

Inflationsprognosen mit vektorautoregressiven Modellen

Thomas J. Jordan, Leiter Ressort Forschung,
Schweizerische Nationalbank, Zürich

Peter Kugler, Professor, Wirtschaftswissenschaftliches Zentrum,
Universität Basel

Carlos Lenz, Assistenzprofessor, Wirtschaftswissenschaftliches Zentrum,
Universität Basel

Marcel R. Savioz, Wissenschaftlicher Berater, Ressort Forschung,
Schweizerische Nationalbank, Zürich

Inflationsprognosen nehmen im neuen Konzept der Schweizerischen Nationalbank (SNB) eine Schlüsselrolle ein. Zum einen stellen sie die Grundlage für die geldpolitischen Entscheidungen dar, zum anderen sind sie ein wichtiges Mittel der Kommunikation mit der Öffentlichkeit. Um die Transparenz ihrer Geldpolitik weiter zu verbessern, hat die Nationalbank beschlossen, verschiedene Studien zu publizieren, die darüber Aufschluss geben sollen, wie die Inflationsprognosen berechnet werden und wie sie die geldpolitischen Entscheidungen beeinflussen. In diesem Zusammenhang sind bereits die Artikel von Jordan und Peytrignet (2001) und von Stalder (2001) im Quartalsheft 2001–2 publiziert worden.

In Jordan und Peytrignet (2001) wird ausführlich diskutiert, weshalb die Nationalbank ihre Politik auf Inflationsprognosen abstützt, wie die Prognosen die geldpolitischen Entscheidungen beeinflussen und wie die publizierten Prognosen zu interpretieren sind. Die Studie macht unter anderem zwei Punkte deutlich. Erstens handelt es sich bei der publizierten Prognose um eine bedingte Prognose, die einen über dem Prognosezeitraum konstanten kurzfristigen Zinssatz unterstellt. Zweitens ist die publizierte Prognose eine Konsensprognose, die auf einer breiten Erfassung aller relevanten Informationen beruht. Dabei kommen insbesondere auch makroökonomische Modelle zum Einsatz.

Unter den Modellen, die die SNB zum Einsatz bringt, figurieren traditionelle ökonomische Modelle und sogenannte vektorautoregressive Modelle (VAR-Modelle). Die traditionellen ökonomischen Modelle gehen von einer expliziten Vorstellung über die Zusammenhänge der Wirtschaft aus und die Beziehungen zwischen den Variablen werden mit Hilfe von Verhaltensgleichungen modelliert. Bei den VAR-Modellen wird hingegen davon ausgegangen, dass die Kenntnisse über die wirtschaftlichen Zusammenhänge beschränkt sind. Aus diesem Grund wird den Modellen entweder keine oder nur eine minimale Struktur auferlegt. Im Unterschied zu den traditionellen Modellen, für die eine Unterscheidung der Variablen in exogene und endogene Variablen charakteristisch ist, werden in VAR-Modellen alle Variablen als endogen behandelt.

In Stalder (2001) wird das grösste von der SNB verwendete traditionelle Strukturmodell – ein nekeynesianisches Modell – detailliert beschrieben. Der vorliegende Aufsatz knüpft an die dort begonnene Darstellung der SNB-Prognosemodelle an und beschreibt die von der SNB entwickelten VAR-Modelle. Diese lassen sich in zwei Gruppen unterteilen: nicht-strukturelle und strukturelle VAR-Modelle.

Nicht-strukturelle VAR-Modelle sind dadurch gekennzeichnet, dass sie keine Struktur der wirtschaftlichen Zusammenhänge vorgeben. Insbesondere wird auch der Transmissionsmechanismus der geldpolitischen Impulse nicht explizit modelliert. Die Prognosen solcher Modelle widerspiegeln implizit die Wirkungen einer Geldpolitik, die dem historischen Durchschnittsverhalten der Zentralbank entspricht. Dabei wird, je nachdem ob das geldpolitische Instrument eine Modellvariable ist oder nicht, der Verlauf des geldpolitischen Instrumentes explizit oder implizit prognostiziert. Da die prognostizierte Geldpolitik dem historischen Durchschnittsverhalten der Notenbank entspricht und den Prognosen keine spezifischen Bedingungen für den Verlauf der Geldpolitik über den Prognosehorizont auferlegt werden, werden sie oft als unbedingte Prognosen bezeichnet.

Mit nicht-strukturellen VAR-Modellen ist es nicht möglich, Prognosen unter verschiedenen Annahmen für die Geldpolitik zu berechnen. Die Wirkungen einer vom historischen Durchschnittsverhalten der Notenbank abweichenden Politik können nicht simuliert werden, weil nicht-strukturelle VAR-Modelle den Transmissionsmechanismus nicht identifizieren. Diese Begrenzung kann mit der Überführung von nicht-strukturellen in strukturelle VAR-Modelle (SVAR-Modelle) überwunden werden. Ein SVAR-Modell modelliert und erfasst mit wenigen – nicht einschneidenden und nur auf allgemein akzeptierten ökonomischen Grundsätzen basierenden – Restriktionen die Übertragung eines monetären Impulses auf die Wirtschaft. Dadurch wird es möglich, Prognosen für unterschiedliche Verläufe der Geldpolitik zu berechnen. Diese Simulationen stellen sogenannte bedingte Prognosen dar. Sie prognostizieren die Variablen des Modells bedingt auf einen bestimmten Verlauf der Geldpolitik, der vom historischen Durchschnittsverhalten der Notenbank abweichen kann.

Sowohl bedingte als auch unbedingte Prognosen sind wichtige Bestandteile der geldpolitischen Analyse. In der Praxis bilden die unbedingten Prognosen die Ausgangslage der Diskussion. Die unbedingten Prognosen geben einen Eindruck über die Inflationsaussichten, wenn unterstellt wird, dass sich die Notenbank über den Prognosehorizont gemäss ihrem historischen Muster verhalten wird. Um die Auswirkungen geldpolitischer Alternativen zu analysieren und schlussendlich einen geldpolitischen Entscheid zu fällen, braucht die Notenbank indes bedingte Prognosen.

Der Aufsatz ist in vier Teile gegliedert. Im ersten Teil wird eine allgemeine Einführung in die VAR-Methodologie gegeben. Leser, die mit VAR-Modellen vertraut sind, können diesen Teil überspringen und direkt zum zweiten Teil übergehen, wo gezeigt wird, wie unbedingte Prognosen mit nicht-strukturellen VAR-Modellen erstellt werden und insbesondere auf das von der SNB gewählte Vorgehen eingegangen wird. Im dritten Teil werden, die strukturellen VAR-Modelle eingeführt und das von der SNB verwendete Modell erläutert. Anschliessend wird gezeigt, wie mit diesem Modell geldpolitische Simulationen und bedingte Prognosen gemacht werden können. Der vierte Teil enthält Schlussbemerkungen.

1. Eine kurze Einführung in die VAR-Methode

Dieser Abschnitt liefert einige Grundlagen zum Verständnis von VAR-Modellen. In Abschnitt 1.1 geben wir zunächst einen kurzen historischen Überblick und gehen anschliessend auf die für die Abgrenzung der VAR-Methode von traditionellen Modellansätzen zentrale Frage der Identifikation von Modellen ein. In Abschnitt 1.2 wenden wir uns der Grundstruktur von VAR-Modellen zu und geben eine allgemeine, einfache Darstellung dieser Modelle.

1.1 Der historische Hintergrund

Die ersten makroökonomischen Modelle sind bereits in der Vorkriegszeit durch Tinbergen erstellt worden.¹ Die methodologische Grundlage dieser Modelle wurde von Koopmans und seinen Mitarbeitern in den fünfziger Jahren geschaffen, als dieser bei der *Cowles Commission* arbeitete.² Daher wird der ökonomische Ansatz, der den traditionellen ökonomischen Modellen zugrunde liegt, oft als *Cowles Commission Approach* bezeichnet. Diese Modelle sind dynamische simultane Gleichungssysteme mit Verhaltensgleichungen und buchhalterischen Identitäten.

Der Cowles-Commission-Ansatz ist durch eine strikte Trennung der Aufgaben der ökonomischen Theorie und der Ökonometrie gekennzeichnet. Die Aufgabe der ökonomischen Theorie ist es, die Gleichungen des Modells zu spezifizieren, d. h. die Beziehungen zwischen den Variablen festzulegen. Insbesondere wird anhand der Theorie bestimmt, welche Variablen in welchen Gleichungen vorkommen und welche Variablen von diesen Gleichungen ausgeschlossen werden. Der Ökonometrie kommt die Aufgabe zu, die Parameter zu schätzen (allenfalls die funktionale Form der Gleichungen zu wählen) und den Anpassungsprozess (Verzögerungsstruktur) zu bestimmen.

Diese einfache Arbeitsteilung wurde Mitte der siebziger Jahre in Frage gestellt. Die traditionellen makroökonomischen Modelle vermochten damals die weltweite durch den ersten Ölpreisschock ausgelöste Rezession ungenügend zu prognostizieren. Dadurch wurde deutlich, dass – trotz jahrelanger theoretischer Forschung zur Verbesserung der Spezifikation – diese Modelle fehlspezifiziert waren. Die Methode, mit der traditionelle makroökonomische Modelle spezifiziert werden, wurde grundlegend kritisiert.³ Dadurch wurde mit Beginn der siebziger Jahre neuen ökonomischen Forschungsrichtungen der Weg geöffnet.⁴

1 Bereits 1936 präsentierte Tinbergen ein Modell mit 27 Gleichungen für die Niederlande und 1939 ein Modell mit 84 Gleichungen für die Vereinigten Staaten.

2 Siehe Koopmans und Hood (1953).

3 Siehe Sims (1980) und Liu (1960) für eine der ersten Kritiken.

4 Siehe Kirchgässner und Savioz (1997) für einen Überblick.

Eine Möglichkeit bestand darin, den Anspruch an die für die Spezifizierung der Modelle benötigte ökonomische Theorie zu reduzieren. Zeitreihenmodelle eigneten sich dazu besonders, da sie vergleichsweise wenig ökonomische Theorie benötigen. Die Forschung im Bereich der Zeitreihenökometrie erlebte dadurch einen bis heute anhaltenden Aufschwung. Bereits 1970 entwickelten Box und Jenkins die ARIMA-Zeitreihenmodelle. 1975 zeigten Granger und Newbold, dass kleine ARIMA-Zeitreihenmodelle häufig bessere Prognosen lieferten als die grossen und aufwendigen traditionellen ökonomischen Modelle der damaligen Zeit. In den achtziger Jahren sind dann entscheidende konzeptionelle Durchbrüche in der Zeitreihenökometrie erzielt worden, insbesondere mit den Arbeiten von Dickey und Fuller (1981) über integrierte Variablen, Engle und Granger (1987) über kointegrierte Variablen, Sims (1980) über multivariate Zeitreihenmodelle (VAR-Modelle) und Johansen (1988) über den Einbezug von Kointegrationsbeziehungen in multivariate Zeitreihenmodelle. Die Entwicklung dieser Methoden ist wesentlich durch Fragen der monetären Ökonomie und der Geldpolitik motiviert worden.

Die Kritik, die Sims (1980) an den traditionellen makroökonomischen Modellen übte, beschäftigt sich im Kern mit der Art und Weise, wie die Modelle identifiziert werden. Ein Modell muss – unabhängig vom gewählten ökonomischen Ansatz – *identifiziert* werden, damit seine Aussagen eine ökonomische Interpretation zulassen. Erst ein identifiziertes Modell erlaubt bedingte Prognosen. Im Ansatz der Cowles Commission erfolgt die Identifikation über so genannte *Ausschlussrestriktionen*. Mit einer Ausschlussrestriktion wird in einer Gleichung des Modells, die zur Bestimmung einer endogenen Variablen dient, der Koeffizient einer potenziellen Erklärungsvariable a priori auf Null gesetzt. Damit wird dieser Variablen jegliche Erklärungskraft zur Bestimmung der endogenen Variablen abgesprochen. Die Identifikation erfolgt somit gleichzeitig mit der Spezifikation des Modells, indem man auf Kenntnisse der ökonomischen Theorie zurückgreift.

Dieses Vorgehen wird traditionellerweise am Beispiel des Marktes für ein landwirtschaftliches Produkt, beispielsweise Kaffee, illustriert. Das Kaffeeangebot in den Vereinigten Staaten hängt von den Wetterverhältnissen in Brasilien ab. Für die Nachfrage nach Kaffee in den Vereinigten Staaten spielen die Wetterverhältnisse jedoch kaum eine Rolle. Dies erlaubt die Formulierung einer Ausschlussrestriktion, nämlich dass die Wetterverhältnisse nicht in der Gleichung

zur Bestimmung der Kaffeefachfrage erscheinen. Wenn die Wetterverhältnisse in Südamerika variieren, bewegt sich das Marktgleichgewicht entlang der Nachfragekurve nach Kaffee. Die Nachfragekurve nach Kaffee kann somit *identifiziert* und deren Parameter empirisch ermittelt werden. Dies ermöglicht bedingte Vorhersagen über die Auswirkungen einer wetterbedingten Angebotsverknappung auf den Kaffeepreis. Würde nicht nur das Angebot, sondern zusätzlich auch die Nachfrage nach Kaffee vom Wetter in Südamerika abhängen, so wären Preis- und Mengenbewegungen entlang der Nachfragekurve nicht mehr von jenen entlang der Angebotskurve zu unterscheiden. Die Zurückführung von Preis- und Mengenbewegungen auf Bewegungen entlang der Nachfragekurve – wofür man den Begriff Identifikation der Nachfragekurve verwendet – wäre nicht möglich. Die Steigung der Nachfragekurve wäre unbekannt und bedingte Prognosen wären unmöglich oder schwierig.⁵

Sims (1980) kritisierte die Identifikation nach dem Ansatz der Cowles Commission als *unglaublich*. Er argumentierte, dass die ökonomische Theorie kaum überzeugende Ausschlussrestriktionen postuliert. Die Theorie des allgemeinen Gleichgewichts stellt die Wirtschaft vielmehr als ein zusammenhängendes System dar, in dem alles mit allem verbunden ist. Als Beispiel lässt sich die Identifikation eines Systems von Nachfragegleichungen anführen. Eine einzelne Gleichung dürfte schwerlich identifizierbar sein, denn alle Nachfragegleichungen sollten theoretisch die Preise aller Güter als Erklärungsvariablen aufweisen. Tauchen in einem System von Nachfragegleichungen dennoch Ausschlussrestriktionen auf, so vermutet Sims, dass dabei mehr auf eine dem Ökonometriker eigene Version von psychologischer oder soziologischer Theorie als auf die ökonomische Theorie zurückgegriffen wird.⁶

Ein weiterer Grund für das Hinterfragen der Identifikation nach dem Cowles-Commission-Ansatz ist die Rolle der Erwartungen. Sims nimmt das Beispiel des Kaffeemarktes, um zu zeigen, dass bei rationalen Erwartungen die Ausschlussrestriktionen von traditionellen ökonomischen Modellen nicht zulässig und die Modelle daher nicht identifiziert sind. Ungünstige Wetterverhältnisse in Brasilien können Kaffeekäufe von Konsumenten und Händlern auslösen, weil diese Preiserhöhungen antizipieren. Genau wie die Angebotskurve hängt somit auch die Nachfragekurve – rationale Erwartungsbildung vorausgesetzt – vom Wetter in Brasilien ab. Die Preis- und Mengenbewegungen, die der Ökonometriker benutzt, um die

5 Vergleiche Sims (1980) für eine weiterführende Diskussion.

6 Vergleiche Sims (1980, S. 3).

Parameter der Nachfragekurve zu bestimmen, finden deswegen nicht entlang der Nachfragekurve statt. Die geschätzte Gleichung ist nur scheinbar eine Nachfragekurve nach Kaffee und kann nicht als *strukturell* bezeichnet werden. Sie ist in Wahrheit das Ergebnis des Zusammenwirkens von sich verschiebenden Nachfrage- und Angebotskurven und stellt eine *reduzierte Form* des Angebots- und Nachfragemodells dar. Bedingte Prognosen und Aussagen anhand der Koeffizienten einer auf diese Weise geschätzten Gleichung sind deshalb irreführend.

