

Hülsewig, Oliver; Mayr, Johannes; Ulbricht, Dirk

Article

Zur Evaluierung von VAR-Prognosen

ifo Schnelldienst

Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Hülsewig, Oliver; Mayr, Johannes; Ulbricht, Dirk (2007) : Zur Evaluierung von VAR-Prognosen, ifo Schnelldienst, ISSN 0018-974X, ifo Institut für Wirtschaftsforschung an der Universität München, München, Vol. 60, Iss. 07, pp. 19-25

This Version is available at:

<https://hdl.handle.net/10419/164412>

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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

In diesem Beitrag wird die Verwendung von VAR-Modellen für die Prognose des realen Bruttoinlandsprodukts in den Vereinigten Staaten analysiert. Den Ausgangspunkt bildet ein Basismodell, das neben dem realen BIP den Verbraucherpreisindex sowie einen kurzfristigen Geldmarktsatz enthält. Das Basismodell wird anschließend durch die Einbeziehung zusätzlicher makroökonomischer Variablen schrittweise erweitert. Die Prognosen der einzelnen Modelle, die als out-of-sample Prognosen generiert werden, werden für verschiedene Schätzzeiträume anhand unterschiedlicher Maße zur Überprüfung der Anpassungsgüte evaluiert. Es zeigt sich, dass die Prognosen der VAR-Modelle durchschnittlich schlechter als die eines univariaten AR(2)-Prozesses abschneiden. Die Qualität der Prognosen variiert jedoch stark in den betrachteten Schätzzeiträumen, so dass die VAR-Prognosen in einzelnen Perioden die AR(2)-Prognosen deutlich dominieren. Für die Prognose impliziert dies, dass die Verwendung einer einzigen Modellspezifikation äußerst problematisch sein kann.

Überblick

In der Konjunkturprognose des ifo Instituts wird die wirtschaftliche Entwicklung zahlreicher Länder regelmäßig beleuchtet. Im Mittelpunkt steht die Prognose des realen Bruttoinlandsprodukts (BIP), das als Indikator für die konjunkturelle Dynamik dient. Zusätzlich werden die Verwendungskomponenten des BIP, die Verbraucherpreise, die wichtigsten Größen des Arbeitsmarkts sowie die Veränderung der Zinsen prognostiziert.

Die Konjunkturprognose des ifo Instituts stützt sich auf verschiedene Verfahren, bei denen vor allem ökonometrische Schätzansätze – neben Indikatormodellen und der iterativ-analytischen Methode – besonders bedeutend sind (vgl. Nierhaus und Sturm 2003). Vektorautoregressive (VAR) Modelle nehmen innerhalb der verwendeten ökonometrischen Schätzverfahren eine Schlüsselrolle ein, da diese für Prognosen aufgrund ihrer einfachen Struktur recht gut geeignet sind. VAR-Modelle basieren auf zeitreihenanalytischen Verfahren, in denen die arbiträre Unterscheidung zwischen endogenen und exogenen Variablen – die zum Beispiel in ökonometrischen Strukturmodellen wesentlich ist – entfällt. Für die Prognose werden daher keine Informationen aus dem Prognosezeitraum benötigt, weshalb unbedingte Vorhersagen über beliebig weite Zeiträume möglich sind.¹

In diesem Beitrag werden VAR-Modelle für die Vereinigten Staaten evaluiert, mit

deren Hilfe die Entwicklung des BIP prognostiziert werden soll. Den Ausgangspunkt bildet ein Basismodell, das neben dem realen BIP den Verbraucherpreisindex und einen kurzfristigen Geldmarktsatz enthält. Das Basismodell wird durch die Berücksichtigung weiterer makroökonomischer Variablen anschließend schrittweise erweitert. Für die Prognosen werden die Modelle mit Hilfe eines »Growing-Regression«-Ansatzes für unterschiedliche Zeiträume geschätzt, was die Gefahr einer Fehlausewahl aufgrund zufälliger Entwicklungen in bestimmten Zeitperioden vermindert. Die Lag-Struktur der Modelle wird dabei für die einzelnen Schätzzeiträume anhand unterschiedlicher Informationskriterien optimiert. Die Berücksichtigung verschiedener Schätzzeiträume führt dazu, dass auch die jeweiligen Prognosezeiträume alternieren. Die Prognosen der einzelnen Modellspezifikationen, die ausschließlich als out-of-sample Prognosen generiert werden, werden anhand zweier Maße zur Überprüfung der Anpassungsgüte, dem RMSE (Root Mean Squared Error) – also der Wurzel des mittleren quadratischen Prognosefehlers – sowie dem Theil'schen Ungleichheitskoeffizienten (Theil's U), evaluiert. Als Referenzmodell dient dabei ein univariater AR(2)-Prozess, dessen Lag-Struktur ebenfalls anhand der Informationskriterien optimiert wurde.

