitrag soll dazu beitragen. Dazu wird zunächst im Abschnitt 2 der Frage nachgegangen, welche erkenntnistheoretischen Probleme es grundsätzlich bei Prognosen gibt und was diese leisten können. Abschnitt 3 beschreibt den für die folgenden empirischen Untersuchungen verwendeten Datensatz. Im Abschnitt 4 geht es um die Treffgenauigkeit und Rationalität von Prognosen und im Abschnitt 5 soll geklärt werden, warum Prognosen differieren und welche Bestimmungsgründe es für Phasen hoher Divergenz bei den Prognosen geben könnte. Abschnitt 6 fasst die Ergebnisse zusammen.

## 2 Wissenschaftlich fundierte Prognosen oder Blick in den Kaffeesatz?

Eine Prognose ist eine Aussage, die das Eintreten eines bestimmten Ereignisses in der Zukunft beschreibt. Wie genau diese Aussage zustande kommt – sei es durch Intuition des Prognostikers oder durch ein mehr oder minder formal beschreibbares Prognosemodell – ist zunächst einmal unwichtig für das mögliche Eintreffen und damit die Richtigkeit dieser Aussage. Insofern ist es irrelevant, ob die Prognose aus dem Kaffeesatz oder einem ökonomischen Modell stammt.

\* Die Auffassungen in diesem Papier sind ausschließlich die der Autoren und stellen nicht notwendigerweise die des DIW bzw. der Deutschen Bundesbank dar.

\*\* Korrespondenz: Universität Hamburg, Fakultät Wirtschafts- und Sozialwissenschaften, Department Wirtschaft und Politik, und Deutsches Institut für Wirtschaftsforschung (DIW Berlin), Abteilung Konjunktur, E-Mail: [ulrich.fritsche@wiso.uni-hamburg.de](mailto:ulrich.fritsche@wiso.uni-hamburg.de)

<sup>1</sup> Der Aufsatz basiert im Wesentlichen auf gemeinsamen Forschungsarbeiten der Autoren. Vgl. Döpke und Fritsche (2006a, b, c), Döpke und Fritsche (2005), Döpke und Fritsche (2004) sowie Fritsche (2005).

Was den wissenschaftlich arbeitenden Prognostiker von unseriösen Propheten unterscheiden sollte, ist eine logische Begründung der Prognoseaussage; mit anderen Worten: Es geht um die Eigenschaften des Prognosearguments (Tietzel 1989). Eine logisch fundierte Aussage über das zukünftige Auftreten eines Ereignisses wird deduktiv aus einem (oder mehreren) allgemeinen Anwendungsgesetz(en) und einer oder mehreren Anwendungsbedingung(en) abgeleitet.

Nur Gesetz und Anwendungsbedingung(en) zusammen genommen stellen eine hinreichende Bedingung für das Prognoseargument dar. Aussagen, die in diesem Sinne kein vollständiges Prognoseargument oder ein – wie es Tietzel (1989) formuliert – „logisch defektes“ Prognoseargument benennen, sind keine wissenschaftlichen Prognosen, da sie sich der Überprüfbarkeit entziehen.

Nach diesem Kriterium sind auch Konjunkturprognosen zu bewerten. Als wissenschaftlich sind sie nur zu bezeichnen, wenn sie auf einem methodischen Fundament (dem „allgemein gültigen Gesetz“) und konkreten Anwendungsbedingungen (Analyse der Ist-Situation, Darstellung der Annahmen der Prognose) beruhen. Im Allgemeinen erfüllen die meisten veröffentlichten Konjunkturprognosen diese Anforderungen<sup>2</sup>, wenn auch häufig die zugrunde liegenden Annahmen und Modelle nicht vollständig oder nur ungenau benannt werden. Allerdings sind selbst hinreichend wissenschaftlich begründete Prognosen theoretischen Problemen ausgesetzt:

1. Der Geltungsbereich der der Prognose zugrunde liegenden „allgemeinen Aussage“ ist beschränkt. Die Tatsache, dass man eine bisher empirisch bewährte Theorie gefunden hat, besagt nicht, dass diese Theorie auch Gültigkeit für alle zukünftigen Ereignisse haben muss.<sup>3</sup> Ökonometrische Modelle und das Erfahrungswissen der Prognostiker sind vergangenheitsbezogen, die daraus abgeleiteten „Gesetzmäßigkeiten“ können im erkenntnistheoretischen Sinn nur für die Vergangenheit Gültigkeit beanspruchen. So verändern sich bestimmte strukturelle ökonomische Beziehungen ständig. Die empirische Wirtschaftsforschung versucht zwar, durch entsprechende „Strukturbruchtests“ diesem Problem zu entgegnen; es ist jedoch besonders schwer, Veränderungen am Ende des *Ex-post*-Zeitraumes, d. h. an dem für die Prognose besonders relevanten „aktuellen Rand“, zu entdecken.
2. Die Anwendungsbedingungen der Prognosen sind unsicher. Alle Konjunkturprognosen benutzen mehr oder weniger formale Modelle, die dazu dienen, Komplexität zu reduzieren. Die Modelle enthalten alle exogen gesetzte oder a priori formulierte Annahmen, die erfüllt sein müssen, um das Prognoseergebnis herbeizuführen. So kann es eine Annahme über die Entwicklung des Ölpreises sein, die für die Richtigkeit der Prognose gegeben sein muss. Zwar versuchte man schon in den 1960er/1970er Jahren mit immer größeren Modellen mehr und mehr Annahmen zu „endogenisieren“. Letztlich kann man damit aber das Problem nicht lösen, sondern verschiebt es nur auf eine weiterführende Ebene.

<sup>2</sup> Problematisch sind aus theoretischer Perspektive die häufig benutzten Trendextrapolationsmethoden. Solche Verfahren sind – wenn überhaupt – aus erkenntnistheoretischer Sicht nur auf streng geschlossene Systeme anwendbar, nicht jedoch auf offene und sich verändernden Systeme.

<sup>3</sup> Die Entwicklung der Naturwissenschaften im vergangenen Jahrhundert wie auch die berühmte keynesianische und andere Revolutionen und Konterrevolutionen in der Makroökonomie zeigen, dass Denkgebäude ganzer Wissenschaftszweige in relativ kurzer Zeit umgeworfen werden können.

3. Erwartungen und Prognosen sind selbst Bestandteil des zu analysierenden ökonomischen Systems.<sup>4</sup> Deshalb kann es zu sich selbst erfüllenden Prognosen kommen. Das Gerücht über eine nahende Währungskrise könnte so in eine Spekulationswelle münden, die die Währung dann tatsächlich ins Wanken bringt. Ebenso könnte das Gerücht über die Liquiditätskrise einer Geschäftsbank zum Abzug der Spareinlagen und damit zu der erwarteten Liquiditätskrise führen. Auch sich selbst zerstörende Prognosen sind denkbar: Wenn allen Wirtschaftssubjekten eine Vorhersage bekannt ist, könnten die Wirtschaftssubjekte ihr Handeln darauf einrichten und die Prognose träfe nicht mehr zu. Am konsequentesten ist dieses Argument bereits von Morgenstern (1928) vorgetragen worden, der deshalb die Möglichkeit zutreffender wirtschaftlicher Prognosen grundsätzlich bestritt. Freilich ist die Bedeutung sich selbst aufhebender Prognosen auch mit guten Gründen bestritten worden. Schon Löwe (1929) wies darauf hin, dass jede Diagnose einer wirtschaftlichen Situation zumindest implizit prognostischen Charakter trägt und schon daher das Morgensternsche Argument wenig praktische Relevanz hat. Modigliani und Grunberg (1958) zeigten zudem, dass die Hypothese sich selbst aufhebender Prognosen auch theoretisch recht restriktive Annahmen erfordert, die kaum je erfüllt sein dürften (vgl. auch die Darstellung der Diskussion bei Stege 1989).