Der VAR-Ansatz ist von Sims (1980) als Antwort auf die oben diskutierte problematische Spezifikation und Identifikation der traditionellen ökonometrischen Modelle entwickelt worden. Sims zeigte auf, dass mit vektorautoregressiven Zeitreihenmodellen ein alternativer – weniger problematischer – Ansatz der Spezifikation und Identifikation möglich ist.

1.2 Grundzüge von VAR-Modellen

Ein VAR(p) Modell kann folgendermassen dargestellt werden:

$$(1) \quad y_t = D_1 y_{t-1} + D_2 y_{t-2} + \dots + D_p y_{t-p} + \varepsilon_t, \\ E(\varepsilon_t \varepsilon_t') = \Omega.$$

Dabei stellt y ein Vektor von n endogenen Variablen im Zeitpunkt t dar. Beispielsweise kann sich der Vektor aus der Inflationsrate, der Zinssatzveränderung, dem Wachstum der Geldmenge und anderen Variablen zusammensetzen

$$y_t = \begin{bmatrix} \text{Inflationsrate,} \\ \text{Zinssatzveränderung,} \\ \text{Geldmengenwachstum,} \\ \dots \end{bmatrix}.$$

Die Matrizen D_i , für $i=1, \dots, p$, enthalten die Koeffizienten des Modells. Die Ordnung des VAR-Modells, p , gibt an, wie viele verzögerte Vektoren y_{t-i} im Modell vorkommen. Der Vektor ε_t stellt die nicht-systematischen Einflüsse, auch Impulse oder Innovationen genannt, dar. Die Varianz-Kovarianz Matrix Ω gibt die Korrelationsstruktur der nicht-systematischen Einflüsse ε_t wieder.⁷

Im Fall eines bivariaten VAR-Modells der Ordnung 1, das wir zur Illustration mehrmals heranziehen werden, wird die Gleichung (1) zu:

$$y_t = D_1 y_{t-1} + \varepsilon_t \\ \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \\ = \begin{bmatrix} d_{11} y_{1t-1} + d_{12} y_{2t-1} + \varepsilon_{1t} \\ d_{21} y_{1t-1} + d_{22} y_{2t-1} + \varepsilon_{2t} \end{bmatrix}.$$

Der Vektor der endogenen Variablen ($y_t = (y_{1t} \ y_{2t})'$) wird durch das Modell (die Koeffizienten der Matrix D_1), die Ausgangsposition in der Vorperiode ($y_{t-1} = (y_{1t-1} \ y_{2t-1})'$) und die nicht-systematischen Einflüsse ($\varepsilon_t = (\varepsilon_{1t} \ \varepsilon_{2t})'$) erklärt. Jede Variable hängt von der vorausgegangenen Entwicklung sämtlicher Variablen ab. Simultane Beziehungen zwischen den Variablen, d.h. gegenseitige Beeinflussung innerhalb der gleichen Zeitperiode, werden durch die Kovarianzmatrix der Innovationen eingefangen:

$$\Omega = \begin{bmatrix} \text{var}(\varepsilon_{1t}) & \text{cov}(\varepsilon_{1t}, \varepsilon_{2t}) \\ \text{cov}(\varepsilon_{1t}, \varepsilon_{2t}) & \text{var}(\varepsilon_{2t}) \end{bmatrix} = \begin{bmatrix} \sigma_{12}^2 & \sigma_{12} \\ \sigma_{12} & \sigma_{22}^2 \end{bmatrix}.$$

Nehmen wir an, dass die erste Variable des Vektors y_t die Inflationsrate darstellt und die zweite den von der Geldpolitik kontrollierten Zinssatz. Wie sind

⁷ Siehe auch Stock und Watson (2001) für eine einfache Darstellung des VAR-Ansatzes. Um die Notation zu vereinfachen, werden die Variablen in Abweichungen zum Trend ausgedrückt. Aus demselben Grund wird bei der Notation nicht zwischen Parametern und geschätzten Parametern unterschieden.

in diesem Fall die Koeffizienten der Matrix D_1 und Ω zu interpretieren? Eine Innovation im Zinssatz wirkt sich nur dann unmittelbar auf die Inflation aus, wenn die Kovarianz mit der Innovation der Inflation verschieden von Null ist: $\text{cov}(\varepsilon_{1t}, \varepsilon_{2t}) \neq 0$. Hat die Inflation eine Eigendynamik, so ist der erste Koeffizient der ersten Zeile der Matrix D_1 verschieden von Null ($d_{11} \neq 0$). Die Trägheit der Inflationsrate ist durch die Grösse dieses Koeffizienten beschrieben. Weist der Koeffizient d_{11} einen Wert von nahezu Eins aus, so bedeutet dies, dass die Inflation hoch bleiben wird, wenn sie einmal hoch ist. Hat die Geldpolitik einen verzögerten Einfluss auf die Inflation, so muss der Koeffizient d_{12} verschieden von Null sein. Die Stärke der Auswirkung ist durch die Grösse des Koeffizienten gegeben. Die Verzögerung kann daher rühren, dass sich eine Änderung des Zinssatzes erst nach einer gewissen Zeit in den Preisen bemerkbar macht. Sie kann aber auch darauf zurückgeführt werden, dass die Geldpolitik im Zeitpunkt t den Zins in Erwartung der Inflation im Zeitpunkt $t+1$ setzt.

Die Beziehung zwischen Inflation und Zins ist in einer Volkswirtschaft zu komplex, als dass sie nur durch die vier Koeffizienten der Matrix D_1 erfasst werden könnte. Jedoch lassen sich die hoch komplexen dynamischen Muster, die in einer Volkswirtschaft zu beobachten sind, anhand von VAR-Modellen mit mehreren Variablen und höherer Ordnung p erfassen.

Der systematische Teil des VAR(p)-Modells (1) entspricht

$$(2) \quad \hat{y}_t = D_1 y_{t-1} + D_2 y_{t-2} + \dots + D_p y_{t-p}$$

Die Koeffizienten der Matrizen D_i $i=1,2,\dots,p$ in Gleichung (2) dürfen nicht als strukturelle Verhaltensparameter interpretiert werden, da sie komplexe Funktionen von zugrunde liegenden Verhaltenskoeffizienten sind. Die systematische Komponente \hat{y}_t des VAR(p)-Modells ist gleichzeitig die Prognose für die Variable y_t . Dabei wird nur Information verwendet, die im Zeitpunkt $t-1$ vorhanden ist, nämlich die Ausgangsposition y_{t-i} $i=1,\dots,p$ und die Matrizen D_i $i=1,2,\dots,p$. Der Vektor der nicht-systematischen Einflüsse ε_t entspricht somit dem Prognosefehler des Modells:

$$y_t - \hat{y}_t = \varepsilon_t$$

Der Vektor der nicht-systematischen Einflüsse ε_t umfasst die Information, die in der laufenden Periode t neu in die Variablen einfließt. Deswegen werden die Komponenten dieses Vektors *Innovationen* genannt. Im VAR-Ansatz rücken diese Innovationen in den Vordergrund der Analyse. Wenn die

zweite Variable des Vektors y_t wie oben im VAR(1)-Modell den Zinssatz darstellt, so repräsentiert der Innovationsterm der zweiten Gleichung von (1) die unprognostizierbaren Einflüsse auf den Zinssatz. Eine unerwartete Veränderung der Geldpolitik bildet einen solchen unprognostizierbaren Einfluss. Diese Veränderung kann zum Beispiel darauf zurückzuführen sein, dass die Geldpolitik auf Information basiert, die *innerhalb* des laufenden Quartals t anfällt, während der erwartete Zinssatz \hat{y}_{2t} auf Information der Vorperiode $t-1$ beruht. Eine unerwartete Veränderung des Zinssatzes kann auch auf eine unerwartete Veränderung der Inflation zurückzuführen sein.

Die Gleichung (1) zeigt, dass im Zeitpunkt t die endogenen Variablen y_t von ihrer Vergangenheit, y_{t-i} für $i=1,2,\dots,p$, und von der laufenden Innovation ε_t abhängen. Dies gilt auch für die Vorperiode und alle früheren Perioden. Somit lässt sich, wegen des rekursiven Charakters von Gleichung (1), der Vektor der endogenen Variablen ausschliesslich als eine Funktion der vergangenen Innovationen darstellen:⁸

$$(3) \quad y_t = \varepsilon_t + C_1 \varepsilon_{t-1} + C_2 \varepsilon_{t-2} + \dots$$

Gleichung (3) bildet den Kern des VAR-Ansatzes. Sie stellt die endogenen Variablen y als gewichteten Durchschnitt der heutigen und der früheren Innovationen dar und wird daher *Vector-Moving-Average*-Darstellung (VMA-Darstellung) des Modells genannt. Der Einfluss einer Innovation ε_{t-j} auf die endogenen Variablen y_t , i Perioden in der Zukunft, wird durch die Koeffizienten-Matrix C_i ausgedrückt. Die Matrizen C_i stellen die Reaktionen der Variablen y_t auf Innovationen dar und sind komplexe Funktionen der Matrizen D_i , $i=1,2,\dots,p$.

Die Matrizen C_i sind aber schwer zu interpretieren, wenn die Innovationen korreliert sind: Es macht kaum Sinn, die Reaktion einer Variablen auf eine bestimmte Innovation isoliert zu betrachten, wenn mit der Veränderung dieser einen Innovation gleichzeitig Veränderungen aller anderen korrelierten Innovationen zu erwarten sind. Zur Lösung dieses Problems hat Sims den einfachen und pragmatischen Vorschlag einer rekursiven Darstellung der Korrelation der Innovationen gemacht, der heute in VAR-Analysen routinemässig angewandt wird. Dabei wird die Innovation der ersten Variablen im VAR-System, ε_{1t} , als exogene Veränderung interpretiert. Die Innovation der zweiten Variablen, ε_{2t} , ergibt sich als Regression auf die Innovation der ersten Variablen und eine residuale exogene Veränderung. Wenn wir dieses rekursive Verfahren bis zur n -ten Innovation, ε_{nt} , anwenden, erhalten wir die folgende Darstellung

8 Vorausgesetzt wird die Stabilität des VAR-Modells.

der Innovationen als lineare rekursive Funktion von n nicht korrelierten Störvariablen, die in der Folge mit u bezeichnet sind:

$$(4) \quad \begin{aligned} \varepsilon_{1t} &= u_{1t}, \\ \varepsilon_{2t} &= a_{21}u_{1t} + u_{2t}, \\ &\dots \\ \varepsilon_{nt} &= a_{n1}u_{1t} + \dots + a_{n,n-1}u_{n-1t} + u_{nt}. \end{aligned}$$

Mit dem Ersetzen der Innovationen in Gleichung (3) durch die nicht-korrelierten Schocks aus Gleichung (4) erhalten wir eine direkt interpretierbare Darstellung. Diese gibt an, wie sich die autonomen Schocks u auf die Variablen des Modells auswirken (Impulse-Response-Funktion). Die untertrianguläre Matrix A_0 der Koeffizienten in (4) kann auch durch eine sogenannte Choleski-Zerlegung der Kovarianzmatrix Ω der Innovationen berechnet werden. Ein mit Gleichung (4) ergänztes VAR-Modell wird als *rekursives* VAR-Modell bezeichnet.

2 Baseline-Prognosen mit VAR-Modellen

In diesem Abschnitt wenden wir uns den Prognosen zu, die mit nicht-strukturellen VAR-Modellen berechnet werden können. Abschnitt 2.1 enthält einige Vorbemerkungen zur Erstellung von VAR-Prognosen und zur Rolle, die solche Prognosen für die Geldpolitik spielen können. Anschliessend wird in Abschnitt 2.2 die Spezifikation der von der SNB verwendeten nicht-strukturellen VAR-Modelle diskutiert. In Abschnitt 2.3 wird erklärt, wie die Prognosen verschiedener VAR-Modelle kombiniert werden können, um die Prognosen zu verbessern.

2.1 Einige Vorbemerkungen

Die Inflationsprognosen, die sich mit einem nicht-strukturellen VAR-Modell erstellen lassen, können am einfachsten anhand eines VAR(1)-Modells erläutert werden

$$\begin{aligned} y_t &= D_1 y_{t-1} + \varepsilon_t, \\ E[\varepsilon_t \varepsilon_t'] &= \Omega. \end{aligned}$$

Wird die Gegenwart mit T bezeichnet, so ist der Zustand der Wirtschaft im Zeitpunkt $T+1$ und $T+2$ durch

$$y_{T+1} = D_1 y_T + \varepsilon_{T+1}$$

und

$$\begin{aligned} y_{T+2} &= D_1 y_{T+1} + \varepsilon_{T+2} \\ &= D_1 (D_1 y_T + \varepsilon_{T+1}) + \varepsilon_{T+2} \\ &= D_1^2 y_T + D_1 \varepsilon_{T+1} + \varepsilon_{T+2} \end{aligned}$$

gegeben. Die zukünftigen Innovationen ε_{T+1} und ε_{T+2} sind im Zeitpunkt T , in dem die Prognose erstellt wird, unbekannt. Ihr Erwartungswert ist gleich Null:

$$E_T(\varepsilon_{T+1}) = E_T(\varepsilon_{T+2}) = 0.$$

Die Inflationsprognosen für die Zeitpunkte $T+1$ und $T+2$ können damit als

$$\begin{aligned} \hat{y}_{T+1} &= E_T(y_{T+1}) = D_1 y_T \\ \hat{y}_{T+2} &= E_T(y_{T+2}) = D_1^2 y_T \end{aligned}$$

geschrieben werden und die Prognose für den Zeitpunkt $T+j$ ist

$$(5) \quad \hat{y}_{T+j} = D_1^j y_T.$$

Wenn die Prognosen für geldpolitische Zwecke nützlich sein sollen, muss der Prognosehorizont län-

ger sein als die durchschnittliche Wirkungsverzögerung der geldpolitischen Instrumente. Die SNB geht davon aus, dass der grösste Teil der Wirkung eines geldpolitischen Schocks innerhalb eines Zeitraums von drei Jahren erfolgt. Der Prognosehorizont i beträgt deshalb 12 Quartale und die längste Prognose ist folglich

$$\hat{y}_{T+12} = D_1^{12} y_T$$

Bei den mit nicht-strukturellen VAR-Modellen berechneten Prognosen ist die Geldpolitik nicht auf einen bestimmten Kurs festgelegt, sondern entspricht dem historischen Durchschnittsverhalten der Zentralbank. Im Unterschied zu bedingten Prognosen, die einen bestimmten Pfad der Geldpolitik vorgeben, werden solche Baseline-Prognosen als unbedingte Prognosen bezeichnet.⁹ Bei genügend grossem Prognosehorizont i weisen die Prognosen von nicht-strukturellen VAR-Modellen, wie übrigens auch die unbedingten Prognosen von jedem richtig spezifizierten ökonomischen Modell, immer eine Konvergenz zum historischen Durchschnitt der Variablen auf (vgl. Gleichung (5)).¹⁰ Dennoch sind die Prognosen dieser Modelle für kurz- und mittelfristige Zeithorizonte von Interesse. Erstens bilden sie eine erste Grundlage für die Beurteilung des geldpolitischen Handlungsbedarfs. Die Prognosen \hat{y}_{T+i} zeigen die Entwicklung der Inflation auf, so wie sie aufgrund des Ausgangszustandes der Wirtschaft und dem durchschnittlichen geldpolitischen Verhalten der Zentralbank in der Vergangenheit zu erwarten ist. Unbedingte Prognosen geben also einen Hinweis darüber, ob das übliche Verhalten der Geldpolitik genügt, um die Inflation mittelfristig im gewünschten Bereich zu halten.