¹ Eine Prognose wird als unbedingt bezeichnet, wenn vorab keine Annahmen über exogene Variablen getroffen werden; hingegen heißt eine Prognose bedingt, wenn ihre Richtigkeit vom Zutreffen bestimmter Annahmen über die Entwicklung der exogenen Variablen abhängig ist.

Die Ergebnisse zeigen, dass die Prognosen der VAR-Modelle durchschnittlich schlechter als die des univariaten AR(2)-Prozesses abschneiden. Die Qualität der Prognosen variiert jedoch stark in den zugrunde liegenden Schätzzeiträumen, so dass die VAR-Prognosen in einzelnen Perioden die AR(2)-Prognosen deutlich dominieren. Für die Prognose impliziert dies, dass die Verwendung einer einzigen Modellspezifikation äußerst problematisch sein kann.

VAR-Prognosemodell

Allgemein können VAR-Modelle in reduzierter Form folgendermaßen dargestellt werden:

$$Y_t = \mu + A_1 Y_{t-1} + A_2 Y_{t-2} + \dots + A_p Y_{t-p} + \varepsilon_t,$$

wobei Y_t den Vektor der endogenen Variablen, A_i die Matrizen der Koeffizienten ($i=1, \dots, p$), μ den Vektor der deterministischen Terme – bspw. eine Konstante und einen linearen Trend – und ε_t den Vektor der Störterme bezeichnet. Die Anzahl der Lags ist durch p dargestellt.

Für die Prognose mit VAR-Modellen ist die Auswahl der Variablen entscheidend. Da sich VAR-Modelle als eine vereinfachte Darstellung dynamischer Zusammenhänge interpretieren lassen, sollten die endogenen Variablen im Vektor Y_t idealerweise alle Größen umfassen, die relevant sind. Allerdings ist die Anzahl der Variablen aufgrund der beschränkten Datenlage häufig begrenzt, so dass eine Vorauswahl der essentiellen Größen getroffen werden muss (vgl. Winker 2002, 247).

Neben der Auswahl der Modellvariablen ist die Wahl der richtigen Lag-Struktur wesentlich. Für die Festlegung der Lag-Länge werden Informationskriterien herangezogen, die an der Log-Likelihood-Funktion anknüpfen. Die Likelihood-Funktion errechnet die Schätzer der Funktionsparameter so, dass die Wahrscheinlichkeit, mit der die gegebenen funktionalen und verteilungstechnischen Annahmen zu den vorliegenden Daten passen, maximiert wird. Die Aufnahme zusätzlicher Lags führt notwendigerweise zu einer Verbesserung des Likelihood-Wertes; allerdings nehmen dadurch die Freiheitsgrade ab, so dass die Präzision der Regression nachlässt. Die Informationskriterien belegen daher die Aufnahme zusätzlicher Regressoren mit einem Strafterm, wobei dieser je nach Informationskriterium unterschiedlich ist. Die gebräuchlichsten Informationskriterien sind das Bayesianische Kriterium (BIC) und das Akaike Kriterium (AIC), die folgendermaßen definiert sind:

$$\text{BIC} = -2 \ln L(\hat{\theta}) + K \log(T)$$

$$\text{AIC} = -2 \ln L(\hat{\theta}) + 2K,$$

wobei $\hat{\theta}$ der geschätzte Parametervektor ist, $\ln L(\hat{\theta})$ den maximierten Wert der Likelihood-Funktion und K die Anzahl der endogenen Variablen darstellt. T beschreibt die Anzahl der Beobachtungen.

Die Prognose eines VAR-Modells wird durch ein rekursives Verfahren erstellt:²

$$Y_{T+\tau}^p = \hat{A}_1 Y_{T+\tau-1} + \dots + \hat{A}_p Y_{T+\tau-p},$$

wobei $Y_{T+\tau}^p$ die τ -Schritt-Prognose bezeichnet. \hat{A}_i , mit $i=1, \dots, p$ sind die geschätzten Koeffizienten des VAR-Modells. Für $\tau=1$ ergibt sich die Prognose durch das Einsetzen der beobachteten Werte (statische Prognose). Für die weiter in der Zukunft liegenden Prognoseschritte werden die bereits prognostizierten Werte rekursiv herangezogen (dynamische Prognosen).

