

Unternehmens- versus Analystenbefragungen – Zum Prognosegehalt von ifo-Geschäftserwartungen und ZEW-Konjunkturerwartungen

Felix P. Hübner und Michael Schröder¹
Zentrum für Europäische Wirtschaftsforschung (ZEW)

ZEW-Discussion Paper No. 01 - 04

Februar 2001

Abstract

In der vorliegenden Arbeit untersuchen wir die Eignung der ifo-Geschäftserwartungen und der ZEW-Konjunkturerwartungen als Frühindikatoren für die deutsche Industrieproduktion. Anhand von Granger-Kausalitätstests wird gezeigt, dass die auf Umfragen unter Finanzanalysten basierenden ZEW-Konjunkturerwartungen einen signifikanten einmonatigen Vorlauf vor den ifo-Geschäftserwartungen haben. Wir führen dies auf unterschiedliche Erwartungsbildungen der beiden Gruppen zurück. Die ZEW-Konjunkturerwartungen erlauben Prognosen der Industrieproduktion für 3 bis 12 Monate im voraus, die signifikant besser als eine naive Vergleichsprognose und eine Prognose mit den ifo-Erwartungen sind. Die Vorhersagen mit Hilfe der ifo-Geschäftserwartungen sind hingegen auf sehr kurze Sicht besser. Mit Hilfe von Encompassing-Tests wird gezeigt, dass für einen Horizont von drei bis sechs Monaten eine Kombination beider Indikatoren bessere Prognoseergebnisse liefert als die Verwendung der jeweiligen Indikatoren alleine. Eine Analyse der Richtungsprognosen zeigt, dass diese mit den beiden Erwartungsindikatoren erheblich besser möglich sind, als mit der naiven Prognose oder einer Zufallsprognose.

JEL-Klassifikation: C52, C53, E37

Keywords:

Konjunkturerwartungen, Prognosemodelle, Prognosevergleich, Frühindikatoren

L 7, 1
Postfach 10 34 43
D-68034 Mannheim

Tel.: 0621 / 1235 - 140
Fax: 0621 / 1235 - 223
E-mail: schroeder@zew.de

¹ Für hilfreiche Kommentare und Anmerkungen danken wir Herbert S. Buscher. Birgit Sasse danken wir für unterstützende Vorarbeiten. Für alle verbleibenden Fehler und Irrtümer sind wir selbstverständlich selbst verantwortlich.

Non-Technical Summary

Ziel der vorliegenden Arbeit ist der Vergleich von ifo-Geschäftserwartungen und ZEW-Konjunkturerwartungen und die Analyse ihres Nutzens als Frühindikator für die wirtschaftliche Entwicklung in Deutschland. Beide Erwartungsindikatoren werden aus monatlichen Umfragen gebildet, wobei der ifo-Konjunkturtest Unternehmen direkt befragt und der ZEW-Finanzmarkttest Analysten aus Banken, Versicherungen und großen Industrieunternehmen als Teilnehmer hat. Untersucht wird in der Studie, inwiefern sich die Erwartungsbildung der Analysten von denen der Unternehmen unterscheidet und welche besser für eine Konjunkturprognose geeignet sind.

Sowohl ifo-Geschäftserwartungen als auch ZEW-Konjunkturerwartungen weisen einen signifikanten Vorlauf zur Jahresveränderungsrate der Industrieproduktion in Deutschland auf. Dieser beträgt für die ZEW-Konjunkturerwartungen bis zu 6 Monate, während bei den ifo-Erwartungen ein Vorlauf von maximal 4 Monaten vorliegt. Darüberhinaus zeigen Granger-Kausalitätstests, dass der ZEW-Indikator einen signifikanten Vorlauf von einem Monat vor dem ifo-Indikator hat.

Die anschließend durchgeführten out-of-sample Prognosetests ergeben, dass mit beiden Frühindikatoren die zukünftige Entwicklung der Industrieproduktion deutlich besser vorausgesagt werden kann als mit einer naiven Prognose, die lediglich die eigenen verzögerten Werte der Industrieproduktion beinhaltet. Eine Prognose mit den ZEW-Konjunkturerwartungen liefert im Bereich zwischen 3 und 12 Monaten signifikant bessere Ergebnisse als die naive Prognose. Bei Verwendung der ifo-Geschäftserwartungen lässt sich diese Signifikanz nur hinsichtlich der 6-Monatsprognosen nachweisen. Ein direkter Vergleich beider Indikatoren führt zu dem Ergebnis, dass die Prognose mit dem ifo-Indikator auf Sicht von einem Monat leicht besser ist, während Prognosen mit den ZEW-Erwartungen für die übrigen Zeiträume überlegen sind. Mit Hilfe von Encompassing-Tests wird darüberhinaus der Nachweis erbracht, dass bei 3- und 6-Schrittprognosen eine Kombination beider Indikatoren eine bessere Prognose liefert als die Verwendung jeweils beider Indikatoren alleine. Abschließend zeigt eine Untersuchung der Qualität der Richtungsprognosen, dass die Verwendung der Erwartungsdaten aus der ifo- und der ZEW-Umfrage jeweils zu besseren Ergebnissen führt als die naive Prognose. Beide Indikatoren sind bei den Richtungsprognosen als gleichwertig einzustufen.

Folgende Ergebnisse lassen sich aus der Untersuchung hinsichtlich der Erwartungsbildung von Unternehmen und Analysten ableiten:

- Die Erwartungsbildung von Unternehmen und Finanzanalysten scheint sich insbesondere hinsichtlich der Prognosegüte über verschiedene Zeiträume zu unterscheiden. Während mit den auf Unternehmensbefragungen basierenden ifo-Geschäftserwartungen für sehr kurze Zeiträume (1 Monat) ein besseres Progno-

seergebnis erreicht werden kann, erlauben die aus dem ZEW-Finanzmarkttest abgeleiteten Konjunkturerwartungen von Analysten bessere Prognosen auf mittlere und längere Sicht (3 bis 12 Monate).

- Bei mittelfristigen Prognosen (3 und 6 Monate) enthalten die Erwartungsindikatoren von ifo und ZEW einander ergänzende Informationen. Es ist daher sinnvoll, für diese Prognosehorizonte eine Kombination aus beiden Erwartungsindikatoren für die Einschätzung der zukünftigen Konjunkturlage zu nutzen.
- Im direkten Vergleich der Indikatoren zeigt sich, dass die ZEW-Konjunkturerwartungen den ifo-Geschäftserwartungen um etwa einen Monat signifikant vorauslaufen. Die ZEW-Erwartungsdaten erlauben daher eine kurzfristige Vorhersage der ifo-Geschäftserwartungen, die insbesondere für die Finanzmärkte von Interesse sein dürfte.

1 Einleitung

Frühindikatoren für die Konjunkturlage spielen für die Finanzmärkte eine entscheidende Rolle. Sie haben sowohl auf die Aktien- und Wechselkurse, als auch auf die Zinsentwicklung einen nicht unerheblichen Einfluss, indem sie die Erwartungen der Marktteilnehmer mitbestimmen. Da die Veröffentlichung der tatsächlichen Zahlen zum Bruttoinlandsprodukt eines Landes nur mit einer erheblichen Zeitverzögerung geschieht, werden Frühindikatoren auf den Kapitalmärkten viel beachtet.

Beispiele für solche Frühindikatoren sind zum einen Daten aus der öffentlichen Statistik, wie etwa die Auftragseingänge, oder Indikatoren, die aus der Befragung von Unternehmen direkt abgeleitet sind. In den USA etwa ist die Befragung von Einkaufsmanagern (NAPM-Index) als Konjunkturindikator verbreitet. Der für Deutschland meistbeachtete und älteste konjunkturelle Frühindikator ist der ifo-Geschäftsklima-Index für Westdeutschland, der auf dem ifo-Konjunkturtest beruht. Er wird monatlich vom ifo Institut für Wirtschaftsforschung, München, veröffentlicht.

Einen alternativen Ansatz verfolgt das Zentrum für Europäische Wirtschaftsforschung (ZEW), Mannheim, in seinem ebenfalls monatlich durchgeführten Finanzmarkttest. Die Adressaten dieser Umfrage sind Finanzanalysten und institutionelle Anleger aus Banken, Versicherungen und ausgewählten Industrieunternehmen. Seit Ende 1991 werden in dieser Umfrage die Einschätzungen der Teilnehmer bezüglich der Konjunkturlage in den wichtigsten Industriestaaten sowie der Entwicklung wichtiger Finanzmarktindikatoren wie Zinsen, Wechselkurse und Börsenindizes abgefragt und die Ergebnisse in Form des ZEW-Finanzmarktreports veröffentlicht.

Die Befragungsmethode und der zeitliche Ablauf der Befragung des ZEW-Finanzmarkttests sind sehr ähnlich denen des ifo-Konjunkturtests. Daher sollten auch die in den beiden Indikatoren potenziell berücksichtigten Informationen ungefähr gleich sein. Allerdings sind die jeweiligen Teilnehmergruppen ganz unterschiedlich: Bei der ifo-Umfrage geben die Teilnehmer Auskunft über die zukünftige Entwicklung der Ertragslage des Unternehmens, in dem sie tätig sind, während bei der Umfrage des ZEW Analysten nach ihrer Einschätzung der Gesamtkonjunktur befragt werden. Die Teilnehmer der ifo-Umfrage haben eine größere Detailkenntnis der Geschäftsentwicklung ihres Unternehmens und vermutlich auch der Branchenkonjunktur, die Teilnehmer der ZEW Umfrage dagegen haben besonders Erfahrung in der Prognose der Gesamtkonjunktur.

In unserer Studie untersuchen wir, in welcher Beziehung die Erwartungen der zwei unterschiedlichen Teilnehmergruppen zueinander stehen. Es könnte sein, dass der Informationsgehalt der Erwartungen der einen Teilnehmergruppe vollständig in demjenigen der anderen Teilnehmergruppe enthalten ist. In diesem Fall sollten nur die Erwartungen mit dem umfassenderen Informationsgehalt für eine Prognose ver-

wendet werden. Es könnte sich allerdings auch zeigen, dass die Erwartungen einander ergänzende Informationen enthalten und somit beide für die Beurteilung der zukünftigen Konjunkturlage herangezogen werden sollten.

Ein direkter Test des Informationsgehaltes der Erwartungen ist nicht möglich. Daher untersuchen wir die Prognosegüte unter Verwendung der zwei unterschiedlichen Erwartungsindikatoren und ziehen aus dem Ergebnis einen indirekten Rückschluss auf die in den Erwartungen enthaltenen Informationsmengen und deren Beziehungen. Wir führen dazu Analysen der Zeitreiheneigenschaften und der Prognosefähigkeit der ifo-Geschäftserwartungen und ZEW-Konjunkturerwartungen durch.² Die Ergebnisse sollen zeigen, ob die direkte Befragung von Unternehmen oder die Befragung von Finanzanalysten bessere Prognosedaten für die zukünftige Konjunktur zulässt.

