

Discussion Paper No. 06-004

**Wie genau sind die Konjunkturprognosen
der Institute für Deutschland?**

Marcus Kappler

ZEW

Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

Discussion Paper No. 06-004

Wie genau sind die Konjunkturprognosen der Institute für Deutschland?

Marcus Kappler

Download this ZEW Discussion Paper from our ftp server:

<ftp://ftp.zew.de/pub/zew-docs/dp/dp06004.pdf>

Die Discussion Papers dienen einer möglichst schnellen Verbreitung von neueren Forschungsarbeiten des ZEW. Die Beiträge liegen in alleiniger Verantwortung der Autoren und stellen nicht notwendigerweise die Meinung des ZEW dar.

Discussion Papers are intended to make results of ZEW research promptly available to other economists in order to encourage discussion and suggestions for revisions. The authors are solely responsible for the contents which do not necessarily represent the opinion of the ZEW.

Das Wichtigste in Kürze

Konjunkturprognosen werden regelmäßig von Instituten, Banken und Versicherungen veröffentlicht, wie beispielsweise im Rahmen der halbjährlichen Gemeinschaftsdiagnose der Wirtschaftsforschungsinstitute. Vor allem die sogenannten Punktprognosen zu wichtigen gesamtwirtschaftlichen Kennzahlen erhalten dabei besondere Aufmerksamkeit. Die statistische Unschärfe, die mit diesen Prognosen verbunden ist, wird allerdings kaum kommuniziert.

Die vorliegende Arbeit diskutiert Ursachen für Unsicherheiten von Konjunkturprognosen und demonstriert die Berechnung von empirischen Prognoseintervallen. Die Verwendung empirischer Prognoseintervalle für eine Beurteilung der Signifikanz von Prognoserevisionen wird für den Median der Prognosen vier deutscher Wirtschaftsforschungsinstitute für die Entwicklung folgender volkswirtschaftlicher Kennziffern exemplarisch veranschaulicht: Reales Bruttoinlandsprodukt, private und öffentliche Konsumausgaben, Ausrüstungsinvestitionen, Ausfuhren und Einfuhren, Zahl der Erwerbstätigen sowie Verbraucherpreise. Der Analyse liegen Beobachtungen von 1980 bis 2004 zugrunde. Es zeigt sich, dass vor allem geringfügige Prognoserevisionen meistens keine wesentlich neuen Einschätzungen hinsichtlich der zukünftigen konjunkturellen Entwicklung signalisieren.

Für eine Beurteilung der grundsätzlichen Prognoseschwierigkeit wurden die Medianprognosen der Institute in einem weiteren Untersuchungsschritt alternativen Zeitreihenmodellen vom Typ ARIMA gegenübergestellt. Die Institutsprognosen weisen gegenüber den Zeitreihenmodellen durchweg eine höhere Güte auf, wenn die Wurzel des mittleren quadratischen Prognosefehlers als Maßstab herangezogen wird. Allerdings sind deutliche Unterschiede in der Prognosepräzision der betrachteten volkswirtschaftlichen Kennziffern feststellbar.

Wie genau sind die Konjunkturprognosen der Institute für Deutschland?

Marcus Kappler*

Zentrum für Europäische Wirtschaftsforschung (ZEW) GmbH †

Dezember 2005

Zusammenfassung

Die vorliegende Arbeit diskutiert Ursachen für Unsicherheiten von Konjunkturprognosen und demonstriert die Berechnung von empirischen Prognoseintervallen. Die Verwendung empirischer Prognoseintervalle für eine Beurteilung der Signifikanz von Prognoserevisionen wird für den Median der Prognosen vier deutscher Wirtschaftsforschungsinstitute für die Entwicklung folgender volkswirtschaftlicher Kennziffern exemplarisch veranschaulicht: Reales Bruttoinlandsprodukt, private und öffentliche Konsumausgaben, Ausrüstungsinvestitionen, Ausfuhren und Einfuhren, Zahl der Erwerbstätigen sowie Verbraucherpreise. Der Analyse liegen Beobachtungen von 1980 bis 2004 zugrunde. Es zeigt sich, dass vor allem geringfügige Prognoserevisionen meistens keine wesentlich neuen Einschätzungen hinsichtlich der zukünftigen konjunkturellen Entwicklung signalisieren. Für eine Beurteilung der grundsätzlichen Prognoseschwierigkeit wurden die Medianprognosen der Institute in einem weiteren Untersuchungsschritt alternativen Zeitreihenmodellen vom Typ ARIMA gegenübergestellt. Die Institutsprognosen weisen gegenüber den Zeitreihenmodellen durchweg eine höhere Güte auf, wenn die Wurzel des mittleren quadratischen Prognosefehlers als Maßstab herangezogen wird. Allerdings sind deutliche Unterschiede in der Prognosepräzision der betrachteten volkswirtschaftlichen Kennziffern feststellbar.

*Besonders möchte ich Prof. Dr. Dr. h.c. mult. Wolfgang Franz danken, der den Anstoß für die Untersuchung von Prognoserevisionen gab. Außerdem möchte ich mich bei Steffen Osterloh für die Unterstützung bei der Prognoseauswertung sowie bei Dr. Friedrich Heinemann und Dr. Michael Schröder für hilfreiche Kommentare und Anregungen bedanken. Alle verbleibenden Unzulänglichkeiten gehen zu meinen Lasten.

†Postfach 103443, 68034 Mannheim, E-Mail: kappler@zew.de

1 Einleitung

Konjunkturprognosen werden regelmäßig von Instituten, Banken und Versicherungen veröffentlicht, wie beispielsweise im Rahmen der halbjährlichen Gemeinschaftsdiagnose der Wirtschaftsforschungsinstitute. Vor allem die sogenannten Punktprognosen zu wichtigen gesamtwirtschaftlichen Kennzahlen erhalten dabei besondere Aufmerksamkeit. Die statistische Unschärfe, die mit diesen Prognosen verbunden ist, wird allerdings kaum kommuniziert. Dass Konjunkturprognosen mit Unsicherheit behaftet sind, liegt in der Natur der Sache. Erstaunlich ist allerdings, dass die Prognostiker in der Regel keine expliziten Schwankungsbreiten ihrer Einschätzungen angeben, obwohl dafür eigentlich Methoden der Statistik zur Verfügung stehen. Konjunkturprognosen sind für eine Reihe von Wirtschaftsakteuren, deren Entscheidungen maßgeblich von zukünftigen Entwicklungen abhängen, von großem Interesse. Als typische Beispiele seien die Haushaltsplanungen des Staates sowie Investitionsentscheidungen von Unternehmen genannt. Für die Empfänger von Prognosen kann es wichtig sein, die Unsicherheiten, mit denen die Prognosen behaftet sind, zu kennen.