Anpassungsgüte

Die Beurteilung der Anpassungsgüte der Prognosen basiert auf dem RMSE, welcher die Wurzel des mittleren quadratischen Prognosefehlers beschreibt:

$$\text{RMSE} = \sqrt{\frac{\sum_{t=T+1}^{T+h} (\hat{y}_t - y_t)^2}{h}},$$

wobei h (mit $h=1, \dots, H$) den zugrunde liegenden Prognosehorizont bezeichnet. Der RMSE berechnet sich folglich auf Basis der einzelnen Prognosefehler zum Zeitpunkt t , also der Differenz aus dem prognostizierten Wert \hat{y}_t und dem tatsächlichen Wert der Variable y_t . Da die Prognosefehler sowohl positive als auch negative Werte annehmen können, werden über einen bestimmten Zeitraum hinweg die quadrierten Differenzen zur Berechnung des Gütemaßes verwendet. Mit Hilfe des RMSE können dann verschiedene Prognosen für eine Zielvariable miteinander verglichen werden. Je höher der RMSE einer Prognose ausfällt, umso geringer ist die Güte der Vorhersage.

Die relative Qualität eines Prognosemodells zu anderen Modellen lässt sich anhand des Theil'schen Ungleichheitskoeffizienten überprüfen. Der Wert von Theil's U berechnet sich dabei aus dem Verhältnis zwischen dem RMSE des Prognosemodells und dem RMSE eines Referenzmodells:

$$\text{Theil's } U = \frac{\sqrt{\frac{\sum_{t=T+1}^{T+h} (\hat{y}_t - y_t)^2}{h}}}{\sqrt{\frac{\sum_{t=T+1}^{T+h} (\hat{y}_t^{\text{Referenz}} - y_t)^2}{h}}}.$$

² Für die Prognose wird der Störterm durch seinen Erwartungswert, der den Wert null annimmt, ersetzt.

Als Referenzmodell kann hierbei beispielsweise ein AR-Prozess der zu prognostizierenden Variable verwendet werden. Für Werte von Theil's U kleiner (größer) eins zeigt das spezifizierte Prognosemodell folglich eine bessere (schlechtere) durchschnittliche Prognosequalität als das jeweilige Referenzmodell.

Wahl der Modellvariablen

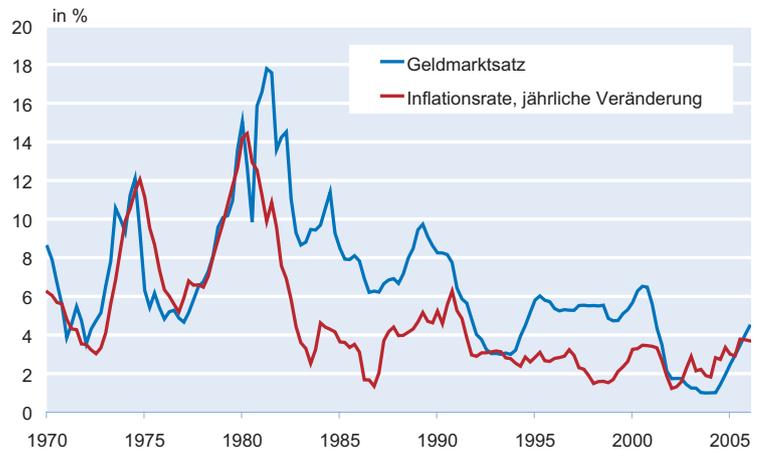
Mit Hilfe von VAR-Modellen soll im Folgenden die Entwicklung des realen BIP in den USA prognostiziert werden. Wegen ihrer dominierenden Stellung im internationalen Handel hat die konjunkturelle Lage der US-amerikanischen Wirtschaft entscheidenden Einfluss auf die aktuelle und zukünftige wirtschaftliche Entwicklung in Deutschland sowie im Euroraum. Aus diesem Grund wird der Prognose wichtiger Aggregate der US-Wirtschaft im Rahmen von Konjunkturprognosen stets ein hohes Gewicht beigemessen.

Für die Prognose des BIP sind bei der Auswahl der relevanten Größen zwei Ziele gegeneinander abzuwägen. Zum einen sollten alle Variablen, die einen Einfluss auf die realwirtschaftliche Entwicklung ausüben, berücksichtigt werden. Zum anderen steigt mit der Anzahl der Variablen in den Einzelgleichungen des VAR-Modells jedoch die Zahl der zu schätzenden Koeffizienten um den Faktor der Anzahl der Variablen sowie der festgelegten Lag-Länge. Der daraus resultierende Verlust an Freiheitsgraden verringert analog die Schärfe der Schätzungen.

Den Ausgangspunkt für die Prognose bildet ein Basismodell, welches das reale BIP, den Verbraucherpreisindex (P) und einen kurzfristigen Geldmarktsatz (R) enthält.³ Die Wahl der Variablen basiert zunächst auf der vereinfachten Vorstellung, dass das reale BIP maßgeblich durch die Entwicklung der Preise und der Zinsen beeinflusst wird, wobei mögliche Wechselwirkungen explizit berücksichtigt werden.