Es gibt zahlreiche Studien, die die Prognosefähigkeit von ifo-Indikatoren untersuchen. Strigel (1985) gibt eine mehr deskriptive Übersicht über die Vorlaufseigenschaften verschiedener ifo-Indikatoren bezogen auf die Industrieproduktion. Wolters/Lankes (1989) untersuchen die Eignung des ifo-Geschäftsklimas als Konjunkturindikator mittels Spektralanalyse und Granger-Kausalitätstest. Das Geschäftsklima sowie dessen Teilkomponenten Geschäftslage und Geschäftserwartungen haben ein ähnliches zyklisches Verhalten wie die Referenzzeitreihe (= Jahreswachstumsrate des Nettoproduktionsindex) und weisen signifikante aber nicht vollständig stabile Vorläufe auf. Funke (1997) untersucht verschiedene vorlaufende Indikatoren für Rezessionen in Deutschland mit Hilfe von Probit-Modellen. Neben der Differenz zwischen lang- und kurzfristigen Zinsen erwies sich vor allem ein vorlaufender Indikator der OECD als gut geeignet für die Prognose von Rezessionen. Das ifo-Geschäftsklima wies zwar auch einen signifikanten Vorlauf auf, wurde jedoch von anderen Variablen deutlich übertroffen. Fritsche (1999) findet anhand von Granger-Kausalitätstests und Stabilitätstests, dass ifo-Geschäftsklima und ifo-Geschäftserwartungen einen signifikanten, aber nicht stabilen Vorlauf zur Industrieproduktion aufweisen. Langmantel (1999) zeigt auf, dass sich das ifo-Geschäftsklima für quantitative Prognosen des Bruttoinlandsprodukts eignet. In der Studie von Fritsche/Stephan (2000) sind die Prognoseeigenschaften der ifo-Indikatoren sowie die Auftragseingänge sogar deutlich besser als monetäre Indikatoren wie z.B. Zinsspreads. Allerdings finden die Autoren bei den Prognosetests Instabilitäten hinsichtlich der Prognosegüte. Die Studien zeigen insgesamt, dass die ifo-

² Der ifo-Geschäftsklimaindex ist ein geometrisches Mittel aus der Einschätzung der Geschäftslage und der Geschäftserwartungen. Daher sollte der Geschäftsklimaindex einen geringeren zeitlichen Vorlauf zur Konjunktur aufweisen als die Geschäftserwartungen alleine. Dies konnte auch durch von uns durchgeführte Granger-Kausalitätstests bestätigt werden. Wir verwenden daher in unserer Studie ausschließlich die ifo-Geschäftserwartungen und nicht das Geschäftsklima.

Indikatoren eine relativ gute Prognosefähigkeit bezüglich der zukünftigen Konjunkturlage aufweisen. Die ZEW-Konjunkturerwartungen wurden dagegen bislang noch nicht untersucht.

Unsere Studie unterscheidet sich in mehrfacher Hinsicht von den bisherigen Untersuchungen. Die zentrale Zielsetzung ist der Vergleich der Prognosefähigkeit von Stimmungsindikatoren, die die Meinung ganz unterschiedlicher Personengruppen abbilden. Die dahinter stehende Frage ist: Ist es besser, die Meinungen zur Entwicklung der einzelnen Unternehmen abzufragen und daraus einen aggregierten Indikator zu erstellen (= ifo-Ansatz) oder können volkswirtschaftliche Analysten, die die Gesamtwirtschaft und die Wirtschaftspolitik untersuchen, genauere Konjunkturprognosen erstellen (= ZEW-Ansatz)?

In Bezug auf die methodische Vorgehensweise ist unsere Studie zum Teil sehr ähnlich wie bisherige Untersuchungen. Wir verwenden zunächst Korrelationen und Granger-Kausalitätstests, um die Vorlaufeigenschaften im Schätzzeitraum zu analysieren. Als Referenzzeitreihe für die Konjunktur dient uns wie bei den meisten der bisherigen Studien ein Index der Industrieproduktion.³ Danach erstellen wir Prognosen auf Basis von vektorautoregressiven Modellen. Diese Vorgehensweise ist ähnlich derjenigen von Fritsche/Stephan (2000).

Bei der Bewertung der Prognosegüte unterscheiden sich allerdings die Verfahren. Zum Modellvergleich werden die von Diebold/Mariano (1995) und Harvey/Leybourne/Newbold (1997) entwickelten Tests auf Unterschiede im mittleren quadratischen Prognosefehler verwendet. Dadurch können wir Unterschiede der Prognosegüte auf Signifikanz testen. Außerdem verwenden wir einen von Harvey/Leybourne/Newbold (1998) verbesserten Encompassing-Test, um zu überprüfen, ob eine Kombination von Prognosen auf Basis der ifo- und ZEW-Indikatoren sinnvoll ist. Abschließend werden die Prognosen der Richtungsveränderungen der Konjunktur ausgewertet. Dadurch wird überprüft, ob die Prognosen in der Lage sind, Wendepunkte korrekt vorherzusagen. Für diese Analyse verwenden wir allerdings kein Probitmodell wie Funke (1997), sondern einen Chi-Quadrat-Unabhängigkeitstest und können damit sowohl die Prognosegüte für Aufschwung- als auch für Abschwungphasen untersuchen.

Um die von früheren Studien gefundene Instabilität von Schätzungen und Prognosegüte zu berücksichtigen, führen wir rollierende Regressionen durch, bei denen die Modelle in jedem Monat neu geschätzt werden. Außerdem wird die gesamte Pro-

³ Die Industrieproduktion als Referenzvariable ist aufgrund ihrer monatlichen Veröffentlichung für unsere Untersuchungen besser geeignet als das nur quartalsweise verfügbare Bruttoinlandsprodukt.

gnoseperiode in zwei Teilperioden aufgeteilt und die Prognosegüte jeweils getrennt ermittelt.

Die Studie ist folgendermaßen aufgebaut. Kapitel 2 beschreibt zunächst die beiden Erwartungsindikatoren und stellt die wichtigsten Unterschiede heraus. Kapitel 3 untersucht die Stationarität der verwendeten Zeitreihen und analysiert mit Hilfe von Korrelationen und Granger-Kausalitätstests den Zusammenhang zwischen den beiden Indikatoren untereinander sowie zwischen den Indikatoren und der Industrieproduktion. In Kapitel 4 werden Prognosemodelle unter Verwendung der Erwartungsindikatoren erstellt und die Prognoseergebnisse miteinander verglichen. Insbesondere wird getestet, ob Unterschiede im mittleren quadratischen Prognosefehler signifikant sind. Außerdem werden Encompassing-Tests angewandt und die Prognosen der Richtungsveränderungen der Konjunktur ausgewertet. Kapitel 5 schließt mit einer Interpretation der Ergebnisse ab.

2 Ifo-Konjunkturtest und ZEW-Finanzmarkttest im Vergleich

Den Ifo-Geschäftserwartungen und den ZEW-Konjunkturerwartungen liegen die jeweiligen Umfragen Ifo-Konjunkturtest und ZEW-Finanzmarkttest zugrunde. Diese sind sich in vielerlei Hinsicht ähnlich, jedoch entscheiden sie sich auch in einigen Punkten. Bevor daher der Verlauf der Erwartungsindizes verglichen werden soll, werden an dieser Stelle zunächst die Umfragemethoden und die Datenaufbereitung erklärt.

2.1 Die Umfragen

Sowohl beim Ifo-Konjunkturtest als auch beim ZEW-Finanzmarkttest handelt es sich um monatliche Umfragen, jedoch mit durchaus unterschiedlichem Teilnehmerkreis. Das Ifo-Institut befragt seit 1949 deutsche Manager und Unternehmer nach ihren Einschätzungen bezüglich der zukünftigen Geschäftslage in ihrem Unternehmen und ihrer Branche.⁴ Mittlerweile liegt die Teilnehmerzahl bei über 7.000 Unternehmen. Für die Ermittlung des Geschäftsklima-Index werden die Antworten nach der Bedeutung der Branchen gewichtet und aggregiert. Seit September 1990 werden auch Unternehmen in Ostdeutschland mit in das Panel einbezogen. Der Geschäftsklima-Index wird durch das geometrische Mittel der Antworten bezüglich der Geschäftslage und den Geschäftserwartungen gebildet, saisonbereinigt und auf das derzeitige Basisjahr 1991 bezogen. Die Ergebnisse für West- und Ostdeutschland werden getrennt ausgewiesen.

⁴ Vgl. Strigel, W.H. (1997), S. 1. Unterschieden werden die Bereiche Verarbeitendes Gewerbe, Bau, Groß- und Einzelhandel.

Im ZEW-Finanzmarkttest werden seit Dezember 1991 monatlich etwa 350 Analysten und institutionelle Anleger nach ihren Einschätzungen und Erwartungen von Finanzmarkt-Kennzahlen befragt.⁵ Der größte Anteil der Teilnehmer stammt dabei mit 77% aus Banken, gefolgt von Versicherungsunternehmen (14%) und großen Industriefirmen (9%). Über die Hälfte der befragten Experten ist im Bereich Vermögensberatung und Anlagemanagement tätig. Der restliche Anteil teilt sich auf volkswirtschaftliche Abteilungen/Wertpapier-Research, Finanzbuchhaltung und die Geschäftsführungsebene auf.

Im einzelnen geben die Teilnehmer ihre Erwartungen wieder, die sie für wichtige internationale Finanzmärkte hinsichtlich der Konjunktur, der Inflationsrate, der kurz- und langfristigen Zinsen, der Aktienkurse und der Wechselkurse haben. Bei den Finanzmärkten handelt es sich um die Märkte in Deutschland, den USA, Japan, Großbritannien, Frankreich und Italien. Seit 1999 werden auch Daten für den Euro-Raum mit einbezogen.

Sowohl ifo als auch ZEW befragen ihre Teilnehmer lediglich nach qualitativen Tendenzsicherungen bezüglich der jeweiligen Variablen. Auf die Erwartungen bezogen wird entsprechend nach der Veränderungsrichtung gefragt, also ob sich die zukünftige Wirtschaftslage verbessern, verschlechtern oder ob sie gleich bleiben wird. Beim ifo-Konjunkturtest werden die Antworten mit der Bedeutung der Branchen gewichtet.⁶ Auch im Prognosehorizont der Umfrage unterscheiden sich beide Tests nicht. So werden jeweils von den Teilnehmern ihre Erwartungen für die nächsten sechs Monate abgefragt.

Der Umfragezeitraum beider Tests ist nahezu identisch, was beide Indikatoren gut vergleichbar macht. Die Teilnehmer haben nach der monatlichen Versendung der Fragebögen jeweils etwa zwei Wochen Zeit, um ihre Einschätzungen abzugeben. Da die Veröffentlichung der Ergebnisse beider Tests häufig in derselben Woche stattfindet, überlappen sich die jeweiligen Befragungszeiträume zu einem großen Teil.⁷ Für die nachfolgenden empirischen Vergleiche der Frühindikatoreigenschaften kann daher angenommen werden, dass den Antworten der Umfrageteilnehmer beider Tests in etwa dieselben öffentlich zugänglichen Informationen zu Grunde liegen.

⁵ Für Informationen zum ZEW-Finanzmarkttest siehe auch <http://www.zew.de/de/publikationen/finanzmarktreport.php3>

⁶ Vgl. ifo (1999-2000).

⁷ Auch wenn die Umfragezeiträume nicht in jedem Monat zeitlich identisch sind, so gibt es dennoch keinen signifikanten permanenten Vorlauf eines der beiden Indikatoren.

