ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Benner, Joachim; Meier, Carsten-Patrick

Article — Digitized Version

Was leisten Stimmungsindikatoren für die Prognose des realen Bruttoinlandsprodukts in Deutschland? Eine Echtzeit-Analyse

Die Weltwirtschaft

Provided in Cooperation with:
Kiel Institute for the World Economy – Leibniz Center for Research on Global Economic Challenges

Suggested Citation: Benner, Joachim; Meier, Carsten-Patrick (2005) : Was leisten Stimmungsindikatoren für die Prognose des realen Bruttoinlandsprodukts in Deutschland? Eine Echtzeit-Analyse, Die Weltwirtschaft, ISSN 0043-2652, Springer, Heidelberg, Iss. 3, pp. 341-355

This Version is available at:
https://hdl.handle.net/10419/3725

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Was leisten Stimmungsindikatoren für die Prognose des realen Bruttoinlandsprodukts in Deutschland? Eine Echtzeit-Analyse

Von **Joachim Benner** und **Carsten-Patrick Meier**

Motivation

Frühindikatoren sind für Konjunkturforscher und Finanzmarktteilnehmer ein wichtiges Instrument, um die zukünftige wirtschaftliche Entwicklung abzuschätzen. Da die Daten der volkswirtschaftlichen Gesamtrechnung für das aktuelle Quartal in Deutschland erst mit etwa zwei Monaten Verzögerung vorliegen, kommt den Konjunkturindikatoren hier eine besonders wichtige Rolle zu. Aber auch zur Prognose der allernächsten Zukunft dürften sie die wichtigste Informationsquelle darstellen. Konjunkturindikatoren können realwirtschaftliche Größen, wie etwa Auftragseingänge oder sektorale Produktionszahlen, aber auch monetäre Größen, wie Zinsen oder Geldmengenaggregate, sein. Noch informativer sind Umfragedaten zur Einschätzung von gegenwärtiger oder zukünftiger Lage. In ihnen sind alle wesentlichen Fakten und Meinungen enthalten.

Wesentlich für einen Indikator ist, dass er die nachfolgende Entwicklung der vorherzusagenden gesamtwirtschaftlichen Produktionsdaten mit einem möglichst weiten und stabilen Vorlauf zumindest teilweise vorwegnimmt. Insbesondere bei der Vorhersage von konjunkturellen Wendepunkten können Frühindikatoren dem Konjunkturforscher hilfreich sein. Eine Zunahme der Auftragseingänge oder etwa sich aufhellende Stimmungsindikatoren können beispielsweise ein Anzeichen für einen bevorstehenden Anstieg des Bruttoinlandsprodukts nach einer Rezessionsphase sein. Bekannteste und zuverlässigste Indikatoren für die deutsche Konjunktur waren in der Vergangenheit die vom Münchner ifo Institut für Wirtschaftsforschung durch monatliche Umfragen bei über 7 000 Entscheidungsträgern in der gewerblichen Wirtschaft erhobenen „Konjunkturtest"-Ergebnisse.

Eine Reihe von Studien hat sich in den vergangenen Jahren mit den Prognoseeigenschaften verschiedener Konjunkturindikatoren in Deutschland befasst.[1] Dabei wurde aber dem Aspekt der „Echtzeit"-Analyse wenig Bedeutung beigemessen. Das vorliegende Papier untersucht die Prognoseeigenschaft einer Reihe von Stimmungsindikatoren für das Bruttoinlandsprodukt in Deutschland. Die Spezifikation des Prognosemodells erfolgt dabei rekursiv – dies bedeutet, dass für jedes Quartal im untersuchten Zeitraum das Modell neu bestimmt wird. Außerdem wird ein Echtzeit(„real-time")-Datensatz verwendet. Durch dieses Vorgehen wird gewährleistet, dass nur Informationen in das Modell eingehen, die zum Zeitpunkt einer Prognoseerstellung auch tatsächlich bekannt waren. In einer realen Prognosesituation liegen Daten nur zum jeweiligen Revisionsstand (z.B. Erst-

[1] So etwa Benner und Meier (2004), Breitung und Jagodzinski (2001), Döpke et al. (1994), Fritsche (1999, 2000), Fritsche und Stephan (2000), Funke (1997), Hüfner und Schröder (2002) sowie Langmantel (1999).

© 2005 Institut für Weltwirtschaft an der Universität Kiel DOI: 10.1007/s10289-005-0019-9

veröffentlichung, erste Revision usw.) vor, die insbesondere am für die Prognose wichtigen aktuellen Rand vom später revidierten Stand abweicht. Die Ergebnisse der Prognosen werden dann mit denen eines Referenzmodells verglichen.

Daten

Die vom Statistischen Bundesamt veröffentlichten Zahlen des Bruttoinlandsprodukts und ihre Komponenten sind mit einer Reihe von Ungenauigkeiten behaftet. Revisionen ergeben sich sowohl durch Veränderungen im Datenstand als auch durch Veränderungen der Saisonfaktoren des Saisonbereinigungsverfahren, die am aktuellen Rand vergleichsweise unsicher zu schätzen sind. Oftmals stehen daher die endgültigen Werte erst Jahre später fest. Zur Ausschaltung der in Verbindung mit Revisionen auftretenden Verzerrungen bei der Beurteilung der Prognosegüte bedarf es daher Echtzeit-Daten, da zum Zeitpunkt der Prognoseerstellung nur die erstveröffentlichten Daten vorliegen.[2]

In dieser Untersuchung verwenden wir daher einen Echtzeit-Datensatz für das reale saisonbereinigte Bruttoinlandsprodukt, das von der Deutschen Bundesbank in den „Saisonbereinigten Wirtschaftszahlen" veröffentlicht wird. In das jeweilige Prognosemodell gehen somit nur die zum damaligen Zeitpunkt bekannten Werte ein. Die erste Datenreihe entstammt aus dem I. Quartal 1973, die letzte aus dem I. Quartal 2004. Insgesamt besteht der Datensatz aus 128 Zeitreihen, die bis zum I. Quartal 1962 zurückreichen und zu verschiedenen Zeitpunkten enden. So reicht die erste Zeitreihe beispielsweise bis zum IV. Quartal 1972, während die letzte Reihe im III. Quartal 2004 endet. Der Datensatz entspricht somit einer 171×128-Matrix. Für die hier verwendeten Konjunkturindikatoren kann auf eine Echtzeit-Betrachtung verzichtet werden, da sie keinen Revisionen unterliegen.