Um die Prognose des Basismodells zu verbessern, ist eine Erweiterung um zusätzliche makroökonomische Variablen denkbar. Für die Binnennachfrage sind der private Konsum (C), die Unternehmensinvestitionen (I) und die Staatsausgaben (G) entscheidend; analog wird der Außenhandel durch die Exporte (X) und die Importe (M) bestimmt. Die Lage am Arbeitsmarkt kann durch die Entwicklung der Arbeitslosen-

Inflationsrate und kurzfristiger Geldmarktsatz



quote (U) beschrieben werden.⁴ Der real effektive Wechselkurs (RER) lässt sich als Indikator für die relative Wettbewerbsfähigkeit der US-amerikanischen Wirtschaft interpretieren.⁵ Zusätzlich kann die Entwicklung des Ölpreises, beispielsweise der Preis für die Sorte Brent (BRENT), für die realwirtschaftliche Expansion bedeutend sein.⁶

Um die Anzahl der Regressoren der VAR-Modelle möglichst gering zu halten, wird das Basismodell im Folgenden jeweils nur um eine der oben genannten Variablen zusätzlich erweitert. Dabei werden alle Variablen als endogene Größen behandelt.⁷

Untersuchungszeitraum

Als Datenbasis für die Schätzung der VAR-Modelle und die darauf aufbauenden Prognosen des BIP werden saisonbereinigte Quartalswerte der amerikanischen Wirtschaft über einen Zeitraum von 1970 (1. Quartal) bis 2006 (1. Quartal) verwendet. Bereits die visuelle Analyse der Inflationsrate – gemessen an der Veränderung des Verbraucherpreisindex – und des kurzfristigen Geldmarktsatzes deutet jedoch an, dass sich die US-amerikanische Wirtschaft ab Mitte der achtziger Jahre grundlegend gewandelt hat (vgl. Abbildung).

Die Variabilität beider Zeitreihen ist im Zeitraum vom 1. Quartal 1970 bis zum 4. Quartal 1984 wesentlich deutlicher ausgeprägt als danach. So nehmen die Standardabweichungen der Inflationsrate und des Geldmarktsatzes von 3,0 auf 1,0% bzw. von 3,7 auf 2,3% ab. Gleichfalls ist zu beobachten

³ Die Quelle des Verbraucherpreisindex ist das Bureau of Labor Statistics, die des kurzfristigen Geldmarktsatzes die US Federal Reserve. Das BIP stammt, wie der überwiegende Teil der übrigen Daten, aus den Veröffentlichungen des US Departments of Commerce. Wo dies nicht der Fall ist, wird in einer separaten Fußnote darauf hingewiesen.

⁴ Quelle: Bureau of Labor Statistics.

⁵ Quelle: OECD Main Economic Indicators.

⁶ Quelle: Ecwin, Dated FOB North West Europe Spot Close.

⁷ Die Verwendung von BRENT als endogene Variable kann dadurch begründet werden, dass die USA als größter Nettoölverbraucher der Welt einen wesentlichen Einfluss auf die Veränderung des Ölpreises ausüben (vgl. hierzu auch Bernanke et al. 1997).

ten, dass auch die durchschnittliche Höhe der beiden Größen gesunken ist. Der Mittelwert des Geldmarktsatzes nimmt von 8,2 auf 5,1% ab, und der Mittelwert der Inflationsrate fällt von 7,0 auf 3,0%.

Die Analyse wird daher – angesichts des wahrscheinlichen Strukturbruchs – auf den aktuelleren Zeitraum ab 1985 (1. Quartal) beschränkt. Zwar könnten durch Hinzunahme der älteren Beobachtungen Freiheitsgrade gewonnen werden, allerdings ist durch den möglichen Strukturbruch eine erhebliche Verzerrung der Schätzergebnisse zu erwarten.

Eine Schätzung von VAR-Modellen in Niveauebenen der endogenen Variablen birgt die Gefahr einer Scheinkorrelation (vgl. Granger und Newbold 1974). Engle und Granger (1987) haben gezeigt, dass bei Vorliegen von Kointegrationsbeziehungen auch eine Regression von nichtstationären Variablen in Niveauebenen zu konsistenten Ergebnissen führt. Kann die Stationarität der Variablen nicht statistisch bestätigt und auch die Existenz von Kointegrationsbeziehungen nicht nachgewiesen werden, so ist eine Schätzung in Differenzen notwendig.⁸ Aufgrund der geringen Mächtigkeit der Stationaritätstests werden im Folgenden die Schätzungen sowohl in Niveaus als auch in ersten Differenzen durchgeführt.⁹

Da einzelne Zeitreihen (bspw. das BIP) trendbehaftet sein können, kann für die Schätzung der VAR-Modelle in Niveauebenen die Aufnahme einer linearen Komponente sinnvoll sein und die Ergebnisse gegebenenfalls verbessern.