Zweitens sind die Prognosen von nicht-strukturellen VAR-Modellen atheoretisch. Die Aussagen über einen eventuellen Handlungsbedarf basieren somit nicht auf einem spezifischen theoretischen Modell und bestehen unabhängig von der Auffassung über die Struktur und die Funktionsweise einer Volkswirtschaft. Aufgrund ihres atheoretischen Charakters dienen diese Prognosen häufig als *Benchmark* für den Vergleich mit anderen Prognosen.

9 Dabei muss man sich aber bewusst sein, dass an sich jede Prognose in irgendeiner Art und Weise bedingt ist. Auch die Prognosen der nicht-strukturellen VAR-Modelle sind auf die vorhandene Information bedingt. Aus diesem Grund wird zunehmend der Begriff Baseline-Prognose jenem der unbedingten Prognose vorgezogen (vergleiche dazu Canova (1995, S. 100)).

10 Der Grund liegt darin, dass das Modell nur mit Daten eines geldpolitischen Regimes geschätzt wird, in dem die Inflation das oberste geldpolitische Ziel darstellt. Die historische durchschnittliche Inflation sollte ferner nicht stark vom gewünschten Niveau abweichen. Unter dieser Bedingung ist Stabilität gegeben und dies impliziert, dass $\lim_{j \rightarrow \infty} D_1^j = 0$ gilt.

2.2 Spezifikation der VAR-Modelle

Die Nationalbank verwendet für Baseline-Prognosen nicht ein einzelnes VAR-Modell, sondern bestimmt zunächst die VAR-Modelle, die in der jüngeren Vergangenheit die besten Prognosen geliefert haben, und berechnet anschliessend eine kombinierte VAR-Prognose.

Die Spezifikation eines VAR-Modells beruht auf vier Entscheidungen: (i) Wahl der Variablen, (ii) Bestimmung der Ordnung des VAR-Modells, (iii) Bestimmung von allfälligen deterministischen Komponenten und (iv) Behandlung der Trendeigenschaften der Variablen (Integrationsgrad und Kointegrationseigenschaften).

Die von der SNB verwendeten nicht-strukturellen VAR-Modelle sind von der Ordnung 4 bis 5. Diese Ordnung reicht aus, um die Dynamik der Variablen einzufangen und damit Modellgleichungen zu erhalten, die keine Autokorrelation der Residuen mehr aufweisen. Jede VAR-Gleichung enthält eine Konstante und saisonale Dummy-Variablen (bei nicht saisonbereinigten Daten). Darüber hinaus werden keine deterministischen Komponenten in die VAR-Modelle eingefügt. Weil mögliche gemeinsame Trends und Kointegrationsbeziehungen von integrierten Variablen in Modellen in ersten Differenzen verloren gehen, werden alle Modelle sowohl in Niveauform als auch in ersten Differenzen geschätzt.¹¹

Bei der Wahl der Variablen wird zunächst eine Gruppe von Variablen mit potenzieller Information über den zukünftigen Verlauf des Preisniveaus und der Inflation bestimmt. Aufgrund der Theorie des Transmissionsprozesses und empirischer Untersuchungen (siehe z.B. Jordan (1999)) gehören neben der Inflation zumindest 10 Variablen zu dieser Gruppe: das am Konsumentenpreisindex (KPI) gemessene Preisniveau, das reale Bruttoinlandprodukt (BIP), die Geldmengen M_1 , M_2 und M_3 , die inländischen Bankkredite, die handelsgewichteten nominellen und realen Wechselkursindizes, ein kurzfristiger Zinssatz (z.B. Dreimonats-Libor) und ein langfristiger Zinssatz (z.B. Bundesobligationenrendite) sowie die Differenz zwischen diesen beiden. Die Ergebnisse von Stationaritätstests zeigen, dass die ersten Differenzen dieser Variablen als stationär angesehen werden können.

Ein erster Eindruck vom Informationsgehalt dieser Variablen für den zukünftigen Verlauf der Inflation kann mit Granger-Kausalitätstests vermit-

11 Siehe dazu auch Sims, Stock und Watson (1990).

telt werden. Der Granger-Kausalitätstest zeigt, ob die Hinzunahme einer der zehn Variablen die Prognose der Inflation, die allein auf der vergangenen Inflationsentwicklung basiert, verbessert. Verbessert die betreffende Variable die Prognose, so besitzt sie Information über die zukünftige Inflation und wird als Granger-kausal für die Inflation bezeichnet. Die Ergebnisse der Granger-Kausalitätstests sind in Tabelle 2.1 zusammengefasst.

Die Ergebnisse zeigen, dass die meisten Variablen einen positiven Beitrag zur Prognose der Inflation leisten. Der Wechselkurs und die Zinsspanne allein sind nicht Granger-kausal für die Inflation. Die Erfahrung zeigt jedoch, dass der Wechselkurs und die Zinsspanne zusammen mit einer anderen Variable kombiniert die Inflationsprognosen verbessern können.

Jede Prognoserunde beginnt mit einer Voruntersuchung der Prognoseeigenschaften aller VAR-Modelle, die aus der Gruppe der informativen Variablen gebildet werden können (maximal 5 Variablen pro Modell). Diese Voruntersuchung erfolgt in zwei Schritten:

1. Prognosen für die Jahresinflation werden für Prognosehorizonte bis zu drei Jahren berechnet. Diese Prognosen beziehen sich auf die jüngere Vergangenheit. Es sind *Out-of-sample*-Prognosen, d.h. die Schätzung der VAR-Modelle basiert auf Daten, die allesamt vor dem Beginn des Prognosehorizontes liegen.
2. Die *Out-of-sample*-Prognosen werden anhand des Root Mean Squared Error (RMSE) evaluiert und die besten VAR-Modelle selektiert.¹²

Vier Regularitäten über die Prognosegüte von VAR-Modellen, die in Jordan (1999) nachgewiesen worden sind, haben sich seit der Implementierung des neuen geldpolitischen Konzepts Anfang 2000 bestätigt:

1. Der RMSE der prognostizierten Jahresinflationsrate nimmt mit dem Prognosehorizont zu. Mit saisonbereinigten Daten und für einen Prognosehorizont von einem Jahr weisen die besten Prognosen einen RMSE von unter 1,0 Prozentpunkten der Jahresinflationsrate auf. Für einen Prognosehorizont von drei Jahren beträgt der RMSE der besten VAR-Modelle weniger als 1,5 Prozentpunkte.
2. Für langfristige Prognosen liefern VAR-Modelle mit Niveauvariablen bessere Ergebnisse als VAR-Modelle mit Variablen in Differenzen.
3. Die besten Prognosen werden mit VAR-Modellen mit drei bis vier Variablen erzielt.¹³
4. Kreditaggregate, Zinsen und Geldaggregate sind die Variablen, die am häufigsten in den VAR-Modellen mit den besten langfristigen Prognosen vorkommen.

Aufgrund der Voruntersuchung werden die Modelle selektioniert, welche für die Prognosen verwendet werden. Neben der Selektion der besten VAR-Modelle dient die Voruntersuchung dazu, allfällige Strukturbrüche zu erkennen. Weil die Selektion in jedem Quartal neu stattfindet, sollten jeweils diejenigen Modelle gewählt werden, die nach einem allfälligen Strukturbruch die besten Prognoseeigenschaften aufweisen. Bisher wurde festgestellt, dass sich die Zusammensetzung der selektierten Gruppe der besten VAR-Modelle im Laufe der Zeit nur wenig ändert.

Granger-Kausalitätstest

Tabelle 2.1

| Variable | Lags = 4 | | Lags = 6 | |
|---|-------------|----------|-------------|----------|
| | F-Statistik | p-Wert | F-Statistik | p-Wert |
| Reales BIP | 2,12 | 0,085* | 1,87 | 0,096* |
| Geldmenge M ₁ | 3,42 | 0,012** | 2,08 | 0,064* |
| Geldmenge M ₂ | 3,58 | 0,009*** | 1,98 | 0,078* |
| Geldmenge M ₃ | 1,58 | 0,187 | 1,92 | 0,087* |
| Inländische Bankkredite | 3,51 | 0,010*** | 2,06 | 0,067* |
| Handelsgewichteter nomineller Wechselkurs | 1,57 | 0,189 | 1,47 | 0,199 |
| Handelsgewichteter realer Wechselkurs | 1,47 | 0,219 | 1,37 | 0,235 |
| Dreimonate-Libor | 2,46 | 0,051* | 2,46 | 0,030** |
| Bundesobligationenrendite 10 Jahre | 4,80 | 0,001*** | 3,41 | 0,005*** |
| Zinsspanne | 0,98 | 0,425 | 1,43 | 0,211 |

Verwendet wurden differenzierte Variablen mit einer Lag-Länge von 4 und 6. Die Nullhypothese lautet, dass keine Granger-Kausalität vorhanden ist. Das Verwerfen der Nullhypothese wird mit Sternen gekennzeichnet. Ein, zwei oder drei Sterne bedeuten, dass die Hypothese auf dem 10%-, 5%- bzw. 1%-Signifikanzniveau verworfen wird. Der Test erfolgte paarweise ohne Berücksichtigung allfälliger Kointegrationsbeziehungen.

12 Der RMSE ist die Wurzel des durchschnittlichen quadrierten Prognosefehlers. Die RMSE von *Out-of-sample*-Prognosen von VAR und traditionellen ökonomischen Modellen sind nicht ohne weiteres vergleichbar. VAR-Prognosen sind vollständig dynamisch. Deshalb ist der RMSE eines traditionellen ökonomischen

Modells nur dann mit demjenigen eines VAR-Modells vergleichbar, wenn die exogenen Variablen, die in das traditionelle ökonomische Modell eingehen, auch prognostiziert werden.

13 Die Prognosen von VAR-Modellen mit fünf Variablen schneiden schlechter ab als diejenigen von VAR-Modellen mit drei und vier Variablen. Es muss vermutet werden, dass die höhere Prognosegenauigkeit auf mangelnde Freiheitsgrade bei der Schätzung der Modelle zurückzuführen ist.

2.3 Kombination von Prognosen aus verschiedenen VAR-Modellen

Nach der Evaluation der VAR-Modelle werden die damit errechneten Prognosen zu einer einzelnen Prognose kombiniert. Nehmen wir zum Beispiel an, dass in der Voruntersuchung drei VAR-Modelle VAR1, VAR2 und VAR3 aufgrund ihrer Out-of-sample-Prognosen für den Prognosehorizont i selektiert wurden. Die Modelle werden mit Daten, die bis zum aktuellen Rand reichen, zunächst neu geschätzt. Anschliessend werden die Inflationsprognosen $\hat{\pi}_{VAR1, T+i}$, $\hat{\pi}_{VAR2, T+i}$ und $\hat{\pi}_{VAR3, T+i}$ für den Prognosehorizont i gerechnet. Die kombinierte Prognose entspricht dann dem gewichteten Durchschnitt dieser Prognosen:

$$\hat{\pi}_{T+i} = w_1 \hat{\pi}_{VAR1, T+i} + w_2 \hat{\pi}_{VAR2, T+i} + w_3 \hat{\pi}_{VAR3, T+i}$$

wobei w_1 , w_2 und w_3 die Gewichte sind, die den einzelnen Prognosen zugeordnet werden. Es existieren verschiedene Methoden, die Prognosen zu gewichten. In der einfachen Durchschnittsmethode wird jeder Prognose das gleiche Gewicht beigemessen, wobei sich die Gewichte auf Eins summieren. Die Gewichte können auch aufgrund der Performance der verschiedenen Modelle in der Out-of-sample-Prognose der Voruntersuchung bestimmt werden. In der Kleinstquadratmethode werden die Gewichte anhand einer Regression der realisierten Inflationsraten auf die unterschiedlichen VAR-Prognosen ermittelt. Die SNB verwendet zurzeit die Methode des einfachen Durchschnitts. Dieser Prozess wird für die Prognosehorizonte $i = 2, 4, 6, 8, 10, 12$ durchgeführt. Für jeden Prognosehorizont werden die Prognosen der jeweils besten VAR-Modelle kombiniert.

Der Präzisionsgewinn von kombinierten Prognosen resultiert aus einem Diversifikationseffekt. Mit Ausnahme des Falles perfekter positiver Korrelation zwischen den einzelnen Prognosefehlern weist eine als gewichteter Durchschnitt ermittelte Prognose eine Fehlervarianz auf, die kleiner ist als die durchschnittliche Fehlervarianz der einzelnen Prognosen. Dieser Diversifikationseffekt kommt nur bei Prognosen zum Tragen, bei denen der Erwartungswert des Fehlers null ist. Da den VAR-Modellen keine Restriktionen auferlegt worden sind, darf dies angenommen werden. Die kleinere Varianz des Prognosefehlers bedeutet auch, dass der RMSE der kombinierten Prognosen kleiner ist als der durchschnittliche RMSE der einzelnen Prognosen.¹⁴

Neben dem Präzisionsgewinn weisen kombinierte Prognosen weitere wichtige Vorteile auf:

1. Durch das Kombinieren von Prognosen wird eine *Aggregation der Information* erzielt. Während in VAR-Prognosen nur die Information von relativ wenigen Variablen berücksichtigt werden kann, können kombinierte Prognosen diese Beschränkung umgehen.
2. Mit dem Kombinieren von Prognosen stützt sich die Prognose nicht auf ein einziges VAR-Modell. Damit wird die Problematik der *Modellunsicherheit* reduziert.
3. Werden die Gewichte für das Kombinieren geschätzt – zum Beispiel anhand der Kleinstquadratmethode – geben die Gewichte Hinweise darauf, welche VAR-Prognosen Information enthalten, die nicht bereits in den anderen VAR-Prognosen vorhanden ist. Ferner zeigt eine Veränderung der Gewichte Strukturbrüche auf.

In Jordan und Savioz (2001) wird gezeigt, dass der RMSE von Inflationsprognosen mit Kombinationen ganz erheblich reduziert werden kann. Die Methode des einfachen Durchschnitts reduziert den RMSE für Prognosen mit einem einjährigen Horizont um mehr als 10%. Für Prognosehorizonte von zwei und drei Jahren erweist sich die Kleinstquadratmethode als geeigneter. Die Reduktion des RMSE beträgt wiederum etwa 10%. Diese Angaben beziehen sich auf einen Vergleich des durchschnittlichen RMSE der VAR-Prognosen mit dem durchschnittlichen RMSE der kombinierten VAR-Prognose. Werden hingegen die besten VAR-Prognosen mit den besten kombinierten VAR-Prognosen verglichen, ist der Effizienzgewinn nochmals grösser. Für Inflationsprognosen mit einem Horizont von zwei oder drei Jahren liegt der RMSE der besten mit der Kleinstquadratmethode gewichteten kombinierten VAR-Prognosen um 30% tiefer als der durchschnittliche RMSE der besten VAR-Modelle.