Zu beachten ist, dass wir einige Annahmen beim Erstellen des Datensatzes treffen mussten. Da in den „Saisonbereinigten Wirtschaftszahlen" nur Zahlen für die letzten zehn Jahre veröffentlicht werden, haben wir für die davor liegenden Zeitpunkte die letzten zuvor veröffentlichten Werte verwendet. Es ist jedoch nicht davon auszugehen, dass diese Zahlen großen Revisionen unterzogen wurden. Bis in die frühen neunziger Jahre entsprechen unsere Daten dem Bruttosozialprodukt, danach dem Bruttoinlandsprodukt. Der Unterschied zwischen den beiden Reihen ist für Deutschland allerdings recht klein. Zu beachten ist außerdem, dass zwischen dem II. Quartal 1975 und dem I. Quartal 1991 keine arbeitstägliche Bereinigung stattgefunden hat. Wir haben daher eine eigene Arbeitstagebereinigung vorgenommen.[3] Des Weiteren beinhalten die Daten bis zum I. Quartal 1995 nur Westdeutschland. Für den Zeitraum danach verwenden wir Zahlen für Gesamtdeutschland. Um einen Sprung in den Daten zu vermeiden, haben wir die Zeitreihen verkettet und für die Zeit vor 1995 die Änderungsraten von West-

[2] Eine Darstellung der Bedeutung von Datenrevisionen findet sich beispielsweise in Croushore und Stark (2002), Koenig et al. (2001) und Runkle (1998).

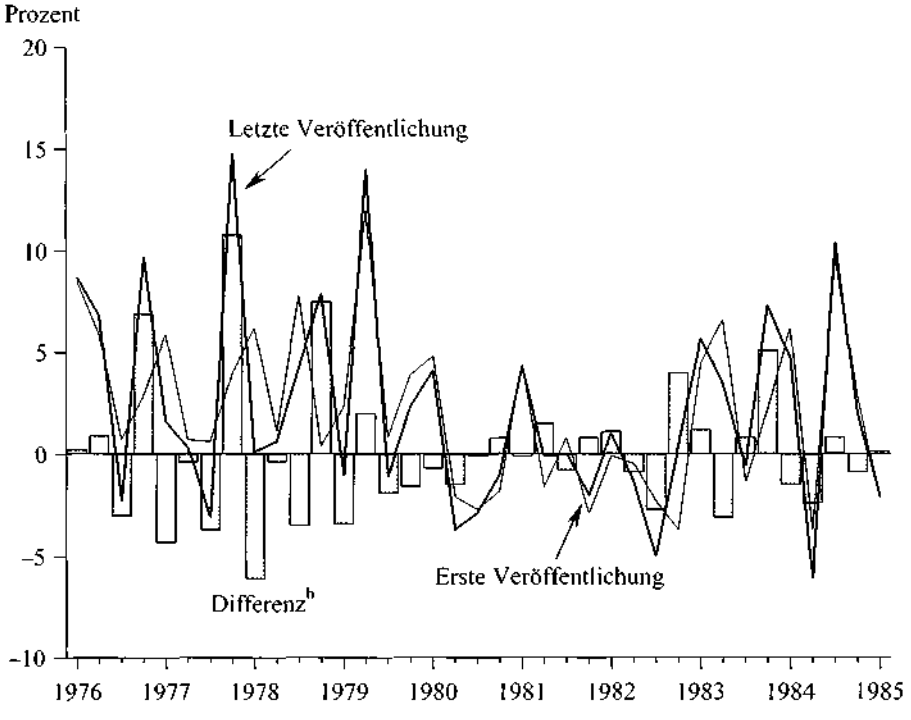
[3] Für diesen Zeitraum liegen offizielle arbeitstäglich bereinigte Änderungsraten des Bruttosozialprodukts vor. Im Vergleich zu den von uns bereinigten Raten sind die Unterschiede gering.

deutschland verwendet (vgl. hierzu auch das Vorgehen des Sachverständigenrats 2001: 459).

Die Bedeutung von Echtzeit-Daten wird in Abbildung 1 deutlich. Dargestellt ist die Veränderung des Bruttoinlandsprodukts gegenüber dem Vorquartal in laufender Jahresrate für den Zeitraum I/1976–I/1985. Die dünne Linie entspricht den Werten der Erstveröffentlichung im II. Quartal 1985, die zweite Linie bildet den Wert der letzten Veröffentlichung ab. Die Balken stellen die Differenz beider Reihen in Prozentpunkten dar. Es ist deutlich zu erkennen, dass einige Quartale erheblichen Revisionen unterzogen wurden.

Voraussetzung für einen Frühindikator ist, dass es einen theoretisch begründbaren Zusammenhang mit der konjunkturellen Entwicklung gibt, um Scheinkorrelationen auszuschließen. Zudem sollte das Messkonzept im Zeitablauf unverändert bleiben, um Strukturbrüche in der Zeitreihe zu vermeiden. Auch ist die Aussagekraft eines Indikators, der permanenten Änderungen des Messverfahrens unterzogen wird, gering.

Abbildung 1:
Bruttoinlandsprodukt[a] I/1976–I/1985 (erste und letzte amtliche Veröffentlichung)



[a] Veränderung gegenüber dem Vorquartal auf Jahresbasis. — [b] In Prozentpunkten.

Quelle: Eigene Berechnungen.

Die vorliegende Untersuchung verwendet zur Prognose des realen Bruttoinlandsprodukts eine Reihe von Stimmungsindikatoren. Hierbei handelt es sich um die ifo-Geschäftserwartungen, die ifo-Geschäftslage sowie das ifo-Geschäftsklima.[4] Die ifo-Inikatoren werden monatlich durch Befragung von etwa 7 000 Unternehmen der gewerblichen Wirtschaft erhoben. Die Unternehmen geben dabei eine Einschätzung ihrer gegenwärtigen Lage und ihrer Erwartungen für die nächsten sechs Monate ab. Aus dem Mittelwert beider Befragungen wird das Geschäftsklima ermittelt. Der Untersuchungszeitraum erstreckt sich vom I. Quartal 1965 bis zum III. Quartal 2004. Die mithilfe der Indikatoren erstellten Prognosen beginnen ab dem I. Quartal 1976; ab diesem Zeitpunkt verwenden wir dann für das reale saisonbereinigte Bruttoinlandsprodukt Echtzeit-Daten. Zum Vergleich haben wir die Modelle auch noch mit den endgültigen Daten, also nicht in Echtzeit, geschätzt.