Was folgt daraus? Prognosen sind schon aus grundsätzlichen Gründen unvollkommen. Aus einem prognostischen Fehlschlag sollte zwar, wenn möglich, gelernt werden, aber ein Rückschluss auf die Gültigkeit der zugrunde liegenden Theorie ist nicht ohne weiteres möglich. Dennoch ist der Versuch einer Prognoseerstellung keineswegs zwecklos. Eine wissenschaftlich fundierte Vorhersage reduziert die Unsicherheit über die Zukunft. Die Abschätzung der Auswirkungen unserer Entscheidungen kann – bei aller Beschränktheit – nur auf Basis fundierter Prognosen erfolgen. Allerdings müssen wissenschaftliche Prognosen auch ebensolchen Kriterien genügen und einer entsprechenden Überprüfung standhalten.

### 3 Daten

Im Folgenden werden verschiedene öffentlich zugängliche Prognosen nationaler und internationaler Institutionen untersucht. Die verwendeten Daten und Quellen sind im Anhang dieses Aufsatzes bzw. bei Döpke und Fritsche (2004) ausführlich erläutert. Es wurden für den Zeitraum von 1970 bis 2004 Prognosen für das Wachstum des Bruttoinlandsprodukts (1983 bis 1989) bzw. Bruttoinlandsprodukts (alle anderen Jahre) und Inflationsprognosen (Deflator des privaten Verbrauchs) untersucht. Als Referenzmaßstab dienten die Erstveröffentlichungen der Wachstumsraten des Bruttoinlands- bzw. Bruttoinlandsprodukts sowie die letzten verfügbaren Daten zur Inflationsentwicklung.<sup>5</sup>

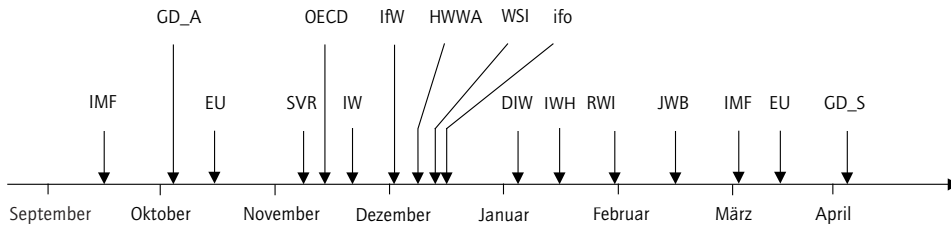
Die untersuchten Prognosen werden in der Regel im Winterhalbjahr in einer regelmäßigen Abfolge veröffentlicht und beziehen sich immer auf das kommende Kalenderjahr bzw. bei

<sup>4</sup> Das Phänomen ist von Oskar Morgenstern und Karl Popper ausführlich diskutiert worden. Vgl. Tietzel (1989: 554ff.).

<sup>5</sup> Die Wahl bezieht sich auf die jeweils schlagzeilenträchtige Zahl für das „Wachstum“ – hier änderte sich das Konzept des Statistischen Bundesamtes in den vergangenen Jahrzehnten. Bei der Inflationszahl gab es in einigen Fällen keine explizite Erläuterung, welcher Preisindex zugrunde lag. In solchem Falle wurde davon ausgegangen, dass der genannte Deflator gemeint war. Es wird davon ausgegangen, dass das Problem der Datenrevision bei Inflationszahlen vernachlässigbar ist.

Abbildung 1

## Institutionen und typische Veröffentlichungszeitpunkte



IMF = Internationaler Währungsfonds, GD\_A = Gemeinschaftsdiagnose Herbst, EU = Eu-Kommission, SVR = Sachverständigenrat, IW = Institut der Deutschen Wirtschaft (IW), OECD = Organisation für Wirtschaftliche Zusammenarbeit und Entwicklung, Paris, IfW = Institut für Weltwirtschaft, Kiel, HWWA = HWWA-Institut, Hamburg, WSI = Wirtschafts- und Sozialwissenschaftliches Institut der Hans-Böckler-Stiftung (WSI), 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, GD\_S = Gemeinschaftsdiagnose Frühjahr.

Quelle: DIW Berlin.

einem Veröffentlichungstermin nach dem Jahreswechsel auf das dann laufende Jahr. Abbildung 1 zeigt die nach unserer Kenntnis typische Abfolge der Veröffentlichungszeitpunkte der untersuchten Institutionen auf einem Zeitstrahl. Die frühesten in die Untersuchung einbezogenen Prognosen datieren aus dem September, während die am spätesten publizierten Vorhersagen im erst im Frühjahr erscheinen.

Aus der Abbildung ist ersichtlich, dass in die Untersuchung Prognosen eingehen, die sich alle auf das gleiche Kalenderjahr beziehen aber wegen unterschiedlicher Veröffentlichungstermine verschiedene Zeithorizonte betreffen. Der Prognosefehler sollte, neben anderen Faktoren, auch vom Prognosehorizont abhängen. Um dies zu verdeutlichen, wird in Abbildung 2 die Abhängigkeit der Prognosefehler (hier der mittlere absolute Fehler der Prognosen der Institutionen) von der Rangfolge der Veröffentlichungen gezeigt. Es ergibt sich das vermutete Bild: Frühe Veröffentlichungen mit entsprechend längerem Horizont haben einen höheren mittleren Fehler als später veröffentlichte.

Dieser in der Sache wenig überraschende Befund widerspricht zumindest nicht der Rationalitätsanforderung an Prognosen, solange die Prognose erwartungstreu ist. Die vom Horizont abhängige Effizienz der Prognosen muss jedoch bei der Konstruktion der Tests und der Wahl angemessener Untersuchungsmethoden in Betracht gezogen werden.

## 4 Treffgenauigkeit und Rationalität

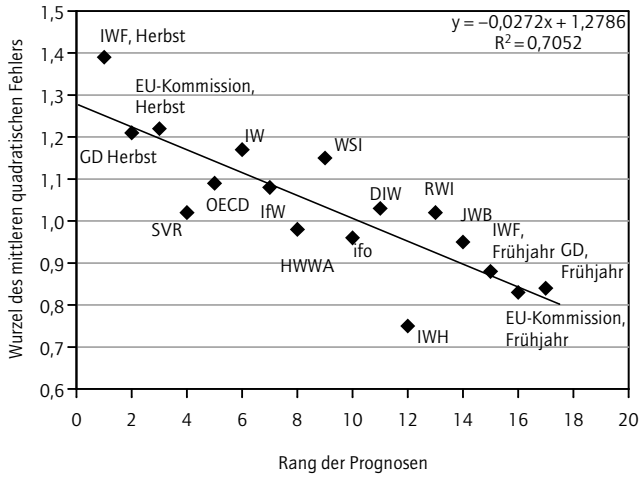
### 4.1 Zur Treffgenauigkeit

Wie Abbildung 2 gezeigt hat, ist empirisch die Treffgenauigkeit von Prognosen zumindest vom Prognosehorizont abhängig: Je weiter die Prognose in die Zukunft reicht, desto größer die Fehlerwahrscheinlichkeit. Wie aber ist die relative Treffgenauigkeit zwischen den

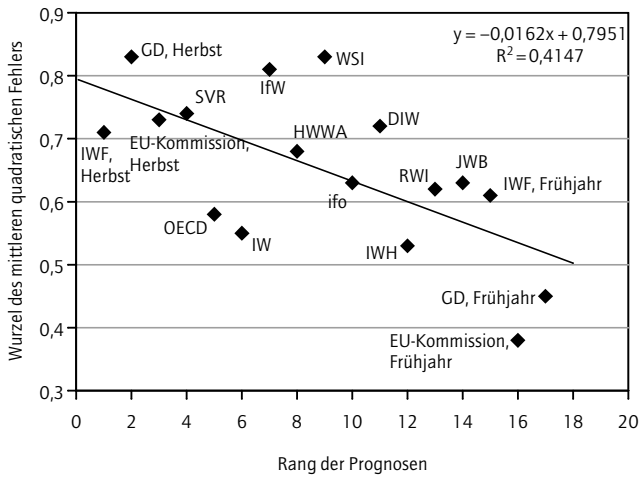
Abbildung 2

## Rang der Prognosen und Prognosefehler

### Wachstumsprognosen



### Inflationsprognosen



Quelle: Döpke und Fritsche (2006b).