Ergebnisse

Die Schätzung der VAR-Modelle für die USA zielt darauf ab, möglichst exakte Prognosen für die Entwicklung des realen BIP zu generieren. Für die Beurteilung der Prognosen scheint es sinnvoll zu sein, die einzelnen Modelle nicht lediglich anhand ihres Abschneidens zu einem bestimmten Zeitpunkt zu beurteilen, da hierbei auch zufällige Übereinstimmungen eine entscheidende Rolle spielen können. Um robuste Ergebnisse zu gewährleisten, ist vielmehr eine Beurteilung der Modelle anhand ihrer Prognosegüte zu verschiedenen Zeitpunkten der Vergangenheit zweckmäßig.

Für jedes der ausgewählten Modelle werden daher 17 achtperiodige out-of-Sample Prognosen erstellt und anhand der resultieren-

den RMSE bewertet, wobei das Verfahren der »Growing Regression« angewandt wird. Die Parameter der VAR-Modelle werden in einem ersten Schritt zunächst auf Basis von 61 Beobachtungen, d.h. mit Daten von 1985 (1. Quartal) bis 2000 (1. Quartal), geschätzt. Aufbauend auf den geschätzten Koeffizienten werden die folgenden acht Quartale prognostiziert und die daraus resultierenden Prognosefehler berechnet. Der Schätzzeitraum wird anschließend um ein Quartal erweitert und darauf basierend jeweils erneut eine achtperiodige Prognose erstellt. Der maximale Schätzzeitraum für die VAR-Modelle der letzten Iteration umfasst die Jahre 1985 (1. Quartal) bis 2004 (1. Quartal), so dass die letzte Prognose im 1. Quartal 2006 endet. Die Lag-Struktur der verschiedenen VAR-Modelle wird für jeden Schätzzeitraum mittels der Informationskriterien jeweils flexibel optimiert.

In Tabelle 1 sind die mit Hilfe des AIC und BIC bestimmten durchschnittlichen Lag-Längen über die Schätzzeiträume für die Modelle sowohl in Niveaus – mit und ohne linearen Trend – als auch in Differenzen wiedergegeben.

Die Betrachtung der Lag-Längen beider Kriterien zeigt, dass das BIC die Aufnahme zusätzlicher Regressoren stärker be-

⁹ Für die Schätzung von VAR-Modellen makroökonomischer Variablen wird zumeist die Verwendung von logarithmierten Zeitreihen präferiert. Die Logarithmierung der Variablen wird dabei mit einer Reduktion der Heteroskedastizität sowie dem Einfluss von Ausreißern gerechtfertigt. Zudem können die ersten Differenzen der logarithmierten Zeitreihen direkt als Wachstumsraten interpretiert werden. Analog dieser Argumentation werden auch im Folgenden die Schätzungen ausschließlich auf Basis logarithmierter Werte durchgeführt. Zur Erstellung von Prognosen der nicht-transformierten BIP-Zeitreihe auf Basis von Schätzungen in logarithmierten Werten ist eine Rücktransformation der resultierenden Modellprognosen notwendig. Die einfache Bildung von Exponenten der prognostizierten logarithmierten Werte führt dabei zu systematischen Verzerrungen, so dass ein direkter Vergleich mit den Prognosen auf Basis nicht-transformierter Zeitreihen keine interpretierbaren Ergebnisse liefert. Arino und Franses (1996) zeigen ein Verfahren zur Bereinigung der Prognosen um diese Abweichungen. Aufgrund der mangelnden direkten Vergleichbarkeit wird im Folgenden auf einen Vergleich der Verfahren verzichtet und die Evaluation der Prognosen ausschließlich auf Basis logarithmierter Werte diskutiert.

Tab. 1
Durchschnittliche Lag-Längen über die Schätzzeiträume nach AIC und BIC

	AIC			BIC		
	Differenzen	Niveau		Differenzen	Niveau	
Trend	–	–	x	–	–	x
X	1.00	5.59	4.00	1.00	1.71	1.65
M	1.00	6.53	6.41	1.00	1.76	1.76
U	1.00	2.53	2.47	1.00	1.06	1.00
G	1.82	3.65	3.41	1.00	1.47	1.47
Basis	1.12	4.12	3.71	1.00	2.00	2.00
RER	1.00	3.59	4.47	1.00	1.53	1.53
BRENT	2.00	3.41	4.41	1.00	2.00	2.00
CONS	2.76	3.76	4.76	1.00	2.00	1.82
I	3.88	4.71	7.29	1.00	1.88	2.00

⁸ Bei einem Integrationsgrad höherer Ordnung müssen entsprechend mehrmals Differenzen gebildet werden.