Auswahl der Prognosemodelle

Wir verwenden zwei Verfahren für die Schätzung des Prognosemodells. Beide basieren auf Vektorautoregressionen für die Veränderungsrate des Bruttoinlandsprodukts gegenüber dem Vorquartal,[5] ΔBIP_t, und die jeweilige Indikatorvariable X (z.B. ifo-Geschäftserwartungen). Das einfache lineare Prognosemodell lautet

$$\Delta BIP_t = a_1 + \sum_{i=1}^{4} \beta_{1i} \Delta BIP_{t-i} + \sum_{j=1}^{4} \delta_{1j} X_{t-j} + \varepsilon_{1t}, \tag{1}$$

$$X_t = a_2 + \sum_{k=1}^{4} \beta_{2k} X_{t-k} + \sum_{n=1}^{4} \delta_{2n} \Delta BIP_{t-n} + \varepsilon_{2t}. \tag{2}$$

Hierbei sind β_1, β_2, δ_1 und δ_2 die zu schätzenden Parameter, α_1 und α_2 die Konstanten und ε_1 bzw. ε_2 die Störterme. Die Schätzung erfolgt mittels des Kleinsten-Quadrate-Verfahrens. Wir verwenden dabei den ersten bis vierten Lag der entsprechenden Variablen. Als Referenzmodell dient ein einfaches autoregressives Modell für die Veränderung des realen Bruttoinlandsprodukts:

$$\Delta BIP_t = a + \sum_{i=1}^{4} \beta_i \Delta BIP_{t-i} + \varepsilon_t. \tag{3}$$

Die Prognosegüte vektorautoregressiver Modelle hängt entscheidend von der Spezifikation der Verzögerungsstruktur ab. Daher verwenden wir ein weiteres Modellauswahlverfahren. Die Gleichung für die Indikatorvariable ist in Fehler-Korrektur-Form spezifiziert.[6] Da bei dieser Modellspezifikation die Verzöge-

---

[4] Neben den ifo-Indikatoren gibt es noch eine Reihe weiterer Stimmungsindikatoren. Eine Untersuchung dieser Indikatoren ist nach dem hier verwendeten Verfahren nicht möglich, da sie erst für einen zu kurzen Zeitraum zur Verfügung steht, um eine verlässliche Aussage über ihre Prognoseeigenschaft zu machen.

[5] $\Delta BIP_t = (BIP_t - BIP_{t-1}) \div BIP_{t-1}$.

[6] Siehe hierzu auch Benner und Meier (2004).

rungen in Veränderungsraten in die Gleichung eingehen, sollte das Multikollinearitätsproblem geringer und damit die Modellselektion besser sein. Die Verzögerungsstruktur wird nun nicht mehr fest vorgegeben, sondern mithilfe eines Modellauswahlprozesses ermittelt. Das selektive Prognosemodell lautet:

$$\Delta BIP_t = a_1 + X_{t-1} + \sum_{i=1}^{p_1} \beta_{1i} \Delta BIP_{t-1} + \sum_{j=1}^{p_2} \delta_{1i} \Delta X_{t-j} + \varepsilon_{1t}, \tag{4}$$

$$X_t = a_2 + X_{t-1} + \sum_{k=1}^{p_3} \beta_{2k} \Delta X_{t-k} + \sum_{n=1}^{p4} \delta_{2k} \Delta BIP_{t-n} + \varepsilon_{2t}. \tag{5}$$

Bei dem Modellauswahlprozess wird eine variable autoregressive Verzögerungsstruktur auf Basis des Bayesianischen Informationskriterium (BIC)[7] ermittelt. Bei der Modellselektion stützen wir uns auf ein Parameterreduktionsverfahren, das auf der sukzessiven Eliminierung von Koeffizienten mit niedrigen empirischen t-Werten beruht. Brüggemann und Lütkepohl (2001) zeigen, dass ein derartiger Selektionsprozess in gleicher Weise vorgeht wie eine sukzessive Variablenelimination nach einem Zeitreihenkriterium wie dem BIC. Bei geeigneter Wahl der kritischen t-Werte führt dieses Verfahren somit zu der gleichen Verzögerungsstruktur wie die deutlich aufwendigere Suche nach dem geringsten BIC-Wert über alle möglichen Kombinationen von Koeffizienten. Als Nebenbedingung fordern wir dabei, dass das auf diese Weise gefundene Modell autokorrelationsfreie Residuen aufweisen muss.[8]

Vergleich der Güte der Modellprognosen

Zur Beurteilung der Leistungsfähigkeit von Prognosemodellen werden üblicherweise rollierende Prognosen außerhalb des für die Schätzung der Parameter verwendeten Stichprobenbereichs, sogenannte Out-of-sample-Prognose, herangezogen. Dadurch wird sichergestellt, dass die Modellprognose nur auf Informationen beruht, die auch in einer echten Prognosesituation zu Beginn des jeweiligen Prognosezeitraums zur Verfügung gestanden hätten. Für eine Anzahl von Out-of-sample-Prognosen lassen sich dann Maße wie die mittlere Abweichungen zu der zu prognostizierenden Größe als Beurteilungskriterium berechnen. Die Modellspezifikation erfolgt dabei rekursiv. In diesem Fall wird zu jedem Zeitpunkt $T_1 + j$ mit den bis dahin zu Verfügung stehenden Daten der Modellselektionsprozess durchgeführt und auf der Basis des so gefundenen Modells die Prognose berechnet.

---

[7] Das BIC wird berechnet als $T \ln(\hat{\varepsilon}'\hat{\varepsilon}) + p \ln(T)$, wobei T die Anzahl der Beobachtungen, $\hat{\varepsilon}'\hat{\varepsilon}$ die geschätzte Residuenquadratsumme und P die Anzahl der Parameter des geschätzten Modells darstellen.

[8] Kilian (2001) stellt in einer Simulation fest, dass das BIC Modelle auswählt, die zu gering parametrisiert sind, um die Dynamik der unterstellten Prozesse korrekt abzubilden. Er empfiehlt für kleinere Stichproben die Verwendung des AIC. Er berücksichtigt allerdings nicht die Verwendung von Tests auf Freiheit von Autokorrelation als weiteres Modellselektionskriterium.