Institutionen zu bewerten und welchen Kriterien von Rationalität entsprechen die untersuchten Prognosen?

Um diese Fragen zu beantworten, wurden zunächst einige gängige Prüfmaße berechnet.

#### 1. Der mittlere Fehler

$$ME = \frac{1}{T} \sum_{t=1}^T e_{t+1,t}, \text{ mit } e_{t+1,t} = y_{t+1} - \hat{y}_{t+1,t} \quad (1)$$

als Prognosefehler in jeder Periode wird 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 korrespondiert mit Unter- bzw. Überschätzungen der Variablen.

#### 2. Der mittlere absolute Fehler ergibt sich als

$$MAE = \frac{1}{T} \sum_{t=1}^T |e_{t+1,t}|. \quad (2)$$

#### 3. Die Wurzel des mittleren quadratischen Fehlers beträgt

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T e_{t+1,t}^2}. \quad (3)$$

Die Tabellen 1 und 2 enthalten die entsprechenden Kennzahlen für die Wachstums- und Inflationsprognosen. Die mittleren Fehler für die Wachstumsprognosen sind in den meisten Fällen negativ, was eine leichte Tendenz zur Wachstumsüberschätzung signalisiert. Einfache Regressionen der Fehler auf eine Konstante deuten allerdings darauf hin, dass die Überschätzungstendenz beim Wachstum in keinem Fall statistisch signifikant von null verschieden ist. Für die Inflationsprognosen liegt der mittlere Fehler in allen Fällen nahe null. Die Wurzeln der mittleren quadratischen Fehler der Prognosen sind relativ groß – zwischen 1,1 und 1,8 Prozentpunkten für Wachstumsprognosen und 0,5 und 1,2 Prozentpunkten für Inflationsprognosen.

### 4.2 Rationalitätstests

Außerdem wurden die Vorhersagen auf Unverzerrtheit und Effizienz geprüft. Ist eine Prognose sowohl unverzerrt als auch effizient, d.h. nutzt sie alle zum Prognosezeitpunkt zur Verfügung stehenden Informationen aus, wird auch von einer rationalen Prognose gesprochen. Als Test auf die Unverzerrtheit der Prognose wird häufig die so genannte Mincer-Zarnowitz-Regression angewendet. Dabei wird die Gleichung

$$y_{t+1} = \beta_0 + \beta_1 \hat{y}_{t+1,t} + u_{t+1} \quad (4)$$

mit  $y_{t+1}$  als realisiertem Wert und  $\hat{y}_{t+1,t}$  als der entsprechenden Prognose in Periode t für die Periode t+1 geschätzt (Diebold 2001: Kap. 11). Die Hypothese der Unverzerrtheit ist dann gleich bedeutend mit einer Konstanten von null und einem Koeffizienten bezüglich der

Tabelle 1

## Güte, Unverzerrtheit und Rationalität der Wachstumsprognosen

Institution	Mittlerer Fehler	Mittlerer absoluter Fehler	Wurzel des mittleren quadratischen Fehlers	Test auf Unverzerrtheit <sup>1</sup>	Test auf „schwache“ Rationalität <sup>1</sup>	Test auf „starke“ Rationalität unter Zuhilfenahme des kurzfristigen Zinssatzes <sup>1</sup>	Test auf „starke“ Rationalität unter Zuhilfenahme des realen Außenwertes <sup>1</sup>	Test auf „starke“ Rationalität unter Zuhilfenahme des Ölpreises <sup>1</sup>	Test auf „starke“ Rationalität unter Zuhilfenahme der OECD-Industrieproduktion <sup>1</sup>
IWF, Herbst	-0,71	1,39	1,83	0,08	0,96	0,01	0,14	0,09	0,13
GD, Herbst	-0,39	1,21	1,65	0,33	0,86	0,06	0,29	0,28	0,37
Europäische Kommission, Herbst	-0,46	1,22	1,67	0,26	0,88	0,02	0,33	0,21	0,24
Sachverständigenrat	-0,36	1,02	1,45	0,35	0,95	0,09	0,31	0,54	0,34
OECD	-0,35	1,09	1,58	0,42	0,60	0,08	0,49	0,30	0,29
IW	-0,34	1,17	1,71	0,30	0,14	0,09	0,48	0,30	0,42
IFW	-0,26	1,08	1,55	0,35	0,95	0,13	0,30	0,32	0,36
HWWA	-0,22	0,98	1,38	0,67	0,57	0,29	0,61	0,56	0,63
WSI	-0,14	1,15	1,58	0,51	0,71	0,12	0,44	0,33	0,38
ifo	-0,13	0,96	1,33	0,86	0,44	0,40	0,86	0,50	0,64
DIW Berlin	0,02	1,03	1,51	0,50	0,62	0,09	0,46	0,18	0,50
IWH	-0,40	0,75	1,00	0,42	0,17	0,64	0,23	0,61	0,28
RWI	-0,30	1,02	1,34	0,40	0,65	0,15	0,51	0,55	0,56
Jahreswirtschaftsbericht	-0,26	0,95	1,40	0,56	0,38	0,29	0,58	0,35	0,66
IWF, Frühjahr	-0,19	0,88	1,22	0,40	0,12	0,20	0,49	0,21	0,65
Europäische Kommission, Frühjahr	-0,23	0,83	1,11	0,42	0,44	0,24	0,67	0,25	0,76
GD, Frühjahr	-0,24	0,84	1,18	0,48	0,91	0,20	0,66	0,24	0,58

<sup>1</sup> Angegeben ist die Wahrscheinlichkeit einer falschen Ablehnung der Nullhypothese.

Für die Tests auf starke Rationalität wurden die exogenen Variablen mittelwertbereinigt, um Verzerrung bei der Schätzung der Konstanten zu vermeiden.

Quelle: Döpke und Fritsche (2006b).

Tabelle 2

**Güte, Unverzerrtheit und Rationalität der Inflationsprognosen**

Institution	Mittlerer Fehler	Mittlerer absoluter Fehler	Wurzel des mittleren quadratischen Fehlers	Test auf Unverzerrtheit <sup>1</sup>	Test auf „schwache“ Rationalität <sup>1</sup>	Test auf „starke“ Rationalität unter Zuhilfenahme des kurzfristigen Zinssatzes <sup>1</sup>	Test auf „starke“ Rationalität unter Zuhilfenahme des realen Außenwertes <sup>1</sup>	Test auf „starke“ Rationalität unter Zuhilfenahme des Ölpreises <sup>1</sup>	Test auf „starke“ Rationalität unter Zuhilfenahme der OECD-Industrieproduktion <sup>1</sup>
IWF, Herbst	-0,11	0,71	0,95	0,59	0,09	0,73	0,58	0,63	0,79
GD, Herbst	0,03	0,83	1,07	0,23	0,00	0,29	0,26	0,27	0,18
Europäische Kommission, Herbst	0,07	0,73	0,97	0,13	0,00	0,22	0,22	0,19	0,12
Sachverständigenrat	0,06	0,74	0,97	0,09	0,01	0,20	0,07	0,12	0,18
OECD	-0,05	0,58	0,79	0,85	0,30	0,89	0,74	0,95	0,95
IW	-0,03	0,55	0,77	0,27	0,26	0,27	0,29	0,44	0,44
IFW	0,19	0,81	1,20	0,22	0,00	0,16	0,21	0,20	0,18
HWWA	0,05	0,68	0,95	0,31	0,01	0,46	0,23	0,39	0,49
WSI	0,01	0,83	1,09	0,92	0,00	0,93	0,68	0,74	0,88
ifo	-0,02	0,63	0,87	0,66	0,18	0,77	0,50	0,73	0,76
DIW Berlin	0,21	0,72	0,95	0,17	0,02	0,33	0,29	0,27	0,31
IWH	0,04	0,53	0,74	0,22	0,42	0,00	0,22	0,42	0,30
RWI	0,13	0,62	0,86	0,03	0,00	0,08	0,02	0,02	0,07
Jahreswirtschaftsbericht	0,16	0,63	0,91	0,42	0,03	0,70	0,50	0,75	0,80
IWF, Frühjahr	-0,16	0,61	0,86	0,52	0,01	0,46	0,12	0,49	0,59
Europäische Kommission, Frühjahr	-0,05	0,38	0,48	0,62	0,15	0,78	0,80	0,66	0,59
GD, Frühjahr	0,05	0,45	0,60	0,52	0,19	0,61	0,33	0,86	0,55

<sup>1</sup> Angegeben ist die Wahrscheinlichkeit einer falschen Ablehnung der Nullhypothese.