In Abbildung 2.1 werden Prognosen von nicht-strukturellen VAR-Modellen für die Periode vom zweiten Quartal 2000 bis zum vierten Quartal 2001 gezeigt. Dabei wurde nur Information verwendet, die im jeweiligen Zeitpunkt der Prognose vorhanden war. Die Prognosen wurden nach dem beschriebenen zweistufigen Verfahren ermittelt.

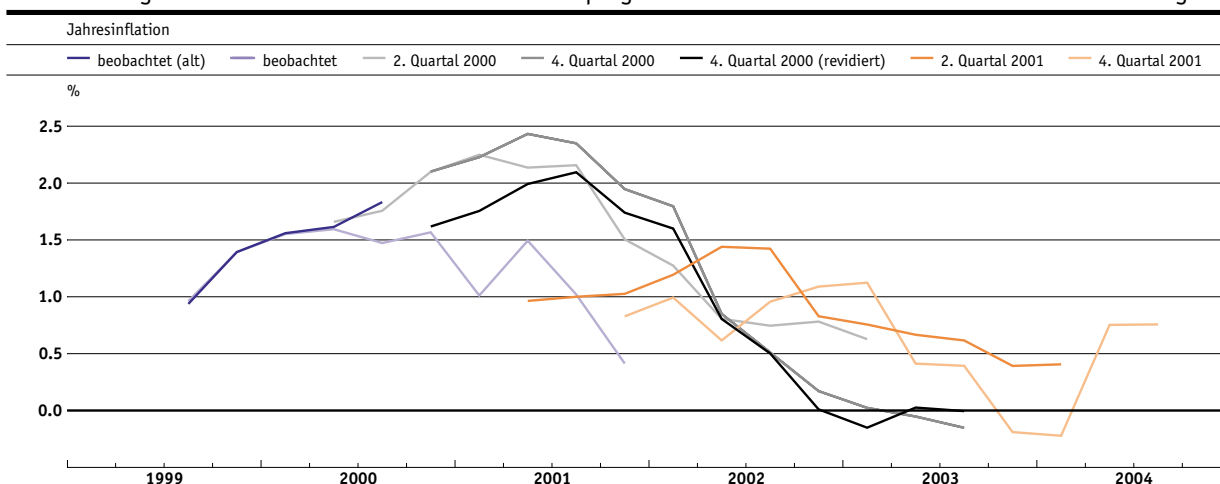
¹⁴ Der RMSE ist die Summe der Varianz des Prognosefehlers und der quadrierten Verzerrung der Prognose. Für unverzerrte Prognosen ist der RMSE folglich eine Funktion der Varianz des Prognosefehlers.

Die Abbildung gibt die am saisonbereinigten KPI gemessene Jahresinflation wieder. Im vierten Quartal 2000 wurde der KPI korrigiert, weshalb auch die Jahresinflation angegeben wird, wie sie vor dieser Korrektur ermittelt worden war. Aus der Abbildung 2.1 ist ersichtlich, dass die Prognosen frühzeitig eine starke Abflachung der Inflation aufgezeigt haben. Das prognostizierte Überschreiten der Inflation über zwei Prozent Anfang 2001 ist teilweise auf die Berechnung des KPI vor der Korrektur zurückzuführen. Dies wird deutlich, wenn die revidierte Prognose, die aufgrund des korrigierten KPI erstellt wurde, mit der Prognose verglichen wird, die vor der Korrektur gerechnet wurde. Insgesamt fällt auf, dass die VAR-Prognosen frühzeitig die Abschwächung der Inflation, die noch bis im zweiten Quartal 2001 am Verlauf des KPI nicht abzulesen war, aufgezeigt haben.

Wir haben bereits erwähnt, dass die Prognosen der nicht-strukturellen VAR-Modelle zu einer tiefen Inflation konvergieren werden, wenn die Schätzung anhand von Daten erfolgt, die einem auf die Erhaltung der Preisstabilität ausgerichteten geldpolitischen Regime entstammen. Die wichtigste Information dieser Prognosen ist deshalb nicht die Konvergenz zu einer niedrigen Inflationsrate, sondern die Dynamik der Anpassung der Inflation zum historischen Niveau. Die Prognosen in Abbildung 2.1 konvergieren innerhalb eines Zeithorizontes von drei Jahren etwas unterhalb von einem Prozent. Ein geldpolitischer Handlungsbedarf über die durchschnittliche Reaktion hinaus würde angezeigt, wenn die Inflationsprognosen innerhalb des Prognosehorizontes keine Tendenz zur Konvergenz im Bereich der Preisstabilität oder sogar eine explosive Tendenz aufweisen würden.

Baseline Prognosen von nicht-strukturellen Inflationsprognosen

Abbildung 2.1



3 Strukturelle VAR-Modelle (SVAR)

In diesem Kapitel betrachten wir die Frage, wie die zukünftigen Auswirkungen alternativer Gangarten der Geldpolitik im Rahmen von VAR-Modellen analysiert werden können. Das kann offensichtlich nicht anhand von Baseline-Prognosen geschehen, da in diesem Rahmen von einer dem historischen Durchschnitt entsprechenden Geldpolitik ausgegangen wird. Jedoch kann diese Fragestellung unter Einbezug von relativ wenig strukturellen Informationen über die Wirkungsweise der Geldpolitik anhand von sogenannten bedingten Prognosen in SVAR-Modellen angegangen werden. In den Abschnitten 3.1 und 3.2 wird der Ansatz in seinen Grundzügen dargestellt. Das von der SNB eingesetzte SVAR-Modell und die verschiedenen Arten von bedingten Prognosen, die mit diesem Modell ermittelt werden können, werden in den Abschnitten 3.3 und 3.4. behandelt.

3.1 Prognosen für geldpolitische Simulationen

Dieser Abschnitt befasst sich mit der Frage, wie VAR-Modelle bzw. SVAR-Modelle für geldpolitische Simulationen verwendet werden können. Konkret geht es darum, die Variablen in einem VAR-Modell unter der Annahme zu prognostizieren, dass ein von der Geldpolitik kontrolliertes Instrument einem gegebenen Pfad über die Zeit folgt oder eine Zielvariable zu einem zukünftigen Zeitpunkt einen vorgegebenen Erwartungswert erreicht. Solche sogenannten *bedingten* Prognosen können dazu benutzt werden, alternative geldpolitische Entscheidungen zu evaluieren, indem z.B. die Wirkungen der Festlegung des zukünftigen Zinssatzes auf die Prognosen von Inflation, Output- und Geldmengenwachstum untersucht werden.

Bei der Berechnung bedingter Prognosen mit VAR-Modellen ist es zentral, die gleichzeitige Korrelation der Variablen untereinander zu beachten. Dies wird im Folgenden anhand eines VAR(1)-Modells dargestellt. Die reduzierte Form dieses Modells entspricht

$$y_t = Dy_{t-1} + \varepsilon_t, \\ E(\varepsilon_t \varepsilon_t') = \Omega.$$

für $t=1,2,\dots,T$, wobei der Vektor $y_t = (y_{1t}, y_{2t})'$ die zu prognostizierenden Variablen enthält. Aus der Schätzung sind die Koeffizientenmatrix D und die Kova-

rianzmatrix der Residuen Ω bekannt. Beide Matrizen haben die Dimension 2×2 . Es sei darauf hingewiesen, dass die Innovationen ε_t im Allgemeinen untereinander korreliert sind, d. h. Ω ist nicht diagonal.

Die wesentlichen Elemente zur Berechnung der bedingten Prognose können anhand der 1-Schritt-Prognose illustriert werden. Ausgehend vom heutigen Zeitpunkt T lautet die unbedingte 1-Schritt-Prognose für y_t im Rahmen dieses Modells:

$$\hat{y}_{T+1} = Dy_T$$

Der Prognosefehler ist also ε_{t+1} und entspricht dem Vektor der Innovationen zum Zeitpunkt $T+1$. Bei der *unbedingten* Prognose mit VAR-Modellen wird ε_{t+1} (und allenfalls die Prognosefehler für längere Zeithorizonte) gleich seinem Erwartungswert von null gesetzt. Dabei wird davon ausgegangen, dass jenseits des Zeitpunkts T keine Information über die Innovationen verfügbar ist.

Wenn nun eine der Variablen in y_t durch die Geldpolitik kontrolliert werden kann, z. B. der Zinssatz y_{2t} , ist natürlich zu fragen, wie sich die Prognose der anderen Variable, also y_{1t} (z. B. die Inflationsrate) verändert, wenn der zukünftige Verlauf der geldpolitischen Variable fest vorgegeben wird. Die Strategie zur Beantwortung dieser Frage besteht darin, ausgehend vom Zeitpunkt T die Geldpolitik so zu verändern, dass die geldpolitische Variable $y_{2, T+1}$ den angestrebten Wert annimmt. Danach wird untersucht, wie dies die Prognose der anderen Variable beeinflusst. Solche bedingten Prognosen erlauben es dann, alternative geldpolitische Szenarien zu evaluieren.

Formal lässt sich diese Strategie am einfachsten anhand des Modells mit zwei Variablen für den Zeitpunkt $T+1$ illustrieren:

$$\begin{bmatrix} y_{1, T+1} \\ y_{2, T+1} \end{bmatrix} = \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix} \begin{bmatrix} y_{1, T} \\ y_{2, T} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1, T+1} \\ \varepsilon_{2, T+1} \end{bmatrix}.$$

Wenn $y_{2, t}$ die geldpolitische Variable ist, geht es also darum, $\varepsilon_{2, T+1}$ so zu wählen, dass $y_{2, T+1}$ bei gegebenen $y_{1, T}$ und $y_{2, T}$ den gewünschten Wert $y_{2, T+1}^*$ annimmt. Zu beachten ist nun allerdings, dass $\varepsilon_{1, T+1}$ nicht einfach gleich Null gesetzt werden darf. Da die Elemente von ε_t im Allgemeinen untereinander korreliert sind, muss die historische Korrelation zwischen den Innovationen berücksichtigt und $\varepsilon_{1, T+1}$ entsprechend festgelegt werden. Anders ausgedrückt bedeutet dies, dass die vergangene Information über die gleichzeitige Korrelation zwischen den beiden Elementen von y_t ausgenutzt werden kann, um die Prognose von $y_{1, T+1}$ zu verbessern, wenn $y_{2, T+1}$ bekannt ist.

Die Innovationen der geldpolitischen Variable $\varepsilon_{2,t}$ sind teilweise auf systematische Reaktionen der Geldpolitik auf unerwartete Veränderungen der anderen Variablen und teilweise auf geldpolitische Schocks zurückzuführen. Für die Berechnung einer bedingten Prognose sind aber nur die geldpolitischen Schocks massgebend, da zum Zeitpunkt der Simulation keine Information über unerwartete Veränderungen der anderen Variablen vorhanden ist. Der geldpolitische Schock wird nun derart bestimmt, dass die geldpolitische Variable den von der Notenbank gewünschten Wert annimmt. Dabei muss die aufgrund der vergangenen Werte aller Variablen erwartete Veränderung der geldpolitischen Variable berücksichtigt werden. Bei den bedingten Prognosen geht es also darum, direkt die Wirkungen der geldpolitischen Schocks im Prognosezeitraum zu erfassen. Die strukturelle Form des VAR-Modells, die der reduzierten Form zugrunde liegt, erfasst nun diese Wirkungen explizit. Es bietet sich deshalb an, die bedingten Prognosen direkt mittels der strukturellen Form des Modells zu berechnen.

Die strukturelle Form des VAR(1)-Modells kann allgemein als

$$B_0 y_t = B_1 y_{t-1} + u_t, \\ E(u_t u_t') = I$$

geschrieben werden, wobei $D = B_0^{-1} B_1$ und $\varepsilon_t = B_0^{-1} u_t$ gilt. Die Innovationen der strukturellen Form u_t , deren Varianz auf eins normalisiert ist, können nun als strukturelle Schocks interpretiert werden, weil sie untereinander unkorreliert sind. Die strukturellen Schocks können damit – formal aber nicht ökonomisch – eindeutig der Veränderung einer Variablen zugeordnet werden. Ihre Wirkung auf die jeweils andere Variable folgt aus den gleichzeitigen Interdependenzen zwischen den Variablen, die in der Matrix B_0 zum Ausdruck kommen.

Das System im Zeitpunkt $T+1$ kann wie folgt dargestellt werden:

$$y_{T+1} = D y_T + A_0 u_{T+1}, \\ A_0 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = B_0^{-1}.$$

Nun kann die oben beschriebene Strategie auf die Festlegung der strukturellen Schocks angewandt werden: $u_{2,T+1}$ wird so gewählt, dass $y_{2,T+1}$ den gewünschten Wert annimmt und $u_{1,T+1}$ wird gleich null gesetzt. Dies ist jetzt erlaubt, weil die strukturellen Schocks unkorreliert sind. So gesehen kann $u_{2,T+1}$ als geldpolitischer Schock interpretiert werden, der zum Ziel hat, $y_{2,T+1}$ auf den angestrebten

Wert $y_{2,T+1}^*$ zu bringen, und dessen Auswirkungen auf die andere Variable untersucht werden sollen. Der benötigte geldpolitische Schock $u_{2,T+1}$ entspricht

$$u_{2,T+1} = \frac{1}{a_{22}} (y_{2,T+1}^* - E_T y_{2,T+1}).$$

Die bedingte Prognose für $y_{1,T+1}$ lautet dann

$$\hat{y}_{1,T+1} = d_{11} y_{1,T} + d_{12} y_{2,T} + a_{12} u_{2,T+1}.$$

Der entscheidende Schritt für die Bestimmung der bedingten Prognose besteht darin, die strukturelle Form des VAR-Modells bzw. die Matrix B_0 bzw. A_0 zu finden und damit aus den Schocks der reduzierten Form die strukturellen Schocks zu bestimmen. Das dazu notwendige Vorgehen wird im nächsten Abschnitt beschrieben.

Zunächst sei darauf hingewiesen, dass das oben beschriebene Verfahren dynamisch angewendet werden kann, wenn man an einer bedingten Prognose über einen längeren Zeitraum interessiert ist. In diesem Fall ist der Zeitpfad der geldpolitischen Variable für die gewünschte Zahl von Perioden vorgegeben und es wird die bedingte Prognose für die andere Variable gesucht. Zu diesem Zweck wird die soeben berechnete 1-Schritt-Prognose für y_{T+1} verwendet. Anhand von

$$y_{T+2} = D \hat{y}_{T+1} + A_0 u_{T+2}$$

wird dann mit Hilfe des vorgegebenen Wertes $y_{2,T+2}^*$ und der bedingten Prognose für y_{T+1} der geldpolitische Schock $u_{2,T+2}$ berechnet und $u_{1,T+2}$ wiederum auf null gesetzt. Dies erlaubt die Erstellung der bedingten Prognose für $y_{1,T+2}$. Dieses Vorgehen kann wiederholt werden, bis der gewünschte Prognosehorizont erreicht ist.

3.2 Die Identifikation von strukturellen Schocks in VAR-Modellen

Im vorangehenden Abschnitt wurde auf die zentrale Bedeutung der strukturellen Form eines VAR-Modells für die Berechnung der bedingten Prognose hingewiesen. In diesem Abschnitt geht es nun darum, aus der gegebenen reduzierten Form eines VAR-Modells dessen strukturelle Form zu identifizieren und damit die Voraussetzung zu schaffen, um bedingte Prognosen auf einfache Weise zu berechnen.