Die Out-of-sample-Prognosen werden ermittelt, indem das Modell zunächst anhand von Daten bis zum Zeitpunkt T_1 geschätzt wird – wobei T_1 deutlich vor dem Ende der verfügbaren Beobachtungen der zu prognostizierenden Zeitreihe in T liegt – und dann über den Zeitraum $T_1 + 1$ bis $T_1 + n$ Prognosen berechnet werden. Anschließend wird der Stützbereich um eine Beobachtung erweitert, die Schätzung erneuert und die Prognose für $T_1 + 2$ bis $T_1 + n + 1$ ermittelt. Setzt man diesen Prozess fort bis $T_1 = T - 1$, so erhält man eine Zeitreihe mit $T - T_1$ 1-Schritt-Prognosen, eine weitere Reihe mit $T - T_1 - 1$ 2-Schritt-Prognosen usw. sowie eine Reihe mit $T - T_1 - n + 1$ n-Schritt-Prognosen. Die Abweichung zwischen der Prognose zum Zeitpunkt $T_1 + j$ über k Perioden ($k = 1, \ldots, n$) und dem tatsächlichen Wert der zu prognostizierenden Größe zu diesem Zeitpunkt $T_1 + j + k$ wird als Prognosefehler $e(T_1 + j,k)$ bezeichnet. Im Folgenden vergleichen wir die Prognosegüte der Modelle anhand der Fehler von Out-of-sample-Prognosen für die Veränderungsrate des Bruttoinlandsprodukts gegenüber dem Vorjahr über Prognosehorizonte von 1, 2, 3, 4 und 8 Quartale.

Zur Beurteilung der Prognosegüte berechnen wir das Theilsche U, also den Quotienten aus der Wurzel aus dem mittleren quadratischen Prognosefehler (RMSE) des Prognosemodells und dem RMSE einer Vergleichsprognose. Die Vergleichsprognose basiert in dieser Untersuchung auf einem autoregressiven Modell („naive Prognose"). Der RMSE einer Prognose zum Zeitpunkt T_{1+j} für den Prognosehorizont k ist definiert als

$$RMSE = \sqrt{\frac{1}{T - T_1 - k + 1} \sum_{j=1}^{T-T_1-k+1} e^2(T_1 + j,k)}, \tag{6}$$

das Theilsche U ergibt sich als $U = \frac{RMSE_{Modell}}{RMSE_{naiv}}$.

Eine dem autoregressiven Modell überlegene Prognose weist dabei einen Wert kleiner als eins auf. Ihr Prognosefehler ist also kleiner als der des naiven Prognosemodells. Um zu erkennen, ob die Unterschiede zwischen den *RMSE* auch statistisch signifikant sind, berechnen wir außerdem den Test auf Gleichheit der Prognosefehler von Diebold und Mariano (1995), allerdings, der Literatur folgend, mit der Modifikation von Harvey et al. (1997).

Prognosegüte mit Echtzeit-Daten

Tabelle 1 zeigt die Prognoseergebnisse bei rekursiver Auswahl der Modellstruktur für Prognosehorizonte von ein, zwei, drei, vier und acht Quartalen mit Echtzeit-Daten. Die Prognosen beginnen im I. Quartal 1976 und enden im III. Quartal 2004. Als Maßstab für die Prognosegüte und die Berechnung des RMSE wurden die letzten uns bekannten Daten berechnet (III. Quartal 2004).[9] In Abbildung 2 sind exemplarisch die Ergebnisse des selektiven Prognosemodells

[9] Stand: 4. Quartal 2004.

Tabelle 1:
RMSE und Theils *U* für den Zeitraum I/1976–III/2004 (Echtzeit-Daten, Vergleich zur letzten amtlichen Veröffentlichung)[a]

Prognoseschritte	Anzahl der Quartale				
	1	2	3	4	8
ifo-Geschäftslage$_{linear}$	1,265	1,626	1,902	2,327	2,304
	0,95	0,97	1,05	1,18	1,10
	(0,243)	(0,576)	(0,511)	(0,082)	(0,020)
ifo-Geschäftsklima$_{linear}$	1,212	1,479	1,757	2,195	2,338
	0,91	0,88	0,97	1,11	1,11
	(0,049)	(0,026)	(0,656)	(0,189)	(0,005)
ifo-Geschäftserwartungen$_{linear}$	1,154	1,379	1,599	1,957	2,413
	0,87	0,82	0,88	0,99	1,15
	(0,001)	(0,001)	(0,102)	(0,929)	(0,029)
Autoregressiv$_{linear}$	1,329	1,680	1,812	1,973	2,101
ifo-Geschäftslage$_{selektiv}$	1,302	1,613	1,798	2,187	2,236
	0,98	0,99	1,00	1,08	1,08
	(0,627)	(0,737)	(0,963)	(0,342)	(0,205)
ifo-Geschäftsklima$_{selektiv}$	1,245	1,481	1,664	1,991	2,236
	0,94	0,91	0,92	0,98	1,08
	(0,131)	(0,053)	(0,181)	(0,785)	(0,011)
ifo-Geschäftserwartungen$_{selektiv}$	1,190	1,363	1,517	1,766	2,086
	0,90	0,83	0,84	0,87	1,01
	(0,007)	(0,002)	(0,029)	(0,145)	(0,738)
Autoregressiv$_{selektiv}$	1,325	1,635	1,803	2,027	2,073

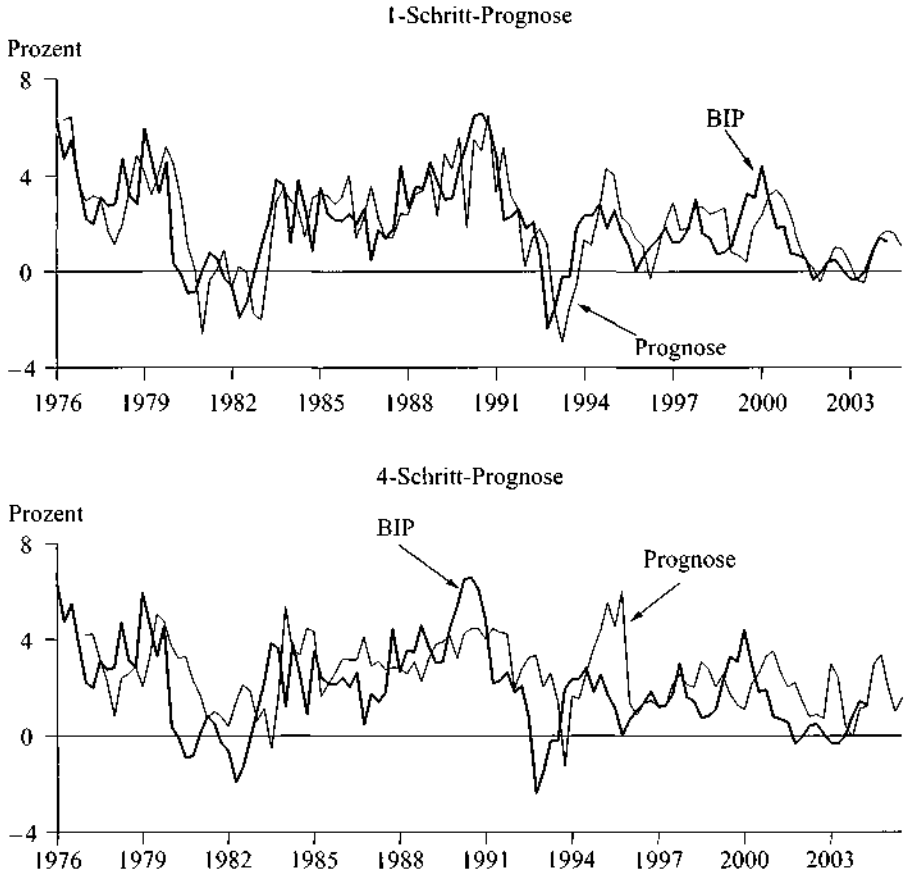
[a] Für jeden Indikator geben die Werte in der ersten Zeile die Wurzel aus dem mittleren quadratischen Prognosefehler (RMSE) für den jeweiligen Prognoseschritt, die zweite Zeile den Quotient aus dem RMSE des Modells und dem RMSE der autoregressiven Prognose (Theils *U*) und die Werte in Klammern darunter das marginale Signifikanzniveau des modifizierten Diebold-Mariano-Tests an.