Borbély und Meier (2005) schlagen eine Bootstrap-Methode zur Erfassung der Unsicherheit modellbasierter Prognosen unter Berücksichtigung von Modellunsicherheiten vor. Am Beispiel von Prognoseintervallen für die Entwicklung der Konsumentenpreisentwicklung in Deutschland wird dieses Vorgehen demonstriert und für eine Analyse der Deflationsrisiken eingesetzt. Es zeigt sich, dass die Prognoseintervalle für die Preisentwicklung sehr rasch mit der Länge des Prognosezeitraums zunehmen und nach wenigen Quartalen bereits die Nulllinie umschliessen. Ebenfalls modellbasiert bei der Analyse von Fehlerquellen und Unsicherheiten gesamtwirtschaftlicher BIP- und Inflationsprognosen geht Heilemann (2004) vor. Sehr knapp zusammengefasst kommt Heilemann (2004) zu dem Ergebnis, dass im Prinzip die gegenwärtigen Fehler beobachteter Konjunkturprognosen die Grenze der Genauigkeit von Konjunkturprognosen markieren, wenn die Simulationsrechnungen des RWI-Konjunkturmodells als Vergleichsgrundlage herangezogen werden. Deutliche Erhöhungen der Treffsicherheit dürften demnach aus heutiger Sicht schwer erreichbar sein. Heilemann (2004) spricht sich dafür aus, die Grenzen der Genauigkeit von Konjunkturprognosen besser zur Kenntnis zu nehmen, da für konjunkturpolitisches Handeln oder Unterlassen die Risikobehaftung von Prognosen von wesentlichem Interesse ist. Der vorliegende Beitrag setzt in gewisser Weise an dieser Empfehlung an. Nicht die Verbesserung von Konjunkturprognosen steht im Zentrum der Betrachtung, sondern die angemessene Erfassung und Darstellung der Unsicherheit von Konjunkturprognosen für Deutschland.

Fritsche (2005) berechnet Prognosegütemaße für BIP- und Inflationsprognosen für Deutschland von mehreren nationalen Instituten und supranationalen Institutionen und führt Tests auf Unverzerrtheit und Rationalität durch. Schwerpunkt der Betrachtung bildet in dieser Studie die Auswertung der einzelnen Institutsprognosen. Das von Fritsche (2005) ermittelte empirische Prognoseintervall zum 95%-Signifikanzniveau der Median-

prognose zur BIP-Entwicklung ist weit und beträgt $\pm 2\%$, das Intervall zum 66%-Niveau immer noch $\pm 1\%$. Auch diese Untersuchung zeigt, dass gesamtwirtschaftliche Prognosen mit einiger Unschärfe verbunden sind.

Diese kurze Einführung in die Literatur stellt einen sehr beschränkten Ausschnitt aus diesem sehr umfangreichen Forschungsgebiet dar und soll lediglich dazu dienen, die unmittelbaren Ansatzpunkte der eigenen Arbeit aufzuzeigen.¹

Welche Argumente für und gegen eine Angabe von Prognoseintervallen sprechen, wie diese näherungsweise berechnet werden können und wie solche Ergebnisse für die Beurteilung von laufenden Prognosen und Prognoserevisionen verwendet werden können, wird im Folgenden näher betrachtet. Vor allem die Nutzung von empirischen Prognoseintervallen für die Einschätzung von Prognoserevisionen wurde in der Literatur bisher noch nicht erwogen.

2 Unterschiedliche Methoden der Konjunkturprognose

Den Prognostikern stehen zur Erfüllung ihrer Aufgabe eine Vielzahl unterschiedlicher Verfahren zur Verfügung, die sich generell in drei Kategorien einordnen lassen: *Indikatoren-gestützte Ansätze*, *makroökonomische Strukturmodelle* und *iterativ-analytische Verfahren*.

Beim Indikatorenansatz werden Vorlaufeigenschaften von Konjunkturindikatoren (z.B. Konjunkturumfragen) zur Prognoseerstellung genutzt. Das monatliche DIW-Konjunkturbarometer beispielsweise schätzt die Veränderungsrate des realen Bruttoinlandsprodukts (BIP) auf Sicht der jeweils folgenden zwei Quartale auf Grundlage von Indikatoren mit Vorlaufeigenschaften. Der monatlich von der EU-Kommission veröffentlichte BIP-Indikator nutzt ebenfalls vorlaufende Informationen einer Reihe von Zeitreihen für eine Abschätzung der Quartalsveränderungsrate des Eurozonen-BIP. In dieses Prognosemodell gehen Frühindikatoren wie die Neuzulassung von PKW, das Verbrauchervertrauen der privaten Haushalte, das Geschäftsklima im Einzelhandel, Stimmungsindikatoren zum Baugewerbe, die Zinsstrukturkurve und einige mehr ein.²

Den makroökonomischen Strukturmodellen liegt eine andere Herangehensweise zugrunde. Hier werden wechselseitige Abhängigkeiten zwischen den Wirtschaftssektoren über umfangreiche Gleichungssysteme abgebildet und für Prognosen herangezogen. Solche Mo-

¹Vgl. u.a. Hinze (2005), Nierhaus (2001), Nierhaus (2005), Döpke und Fritsche (2004), Heilemann und Stekler (2003) und Grömling (2005) für weitere Arbeiten, die sich mit der Güte makroökonomischer Prognosen für Deutschland befassen.

²Die Konstruktion des Prognosemodells sowie eine vollständige Beschreibung der verwendeten Frühindikatoren ist in Grasmann und Keeremann (2001) zu finden.

delle für Deutschland werden etwa von einigen Wirtschaftsforschungsinstituten und der Deutschen Bundesbank eingesetzt, auf supranationaler Ebene werden länderübergreifende Makromodelle beispielsweise von der Europäischen Zentralbank und der OECD verwendet.³

Das iterativ-analytische Verfahren ist für Prognosezeiträume von ein bis zwei Jahren besonders geeignet und wird in Deutschland unter anderem vom Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung und den Wirtschaftsforschungsinstituten eingesetzt. Dieses Verfahren basiert auf dem System der volkswirtschaftlichen Gesamtrechnung. In einem mehrstufigen Prozess werden dabei Prognosen, die über verschiedene Methoden und Herangehensweisen ermittelt werden, solange abgeglichen, bis eine in sich konsistente Vorhersage vorliegt. Vorteil dieses Verfahrens ist seine Flexibilität sowie die Möglichkeit, neben quantitativen auch qualitative Annahmen zu berücksichtigen. Die Entstehung einer Prognose auf Basis des iterativ-analytischen Verfahrens ist für Außenstehende aufgrund der Komplexität und der Vielzahl an eingesetzten Methoden jedoch nicht mehr nachvollziehbar.

