

Einige Prognoseeigenschaften des ifo Geschäftsklimas – Ein Überblick über die neuere wissenschaftliche Literatur

Das ifo Geschäftsklima ist ein viel beachteter Indikator für die konjunkturelle Entwicklung in Deutschland. Es ist daher auch immer wieder Gegenstand von wissenschaftlichen Analysen, in denen verschiedene Eigenschaften des Geschäftsklimas untersucht werden. Im Zentrum des Interesses steht dabei häufig die Verwendung des Indikators zu Prognosezwecken. So widmet etwa der Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung in seinem Jahresgutachten 2006/2007 dem ifo Geschäftsklima einen Abschnitt. Der Beitrag gibt einen Überblick über die wissenschaftliche Literatur, die sich mit dem ifo Geschäftsklima auseinandersetzt und seine Prognoseeigenschaften beurteilt.

Das ifo Institut führt seit seiner Gründung im Jahre 1949 Unternehmensbefragungen durch. Anfang der siebziger Jahre wurde auf Basis des ifo Konjunkturtests das ifo Geschäftsklima als Konjunkturindikator entwickelt. Seither wird das Geschäftsklima für die gewerbliche Wirtschaft Deutschlands vom ifo Institut monatlich veröffentlicht. In seiner langen Geschichte war es regelmäßig Gegenstand von wissenschaftlichen Untersuchungen. Auch das 1960 unter Mitwirkung von ifo gegründete Centre for International Research on Economic Tendency Surveys (CIRET) hat Wissenschaftlern durch die Organisation von Konferenzen und der Herausgabe von Zeitschriften ein Forum geboten, in dem die Methoden und Ergebnisse von Konjunkturumfragen, wie der des ifo Instituts, regelmäßig diskutiert wurden. Der vorliegende Beitrag stellt einen kurzen Literaturüberblick über die wissenschaftlichen Aufsätze der letzten Jahre dar, die sich mit der Prognosequalität des ifo Geschäftsklimas beschäftigt haben. Diskutiert werden nur die Studien, die in den letzten Jahren entstanden sind.

Mitarbeiter des ifo Instituts führten und führen ebenfalls regelmäßig Analysen zu diesem Thema durch. Die Ergebnisse dieser Studien wurden in zahlreichen Büchern, Schnelldienstartikeln und wissenschaftlichen Aufsätzen publiziert. Diese Veröffentlichungen sollen aber nicht im Fokus des vorliegenden Artikels stehen. Namentlich erwähnt sei hier nur eine frühe Arbeit von Strigel (1985), da er als »Vater des Geschäftsklimas« bezeichnet werden kann, sowie ein neues englischsprachiges Handbuch über die ifo Konjunkturumfragen von Goldrian (2007).

Konjunkturindikatoren, wie das ifo Geschäftsklima, finden auch in der Öffentlichkeit große Aufmerksamkeit und sind Entscheidungs- und Orientierungshilfe in vielen Bereichen. Dies stellt an jeden Indikator hohe Qualitätsmerkmale. Von einem geeigneten Frühindikator kann man verlangen, dass er (vgl. auch Abberger 2006, Fritsche und Stephan 2002 oder Hinze 2003):

1. möglichst zeitnah vorliegt,
2. nach der Publikation keinen größeren Revisionen unterliegt,
3. einen Vorlauf vor der zur beurteilenden Zielgröße besitzt; d.h. er sollte frühzeitig Signale über den kommenden Konjunkturverlauf geben,
4. eine stabile Vorlaufbeziehung besitzt, so dass relativ sicher abgeschätzt werden kann, wie frühzeitig das Signal des Indikators erfolgt und
5. ein möglichst deutliches Konjunktur-signal enthält; d.h. nicht interessierende kurzfristige Schwankungen in der Zeitreihe sollten das eigentlich interessierende konjunkturelle Signal möglichst wenig überdecken.

Die ersten beiden Eigenschaften aus diesem Katalog werden vom ifo Geschäftsklima und den weiteren Ergebnissen der ifo Konjunkturumfragen offensichtlich erfüllt. Die Resultate der ifo Konjunkturumfragen werden in der letzten Woche eines betrachteten Monats publiziert und in der Regel im Nachhinein nicht mehr revidiert.

Die letzten drei Eigenschaften sowie unterschiedliche Anwendungsmöglichkeiten sind in verschiedenen wissenschaftlichen Artikeln untersucht worden. Diese gehen

insbesondere der Frage nach, ob ein Indikator überhaupt Gleich- bzw. Vorlauf bzgl. der Konjunktur (gemessen an einer Referenzreihe) besitzt oder ob er Wendepunkte zuverlässig vorhersagen kann. Der am häufigsten untersuchte Gegenstand ist, ob mit Hilfe eines Konjunkturindikators eine quantitative Prognose (meist einer Wachstumsrate) durchgeführt werden kann.

Der Sachverständigenrat hat in seinem Gutachten 2006/2007 in einer speziellen Untersuchung festgestellt, dass das ifo Geschäftsklima keine signifikante quantitative Prognosekraft bzgl. der Wachstumsrate des Bruttoinlandsprodukts besitzt. Dies bedeutet, dass mit Hilfe des ifo-Indikators ein autoregressives Modell als Messlatte nicht signifikant verbessert werden kann. Ein Blick in die wissenschaftliche Literatur zeigt jedoch, dass die verwendeten Methoden und Referenzreihen sich durchaus unterscheiden. Je nach Untersuchungsanordnung fällt die Beurteilung eines Indikators unterschiedlich aus.