Es wurde bereits darauf hingewiesen, dass nur die Innovationen von strukturellen VAR-Modellen als strukturelle Schocks interpretiert werden dürfen. Die Innovationen der reduzierten Form sind im Allgemeinen untereinander korreliert, was es unmöglich macht, ihnen eine strukturelle Interpretation zu geben. Nur unkorrelierte Schocks stellen eigenständige Quellen von Veränderungen dar, während korrelierte Schocks immer die Wirkung verschiedener Quellen enthalten. Wie aus den Ausführungen im vorangehenden Abschnitt ersichtlich ist, besteht das zentrale Problem bei der Identifikation eines strukturellen VAR-Modells darin, die Matrix B_0 so zu bestimmen, dass sie erlaubt, aus der (beobachtbaren) reduzierten Form die (unbeobachtbare) strukturelle Form zu berechnen. Die Matrix B_0 muss die Bedingung $B_0\Omega B_0' = I$ erfüllen, was sicherstellt, dass die Innovationen u_t der strukturellen Form untereinander unkorreliert sind. Weil die Matrix B_0 n^2 Elemente hat, die obige Bedingung aber wegen der Symmetrie der Kovarianzmatrix Ω nur $n(n+1)/2$ Restriktionen liefert, sind zusätzlich mindestens $n(n-1)/2$ Restriktionen nötig, um B_0 zu bestimmen. So sind zum Beispiel für ein VAR-Modell mit vier Variablen sechs Restriktionen notwendig. Diese zusätzlichen Restriktionen, die üblicherweise als identifizierende Restriktionen bezeichnet werden, sollten sinnvollerweise ökonomisch fundiert sein. Damit wird sichergestellt, dass die Innovationen der strukturellen Form tatsächlich eine strukturelle ökonomische Bedeutung haben.

Die traditionelle Form identifizierender Restriktionen, die wir schon in Abschnitt 1.2 angesprochen haben, beruht darauf, dass die Matrix B_0 bzw. A_0 untertriangulär ist (sogenanntes rekursives VAR-Modell). Dies bedeutet, dass die strukturellen Schocks über ihre zeitliche Wirkung auf die Variablen identifiziert werden: Die erste Variable in y_t wird gleichzeitig nur vom ersten Schock in u_t beeinflusst, die anderen Schocks wirken mit Verzögerung. Die zweite Variable wird gleichzeitig nur von den ersten beiden Schocks beeinflusst usw. Es kann z. B. unterstellt werden, dass

das Outputwachstum und die Inflation nur mit einer gewissen Verzögerung auf geldpolitische Schocks reagieren. Dieses zeitliche Restringieren der Schockwirkungen kann natürlich nur ökonomisch sinnvoll sein, wenn die Beobachtungsfrequenz der Variablen relativ hoch ist. Ausserdem ist es nicht immer möglich, ökonomisch sinnvolle Restriktionen zu finden, die eine untertrianguläre Struktur von B_0 implizieren.

Diesbezügliche Kritik an der traditionellen Form identifizierender Restriktionen in VAR-Modellen durch Cooley und LeRoy (1985) hat zu einer ganzen Reihe weiterer, über die Untertriangulärheit von B_0 hinausgehender Ansätze geführt, die einerseits die kurzfristige Wirkung einzelner Schocks und andererseits ihre langfristige Wirkung (oder Kombinationen von beidem) restringieren.¹⁵ In diesem Rahmen kann beispielsweise davon ausgegangen werden, dass der Zinssatz und die Geldmenge gleichzeitig vom geldpolitischen Schock und einem Geldnachfrageschock beeinflusst werden und diese beiden Schocks keinen laufenden Einfluss auf Inflation und Output ausüben. Daraus ergibt sich in einem VAR-Modell mit vier Variablen (Inflation, Outputwachstum, Geldmengenwachstum, Zinssatz) die folgende nichttrianguläre Struktur:

$$A_0 = \begin{bmatrix} x & x & 0 & 0 \\ x & x & 0 & 0 \\ x & x & x & x \\ x & x & x & x \end{bmatrix}$$

Mit diesen vier kurzfristigen Restriktionen ist das Modell noch nicht identifiziert, da wir mindestens sechs Beschränkungen brauchen. Dies kann durch zusätzliche langfristige Restriktionen erreicht werden. So ist es beispielsweise möglich zu postulieren, dass geldpolitische Schocks im Sinne der Neutralität des Geldes langfristig keine Wirkungen auf den Output haben. Bei langfristigen Restriktionen zur Identifikation von strukturellen VAR-Modellen ist zu beachten, dass sie nur auf Variablen angewandt werden können, die integriert von der Ordnung Eins (oder höher) sind. Dies bedeutet, dass es Schocks gibt, die eine permanente Wirkung auf diese Variablen haben. Wenn dies der Fall ist, kann aus der strukturellen Form für ein VAR(1)-Modell in ersten Differenzen,

$$B_0\Delta y_t = B_1\Delta y_{t-1} + u_t,$$

die Vektor-Moving-Average-Darstellung,

$$\Delta y_t = A_0 u_t + A_1 u_{t-1} + A_2 u_{t-2} + \dots,$$

bestimmt werden. Diese Darstellung, die auch als Impulse-Response-Funktion bezeichnet wird, reprä-

¹⁵ Bernanke (1986), Blanchard und Watson (1986) sowie Sims (1986) führten kurzfristige identifizierende Restriktionen ein, die nicht auf der Untertriangulärheit von B_0 basieren. Shapiro und Watson (1988), Blanchard und Quah (1989) sowie King et al. (1991) sind frühe Beispiele für

die Identifikation von Schocks mittels langfristiger Restriktionen. Galí (1992) kombinierte als erster kurz- und langfristige Restriktionen.

sentiert die dynamische Wirkung der gleichzeitigen und der vergangenen strukturellen Schocks auf die erste Differenz der Variablen. Diese Wirkungen kommen in den Matrizen A_0, A_1, A_2, \dots zum Ausdruck. Die Impulse-Response-Funktion für das Niveau der Variablen lautet dann:

$$y_t = y_0 + A_0 u_t + (A_0 + A_1) u_{t-1} + (A_0 + A_1 + A_2) u_{t-2} + \dots + (A_0 + A_1 + \dots + A_{t-1}) u_1.$$

Damit ist klar, dass die Matrix $A(1) = A_0 + A_1 + \dots + A_{t-1}$ die langfristigen Effekte der strukturellen Schocks auf das Niveau der Variablen repräsentiert. Die identifizierenden Restriktionen werden dadurch gewonnen, dass gewisse Elemente der Matrix $A(1)$ gleich null gesetzt werden. Da zwischen $A(1)$ und der Matrix B_0 bzw. A_0 eine direkte (wenn auch komplizierte) Beziehung besteht, lassen sich ökonomische a priori Vorstellungen über die Form von $A(1)$ auf B_0 übertragen und damit die strukturelle Form identifizieren.

Es sei hier auch darauf hingewiesen, dass es aus ökonomischen Gründen angebracht sein kann, mehr als die unbedingt notwendigen Restriktionen zu verwenden, um ein sogenannt überidentifiziertes strukturelles VAR-Modell zu definieren. Die Gültigkeit der überidentifizierenden Restriktionen kann statistisch getestet werden.

Das Ziel der Identifikation von strukturellen VAR-Modellen besteht darin, ökonomisch interpretierbare Schocks zu erkennen. Daher ist es nützlich, die Wirkungen der identifizierten Schocks zu untersuchen, um allfällige Widersprüche zwischen empirisch geschätzten Schockwirkungen und A-priori-Vorstellungen aus der ökonomischen Theorie zu erkennen und gegebenenfalls das Modell zu modifizieren. Das Instrument zur Evaluation von strukturellen VAR-Modellen ist die Impulse-Response-Funktion, welche die zeitliche Reaktion der Variablen auf die identifizierten Schocks abbildet. Die Untersuchung der Impulse-Response-Funktion liefert einerseits Erkenntnisse über die zeitliche Struktur von Schockwirkungen, was z. B. allfällige Verzögerungen bei der Wirkung der Geldpolitik aufzeigt. Andererseits müssen die identifizierten Schocks, wenn sie zur bedingten Prognose verwendet werden sollen, den theoretischen Vorstellungen bezüglich ihrer Wirkung entsprechen. Es wäre z. B. ökonomisch unglaubwürdig, wenn die zur Simulation der Geldpolitik verwendeten Schocks die Eigenschaft hätten, dass eine expansive Geldpolitik die Inflation langfristig reduzieren würde.

Die Bezeichnung strukturelle Vektorautoregression lässt möglicherweise den Eindruck entstehen, dass diese Modellklasse in einer engen Beziehung zu traditionellen makroökonomischen Strukturmodellen steht. Unsere obigen Ausführungen zeigen jedoch, dass das nur äusserst beschränkt der Fall ist. In strukturellen VAR-Modellen werden nur einige wenige ökonomisch begründete, die kurze oder lange Frist betreffende Restriktionen eingeführt. Ansonsten bleibt die Beziehung zwischen den Variablen offen und wird aus den Daten heraus modelliert, was als grosser Vorteil anzusehen ist. In traditionellen Strukturmodellen hingegen muss die Wirkung der Geldpolitik durch eine detaillierte Wiedergabe der dynamischen Übertragung von Veränderungen der Zinssätze und des Wechselkurses auf eine Vielzahl von Nachfragekomponenten und deren Deflatoren dargestellt werden (Transmissionsmechanismus). Das ist angesichts der vielen konkurrierenden Theorien über diesen Prozess sicher als Nachteil zu werten.

3.3 Das von der SNB eingesetzte SVAR-Modell

In diesem Abschnitt wird das von der SNB für die geldpolitischen Analysen eingesetzte SVAR-Modell vorgestellt. Das Modell umfasst vier Variablen, nämlich die am KPI gemessene Inflationsrate, die Wachstumsrate des realen BIP, die Wachstumsrate der Geldmenge M_1 und die Veränderung des Dreimonats-Libor. Der Vektor der im Modell berücksichtigten Variablen lässt sich somit als

$$y_t' = (\Delta \log p_t, \Delta \log y_t, \Delta \log m_t, \Delta r_t)$$

schreiben, wobei die Differenzen der Logarithmen den Veränderungsraten der Variablen entsprechen.

Wie können wir die Auswahl der Variablen begründen? Inflationsrate und Zinssatz müssen offensichtlich berücksichtigt werden, da sie als Endziel beziehungsweise als operatives Ziel in der geldpolitischen Strategie der SNB eine zentrale Rolle spielen. Die BIP-Wachstumsrate ist auch unabdingbar, da diese Grösse bei der Bestimmung der Gangart der Geldpolitik ebenfalls beachtet werden sollte. Die Geldmenge M_1 schliesslich wurde aufgenommen, da sich die monetären Operationen sehr schnell in Veränderungen der Zinsen und Geldmengen äussern. Mit anderen Worten bedeutet das, dass die Geldmenge wichtige Informationen zur Identifikation eines geldpolitischen Schocks liefert. Einheitswurzeltests zeigen, dass diese Variablen nicht-stationär sind und somit in das VAR-Modell in ersten Differenzen eingehen. Zudem kann die Hypothese keiner Kointegration zwischen den betrachteten Variablen nicht verworfen werden. Die letztere Annahme widerspricht der Vorstellung einer stabilen langfristigen Geldnachfragefunktion, d. h. einer stabilen Beziehung zwischen der Geldmenge, dem Preisniveau, dem Realeinkommen und dem Zinsniveau. Für die Schweiz und übrigens auch für viele andere Länder gilt aber, dass eine derartige langfristige Gleichgewichtsbeziehung nur für breiter definierte Geldaggregate wie M_3 , aber nicht für M_1 existiert. Eine instabile Nachfragefunktion für M_1 bedeutet, dass es Geldnachfrageschocks mit permanenter Wirkung auf die reale Geldhaltung gibt. Dieser Sachverhalt erlaubt es, die geldpolitischen Schocks besser zu identifizieren.

Viele Leser werden bei dieser Liste von Variablen sicherlich den Wechselkurs vermissen. Diese Grösse hat bei der Analyse der Geldpolitik in einer kleinen offenen Volkswirtschaft eine grosse Bedeutung. Sie wird hier jedoch aus zwei Gründen nicht direkt berücksichtigt. Erstens ist der Wechselkurs eine

sehr volatile Zeitreihe, die mit Strukturbrüchen (Einführung und Aufgabe eines Wechselkursziels Ende der siebziger Jahre, Einführung des Euro Ende der neunziger Jahre) behaftet und daher kaum angemessen in einem linearen VAR-Modell darzustellen ist. Zweitens reflektieren die Impulse-Response-Funktionen implizit auch die Wirkung von nicht direkt berücksichtigten Variablen auf die Inflation und das Realwachstum. Es ist gerade ein Vorteil des SVAR-Ansatzes, dass Zusammenhänge nicht a priori vorgegeben werden müssen, sondern sich aus der dynamischen Interaktion der Variablen ergeben. Zur empirischen Untermauerung dieser Sichtweise kann hier festgehalten werden, dass sich die dynamischen Reaktionen des Preisniveaus und des BIP auf einen geldpolitischen Schock in einem SVAR-Modell mit Wechselkurs nicht wesentlich von denjenigen in einem SVAR-Modell ohne Wechselkurs unterscheiden und im Modell mit fünf Variablen nur ungenauer geschätzt werden.

Ein paar Bemerkungen zur Wahl des Geldmengenaggregats sind notwendig. Die Giroguthaben bzw. die Reserven der Banken bei der SNB sind auf den ersten Blick der geeignetste Mengenindikator für geldpolitische Impulse. Diese Variable ist aber auch sehr volatil und mit Strukturbrüchen behaftet. Die Einführung des SIC und der neuen Liquiditätsvorschriften Anfang 1988 haben zu dramatischen Veränderungen der Gironachfrage geführt. Zudem hat der in den späten neunziger Jahren erfolgte Wechsel von den Giroguthaben zum Dreimonats-Libor als operationellem Ziel der Geldpolitik die Zeitreiheneigenschaften der Giroguthaben verändert. Ein weiterer Strukturbruch erfolgte mit der Einführung des Intraday-Repos, welcher den Bedarf der Banken an Giroguthaben weiter senkte. Die Zeitreiheneigenschaften der Giroguthaben ändern sich daher so stark über die Zeit, dass diese Reihe nicht für das Modell verwendet werden kann. Wir verwenden daher die Geldmenge M_1 , welche eine gute Approximation an die eigentliche Transaktionskassa darstellt.

Als nächstes müssen wir uns mit den Annahmen zur Identifikation des Modells und insbesondere des geldpolitischen Schocks, der für die Berechnung der bedingten Prognosen benötigt wird, beschäftigen. Hier muss vorausgeschickt werden, dass wir aus unseren vier Variablen vier Schocks identifizieren können: Ein Angebots- oder Produktivitätsschock, ein Preisschock, ein geldpolitischer Schock und schliesslich ein Nominalzinsschock. Dazu wurden kurz- und langfristige Restriktionen verwendet:

- Der geldpolitische Schock und der Nominalzinschock wirken sich im laufenden Quartal nicht auf Inflation und Realwachstum aus. Formal bedeutet das, dass die Matrix A_0 und damit auch ihre Inverse B_0 blocktriangulär sind (alle vier Elemente oben rechts sind null).
- Langfristig wird das reale BIP nur durch Angebotschocks beeinflusst, während der Nominalzinssatz in der langen Frist zusätzlich durch Nominalzinsschocks geprägt ist. Die Geldmenge M_1 und das Preisniveau werden auch langfristig durch alle vier Schocks beeinflusst. Dies impliziert, dass die Matrix der Summe der strukturellen Impulse-Response-Koeffizienten $A(1)$ an fünf Stellen den Wert null aufweist.