3 Die Zukunft ist unsicher

Allen Verfahren ist jedoch gemein, dass sie nur zuverlässige Prognose liefern können, solange die unterstellten ökonomischen Zusammenhänge stabil sind und sich die getroffenen Annahmen als richtig erweisen. Da Prognosen gewöhnlich unter bestimmten Annahmen entstehen (z.B. Annahmen über den Ölpreis, über die Geld- und Fiskalpolitik oder Wechselkurse) spricht man auch von bedingten Prognosen. Wirtschaftliche Rahmenbedingungen ändern sich jedoch regelmäßig und sind nur bedingt vorhersehbar. Bestimmte Schocks, wie beispielsweise Naturkatastrophen, sind überhaupt nicht vorhersehbar. Zudem sorgen regelmäßige Revisionen der amtlichen Statistik für rückwirkende Änderungen der Historie von volkswirtschaftlichen Daten, so dass die Ausgangslage der Prognose nicht mehr gültig ist. Prognosen unterliegen also Fehlerrisiken, die auch mit zunehmender Verbesserung der Verfahren nicht auszuschließen sind und gegenüber denen der Prognostiker machtlos ist. Eine Angabe von Schwankungsbereichen, in den z.B. eine prognostizierte BIP-Entwicklung anzusiedeln ist, würde der Unsicherheit der Prognose deswegen angemessen Rechnung tragen.

Die Unsicherheit über eine Prognose wird üblicherweise über den erwarteten mittleren quadratischen Prognosefehler (Mean Squared Error, MSE) ausgedrückt:

$$MSE(Y_{t+h|t}^P) = E(Y_{t+h} - Y_{t+h|t}^P)^2 \quad (1)$$

³Ausführlich dokumentiert wird das makroökonomische Modell der Bundesbank in Deutsche Bundesbank (2000). Richardson (1988) beschreibt das INTERLINK-Modell der OECD.

$Y_{t+h|t}^P$ ist die zum Zeitpunkt t erstellte Prognose für die Variable Y_{t+h} und $E(\dots)$ bezeichnet den Erwartungswert. Mit h wird der Prognosehorizont angegeben. Bei erwartungstreuen Prognosen ist der MSE gleich der Prognosefehlervarianz.⁴ Der MSE ist die zentrale Größe für die Berechnung von Prognoseintervallen.

In der Praxis ist es allerdings meist schwierig, eine exakte Quantifizierung für (1) zu finden. Theoretische Ausdrücke für den MSE existieren zwar für eine Reihe von Prognosemodellen (z.B. AR-Modelle und ARIMA-Modelle). Für komplizierte Mehrgleichungssysteme, die darüber hinaus nicht-lineare Zusammenhänge modellieren oder Modelle, in denen Entwicklungen über exogene Variablen auf Basis von Erfahrungswerten (z.B. konkrete Annahmen über den Verlauf des Ölpreises) vorgegeben werden, können in der Regel keine theoretischen Lösungen für (1) gefunden werden (vgl. Chatfield, 1993). Das in der Praxis häufig eingesetzte iterativ-analytische Verfahren schließt im Prinzip die Herleitung eines analytischen MSE aus.

Wenn analytische Formeln für (1) fehlen, können jedoch empirische MSE berechnet werden, die auf der Beobachtung der Prognosefehler basieren. Für eine unverzerrte Prognose gilt, dass $E(Y_{t+h} - Y_{t+h|t}^P) = 0$ und $E(Y_{t+h} - Y_{t+h|t}^P)^2 = Var(Y_{t+h} - Y_{t+h|t}^P)$ ist, wobei $Var(\dots)$ die Varianz bezeichnet. Wenn die Prognosefehler normalverteilt sind, kann ein Prognoseintervall für eine Vertrauenswahrscheinlichkeit von $100 * (1 - \alpha)\%$ wie folgt berechnet werden:

$$Y_{t+h|t}^P \pm z_{\alpha/2} \sqrt{Var(Y_{t+h} - Y_{t+h|t}^P)} \quad (2)$$

α ist die vorgegebene Irrtumswahrscheinlichkeit und $z_{\alpha/2}$ das entsprechende Quantil der Standardnormalverteilung, welches mit der Wurzel des durchschnittlichen quadratischen Prognosefehlers (RMSE, Root Mean Squared Error) multipliziert wird. Man beachte die Ähnlichkeit von (2) zu der Berechnung von Konfidenzintervallen bei der Schätzung unbekannter, aber fester Parameter, beispielsweise mit der Methode der Kleinsten Quadrate. Die klassische Interpretation eines solchen Konfidenzintervalls ist wie folgt (vgl. Hogg und Craig, 1995, S. 269): Vor der wiederholten Durchführung des unabhängigen Zufallsexperiments beträgt die Wahrscheinlichkeit, dass das Konfidenzintervall den unbekanntem, aber festen Parameter enthält, $100 * (1 - \alpha)\%$. Im Gegensatz dazu ist ein Prognoseintervall eine Schätzung der Bandbreite eines unbekanntem zukünftigen Wertes, der als Zufallsvariable zum Zeitpunkt der Prognoseerstellung betrachtet werden kann. Durch (2) werden die unteren und oberen Werte dieses Intervalls vorgegeben.

4 Empirische Prognoseintervalle

Die Berechnung von Prognoseintervallen wird im Folgenden für Prognosen des DIW, des ifo, des IfW und des RWI für eine Auswahl an volkswirtschaftlichen Kennziffern durchgeführt. Betrachtet werden die Mediane der Prognosen dieser Institute für die Jahre 1980

⁴Vgl. z.B. Schröder (2002), S. 449

bis 2004, da die Prognosen der Institute nicht einzeln ausgewertet werden sollen. Die Prognosen weisen Horizonte von 6, 12 und 18 Monaten auf. Diese Horizonte entsprechen ungefähr der Veröffentlichungsstruktur der Institutsprognosen, wenn die halbjährlichen Prognosepublikationen im Rahmen der Gemeinschaftsdiagnose nicht berücksichtigt werden.⁵ Eine Prognose mit einem Horizont von 12 Jahren wird im Dezember des vorangehenden Jahres oder im Januar des Prognosejahres veröffentlicht. Die Prognosen mit Horizonten von 6 und 18 Monaten werden üblicherweise im Juni oder Juli veröffentlicht und prognostizieren die Entwicklung des laufenden und des nächsten Jahres. Da die Institutsprognosen nicht alle zum exakt gleichen Zeitpunkt veröffentlicht werden, sondern gelegentlich auch im Zeitverlauf variieren, ist der Informationsstand nicht identisch. Aus diesem Grunde ist eine strenge Gegenüberstellung und Gütebewertung der Prognosen nicht sachgerecht. Auch dies ist ein Grund für die folgende Einschränkung der Analyse auf die Medianprognose, da dadurch das Problem unterschiedlicher Informationsumgangen wird.