Quelle: Eigene Berechnungen.

unter Verwendung der ifo-Geschäftserwartung für 1- und 4-Schritt-Prognosen dargestellt. Die dicke Linie bildet dabei die Veränderung des Bruttoinlandsprodukt der letzten uns bekannten Veröffentlichung ab. Deutlich zu erkennen ist, dass die Prognosegüte mit zunehmendem Prognosehorizont abnimmt.

Für 1- und 2-Schritt-Prognosen führt bei der linearen Schätzung auf einem Signifikanzniveau von 5 Prozent das ifo-Geschäftsklima und die ifo-Geschäftserwartungen zu einer besseren Prognose als das autoregressive Modell. Legt man ein großzügigeres Niveau von 15 Prozent zugrunde, so gelingt dies den ifo-Geschäftserwartungen auch für 3-Schritt-Prognosen. Die ifo-Geschäftslage führen nicht zu einer signifikanten Verbesserung der Prognosegüte. Den geringsten Prognosefehler insgesamt für Prognosen bis zu vier Quartalen liefern die Geschäftserwartungen. Ab dem IV. Quartal ist keines der Prognosemodelle

Abbildung 2:
Prognosemodell mit den ifo-Geschäftserwartungen im Vergleich zur tatsächlichen Veränderung des realen Bruttoinlandsprodukts 1976–2003



Quelle: Eigene Berechnungen.

signifikant besser als das autoregressive Modell. Für die Lage und das Klima liefern die Prognosemodelle sogar einen größeren RMSE. Für Prognosen über mehr als ein Jahr bis zu zwei Jahren gilt dies für alle Indikatoren.

Für das selektive Verfahren ist das Bild ähnlich. Insgesamt führt das selektive Modellauswahlverfahren zu geringeren Prognosefehlern als die einfache lineare Schätzung. Dies gilt sowohl für die Prognosemodelle als auch für das autoregressive Modell. Den geringsten Prognosefehler erhält man bei 1- bis 4-Schritt-Prognosen auch hier unter Verwendung der Geschäftserwartungen. Auf einem Signifikanzniveau von 5 Prozent sind sie hier sogar signifikant besser als das naive Modell. Hingegen liefert nun das Geschäftsklima bei 1-Schritt-Prognosen

keine signifikant überlegenen Prognosen. Dies ist darauf zurückzuführen, dass das autoregressive Modell bei selektiver Modellauswahl relativ besser geworden ist.

In einem weiteren Schritt haben wir die RMSE der Prognosemodelle auf Basis der erstveröffentlichten Daten berechnet. Maßstab der Prognosegüte ist nun also, wie gut es dem Modell gelingt, die zuerst von der Bundesbank veröffentlichte Zahl für das saisonbereinigte reale Bruttoinlandsprodukt vorherzusagen. Die Ergebnisse sind in Tabelle 2 abgebildet. Insgesamt sind die Prognosefehler geringer. Offenbar können für Prognosen bis zu einem Jahr die erstveröffentlichten Zahlen besser vorhergesagt werden. Ab dem vierten Schritt werden die Prognosefehler jedoch größer; die Prognose der revidierten Zahlen gelingt dann also besser. Die Geschäftserwartungen liefern für die Prognosen über ein, zwei und drei Quartale erneut signifikant bessere Ergebnisse als das naive Modell. Mit dem

Tabelle 2:
RMSE und Theils *U* für den Zeitraum I/1976–III/2004 (Echtzeit-Daten, Vergleich zur ersten amtlichen Veröffentlichung)[a]

Prognoseschritte	Anzahl der Quartale				
	1	2	3	4	8
ifo-Geschäftslage$_{linear}$	0,947	1,418	1,793	2,278	2,384
	0,92	0,95	1,01	1,13	1,09
	(0,169)	(0,366)	(0,865)	(0,142)	(0,018)
ifo-Geschäftsklima$_{linear}$	0,932	1,323	1,668	2,164	2,411
	0,91	0,88	0,94	1,08	1,10
	(0,122)	(0,042)	(0,396)	(0,314)	(0,008)
ifo-Geschäftserwartungen$_{linear}$	0,916	1,270	1,557	1,958	2,456
	0,89	0,85	0,88	0,97	1,12
	(0,029)	(0,005)	(0,080)	(0,756)	(0,080)
Autoregressiv$_{linear}$	1,028	1,500	1,770	2,010	2,188
ifo-Geschäftslage$_{selektiv}$	1,002	1,442	1,746	2,179	2,302
	0,94	0,96	0,98	1,04	1,06
	(0,198)	(0,410)	(0,733)	(0,597)	(0,382)
ifo-Geschäftsklima$_{selektiv}$	0,947	1,314	1,570	1,974	2,340
	0,89	0,88	0,88	0,94	1,08
	(0,036)	(0,026)	(0,061)	(0,439)	(0,020)
ifo-Geschäftserwartungen$_{selektiv}$	0,933	1,250	1,462	1,773	2,180
	0,88	0,83	0,82	0,85	1,00
	(0,014)	(0,007)	(0,018)	(0,098)	(0,911)
Autoregressiv$_{selektiv}$	1,063	1,498	1,785	2,090	2,175

[a] Für jeden Indikator geben die Werte in der ersten Zeile die Wurzel aus dem mittleren quadratischen Prognosefehler (RMSE) für den jeweiligen Prognoseschritt, die zweite Zeile den Quotient aus dem RMSE des Modells und dem RMSE der autoregressiven Prognose (Theils *U*) und die Werte in Klammern darunter das marginale Signifikanzniveau des modifizierten Diebold-Mariano-Tests an.

