Sind Kombinationen von Prognosen aus VAR-Modellen sinnvoll? Eine empirische Analyse mit Inflationsprognosen für die Schweiz

Thomas J. Jordan und Marcel R. Savioz, Forschung,
Schweizerische Nationalbank, Zürich

Wenn ihr durchschauen könnt die Saat der Zeit,
Und sagen: dies Korn sprosst und jenes nicht,
So sprecht zu mir (...). Shakespeare.[1]

1 Einleitung[2]

Anfang 2000 führte die Schweizerische Nationalbank (SNB) ein neues geldpolitisches Konzept ein. Das neue Konzept basiert auf einer expliziten Definition der Preisstabilität und verwendet eine Inflationsprognose als Hauptindikator für die geldpolitische Entscheidungsfindung.[3] Zwar hat dieser vorausschauende Ansatz der SNB gewisse Ähnlichkeiten mit dem 2-Säulen-Konzept der Europäischen Zentralbank und der strikten Inflationszielstrategie, wie sie z. B. die Bank of England verfolgt. Dennoch weist es Eigenschaften auf, die es zu einem eigenständigen geldpolitischen Konzept machen.[4]

Diese neuen geldpolitischen Strategien haben eine Gemeinsamkeit: die aus ihnen hergeleiteten geldpolitischen Beschlüsse sind alle vorausschauend. Dies ist durch die grosse zeitliche Verzögerung bedingt, mit der sich ein monetärer Impuls auf Produktion und Preise auswirkt. In der Schweiz geht man beispielsweise davon aus, dass zwischen dem monetären Impuls und der Hauptauswirkung auf die Preise ein bis drei Jahre vergehen. Für Zentralbanken sind deshalb Inflationsprognosen mit einem Prognosehorizont von ein bis drei Jahren eine besonders wichtige Entscheidungsgrundlage. In letzter Zeit sind die Forschungsanstrengungen auf diesem Gebiet entsprechend intensiviert worden. Insbesondere geht es um die Erstellung langfristiger Inflationsprognosen mit verschiedenen Methoden und die Analyse ihrer Eigenschaften und Güte. Dieser Aufsatz ist ein empirischer Beitrag zur einschlägigen Literatur.

Das Erstellen einer präzisen Inflationsprognose mit einem Horizont von ein bis drei Jahren ist ein schwieriges Unterfangen. Angesichts von Unsicherheiten bezüglich der tatsächlichen wirtschaftlichen Zusammenhänge und des monetären Transmissionsmechanismus sehen sich die Zentralbanken gezwungen, bei der Inflationsprognose verschiedene Ansätze zu verwenden und sich nicht auf ein einziges Modell zu stützen.[5] Einer dieser Ansätze beruht auf grossen makroökonomischen Strukturmodellen.[6] Der Vorteil dieser Modelle liegt darin, dass Prognosen für eine Vielzahl von Variablen erstellt werden und dass eine klare ökonomische Intuition hinter der Dynamik der Prognose geliefert wird. Problematisch bei grossen strukturellen Modellen sind die restriktiven Annahmen, die getroffen werden müssen, um die wirtschaftlichen Zusammenhänge zu identifizieren.[7] Ein zweiter Ansatz sind die vektorautoregressiven Modelle (VAR-Modelle).[8] VAR-Modelle nutzen die in makroökonomischen Zeitreihen enthaltenen Informationen, ohne dabei starke Restriktionen bezüglich der wirtschaftlichen Zusammenhänge zu bedingen. Bei VAR-Modellen kommt es demnach weniger zu einer Datenverfälschung, da sie bezüglich der Struktur der Wirtschaft und des Transmissionsmechanismus weniger – zwangsläufig – inkorrekte Restriktionen auferlegen. Ein weiterer Vorteil von VAR-Modellen ist, dass keine Annahmen für den Verlauf exogener Variablen während des Prognosehorizonts getroffen werden müssen. Alle Variablen in VAR-Modellen sind endogen und die dynamischen Prognosen einfach zu berechnen. Bei der Schätzung von VAR-Modellen besteht jedoch oft das Problem zu geringer Freiheitsgrade. Die Anzahl der zu schätzenden Parameter wird schnell einmal sehr gross, sobald mehr Variablen in ein VAR-Modell einbezogen werden. Angesichts der begrenzten Länge der Quartalszeitreihen in der Schweiz können beispielsweise VAR-Modelle häufig mit lediglich drei bis fünf Variablen geschätzt werden. Da VAR-Modelle auf einige wenige Variablen beschränkt werden müssen, können makroökonomisch relevante Informationen zum Teil unberücksichtigt bleiben. Eine Möglichkeit, dieses Problem zu umgehen, ist die Verwendung mehrerer kleiner VAR-Modelle, deren Prognosen dann kombiniert werden.[9] Als Ziel dieser Studie soll analysiert werden, ob die Kombination von Prognosen aus verschiedenen kleinen VAR-Modellen die Güte von Einzelprognosen verbessern kann, da Informationen aus mehr Variablen berücksichtigt werden. Dies wird anhand von Inflationsprognosen für die Schweiz untersucht.[10]