Amtliche Zahlen zur Volkswirtschaftlichen Gesamtrechnung werden üblicherweise mehrmals revidiert, so daß zur Überprüfung der Prognosegenauigkeit mehrere Werte zur Verfügung stehen. Aufgrund der zeitlichen Nähe zum Prognosezeitraum und weil die Auswirkungen statistischer Revisionen von den Prognostikern kaum präzise abzuschätzen sind, werden in Übereinstimmung mit der Literatur den Prognosen die ersten Veröffentlichungen des Statischen Bundesamtes als realisierte Werte gegenübergestellt. Da bis etwa 1991 das Augenmerk auf dem Bruttosozialprodukt (BSP) lag, wird bis einschließlich 1991 die Veränderungsrate des BSP und ab 1992 die Veränderungsrate des BIP verwendet. Von 1980 bis 1994 beziehen sich die Daten auf das frühere Bundesgebiet, ab 1995 wird Deutschland betrachtet. Tabelle 1 gibt eine Übersicht über die analysierten Daten.

Vor der Berechnung der Prognoseintervalle werden die Prognosefehler der Mediane der Institutsprognosen auf Erwartungstreue und Normalverteilung getestet. Für den Test auf Erwartungstreue werden Holden und Peel (1990) folgend die Prognosefehler auf eine Konstante regressiert. Diese Konstante gibt die durchschnittliche Abweichung der Prognosen von der Realisation an. Nur wenn die Konstante signifikant von null verschieden ist, wird die Hypothese der Erwartungstreue der Prognose verworfen. Für den Test auf die Normalverteilung der Prognosefehler wurde die Test-Statistik von Jarque-Bera berechnet. Unter der Nullhypothese dieses Tests sind die Prognosefehler normalverteilt.

Die in Tabelle 1 ausgewiesene Konstante der Regression entspricht dem mittleren Prognosefehler (ME, Mean Error). Die Prognosen für den privaten Konsum sind ab einem Prognosehorizont von einem Jahr als nicht erwartungstreu (verzerrt) zu betrachten. Sowohl die Prognosen mit einem Horizont von 12 Monaten als auch die Prognosen mit einem Horizont von 18 Monaten waren im Durchschnitt um 0,39 beziehungsweise um 0,64 Prozentpunkte über den tatsächlichen Jahresveränderungsraten. Die Medianprognosen der

⁵In Fritsche (2005), S. 363, findet sich eine Abbildung, die die typischen Veröffentlichungszeitpunkte der Prognosen der nationalen und internationalen Institutionen auf einem Zeitstrahl abbildet.

Tabelle 1: Test auf Erwartungstreue der Medianprognosen und Normalverteilung der Prognosefehler

Variable	Horizont	RMSE ¹	ME (Mean Error) ²	Jarque-Bera ³
Bruttoinlandsprodukt (real)	6 Monate	0,48	0,07	1,00
	12 Monate	1,06	-0,05	0,33
	18 Monate	1,51	-0,32	1,18
Private Konsumausgaben (real)	6 Monate	0,53	-0,16	0,25
	12 Monate	0,99	-0,39***	1,51
	18 Monate	1,32	-0,64***	0,49
Konsumausgaben des Staates (real)	6 Monate	0,94	0,13	3,07
	12 Monate	0,99	0,08	1,10
	18 Monate	1,05	0,04	1,33
Ausrüstungsinvestitionen (real)	6 Monate	3,31	-1,68**	1,10
	12 Monate	4,88	-1,48	1,45
	18 Monate	6,51	-2,44**	3,80
Ausfuhr (real)	6 Monate	1,74	-0,05	0,04
	12 Monate	4,07	0,78	0,04
	18 Monate	5,03	0,39	0,75
Einfuhr (real)	6 Monate	2,19	-0,47	5,14*
	12 Monate	3,53	-0,12	17,79***
	18 Monate	4,62	-0,53	15,20***
Erwerbstätige	6 Monate	0,58	-0,22	84,85***
	12 Monate	0,68	-0,35	0,88
	18 Monate	0,84	-0,57	0,89
Verbraucherpreise	6 Monate	0,30	-0,01	14,36***
	12 Monate	0,74	-0,08	5,75*
	18 Monate	0,90	-0,09	0,99

Anmerkungen: ¹Wurzel des durchschnittlichen quadratischen Prognosefehlers (RMSE, Root Mean Squared Error), ²Angaben in Jahresveränderungsraten und Prozent, ³Jarque-Bera-Statistik ist Chi-Quadratverteilt mit 2 Freiheitsgraden, Stützbereich 1980 bis 2004, ***/**/* Signifikanz zum 1%/5%/10%-Niveau

Institute haben also die Entwicklung zum privaten Konsum über die nächsten ein bis eineinhalb Jahre im Durchschnitt zu optimistisch eingeschätzt. Eine ähnliche Verzerrung ist für die Prognosen zur Entwicklung der Ausrüstungsinvestitionen festzustellen. Hier sind die durchschnittlichen Prognosefehler für die Horizonte 6 Monate und 18 Monate signifikant von Null verschieden, der durchschnittliche Prognosefehler auf Sicht von 12 Monaten ist insignifikant zum 10%-Niveau, allerdings ebenfalls positiv. Für die restlichen Größen können die Prognosen als erwartungstreu erachtet werden. Sie erfüllen somit eine wichtige Eigenschaft optimaler Prognosen.