Dieser Artikel stellt einen kurzen Überblick über die wissenschaftliche Literatur dar, in der das ifo Geschäftsklima oder seine beiden Komponenten, Lagebeurteilung und Geschäftserwartungen, Gegenstand der Forschung sind. In den Artikeln werden meist auch noch andere Indikatoren auf ihre Eigenschaften untersucht, wir konzentrieren uns jedoch ausschließlich auf die ifo-Indikatoren und ihre Eigenschaften. Ein Vergleich wird jeweils nur zum Referenzmodell gezogen. Für weitere Informationen sei auf die angegebene Literatur verwiesen.

Da Konjunkturindikatoren generell nicht nur für verschiedene Zwecke eingesetzt werden, wie etwa der Wendepunktprognose, der Richtungsprognose oder der quantitativen Prognose, sondern sich auch innerhalb der verschiedenen Zielsetzungen die Vorgehensweisen und die Untersuchungsdesigns erheblich unterscheiden können, versuchen wir im Folgenden eine grobe Kategorisierung der Ansätze. Abgesehen von den grundsätzlichen Prognosezielen, unterscheiden sich die zur Prognose ausgewählten Referenzzeitreihen und die zur Prognose eingesetzten Methoden.

Die Referenzreihe

Zunächst dokumentieren wir die Referenzreihe in den jeweiligen Artikeln, die den Konjunkturverlauf darstellen soll. Wie sich noch zeigen wird, sind die Ergebnisse durchaus sensitiv bzgl. der gewählten Referenzreihe und des Schätzzeitraumes. Die Ergebnisse in der Tabelle zeigen, dass als Referenzreihe meist der Index der Industrieproduktion verwendet wird und hier insbesondere die Wachstumsrate. Der Grund dafür ist, dass der Index monatlich erhoben wird und nahezu alle Indikatoren ebenfalls monatlich veröffentlicht werden. Darüber hinaus wird das verarbeitende Gewer-

be häufig als der Konjunkturzyklen generierende Bereich angesehen. Wird das Bruttoinlandsprodukt als Referenz verwendet, müssen die Indikatoren auf Quartalsebene aggregiert werden. Dies geschieht meistens mit einem Drei-Monatsdurchschnitt. Erst neuere Zeitreihenmodelle wie in Mittnik und Zadrozny (2005) erlauben es, gleichzeitig unterschiedliche Frequenzen zu verwenden.

Es ist wichtig, darauf hinzuweisen, dass sowohl die Veränderungsrate zum Vormonat (Vorquartal) als auch die Jahreswachstumsrate als Referenz dienen können. Warum eine bestimmte Wachstumsrate gewählt wurde, wird allerdings häufig nicht inhaltlich begründet. Neben ökonometrischen Überlegungen (Aspekt der Einheitswurzel, Phasenverschiebung) können auch Fragen der inhaltlichen Interpretation die Auswahl einer bestimmten Wachstumsrate beeinflussen (vgl. hierzu Nierhaus 1999). Neben der Bildung von Wachstumsraten können aus Zeitreihen, aber auch mit anderen Methoden konjunkturelle Komponenten, die dann prognostiziert werden sollen, extrahiert werden. Eine Vorgehensweise ist etwa, die Abweichung von einem zu schätzenden Trend zu verwenden. Wieder andere Ansätze versuchen, eine möglichst glatte Konjunkturkomponente zu ermitteln, um eine einfachere Analyse des Zustandes der Konjunktur zu erreichen. Ein Beispiel für diese Herangehensweise ist der Indikator EuroCOIN des Centre for Economic Policy Research (vgl. <http://www.cepr.org/data/eurocoin/>).

Alle in der Tabelle aufgeführten Studien verwenden Zeitreihen, die »endgültig« sind, d.h. sie unterliegen keiner Revision mehr. Die Industrieproduktion und das Bruttoinlandsprodukt werden zunächst mit einer Schätzung für den entsprechenden Monat und dem Quartal angegeben und unterliegen danach mehreren Revisionen. Ein Datensatz, der aus einer ersten Schätzung und den darauf folgenden Revisionen besteht, wird als Real-Time-Datensatz bezeichnet (vgl. Döpke 2004). Ein guter Indikator sollte in der Lage sein, diese Revisionen vorherzusagen, um zeitnah eine Einschätzung über die Entwicklung makroökonomischer Größen zu liefern. In dem Artikel von Jacobs und Sturm (2005) wird ein solcher Real-Time-Datensatz für die Industrieproduktion verwendet.

Die Methoden

Neben der verwendeten Referenzreihe spielen die eingesetzten ökonometrischen Methoden eine Rolle in der Beurteilung der Eigenschaften von Indikatoren. Im Folgenden werden wir diese kurz darstellen. Für weitere Informationen siehe Hamilton (1994), Nierhaus und Sturm (2005) oder Lütkepohl (2005).

Ob ein Indikator einen grundsätzlichen Erklärungsgehalt bzgl. einer Referenzreihe besitzt, wird meist mit dem Kon-