Der Aufsatz gliedert sich wie folgt: Im zweiten Teil wird die Rolle unbedingter Prognosen im geldpolitischen Entscheidungsprozess diskutiert. Im dritten Teil werden die verschiedenen Methoden zur Kombination von Prognosen erläutert. Im vierten Teil werden die Zeitreiheneigenschaften der Variablen und die verschiedenen VAR-Modelle, die für Prognosen eingesetzt werden, analysiert. Im fünften Teil wird anhand von Out-of-sample-Ergebnissen erörtert, ob kombinierte Prognosen besser sind als Einzelprognosen. Der sechste Teil enthält Schlussbemerkungen.

1 Zitat aus Granger (1989, S. 153), übersetzt von D. Tieck.
2 Dieser Aufsatz profitierte von wertvollen Kommentaren und Anregungen von Caesar Lack, Jean-Marc Natal, Samuel Reynard, Enzo Rossi, Martin Schlegel und Peter Stalder sowie von den Diskussionen mit den Teilnehmern des Jahrestreffens der Schweizerischen Gesellschaft für Volkswirtschaft und Statistik 2001 in Genf und des Meetings der «Arbeitsgruppe Prognoseverfahren der Gesellschaft für Operations Research» an der Eichstätt-Universität in Ingolstadt im Jahre 2000. Für allfällige Fehler sind die Autoren alleine verantwortlich.
3 Vgl. Jordan und Peytrignet (2001) für eine Darstellung der Rolle von Inflationsprognosen im neuen geldpolitischen Konzept.
4 Vgl. Baltensperger, Fischer und Jordan (2002) für eine Diskussion der charakteristischen Merkmale des geldpolitischen Konzepts der Schweiz.
5 Vgl. Kirchgässner und Savioz (1997) für eine Diskussion verschiedener ökonometrischer Ansätze.
6 Vgl. Stalder (2001) für die Darstellung des von der SNB verwendeten grossen strukturellen Modells.
7 Die klassische Kritik grosser struktureller Modelle wurde von Sims (1980) formuliert.
8 Vgl. Jordan, Kugler, Lenz und Savioz (2002) für eine Beschreibung der VAR-Modelle, die bei der SNB für bedingte und unbedingte Prognosen eingesetzt werden.
9 Eine andere Möglichkeit wäre, die Bayesianische VAR-Methodologie zu verwenden oder einige Koeffizienten nach entsprechenden Tests auf Null zu setzen.
10 Empirische Ergebnisse zu BIP-Prognosen für die Schweizer Volkswirtschaft finden sich in Ruoss und Savioz (2002).

2 Prognosearten in der Geldpolitik: Bedingte und unbedingte Prognosen

In der Geldpolitik werden zwei Arten von Inflationsprognosen erstellt: Einerseits werden bedingte Prognosen hergeleitet, die auf der Annahme eines bestimmten zukünftigen Verlaufs der Geldpolitik beruhen. Diese ermöglichen es der Zentralbank, die Auswirkungen verschiedener geldpolitischer Alternativen abzuschätzen. Anderseits werden unbedingte Prognosen errechnet. Diese zeigen den Inflationsverlauf, wenn die geldpolitische Haltung während des Prognosehorizonts ebenfalls explizit oder implizit prognostiziert wird.

Unbedingte Inflationsprognosen werden aus drei Hauptgründen erstellt. Erstens dienen sie als Benchmark für die in der Vergangenheit beobachtete Reaktion der Zentralbank auf die makroökonomische Lage. Solche Inflationsprognosen sind besonders wichtig und aufschlussreich, da die bei bedingten Prognosen zugrunde liegenden Annahmen über den Verlauf des geldpolitischen Instruments – z.B. ein über den Prognosehorizont konstanter Zinssatz – meist unrealistisch sind. Unbedingte Prognosen sind deshalb wichtige Indikatoren für die allgemeinen Inflationsaussichten. Zweitens müssen die verschiedenen Modelle bezüglich der Qualität ihrer unbedingten Prognosen beurteilt werden, da bedingte Inflationsprognosen aufgrund ihrer hypothetischen Natur nicht auf ihre Güte hin getestet werden können. Eine Zentralbank muss deshalb für alle ihre Modelle – auch für strukturelle – unbedingte Prognosen durchführen, um sie zu validieren. Drittens erlauben unbedingte Prognosen auch den Vergleich mit externen, ausserhalb der Zentralbank erstellten Prognosen. Die geldpolitischen Entscheidungsträger können somit beurteilen, ob zwischen der Markteinschätzung der Inflationsdynamik und der eigenen Analyse Unterschiede bestehen.