Diese Annahmen lassen sich einerseits durch die Existenz von kurzfristigen Rigiditäten und Friktionen begründen, die zu einer verzögerten Wirkung der Geldpolitik auf Produktion und Preise führen. Andererseits reflektieren sie die langfristige Neutralität des Geldes, gemäss der sich Geldangebotsstörungen langfristig nur im Preisniveau und der nominellen Geldmenge widerspiegeln.

Formal bedeuten diese Annahmen, dass unser Modell mit diesen neun Nullrestriktionen überidentifiziert ist: Bei vier Variablen sind gemäss den Ausführungen in Abschnitt 3.2 mindestens sechs Restriktionen notwendig, um eine Identifikation der strukturellen Schocks zu erreichen. Statistische Tests dieser überidentifizierenden Restriktionen zeigen, dass die Daten nicht gegen die getroffenen Annahmen sprechen. Es zeigt sich sogar, dass für die Matrix B_0 (und damit A_0) eine untertrianguläre Struktur mit den Daten vereinbar ist.

Die Schätzung dieses Modells mit Quartalsdaten der Periode 1974:2–2002:2 führt zu den in Abbildung 3.1 dargestellten Impulse-Response-Verläufen für einen geldpolitischen Schock. Dabei ist zu berücksichtigen, dass die kumulierten Effekte, d.h. die Effekte auf das Niveau der Variablen und nicht die Veränderungen, mit den dazugehörigen aus Bootstrap-Wiederholungen berechneten Vertrauensintervallen dargestellt sind. Ein geldpolitischer Schock in der Grösse einer Standardabweichung bewirkt einen sofortigen Anstieg der Geldmenge M_1 von 1,5% und eine sofortige Senkung des Dreimonats-Libors um knapp 0,5 Prozentpunkte (vgl. Panel oben rechts bzw. unten rechts). Dieser sogenannte Liquiditätseffekt vermindert sich wegen steigender Inflationserwartungen in den nachfolgenden vier Quartalen auf null. Anschliessend steigt der Zinssatz über das ursprüngliche Niveau (maximal rund 0,1 Prozentpunkte), bis nach etwa vier Jahren wieder das alte Zinsniveau

erreicht wird. Das reale BIP reagiert nach einer anfänglich negativen Reaktion verzögert positiv auf die expansive Geldpolitik, wobei nach acht Quartalen eine kumulierte Erhöhung der Wachstumsrate von 0,1 Prozentpunkten zu registrieren ist (vgl. Panel unten links). Im folgenden Jahr geht dieser Effekt praktisch wieder auf den Langfristwert von null zurück. Das Niveau der Konsumentenpreise reagiert nur stark verzögert auf die expansive Geldpolitik: Es dauert gut drei Jahre, bis sich die Wirkung des expansiven geldpolitischen Schocks vollständig auf das Preisniveau übertragen hat. Die Erhöhung des Preisniveaus beträgt rund 0,4 Prozent. Die rückläufige Wirkung bei den Konsumentenpreisen nach rund drei bis vier Quartalen ist auf die Verminderung der Mieten als Folge der Zinssenkung zurückzuführen (vgl. Panel oben links).

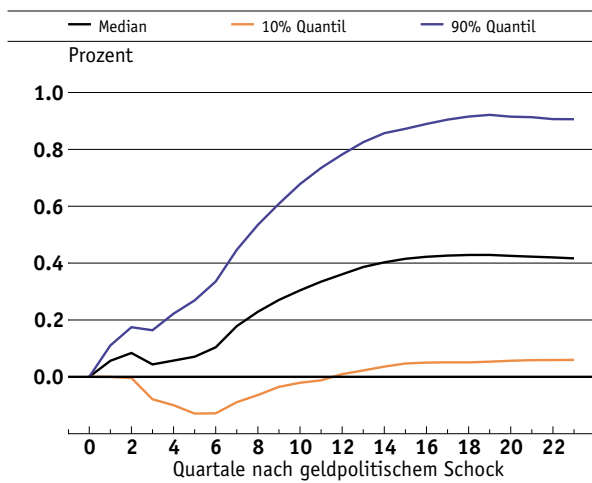
Die abgebildeten Wirkungen der geldpolitischen Schocks erscheinen weitgehend plausibel. Sie zeigen, dass die Zentralbank in der kurzen Frist nur Geldmenge und Zinsniveau beeinflussen kann. In der mittleren Frist (nach ein bis zwei Jahren) beeinflusst die Geldpolitik die Realwirtschaft und in der langen Frist (nach drei und mehr Jahren) bleibt nur noch der Effekt auf das Preisniveau. Auf den ersten Blick erstaunlich ist der im zweiten und dritten Quartal negative BIP-Effekt einer expansiven Geldpolitik. Neben der Ungenauigkeit der Schätzung (vgl. Vertrauensintervall) bietet sich als Erklärung der sogenannte J-Kurven-Effekt an: Die mit einer expansiven Geldpolitik einhergehende reale Abwertung des Franken führt kurzfristig zu einer Reduktion des realen Aussenbeitrags zum BIP, da die realen Exporte und Importe nur schwach auf die Abwertung reagieren.

3.4 Bedingte Inflationsprognosen

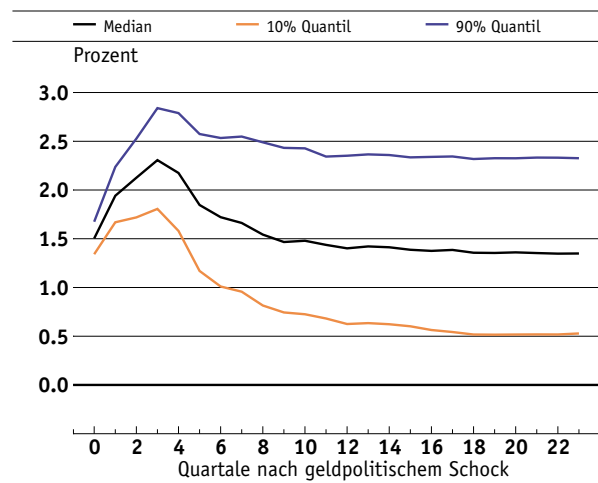
In diesem Abschnitt soll der Einsatz dieses Modells bei der Evaluation der Geldpolitik illustriert werden. Dabei gilt es zuerst der Lucas-Kritik an ökonomischen Politikevaluationen Rechnung zu tragen.¹⁶ Diese besagt, dass bei einer systematischen Änderung der Politik die in die Erwartungsbildung der privaten Wirtschaftssubjekte eingeht, die unter einem anderen Regime geschätzten ökonomischen Modelle nicht mehr gültig sind. Daher müssen wir uns die Frage stellen, ob seit 1974 eine systematische Änderung der schweizerischen Geldpolitik stattgefunden hat. Nach unserer Auffassung kann diese Frage negativ beantwortet werden. Die Politik der SNB war seit dem Übergang zu flexiblen Wechselkur-

16 Für eine ausführliche Diskussion der Relevanz der Lucas-Kritik bezüglich VAR-Modellen sei der Leser auf Leeper und Zha (1999) verwiesen.

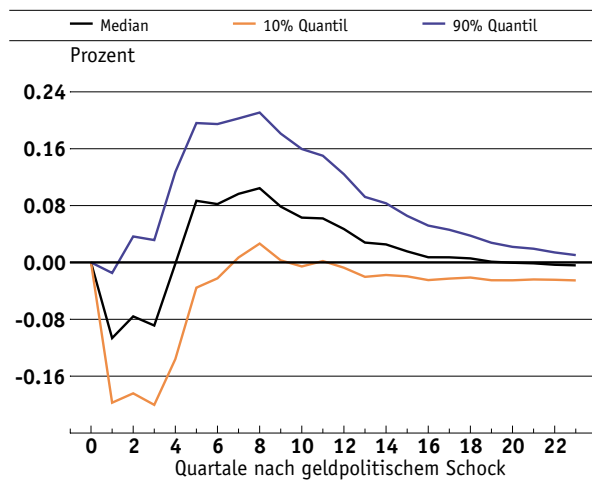
Reaktion Konsumentenpreise



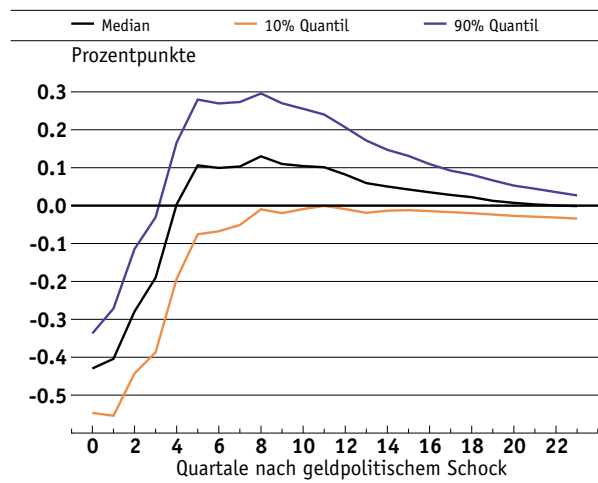
Reaktion Geldmenge M1



Reaktion BIP



Reaktion Dreimonate-Libor



sen auf das Ziel der Preisstabilität im Sinne einer tiefen Inflationsrate ausgerichtet. Zwar haben sich die Operationsprozeduren und die anvisierten Zwischenziele verändert, es kann jedoch davon ausgegangen werden, dass diese Änderungen zwar einen systematischen Einfluss auf die Zeitreiheigenschaften der Bankreserven und der ganz kurzfristigen Zinssätze (Tagesgeldsatz, Repo-Satz), nicht aber auf Inflation, M_1 -Wachstum, BIP-Wachstum und Dreimonats-Libor gehabt haben.

a) Bedingte Prognosen mit konstantem Zinssatz

Nach diesen Vorbemerkungen wollen wir zuerst die Konsequenzen einer Fixierung des Zinssatzes auf ein vorgegebenes Niveau r^* über die nachfolgenden drei Jahre betrachten. In einem ersten Schritt wird die unbedingte Prognose für die Periode $T+1$ mit dem SVAR-Modell berechnet. Der damit erhaltene Wert für den Zinssatz wird in der Regel von der Zielgrösse r^* abweichen. Daher wird der geldpolitische Schock für die Periode $T+1$ derart bestimmt, dass der Zinssatz auf das gewünschte Niveau zu liegen kommt. Formal lässt sich das mit dem laufenden Impulse-Response-Koeffizienten a_{43} des Zinssatzes (der vierten Variablen im System) auf den geldpolitischen Schock (dritter Schock im System) folgendermassen schreiben:

$$u_{3, T+1} = \frac{1}{a_{43}} (r^* - E_T r_{T+1}).$$

Die bedingte Prognose für $T+1$ lautet damit

$$\hat{y}_{T+1} = E_T y_{T+1} + a_{33} u_{3, T+1},$$

wobei der Vektor a_3 der dritten Spalte der Matrix A_0 entspricht. Im nächsten Schritt wird, ausgehend von der bedingten Prognose für $T+1$, anhand des SVAR-Modells eine unbedingte Prognose für $T+2$ berechnet und diese Prognose analog zum Vorgehen in Periode $T+1$ dem vorgegebenen Zinssatz angepasst. Dieses Verfahren wird bis zum Erreichen des Prognosehorizonts (in unserem Fall zwölf Quartale) wiederholt.

Zur Illustration dieses Verfahrens betrachten wir die gemäss dem Modell in der Periode 1995:4 erwarteten Auswirkungen verschiedener konstanter Zinssätze auf die Volkswirtschaft im Zeitraum 1996:1 bis 1998:4. Wir berechnen also die bedingte Prognose am Ende des Jahres 1995 unter der Annahme, dass die SNB den Zinssatz in den nachfolgenden drei Jahren konstant bei 1,5%, 2,0% und 2,5% hält. Abbildung 3.2 zeigt den mit diesen unterschiedlichen kurzfristigen Zinssätzen zu erwartenden Verlauf der Inflationsrate, der Wachstumsraten des realen BIP und der Geldmenge M_1 . Mit Ausnahme des Zinssatzes

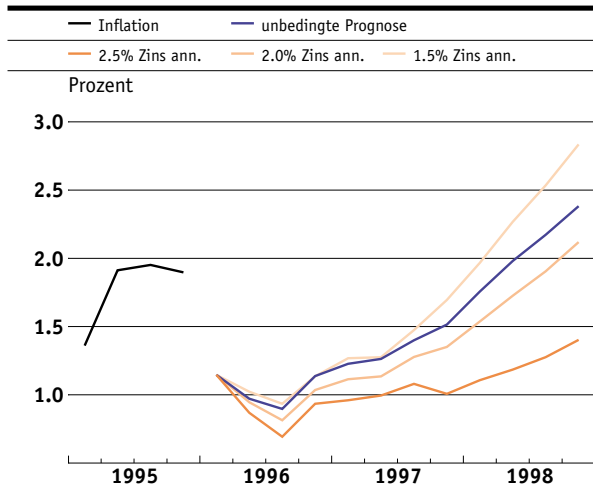
sind alle Variablen als Veränderungsraten gegenüber dem Vorjahresquartal definiert. Neben den bedingten Prognosen ist auch eine unbedingte Vorhersage abgebildet, bei der die Veränderung des Zinssatzes mitprognostiziert wird und keine geldpolitischen Schocks enthalten sind.

Aus der Abbildung 3.2 ist ersichtlich, dass die prognostizierte Inflation die 2%-Marke ab der zweiten Jahreshälfte 1998 übersteigt, wenn das Zinsniveau Anfang 1996 auf 1,5% gesetzt wird. Wenn das Zinsniveau unverändert bei 2% belassen wird, ist mit diesem Problem erst ab Ende 1998 zu rechnen. Um die Inflationsprognose für die drei Jahre immer deutlich unter 2% zu halten, ist eine restriktivere Gangart der Geldpolitik, d.h. eine Anhebung des Zinssatzes auf 2,5% nötig.

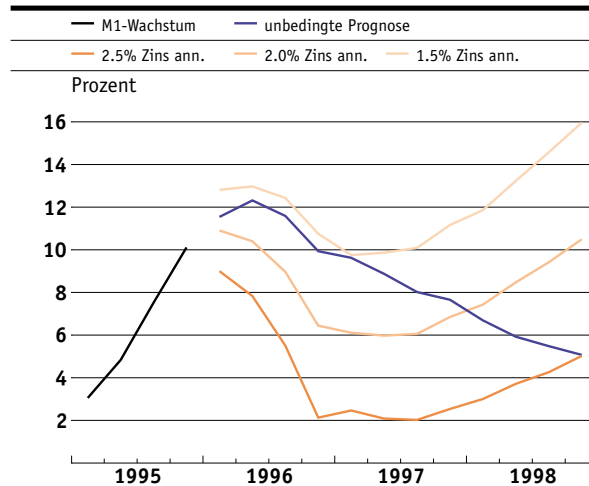
Die bedingte Prognose für die BIP-Wachstumsrate verhält sich spiegelbildlich zur Inflationsentwicklung: Die unter der Zinsannahme von 1,5% errechnete Vorhersage liegt um 0,2 Prozentpunkte höher als diejenige, die unter der Zinsannahme von 2,0% resultiert. Diese Differenz ist durch ein starkes Sinken des Realzinsniveaus bedingt. Wegen der verzögerten Wirkung auf Realwachstum und Inflation zeigen sich die vollen Konsequenzen dieser Politik erst ausserhalb des betrachteten Prognosehorizonts, d.h. in den Jahren 1999 und 2000.