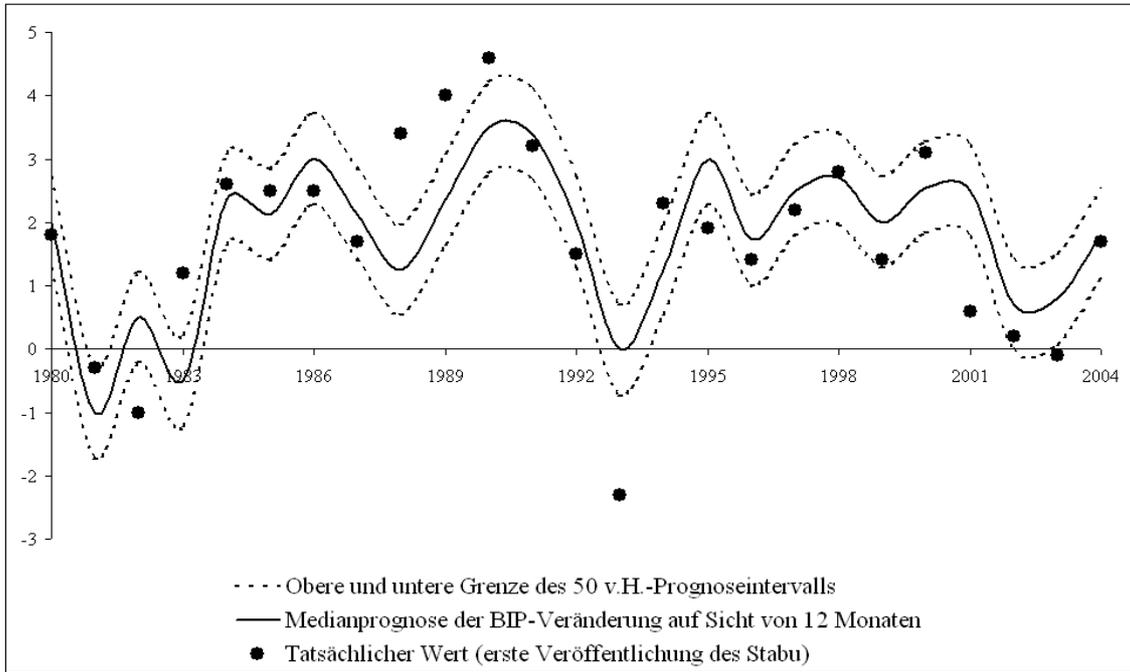
Die letzte Spalte der Tabelle 1 zeigt die Ergebnisse der Tests auf Normalverteilung der Prognosefehler. Hier wird die Jarque-Bera-Teststatistik der Prognosefehler für die unterschiedlichen Prognosehorizonte ausgewiesen. Eine Normalverteilung der Prognosefehler kann bis auf die Einfuhrentwicklung, die Erwerbstätigenentwicklung sowie die Verbraucherpreisentwicklung für alle getesteten VGR-Prognosen und für alle drei Prognosehorizonte festgestellt werden. Die Ablehnung der Normalverteilung für die Prognosefehler der Einfuhren stehen in Zusammenhang mit statistischen Ausreißern im Jahr 1993, die auf eine Änderung in der Erfassung des Außenhandels zurückgehen. Bei einer Auswertung ohne Berücksichtigung der Prognosefehler im Jahr 1993 kann auch für diese Größe die Normalverteilungshypothese nicht abgelehnt werden.

Ebenfalls ausgewiesen in Tabelle (1) sind die durchschnittlichen quadratischen Prognosefehler (RMSE), welche erwartungsgemäß für alle Prognosen mit zunehmendem Horizont größer werden. Der RMSE ist zentraler Ausgangspunkt zur Erstellung empirischer Prognoseintervalle. Gemäß Formel (2) wird der RMSE für den jeweiligen Prognosehorizont mit den entsprechenden Quantilen der Normalverteilung multipliziert, um einen oberen und unteren Wert für die Intervalle zu erhalten.

Abbildung 1 demonstriert diese Berechnung für die Medianprognose der Institute zur realen BIP-Entwicklung auf Sicht von 12 Monaten. In der Grafik wurde ein Prognoseintervall für eine Vertrauenswahrscheinlichkeit von 50% abgetragen.

Die Angabe von 50%-Intervallen ist aus praktischer Sicht gegenüber den üblichen Intervallen von 95% zweckmäßiger (vgl. Granger, 1996): Prognoseintervalle auf dem 95%-Niveau haben eine relativ große Spannweite und sind deshalb kaum noch zu vermitteln. Das 95%-Intervall für die BIP-Prognose im Jahr 2004 würde beispielsweise eine Spanne -0,4 v.H. bis 3,8 v.H. aufweisen. Im Gegensatz dazu erstreckt sich das 50%-Prognoseintervall für das Jahr 2004 über eine Spanne von 1,1% bis 2,5%. Darüber hinaus sind 50%-Intervalle robuster gegenüber der Verteilungsannahme von Prognosefehlern und Ausreißern. Die Plausibilität des Intervalls kann zudem leicht anhand der historischen Fehler beurteilt werden. Wie aus Abbildung 1 hervorgeht, wird ungefähr die Hälfte der tatsächlichen Werte von dem Prognoseintervall abgedeckt. Je weiter die Prognosen in die Zukunft reichen, desto größer ist die Prognoseunsicherheit, wie an den steigenden Werten für die RMSE

Abbildung 1: BIP-Prognosen und Prognoseintervalle



in Tabelle 1 ersichtlich ist. Deshalb werden auch die Prognoseintervalle mit wachsendem Prognosehorizont zunehmend breiter.

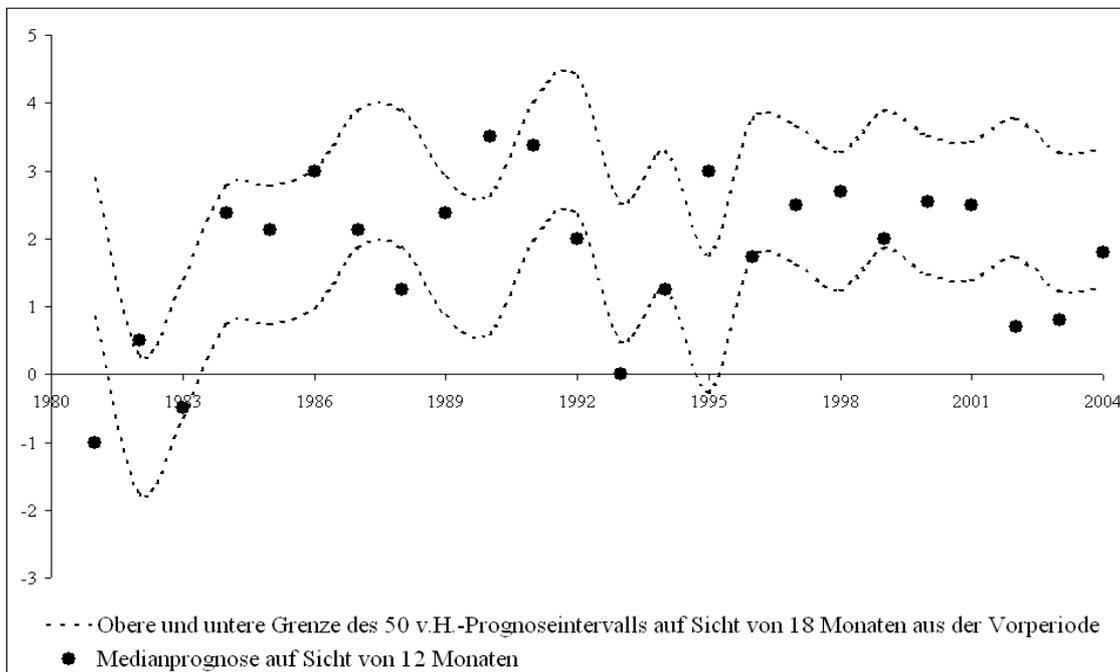
5 Signifikanz von Prognoserevisionen

Volkswirtschaftliche Prognosen werden regelmäßig revidiert, sobald neue Informationen über die aktuelle Konjunktursituation und den weiteren Ausblick vorliegen. Neubewertungen sind bei geändertem Informationsstand geradezu notwendig um die wichtige Eigenschaft der Informationseffizienz von Prognosen aufrecht zu erhalten. Revisionen ändern die entsprechende Punktprognose, aber auch das daraus abgeleitete Prognoseintervall.