2.2 Berechnungsmethoden

Die prozentualen Ergebnisse der verschiedenen Antwortkategorien werden in beiden Umfragen saldiert. In dem jeweiligen Saldenwert kommt letztlich die Differenz zwischen den Anteilen an positiven und negativen Antworten zum Ausdruck. Diese Werte werden vom ZEW monatlich veröffentlicht. Die Salden aus dem ifo-Konjunkturtest bilden die Grundlage für die Berechnung der ifo-Geschäftserwartungen und werden ebenfalls monatlich getrennt für West- und Ostdeutschland veröffentlicht. Die im ZEW-Finanzmarkttest berechneten Saldenwerte korrespondieren dabei zeitlich mit den jeweiligen Salden der ifo-Geschäftserwartungen. Für die Analysen in den folgenden Kapiteln werden jeweils die vom ifo für Westdeutschland veröffentlichten Erwartungen für die zukünftige Geschäftslage (= Geschäftserwartungen) berücksichtigt und mit den jeweiligen ZEW-Konjunkturerwartungen verglichen. Diese Vorgehensweise liegt zum einen darin begründet, dass im ZEW-Finanzmarkttest keine Unterscheidung zwischen Ost- und Westdeutschland getroffen wird. Zum anderen werden die ifo-Geschäftserwartungen nach West- und Ostdeutschland getrennt publiziert, und es ist für Externe nicht möglich, einen Gesamtindikator zu erstellen.

Als Referenzgröße zur Abbildung der konjunkturellen Entwicklung verwenden wir einen Index der Industrieproduktion für Gesamtdeutschland. Diese Zeitreihe⁸ erfasst die Mengenentwicklung im verarbeitenden Gewerbe. Sie wird von der Deutschen Bundesbank ermittelt und saisonbereinigt. Die Wahl einer monatlichen Produktionszeitreihe liegt nahe, da bei Verwendung des Bruttoinlandsproduktes zu wenig Beobachtungen für Schätzung und Prognosetest zur Verfügung stünden. Es könnte sein, dass eine Produktionszeitreihe für West-Deutschland besser zum Test der ifo-Geschäftserwartungen passt. Allerdings ist der Anteil der Produktion Ostdeutschlands an der Gesamtproduktion relativ niedrig, so dass dadurch keine größere Verzerrung zu Ungunsten des ifo-Indikators auftreten sollte.

Alle Berechnungen der Studie beziehen sich maximal auf den Zeitraum Dezember 1991 bis September 2000. Da die ZEW-Erwartungsdaten seit Dezember 1991 erhoben werden, war dies der zu Beginn der Studie größtmögliche gemeinsame Untersuchungszeitraum.

3 Untersuchung der Vorlaufeigenschaften

Um Konjunkturveränderungen frühzeitig erkennen zu können, werden Indikatoren gesucht, die einen Vorlauf vor der tatsächlichen Entwicklung haben und damit für die Prognose der zukünftigen Wirtschaftslage von Wert sind. Die ifo-Geschäftserwartungen können zurecht als ein derartiger Vorlaufindikator für Deutschland be-

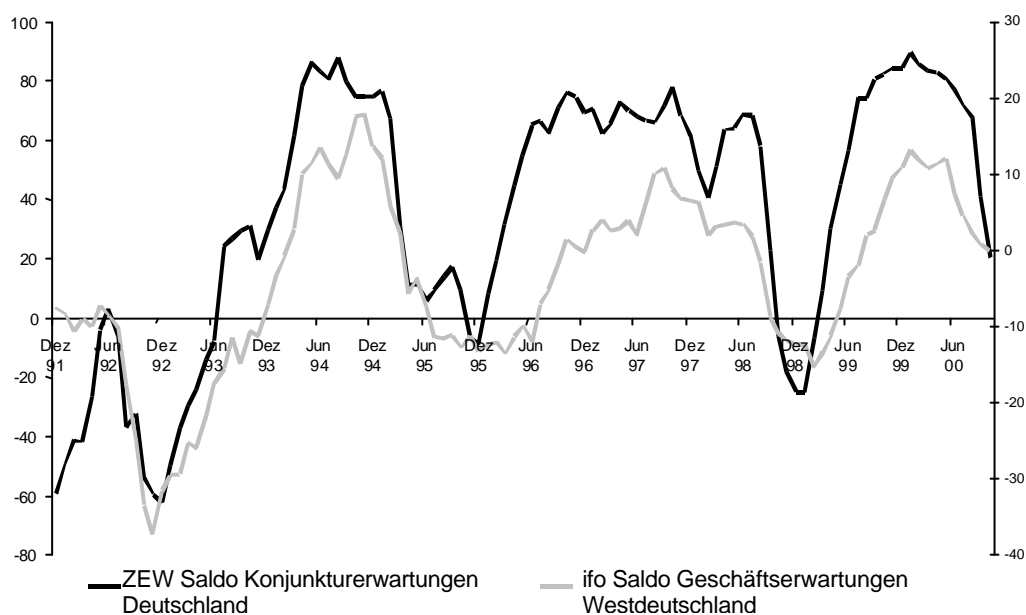
⁸ Quelle: Thomson Financial Datastream (Kürzel: BDIP7500G).

trachtet werden.⁹ Es stellt sich nun die Frage, ob der im ZEW-Finanzmarkttest ermittelte Konjunkturindikator ebenfalls ähnliche Vorlaufeigenschaften aufweist.

3.1 Grafische Darstellung

Im folgenden werden die beiden Erwartungsindikatoren einander gegenübergestellt. Abbildung 1 zeigt den Verlauf des ZEW-Indikators Konjunkturerwartungen im Vergleich zum Indikator ifo-Geschäftserwartungen. Die Abbildung zeigt, dass beide Zeitreihen – abgesehen von den ersten Monaten – optisch einen sehr ähnlichen Verlauf haben.

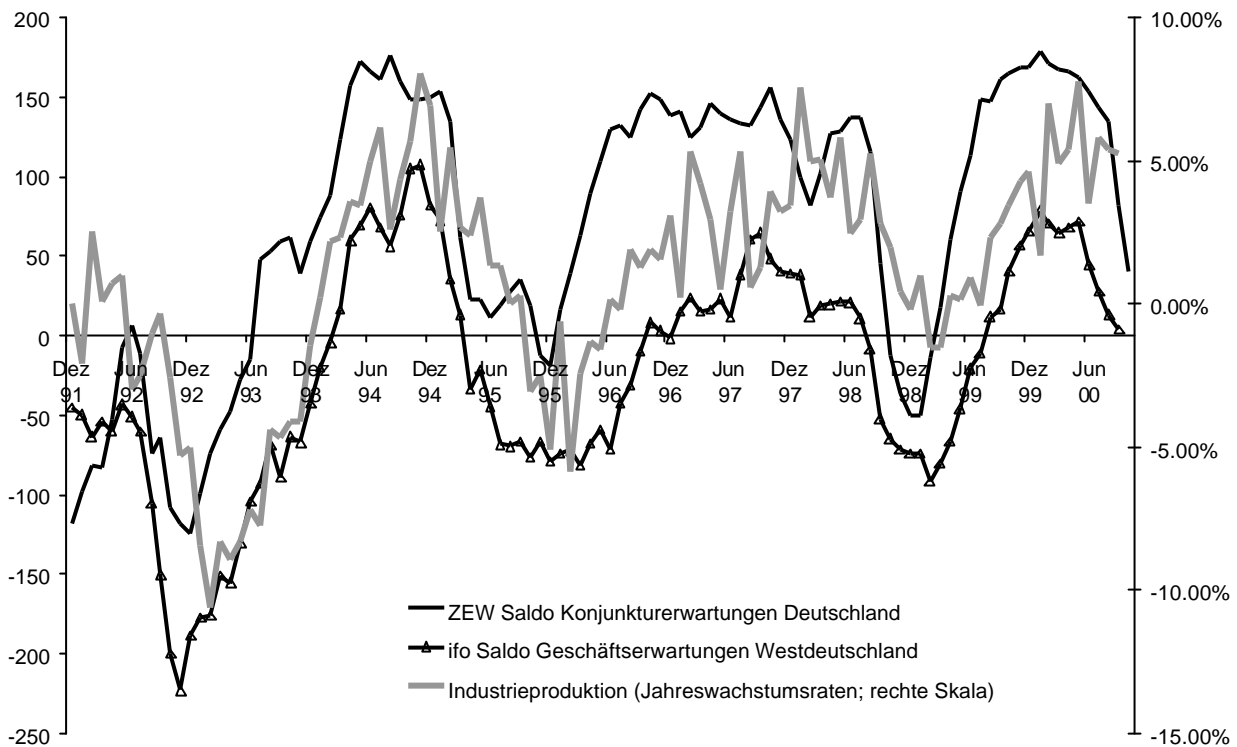
Abbildung 1: ZEW-Konjunkturerwartungen und ifo-Geschäftserwartungen



Weiterhin lässt sich aus der Grafik ein leichter Vorlauf des ZEW-Indikators vor dem ifo-Indikator erkennen. Betrachtet man die Wendepunkte beider Indikatoren, so lässt sich beispielsweise der letzte konjunkturelle Wendepunkt des ZEW-Indikators im Januar 2000 feststellen. Der ifo-Indikator dagegen hat seinen klaren Wendepunkt erst im Mai 2000, während der Index im Januar nur einen leichten und vorübergehenden Rückgang aufweist. Insgesamt gibt diese Grafik ein erstes Indiz dafür, dass der ZEW-Indikator einen gewissen Vorlauf vor dem ifo-Indikator aufweisen könnte.

⁹ Vgl. beispielsweise Wolters/Lankes (1989), Funke (1997), Fritsche (1999) und Fritsche/Stephan (2000).

Abbildung 2: ZEW und ifo-Indikator im Vergleich zu den Jahreswachstumsraten der Industrieproduktion



Anmerkung: Die Erwartungsreihen wurden jeweils umskaliert (bei ifo durch Versechsfachung und bei ZEW durch Verdopplung der Saldenwerte).

Aus Abbildung 2 erkennt man, dass beide Indikatoren einen Vorlauf vor der Industrieproduktion haben. Dennoch sind die Aussagen nur durch eine grafische Untersuchung nicht repräsentativ. Im folgenden sollen daher statistische Belege gefunden werden.

3.2 Überprüfung der Stationarität

Eine notwendige Voraussetzung für aussagekräftige Resultate bei Korrelationen und Regressionen ist die korrekte Berücksichtigung der Stationaritätseigenschaften der zu untersuchenden Zeitreihen. Daher haben wir alle relevanten Zeitreihen – ZEW- und ifo-Indikator sowie die Jahresveränderungsrate der Industrieproduktion – auf Stationarität untersucht. Für diese Analyse wurde der gesamte Zeitraum von Dezember 1991 bis September 2000 herangezogen.

Eine Anwendung des Augmented Dickey Fuller (ADF)-Tests ergibt eine klare Ablehnung der Nullhypothese „Nicht-Stationarität“ für die beiden Erwartungsindikatoren. Für die ifo-Geschäftserwartungen wird die Nullhypothese mit einer Irrtumswahrscheinlichkeit von 1% abgelehnt. Die dabei verwendete ADF-Gleichung enthält die Lags 1, 2 und 14 der ersten Differenzen des ifo-Indikators sowie den um eine Periode verzögerten ifo-Indikator selbst. Der t-Wert des verzögerten ifo-Indikators

beträgt $-2,73$. Die Lag-Strukturen in dieser Gleichung und in allen weiteren Gleichungen und Modellen dieser Studie wurden mit Hilfe des Bayesianischen Informationskriteriums (BIC)¹⁰ ermittelt. Bei den ZEW-Konjunkturerwartungen wird die Nullhypothese ebenfalls auf einem Signifikanzniveau von 1% verworfen. Die Gleichung enthält dabei eine Konstante, die Lags 1 und 13 der ersten Differenzen der ZEW-Erwartungen sowie die um eine Periode verzögerten Erwartungen. Der t-Wert der verzögerten Erwartungen beträgt $-3,49$.