Konjunkturelle Referenzreihen und Untersuchungszeiträume

Artikel	Referenzreihe	Zeitraum
Döpke, Krämer und Langfeldt (1994)	Industrieproduktion, saisonal, arbeitstäglich und um Trend bereinigt.	1966:01–1992:12
Funke (1997)	Bruttoinlandsprodukt, Quartalswachstumsraten.	1971:IV–1995:IV
Döpke (1999)	Industrieproduktion, gleitender Durchschnitt von sieben Monaten, Ausreißer wurden ersetzt.	1962:01–1998:12
Langmantel (1999)	Reales Bruttoinlandsprodukt, jährliche Wachstumsraten.	1965:I–1998:IV
Breitung und Jagodzinski (2001)	Industrieproduktion, saisonal und arbeitstäglich bereinigt, jährliche Wachstumsraten (log Differenzen).	1991:01–2001:06
Fritsche und Stephan (2002)	Industrieproduktion (ohne Bau), jährliche Wachstumsraten (log Differenzen).	1978:01–1998:12
Hüfner und Schröder (2002)	Industrieproduktion, saisonbereinigt, jährliche Wachstumsraten.	1991:12–2000:12
Hinze (2003)	Industrieproduktion, Bruttoinlandsprodukt, Vorjahresveränderungsraten.	1992–2002
Bandholz und Funke (2003)	Reales Bruttoinlandsprodukt, trendbereinigt (Band-Pass-Filter), Output-Gap.	1971:I–2002:II
Benner und Maier (2004)	Industrieproduktion, jährliche Wachstumsraten.	1991:01 2000:12 1968:01–2000:10
Schuhmacher und Dreger (2004)	Bruttoinlandsprodukt, jährliche Wachstumsraten (log Differenzen).	1978:II–2002:I
Fritsche und Kouzine (2005)	Industrieproduktion, Originalreihe und gleitender Durchschnitt von sieben Monaten, Ausreißer wurden ersetzt	1978:01–1999:12
Dreger und Schuhmacher (2005)	Industrieproduktion, saisonal nicht bereinigt, jährliche Wachstumsraten (log Differenzen).	1992:01–2004:12
Sachverständigenrat (2005)	Industrieproduktion, Monats- und Jahreswachstumsraten (log Differenzen).	1994:01–2005:02
Mitnik und Zadrozny (2005)	Bruttoinlandsprodukt, erste Differenzen.	1970:I–2003:IV
Jacobs und Sturm (2005)	Industrieproduktion, Real-Time Datensatz.	1995:12–2003:08
Nierhaus und Sturm (2005)	Industrieproduktion, saisonal nicht bereinigt, Jahreswachstumsraten (log Differenzen).	1991:01–2002:12
Kholodilin und Siliverstovs (2006)	Bruttoinlandsprodukt, annualisierte Jahreswachstumsraten.	1991:II–2004:IV
Sachverständigenrat (2006)	Bruttoinlandsprodukt, saison- und kalenderbereinigt, Jahreswachstumsraten.	1970:I–2005:IV

Quelle: Zusammenstellung des ifo Instituts.

zept der Granger-Kausalität analysiert. Bei diesem Verfahren wird untersucht, ob der Indikator einen signifikanten Erklärungsgehalt bzgl. der Referenzreihe besitzt. Es ist jedoch wichtig, darauf hinzuweisen, dass Granger-Kausalität nicht Kausalität im philosophischen oder logischen Sinne meint.

Zur Vorhersage von Rezessionen und Aufschwüngen werden meist Probit-Schätzungen und Markov-Switching-Modelle verwendet. Zunächst werden die Wendepunkte einer Zeitreihe, und somit auch ihre Rezessionsphasen, bestimmt. Danach wird eine neue Zeitreihe konstruiert. Befindet sich die Volkswirtschaft in einer Rezession, so wird der Zeitreihe zu diesem Zeitpunkt der Wert 1 zugeordnet, ansonsten 0. Probit-Modelle sind in der Lage, solche so genannten binären Modelle zu schätzen. In diesem Rahmen wird geprüft, inwieweit der Konjunkturindikator die Rezessionen vorhersagen kann. Im Rahmen der Markov-Switching-Modelle werden Wahrscheinlichkeiten berechnet, mit der sich die Volkswirtschaft in einer Rezession befindet.

Zur quantitativen Prognose werden entweder univariate oder multivariate Zeitreihenmodelle verwendet. Das am häufigsten eingesetzte Benchmark-Modell ist der autoregressive Prozess (AR-Prozess). Hier wird die Zeitreihe durch eigene verzögerte Werte erklärt. Gute Konjunkturindikatoren verbessern die Prognose dieses Modells, indem verzögerte Indikatorwerte in das Modell einfließen. Im multivariaten Fall wird meist ein bivariates vektorautoregressives (VAR-) Modell geschätzt. Dieses besteht aus zwei Gleichungen, in denen die zwei Zeitreihen wechselseitig abhängig von verzögerten Werten dargestellt werden. Diese Gleichungen (Vektoren) werden dann simultan geschätzt und zur Prognose verwendet. Der Vorteil dieser Art multivariater Zeitreihenmodelle ist, dass mögliche Feedback-Reaktionen explizit modelliert und geschätzt werden können.

Die oben vorgestellten Zeitreihenmodelle zur quantitativen Prognose sind linearer Natur. In jüngerer Zeit werden jedoch auch für die quantitative Prognose mehr und mehr nichtlineare Zeitreihenmodelle, wie z.B. autoregressive Mar-

kov-Switching-Modelle, verwendet. Für Deutschland gibt es noch keine explizite Analyse der Prognosemöglichkeiten dieser Modelle. Für eine Anwendung auf den Euroraum siehe Clavaria et al. (2006).

Die quantitative Prognose von Zeitreihen kann auf zwei Arten erfolgen. Bei der In-Sample-Prognose wird das Modell mit allen zur Verfügung stehenden Daten geschätzt, um dann einzelne Zeitpunkte innerhalb des Samples vorherzusagen. Bei der Out-of-Sample-Prognose wird das Modell mit Daten bis zu einem bestimmten Zeitpunkt t geschätzt und dann eine Prognose über diesen Zeitraum hinaus (außerhalb des benutzten Samples) berechnet. Die Beurteilung der Qualität von Indikatoren kann sich je nach verwendeter Prognoseart stark unterscheiden. So kann es sein, dass ein Indikator sehr gute Prognosen bei der In-Sample-Schätzung liefert, jedoch bei der Out-Of-Sample-Prognose sehr schlecht abschneidet. Aufgrund der größeren Realitätsnähe von Out-Of-Sample-Prognosen werden diese am häufigsten in der wissenschaftlichen Literatur verwendet.