Tab. 2
RMSE für die einzelnen Modelle

	Differenzen	Rang	Niveau (ohne Trend)	Rang	Niveau (mit Trend)	Rang
X	387.28	1	503.60	5	608.30	6
M	390.19	2	700.84	9	662.45	8
U	391.99	3	589.90	8	695.39	9
G	394.20	4	476.05	3	484.24	2
Basis	399.13	5	353.90	1	461.07	1
RER	406.49	6	495.52	4	571.62	5
BRENT	426.96	7	547.85	7	615.45	7
CONS	476.05	8	525.29	6	514.30	4
I	477.93	9	461.68	2	493.65	3
∅	416.69		517.18		567.39	

strafft und somit sparsame Modelle bevorzugt. Während sich für die Niveauschätzungen die durchschnittliche Verzögerungsstruktur unter Anwendung des AIC zwischen 2,47 und 7,29 bewegt, resultiert für das BIC lediglich eine Struktur von maximal zwei Lag-Längen. Bezüglich der Schätzung in Differenzen verhält es sich entsprechend. Die Spanne reicht von 1 bis 3,88 für das AIC gegenüber einer starren Struktur von einem Lag im Falle des BIC. Die durch das BIC bestimmte relativ kurze Lag-Struktur bildet die Dynamik der Residuen jedoch nur unzureichend ab, so dass weiterhin Autokorrelation auftritt. Um die damit verbundenen Verzerrungen in den Schätzungen zu vermeiden, wird im Folgenden für die Festlegung der Lag-Strukturen lediglich das AIC berücksichtigt.

Als Kriterium für die Auswahl der für Prognosezwecke am besten geeigneten Modelle kann wie oben bereits beschrieben der RMSE herangezogen werden. Um die Vergleichbarkeit der Ergebnisse zu gewährleisten, sind für die Modelle in Differenzen die Prognosewerte in Niveaus rücktransformiert worden.

Tabelle 2 gibt die, über die verschiedenen Schätzzeiträume und Prognosehorizonte ermittelten, durchschnittlichen RMSE für die neun untersuchten Modellspezifikationen wieder. Die zweite Spalte bildet die Ergebnisse für die Schätzungen in Differenzen ab, die vierte und sechste Spalte zeigen die Ergebnisse für die Niveauschätzungen ohne bzw. unter Berücksichtigung eines linearen Trends. Der jeweilige Rang innerhalb einer Modellklasse ist in den Spalten hinter den RMSE-Werten angegeben.

Sowohl für die Differenzen- als auch die Niveauschätzungen zeigt die Basisspezifikation (Basis), d.h. unter ausschließlicher Berücksichtigung von BIP, Preisniveau und kurzfristigen Zinsen, relativ gute Ergebnis-

se. Lediglich für die Schätzungen in Differenzen führt eine Erweiterung um eine zusätzliche Variable für einige Spezifikationen zu einer Verbesserung der Prognosegüte. Insbesondere die Mitaufnahme einer Variablen aus der Gruppe der Exporte (X), der Importe (M), der Arbeitslosenquote (U) und der Staatsausgaben (G) erhöht die Güte der Modelle zur Prognose. Die insgesamt beste Spezifikation stellt hierbei das um die Exporte erweiterte Basismodell in ersten Differenzen dar.

Die relative Prognosegüte der Modelle lässt sich mit Hilfe von Theil's U, welches die RMSEs der Modelle ins Verhältnis zu dem

RMSE eines Referenzmodells stellt, beurteilen. Im Folgenden wird ein univariater AR-Prozess in Niveaufwerten mit linearem Trend als Referenzmodell verwendet. Um eine bessere Vergleichbarkeit zu gewährleisten, wird dessen Verzögerungsstruktur ebenfalls auf Basis des AIC festgelegt. Hierbei ergibt sich über sämtliche Schätzzeiträume eine konstante Lag-Struktur von zwei Verzögerungen. Der über die Zeit und die Horizonte gemittelte, durchschnittliche RMSE des AR(2)-Prozesses beträgt 387,46. Die auf Grundlage dieses Wertes ermittelten Theil's U der einzelnen Modelle sind in Tabelle 3 dargestellt. Die Spaltenbeschreibungen sind analog zu Tabelle 2 gewählt.