Die Nicht-Stationarität der Jahresveränderungsrate der Industrieproduktion konnte für den Untersuchungszeitraum jedoch nicht verworfen werden. Daher wurde auch der KPSS-Test von Kwiatkowski/Phillips/Schmidt/Shin (1992) angewandt. Dieser Test hat die Nullhypothese „Stationarität“. Für die Jahresveränderungsrate der Industrieproduktion lagen die entsprechenden Teststatistiken weit entfernt von den kritischen Werten. Selbst bei einem Signifikanzniveau von 10% konnte die Nullhypothese der Stationarität nicht verworfen werden.¹¹ Für die weiteren Untersuchungsschritte dieser Studie wird daher davon ausgegangen, dass die Jahresveränderungsrate der Industrieproduktion trendstationär ist. Eine Anwendung des KPSS-Tests auf den Index der Industrieproduktion ergab eine klare Ablehnung der Stationaritätshypothese mit einem Signifikanzniveau von 1%.

Bei der Ermittlung von Zusammenhängen zwischen den in dieser Studie berücksichtigten Variablen und der Erstellung von Prognosemodellen werden alle Variablen – ZEW- und ifo-Erwartungen, Jahresveränderungsrate der Industrieproduktion – ohne weitere Differenzenbildung direkt verwendet.

3.3 Kreuzkorrelationen und Granger-Kausalitätstests

Ein erster Schritt zur Bestimmung der Beziehungen zwischen den drei Variablen besteht in der Berechnung von Kreuzkorrelationen. Zu beachten ist, dass dabei – im Gegensatz zu Granger-Kausalitätstests – die Autokorrelation der zu erklärenden Variablen nicht berücksichtigt wird. Daher können Kreuzkorrelationen nur Hinweise über die Zusammenhänge vermitteln, die jedoch noch mit anderen Methoden weiter

¹⁰ Eine andere Bezeichnung ist: Schwarz-Kriterium (SC). Im Unterschied zu dem auch häufiger verwendeten Akaike-Informationskriterium (AIC) ist das BIC konsistent in Bezug auf die korrekte Lagstruktur. Es führt typischerweise zu sparsameren Lag-Strukturen als das AIC.

¹¹ Bei der Anwendung des KPSS-Tests wurde die Basisgleichung des Tests mit Konstante und Trend geschätzt. Beide Koeffizienten sind hochsignifikant. Der KPSS-Test setzt die Schätzung einer um Heteroskedastie und Autokorrelation bereinigten Residualvarianz dieser Basisgleichung mit dem Verfahren von Newey/West (1987) voraus. Ab einem berücksichtigten Lag von 2 in der Schätzgleichung der Residualvarianz liegt die Teststatistik unterhalb des kritischen Wertes für ein Signifikanzniveau von 10%. Vgl. Kwiatkowski/Phillips/Schmidt/Shin (1992) für eine Darstellung des Testverfahrens.

untersucht werden müssen. Tabelle 1 zeigt die Kreuzkorrelationen der beiden Indikatoren ZEW-Konjunkturerwartungen und ifo-Geschäftserwartungen.

Tabelle 1: Kreuzkorrelationen zwischen den Erwartungsindikatoren

Lag	ZEWERW (t - Lag) → IFOERW (t)	IFOERW (t - Lag) → ZEWERW (t)
0	0,8598	0,8598
1	0,8765	0,8007
2	0,8584	0,7124
3	0,8229	0,5969
4	0,7720	0,4724
5	0,7116	0,3473

Generell auffallend ist die hohe Korrelation zwischen beiden Indikatoren. Die höchste Korrelation mit einem Wert von 0,8765 besteht zwischen dem ifo-Indikator und dem ZEW-Indikator verzögert um einen Monat. Somit legen die Korrelationsergebnisse einen Vorlaufcharakter des ZEW-Indikators nahe. Selbst eine Verzögerung des ZEW Indexes um 3 Monate liefert noch einen Korrelationskoeffizienten von 0,8229. Umgekehrt nehmen die Korrelationskoeffizienten bei Verzögerungen des ifo-Indikators deutlich ab. Ein Vorlauf der ifo Erwartungen vor den ZEW Erwartungen scheint somit nicht zu bestehen.

Tabelle 2: Kreuzkorrelationen zwischen den Erwartungsindikatoren und der Jahresveränderungsrate der Industrieproduktion

Lag	IFOERW (t - Lag) → INDPROD (t)	ZEWERW (t - Lag) → INDPROD (t)
0	0,8327	0,6696
1	0,8693	0,7400
2	0,8879	0,7744
3	0,8753	0,8163
4	0,8456	0,8098
5	0,7827	0,7905
6	0,6951	0,7613

Betrachtet man die Kreuzkorrelationen zwischen der Jahresveränderungsrate der Industrieproduktion und den beiden Frühindikatoren (Tabelle 2), so zeigt sich, dass der ifo-Indikator bei einer Verzögerung von zwei Monaten mit 0,8879 die höchste Korrelation mit der Industrieproduktion aufweist. Industrieproduktion und die ZEW Erwartungen sind bei einem Lag von drei Monaten am stärksten korreliert (0,8163). Dieses Ergebnis stimmt somit überein mit der vorherigen Feststellung, dass der ZEW-Indikator einen leichten Vorlauf von etwa einem Monat vor dem ifo-Indikator hat. Allerdings ist die Korrelation zwischen ifo-Indikator und Industrieproduktion bis zu einer Verzögerung von vier Monaten deutlich höher als diejenige des ZEW-

Indikators mit der Industrieproduktion, während danach der ZEW-Indikator eine höhere Korrelation aufweist.

Im nächsten Schritt sollen die Zusammenhänge zwischen ifo- und ZEW-Erwartungen sowie der Jahresveränderungsrate der Industrieproduktion mit Hilfe von Granger Kausalitätstests untersucht werden. Der Granger Kausalitätstest¹² gibt Aufschluss darüber, ob es möglich ist mit Hilfe einer Variablen X die Prognose der Variable Y zu verbessern. Der von uns verwendete Schätzansatz ist¹³

$$(1) \quad Y_t = a + \sum_i \mathbf{b}_i Y_{t-i} + \mathbf{d}_j X_{t-j} + \mathbf{e}_t, \quad \text{mit } j \geq 1$$

Wenn die Hinzunahme der Variablen X mit Lag j in die Regressionsgleichung zu einer signifikanten Parameterschätzung für δ_j führt, dann ist X_{t-j} granger-kausal zu Y. Dies kann mittels eines t-Tests überprüft werden.¹⁴ Die autoregressiven Lags von Y werden mit Hilfe des Bayesianischen Informationskriteriums (BIC) für alle drei untersuchten Variablen bestimmt. Tabelle 3 gibt einen Überblick über die Schätzergebnisse, wenn zunächst nur die autoregressiven Lags einbezogen werden:

Tabelle 3: Die Eigenschaften der Basisgleichungen $Y_t = a + \sum_i \mathbf{b}_i Y_{t-i} + \mathbf{e}_t$

	ifo-Geschäfts- erwartungen	ZEW-Konjunktur- erwartungen	Industrieproduktion (Verände- rungsrate zum Vorjahresmonat)
Lags i	1, 12	1, 2	1, 2, 12, 13
Koeffizienten \mathbf{b}_i (t-Werte in (..))	0.91, -0.11 (41.5), (-5.1)	1.47, -0.53 (17.0), (-6.3)	0.59, 0.37, -0.49, 0.37 (7.1), (4.5), (-5.9), (4.3)
DW-Werte	1.72	2.00	2.10
R ²	0.95	0.95	0.81

Anmerkung: Schätzzeitraum ist Dez. 1991 bis Sept. 2000

¹² Vgl. Granger (1969).

¹³ Im ursprünglichen Granger-Kausalitätstests werden für die Variable X die Lags von 1 bis j *en bloc* hinzugefügt und mit einem F-Test auf Signifikanz untersucht. Unsere Fragestellung ist jedoch, welche einzelnen Lags signifikant sind. Mit unserem Ansatz können wir auch den maximal signifikanten Lag ermitteln. Fritsche/Stephan (2000) schlagen vor, in Gleichung (1) nicht die t-Werte von X_{t-j} , sondern das BIC zu berechnen und mit dem BIC ohne X_{t-j} zu vergleichen. Die Anwendung dieses Verfahrens bestätigt im wesentlichen das von uns ermittelte Bild, allerdings sind die Ergebnisse relativ unscharf und schlechter zu interpretieren.

¹⁴ Den t-Werten liegt eine nach dem Newey-West Verfahren berechnete heteroskedastie- und autokorrelationskonsistente Varianz zu Grunde. Dabei wurde eine mögliche Autokorrelation der Residuen bis zum Lag 12 berücksichtigt. Die t-Werte ohne Newey-West Korrektur unterscheiden sich meistens nur gering von den in den Tabellen 3 bis 5 angegebenen Werten. Vgl. Newey/West (1987).

Bei den Granger-Kausalitätstests fügen wir zu diesen Basisgleichungen weitere Variablen mit verschiedenen Lags hinzu und überprüfen die Signifikanz der Koeffizienten dieser zusätzlichen Variablen. In Bezug auf die ifo-Geschäftserwartungen wird geprüft, ob die ZEW-Konjunkturerwartungen eine Prognosefähigkeit besitzen und entsprechend umgekehrt bezüglich der ZEW-Erwartungen. Dann wird die Prognosefähigkeit der beiden Erwartungsindikatoren bezüglich der Industrieproduktion untersucht. Tabelle 4 zeigt zunächst die Zusammenhänge zwischen ifo- und ZEW-Erwartungen für verschiedene Lags der jeweils anderen Erwartungsvariablen:

Tabelle 4: Granger-Kausalitätstest für den Zusammenhang zwischen ifo- und ZEW-Erwartungsdaten, Zeitraum: Dez. 1991 bis Sept. 2000

Parameterschätzung für Lag j, t-Werte in (..)		
Lags (j)	ifo (= X_{t-j}) \rightarrow ZEW (= Y_t)	ZEW (= X_{t-j}) \rightarrow ifo (= Y_t)
1	0.15 (1.18)	0.039 (2.41) ***
2	-0.08 (-0.89)	0.0066 (0.42)
3	-0.27 (-2.59)****	0.0044 (0.26)

Anmerkung: Signifikanzniveaus **** = 1%, *** = 2.5%, ** = 5%, * = 10%

Es zeigt sich, dass die ZEW-Konjunkturerwartungen signifikant um einen Monat den ifo-Geschäftserwartungen vorlaufen. Bei den ifo-Geschäftserwartungen ist Lag 3 klar signifikant, allerdings mit dem falschen Vorzeichen. Alle möglichen weiteren Lags (4, 5, ...) in den obigen Schätzgleichungen sind entweder insignifikant oder signifikant mit negativem Vorzeichen. Die Ergebnisse aus den Granger-Kausalitätstests bestätigen somit den ersten Eindruck, den die grafischen Darstellungen und die Kreuzkorrelationen vermittelt haben.