Quelle: Eigene Berechnungen.

Geschäftsklima gelingt dies mit dem selektiven Prognosemodell ebenfalls. Auf einem Signifikanzniveau von 10 Prozent erzeugen die ifo-Geschäftserwartungen sogar Prognosen über ein Jahr geringere Prognosefehler. Die Geschäftslage liefert wiederum in keinem Prognoseschritt signifikant bessere Ergebnisse als das autoregressive Modell. Tendenziell schneiden die Indikator-Modelle gegenüber dem autoregressiven Modell bei Prognosen bis zu vier Quartalen besser ab. Danach dominiert das naive Modell.

Prognosegüte mit kontemporärer Korrelation

Ein Vorteil der Vertrauensindikatoren ist, dass sie bereits am Ende eines Quartals zur Verfügung stehen, während das reale Bruttoinlandsprodukt für ein Quartal

Tabelle 3:
RMSE und Theils *U* für den Zeitraum I/1976–III/2004 (Echtzeit-Daten, Vergleich zur letzten amtlichen Veröffentlichung, kontemporäre Korrelation)[a]

Prognoseschritte	Anzahl der Quartale				
	1	2	3	4	8
ifo-Geschäftslage$_{linear}$	1,199	1,418	1,726	2,042	2,294
	0,90	0,84	0,95	1,03	1,09
	(0,116)	(0,035)	(0,593)	(0,763)	(0,032)
ifo-Geschäftsklima$_{linear}$	1,204	1,375	1,521	1,850	2,300
	0,91	0,82	0,84	0,94	1,09
	(0,083)	(0,006)	(0,052)	(0,549)	(0,014)
ifo-Geschäftserwartungen$_{linear}$	1,182	1,340	1,368	1,660	2,364
	0,89	0,80	0,75	0,84	1,13
	(0,020)	(0,002)	(0,006)	(0,161)	(0,032)
Autoregressiv$_{linear}$	1,329	1,680	1,812	1,973	2,101
ifo-Geschäftslage$_{selektiv}$	1,206	1,428	1,624	1,855	2,191
	0,91	0,87	0,90	0,92	1,06
	(0,158)	(0,092)	(0,295)	(0,461)	(0,255)
ifo-Geschäftsklima$_{selektiv}$	1,206	1,386	1,475	1,631	2,133
	0,91	0,85	0,82	0,80	1,03
	(0,077)	(0,026)	(0,037)	(0,075)	(0,258)
ifo-Geschäftserwartungen$_{selektiv}$	1,167	1,326	1,338	1,511	2,112
	0,88	0,81	0,74	0,75	1,02
	(0,007)	(0,007)	(0,007)	(0,028)	(0,360)
Autoregressiv$_{selektiv}$	1,325	1,635	1,803	2,027	2,073

[a] Für jeden Indikator geben die Werte in der ersten Zeile die Wurzel aus dem mittleren quadratischen Prognosefehler (RMSE) für den jeweiligen Prognoseschritt, die zweite Zeile den Quotient aus dem RMSE des Modells und dem RMSE der autoregressiven Prognose (Theils *U*) und die Werte in Klammern darunter das marginale Signifikanzniveau des modifizierten Diebold-Mariano-Tests an.

Quelle: Eigene Berechnungen.

erst zwei Monate später veröffentlicht wird. Um diesem Informationsvorsprung Rechnung zu tragen, haben wir im Folgenden für den Indikator eine kontemporäre Korrelation zugelassen. Wiederum haben wir zwei Schätzverfahren durchgeführt. Die Gleichung für die Veränderung des Bruttoinlandsprodukts im selektiven Modellansatz lautet nun:

$$(7) \qquad \Delta BIP_t = \alpha_1 + X_{t-1} + \sum_{i=1}^{P_1} \beta_{1i} \Delta BIP_{t-i} + \sum_{j=0}^{P_2} \delta_{1j} \Delta X_{t-j} + \varepsilon_{1t}.$$

Der Laufindex j beginnt in diesem Fall bei null. Analog wurde auch beim einfachen linearen Modell verfahren.

In Tabelle 3 sind die Ergebnisse auf Basis der letzten bekannten amtlichen Veröffentlichung dargestellt. Wie zu erwarten, sind die RMSE der Prognosemodelle

Tabelle 4:
RMSE und Theils U für den Zeitraum I/1976–III/2004 (Echtzeit-Daten, Vergleich zur ersten amtlichen Veröffentlichung, kontemporäre Korrelation)[a]

Prognoseschritte	Anzahl der Quartale				
	1	2	3	4	8
ifo-Geschäftslage$_{linear}$	0,869	1,168	1,560	1,969	2,377
	0,85	0,78	0,88	0,98	1,09
	(0,051)	(0,013)	(0,217)	(0,855)	(0,024)
ifo-Geschäftsklima$_{linear}$	0,930	1,181	1,444	1,802	2,381
	0,90	0,79	0,82	0,90	1,09
	(0,183)	(0,009)	(0,039)	(0,285)	(0,011)
ifo-Geschäftserwartungen$_{linear}$	0,970	1,222	1,405	1,669	2,414
	0,94	0,81	0,79	0,83	1,10
	(0,357)	(0,012)	(0,023)	(0,101)	(0,082)
Autoregressiv$_{linear}$	1,028	1,500	1,770	2,010	2,188
ifo-Geschäftslage$_{selektiv}$	0,891	1,184	1,477	1,783	2,248
	0,84	0,79	0,83	0,85	1,03
	(0,032)	(0,026)	(0,247)	(0,247)	(0,515)
ifo-Geschäftsklima$_{selektiv}$	0,924	1,199	1,408	1,614	2,232
	0,87	0,80	0,79	0,77	1,03
	(0,042)	(0,015)	(0,034)	(0,049)	(0,205)
ifo-Geschäftserwartungen$_{selektiv}$	0,944	1,213	1,345	1,535	2,189
	0,89	0,81	0,75	0,73	1,01
	(0,054)	(0,017)	(0,018)	(0,025)	(0,753)
Autoregressiv$_{selektiv}$	1,063	1,498	1,785	2,090	2,175

[a] Für jeden Indikator geben die Werte in der ersten Zeile die Wurzel aus dem mittleren quadratischen Prognosefehler (RMSE) für den jeweiligen Prognoseschritt, die zweite Zeile den Quotient aus dem RMSE des Modells und dem RMSE der autoregressiven Prognose (Theils U) und die Werte in Klammern darunter das marginale Signifikanzniveau des modifizierten Diebold-Mariano-Tests an.