Zur Beurteilung der Prognosequalität eines Modells gibt es in der Literatur verschiedene Maße. Das am häufigsten verwendete ist die Wurzel aus dem durchschnittlichen quadratischen Prognosefehler (Root mean squared error = RMSE). Je niedriger dieser ist, desto besser ist die Prognose. Da dieses Maß jedoch skalenabhängig ist, setzt man es meist ins Verhältnis zu einem Referenzmodell (manchmal auch als »naive Prognose« bezeichnet). Dieses skalenunabhängige Maß wird auch als Theil's U bezeichnet. Ein Wert von kleiner als 1 bedeutet, dass das Prognosemodell besser ist als das Referenzmodell. In den neunziger Jahren sind Tests entwickelt worden, die untersuchen, ob der Unterschied der Prognosefehler zwischen zwei Modellen statistisch signifikant ist. Der am häufigsten verwendete Test in der Prognoseliteratur ist der Diebold-Mariano-Test.

Die Kohärenz mit der Referenzreihe und der konjunkturelle Vorlauf

Bei der Untersuchung von Korrelationskoeffizienten des ifo Geschäftsklimas im verarbeitenden Gewerbe mit der Referenzreihe (vgl. jeweils die Tabelle) in Döpke, Krämer und Langfeldt (1994) weist ifo einen Vorlauf von neun Monaten aus. Ebenso ist der Test auf Granger-Kausalität signifikant positiv. Die Autoren führen auch einen statistischen Test durch, der zeigt, dass der Zusammenhang zwischen den Zeitreihen über den Zeitverlauf konstant ist. Die Methode der Kreuzkorrelationen wird auch in Breitung und Jagodzinski (2001) verwendet. Ähnlich wie im Fall der Wendepunkte besitzt das ifo Geschäftsklima einen Nachlauf von einem Monat und die Erwartungen einen Vorlauf von zwei Monaten. Der Test auf Granger-Kausalität fällt ebenfalls positiv aus. Eine Spektralanalyse in Fritsche und Stephan (2002) zeigt, dass sowohl

das Geschäftsklima als auch die Erwartungen für die Industrie eine hohe Kohärenz mit der Referenzzeitreihe aufweist. Darüber hinaus zeigen die Autoren, dass die ifo-Indikatoren signifikant Granger-kausal sind und somit einen zusätzlichen Erklärungsgehalt besitzen. Zu den gleichen Ergebnissen kommen auch Hüfner und Schröder (2002), Hinze (2003) sowie Dreger und Schumacher (2005).

Langmantel (1999) zeigt mit Hilfe des Konzepts der Granger-Kausalität, dass das ifo Geschäftsklima einen signifikanten Zusammenhang mit der Jahreswachstumsrate des realen Bruttoinlandsprodukts besitzt. Hinze (2003) zeigt, dass die ifo Geschäftserwartungen auf Basis des Konzepts der Kreuzkorrelationen einen effektiven Vorlauf von zwei Quartalen aufweisen.

Die Vorhersage von konjunkturellen Wendepunkten

Nach Döpke, Krämer und Langfeldt (1994) besitzt das ifo Geschäftsklima einen durchschnittlichen Vorlauf von 5,33 Monaten bei den unteren konjunkturellen Wendepunkten und 5,5 Monate bei den oberen Wendepunkten. Dagegen besitzt das ifo Geschäftsklima in Breitung und Jagodzinski (2001) einen durchschnittlichen Nachlauf von zwei Monaten und die Geschäftserwartungen einen Vorlauf von zwei Monaten gegenüber der Jahreswachstumsrate der Industrieproduktion.

Einen anderen Ansatz verfolgt Funke (1997). Mit Hilfe eines Probit-Modells wird die Fähigkeit des ifo Geschäftsklimas (in Jahreswachstumsraten) für die Vorhersage von Rezessionen untersucht. Es zeigt sich, dass der Indikator eine Rezession im Durchschnitt mit einem Quartal (bezogen auf die Quartalswachstumsrate des BIP) Vorlauf anzeigt.

Döpke (1999) erweitert den Ansatz von Funke (1997) um eine dynamische Komponente. Angewendet auf die Industrieproduktion kommt der Autor zu dem Schluss, dass das ifo Geschäftsklima und seine Bestandteile sehr schlecht abschneiden und wenn, dann nur für eine sehr kurzfristige Prognose von Rezessionen geeignet sind. Am besten würden noch die Erwartungen abschneiden.

Fritsche und Kouzine (2005) verwenden neben der Probit-Analyse auch den Markov-Switching-Ansatz. Die Autoren kommen zu dem Schluss, dass im Rahmen des Probit-Verfahrens die ifo Erwartungen einen Vorlauf von drei Monaten aufweisen. Jedoch verwenden die Autoren im Vergleich zu Döpke (1999) die Zeitreihe der Hersteller von Zwischenprodukten und nicht die Gesamtreihe des verarbeitenden Gewerbes. Wird jedoch das dynamische Modell verwendet, steigt der Vorlauf auf sieben Monate. Diese Ergebnisse werden durch den Markov-Switching-Ansatz bestätigt.

Bandholz und Funke (2003) verwenden ein dynamisches Faktormodell, um einen eigenen Konjunkturindikator zu konstruieren. Zur Berechnung dieses Index nutzen die Autoren auch die Zeitreihe der Beurteilung des Fertigwarenbestandes im verarbeitenden Gewerbe aus dem ifo Konjunkturtest. Die Autoren zeigen im Rahmen eines Markov-Switching-Ansatzes, dass dieser Indikator besser in der Lage ist, Rezessionen vorherzusagen, als das ifo Geschäftsklima.