Ob solche Revisionen im Bereich der üblichen Prognoseunsicherheit anzusiedeln sind, kann anhand einer Gegenüberstellung des neuen Prognosewertes und des Prognoseintervalls vor Revision beurteilt werden. Formal läßt sich dies wie folgt ausdrücken: Eine nach r Perioden revidierte Prognose $Y_{t+h-r|t+r}$ für einen Horizont von $h - r$ ist mit einer Vertrauenswahrscheinlichkeit von $100 * (1 - \alpha)\%$ nicht signifikant, falls diese Revision von dem Prognoseintervall der ursprünglichen Prognose $\left(y_{t+h|t} - z_{\alpha/2} \sqrt{Var(Y_{t+h} - Y_{1+h|t}^P)}, y_{t+h|t} + z_{\alpha/2} \sqrt{Var(Y_{t+h} - Y_{1+h|t}^P)}\right)$ umschlossen wird.

Abbildung 2 verdeutlicht dies, indem neben dem 50 v.H.-Prognoseintervall der BIP-Veränderung auf Sicht von 18 Monaten aus der Vorperiode die neue Prognose auf Sicht von 12 Monaten abgetragen wird. Letztere entspricht der Prognose, die revidiert wurde.

Abbildung 2: Signifikanz von Prognoserevisionen



Aus der Abbildung 2 geht hervor, dass ca. 45% der revidierten Prognosen insignifikant zum 50%-Niveau waren. Ein Signifikanztest zu den üblichen 90%- oder 95%-Niveaus mit entsprechend weiteren Prognoseintervallen würde die Mehrzahl der Prognoserevisionen als statistisch insignifikant erscheinen lassen. Eine systematische Übersicht über den Anteil signifikanter Prognoserevisionen über den Beobachtungszeitraum von 1980 bis 2004 gibt Tabelle 2. Ausgangspunkt für die Beurteilung einer Prognoserevision für das laufende Jahr ist eine Prognose, die zur Mitte des Jahres erstellt wurde und die Veränderungsrate der entsprechenden Variablen für das laufende Jahr prognostiziert. Diese Prognose revidiert die Prognose, die zu Beginn des Jahres abgegeben wurde und ebenfalls auf die Veränderungsrate des laufenden Jahres abzielte. Liegt die Prognose aus der Jahresmitte im Prognoseintervall, welches sich durch die Prognose zu Jahresbeginn ergeben hat, ist die Revision als insignifikant anzusehen. Entsprechend werden die Anteile der insignifikanten Revisionen für das folgende Jahr ermittelt: Hier wird die Prognose zu Beginn des Jahres dem Prognoseintervall der Prognose aus der Mitte des Vorjahres gegenübergestellt. In der Tabelle sind die entsprechenden Ergebnisse für Vertrauenswahrscheinlichkeiten von 50% und 90% dargestellt.

Legt man eine Vertrauenswahrscheinlichkeit von 90% zugrunde, wie es für statistische Hypothesentests üblich ist, kann nur ein verschwindend geringer Anteil von Revisionen der Medianprognosen als signifikant bezeichnet werden. Selbst bei einer Vertrauenswahrscheinlichkeit von 50% war die Mehrzahl der Prognoserevisionen für die hier betrachteten volkswirtschaftlichen Kennziffern insignifikant. Insbesondere geringfügige Revisionen von

Tabelle 2: Anteile signifikanter Prognoserevisionen an den gesamten Revisionen über den Zeitraum 1980 bis 2004

	für das laufende Jahr		für das folgende Jahr	
	50%-Niveau	90%-Niveau	50%-Niveau	90%-Niveau
Bruttoinlandsprodukt (real)	0,36	0,04	0,42	0,04
Private Konsumausgaben (real)	0,40	0,04	0,21	0,00
Konsumausgaben des Staates (real)	0,24	0,00	0,46	0,00
Ausrüstungsinvestitionen (real)	0,20	0,00	0,21	0,00
Ausfuhr (real)	0,40	0,00	0,21	0,00
Einfuhr (real)	0,24	0,00	0,08	0,00
Erwerbstätige	0,44	0,04	0,21	0,04
Verbraucherpreise	0,56	0,08	0,08	0,00

Prognosen signalisieren also meistens keine wesentlich neuen Einschätzungen hinsichtlich der zukünftigen konjunkturellen Entwicklung.

6 Grundsätzliche Prognoseschwierigkeiten

Die generellen Unsicherheiten von Konjunkturprognosen können durch den RMSE und die daraus abgeleiteten Prognoseintervalle nur unzureichend ausgedrückt werden, da der RMSE eines Prognosemodells für sich allein schwierig zu interpretieren ist. Eine Normalisierung ist deshalb nötig, um die Größenordnung der Variablen zu berücksichtigen. Dafür wird üblicherweise der RMSE des zu überprüfenden Prognosemodells dem RMSE eines Referenzmodells gegenübergestellt. Der Quotient aus beiden RMSE ergibt den Theil'schen Ungleichheitskoeffizienten (Theil's U). Wenn das Prognosemodell besser als das Referenzmodell im Sinne eines geringeren durchschnittlichen quadratischen Prognosefehlers abschneidet, sollte Theil's U kleiner als eins sein.⁶

Im Folgenden wird Theil's U herangezogen, um zu beurteilen, welche volkswirtschaftlichen Kennziffern grundsätzlich besser zu prognostizieren sind als andere. Die isolierte Betrachtung der Prognosefehler hilft hier nicht weiter, da nicht eindeutig ist, welche generelle Streuung mit der Zielgröße einer bestimmten Prognose verbunden ist. Die Kenntnis dieser grundsätzlichen Unsicherheiten kann für die Empfänger von Prognosen durchaus hilfreich sein. Beispielsweise ist es für die Steuerschätzung interessant zu wissen, ob die Bemessungsgrundlage für Konsumsteuern und Einkommensteuern mit unterschiedlicher Genauigkeit prognostiziert werden. Für die Arbeitsmarktpolitik oder für die Abschätzung

⁶Formal ist Theil's U definiert als $\frac{RMSE_{Prognosemodell}}{RMSE_{Referenzmodell}}$, vgl. z.B. Schröder (2002), S. 449

der Einnahmen der Sozialversicherungen ist eine Kenntnis über die Prognosequalität der Erwerbstätigenentwicklung und der Arbeitslosigkeit von entscheidender Bedeutung.