Tabelle 5: Granger-Kausalitätstest für den Zusammenhang zwischen Erwartungsdaten und der Jahresveränderungsrate der Industrieproduktion, Zeitraum: Dez. 1991 bis Sept. 2000

Parameterschätzung für Lag j, t-Werte in (..)		
Lags (j)	ifo (= X_{t-j}) \rightarrow Industrieprod. (= Y_t)	ZEW (= X_{t-j}) \rightarrow Industrieprod. (= Y_t)
1	0.0014 (11.8) ****	0.00021 (3.4) ****
2	0.0016 (8.2) ****	0.00022 (3.1) ****
3	0.0013 (3.2) ****	0.00025 (3.6) ****
4	0.0011 (2.8) ****	0.00020 (3.3) ****
5	0.0006 (1.3)	0.00019 (3.1) ****
6	0.00013 (0.3)	0.00014 (1.8) *
7	-0.00019 (-0.6)	0.00007 (0.89)

Anmerkung: Signifikanzniveaus: **** = 1%, *** = 2.5%, ** = 5%, * = 10%

Tabelle 5 zeigt Granger-Kausalitätstests zwischen den Erwartungsvariablen und der Jahresveränderungsrate der Industrieproduktion. Sowohl ifo-Geschäftserwartungen als auch ZEW-Konjunkturerwartungen sind im Schätzzeitraum bis zu einem Vorlauf von vier Monaten hoch signifikant. Die t-Werte für die ifo-Erwartungen sind dabei für Lags von 1 und 2 Monaten erheblich höher als diejenigen der ZEW-Erwartungen und danach niedriger. Dagegen weisen die ZEW-Erwartungen auch bei einem Vorlauf von fünf Monaten noch einen sehr signifikanten und bei sechs Monaten einen schwach signifikanten Koeffizienten auf. Beide Erwartungsindikatoren sollten auf Basis dieser Ergebnisse gut geeignet sein, um kurz- und mittelfristige Prognosen der Industrieproduktion zu verbessern. Dabei dürften die ifo-Erwartungen bei sehr kurzfristigen Prognosen zu besseren Ergebnissen führen, während Prognosen auf Basis der ZEW-Erwartungen bei mittelfristigen Prognosehorizonten überlegen sein sollten.

4 Bewertung der Prognosegüte

Die Resultate der Granger-Kausalitätstests sprechen dafür, dass die ifo- und ZEW-Erwartungsindikatoren einen sinnvollen Beitrag für die Prognose der Industrieproduktion liefern können. Die ZEW-Erwartungen könnten bei der Prognosegüte sogar insgesamt leicht besser sein. Granger-Kausalitätstests treffen jedoch nur eine Aussage über die Eigenschaften innerhalb des Schätzzeitraums. Die Bewertung der out-of-sample Prognosen kann dagegen durchaus zu anderen Ergebnissen führen, da sich dabei auch mögliche Instabilitäten der Modellstruktur auswirken können. Dieses Kapitel konzentriert sich auf die Erstellung von out-of-sample Prognosen der Industrieproduktion mit den beiden Erwartungsindikatoren. Die Bewertung der Prognosegüte erfolgt hauptsächlich über einen Vergleich der mittleren quadratischen Prognosefehler für verschiedene Prognosehorizonte. Dabei wird ein von Harvey/Leybourne/Newbold (1997) weiterentwickelter Test eingesetzt, der in seiner Grundform auf Diebold/Mariano (1995) zurückgeht. Außerdem werden Tests auf Unverzerrtheit der Prognose durchgeführt. Encompassing-Tests schließlich dienen dazu herauszufinden, ob möglicherweise eine Kombination der Erwartungsindikatoren zu besseren Ergebnissen führt als die Einzelprognosen. Der dabei angewandte Encompassing-Test wurde von Chong/Hendry (1986) entwickelt und von Harvey/Leybourne/Newbold (1998) verbessert.

Die bisher genannten Prognosetests untersuchen alle die Güte der Punktprognosen. Für die praktische Anwendung der Prognoseergebnisse mag es jedoch ausreichend sein, nur die Prognoserichtung auszuwerten. Daher untersuchen wir in einem abschließenden Schritt die Güte der Richtungsprognosen für verschiedene Prognosehorizonte mit Hilfe eines Chi-Quadrat-Unabhängigkeitstests.

Bei allen genannten Tests dient eine naive Prognose, die nur Lags der Industrieproduktion einbezieht, als Ausgangsbasis. Die Prognosen auf Basis der ifo-

Geschäftserwartungen und der ZEW-Konjunkturerwartungen werden sowohl mit der naiven Prognose als auch untereinander verglichen. Beim Test auf Unterschiede im mittleren quadratischen Prognosefehler und bei der Überprüfung der Richtungsprognosen werden alle Untersuchungen für drei Prognosezeiträume durchgeführt: 1. Gesamtzeitraum: Januar 1994 bis September 2000, 2. erste Hälfte des Gesamtzeitraums: Januar 1994 bis Mai 1997 und 3. zweite Hälfte des Gesamtzeitraums: Juni 1997 bis September 2000. Dies dient dazu, ein Gefühl für die zeitliche Stabilität der Prognosequalität zu bekommen. Der Startzeitpunkt für die Prognosen, Januar 1994, ist so gewählt, dass eine gerade noch ausreichende Datenhistorie von 2 Jahren für die erste Schätzung zur Verfügung steht.

4.1 Prognosemodelle

Die Prognosen werden mit vektorautoregressiven Modellen durchgeführt. Dieser Ansatz erlaubt die Berücksichtigung der dynamischen Struktur der Zeitreihen und die gleichzeitige Prognose sowohl der Industrieproduktion als auch der vorlaufenden Erwartungsvariablen. Die Prognosemodelle haben folgende allgemeine Struktur:

$$\begin{aligned} \Delta IP_t &= a_1 + \sum_i \mathbf{b}_{1i} \Delta IP_{t-i} + \sum_j \mathbf{d}_{1j} X_{t-j} \\ (2) \quad X_t &= a_2 + \sum_k \mathbf{b}_{2k} X_{t-k} + \sum_n \mathbf{d}_{2n} \Delta IP_{t-n} \end{aligned}$$

Die Jahresveränderungsrate der Industrieproduktion (**DIP**) wird durch die verzögerte Industrieproduktion und die verzögerte Erwartungsvariable (**X**) prognostiziert. Entsprechend wird die Erwartungsvariable durch die eigenen Verzögerungen und die verzögerte Industrieproduktion prognostiziert. Die Lagstruktur der eigenen Verzögerungen entspricht den Basismodellen, die in Tabelle 3 dargestellt sind. Die Lagstruktur der jeweils anderen Variablen wird für das gesamte Gleichungssystem mittels des Bayesianischen Informationskriteriums festgelegt. In allen betrachteten Modellen ergibt sich dabei ein optimaler Lag von 1. Ausserdem wird als Vergleichsmaßstab der Tests ein sogenanntes naives Prognosemodell verwendet, das ausschließlich die eigenen Verzögerungen der Industrieproduktion für die Prognosen heranzieht und der Basisgleichung aus Tabelle 3 entspricht.

Es ist hier nicht das Ziel, das bestmögliche Prognosemodell für die Industrieproduktion zu finden. Denn dazu wäre es notwendig, zusätzliche Variablen wie z.B. die Steigung der Zinsstrukturkurve oder Auftragseingänge auf ihre Brauchbarkeit für die Prognose hin zu untersuchen und gegebenenfalls in das Modell einzubeziehen. Die in dieser Studie verwendeten Prognosemodelle stellen vielmehr optimale Prognosemodelle auf Basis der beschriebenen Auswahlkriterien dar. Da das naive Modell der Gleichung der Industrieproduktion des Gesamtmodells entspricht, kann durch einen Vergleich der Prognosegüte mit dem naiven Modell herausgefunden werden, wie

groß der zusätzliche Nutzen für die Prognose ist, der durch die Hinzunahme der Erwartungsvariablen entsteht. Das Ziel ist zu untersuchen, ob ZEW- und ifo-Erwartungen eine signifikante Verbesserung zu einem optimalen naiven Modell darstellen und herauszufinden, welche der Erwartungsindikatoren zu besseren Prognoseresultaten führt.

Mit diesen so festgelegten Prognosemodellen werden rollierende out-of-sample Prognosen für die Prognosehorizonte 1, 3, 6, 9 und 12 Monate erstellt. Das bedeutet z.B. für den Prognosehorizont 12, dass im Monat t eine Schätzung für die gesamte vorliegende Datenhistorie und Prognosen für (t+1) bis(t+12) durchgeführt werden. Im darauf folgenden Monat ist die Datenhistorie für die Schätzung entsprechend um einen Datenpunkt länger. Da zum Zeitpunkt der Ermittlung der Erwartungsindikatoren in der Mitte von Monat t gerade erst der Wert der Industrieproduktion für den Monat (t-2) bekannt ist, bezieht sich somit der erste Prognoseschritt auf (t-1). Eine 12-Schrittprognose geht daher praktisch nur 10 Monate in die Zukunft. Die Prognose stellt allerdings auch für (t-1) und t eine echte Zusatzinformation dar, da die Industrieproduktion für diese Monate nicht offiziell vorliegt.

4.2 Bewertung der Prognosegüte für die Punktprognosen

Tabelle 6 gibt die Quadratwurzel der mittleren quadratischen Prognosefehler (RMSE = Root Mean Squared Error) für den Gesamtzeitraum an. Ergänzend wird auch der für die Prognosebewertung häufig verwendete Theil's U-Wert ausgewiesen, der sich aus dem Verhältnis zum RMSE der naiven Prognose errechnet: $U = RMSE(\text{Prognose Erwartungsmodell})/RMSE(\text{naive Prognose})$. Eine Prognose ist dann als gut zu bezeichnen, wenn ihr mittlerer quadratischer Prognosefehler kleiner ist als derjenige der naiven Prognose. In diesem Fall trägt die Verwendung der Erwartungsvariablen zu einer Verbesserung der naiven Prognose bei. Bei einer guten Prognose ist entsprechend der Theil's U-Wert kleiner als Eins.

Tabelle 6: Vergleich der RMSE- und Theil's U-Werte der einzelnen Prognosen für den Gesamtzeitraum: Januar 1994 bis September 2000

Prognosehorizont	Naive Prognose RMSE	ifo-Geschäftserwartungen		ZEW-Konjunkturerwartungen	
		RMSE	Theil's U	RMSE	Theil's U
1	0.01889	0.01736	0.92	0.01803	0.95
3	0.02511	0.02111	0.84	0.02004	0.80
6	0.04088	0.02933	0.72	0.02702	0.66
9	0.05540	0.04730	0.85	0.03303	0.60
12	0.06609	0.07405	1.12	0.03724	0.56

Für den gesamten Prognosezeitraum von Januar 1994 bis September 2000 ergeben sich mit beiden Erwartungsindikatoren recht gute Resultate. Bei der 1-Schrittprognose sind die ifo-Erwartungen etwas besser als die ZEW-Konjunkturerwartungen. Ab einem Prognosehorizont von drei Monaten weisen allerdings die ZEW-Erwartungen kleinere Werte für den RMSE auf. Die Prognosen mit den ZEW-Erwartungen führen zu Theil's U-Werten, die immer kleiner als 1 sind und damit das naive Modell in jedem Fall verbessern. Interessanterweise werden die RMSE-Werte mit wachsendem Prognosehorizont immer besser und erreichen bei 12 Monaten das beste Ergebnis relativ zur naiven Prognose. Bei Verwendung der ifo-Geschäftserwartungen sind die Prognosen bis zu einem Prognosehorizont von neun Monaten ebenfalls besser als die naive Prognose. Die 12-Monatsprognosen sind jedoch schlechter. Mit Ausnahme der 1-Monatsprognose ist der RMSE mit den ZEW-Erwartungen immer kleiner als bei Prognosen mit den ifo-Erwartungen.