Die quantitative Prognoseeigenschaft

Im Folgenden stellen wir eine Zusammenfassung der Ergebnisse von Out-of-Sample-Prognosen dar. Dabei trennen wir zwischen Industrieproduktion und Bruttoinlandsprodukt. Die Reihenfolge der Artikelennennungen bestimmt sich nach dem Erscheinungsjahr der Veröffentlichungen.

Industrieproduktion

Breitung und Jagodzinski (2001) schätzen ein bivariates VAR-Modell und vergleichen dieses mit einem univariaten AR(p)-Prozess. Rekursiv werden von 1999:01 bis 2001:06 30 Einschrittprognosen berechnet. Wird das komplette bivariate Modell zur Vorhersage benutzt (ohne Restriktionen), so schneiden sowohl das ifo Geschäftsklima als auch die Erwartungskomponente schlechter als das Referenzmodell ab. Wird jedoch ein reduziertes Modell zur Prognose verwendet, indem die insignifikanten Parameter weggelassen werden, so sind beide Indikatoren besser als das Referenzmodell. Auch zeigt die Analyse, dass die Prognosen keine systematischen Verzerrungen aufweisen.

Ähnlich gehen auch Fritsche und Stephan (2002) vor. Auch sie schätzen ein bivariates Modell und vergleichen es mit einem univariaten AR(p)-Prozess. Für den Zeitraum von 1991:01 bis 1998:12 werden rekursiv drei und sechs Monatsprognosen durchgeführt. Im Durchschnitt liefert das ifo Geschäftsklima der Investitionsgüterindustrie, des verarbeitenden Gewerbes und die Erwartungen des verarbeitenden Gewerbes sowie der Hersteller von Zwischenprodukte eine bessere Vorhersage als das Referenzmodell.

Wie in den beiden vorherigen Studien verwenden auch Hüfner und Schröder (2002) ein bivariates VAR-Modell, jedoch ist die Lag-Struktur in dem Modell über den Zeitablauf konstant. Als Referenzmodell wird hier der so genannte Random Walk benutzt, d.h. die Vorhersage ist einfach der Wert der Vorperiode. Bei der rekursiven Prognose (1994:01–2000:09) ist der Prognosehorizont jeweils eins, drei, sechs, neun und zwölf Monate. Nach dem Kriterium Theil's U sind die ifo Geschäftserwartungen bis auf den Zwölf-Monatszeitraum besser als die naive Prognose. Die Anwendung des Diebold-Mariono-Tests zeigt jedoch, dass die Unterschiede nur für einen Horizont von sechs Monaten signifikant sind.

Benner und Meier (2004) gehen noch einen Schritt weiter als Hüfner und Schröder (2002). Sie erweitern das bivariate VAR- zu einem Vector-Error-Correction-Modell (vgl. Lütkepohl 2005) und bestimmen die Modellstruktur rekursiv. Darüber hinaus verwenden sie nur die Informationen, die tatsächlich zum Prognosezeitpunkt vorliegen. Dagegen nehmen Hüfner und Schröder (2002) an, dass die ifo Geschäftserwartungen für den Prognosezeitraum schon bekannt sind (also die tatsächlichen Werte benutzen). Bei zunächst konstanter Modellstruktur erweisen sich die ifo Geschäftserwartungen schlechter als das autoregressive Referenzmodell (Prognosezeitraum: 1994:01–2000:0). Wird das Schätzmodell jedoch bei jeder Prognose neu bestimmt, so weisen die ifo Erwartungen für alle Prognosehorizonte niedrigere RMSEs aus. Der Unterschied ist aber nach dem Diebold-Mariono-Test nicht signifikant. Wird der Prognosezeitraum zu 1996:01–2002:10 variiert, so sind die ifo Erwartungen für die Prognosehorizonte drei, sechs, neun und zwölf Monate besser. Der Unterschied für den Neun-Monatszeitraum ist auch statistisch signifikant. Hier zeigt sich, dass schon eine Veränderung des Schätzzeitraumes zu einer unterschiedlichen Bewertung der Prognosegüte eines Indikators führen kann. In einer letzten Variation vergrößern die Autoren den Schätzzeitraum. In der Periode von 1968 bis 2002 sind die ifo Geschäftserwartungen bis auf dem Zwölf-Monatszeitraum immer signifikant besser als das autoregressive Referenzmodell.

Der Sachverständigenrat (2005) prognostiziert in seinem Gutachten sowohl die Monats- als auch die Jahreswachstumsrate mit Hilfe einer univariaten Gleichung. Im Rahmen einer konstanten Modellspezifikation schneiden die ifo Erwartungen besser als die naive Prognose (AR(p)-Prozess) für die Jahreswachstumsrate der Industrieproduktion ab. Jedoch ist dieser Unterschied nach dem Diebold-Mariono-Test nicht signifikant. Wird die Wachstumsrate zum Vormonat prognostiziert, ergibt sich, dass die ifo Erwartungen für einen Prognosehorizonten von einem, vier und sechs Monate signifikant bessere Ergebnisse als das Referenzmodell liefern.

Zu einem anderen Ergebnis kommen Dreger und Schumacher (2005). Mit Hilfe eines bivariaten VAR-Modells liefert das ifo Geschäftsklima im Vergleich zu dem autoregressiven Referenzmodell nie eine statistisch signifikante bessere Prognose. Leider werden keine RMSEs oder Theil's U dargestellt, um festzustellen, ob der Indikator bessere Vorhersagen liefert.