Für die Tests auf starke Rationalität wurden die exogenen Variablen mittelwertbereinigt, um Verzerrung bei der Schätzung der Konstanten zu vermeiden.

Quelle: Döpke und Fritsche (2006b).



Prognose von 1. Mit anderen Worten: die Prognostiker machen keinen *systematischen* Fehler. Die Nullhypothese lautet also

$$H_0 : \beta_0 = 0 \wedge \beta_1 = 1. \quad (5)$$

Die entsprechende Zeile in der Tabelle enthält den p-Wert eines entsprechenden F-Tests.

Ein Test auf Effizienz der Prognose kann direkt aus der Mincer-Zarnowitz-Regression abgeleitet werden, indem diese um exogene Variablen erweitert wird. Es wird also die Gleichung

$$y_{t+1} = \beta_0 + \beta_1 \hat{y}_{t+1,t} + \beta_2 X_{t-1} + u_{t+1} \quad (6)$$

geschätzt, wobei  $X_{t-1}$  für eine zum Zeitpunkt der Prognoseerstellung bekannte exogene Variable steht. Ein F-Test dient dann zur Überprüfung der Hypothese

$$H_0 : \beta_0 = 0 \wedge \beta_1 = 1 \wedge \beta_2 = 0 \quad (7)$$

(siehe Holden und Peel 1990).

Für eine effiziente Prognose muss gelten, dass es nicht möglich ist, eine Variable zu finden, die die Prognosefehler systematisch erklärt. Man kann eine „schwache“ und eine „starke“ Version dieses Tests unterscheiden. Die „schwache“ Version des Tests bezieht sich auf einen Test auf Autokorrelation der Prognosefehler. Wenn keine Variable den Prognosefehler erklärt, sollten auch verzögerte Prognosefehler keinen Informationsgehalt für die Erklärung der heutigen Prognosefehler bieten (Kirchgässner 1994). Definiert man

$$y_{t+1} - \hat{y}_{t+1,t} = e_{t+1} \quad (8)$$

und schätzt die Gleichung

$$y_{t+1} = \beta_0 + \beta_1 \hat{y}_{t+1,t} + \beta_2 e_t + u_t \quad (9)$$

so lautet die Nullhypothese des Tests:

$$H_0 : \beta_0 = 0 \wedge \beta_1 = 1 \wedge \beta_2 = 0 \quad (10)$$

Die entsprechende Zeile in der Tabelle enthält wieder den entsprechenden p-Wert.

Verzögerte Prognosefehler stellen nur eine Teilmenge der dem Prognostiker zur Verfügung stehenden Informationsmenge dar. Eine „starke“ Version des Tests – oftmals auch als Tests auf „rationale Erwartungsbildung“ oder „starke“ Rationalität bezeichnet – prüft, ob Prognosefehler mit irgendeiner, zum Zeitpunkt der Prognose bekannten Information systematisch erklärt werden können. Wenn es möglich ist eine solche informative Variable, zu finden, dann kann die Prognose nicht als rational bezeichnet werden, weil nicht alle zur Verfügung stehenden Information effizient genutzt wurden. Einige Autoren (z.B. Tichy 1994) weisen darauf hin, dass es unklar ist, welche Informationen dem Prognostiker zum Zeitpunkt der Prognosen genau bekannt sind. So stehen am Jahresende, wo ein großer Teil der hier untersuchten Prognose erstellt wird, die Daten des vierten Quartals für die

meisten relevanten Zeitreihen noch gar nicht zur Verfügung. Deshalb wurden Variablen aus dem zweiten Quartal des Jahres benutzt, in welchem die Vorhersage erstellt wurde. Konkret wurde das Konzept der strengen Rationalität mit folgenden exogenen Variablen getestet (die entsprechende Zeile in der Tabelle enthält wieder den p-Wert):<sup>6</sup>

- kurzfristiger Zinssatz (Dreimonatsgeld),
- realer effektiver Wechselkurs,
- Ölpreisentwicklung
- Industrieproduktion in der OECD.

Die Testresultate – siehe Tabellen 1 und 2 – zeigen, dass Wachstumsprognosen im Allgemeinen den Kriterien von Unverzerrtheit und „schwacher“ Rationalität genügen. Bei den Inflationsprognosen wird das Kriterium der „schwachen“ Rationalität vielfach abgelehnt – was aber auf bestimmte Zeitreiheneigenschaften der Inflationsrate zurückzuführen sein kann.<sup>7</sup> Tests auf „starke“ Rationalität zeigen, dass in einigen Fällen exogene Variablen einen Informationsgehalt bezüglich der Prognosefehler haben, insbesondere scheinen monetäre Faktoren nicht von allen Prognostikern in ausreichendem Maße berücksichtigt zu werden. Dies entspricht auch den Ergebnissen von Hagen und Kirchgässner (2001), Kirchgässner und Savioz (2001) sowie Harvey (1991).

Gegen die Verwendung individueller Rationalitätstests ist eingewendet worden, dass die Anwendung von Panel-Schätzmethoden eine effizientere Testmethode sein würde (Keane und Runkle 1990). Deshalb wurde die erweiterte Mincer-Zarnowitz-Regression als Panel in der Form

$$y_{t+1,i} = \beta_0 + \beta_1 \hat{y}_{t+1,t,i} + \beta_2 X_{t-1} + u_{t+1,i} \quad (11)$$

mit  $i$  als Index für die Prognostiker geschätzt. Um ein balanciertes Panel zu erhalten, wurde das Institut für Wirtschaftsforschung in Halle aus der Berechnung herausgenommen und der Untersuchungszeitraum wieder auf die Periode 1974–2004 eingegrenzt.

Als Vergleichsmaßstab wurden bei Döpke und Fritsche (2006a) die Koeffizienten der Gleichung mit üblichen KQ-Schätzern (mit gemeinsamen und verschiedenen Konstanten für die Institutionen) geschätzt. Keane und Runkle (1990) argumentieren jedoch, dass bei der Auswahl der Schätzer eine mögliche Heteroskedastie zu berücksichtigen sei. Deshalb wurden von ihnen GLS-Schätzer, die das Problem nichtkonstanter Varianz im Querschnitt konsistent berücksichtigen, verwendet. Da eine enge Korrelation der Prognosefehler zwi-

**6** Alle nichtstationären Variablen wurden über Vorjahresvergleiche in stationäre Variablen überführt. In Bezug auf die Inflationsrate ist aber zweifelhaft, ob diese Transformation ausreicht. Möglicherweise ist das Preisniveau I(2) und damit auch noch die Inflationsrate instationär. Zumindest ist sie jedoch sehr persistent, so dass ein AR(1)-Prozess die Inflation bereits recht gut zu prognostizieren vermag.