Die Auswahl von bedingten Prognosen mit konstanten Zinssätzen erleichtert die Abschätzung des geldpolitischen Handlungsbedarfs. Die oben dargestellten Ergebnisse zeigen aber auch die Problematik der Fixierung des Nominalzinssatzes über drei Jahre bei sich verändernder Inflation auf. Es ist klar, dass die SNB den Zinssatz unter diesen Umständen nicht unverändert lassen würde. Um zu erreichen, dass die Inflationsrate im preisstabilen Bereich von 0%-2% bleibt, müssten die geldpolitischen Zügel sowohl bei der 1,5%- als auch bei der 2,0%-Politik in den letzten Quartalen der betrachteten Periode stärker angezogen werden.

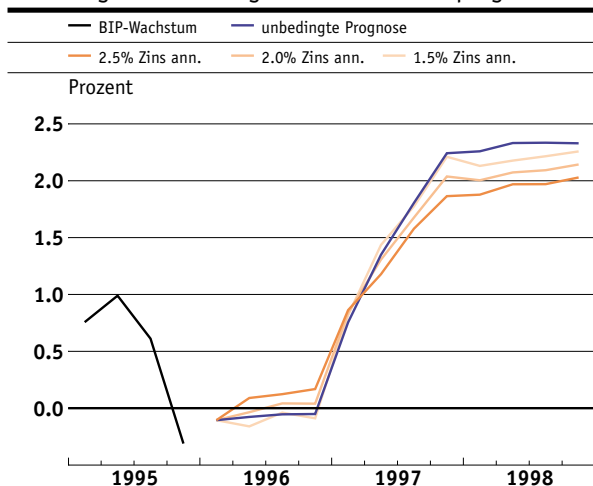
Unbedingte und bedingte Inflationsprognosen



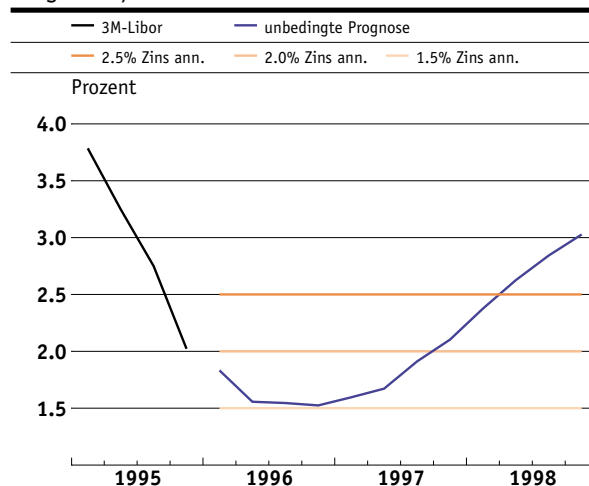
Unbedingte und bedingte M1-Wachstumsprognosen



Unbedingte und bedingte BIP-Wachstumsprognosen



Prognosen/Annahmen für den 3M-Libor



b) Bedingte Prognosen mit variablem Zinssatz bei rollender Zielinflation

Das SVAR-Modell kann auch verwendet werden, um die Konsequenzen für die Zinssetzung im Hinblick auf das Erreichen einer durchschnittlichen Inflationsrate von beispielsweise 2% über die nachfolgenden drei Jahre zu simulieren. Im Rahmen dieser Simulation gilt es die Geldpolitik so anzupassen, dass das prognostizierte Preisniveau im Durchschnitt der jeweils nachfolgenden drei Jahre um ungefähr 0,5% pro Quartal (2% im Jahr) steigt. Diese Bedingung kann analog zum Verfahren bei gegebenem konstanten Zinssatz in die SVAR-Prognose eingebaut werden.

Der Ausgangspunkt besteht wieder in einer unbedingten Prognose für das Preisniveau über die nachfolgenden zwölf Quartale. Wenn diese Prognose über dem durch eine Inflationsrate von 0,5% pro Quartal gegebenen Zielwert ($\pi^*=0,5\%$) liegt, muss die Geldpolitik angepasst werden. Der Betrag des nötigen Schocks lässt sich aus der in Abbildung 3.1 dargestellten und mit $AA_{13}(12)$ bezeichneten Reaktion des Preisniveaus (der ersten Variablen im System) nach zwölf Quartalen auf den geldpolitischen Schock (den dritten Schock im System) berechnen:

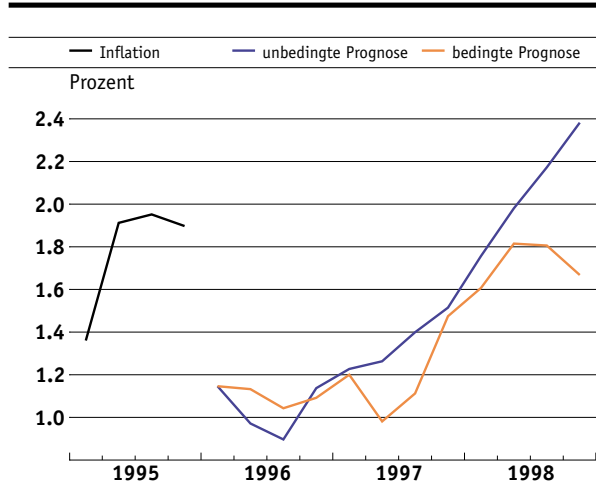
$$u_{3, T+1} = \frac{1}{AA_{13}(12)} ((\log p_{T+12} \pi^*) - E_T \log p_{T+12}).$$

Der geldpolitische Schock wird also derart berechnet, dass gemäss Modell in $T+12$ die in der Simulation angestrebte durchschnittliche Zielinflation von π^* pro Quartal erreicht wird. Anhand der so berechneten geldpolitischen Schocks wird nun genau gleich wie bei konstantem Zinssatz die Prognose für $T+1$ angepasst und das gesamte Verfahren sequentiell für $T+1$, $T+2$, ..., $T+12$ wiederholt.

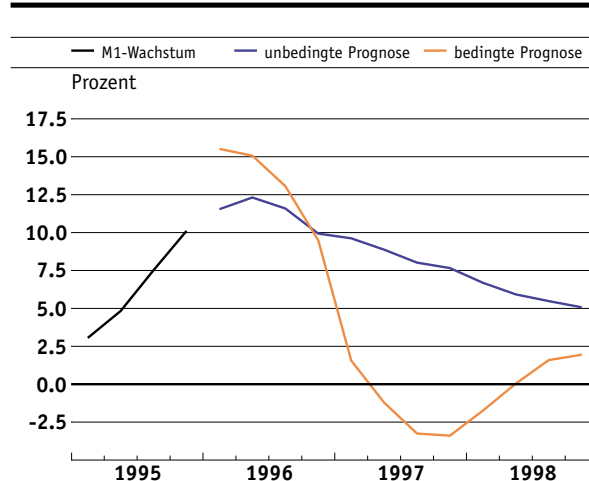
Zur Illustration dieses Verfahrens wollen wir wieder die bedingte Prognose für die Jahre 1996 bis 1998 betrachten. Die in Analogie zur Darstellung der bedingten Prognosen für den Fall mit konstanten Zinssätzen wiedergegebenen Ergebnisse finden sich in Abbildung 3.3. Der Zinssatz wird in den ersten Prognosequartalen deutlich unter 2% gesenkt, da gemäss der unbedingten Prognose die durchschnittliche Inflation über drei Jahre unter 2% pro Jahr liegt. Mit dem damit verbundenen Anstieg der für das Jahr 1997 zu erwartenden Dreijahresinflation wird die Geldpolitik restriktiv gestaltet (ein Anheben des Zinssatzes auf 3%). Damit wird der steigende Inflationstrend im letzten Jahr der Prognoseperiode gebrochen.

Bei der hier dargestellten Simulation wird davon ausgegangen, dass die Geldpolitik bei der Festlegung des geplanten geldpolitischen Schocks für die zwölf Quartale des Planungshorizonts immer die durchschnittliche Inflation der nächsten drei Jahre im Auge hat. Dies impliziert, dass die Geldpolitik in der Regel die angestrebte Inflation nicht genau trifft. Der beispielsweise für den Zeitpunkt $T+3$ bezüglich der durchschnittlichen Inflation der Quartale $T+3$ bis $T+14$ festgelegte Schock hat natürlich auch einen Einfluss auf die durchschnittliche Inflation von $T+1$ und $T+12$. Dies gilt auch für die Schocks in den nachfolgenden Zeitpunkten. Daher liegt in unserem Beispiel in Abbildung 3.3 die Inflationsrate für die drei explizit betrachteten Jahre immer unter 2%.

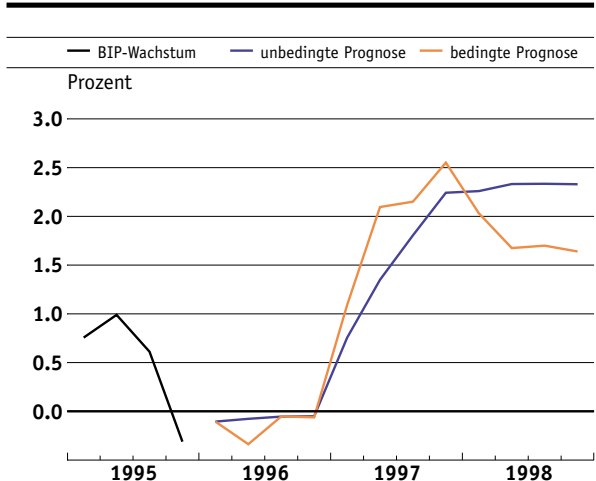
Unbedingte und bedingte Inflationsprognosen



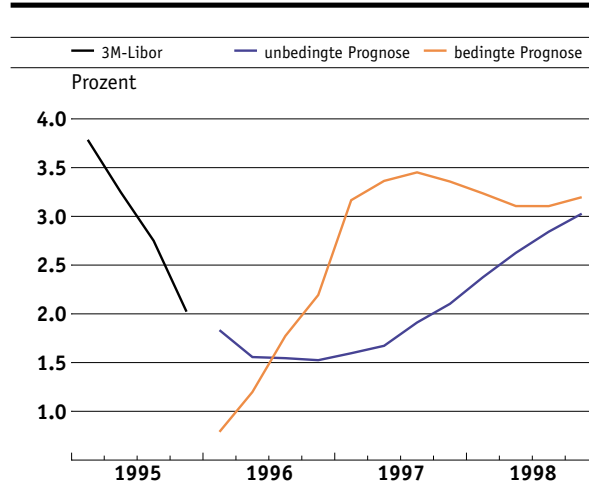
Unbedingte und bedingte M1-Wachstumsprognose



Unbedingte und bedingte BIP-Wachstumsprognose



Unbedingte und bedingte 3M-Libor-Prognose



c) Bedingte Prognosen mit variablem Zinssatz bei fester Zielinflation

Das SVAR-Modell kann auch auf den Fall angewendet werden, dass die im betrachteten Zeitraum von drei Jahren erwartete durchschnittliche Inflation exakt angepeilt wird. Dieser Fall hat allein illustrativen Charakter, da die SNB nicht eine Durchschnittsinflation über einen bestimmten Zeitraum anstrebt, sondern bemüht ist, die Inflation möglichst immer unter 2% zu halten.

Das Simulationsverfahren ist unter diesen Umständen leicht anders. Der Geldpolitik wird mehr Spielraum gegeben, da zur Erreichung des Ziels die geldpolitischen Schocks über den ganzen Prognosehorizont eingesetzt werden können. Als Ausgangspunkt dient wieder die Differenz zwischen dem erwarteten Preisniveau in der Periode $T+12$ und dem angestrebten Zielpreisniveau bei einer durchschnittlichen Inflationsrate von π^* ,

$$d = (\log p_{T+12} \pi^*) - E_T \log p_{T+12}.$$

Es gilt nun die Folge der geldpolitischen Schocks von $T+1$ bis $T+12$ derart zu bestimmen, dass das erwartete Preisniveau in der Periode $T+12$ dem Zielpreisniveau der Simulation entspricht. Dies soll dadurch geschehen, dass die Summe der quadrierten Schocks minimal wird. Mathematisch entspricht das der folgenden Optimierung unter einer Nebenbedingung:

$$\sum_{i=1}^{12} u_{3, T+i}^2 \rightarrow \min,$$

wobei

$$\sum_{i=1}^{12} AA_{13}(12-i) u_{3, T+i} = d.$$

Die Nebenbedingung bedeutet, dass die Summe der mit den relevanten Impulse-Response-Koeffizienten gewichteten Summe der geldpolitischen Schocks gerade der Differenz zwischen dem unbedingt prognostizierten Preisniveau und dem Zielpreisniveau entspricht. Die Lösung dieser Optimierungsaufgabe lässt sich leicht als

$$u_{3, T+i} = \frac{AA_{13}(12-i)}{\sum_{j=0}^{11} [AA_{13}(j)]^2} d$$

ermitteln. Da die Reaktion des Preisniveaus auf den geldpolitischen Schock mit der Zeit tendenziell zunimmt (vgl. Abbildung 3.1), werden die optimalen geldpolitischen Schocks über den Prognosehorizont tendenziell kleiner.

Abbildung 3.4 illustriert dieses Verfahren für die Periode 1996:1 bis 1998:4. Es zeigt sich, dass die unbedingte Prognose nur unwesentlich modifiziert werden muss, um die durchschnittliche Zielinflation der Simulation von exakt 2% zu erreichen. Da die Inflation in den ersten Prognosequartalen sehr tief ist und nahe bei 1% liegt, wird die Geldpolitik relativ zur unbedingten Prognose in den ersten acht Quartalen gelockert, was mit einem Anstieg der erwarteten Inflationsrate auf über 2,5% verbunden ist. Dieses Ergebnis folgt hier zwangsläufig aus der Ausgangslage mit einer Inflationsrate deutlich unter 2%, die einen Anstieg der Inflationsrate zu der in der Simulation gesetzten Zielerreichung nötig macht. Natürlich würde das Eintreten des prognostizierten Inflationsverlaufs im Verlauf dieser Jahre zu einer zunehmend restriktiveren Geldpolitik führen.

d) Business as Usual oder Modest Policy Shocks

Wie schon einleitend festgehalten wurde, ist bei bedingten Prognosen zu berücksichtigen, dass keine systematische Änderung der Geldpolitik simuliert werden darf. Bezogen auf unsere bedingten Prognosen bedeutet dies, dass die betrachteten geldpolitischen Schocks nicht systematischer Natur sein dürfen und mit der historischen Erfahrung kompatibel sein müssen. Leeper und Zha (1999) bezeichnen Schocks als *modest* und damit als zulässig, wenn diese im Rahmen des *business as usual* der Zentralbank immer wieder auftreten.