Tabelle 7 gibt Auskunft darüber, ob die RMSE tatsächlich signifikant verschieden sind. Die in der Tabelle angegebenen Werte sind die Signifikanzniveaus des modifizierten Diebold/Mariano-Tests, mit dem Prognosefehler auf Unterschiede im RMSE untersucht werden können. Dieser asymptotische Test wurde von Diebold/Mariano (1995) entwickelt und ist für relativ allgemeine Verteilungseigenschaften der Prognosefehler anwendbar. Der Test wurde von Harvey/Leybourne/Newbold (1997) insbesondere für kleine Stichprobenumfänge verbessert. Eine ausführlichere Beschreibung des Tests findet sich in Teil 1 des Anhangs. Die in Tabelle 7 angegebenen Signifikanzniveaus beziehen sich auf diesen weiterentwickelten Test.

Tabelle 7: Signifikanzniveaus (in %) des modifizierten Diebold/Mariano-Tests auf Unterschiede im mittleren quadratischen Prognosefehler

	Ifo vs. Naive Prognose			ZEW vs. Naive Prognose			ZEW vs. ifo		
	Gesamt	1. H.	2. H.	Gesamt	1. H.	2. H.	Gesamt	1. H.	2. H.
1	22.4	83.6	98.3	36.1	97.0	98.3	44.6	86.5	99.9
3	14.2	63.1	96.2	5.8	49.0	96.6	44.4	83.3	99.6
6	7.7	11.7	90.5	4.8	6.3	90.0	45.3	73.5	99.5
9	21.9	28.3	86.8	6.9	0.8	83.8	21.1	6.4	97.0
12	64.4	18.1	85.2	9.5	0.2	79.6	29.0	1.1	94.2

Anmerkung: Signifikanzniveaus < 10% sind **fett** gedruckt.

Die Ergebnisse des Tests zeigen, dass die Prognosen unter Verwendung der ZEW-Erwartungen für die Gesamtperiode meistens signifikant besser sind als die naive Prognose. Dagegen ist die Prognose mit den ifo-Erwartungen bis auf eine Ausnahme – den Prognosehorizont von sechs Monaten - nicht signifikant von der naiven Prognose verschieden. Im direkten Vergleich ifo vs. ZEW ergeben sich allerdings für die Gesamtperiode keine statistisch signifikanten Unterschiede. Dies zeigt auch, dass nur relativ grosse Unterschiede im RMSE vom Test als signifikante Differenz α -

kannt werden: Obwohl die ZEW-Erwartungen für die Gesamtperiode fast immer zu kleineren RMSE-Werten führten als bei Verwendung der ifo-Erwartungen, wird dies vom Testverfahren nur in wenigen Fällen als systematische Verbesserung ausgewiesen.

Die folgenden Tabellen 8 und 9 beziehen sich jeweils auf die erste und die zweite Hälfte des Prognosezeitraums und geben wie Tabelle 6 die Werte für RMSE und Theil's U für verschiedene Prognosehorizonte an. Die Signifikanzniveaus für den Vergleich der RMSE sind in Tabelle 7 aufgeführt.

Die ZEW-Erwartungen führen sowohl in der ersten als auch in der zweiten Hälfte der Prognoseperiode fast immer zu besseren Prognosen als die ifo-Erwartungen. In der ersten Hälfte sind diese Unterschiede besonders ausgeprägt, dagegen nähert sich die Prognosegüte in der zweiten Hälfte deutlich an. Nur für die 9- und 12-Monatsprognosen in der ersten Hälfte weist die Teststatistik des modifizierten Diebold/Mariano-Tests diese Unterschiede als signifikant aus. In allen anderen Fällen, insbesondere für die zweite Hälfte, liegen keine statistisch signifikanten Unterschiede der Prognosegüte vor. Allerdings bleibt festzuhalten, dass die ZEW-Erwartungen insgesamt betrachtet meistens zu geringeren mittleren quadratischen Prognosefehlern führen.

Tabelle 8: Vergleich der RMSE- und Theil's U-Werte der einzelnen Prognosen für die erste Hälfte des Gesamtzeitraums: Januar 1994 bis Mai 1997

Prognosehorizont	Naive Prognose RMSE	ifo-Geschäftserwartungen		ZEW-Konjunkturerwartungen	
		RMSE	Theil's U	RMSE	Theil's U
1	0.01929	0.01771	0.92	0.01902	0.99
3	0.02918	0.02468	0.85	0.02247	0.77
6	0.05028	0.03446	0.69	0.03046	0.61
9	0.07192	0.06300	0.88	0.03904	0.54
12	0.08945	0.10753	1.20	0.04618	0.52

Tabelle 9: Vergleich der RMSE- und Theil's U-Werte der einzelnen Prognosen für die zweite Hälfte des Gesamtzeitraums: Juni 1997 bis September 2000

Prognosehorizont	Naive Prognose RMSE	ifo-Geschäftserwartungen		ZEW-Konjunkturerwartungen	
		RMSE	Theil's U	RMSE	Theil's U
1	0.01869	0.01720	0.92	0.01719	0.92
3	0.02065	0.01719	0.83	0.01761	0.85
6	0.03037	0.02418	0.80	0.02387	0.79
9	0.03694	0.02947	0.80	0.02755	0.75
12	0.04091	0.03292	0.80	0.02936	0.72

Die leitende Fragestellung dieser Studie ist, ob die unternehmensnahen Umfrageteilnehmer des ifo-Konjunkturtest oder die vom ZEW befragten Analysten genauere Erwartungen über zukünftige Konjunktorentwicklungen haben. Die bisherige Auswertung der Prognosefehler gibt klare Hinweise darauf, dass die Verwendung der ZEW-Erwartungen ab einem Prognosehorizont von drei Monaten zu besseren Prognosen führt. Dies lässt auch einen indirekten Rückschluss auf die Güte der Erwartungsbildung der jeweiligen Umfrageteilnehmer zu. Eine wichtige Teilfrage wird mit den Tests auf Unterschiede im RMSE jedoch nicht berührt, nämlich ob die Informationen, die den ifo-Erwartungen zugrunde liegen, vollständig in den ZEW-Erwartungen enthalten sind. Wenn die ifo-Erwartungsdaten Informationen enthalten, die in den ZEW-Erwartungsdaten nicht enthalten sind, dann könnte es nützlich sein, für die Konjunkturprognose beide Erwartungsindikatoren zu kombinieren.

Eine Antwort darauf kann mit Encompassing-Tests erreicht werden. Ein häufig gewählter Ansatz ist dabei derjenige von Chong/Hendry (1986). Ausgangspunkt ist für unseren Fall die folgende Gleichung: $\Delta IP_t = (1 - \lambda) \cdot \Delta \hat{IP}_{ZEW,t} + \lambda \cdot \Delta \hat{IP}_{ifo,t} + \mathbf{e}_t$. $\Delta \hat{IP}$ bezeichnet dabei die Prognose der Industrieproduktion. Mit dieser Gleichung kann überprüft werden, ob die ZEW-Prognose die ifo-Prognose dominiert oder ob die umgekehrte Beziehung vorliegt: Wenn λ den Wert Null annimmt, wäre die ZEW-Prognose die bessere und bei einem Wert von Eins entsprechend die ifo-Prognose. Für alle anderen Werte ist dagegen eine Kombination der Prognosen sinnvoll. Eine einfache Umformung dieser Gleichung ergibt: $f_{ZEW,t} = \lambda \cdot (f_{ZEW,t} - f_{ifo,t}) + \mathbf{e}_t$. Dabei sind $f_{ZEW,t}$ und $f_{ifo,t}$ die jeweiligen Prognosefehler auf Basis der ZEW- bzw. der ifo-Erwartungen. Statt einem direkten Test dieser Gleichung schlagen Harvey/Leybourne/Newbold (1998) eine ähnliche Vorgehensweise wie bei dem oben beschriebenen modifizierten Diebold/Mariano-Test vor. Der Test untersucht die Signifikanz des Koeffizienten λ für relativ allgemeine Verteilungseigenschaften der Prognosefehler. Eine ausführlichere Beschreibung des Tests und der Teststatistik findet sich in Teil 2 des Anhangs. Die Tests müssen in unserem Fall zweimal durchgeführt werden: einmal für die Null-Hypothese, dass die Prognosen mit den ZEW-Erwartungen alleine am besten sind und dann für die Null-Hypothese aus Sicht der ifo-Erwartungen. Tabelle 10 zeigt die Werte der Teststatistik und die Beurteilung der Signifikanz für verschiedene Prognosehorizonte.

In zahlreichen Fällen verwirft der Encompassing-Test die Null-Hypothese, dass eine Kombination nicht sinnvoll ist. Bei der 1-Schrittprognose ergibt sich, dass die Prognosen mit den ZEW-Erwartungen um die ifo-Prognosen ergänzt werden sollten, dagegen aber nicht umgekehrt. Da der RMSE für die 1-Schrittprognose bei Verwendung der ifo-Erwartungen am kleinsten ist, zeigt sich hier also eine Überlegenheit der ifo-Erwartungsdaten. Bei Prognosehorizonten von drei und sechs Monaten könnten durch eine Kombination sowohl die Prognoseergebnisse auf Basis der ZEW- als auch der ifo-Erwartungen verbessert werden. Bei diesen Horizonten füh-

ren die Prognosen mit den ZEW-Erwartungen zwar zu deutlich niedrigeren RMSE-Werten als die ifo-Erwartungen, die zusätzliche Nutzung der ifo-Erwartungen könnte die Prognosen jedoch noch genauer machen.

Tabelle 10: Ist eine Kombination der Einzelprognosen sinnvoll? Ergebnisse der Encompassing-Tests, Gesamtzeitraum: Jan. 1994 bis Sept. 2000

Prognosehorizont	H ₀ : ifo-Prognose alleine? Wert der Teststatistik	H ₀ : ZEW-Prognose alleine? Wert der Teststatistik
1	1.33	2.74****
3	3.27****	3.04****
6	2.12**	2.50***
9	1.41	1.56
12	1.08	-1.69*

Signifikanzniveaus: 1% = ****, 2.5% = ***, 5% = **, 10% = *

Bei einem Prognosehorizont von 12 Monaten tritt beim Test der ZEW-Erwartungen kurioserweise ein negatives Vorzeichen auf: In diesem Fall sollten also die Prognosen auf Basis der ifo-Erwartungen mit einem negativen Gewicht versehen werden. Beim Prognosehorizont von 9 Monaten tritt ein widersprüchliches Ergebnis auf, da angeblich keine der Erwartungsprognosen durch eine Kombination verbessert werden kann. Dies widerspricht auch den Ergebnissen des modifizierten Diebold/Mariano-Tests zum Vergleich der RMSE-Werte, bei dem die ZEW-Erwartungen deutlich besser abschneiden.

Was sind die praktischen Schlussfolgerungen aus den Encompassing-Tests? Für die Prognosehorizonte 3 und 6 scheinen in den Prognosen mit den ZEW- und ifo-Erwartungen Informationen enthalten zu sein, die die Prognosen auf Basis der jeweils anderen Erwartungsdaten sinnvoll ergänzen können. Dies gibt jedoch keinen klaren Hinweis auf die Art der Kombination, die zu der größtmöglichen Ergebnisverbesserung führt. Sinnvolle Kombinationen könnten beispielsweise darin bestehen, dass beide Erwartungsindikatoren in das Prognosemodell eingehen oder auch dass ein gemeinsamer Indikator aus den ursprünglichen Umfragedaten erstellt wird. Um einen ersten Eindruck von dem Verbesserungspotential einer kombinierten Prognose zu gewinnen, wird das Prognosemodell (2) um eine dritte Gleichung erweitert, so dass nun beide Erwartungsindikatoren in das Prognosemodell eingehen. Das Bayesianische Informationskriterium weist eine Laglänge von 1 als optimal aus. Mit diesem Prognosemodell werden wieder rollierende Regressionsschätzungen und die 1- bis 12-Monatsprognosen erstellt. Tabelle 11 fasst die Ergebnisse dieses erweiterten Prognosemodells zusammen.