Quelle: Eigene Berechnungen.

unter Verwendung der Stimmungsindikatoren kleiner als im Fall ohne kontemporäre Korrelation. Da das autoregressive Modell unverändert geblieben ist, sind somit auch die Werte für das Theilsche *U* geringer. Die selektiven Prognosemodelle liefern bei 1- bis 4-Schritt-Prognosen für die Geschäftserwartungen signifikant bessere Ergebnisse als das naive Modell; für das Geschäftsklima gilt dies auf einem Signifikanzniveau von 10 Prozent. Ähnlich sind die Ergebnisse im einfachen linearen Modell. Prognosen über vier Quartale sind hier allerdings nicht mehr signifikant besser. Auch mithilfe der ifo-Geschäftslage gelingt es, im Fall von 2-Schritt-Prognosen signifikant bessere Ergebnisse zu erzielen.

Die Ergebnisse im Vergleich zu den erstveröffentlichten Daten sind in Tabelle 4 abgebildet. Wie schon im zuvor untersuchten Fall sind auch hier die Prognosefehler tendenziell geringer. Die Geschäftslage scheint hier besonders gut geeignet zu sein, die erste Veröffentlichung für das Bruttoinlandsprodukt des aktuellen Quartals zu prognostizieren; sie liefert hier den geringsten RMSE. Über einen Prognosehorizont von drei bis vier Quartalen erreicht man hingegen mit den Geschäftserwartungen die besten Resultate.

Prognosegüte mit endgültigen Daten

In einem letzten Schritt haben wir nun für die Untersuchung die uns bekannten endgültigen Daten herangezogen. Die Schätzungen erfolgen also nicht mit Echtzeit-Daten, sondern den ex post heute bekannten Daten.

In Tabelle 5 sind die Ergebnisse (mit kontemporärer Korrelation) dargestellt. Die Ergebnisse sind ähnlich wie in der Untersuchung mit Echtzeit-Daten. Die Prognosefehler der Indikatormodelle sind auch hier geringer, wenn man die frühere Veröffentlichung der Indikatoren berücksichtigt. Die ifo-Geschäftserwartungen haben bei Prognosen bis zu einem Jahr tendenziell die geringsten RMSE. Im kontemporären Fall erzeugt die Geschäftslage für 1-Schritt-Prognosen die beste Prognose. Bei Prognosen über einen längeren Zeitraum ist hingegen das autoregressive Modell überlegen. Bei Prognosen über drei und vier Quartale scheinen die Indikatormodelle relativ zu den autoregressiven Modellen etwas schlechter abzuschneiden. Insgesamt sind die Prognosefehler tendenziell geringer als bei Verwendung von Echtzeit-Daten. Insbesondere das autoregressive Modell erzeugt durchweg niedrigere Prognosefehler. Zieht man also für die Analyse der Prognosegüte der Stimmungsindikatoren keine Echtzeit-Daten heran, so würde man ihre Prognoseleistung gegenüber dem Referenzmodell eher unterschätzen.

Besonders wichtig für die Untersuchung der Prognoseeigenschaften von Frühindikatoren ist eine rekursive Spezifikation des Prognosemodells. Bestimmt man das Prognosemodell hingegen über den gesamten zu untersuchenden Zeitraum (full sample) und nicht für jedes Quartal neu, ergeben sich deutlich niedrigere Prognosefehler. In Tabelle 6 sind die Ergebnisse, die sich bei einer Full-sample-Spezifikation ergeben, anhand des linearen Prognosemodells mit endgültigen Daten dargestellt. Die Prognosefehler sind im Vergleich zu einer rekursiven Spezifikation (siehe Tabelle 1) mit Echtzeit-Daten durchweg niedriger. Betrachtet man die Prognosegüte der Indikatormodelle, so ist diese deutlich besser als bei

Tabelle 5:
RMSE und Theils *U* für den Zeitraum I/1976–III/2004 (Ex-post-Daten, kontemporäre Korrelation)[a]

Prognoseschritte	Anzahl der Quartale				
	1	2	3	4	8
ifo-Geschäftslage$_{linear}$	0,866	1,190	1,588	1,988	2,255
	0,85	0,85	0,96	1,00	1,07
	(0,061)	(0,094)	(0,719)	(0,969)	(0,058)
ifo-Geschäftsklima$_{linear}$	0,871	1,129	1,417	1,816	2,230
	0,85	0,81	0,86	0,91	1,05
	(0,013)	(0,016)	(0,148)	(0,384)	(0,087)
ifo-Geschäftserwartungen$_{linear}$	0,887	1,097	1,266	1,653	2,308
	0,87	0,77	0,77	0,83	1,09
	(0,010)	(0,003)	(0,026)	(0,112)	(0,096)
Autoregressiv$_{linear}$	1,019	1,394	1,652	1,997	2,117
ifo-Geschäftslage$_{selektiv}$	0,867	1,155	1,503	1,873	2,205
	0,82	0,82	0,90	0,93	1,06
	(0,027)	(0,041)	(0,345)	(0,538)	(0,204)
ifo-Geschäftsklima$_{selektiv}$	0,887	1,127	1,354	1,654	2,225
	0,84	0,80	0,81	0,84	1,07
	(0,008)	(0,010)	(0,055)	(0,088)	(0,032)
ifo-Geschäftserwartungen$_{selektiv}$	0,892	1,090	1,211	1,532	2,113
	0,85	0,77	0,73	0,76	1,02
	(0,001)	(0,002)	(0,010)	(0,028)	(0,518)
Autoregressiv$_{selektiv}$	1,051	1,408	1,670	2,010	2,079

[a] Für jeden Indikator geben die Werte in der ersten Zeile die Wurzel aus dem mittleren quadratischen Prognosefehler (RMSE) für den jeweiligen Prognoseschritt, die zweite Zeile den Quotient aus dem RMSE des Modells und dem RMSE der autoregressiven Prognose (Theils *U*) und die Werte in Klammern darunter das marginale Signifikanzniveau des modifizierten Diebold-Mariano-Tests an.

Quelle: Eigene Berechnungen.

rekursiver Spezifikation. Wählt man keine „realistische" Modellspezifikation, überschätzt man also tendenziell die Prognoseleistung der Frühindikatoren.

Fazit

In dem vorliegenden Papier wurde die Prognosegüte alternativer Konjunkturindikatoren mittels Echtzeit-Daten analysiert. Unserer Ansicht nach kann nur so die Prognosegüte eines Indikators bewertet werden, da in die Untersuchung nur die zum Zeitpunkt einer Prognoseerstellung tatsächlich vorliegenden Informationen eingehen.