Mit Ausnahme des Basismodells in Niveaufwerten ohne Trendkomponente zeigen sämtliche Modellspezifikationen Werte von Theil's U größer eins und somit relativ zum AR-Prozess eine schwächere Prognosegüte. Dieses Ergebnis stützt die in der aktuellen Forschung weit verbreitete Feststellung, dass ein univariater AR(2)-Prozess bereits eine sehr gute Prognosegüte bezüglich des Bruttoinlandsproduktes gewährleistet. Insgesamt zeigen die Ergebnisse, dass über die hier untersuchten Spezifikationen Prognosen auf Basis von Differenzen der logarithmierten Werte im Mittel der untersuchten Schätzzeiträume

Tab. 3
Theil's U für die einzelnen Modelle

	Differenzen	Rang	Niveau (ohne Trend)	Rang	Niveau (mit Trend)	Rang
X	1.00	1	1.30	5	1.57	6
M	1.01	2	1.81	9	1.71	8
U	1.01	3	1.52	8	1.79	9
G	1.02	4	1.23	3	1.25	2
Basis	1.03	5	0.91	1	1.19	1
RER	1.05	6	1.28	4	1.48	5
BRENT	1.10	7	1.41	7	1.59	7
CONS	1.23	8	1.36	6	1.33	4
I	1.23	9	1.19	2	1.27	3
∅	1.08		1.33		1.46	

Tab. 4
Rangfolge der Modelle über die Horizonte (Mittel der Schätzzeiträume)

Horizonte	1 Q	2 Q	3 Q	4 Q	5 Q	6 Q	7 Q	8 Q
X	6	4	4	2	4	2	2	1
M	1	3	3	1	2	4	3	6
U	7	2	2	3	1	3	6	3
G	2	7	6	4	5	5	4	2
Basis	5	5	5	5	6	6	5	5
RER	8	6	8	7	7	7	7	7
BRENT	3	8	7	8	8	8	8	8
CONS	9	9	9	10	10	10	10	10
I	10	10	10	9	9	9	9	9
AR	4	1	1	6	3	1	1	4

me sowie Prognosehorizonte deutlich besser abschneiden als auf Basis von Niveauwerten. Aus diesem Grund werden die Niveaumodelle im Folgenden nicht weiter berücksichtigt.

Beurteilung der Robustheit

Für die Auswahl lediglich eines für Prognosezwecke besten Modells ist die Robustheit der Ergebnisse von entscheidender Bedeutung. Das präferierte Modell sollte optimalerweise den Alternativmodellen sowohl über jeden Prognosehorizont als auch zu jedem Zeitpunkt überlegen sein.

Um zunächst einen Einblick in die Stabilität der Rangfolge der Modelle über die Prognosehorizonte zu gewinnen, ist in Tabelle 4 die Rangfolge der Modelle für die einzelnen Horizonte abgebildet. Im Gegensatz zum obigen Ansatz ist der RMSE hierbei über die Schätzzeiträume, nicht aber über die Horizonte gemittelt. Bemerkenswert ist hierbei insbesondere, dass die Prognosegüte der einzelnen Modelle je nach dem zu prognostizierenden Horizont stark alterniert.

Die Robustheit der Rangfolge über die Schätzzeiträume ist in Tabelle 5 dargestellt. Analog zu dem vorherigen Vorgehen sind die zugrunde liegenden RMSE hierbei nicht über die Zeiträume, sondern über die Horizonte gemittelt. Als zusätzliche Spezifikation ist hierbei der oben dargestellte AR-Prozess berücksichtigt.

Es zeigt sich, dass die Rangfolge auch über die Schätzzeiträume stark variiert. Der im Durchschnitt dominierende AR(2)-Prozess zeigt in weniger als der Hälfte der Schätzzeiträume die beste Prognosegüte. In gleicher Häufigkeit wird er von mindestens sechs Modellen geschlagen und belegt zu vier Zeitpunkten sogar den letzten Rangplatz. Auch die Berechnung des durchschnittlichen Rangplatzes der Modelle lässt an der Überlegenheit des AR(2)-Prozesses zweifeln. Den im Ø besten Rangplatz zeigt das um die Arbeitslosenquote erweiterte Basismodell. Für die Prognose bedeutet dies, dass die Verwendung einer einzelnen Modellspezifikation äußerst problematisch sein kann.