In der Literatur wird für die Berechnung des Theil's U häufig ein Random Walk als Referenzmodell verwendet oder direkt die Standardabweichung der Zielgröße ins Verhältnis zum RMSE des Prognosemodells gesetzt⁷. Im letzten Fall werden die Prognosen also mit einer Methode verglichen, die lediglich auf Durchschnittswerten der realisierten Zielgröße aufbaut. Bei einem Random Walk als Prognosemodell ist die optimale Prognose gleich dem Wert der letzten Beobachtung. Daher wird die Random Walk-Prognose häufig auch als naive Prognose bezeichnet, da keine weitere Information außer der letzten Realisation der Zielgröße Eingang in die Vorhersage findet.

Für die Berechnung der Theil's U-Werte in Tabelle 3 wurden als alternative Referenz zu den Median-Prognosen der Wirtschaftsforschungsinstitute hingegen ARIMA-Modelle eingesetzt, da hier ein ambitionierterer Maßstab als Referenz herangezogen wird, um zu aussagefähigeren Gütemaße für die Institutsprognosen zu kommen. Diese Zeitreihenmodelle wurden mit Quartalsdaten des Statistischen Bundesamts geschätzt und auf Basis des modifizierten Schwarz-Informationskriteriums spezifiziert. Entsprechend der Abgrenzung der Institutsprognosen wurden bis 1994 Daten für das frühere Bundesgebiet und danach für Deutschland verwendet. Der statistische Stützbereich für die ARIMA-Schätzungen beginnt im 1. Quartal 1968 und endet in dem jeweiligen Quartal, in dem die Prognose erstellt wird. Theil's U-Werte wurden für Prognosehorizonte von 6, 12 und 18 Monaten berechnet. In der Tabelle 3 sind Durchschnitte der Theil's U-Werte über diese drei Horizonte angegeben, um die Präzision der Prognosen insgesamt zu erfassen.

Tabelle 3: Güte der Institutsprognosen

	Theil's U ¹	Standardfehler
Verbraucherpreise	0,45	1,55
Konsumausgaben des Staates (real)	0,48	1,23
Bruttoinlandsprodukt (real)	0,56	1,57
Private Konsumausgaben (real)	0,57	1,51
Ausfuhr (real)	0,58	4,61
Ausrüstungsinvestitionen (real)	0,74	6,47
Einfuhr (real)	0,75	4,45
Erwerbstätige	0,82	2,58

Anmerkungen: ¹ARIMA-Modelle als Benchmark, arithmetisches Mittel der 6-, 12- und 18-Monatsprognosen

⁷vgl. u.a. Heilemann und Stekler (2003), Ruoss und Savioz (2002), Pons (2000) oder Ash, Smyth und Heravi (1997) für Untersuchungen in dieser Richtung. Carnot, Koen und Tissot (2005), S. 238f beschreiben weitere Varianten und Anwendungsmöglichkeiten des Theil'schen Ungleichheitskoeffizienten.

Kein ARIMA-Modell kann exaktere Vorhersagen treffen als die Medianprognosen der Institute, wenn die Wurzel des durchschnittlichen quadratischen Prognosefehlers als Maßstab herangezogen wird, da der Theil'sche Ungleichheitskoeffizient keiner Variablen größer als eins ist. Dieses Ergebnis spricht für die Güte der Medianprognose der Institute. Eindeutig zu erkennen ist auch die unterschiedliche Genauigkeit, mit der die verschiedenen VGR-Kennziffern prognostiziert werden. Am besten gelingt die Prognose der Verbraucherpreise, während die Entwicklung der Erwerbstätigen am ungünstigsten im Vergleich zum Referenzmodell ausfällt. Die Unterschiede in der Prognosegüte des Bruttoinlandsprodukts, des privaten Verbrauchs und der Ausfuhren sind gering. Mehr Schwierigkeiten bereiten die Prognosen zur Entwicklung der Ausrüstungsinvestitionen sowie zur Einfuhrentwicklung. Generell ist festzustellen, dass die Prognoseschwierigkeit mit zunehmender Streuung der Zielgröße steigt. Die Verwendung eines relativ einfachen Zeitreihenmodells wird folglich mit steigender Varianz der zu prognostizierenden Variablen zunehmend zweckmäßiger. Alternative Methoden, die bedeutend mehr Aufwand für die Prognoseerstellung erfordern, sind für solche un stetigen Zielgrößen kaum mehr in der Lage, wesentlich bessere Prognosen zu liefern.

7 Schlussfolgerungen

Die Berechnung von empirische Prognoseintervallen sowie eine Analyse der Signifikanz von Prognoserevisionen auf Grundlage dieser Intervalle wurde für den Median der Prognosen vier deutscher Wirtschaftsforschungsinstitute für eine Reihe volkswirtschaftlicher Kennziffern demonstriert. Es zeigt sich, dass vor allem geringfügige Prognoserevisionen meistens keine wesentlich neuen Einschätzungen hinsichtlich der zukünftigen konjunkturellen Entwicklung signalisieren. Für eine Beurteilung der grundsätzlichen Prognoseschwierigkeit makroökonomischer Größen wurden die Medianprognosen der Institute in einem weiteren Untersuchungsschritt alternativen Zeitreihenmodellen gegenübergestellt. Die Institutsprognosen weisen gegenüber den Zeitreihenmodellen durchweg eine höhere Güte auf, wenn die Wurzel des mittleren quadratischen Prognosefehlers als Maßstab herangezogen wird. Allerdings sind deutliche Unterschiede in der Prognosepräzision der betrachteten volkswirtschaftlichen Kennziffern feststellbar.