Nierhaus und Sturm (2005) zeigen mit Hilfe eines univariaten Modells, dass das ifo Geschäftsklima und seine Bestandteile besser als das Referenzmodell prognostizieren. Das verwendete Modell unterscheidet sich gegenüber den anderen dadurch, dass hier auch saisonale Einflüsse model-

liert werden. Einen statistischen Test auf systematische Unterschiede der Prognosen geben die Autoren nicht an.

Den Aspekt von Revisionen der amtlichen Statistik und deren Vorhersage untersuchen Jacobs und Sturm (2005). Sie zeigen mit Hilfe eines Real-Time-Datensatzes, dass die ifo-Zeitreihen zur Beurteilung der Produktion gegenüber dem Vormonat und die aktuelle Geschäftslage insbesondere die erste Revision der Industrieproduktion vorherzusagen hilft.

Bruttoinlandsprodukt

Während das Bruttoinlandsprodukt quartalsweise erhoben wird, wird das ifo Geschäftsklima monatlich veröffentlicht. Da die Standardzeitreihenmodelle jedoch voraussetzen, dass die Zeitreihen mit gleicher Frequenz vorliegen, muss eine Transformation erfolgen. Entweder wird das Bruttoinlandsprodukt auf Monatebene interpoliert oder der Indikator auf Quartalsebene aggregiert. Letzteres erfolgt meist über die Durchschnittsbildung über drei Monate. In der Literatur hat sich letzteres Verfahren durchgesetzt. Es muss jedoch darauf hingewiesen werden, dass durch die Aggregation Informationen verloren gehen.

Langmantel (1999) zeigt mit Hilfe eines univariaten Schätzmodells, dass die jährliche Wachstumsrate des realen Bruttoinlandsprodukts mit Hilfe des ifo Geschäftsklimas prognostiziert werden kann. Der Autor verwendet jedoch keine statistischen Maße, um die Prognosequalität zu beurteilen. Er erklärt verbal, warum bestimmte Abweichungen vom tatsächlichen Wert eingetreten sind.

Hinze (2003) kommt zu dem Schluss, dass für die Jahre 2001 und 2002 die ifo Geschäftserwartungen keine befriedigenden Ergebnisse im Rahmen eines bivariaten VAR-Modells liefern, obwohl der durchschnittliche relative RMSE bzgl. des Referenzmodells kleiner als 1 ist.

Schuhmacher und Dreger (2004) verwenden ein univariates Modell, um die Jahreswachstumsrate des Bruttoinlandsprodukts vorherzusagen. Im Vergleich zu einem AR(p)-Prozess als Referenzmodell besitzt das ifo Geschäftsklima für einen Prognosehorizont von vier Quartalen eine bessere Prognosegüte, jedoch erweist sich erneut, dass dieser Unterschied nicht statistisch signifikant ist.

Einen komplett anderen Ansatz verfolgen Mitnik und Zadrozny (2005). Sie verwenden ein neues multivariates Zeitreihenmodell, welches erlaubt, Zeitreihen unterschiedlicher Frequenzen gleichzeitig zu schätzen. Dies hat den Vorteil, dass die Indikatoren nicht aggregiert werden müssen und so keine Informationen verloren gehen. Mit Hilfe dieses Ansatzes zeigen Mitnik und Zadrozny (2005), dass das ifo Geschäftsklima und seine Bestandteile erheblich zur Verbes-

serung der Prognose des Bruttoinlandsprodukts beitragen. Das Prognosemaß Theil's U ist teilweise weit unter 1. Einen formalen Test auf systematische Verbesserung führen die Autoren jedoch nicht durch.

Ein univariates Prognosemodell wird in Kholodilin und Siliverstovs (2006) zur Prognose der annualisierten Wachstumsrate des BIP eingesetzt. Da sie einen Strukturbruch im Jahr 2001 in der Zeitreihe feststellen, untersuchen sie sowohl die Prognoseleistung vor dem Jahr 2001, für den Zeitraum 2001–2004 als auch für den gesamten Zeitraum. Es zeigt sich, dass das ifo Geschäftsklima und die Erwartungen für den Gesamtzeitraum für einen Prognosehorizont von ein und zwei Quartalen bessere Prognosen liefert als das Referenzmodell, was hier der langfristige Durchschnitt der Wachstumsrate der Bruttoinlandsprodukts ist. Die Analyse wurde auch mit den ersten Differenzen durchgeführt, und es zeigt sich, dass diese noch besser sind als die Ursprungszeitreihe. In dem Prognosezeitraum 1998:I bis 2001:I ist die Vorhersagequalität schlechter als im Bereich 2001:II bis 2004:IV. Die ifo-Indikatoren sind für die verschiedenen Zeiträume unter den besten sieben Indikatoren von 58 (inklusive der ersten Differenzen). Es zeigt sich, dass insbesondere die Differenzen erheblich zur Verbesserung der Prognose des BIP beitragen. Einen Test auf Signifikanz der Ergebnisse mit Hilfe des Diebold-Mariano-Tests führen die Autoren nicht durch, da sie der Ansicht sind, dass der Stichprobenumfang zu klein ist.

Das Gutachten des Sachverständigenrates (2006) kommt mit Hilfe eines univariaten Modells zu dem Schluss, dass das ifo Geschäftsklima und seine Bestandteile keine signifikant besseren Prognosen als das autoregressive Referenzmodell liefern. Die Ergebnisse seien »befriedigend«, da die Theil's U für bis zu sechs Quartale kleiner als 1 seien, aber nicht signifikant sind (Diebold-Mariano-Test). Die Prognosequalität hat zum Ende der Zeitreihen eher abgenommen. Es ist wichtig, darauf hinzuweisen, dass das Gutachten nur die Ergebnisse des ifo Konjunkturtest West verwendet. Ab 1991 beinhaltete die Referenzzeitreihe aber das gesamtdeutsche Bruttoinlandsprodukt.