Aus den hier gewonnenen Ergebnissen kann man schließen, dass die ifo-Indikatoren nützliche Informationen für die Prognose des realen Bruttoinlands-

Tabelle 6:
RMSE und Theils *U* für den Zeitraum I/1976–III/2004 (Full-sample-Spezifikation, Ex-post-Daten)[a]

Prognoseschritte	Anzahl der Quartale				
	1	2	3	4	8
ifo-Geschäftslage$_{linear}$	0,834	1,104	1,392	1,758	1,891
	0,89	0,89	0,93	1,00	1,02
	(0,004)	(0,020)	(0,187)	(0,923)	(0,206)
ifo-Geschäftsklima$_{linear}$	0,830	1,033	1,294	1,612	1,873
	0,88	0,83	0,87	0,91	1,01
	(0,003)	(0,006)	(0,061)	(0,211)	(0,542)
ifo-Geschäftserwartungen$_{linear}$	0,811	0,993	1,239	1,527	1,895
	0,86	0,80	0,83	0,87	1,02
	(0,001)	(0,004)	(0,043)	(0,114)	(0,266)
Autoregressiv$_{linear}$	0,941	1,243	1,489	1,766	1,850

[a] Für jeden Indikator geben die Werte in der ersten Zeile die Wurzel aus dem mittleren quadratischen Prognosefehler (RMSE) für den jeweiligen Prognoseschritt, die zweite Zeile den Quotient aus dem RMSE des Modells und dem RMSE der autoregressiven Prognose (Theils *U*) und die Werte in Klammern darunter das marginale Signifikanzniveau des modifizierten Diebold-Mariano-Tests an.

Quelle: Eigene Berechnungen.

produkts enthalten. Wesentlich ist dabei auch die frühzeitige Veröffentlichung der Indikatoren. Für Prognosen bis zu einem Jahr sind die Geschäftserwartungen besonders gut geeignet, während die Geschäftslage hilfreich für die Schätzung des aktuellen Rands ist. Das entspricht auch der zugrunde liegenden Konzeption dieser beiden Stimmungsindikatoren. Für Prognosen über einen Zeitraum von mehr als einem Jahr sind die Frühindikatoren hingegen nicht geeignet. Hier gelingt es ihnen nicht, bessere Ergebnisse zu liefern als ein autoregressiver Prozess. Oft sind sie sogar signifikant schlechter. Unterschiede zeigen sich auch in den Prognosen der Erstveröffentlichung und der revidierten Zahlen. So wurden von den Prognosemodellen für Prognosehorizonte von ein bis vier Quartalen die erstveröffentlichten Werte des Bruttoinlandsprodukts teilweise besser vorhergesagt als die endgültigen Werte. Besonders wichtig bei der Untersuchung der Prognoseleistung von Frühindikatoren ist zudem eine realistische Spezifikation der Prognosemodelle.

Summary

In this paper we compare the forecasting ability of set of sentiment indicators for the GDP in Germany. The analyses is based on a vector autoregressive forecasting model. We estimate the parameters and the model selection recursively. We argue that the use of real-time data is important because those are the data that were available to a forecaster when he developed his model.

Literatur

Benner, J., und C.-P. Meier (2004). Prognosegüte alternativer Frühindikatoren für die Konjunktur in Deutschland. *Jahrbücher für Nationalökonomie und Statistik* 224 (6): 639–652.

Breitung, J., und D. Jagodzinski (2001). Prognoseeigenschaften alternativer Indikatoren für die Konjunkturentwicklung in Deutschland. *Konjunkturpolitik* 47 (4): 292–314.

Brüggemann, R., und H. Lütkepohl (2001). Lag Selection in Subset VAR Models With an Application to a U.S. Monetary System. In R. Friedmann (Hrsg.), *Econometric Studies. A Festschrift in Honour of Joachim Frohn*. Münster.

Croushore, D., und T. Stark (2002). Forecasting with a Real-time Data Set for Macroeconomists. *Journal of Macroeconomics* 24 (4): 507–531.

Diebold, F. X., und R. S. Mariano (1995). Comparing Predictive Accuracy. *Journal of Business and Economic Statistics* 13 (3): 253–263.

Döpke, J., J. W. Krämer und E. Langfeldt (1994). Konjunkturelle Frühindikatoren in Deutschland. *Konjunkturpolitik* 40 (2): 133–153.

Fritsche, U. (1999). Vorlaufeigenschaften von ifo-Indikatoren für Westdeutschland. DIW-Diskussionspapier 179. Deutsches Institut für Wirtschaftsforschung, Berlin.

Fritsche, U. (2001). Do Probit Models Help in Forecasting Turning Points in German Business Cycles? DIW-Diskussionspapier 241. Deutsches Institut für Wirtschaftsforschung, Berlin.

Fritsche, U., und S. Stephan (2000). Leading Indicators of German Business Cycles: An Assessment of Properties. *Jahrbücher für Nationalökonomie und Statistik* 222 (3): 289–315.

Funke, N. (1997). Predicting Recessions: Some Evidence for Germany. *Weltwirtschaftliches Archiv* 133 (1): 91–102.

Hüfner, F. P., und M. Schröder (2002). Prognosegehalt von ifo-Geschäftserwartungen und ZEW-Konjunkturerwartungen: Ein ökonometrischer Vergleich. *Jahrbücher für Nationalökonomie und Statistik* 222 (3): 316–336.

Harvey, D. I., S. J. Leybourne und P. Newbold (1997). Testing the Equality of Prediction Mean Squared Errors. *International Journal of Forecasting* 13 (2): 281–291.

Kilian, L. (2001). Impulse Response Analysis in Vector Autoregressions With Unknown Lag Order. *Journal of Forecasting* 20 (3): 161–179.

Koenig, E. F., S. Dolmas und J. Piger (2001). The Use and Abuse of „Real-time" Data in Economic Forecasting. Research Department Working Paper 00,04. Federal Reserve Bank of Dallas, Tex.

Langmantel, E. (1999). Das ifo-Geschäftsklima als Indikator für die Prognose des Bruttoinlandsproduktes. *ifo Schnelldienst* 16/17: 16–21.

Runkle, D. E. (1998). Revisionist History: How Data Revisions Distort Economic Policy Research. *Federal Reserve Bank of Minneapolis Quarterly Review* 22 (4): 3–12.

Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung (2001). *Für Stetigkeit – gegen Aktionismus. Jahresgutachten 2001/02*. Stuttgart.