Diese Anforderung lässt sich statistisch überprüfen. Erstens können wir das Mittel der simulierten Schocks über die 12 Prognoseperioden ausrechnen,

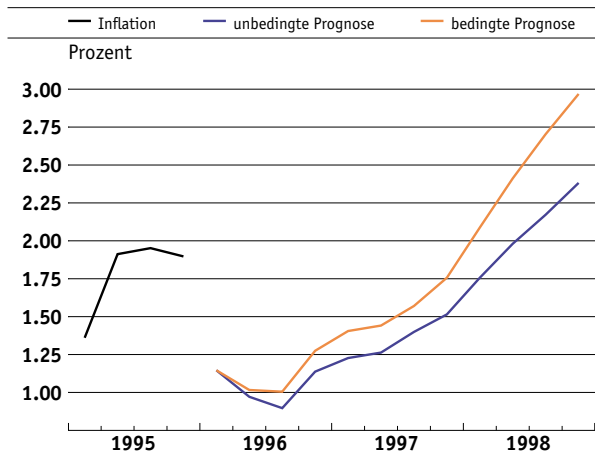
$$\eta(12, T) = \sum_{i=1}^{12} u_{3, T+i} / 12,$$

und überprüfen, ob diese Statistik signifikant von null abweicht. Diese Grösse sollte unter gewissen Annahmen mit Erwartungswert null und Varianz $1/T$ asymptotisch normal verteilt sein. Damit können wir prüfen, ob die simulierten geldpolitischen Schocks vom Vorzeichen her systematisch verzerrt sind. Zweitens können wir durch Summieren der quadrierten Werte überprüfen, ob der Betrag der simulierten Schocks im historischen Mittel liegt. Unter der Nullhypothese zufälliger Schocks mit Varianz von 1 ist

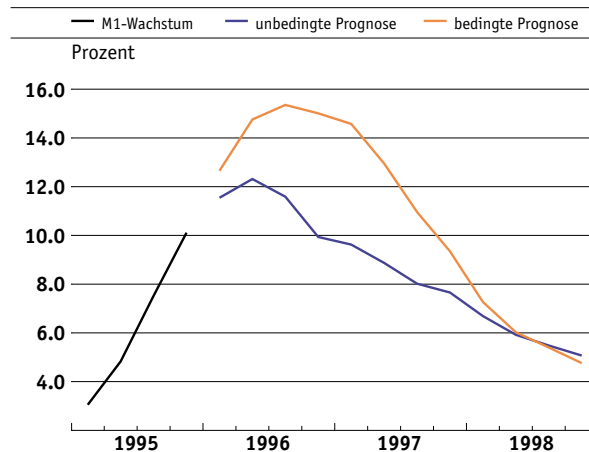
$$Q(12, T) = \sum_{i=1}^{12} u_{3, T+i}^2$$

asymptotisch χ^2 verteilt mit 12 Freiheitsgraden. Wird die Nullhypothese verworfen, bedeutet dies, dass der Betrag der simulierten Schocks zu gross ist. Im vorliegenden Zusammenhang können diese beiden Hypothesen in den beiden oben betrachteten Fällen nicht verworfen werden.

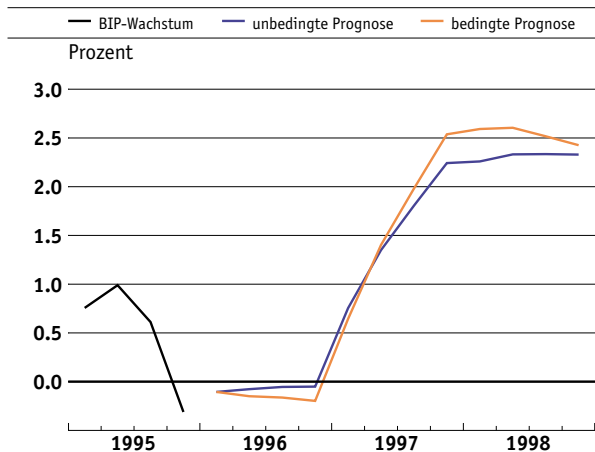
Unbedingte und bedingte Inflationsprognose



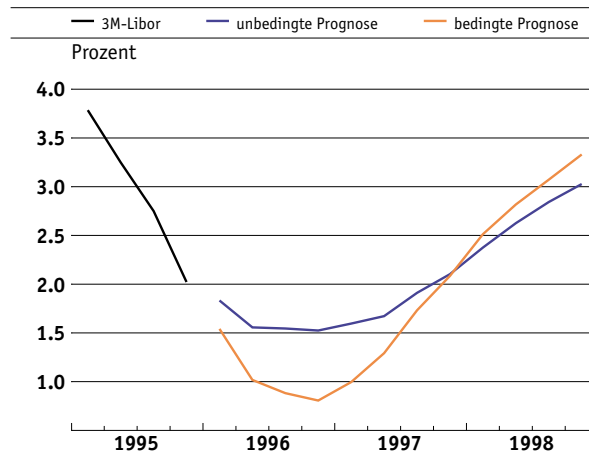
Unbedingte und bedingte M1-Wachstumsprognose



Unbedingte und bedingte BIP-Wachstumsprognose



Unbedingte und bedingte 3M-Libor-Prognose



4. Schlussbemerkungen

In diesem Abschnitt werden einige Grenzen von VAR-Modellen dargestellt und aufgezeigt, wie diese Grenzen allenfalls zu überwinden sind. Es wird dabei unterschieden zwischen Problemen, welche die VAR-Modellierung im Allgemeinen betreffen und Problemen, die sich spezifisch auf die Identifikation struktureller VAR-Modelle beziehen. Generell sind VAR-Modelle durch zwei wesentliche Eigenschaften beschränkt: Erstens sind sie linear und zweitens ist die Zahl der Variablen, die in das Modell eingehen, relativ klein. Diese Eigenschaften haben natürlich einen Einfluss auf die mit VAR-Modellen erzielten Resultate. Bei bedingten Prognosen mit strukturellen VAR-Modellen kommt hinzu, dass auch die Annahmen zur Identifikation der Modelle die Resultate beeinflussen.

Durch die Linearität von VAR-Modellen werden mögliche nichtlineare Beziehungen zwischen den Variablen sowie nichtlineare zeitliche Abhängigkeiten der Variablen nicht erfasst. Insbesondere der zweite Punkt hat in der empirischen Ökonomie in den letzten Jahren eine grosse Beachtung erfahren. Es hat sich herausgestellt, dass viele ökonomische Variablen dadurch gekennzeichnet sind, dass sich ihre Volatilität im Laufe der Zeit verändert. Die Volatilität hat zudem eine spezielle zeitliche Struktur, indem sie von ihrer eigenen Vergangenheit abhängt. Sogenannte GARCH-Modelle, die diese Eigenschaft der Variablen abbilden, haben sich insbesondere bei der univariaten Modellierung von Wechselkursen und Finanzmarktdaten als erfolgreich erwiesen. Die Kombination von VAR- mit GARCH-Modellen stellt eine Möglichkeit dar, nichtlineare Abhängigkeiten in einem multivariaten Kontext zu erfassen. Allerdings sind für die Schätzung solcher Modelle typischerweise sehr viele Beobachtungen nötig, weshalb sie sich in der Anwendung auf Finanzmarktdaten beschränken. Die relativ kleine Frequenz, mit der makroökonomische Daten verfügbar sind, stellt denn auch ein generelles Hindernis bei der Anwendung nichtlinearer Modelle dar.

Normalerweise enthalten VAR-Modelle eine relativ kleine Anzahl von Variablen. Üblich sind zwei bis sechs Variablen, Modelle mit bis zu zehn Variablen sind schon sehr selten. Diese Beschränkung kann sich auf die Prognosequalität der Modelle auswirken: Veränderungen von Variablen, die zwar wichtig, aber nicht im Modell enthalten sind, führen unter Umständen zu grossen Prognosefehlern. Diese Fehlerquelle kann bis zu einem gewissen Grad dadurch eliminiert

werden, dass die Zahl der Variablen, die in das Modell eingehen, vergrössert wird. Allerdings ist dabei zu beachten, dass die Zahl der zu schätzenden Koeffizienten mit dem Quadrat der Anzahl der Variablen zunimmt und bei einem VAR mit sechs Variablen und vier Lags bereits 150 beträgt. Das dabei entstehende Problem wird besonders klar, wenn man sich vor Augen hält, dass bei Quartalsdaten über einen Zeitraum von 40 Jahren nur 160 Beobachtungen zur Verfügung stehen.

Eine Möglichkeit, diesem Problem zu begegnen, besteht darin, sogenannte Bayesianische VAR-Modelle zu schätzen. Bei dieser Methode, die von Litterman (1980) vorgeschlagen wurde, werden bezüglich der Beziehungen zwischen den Variablen und bezüglich der zeitlichen Abhängigkeit der Variablen gewisse Annahmen getroffen. Typischerweise werden die Modelle so beschränkt, dass erstens jede Variable stärker von ihrer eigenen Vergangenheit abhängt als von derjenigen der anderen Variablen. Zweitens wird generell ein abnehmender Einfluss von zeitlich weiter zurückliegenden Beobachtungen unterstellt. Mit solchen Annahmen kann die Zahl der zu schätzenden Koeffizienten massiv reduziert werden. Dadurch wird es möglich, Modelle mit mehr Variablen zu schätzen. Zudem hat sich herausgestellt, dass mit Bayesianischen VAR-Modellen die Qualität unbedingter Prognosen verbessert werden kann.

Neben diesen generellen Vorbehalten, die gegenüber VAR-Modellen vorgebracht werden können, ist bei SVAR-Modellen zu berücksichtigen, dass die Resultate von den Annahmen abhängen, die zur Identifikation der strukturellen Schocks getroffen werden. Wenn die Identifikationsannahmen falsch sind, wird z.B. ein struktureller Schock, der als geldpolitischer Schock identifiziert wurde, auch den systematischen Teil der Geldpolitik repräsentieren oder eine Kombination des geldpolitischen Schocks und anderer Schocks darstellen. Obwohl die Resultate in jedem ökonometrischen Modell durch die Annahmen beeinflusst werden, ist es nützlich zu wissen, wie gross dieser Einfluss ist.

Zur Identifikation struktureller Schocks werden typischerweise Informationen verwendet, die entweder institutionelle Gegebenheiten oder theoretische Vorstellungen widerspiegeln. Beispielsweise können Informationen über das geldpolitische Instrument und dessen kurzfristige Wirkungen oder das Konzept der langfristigen Geldneutralität zur Identifikation von Schocks beigezogen werden. Da über die Gültigkeit dieser identifizierenden Annahmen oft eine gewisse Unsicherheit besteht, haben Faust (1998)

und Uhlig (2001) vorgeschlagen, diese Unsicherheit in die Identifikation der Schocks einfließen zu lassen. Ihre Grundidee besteht darin, zuerst die strukturellen Schocks unter verschiedenen Annahmen zu identifizieren und anschliessend die Wirkungen der Schocks zu vergleichen. So können z.B. bedingte Prognosen bezüglich der Wirkung geldpolitischer Schocks, die auf unterschiedlichen Identifikationsannahmen beruhen, verglichen werden. Dieser Vergleich der Prognosen liefert ein Mass für die Unsicherheit der Prognose, die auf die Unsicherheit bezüglich der Identifikationsannahmen zurückzuführen ist.

Literaturverzeichnis

- Bernanke, B.S. 1986. Alternative Explanations of the Money-Income Correlation. *Carnegie Rochester Conference Series on Public Policy* 25: 49–100.
- Blanchard, O.J. und Quah, D. 1989. The Dynamic Effect of Aggregate Supply and Demand Disturbances. *American Economic Review* 79: 65–77.
- Blanchard, O.J. und Watson, M.W. 1986. Are Business Cycles All Alike? In *The American Business Cycle – Continuity and Change*, Hrsg. R. Gordon. Chicago: University of Chicago Press.
- Canova, F. 1995. Vector Autoregressive Models: Specification, Estimation, Inference and Forecasting. In *Handbook of Applied Econometrics, Macroeconometrics*, Hrsg. M.H. Pesaran und M.Wickens. Cambridge MA: Blackwell.
- Cooley, T.F. und LeRoy, S.F. 1985. A Theoretical Macroeconomics: A Critique. *Journal of Monetary Economics* 16: 283–308.
- Dickey, D. und Fuller, W.A. 1981. Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root. *Econometrica* 49: 1057–72.
- Engle, R.E. und Granger, C.W.J. 1987. Cointegration and Error-Correction: Representation, Estimation and Testing. *Econometrica* 55: 251–76.
- Faust, J. 1988. The Robustness of Identified VAR Conclusions About Money. *Carnegie-Rochester Conference Series on Public Policy* 49: 207–244.
- Friedman, M. 1968. The Role of Monetary Policy. *American Economic Review* 58: 1–17.
- Friedman, M. und Schwartz, A.J. 1963. *A Monetary History of the United States, 1867–1970*. Princeton: Princeton University Press.
- Gali, J. 1992. How Well Does the IS-LM Model Fit Postwar US Data? *Quarterly Journal of Economics* 107: 709–735.
- Johansen, S. 1988. Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control* 12: 231–254.
- Jordan, T.J. 1999. Inflationsprognosen mit VAR-Systemen. Arbeitspapier, Schweizerische Nationalbank.
- Jordan, T.J. 1998. Monetary Policy Shocks in Switzerland: Is There a Liquidity Effect? Arbeitspapier, Schweizerische Nationalbank.
- Jordan, T.J. und Peytrignet, M. 2001. Die Inflationsprognose der Schweizerischen Nationalbank. *Schweizerische Nationalbank Quartalsheft* 19(2): 54–61.
- Jordan, T.J. und Savioz, M.R. 2001. Is there a Point in Combining Inflation Forecasts from VAR-Models? An Empirical Investigation for Switzerland. Arbeitspapier, Schweizerische Nationalbank.
- King, R.G., Plosser, C.I., Stock, J.H. und Watson, M.W. 1991. Stochastic Trends and Economic Fluctuations. *American Economic Review* 81:819–840.
- Kugler, P. und Jordan, T.J. 2000. Vector Autoregressions and the Analysis of Monetary Policy Interventions: The Swiss Case. Arbeitspapier, Schweizerische Nationalbank.
- Kirchgässner, G. und Savioz, M.R. 1997. Empirische Forschung in den Wirtschaftswissenschaften: Ein Überblick. *Homo oeconomicus* 16: 209–247.
- Koopmans, T.C. und Hood, W.C. 1953. The Estimation of Simultaneous Linear Economic Relationships. In *Studies in Econometric Method*, Hrsg. W.C. Hood und T.C. Koopmans, 112–199. New York: John Wiley.
- Leeper, E.M. und Zha, T. 1999. Modest Policy Interventions. Federal Reserve Bank of Atlanta Working Paper 99–122.
- Litterman, R.B. 1980. A Bayesian Procedure for Forecasting with Vector Autoregression. Massachusetts Institute of Technology, Department of Economics Working Paper.
- Liu, T.C. 1960. Underidentification, Structural Estimation, and Forecasting. *Econometrica* 28: 855–865.
- Lucas, R.E. 1976. Macroeconomic Policy Evaluation: A Critique. *Carnegie-Rochester Conference Series on Public Policy* 1: 19–46.
- Phelps, E.S. 1967. Phillips Curves, Expectations of Inflation and Optimal Employment over Time. *Economica* 34: 254–281.
- Shapiro, M.D. und Watson, M.W. 1988. Sources of Business Cycle Fluctuations. *NBER Macroeconomics Annual* 3: 108–148.
- Sims, C.A. 1980. Macroeconomics and Reality. *Econometrica* 48: 1–48.
- Sims, C.A. 1986 Are Forecasting Models Usable for Policy Analysis. *Federal Reserve Bank of Minneapolis, Quarterly Review* 10: 2–16.
- Sims, C.A., Stock, J. und Watson, M. 1990. Inference in Linear Time Series Models with Some Unit Roots. *Econometrica* 58:113–44.
- Stalder, P. 2001. Ein ökonometrisches Makromodell für die Schweiz. *Schweizerische Nationalbank Quartalsheft* 19(2): 63–89.
- Stock, J.H. und Watson, M.W. 2001. Vector Autoregressions. *Journal of Economic Perspectives* 1(4): 101–115.
- Uhlig, H. 2001. What are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure. Arbeitspapier, Humboldt Universität Berlin.
- Zha, T. 1999. Block Recursion and Structural Vector Autoregressions, *Journal of Econometrics* 90: 265–289.