Bei der Bewertung der Studien in dieser Kategorie muss noch beachtet werden, dass – der Diebold-Mariano-Test als Nullhypothese verwendet – die durchschnittlichen Prognosefehler der beiden untersuchten Modelle gleich sind. Wird diese Hypothese abgelehnt, ist eines der Modelle signifikant »besser«. Wird die Nullhypothese nicht abgelehnt, ergibt sich daraus aber noch nicht der Schluss, dass beide Modelle gleich gut prognostizieren. Dies rührt aus der generellen Konstruktionsweise solcher Tests her. Der Fehler, dass der Test die Nullhypothese nicht ablehnt, obwohl sie tatsächlich falsch ist, wird nicht kontrolliert.

Zu Beginn dieses Abschnitts wurde darauf verwiesen, dass der unterschiedlichen Messfrequenz der Zeitreihen Rech-

nung getragen werden muss. Gerade zu diesem Problem-bereich gibt es neuere Forschungsansätze, so dass hier in Zukunft ausgereifere Methoden zum Einsatz kommen könnten.

ifo-Befragungsergebnisse als Bestandteile anderer Indikatoren

Die ifo Konjunkturtestergebnisse werden nicht nur isoliert untersucht, sondern sind auch Bestandteil von zusammengesetzten Indikatoren. In den Handelsblatt-Indikator geht der ifo Geschäftsklimaindex für das verarbeitende Gewerbe mit einem Gewicht von 30% ein. Das ifo Geschäftsklima ist auch Bestandteil des deutschen Composite Leading Indicators der OECD. Letztere wurde u.a. in Kholodilin und Siliverstovs (2006) untersucht.

Die ifo-Indikatoren fließen oftmals in so genannte Faktormodelle ein. In diesen werden einzelne Faktoren aus einer Vielzahl von Zeitreihen extrahiert, die dann als Konjunkturindikatoren verwendet werden. Die ifo-Indikatoren werden in Faktormodellen u.a. in Schumacher und Dreger (2004), Schumacher (2005), Schumacher und Breitung (2006) eingesetzt. In Schumacher und Dreger (2004) wird das ifo Geschäftsklima mit dem Faktormodell verglichen. Das Faktormodell liefert zwar relativ gesehen bessere Ergebnisse bzgl. des Referenzmodells als das Geschäftsklima, jedoch ist dieser Unterschied nicht statistisch signifikant.

Fazit

Das Einsatzgebiet von Konjunkturindikatoren ist sehr weit gesteckt. Es gibt zahlreiche Zielsetzungen, denen solche Indikatoren dienen sollen. Neben den verschiedenen Zielsetzungen werden in wissenschaftlichen Untersuchungen aber auch unterschiedliche Zielgrößen als Referenzzeitreihen und verschiedene methodische Ansätze verwendet. Aus diesem Umstand ergibt sich, dass eine Vielzahl von wissenschaftlichen Untersuchungen zum ifo Geschäftsklima, als einem der prominentesten Konjunkturindikatoren in Deutschland, existiert. Trotz der beachtlichen Anzahl von Studien sind die methodischen Möglichkeiten aber noch lange nicht ausgeschöpft. Die Resultate dieser Forschungsarbeiten sind wichtig, nicht nur für die Prognose an sich, sondern auch für das Verständnis der Indikatoren. Doch auch wenn das ifo Geschäftsklima häufig als ein nützlicher Indikator angesehen wird, sollte man sich bei der Verallgemeinerung einzelner Studienergebnisse in Bescheidenheit üben. Denn die Zielsetzungen, denen die Indikatoren dienen sollen, sind vielseitig. Auf der anderen Seite muss sich jede einzelne Untersuchung auf ein relativ enges Studiendesign beschränken, so dass man aus jeder einzelnen Analyse nicht auf die generelle »Prognosefähigkeit« der Indikatoren schließen soll-

te. Gerade das wird aber gerne getan, indem man aus den Ergebnissen einer spezifischen Versuchsanordnung schließt, welches »der beste«, »der bessere« oder ein »unbrauchbarer« Indikator ist.

Grundsätzlich könnte man versuchen, einen Indikator wie das ifo Geschäftsklima auf eine bestimmte enge Zielsetzung hin zu optimieren. Dies könnte etwa dadurch geschehen, dass die Unternehmensmeldung über ein zielgerichtetes Gewichtungsschema aggregiert wird, so dass sie für einen Zweck optimal ist. Auch die Quantifizierung der dreistufigen kategorialen Unternehmensmeldungen könnte zweckgebunden optimiert werden. Gerade die Quantifizierung kategorialer Umfrageergebnisse ist ein großes eigenständiges Forschungsgebiet. Aus der vorliegenden Literaturdarstellung soll es aber ausgegrenzt bleiben, da es ein eigenständiger Bereich ist. Das ifo Institut verwendet mit der so genannten Saldenmethode ein einfaches, aber sehr effektives Quantifizierungsverfahren. Es hat sich, wie die verwendete Methodik generell, als ein über viele Anwendungen und Zielsetzungen hinweg nützliches Verfahren bewährt. Zudem besitzt es den Vorteil, dass es ein luzides Verfahren ist, das auch durch Dritte nachvollzogen und interpretiert werden kann.

Der vorliegende Überblick über die wissenschaftliche Literatur zeigt, dass das ifo Geschäftsklima für viele Problemfelder der Konjunkturanalyse gewinnbringend eingesetzt werden kann. Diese Vielseitigkeit gepaart mit dem transparenten Konstruktionsprinzip sind sicherlich wesentliche Gründe für die Popularität des Geschäftsklimas.