Schlussbemerkung und Ausblick

In diesem Beitrag wurde die Verwendung von VAR-Modellen für die Prognose des realen Bruttoinlandsprodukts in den USA diskutiert. Zunächst wurde ein Basismodell, das neben dem realen BIP den Verbraucherpreisindex und einen kurzfristigen Geldmarktsatz enthält, spezifiziert. Dieses wurde anschließend durch die Aufnahme zusätzlicher makroökonomischer Variablen erweitert. Für die Prognosen wurden die Modelle mit Hilfe eines »Growing-Regression«-Ansatzes für unterschiedliche Zeiträume geschätzt, wodurch die Gefahr einer Fehlausewahl aufgrund von zufälligen Entwicklungen in bestimmten Zeitperioden vermieden werden kann. Die Lag-Struktur der Modelle wurde für die einzelnen Schätzzeiträume anhand des Akaike Informationskriteriums opti-

Tab. 5
Rangfolge der Modelle über die Schätzzeiträume (Mittel der Horizonte)

Schätzzeiträume	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	Ø
X	6	7	6	6	5	4	3	1	5	5	6	8	5	6	6	3	10	5.41
M	8	10	10	5	2	3	2	7	7	8	7	1	6	8	8	2	6	5.88
U	3	4	2	9	4	9	4	3	2	1	4	3	1	9	7	5	7	4.53
G	2	1	5	8	9	8	7	6	3	2	2	7	2	7	5	7	9	5.29
Basis	4	6	4	7	8	7	5	4	6	4	3	6	3	4	9	9	5	5.53
RER	5	8	7	10	10	10	6	5	4	3	5	4	4	5	10	8	8	6.59
BRENT	9	3	9	4	3	2	10	8	8	6	1	2	8	1	2	10	1	5.12
CONS	7	2	3	3	6	5	9	10	9	10	8	9	9	3	3	4	2	6.00
I	1	9	1	2	7	6	8	9	10	9	9	5	10	2	4	6	3	5.94
AR	10	5	8	1	1	1	1	2	1	7	10	10	7	10	1	1	4	4.71

miert. Die Berücksichtigung verschiedener Schätzzeiträume führte dazu, dass auch die jeweiligen Prognosezeiträume alternierten. Die Qualität der Prognosen wurde anhand zweier Maße, dem RMSE, also der Wurzel des mittleren quadratischen Prognosefehlers, und dem Theil'schen Ungleichheitskoeffizienten evaluiert. Die Ergebnisse zeigten, dass die Prognosen des Basismodells in ersten Differenzen, welches zusätzlich durch die Exporte, die Importe sowie die Arbeitslosenquote erweitert wurde, durchschnittlich am besten abschneiden. Diese Modellvarianten wiesen im Vergleich der hier spezifizierten VAR-Modelle die niedrigsten RMSE aus. Der Vergleich mit einem univariaten AR(2)-Prozess für das reale BIP zeigte jedoch die Überlegenheit einer noch sparsameren Modellspezifikation. Dieses Ergebnis sollte allerdings mit der entsprechenden Vorsicht gesehen werden. Die Überprüfung der Robustheit der Prognosegüte hat ergeben, dass die Qualität der Prognosen der verschiedenen Modellvarianten mit dem zugrunde liegenden Schätzzeitraum stark variiert. Die Prognose anhand eines einzelnen Modells ist daher als problematisch zu sehen.

Für die Erstellung von Prognosen scheint es empfehlenswert zu sein, mehrere Modelle simultan zu betrachten, die sich in der Auswahl der Variablen sowie in der Festlegung der jeweiligen Lag-Struktur unterscheiden. Mit Hilfe eines solchen Model-Averaging-Ansatzes ließe sich dann eine Vorhersage generieren, die sich aus dem gewichteten Durchschnitt der einzelnen Modellprognosen herleiten lässt.¹⁰ Die Robustheit der Prognose kann dadurch verbessert werden.

Literatur

- Altissimo, F. und G.L. Violante (2001), »The non-linear dynamics of output and unemployment in the U.S.«, *Journal of Applied Econometrics* 16(4), 461–486.
- Bernanke, B.S., M. Gertler und M. Watson (1997), »Systematic Monetary Policy and the Effects of Oil Price Shocks«, *Brookings Papers on Economic Activity* (1), 91–142.
- Charemza, W.W. und D.F. Deadman (1992), *New directions in econometric practice*, Edward Elgar Publishing Limited, Cheltenham.
- Engle, R. und C. Granger (1987), »Co-integration and Error Correction: Representation, Estimation and Testing«, *Econometrica* 55, 251–276.
- Erdöl-/Energie-Informationssdienst Nr. 14/05, 4. April 2005.
- Granger, C. und P. Newbold (1987), »Spurious Regressions in Econometrics«, *Journal of Econometrics* 194(2), 111–120.
- Nierhaus, W. und J.-E. Sturm (2003), »Methoden der Konjunkturprognose«, *ifo Schnelldienst* 56(4), 7–23.
- Winker, P. (2002), *Finanzmarkt-Ökonometrie*, Schäffer-Poeschel, Stuttgart.

¹⁰ Die Gewichtung der einzelnen Prognosen könnte sich dabei beispielsweise an dem Kehrwert des RMSE orientieren.