Die Theil's U-Werte des gemeinsamen Modells sind, wie schon die jeweiligen Prognosen mit nur einem Erwartungsindikator, deutlich besser als die Prognosen mit

dem naiven Modell. In einigen Fällen sind die Theil's U-Werte sogar kleiner als diejenigen der besten Einzelprognose.

Tabelle 11: Theil's U-Werte der gemeinsamen Prognosen

Prognosehorizont	Gesamtzeitraum: Jan. 1994 – Sept. 2000	Erste Hälfte: Jan. 1994 – Mai 1997	Zweite Hälfte: Juni 1997 – Sept. 2000
1	0.94	0.95	0.92
3	0.81	0.80	0.83*
6	0.62*	0.56*	0.76*
9	0.62	0.58	0.74*
12	0.65	0.63	0.74

Anmerkung: * = bessere Prognose als die jeweils beste Einzelprognose mit nur einem Erwartungsindikator

Dies ist insbesondere bei der 6-Monatsprognose und teilweise der 3-Monatsprognose der Fall, bei denen auch der Encompassing-Test auf eine Kombination der Erwartungsdaten hingewiesen hat. Meistens jedoch ist die Prognosegüte etwas schlechter als die beste Einzelprognose. Insgesamt scheint die Verwendung beider Erwartungsindikatoren in einem Prognosemodell eine durchaus sinnvolle Alternative zu Prognosen zu sein, die nur auf den ZEW-Erwartungen aufbauen. Allerdings könnte sich eine andere als die hier gewählte Kombination von ifo- und ZEW-Erwartungen als noch besser erweisen.

4.3 Bewertung der Prognosegüte für die Richtungsprognosen

Für die Einschätzung zukünftiger Konjunkturveränderungen ist es häufig schon ausreichend, verlässliche Prognosen der Veränderungsrichtung zu haben. Zumindest aber ist die Richtungsprognose eine nützliche Zusatzinformation, die sich aus den Punktprognosen ermitteln lässt. Die Güte der Richtungsprognosen muss keineswegs direkt mit der Güte der Punktprognosen zusammenhängen, so dass der Modellvergleich im Extremfall auch zu völlig anderen Ergebnissen als den in Kapitel 4.2 dargestellten führen könnte. Ausgangspunkt der Bewertung der Richtungsprognosen ist die Definition der Richtungsprognose selbst. In dieser Studie verwenden wir die folgenden Definitionen:

Richtung der tatsächlichen Jahresveränderungsrate in n Monaten (ΔIP_{t+n}):

"steigen" falls $\Delta IP_{t+n} \geq \Delta IP_t$

"fallen" falls $\Delta IP_{t+n} < \Delta IP_t$

Prognostizierte Richtung auf Basis einer n-Schrittprognose ($\Delta \hat{IP}_{t+n}$):

"steigen" falls $\Delta \hat{IP}_{t+n} \geq \Delta IP_t$

"fallen" falls $\Delta \hat{IP}_{t+n} < \Delta IP_t$

Ein Treffer der Prognose liegt immer dann vor, wenn die Richtungsprognose mit der tatsächlichen Veränderung der Jahresveränderungsrate übereinstimmt. Zur Überprüfung der statistischen Signifikanz der Trefferquoten der drei alternativen Modelle wird ein Chi-Quadrat-Unabhängigkeitstest verwendet. Es handelt sich dabei um ein Standardverfahren, das testet, ob sich die Richtungsprognosen signifikant von einer Zufallsprognose unterscheiden.¹⁵ Die Trefferquoten der drei alternativen Prognosemodelle sind zusammen mit den Testergebnissen in Tabelle 12 dargestellt.

Tabelle 12: Bewertung der Richtungsprognosen: Trefferquoten für die drei Prognosezeiträume

	Naive Prognose			ifo-Prognose			ZEW-Prognose		
	Gesamt	1. H.	2. H.	Gesamt	1. H.	2. H.	Gesamt	1. H.	2. H.
1	62	61	63	80	76	85	77	63	90
3	73	72	74	77	80	74	81	87	74
6	66	66	66	76	76	76	78	76	79
9	68	69	67	74	72	76	74	75	73
12	63	68	57	77	80	74	76	83	69

Anmerkung: **fett** gedruckt bedeutet: signifikant mit 1% oder 5%, nicht fett gedruckt: signifikant mit 10%.

Es zeigt sich, dass alle drei Modelle Richtungsprognosen liefern, die für alle untersuchten Prognosehorizonte und alle drei Zeitperioden immer signifikant besser sind als reine Zufallsprognosen. In fast allen Fällen kann die Null-Hypothese einer Zufallsprognose sogar mit einem Signifikanzniveau von 1% abgelehnt werden. Die Verwendung der Erwartungsdaten führt dabei in jedem Fall zu einer erheblichen Verbesserung der Prognosegüte im Vergleich mit der naiven Prognose. Im Gegensatz zu den Punktprognosen führen hier die ifo-Erwartungen zu genauso guten Prognosen wie die ZEW-Erwartungen. Ein Unterschied zwischen der ersten und der zweiten Hälfte des Prognosezeitraums ist ebenfalls nicht zu erkennen. Interessant ist außerdem, dass die Richtungsprognosen selbst bei den 12-Monatsprognosen zu keinen schlechteren Resultaten führen als bei kurzen Prognosehorizonten.

Eine wichtige Voraussetzung für den praktischen Nutzen der Richtungsprognosen ist, dass sowohl steigende als auch fallende Wachstumsraten der Industrieproduktion gleichermaßen gut vorhergesagt werden. Tabelle 13 zeigt die Trefferquoten getrennt nach Auf- und Abschwungbewegungen. Hier zeigt sich der besondere Nutzen der Erwartungsdaten. Während das naive Prognosemodell Aufschwungbewegungen erheblich unterschätzt, führen ZEW- und ifo-Erwartungsdaten zu deutlich ausgegli-

¹⁵ Vgl. z.B. Diebold/Lopez (1996), S. 256-258 für eine Beschreibung des Tests und der Anwendungen. Eine kurze Beschreibung findet sich auch in Teil 3 des Anhangs unserer Studie.

cheneren Ergebnissen. Die Trefferquoten bei Aufwärtsbewegungen sind zwar immer noch geringer als bei zurückgehenden Wachstumsraten, aber mit 64% bis 80% geben sie relativ verlässliche Prognosen der Veränderungen an.

Tabelle 13: Bewertung der Richtungsprognose: Trefferquoten bei Auf- und Abschwung Gesamtzeitraum: Jan. 1994 bis Dez. 2000

	Naive Prognose		ifo-Prognose		ZEW-Prognose	
	TQ+	TQ-	TQ+	TQ-	TQ+	TQ-
1	47	79	72*	89*	67	86
3	50	97*	68	87	80*	82
6	43	97*	64	94	68*	91
9	44	97*	64	85	64	85
12	35	100*	68*	90	63	93

Anmerkungen: TQ+ : Trefferquote bei Zunahme der Wachstumsrate der Industrieproduktion und TQ- entsprechend bei Abnahme der Wachstumsrate; * bedeutet: höchste Trefferquote für TQ+ bzw. TQ- der jeweiligen Zeile.

Die Trefferquoten in Abschwungphasen sind bei der naiven Prognose meistens besser als bei Verwendung der Erwartungsdaten, allerdings wird die Verbesserung der Symmetrieeigenschaften der Prognosen mit nur einer vergleichsweise unbedeutenden Verschlechterung der Trefferquoten bei Konjunkturabschwächungen erkauft.

Als Fazit bleibt festzuhalten, dass die Verwendung der Erwartungsindikatoren zu Richtungsprognosen führt, die nicht nur absolut gesehen sehr gut sind, sondern auch relativ zum naiven Modell zu deutlichen Verbesserungen der Prognosegüte führen. Bei Richtungsprognosen sind, anders als bei den Punktprognosen, ifo- und ZEW-Erwartungen als gleichwertig einzustufen.

5 Zusammenfassung und Interpretation

Die Untersuchung der Zusammenhänge zwischen ifo-Geschäftserwartungen und ZEW-Konjunkturerwartungen hat gezeigt, dass beide gute Eigenschaften als Vorlaufindikator für die Industrieproduktion in Deutschland haben. Darüberhinaus hat sich der Eindruck bestätigt, dass der ZEW-Indikator einen etwa einmonatigen Vorlauf vor dem ifo-Indikator hat. Durch Korrelationen und Granger Kausalitätstests wurde dieser Vorlauf zunächst innerhalb des Beobachtungszeitraumes statistisch nachgewiesen. Im Vergleich mit der Industrieproduktion hat sich dabei gezeigt, dass die ZEW-Konjunkturerwartungen deren Veränderung signifikant bis zu 6 Monaten vorauslaufen, während die ifo-Erwartungen nur einen signifikanten Vorlauf von bis zu 4 Monaten vor der Industrieproduktion aufweisen.

Die anschließenden out-of-sample Prognosetests ergaben, dass mit beiden Frühindikatoren die zukünftige Industrieproduktionsentwicklung deutlich besser vorausgesagt werden kann als mit einer naiven Prognose, die lediglich die historischen Industrieproduktionswerte enthält. Insbesondere konnte gezeigt werden, dass eine Prognose mit den ZEW-Konjunkturerwartungen im Bereich zwischen 3 und 12 Monaten signifikant bessere Ergebnisse liefert als die naive Prognose. Dies ist bei Verwendung der ifo-Geschäftserwartungen nur hinsichtlich der 6-Monatsprognose der Fall. Ein direkter Vergleich beider Indikatoren führt zu dem Ergebnis, dass die Prognose mit dem ifo-Indikator auf Sicht von einem Monat leicht besser ist, während Prognosen mit den ZEW-Erwartungen für die übrigen Zeiträume überlegen sind.

Abschließend konnte mit Encompassing-Tests der Nachweis erbracht werden, dass bei 3- und 6-Schrittprognosen eine Kombination beider Indikatoren eine bessere Prognose bietet als die Verwendung jeweils beider Indikatoren alleine. Eine weitere Analyse beider Erwartungsindikatoren zeigte, dass diese für die Prognose der Richtungsänderung der Industrieproduktion als gleichwertig einzustufen sind.

Worin könnten nun die Gründe für dieses Ergebnis bestehen? Wie bereits erwähnt, sind die Befragungszeiträume beider Umfragen weitestgehend gleich, so dass der Vorlauf des ZEW-Indikators nicht durch einen zeitlichen Vorteil erklärt werden kann. Den Einschätzungen der Teilnehmer der beiden Umfragen sollten jeweils dieselben Informationen zu Grunde liegen. Dies lässt vermuten, dass die Analysten des ZEW-Tests diese Informationen in einer anderen Weise verarbeiten bzw. andere Informationen benutzen als die Teilnehmer des ifo-Konjunkturtests.