**7** Dies liegt vor allem daran, dass die Inflationsrate die Zeitreiheneigenschaft einer ausgeprägt hohen Persistenz aufweist, d.h. Inflationsraten zwischen zwei aufeinanderfolgenden Zeitpunkten sehr stark miteinander korrelieren. Durchgeführte Tests zeigen, dass die Inflationsrate – zumindest für die hier betrachteten Jahreswerte – jedoch die Eigenschaft der Stationarität aufweisen. Das ist Bedingung für die hier verwendete Spezifikation der Tests.

schen den Prognostikern wahrscheinlich ist, wurden hier Schätzer verwendet, die zusätzlich auch mögliche Korrelation zwischen den Prognosen berücksichtigen („cross section SUR weighted least squares“, sog. Parks-Schätzer). Ein zusätzlicher Vorteil der Panel-Schätzungen liegt darin, dass das Problem der unterschiedlichen Prognosezeitpunkte über Dummies angemessen berücksichtigt werden kann.

Die Resultate der bei Döpke und Fritsche (2006a) ausführlich dargelegten Panel-Regressionen bestätigen die Schlussfolgerungen, die aus den auf die einzelnen Institutionen bezogenen Tests abgeleitet wurden: Es gibt auch in der Paneldimension und unter Berücksichtigung der unterschiedlichen Prognosezeitpunkte Hinweise für eine Verletzung der strengen Rationalitätskriterien, insbesondere bei Einbeziehung monetärer Variablen.<sup>8</sup>

### 4.3 Zur Qualität von Richtungsprognosen

Eine im Zusammenhang mit Konjunkturprognosen häufig gestellte Frage lautet: Liefern uns die Prognosen Informationen bezüglich der Richtungsänderung des Wachstums bzw. der Inflation? Um diese Frage zu untersuchen, wurden die Prognosen nach „Beschleunigung“ und „Verlangsamung“ in einer Kontingenztabelle klassifiziert (vgl. Diebold und Lopez 1996: 257). Das Symbol „ $ii$ “ steht zum Beispiel für eine prognostizierte Beschleunigung, die auch eingetreten ist, „ $ij$ “ für eine prognostizierte aber nicht eingetretene Beschleunigung.

Mit dieser Klassifikation – siehe Tabelle 3 – kann der Informationsgehalt anhand der tatsächlichen Zellbelegungen nach dem Ausdruck

$$\frac{O_{ii}}{O_{ii} + O_{ji}} + \frac{O_{jj}}{O_{jj} + O_{ij}} \quad (12)$$

gemessen werden. Im Falle eines reinen Münzwurfs ergäbe 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. Zusätz-

Tabelle 3

#### Klassifikation von Richtungsfehlern

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

Quelle: Diebold und Lopez (1996).

<sup>8</sup> Allerdings konnten die Tests nicht mit echten „Realtime“-Daten für die Variablen durchgeführt werden. Insofern kann nicht rigoros getestet werden, ob sich die Verletzung der Rationalität nicht erst im nachträglichen Blick auf die Daten zeigt. Jedoch ist auch zu bedenken, dass die Revisionsanfälligkeit von Zinsen null ist.

lich 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 Kontingenztafel lautet

$$\hat{E}_{ij} = O_{i \cdot} O_{\cdot j} / O. \quad (13)$$

Damit kann eine Maßzahl (Pearsons  $\chi^2$ ) nach der Formel

$$\sum_{i,j=1}^2 (O_{ij} - \hat{E}_{ij})^2 / \hat{E}_{ij} \quad (14)$$

konstruiert werden, die unter der Nullhypothese keines Informationsgehaltes  $\chi^2$ -verteilt ist. Die Tabelle 4 enthält den Informationsgehalt sowie die Wahrscheinlichkeit einer falschen Ablehnung der erläuterten Nullhypothese.

Die Ergebnisse zeigen ein für die Prognostiker insgesamt erfreuliches Ergebnis: Alle Prognosen haben über einen längeren Zeitraum einen systematisch höheren Informationsgehalt als auf Zufall basierende Vergleichsprognosen Prognosen („Münzwurf“). Die Ergebnisse der durchgeführten  $\chi^2$ -Tests bestätigen die statistische Signifikanz dieser Aussage.

Tabelle 4

### Auswertung von Richtungsprognosen

	Wachstumsprognosen		Inflationsprognosen	
	Informationsgehalt	Unabhängigkeitstest <sup>1</sup>	Informationsgehalt	Unabhängigkeitstest <sup>1</sup>
IWF, Herbst	1,47	0,01	1,62	0,00
GD, Herbst	1,34	0,02	1,46	0,00
Europäische Kommission, Herbst	1,39	0,02	1,58	0,00
Sachverständigenrat	1,31	0,05	1,51	0,00
OECD	1,65	0,00	1,72	0,00
IW	1,78	0,00	1,33	0,01
IFW	1,41	0,01	1,51	0,00
HWWA	1,36	0,03	1,51	0,00
WSI	1,41	0,01	1,45	0,00
ifo	1,47	0,00	1,40	0,01
DIW Berlin	1,47	0,00	1,50	0,00
IWH	1,83	0,00	1,43	0,03
RWI	1,36	0,03	1,39	0,00
Jahreswirtschaftsbericht	1,41	0,01	1,52	0,00
IWF, Frühjahr	1,39	0,01	1,47	0,00
Europäische Kommission, Frühjahr	1,53	0,00	1,76	0,00
GD, Frühjahr	1,52	0,00	1,76	0,00

<sup>1</sup> Angegeben ist die Wahrscheinlichkeit einer falschen Ablehnung der Nullhypothese.

Quelle: Döpke und Fritsche (2006a).

## 5 Zur Streuung der Prognosen

Neben dem Erklärungsgehalt der einzelnen Prognosen interessiert auch, wann und warum Prognosen differieren. Schließlich wird häufig vermutet, dass die meisten professionellen Prognostiker einen ähnlichen Kenntnisstand haben und weitgehend gleiche Methoden verwenden. Dann kann es den Nutzer dieser Art von Prognosen schon verwundern, wenn – wie in der Vergangenheit durchaus der Fall – Prognostiker in ihrer Einschätzung der künftigen Entwicklung stark unterschiedliche Meinungen vertreten.

Es gibt bisher in der Literatur keinen Konsens, warum Prognosen divergieren und eine beträchtliche Querschnittsstreuung aufweisen. Typischerweise wird beispielsweise die Unsicherheit über das Eintreffen der Vorhersagen eher mit individuellen Prognosen verknüpft (Zarnowitz und Lambros 1987). Allerdings argumentiert beispielsweise Linden (2003), dass die Varianz im Querschnitt der Prognosen ein brauchbares Maß für die allgemein herrschende Unsicherheit sei. Insofern könnte die Dispersion von Prognosen als Maß für die Prognoseunsicherheit dienen und dem Nutzer eine wichtige Entscheidungshilfe geben. Allerdings ist diese Interpretation der Querschnitt-Streuung nicht unumstritten: Bomberger (1996) sowie Rich und Tracy (2003) beispielsweise bezweifeln die Stichhaltigkeit entsprechender Argumente.

Die Abbildung 3 zeigt einige Maße der Streuung der Prognosen im Querschnitt der Prognostiker: die Standardabweichung und – als robustes Maß – den Interquartilsabstand. Zum Vergleich wurde außerdem die unbedingte *Volatilität* des Wachstums bzw. der Inflation (jeweils in einem rollenden 5-Jahresfenster gemessen) eingetragen.

Es fällt ein deutlicher Rückgang der unbedingten Volatilität der betrachteten Größen auf, während sich die Dispersion der Prognosen kaum verändert zu haben scheint. Schon dies weckt Zweifel am Argument, dass die Querschnittsvarianz als Unsicherheitsmaß für die Prognosen dienen könnte. Die Abbildung 3 zeigt auch, dass die Querschnittsstreuung deutlich kleiner ist als die Volatilität der zu prognostizierenden Variablen. Berücksichtigt man, dass die Streuung der Vorhersagen nicht mit dem folgenden Prognosefehler korreliert, wie entsprechende Tests bei Döpke und Fritsche (2006a) gezeigt haben, so wird deutlich, dass die Kennzahlen nicht als Maß für die allgemeine Prognoseunsicherheit taugen.