Literatur

- Abberger, K. (2006), »ifo Geschäftsklima und Produktionsindex im verarbeitenden Gewerbe«, *ifo Schnelldienst* 59(21), 42–45.
- Bandholz, H. und M. Funke (2003), »In Search of Leading Indicators of Economic Activity«, *Journal of Forecasting* 22, 277–297.
- Benner, J. und C-P. Meier (2004), »Prognosegüte alternativer Frühindikatoren für die Konjunktur in Deutschland«, *Jahrbücher für Nationalökonomie und Statistik* 224, 637–652.
- Breitung, J. und D. Jagodzinski (2001), »Prognoseeigenschaften alternativer Indikatoren für die Konjunktur in Deutschland«, *Konjunkturpolitik* 47, 292–314.
- Claveria, O., E. Pons und R. Ramos (2006), »Business and consumer expectations and macroeconomic forecasts«, *International Journal of Forecasting*, forthcoming.
- Döpke, J. (1999), »Predicting Germany's Recessions with Leading Indicators – Evidence from Probit Models«, Kiel Working Paper No. 944.
- Döpke, J., J.W. Krämer und E. Langfeldt (1994), »Konjunkturelle Frühindikatoren in Deutschland«, *Konjunkturpolitik* 40, 133–153.
- Dreger C. und C. Schuhmacher (2005), »Out-of-sample Performance of Leading Indicators for the German Business Cycle – Single vs. Combined Forecasts«, *Journal of Business Cycle Measurement and Analysis* 2, 71–87.
- Fritsche, U. und V. Kouzine (2005), »Prediction of in Germany«, *Jahrbücher für Nationalökonomie und Statistik* 225, 22–43.
- Fritsche, U. und S. Stephan (2002), »Leading Indicators of German Business Cycles – An Assessment of Properties«, *Jahrbücher für Nationalökonomie und Statistik* 222, 289–311.
- Funke, N. (1997), »Predicting Recessions: Some Evidence for Germany«, *Weltwirtschaftliches Archiv* 133, 90–102.
- Goldrian, G. (Hrsg., 2007), *Handbook Of Survey-Based Business Cycle Analysis*, Ifo Economic Policy Series, Edward Elgar, Cheltenham.

- Hamilton, J.D. (1994), *Time Series Analysis*, Princeton University Press, Princeton.
- Hinze, J. (2003), »Prognoseleistung von Frühindikatoren: Die Bedeutung von Frühindikatoren für Konjunkturprognosen – Eine Analyse für Deutschland«, HWWA Discussion Paper, No. 236.
- Hüfner, F.P. und M. Schröder (2002), »Prognosegehalt von ifo-Geschäftserwartungen und ZEW-Konjunkturerwartungen: Ein ökonometrischer Vergleich«, *Jahrbücher für Nationalökonomie und Statistik* 222, 316–336.
- Jacobs, J. und J.-E. Sturm (2005), »Do Ifo Indicators Help Explain Revisions in German Industrial Production«, in: J.-E. Sturm und T. Wollmershäuser (Hrsg.), *Ifo Survey Data in Business Cycle and Monetary Policy Analysis*, Physica-Verlag, Heidelberg, 93–114.
- Kholodilin, K.A. und B. Siliverstovs (2006), »On the Forecasting Properties of the Alternative Leading Indicators for the German GDP: Recent Evidence«, *Jahrbücher für Nationalökonomie und Statistik* 226, 234–259.
- Langmantel, E. (1999), »Das ifo Geschäftsklima als Indikator für die Prognose des Bruttoinlandsproduktes«, *ifo Schnelldienst* 52(16–17), 16–21.
- Lütkepohl, H. (2005), *New Introduction to Multiple Time Series Analysis*, Springer Verlag, Berlin.
- Mitnik, S. und P. Zadrozny (2005), »Forecasting Quarterly German GDP at Monthly Intervals Using Monthly Ifo Business Conditions Data«, in: J.-E. Sturm und T. Wollmershäuser (Hrsg.), *Ifo Survey Data in Business Cycle and Monetary Policy Analysis*, Physica-Verlag, Heidelberg, 19–48.
- Nierhaus, W. (1999), »Aus dem Instrumentenkasten der Konjunkturanalyse: Veränderungsdaten im Vergleich«, *ifo Schnelldienst* 52(27), 11–19.
- Nierhaus, W. und J.-E. Sturm (2005), »Methoden der Wirtschaftsprognose und Konjunkturindikatoren«, in: G. Goldrian (Hrsg.), *Handbuch der umfragebasierten Konjunkturforschung*, ifo Beiträge zur Wirtschaftsforschung Nr. 15, ifo Institut, München, 271–301.
- Sachverständigenrat (2005), *Jahresgutachten 2005/2006*, »Die Chance Nutzen – Reformen mutig voranbringen«, erhältlich unter: www.sachverstaendigenrat-wirtschaft.de.
- Sachverständigenrat (2006), *Jahresgutachten 2006/2007*, »Widerstreitende Interessen – Ungenutzte Chancen«, erhältlich unter: www.sachverstaendigenrat-wirtschaft.de.
- Schuhmacher, C. (2005), »Forecasting German GDP using alternative factor models based on large data sets«, Bundesbank Discussion Paper No. 24.
- Schuhmacher, C. und J. Breitung (2006), »Real-time forecasting of GDP based on a large factor model with monthly and quarterly data«, Bundesbank Discussion Paper, No. 33.
- Schuhmacher, C. und C. Dreger (2004), »Estimating Large-Scale Factor Models for Economic Activity: Do They Outperform Simpler Models«, *Jahrbücher für Nationalökonomie und Statistik* 224, 731–750.
- Strigel, W.H. (1985), »Das ifo Geschäftsklima als Konjunkturindikator – Eine Retrospektive«, *ifo Studien* 31, 29–68.