Die Angabe von expliziten Intervallen für Konjunkturprognosen ist nicht üblich. Unsicherheit wird zumindest in den ausführlicheren Gutachten über qualitative Informationen zu den Prognosenannahmen zum Ausdruck gebracht. Das weitverbreitete iterativ-analytische Prognoseverfahren erschwert zudem die Kalkulation exakter Prognoseintervalle. Einen Ausweg könnten hier empirische Prognoseintervalle liefern. Allerdings kann die Streuung historischer Prognosefehler auch zu starken Fehleinschätzungen über die Prognoseunsicherheit führen, wenn sich die bedingte Varianz der Zielgröße im Zeitverlauf ändert, wie es beispielsweise bei Finanzmarktdaten regelmäßig der Fall ist. Elegantere Alternativen

der Kalkulation von Prognoseintervallen existieren, wenn das wahre Modell der zu prognostizierenden Variablen bekannt ist. Dann können unmittelbar theoretische Prognoseintervalle abgeleitet werden. Üblicherweise ist das wahre Modell des datenerzeugenden Prozesses allerdings nicht bekannt, so dass ein Modell empirisch angepasst werden muss. Da dabei die funktionale Form sowie die Parameter des Modells geschätzt werden müssen, ergibt sich eine zusätzliche Quelle der Prognoseunsicherheit, die die resultierenden Prognoseintervalle dementsprechend verbreitern.⁸ Für die Evaluation externer Prognosen, wie beispielsweise der Prognosen von Wirtschaftsforschungsinstituten oder Banken, scheiden diese Möglichkeiten in der Regel aus, da die eingesetzten Verfahren eine Reproduktion der Prognosen durch Außenstehende nicht ermöglichen. Die Qualität dieser Prognosemethoden kann deshalb nur anhand der veröffentlichten Prognosen und der Prognosefehler beurteilt werden.

Ein weiteres Problem von Prognoseintervallen stellt die Wahl der Vertrauenswahrscheinlichkeit dar, da diese maßgeblich für die Spannweite der Intervalle ist. Ein objektives Kriterium für die Bestimmung dieser Vertrauenswahrscheinlichkeit existiert nicht und Spielräume bei der Berechnung von Prognoseintervallen sind unvermeidlich. Letztendlich muss der Empfänger von Prognosen, für den diese bei seinen Entscheidungen wichtig sind, beurteilen, welches Vertrauen er Konjunkturprognosen zu Grunde legt und mit welchen möglichen Ausprägungen er zu rechnen hat. Empirische Prognoseintervalle, die zugegebenermaßen nur eine grobe Vorgehensweise darstellen, können dafür eine Hilfestellung bieten. Beachtet man alle Einschränkungen im Zusammenhang mit der Berechnung von Prognoseintervallen, könnte eine Konjunkturprognose unter Angabe von Schwankungsbereichen für die Prognostiker eventuell noch schwieriger zu kommunizieren sein.

⁸Chatfield (1993), S. 124 demonstriert die Korrektur für den MSE des zugrundeliegenden Prognoseintervalls, wenn die Parameter des Prognosemodells geschätzt werden. Demnach hängt die Korrektur positiv von der Anzahl der zu schätzenden Parameter ab. Eine geringe Anzahl an Beobachtungen trägt ebenfalls zu einer Zunahme des MSE bei. Die Methode von Borbély und Meier (2005) erlaubt ebenfalls eine Berücksichtigung von Modellunsicherheiten bei der Kalkulation modellbasierter Prognoseintervalle.

Literatur

- ASH, J.C.K, D.J. SMYTH UND S.M. HERAVI (1997): "The accuracy of OECD forecasts of the international economy: balance of payments," *Journal of International Money and Finance*, 16(6), 969–987.
- BORBÉY, D. UND C.-P. MEIER (2005): "Assesing Macroeconomic Forecast Uncertainty: An Application to the Risk of Deflation in Germany," *Kredit und Kapital*, 38(3), 1–23.
- CARNOT, N., V. KOEN UND B. TISSOT (2005): *Economic Forecasting*. Palgrave MacMillan.
- CHATFIELD, C. (1993): "Calculating Interval Forecasts," *Journal of Business Economic Statistics*, 11(2), 121–135.
- DEUTSCHE BUNDESBANK (2000): "Macro-Econometric Multi-Country Model: MEM-MOD," *Frankfurt am Main*.
- DÖPKE, J. UND U. FRITSCH (2004): "Growth and Inflation Forecasts in Germany - An Assessment of Accuracy and Dispersion," DIW Berlin, unveröffentlicht.
- FRITSCH, U. (2005): "Warum Konjunkturprognosen?," *Wochenbericht des DIW*, 22, 361–369.
- GRANGER, C. (1996): "Can We Improve the Perceived Quality of Economic Forecasts?," *Journal of Applied Econometrics*, 11, 455–473.
- GRASMANN, P. UND F. KEEREMAN (2001): "An indicator-based short-term forecast for quarterly GDP in the euro area," *European Commission Economic Papers No. 154*, Brussels.
- GRÖMLING, M. (2005): "Konjunkturprognosen - Verfahren, Erfolgskontrolle und Prognosefehler," unveröffentlicht.
- HEILEMANN, U. (2004): "As Good as it Gets - Limits of Accuracy of Macroeconomic Short Term Forecasts," *Jahrbücher für Nationalökonomie und Statistik*, 224(1+2), 51–64.
- HEILEMANN, U. UND H.O. STEKLER (2003): "Has the accuracy of German macroeconomic forecasts improved?," *Discussion Paper of the German Research Council's Research Centre 475*. 31/03.
- HINZE, J. (2005): "Konjunkturprognosen: Falsche Erwartungen an Treffgenauigkeit," *Wirtschaftsdienst, HWWA*, 85(2), 117–123.

- HOGG, R.V. UND CRAIG, A.T. (1995): *Introduction to Mathematical Statistics*. Prentice-Hall, fifth edn.
- HOLDEN, K. UND D.A. PEEL (1990): “On testing for unbiasedness and efficiency of forecasts,” *The Manchester School of Economic Social Studies*, 58, 120–127.
- NIERHAUS, W. (2001): “Konjunkturprognosen und Prognoserisiko,” *ifo Schnelldienst*, 54(16), 17–21.
- (2005): “Wirtschaftskonjunktur 2004: Prognose und Wirklichkeit,” *ifo Schnelldienst*, 58(3), 26–30.
- PONS, J. (2000): “The Accuracy of IMF and OECD Forecasts for G7 Countries,” *Journal of Forecasting*, 19, 53–63.
- RICHARDSON, P. (1988): “The structure and simulation properties of OECD’s interlink model,” *OECD Economic Studies No. 10*, pp. 57–122.
- RUOSS, E. UND M. SAVIOZ (2002): “Wie gut sind BIP-Prognosen? Eine Untersuchung für die Schweiz,” *Schweizerische Nationalbank, Quartalsheft 3/2002*.
- SCHRÖDER, M. (2002): “Erstellung von Prognosemodellen,” *in: M. Schröder, Finanzmarkt-Ökonometrie, Stuttgart, Schäfer-Poeschel Verlag*, pp. 397–462.