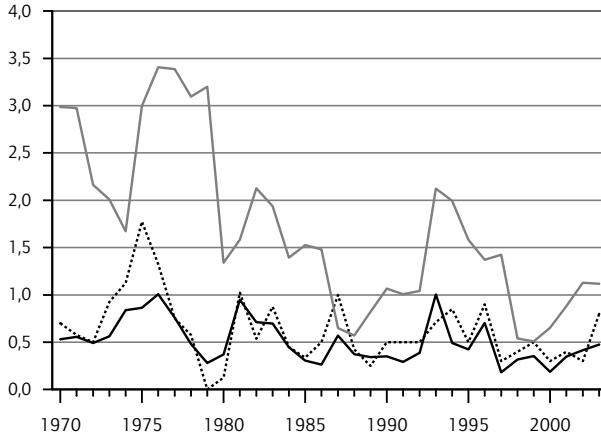
Allerdings fällt bei der Betrachtung von Abbildung 3 auch ins Auge, dass zumindest bei Wachstumsprognosen Phasen hoher Dispersion mit konjunkturellen Schwächephase bzw. mit Rezessionen einhergehen – also genau in diesen Perioden die Meinungen der Experten deutlich auseinander gehen. Abbildung 4 zeigt diesen Zusammenhang: Insbesondere bei Wachstumsprognosen gehen Phasen hoher Dispersion mit Rezessionen einher, während sich bei Inflationsprognosen kein klares Bild ergibt.<sup>9</sup>

<sup>9</sup> Zur Einteilung der Konjunktur wurde auf die Datierung „Economic Cycle Research Institute“ zurückgegriffen, welches für zahlreiche Länder eine der NBER-Prozedur vergleichbare Einteilung der Wirtschaftsentwicklung in Phasen von Hochkonjunktur und Rezession bereitstellt. Um die monatlich vorliegenden Daten für Jahresdaten vergleichbar zu machen, wurden diejenigen Jahre als Rezessionsjahre gezählt, in denen mehr als sechs Monate lang eine Rezession ausgewiesen war. Als Konzept diente das traditionelle Konzept eines Konjunkturzyklus, bei dem der absolute Rückgang/Anstieg der Wirtschaftsleistung eine Rezession/Boom-Klassifikation erlaubt.

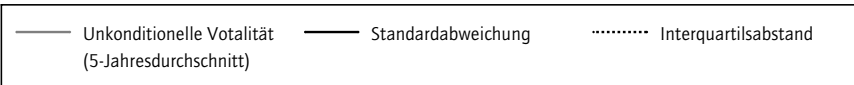
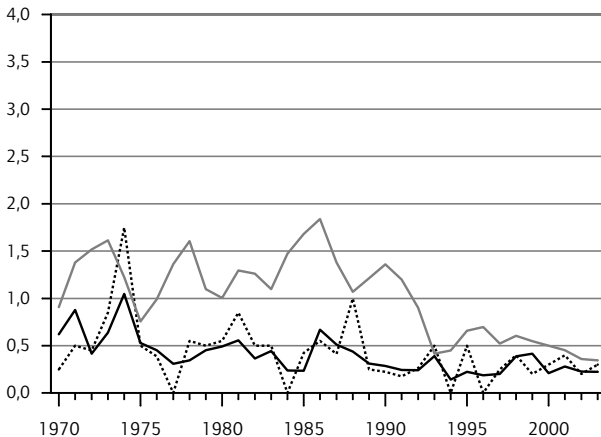
Abbildung 3

### Streuung der Prognosen

#### Wachstumsprognosen



#### Inflationsprognosen



Quelle: Döpke und Fritsche (2006a).

Der Augenschein wird auch durch eine Regressionsanalyse bestätigt, in der Dummy-Variablen für Rezessionsperioden genutzt werden, um die Querschnittsverteilung der Prognostiker zu erklären. Da die endogene Variable nur positive Werte annehmen kann, wurde dazu für den Logarithmus der Querschnittsstandardabweichung der Prognosen folgende

Regression geschätzt, wobei ein Trend zur Berücksichtigung der abnehmenden Volatilität aufgenommen wurde:

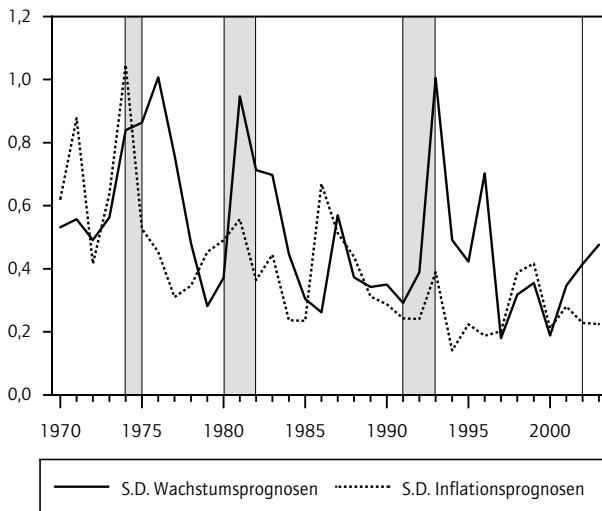
$$\log(\sigma_t^f) = \beta_0 + \beta_1 D_{t-1}^{\text{Re zession}} + \beta_2 D_t^{\text{Re zession}} + \beta_3 D_{t+1}^{\text{Re zession}} + \beta_4 \text{Trend} + u_t \quad (15)$$

Die Ergebnisse in Tabelle 5 – die Residuen sind nach allen Tests unproblematisch – zeigen, dass sich die Prognostiker zumindest bei den Wachstumsprognosen an Wendepunkten tendenziell weniger darüber einig sind, ob und wann die nächste Phase des Konjunkturzyklus eingeleitet wird. Anders formuliert: Solange die Konjunktur im ruhigen Fahrwasser verläuft, ist die Prognose eines unbedingten Erwartungswertes für das Wachstum durchaus sinnvoll und nachvollziehbar – und viele Prognostiker können sich darauf einigen. Werden die Zeiten „turbulenter“, so sind die Einschätzungen über die der konjunkturellen Dynamik zugrunde liegenden Prozesse deutlich unterschiedlicher. Im Folgenden soll kurz auf zwei Aspekte der möglichen Dispersion von Prognosen im Querschnitt eingegangen werden, die ausführlich bei Döpke und Fritsche (2006a) analysiert worden sind: die Beziehung zwischen der Dispersion und makroökonomischen Variablen sowie zwischen der Dispersion und der Volatilität makroökonomischer Variablen.<sup>10</sup>

- *Prognosedispersion und makroökonomische Variablen:* Es gibt eine Tendenz zur Uneinigkeit bei den Prognostikern in Phasen hoher Inflation und hoher Nominal- wie Realzinsen. Hier scheinen unterschiedliche Einschätzungen über die Wirkungskanäle von Geldpolitik eine Rolle zu spielen. Bezüglich der Inflationseinschätzung erweisen sich

Abbildung 4

### Streuung der Prognosen und Konjunkturzyklus



Quelle: Döpke und Fritsche (2006a).

<sup>10</sup> Die exogenen makroökonomischen Variablen sind mit dem Informationsstand des 2. Quartals des ersten Prognosejahres aufgenommen, d.h. mit hoher Sicherheit allen Prognostikern zum Zeitpunkt der Prognoseerstellung bekannt.

ein hohes Wachstum von Reallöhnen und wiederum hohe Nominal- wie Realzinsen sowie die Inflationsrate selbst als hilfreich zur Erklärung der systematischen Komponente der Dispersion zwischen den Prognostikern. Dieses Resultat stimmt mit der These überein, dass eine divergierende Einschätzung über die Dynamik von Inflationsprozessen zur Divergenz zwischen den Prognostikern beiträgt.