Der wohl entscheidende Unterschied zwischen beiden Indikatoren liegt in der Auswahl der Umfrageteilnehmer. Während das ifo-Institut Unternehmen aus verschiedenen Branchen befragt, sind die Teilnehmer des ZEW-Finanzmarkttests überwiegend Bankanalysten und institutionelle Anleger. Es ist anzunehmen, dass diese beiden Gruppen durchaus unterschiedliche Formen der Erwartungsbildung haben und verschiedene Faktoren in ihr Urteil über zukünftige Konjunkturänderungen einfließen. Denkbar ist etwa, dass die befragten Unternehmen im ifo-Konjunkturtest ihre Aussagen primär auf die Erwartungen bezüglich ihrer eigenen Branche stützen. Bankanalysten dagegen betrachten möglicherweise bevorzugt gesamtwirtschaftliche Zusammenhänge. Die verstärkte Beachtung von Faktoren wie etwa Zins- und Wechselkursentwicklung oder die Konjunktorentwicklung in anderen Ländern durch die Analysten könnte dafür verantwortlich sein, dass mit den ZEW-Erwartungen längerfristige Prognosen möglich sind als mit den ifo-Erwartungen.

Auf der anderen Seite weisen die Ergebnisse der ifo-Umfrage auf sehr kurze Sicht (ein Monat) einen stärkeren Zusammenhang mit der Industrieproduktion auf als die ZEW-Daten. Dies wiederum mag dadurch erklärbar sein, dass die Unternehmer die derzeitige Lage und die kürzerfristigen Entwicklungen besser einschätzen können.

Insgesamt lässt sich daher aus unseren Untersuchungen nicht der Schluss ziehen, dass ein Indikator besser als der andere für Konjunkturvorhersagen geeignet ist. Wie die Encompassing-Tests gezeigt haben, erscheint gerade die Kombination beider Daten für die Prognose der Industrieproduktion besser zu sein als die jeweilige Verwendung nur einer Datenreihe. Beide Frühindikatoren enthalten somit Informationen, die nicht in dem jeweils anderen enthalten sind.

Für die Finanzmärkte dürfte allerdings die Information, dass die ZEW-Konjunkturerwartungen einen signifikanten einmonatigen Vorlauf vor den ifo-Geschäftserwartungen für Westdeutschland haben, von Bedeutung sein. Ihnen steht mit den ZEW-Konjunkturerwartungen damit ein Instrument zur Verfügung, mit dem konjunkturelle Schwankungen noch frühzeitiger erkannt werden können.

Literaturverzeichnis

- Chong, Y.Y. und D.F. Hendry (1986), Econometric Evaluation of Linear Macroeconomic Models, *Review of Economic Studies*, vol. 53, S. 671-690.
- Diebold, F.X. und R. S. Mariano (1995): Comparing Predictive Accuracy, *Journal of Business and Economic Statistics* 13, 253-263.
- Diebold, F.X. und J.A. Lopez (1996), Forecast Evaluation and Combination, in G.S. Maddala und C.R. Rao (Hrsg.): *Handbook of Statistics* 14 (8), 241-268.
- Fritsche, U. (1999), Vorlaufeigenschaften von Ifo-Indikatoren für Westdeutschland, DIW Diskussionspapier Nr. 179.
- Fritsche, U. und S. Stephan (2000): Leading Indicators of German Business Cycles: An Assessment of Properties, DIW Diskussionspapier Nr. 207.
- Funke, N. (1997): Predicting Recessions: Some Evidence for Germany, *Weltwirtschaftliches Archiv*, vol. 133, S. 90-102.
- Granger, C.W.J. (1969): Investigating causal relations by econometric models and cross-spectral methods, *Econometrica*, 37, 424-438.
- Harvey, D. I., S.J. Leybourne und P. Newbold (1997): Testing the Equality of Prediction Mean Squared Errors, *International Journal of Forecasting*, vol. 13, S. 281-291.
- Harvey, D. I., S.J. Leybourne und P. Newbold (1998): Tests of Forecast Encompassing, *Journal of Business and Economic Statistics*, vol. 16, S. 254-259.
- ifo (1999-2000): Wie wird der ifo-Geschäftsklima-Index ermittelt ?; <http://www.ifo.de>.
- Kwiatkowski, D., P.C.B. Phillips, P. Schmidt, Y. Shin (1992): Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We that Economic Series Have a Unit Root?, *Journal of Econometrics*, vol. 54, S. 159-178.
- Langmantel, E. (1999): Das ifo-Geschäftsklima als Indikator für die Prognose des Bruttoinlandsproduktes, IFO Schnelldienst, Ausgabe 16/17, S. 16-21.
- Newey, W.K. und K.D. West (1987): A Simple Positive Semi-Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica*, vol. 55, S. 703-708.

- Strigel, W.H. (1985): Das Geschäftsklima als Konjunkturindikator – eine Retrospektive, ifo Studien, vol. 31, S. 29-68.
- Strigel, W.H. (1997): Seit 25 Jahren: Das ifo-Geschäftsklima als Konjunkturindikator; ifo-Schnelldienst 27/97.
- West, K. (2001): Tests for Forecast Encompassing when Forecasts Depend on Estimated Regression Parameters, Journal of Business and Economic Statistics, vol. 19, S. 29-33.
- Wolters, J. und F. Lankes (1989): Das ifo-Geschäftsklima als Konjunkturindikator, ifo Studien, vol. 35, S. 198-209.

Anhang

1. Beschreibung des modifizierten Diebold/Mariano-Tests zum Test auf signifikante Unterschiede im RMSE

Die Teststatistik (*mDM*) hat folgendes Aussehen:¹⁶

$$mDM = K \cdot \frac{\bar{d}}{\sqrt{\hat{V}(\bar{d})}} \sim \text{t-verteilt mit } (n-1) \text{ Freiheitsgraden}$$

mit $K = [(n+1 - 2h + n^{-1}h(h-1)) / n]^{1/2}$, $\bar{d} = \frac{1}{n} \sum_{t=1}^n d_t$, $d_t = f_{1t}^2 - f_{2t}^2 =$ Differenz

der quadrierten Prognosefehler, $f_{it} = \Delta IP_t - \Delta \hat{IP}_{it} =$ Prognosefehler von Modell *i*,

$n =$ Anzahl der Prognosen, $h =$ Prognosehorizont in Monaten, $\hat{V}(\cdot) =$ heteroskedastie- und autokorrelationskonsistente Schätzung der Varianz.

Der ursprüngliche Test von Diebold/Mariano (1995) entspricht *mDM* ohne den Korrekturfaktor *K*. Diebold/Mariano zeigen, dass diese Testgröße asymptotisch standardnormalverteilt ist. Harvey/Leybourne/Newbold (1997) verbessern die Eigenschaften des Tests für kleine Stichprobenumfänge und empfehlen, die t-Verteilung für die kritischen Werte heranzuziehen. Die heteroskedastie- und autokorrelationskonsistente Schätzung der Varianz ermöglicht eine relativ allgemeine Anwendung des Tests bezüglich der Verteilung der Prognosefehler. Insbesondere ist davon auszugehen, dass bei *h*-Schritt Prognosen eine Autokorrelation der Ordnung (*h*-1) vorliegt. Üblicherweise wird die Schätzung von *V* unter Verwendung des Newey-West Verfahrens durchgeführt.¹⁷

¹⁶ Vgl. Harvey/Leybourne/Newbold (1997).

¹⁷ Vgl. Newey/West (1987).

2. Beschreibung des modifizierten Diebold/Mariano-Tests zum Test auf Encompassing

Die formale der Struktur der Teststatistik (*enDM*) ist mit dem in Teil 1 des Anhangs dargestellten Test identisch.¹⁸

$$enDM = K \cdot \frac{\bar{d}}{\sqrt{\hat{V}(\bar{d})}} \sim \text{t-verteilt mit } (n-1) \text{ Freiheitsgraden}$$

Der wesentliche Unterschied zum modifizierten Diebold/Mariano Test von Teil 1 des Anhangs liegt in der Definition der Variablen d_t : $d_t = (f_{ZEW,t} - f_{ifo,t}) \cdot f_{ZEW,t}$. Dabei sind $f_{ZEW,t}$ und $f_{ifo,t}$ die Prognosefehler bei Verwendung der ZEW- bzw. der ifo-Erwartungsindikatoren. Eine Definition aller anderen Variablen findet sich in Teil 1 dieses Anhangs.

Ein Test auf Signifikanz von d_t hat eine ähnliche Aussage wie der Test auf Signifikanz von λ in der Regressionsgleichung $f_{ZEW,t} = \mathbf{I} \cdot (f_{ZEW,t} - f_{ifo,t}) + \mathbf{e}_t$. Allerdings kann der hier verwendete, von Harvey/Leybourne/Newbold (1998) vorgeschlagene Test auch bei Vorliegen von Heteroskedastie und Autokorrelation in den Prognosefehlern angewandt werden.

Die Tests müssen in unserem Fall zweimal durchgeführt werden: einmal für die Null-Hypothese, dass die Prognosen mit den ZEW-Erwartungen alleine am besten sind (= oben dargestellter Fall) und dann für die Null-Hypothese aus Sicht der ifo-Erwartungen mit $d_t = (f_{ifo,t} - f_{ZEW,t}) \cdot f_{ifo,t}$.

West (2001) verweist darauf, dass bei den üblichen Encompassing-Tests, und darunter fällt auch der vorliegende Test, die Varianz der Schätzfehler der Prognosemodelle nicht berücksichtigt wird. Dadurch ist die geschätzte Varianz V generell zu klein, so dass die Null-Hypothese zu häufig abgelehnt wird. Allerdings scheint nach den Simulationen von West (2001) in Situationen, die unserer Anwendung vergleichbar sind, diese Verzerrung für Prognosehorizonte bis etwa 6 Monate vernachlässigbar klein zu sein und könnte erst für grössere Prognosehorizonte eine gewisse Rolle spielen.

¹⁸ Vgl. Harvey/Leybourne/Newbold (1998).

3. Bewertung der Richtungsprognosen mit dem Chi-Quadrat-Unabhängigkeitstest

Basis des Tests ist eine 2x2-Kontingenztabelle, die in der Hauptdiagonalen (N_{11} , N_{22}) die Anzahl der richtigen Prognosen für „steigt“ und „fällt“ enthält und in der Nebendiagonalen (N_{12} , N_{21}) die Anzahl der Wendepunktfehler ausweist. Der Test überprüft, ob sich die Belegungshäufigkeit der 4 Felder (N_{11} , N_{22} , N_{12} und N_{21}) der Kontingenztabelle von einer Zufallsprognose unterscheidet, für die nur die Randsummen ($N_{1.}$, $N_{2.}$, $N_{.1}$ und $N_{.2}$) verwendet werden. Tabelle 14 verdeutlicht den Aufbau der Kontingenztabelle:

Tabelle 14: Aufbau der Kontingenztabelle beim Chi-Quadrat-Unabhängigkeitstest

	Tatsächliche Veränderung „steigt“	Tatsächliche Veränderung „fällt“	Randsummen
Prognose „steigt“	N_{11}	N_{12}	$N_{1.}$
Prognose „fällt“	N_{21}	N_{22}	$N_{2.}$
Randsummen	$N_{.1}$	$N_{.2}$	N (Gesamtsumme)

Die Teststatistik ist:

$$C = \sum_{i,j=1}^2 \frac{(N_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}} \quad \text{mit } \hat{E}_{ij} = N_{i.} \cdot N_{.j} \cdot N$$

C ist asymptotisch Chi-Quadrat-verteilt mit einem Freiheitsgrad. Eine ausführlichere Beschreibung des Tests findet sich beispielsweise in Diebold/Lopez (1996: 256-258).