- *Prognosedispersion und Volatilität makroökonomischer Variablen*: Auch die Volatilität makroökonomischer Variablen ist positiv mit der Dispersion der Prognosen korreliert. Auffällig ist, dass dies insbesondere für die Volatilität von kurzfristigen Zinssätzen gilt. Offenbar teilen die prognostizierenden Institutionen keine einheitliche Einschätzung über die Auswirkungen von Geldpolitik auf die Konjunktur.

Ausgehend von diesen Ergebnissen stellt sich zudem die praktische Frage, ob es sinnvoll ist, immer auf die Mehrheit der Prognostiker zu vertrauen oder lieber einer abweichenden Minderheit Gehör zu schenken. Hier kann noch einmal ein Blick auf die in Abschnitt 4.3 erläuterte Analyse der Richtungsprognosen hilfreich sein. Dazu musste die Situation einer Minderheitenprognose definiert werden. Entsprechend gewertet wurde eine Situation, in der mindestens zwei Prognostiker eine Beschleunigung/Abschwächung voraussagen, während die anderen Prognostiker von einer Abschwächung/Beschleunigung ausgehen.

Aus der Tabelle 6 geht hervor, dass es immerhin neun respektive 14 Jahre gab, in denen eine Minderheit der Prognostiker Richtungsprognosen bezüglich Wachstum und Inflation abgaben, die denen der Mehrheit entgegenstanden. Die Resultate bezüglich des Informationsgehaltes unterscheiden hinsichtlich beider prognostizierter Variablen. Bei den Wachstumsprognosen erreichten die Minderheitsvoten einen nahezu gleichen Informationsgehalt wie die Mehrheitsprognosen. Im Durchschnitt und nur bei Betrachtung dieses Kriteriums

Tabelle 5

### Regressionsergebnisse zu Streuung und Konjunkturzyklus 1970–2004

	Streuung der Wachstumsprognosen	Streuung der Inflationsprognosen
Konstante	-0,28 (0,13)**	0,17
Rezessionsdummy (t-1)	0,43 (0,08)***	-0,15 (0,15)
Rezessionsdummy (t)	0,20 (0,08)***	0,21 (0,10)*
Rezessionsdummy (t+1)	-0,27 (0,13)*	-0,01 (0,13)
Trend	-0,02 (0,004)***	-0,03 (0,005)***
R2	0,55	0,52

t-Werte in Klammern, Korrektur nach Newey-West.

\*\*\* (\*\*, \*) kennzeichnet Ablehnung der Nullhypothese auf 1 % (5 %, 10 %) Niveau.

Quelle: Döpke und Fritsche (2006a).



Tabelle 6

**Mehrheit und Minderheit bei Richtungsprognosen**

1970–2004

<b>Wachstumsprognosen (Anzahl der Minderheitsvoten: 9)</b>		
<b>Fehler Vorhersage / Ergebnis</b>	<b>Mehrheit</b>	<b>Minderheit <sup>1</sup></b>
Beschleunigung / Beschleunigung (II)	2 [16]	1 [2]
Beschleunigung / Abschwächung (IJ)	4 [3]	2 [6]
Abschwächung / Beschleunigung (JI)	2 [6]	4 [5]
Abschwächung / Abschwächung (JJ)	1 [2]	2 [3]
Informationsgehalt: II/(II+IJ) + JJ/(JJ+JI)	1,00 [1,12]	1,00 [0,88]
<b>Inflationsprognosen (Anzahl der Minderheitsvoten: 14)</b>		
<b>Fehler Vorhersage / Ergebnis</b>	<b>Mehrheit</b>	<b>Minderheit <sup>1</sup></b>
Beschleunigung / Beschleunigung (II)	4 [6]	2 [3]
Beschleunigung / Abschwächung (IJ)	2 [2]	6 [10]
Abschwächung / Beschleunigung (JI)	5 [10]	2 [2]
Abschwächung / Abschwächung (JJ)	3 [3]	4 [6]
Informationsgehalt: II/(II+IJ) + JJ/(JJ+JI)	1,29 [1,52]	0,71 [0,49]

<sup>1</sup> Als Minderheitenvotum wurde eine Situation gewertet, in der mindestens zwei Institutionen [in Klammern: eine] nicht mit der Mehrheit der Prognosen bezüglich prognostizierter Beschleunigung/Abschwächung übereinstimmten.

Quelle: Döpke und Fritsche (2006a).

dürfte sich der geneigte Nutzer von Prognosen fragen, welche Prognose er nun wählt.<sup>11</sup> Bei den Inflationsprognosen ist die Mehrheitsprognose deutlich besser.

Zwar sind die Maßzahlen für den Informationsgehalt nicht sehr hoch, aber es scheint – ohne dass diese Hypothese wegen des mangelnden Beobachtungsumfangs einem rigorosen statistischen Test unterzogen werden konnte – die Minderheit „besser“ zu sein, wenn es um die Prognose von Rezessionen geht. Nehmen wir ein negatives Wachstum in einem Jahr als Kriterium für das Vorliegen einer Rezession an, so wurden bei den drei so klassifizierten Rezessionen im vorliegenden Untersuchungszeitraum von der Mehrheit keine einzige erkannt. In zwei von drei Fällen wurde von der Mehrheit sogar eine Beschleunigung der vor der Rezession ohnehin positiven Wachstumsrate prognostiziert.

Unterschiedliche Ansätze und konkurrierende Prognosen haben also durchaus einen „Sinn“, wenn es um Konjunktur – insbesondere konjunkturelle Wendepunkte – geht. Gerade in diesen Phasen zahlt es sich aus, eine Pluralität von Konjunkturprognosen zur Ver-

<sup>11</sup> Allerdings ist bei dem strengen Kriterium von zwei Minderheitsinstitutionen der Informationsgehalt in keinem Fall besser als der eines Münzwurfs.

fügung zu haben. Zukünftige Forschungsarbeiten sollten sich auch auf die Frage konzentrieren, wie die Informationen aus dem verfügbaren Spektrum von Prognosen in einer der Öffentlichkeit verständlicheren Form besser zugänglich gemacht werden können.

## 6 Zusammenfassung

Alles in allem lassen sich einige vorsichtige Schlussfolgerungen über die Qualität von Prognosen und die Aussagekraft der Dispersion von Vorhersagen treffen. Konjunkturprognosen haben, gemessen an den Erwartungen, eine erstaunlich hohe Fehlermarge. Die Wurzel des mittleren quadrierten Fehlers ist mindestens einen Prozentpunkt groß, für viele Institutionen eher zwischen 1,5 und 1,8 Prozentpunkte. Zwar mag dieser Fehler in den vergangenen zwei Jahrzehnten leicht abgenommen haben (vgl. Heilemann und Stekler 2005), er ist aber immer noch beträchtlich. Hier scheint neben weiteren Bemühungen zur Reduktion des Fehlers auch eine weitgehende Aufklärung der öffentlichen Meinung notwendig zu sein. Eine realistische Einschätzung der Höhe des Prognosefehlers auf Seiten der Öffentlichkeit könnte mancher allzu aufgeregten Diskussion über die Revision von Prognosen den Boden entziehen und zur Stabilisierung positiver Erwartungen einen Beitrag leisten.

In diesem Sinne wäre es wünschenswert, wenn es gelänge, die, in der wissenschaftlichen Prognoseliteratur zu einem wichtigen Thema gewordene, Offenlegung und Einbeziehung der Prognoseunsicherheit zu popularisieren. So werden zunehmend Konfidenzbereiche von Prognosen von wirtschaftspolitischen Institutionen veröffentlicht (siehe z.B. die Darstellung der Inflationsprognose der Bank of England über ein „Fan-Chart“ in ihren Monatsberichten). Konstruiert man auf der Basis der historischen Fehler Konfidenzbänder im üblichen statistischen Sinne für die hier untersuchten Prognosen, dann ergäben sich recht große Intervalle (Fritsche 2005, Kappler 2005). Dies wird von manchen Autoren als schwer vermittelbar angesehen (Kappler 2005: 9). Nichtsdestotrotz wäre es wünschenswert, die Unsicherheitsbereiche der Prognosen offen zu legen. Ein Weg bestünde in der Befragung der Prognostiker, so wie es im *Survey of Professional Forecaster* in den USA oder bei der Expertenbefragung der EZB erfolgt. Ein solches Vorgehen würde einerseits zu einer transparenteren Erwartungsbildung in der Öffentlichkeit beitragen, andererseits auch eine kritische und rationale Prognoseüberprüfung erleichtern.

Die Ergebnisse der Rationalitätstests führen einerseits zu einem vorsichtig optimistischen Schluss über die Qualität der Prognosen. Danach sind offenbar systematische Prognosefehler nicht nachweisbar. Dennoch gibt es andererseits zum Zeitpunkt der Prognoseerstellung verfügbare Informationen, die die Prognosegüte systematisch verbessern könnten. Es gibt deutliche Hinweise darauf, dass monetäre Schocks oder geldpolitische Richtungswechsel (gemessen vor der Prognoseerstellung) mit künftigen Prognosefehlern einhergehen, die durch eine stärkere Berücksichtigung der Wirkung von Geldpolitik vermeidbar gewesen wären. Weitere Forschungen über die Ursachen von Verletzungen der Hypothese rationaler Erwartungen könnten Hinweise auf ökonomische Ursachen von Prognosefehlern liefern. Die Prognosedispersion variiert mit dem Konjunkturzyklus, was auf eine deutlich unterschiedliche Einschätzung der konjunkturellen Dynamik an Wendepunkten durch die Prognostiker hindeutet. Die Regressionsanalyse zeigt zudem einen Zusammenhang zwischen Höhe und Volatilität von makroökonomischen Variablen und der Dispersion von Prognosen. Insbesondere die Wirkungen der Geldpolitik werden offenbar sehr un-

terschiedlich eingeschätzt und führen zu divergierenden Erwartungen. Dieses Resultat ist insofern interessant, als es einem sehr einfachen Verständnis von rationalen Erwartungen widerspricht, welches von einem „einheitlichen“ Verständnis der Erwartungsbildung für alle Akteure ausgeht. Zudem zeigen die Ergebnisse, dass es keineswegs immer sinnvoll ist, sich an der Prognose der Mehrheit zu orientieren. Dies zeigt, dass Wettbewerb auch unter Prognostikern sinnvoll ist.

Alles in allem bleibt Prognose ein wichtiges Feld für Wirtschaftsforscher. In ihrer Prognosefähigkeit zeigt sich letztlich die Angemessenheit ökonomischer Theorien zur Erklärung empirischer Phänomene – ohne allerdings die Grenzen der Prognosefähigkeit außer Kraft setzen zu können.

### Literaturverzeichnis

- Bomberger, W.A. (1996): Disagreement as a Measure of Uncertainty. *Journal of Money, Credit, and Banking*, 28, 318–392.
- Der Spiegel (2005): Blindflug der Forscher. *Der Spiegel* vom 2. Mai 2005, 94.
- Diebold, F.X. (2001): Elements of Forecasting. 2. Aufl. Cincinnati, OH, South-Western.
- Diebold, F.X. und J. Lopez (1996): *Forecast Evaluation and Combination*. NBER Technical Working Paper 192. Download unter: [www.nber.org](http://www.nber.org)
- Döpke, J. und U. Fritsche (2004): Growth and Inflation Forecasts in Germany – An Assessment of Accuracy and Dispersion. DIW Discussion Paper 399. Download unter: [www.diw.de/deutsch/produkte/publikationen/diskussionspapiere/docs/papers/dp399.pdf](http://www.diw.de/deutsch/produkte/publikationen/diskussionspapiere/docs/papers/dp399.pdf)
- Döpke, J. und U. Fritsche (2005): Forecast Errors and the Macroeconomy – A Non-Linear Relationship? DIW Discussion Paper 498. Berlin.
- Döpke, J. und U. Fritsche (2006a): When Do Forecasters Disagree? An Assessment of German Growth and Inflation Forecast Dispersion. *International Journal of Forecasting*, 22 (1), 125–135.
- Döpke, J. und U. Fritsche (2006b): Growth and Inflation Forecasts for Germany. A Panel-based Assessment of Accuracy and Efficiency. *Empirical Economics*. Download unter: <http://dx.doi.org/10.1007/s00181-005-0050-5>
- Döpke, J. und U. Fritsche (2006c): Shocking! Do Forecasters Share a Common Belief? *Applied Economics Letters* (im Erscheinen).
- FAZ (2005a): Entlasst die Experten. *Frankfurter Allgemeine Zeitung* vom 19. April 2005, 33.
- FAZ (2005b): Deutlich besser als ein Münzwurf. *Frankfurter Allgemeine Zeitung* vom 6. Mai 2005, 12.
- Fritsche, U. (2005): Warum Konjunkturprognosen? *Wochenbericht des DIW Berlin*, Nr. 22/2005, 361–369.
- Hagen, H.M. und G. Kirchgässner (1997): *Interest Rate Based Forecasts vs. Professional Forecast for German Economic Growth: A Comparison*. Discussion Paper No. 9714. Universität St. Gallen.
- Harvey, C.R. (1991): Interest Rate Based Forecasts of German Economic Growth. *Weltwirtschaftliches Archiv*, 127, 701–718.
- Heilemann, U. und H.O. Stekler (2003): Have Macroeconomic Forecasts Improved? The Case of Germany 1967 to 2001. SFB 475 Universität Dortmund – Technical Report 31/2003. Dortmund.

- Holden, K. und D.A. Peel (1990): On Testing for Unbiasedness and Efficiency of Forecasts. *The Manchester School*, 18 (2), 120–127.
- Kappler, M. (2006). Wie genau sind die Konjunkturprognosen der Institute für Deutschland? ZEW Discussion Paper Nr. 06-004. Mannheim.
- Keane, M.P. und D.E. Runkle (1990). Testing the Rationality of Price Forecasts: New Evidence from Panel Data. *American Economic Review*, 80, 714–735.
- Kirchgässner, G. (1984): Sind die Erwartungen der Wirtschaftssubjekte rational? Eine empirische Untersuchung für die Bundesrepublik Deutschland. *Weltwirtschaftliches Archiv*, 120, 279–300.
- Kirchgässner, G. und M. Savioz (2001): Monetary Policy and Forecasts for Real GDP Growth: An Empirical Investigation for the Federal Republic of Germany. *German Economic Review*, 2 (4), 339–365.
- Linden, S (2003). *Assessment of GDP Forecasts Uncertainty*. European Commission Economic Papers 184. Brüssel.
- Löwe, A. (1929): Besprechung von O. Morgenstern. *Zeitschrift für die gesamte Staatswissenschaft*, 2, 419–423.
- Morgenstern, O. (1928): *Wirtschaftsprognose. Eine Untersuchung ihrer Voraussetzungen und Möglichkeiten*. Wien.
- Stege, C. (1989): *Zur Beurteilung der Prognosetätigkeit der „führenden Wirtschaftswissenschaftlichen Forschungsinstitute“ in der Bundesrepublik Deutschland*. Hamburg.
- Tichy, G. (1994): *Konjunktur: stilisierte Fakten, Theorie, Prognose*. 2., völlig neubearb. Aufl. Berlin.
- Tietzel, M. (1989): Prognoselogik oder: Warum Prognostiker irren dürfen. *Jahrbücher für Nationalökonomie und Statistik*, 206 (6), 546–562.
- Zarnowitz, V. und L.A. Lambros (1987). Consensus and Uncertainty in Economic Prediction. *Journal of Political Economy*, 95, 591–621.