Einfache VAR-Modelle stellen reduzierte Formen dar, d.h. die Parameter von einfachen VAR-Modellen haben keine strukturelle Interpretation. Das Erstellen *bedingter* Prognosen mit einfachen VAR-Modellen ist problematisch, denn ein vorgegebener Verlauf des Zinssatzes ist nicht gleichzusetzen mit einem bestimmten geldpolitischen Kurs.[11] Ausserdem sind die geschätzten Koeffizienten nicht politikinvariant. VAR-Modelle sind vielmehr eine ideale Methode zur Erstellung *unbedingter* Benchmark-Prognosen, weil sie sich nur auf ein Minimum an strukturellen Informationen stützen, d.h. auf die gewählten Variablen und die Länge des Lag. Ausserdem müssen exogene Variablen, anders als bei strukturellen Modellen, nicht prognostiziert werden. In diesem Aufsatz befassen wir uns ausschliesslich mit unbedingten VAR-Inflationsprognosen. Dabei untersuchen wir, ob sich deren Qualität durch Kombinationen verbessern lassen.

3 Kombinationen von Prognosen

Der traditionelle Ansatz zur Erstellung von VAR-Prognosen ist einfach. Der Prognostiker spezifiziert die Variablen und die Anzahl Lags, die ins VAR-Modell einfliessen, und trifft die Annahmen betreffend Integration, Kointegration, Trend und Saisoncharakter der Daten. Danach werden die Prognosen mit dem ausgewählten Modell berechnet. Die in der Makroökonomie üblicherweise begrenzte Länge der Zeitreihen ist im traditionellen Ansatz mit drei grundlegenden Problemen verbunden. Erstens können nur sehr kleine VAR-Modelle, d.h. solche mit wenigen Variablen, für die Prognose beigezogen werden. Dies führt dazu, dass potenziell nützliche Informationen nicht berücksichtigt werden, wenn man sich auf ein einziges Modell beschränkt. Zweitens können die kleinen Freiheitsgrade zu grossen Standardfehlern bei den geschätzten Parametern führen, wodurch die Out-of-sample-Performance der Prognosen in Mitleidenschaft gezogen werden dürfte. Ein weiterer Nachteil des traditionellen Ansatzes liegt darin, dass die Wahl des Modells nicht in erster Linie von der Prognose-Performance in der Vergangenheit abhängt, sondern von der Güte des Fit.

Um diese Nachteile auszuschalten, haben wir ein alternatives Verfahren entwickelt. Dieses besteht aus zwei Schritten: In einem ersten Schritt berechnen wir eine Vielzahl von Prognosen mittels einer Serie von kleinen VAR-Modellen. Die Kleinheit der VAR-Modelle gewährleistet ein Minimum an Freiheitsgraden. Im zweiten Schritt werden die Prognosen verschiedener Modelle gewichtet und zu einer Prognose kombiniert.[12] Die mit diesem Verfahren ermittelten Prognosen nennen wir «kombinierte VAR-Prognosen» (*combined VAR forecast* oder *CVAR*). Es gibt viele verschiedene Methoden, um die Prognosen zu kombinieren. Wir beschränken uns im Folgenden auf die gängigsten.[13]

Es gibt drei Hauptgründe, weshalb kombinierte Prognosen das Ergebnis verbessern sollten: Erstens lassen sich dadurch die Prognosefehler diversifizie-

11 Die Erstellung bedingter Prognosen mit strukturellen VAR-Modellen wird in Kugler und Jordan (2000) sowie in Jordan, Kugler, Lenz und Savioz (2002) behandelt.

12 Siehe Winkler (1989, S. 606) für die Basismotivation, welche dem vorgeschlagenen Verfahren zugrunde liegt: "In most interesting forecasting situations in our uncertain and rapidly changing world, I doubt that such 'true' models are attainable and I think that it is counterproductive to think in terms of 'true' models. The motivation for the combination of forecasts, then, is at its most basic level the simple idea of aggregation of information to achieve a reduction in uncertainty, or an increase in accuracy."

13 Methoden zur Bestimmung der Gewichtung finden sich z.B. in Clemen und Winkler (1986), Clemen (1989) sowie Holden und Peel (1986).

ren, was das Problem ungenauer Schätzungen einzelner Modelle verringert.[14] Zweitens sollten kombinierte Prognosen robuster sein, da sie nicht stark von der Spezifikation eines einzigen Modells abhängen. Ein Spezifikationsfehler eines einzigen Modells dürfte demzufolge viel weniger Schaden anrichten. Drittens stützen sich kombinierte Prognosen auf eine breitere Informationsbasis, da sie mehr Variablen einbeziehen. Dies ist insbesondere für VAR-Prognosen relevant, die in der Regel auf Modellen mit nur wenigen Variablen beruhen. Der Nutzen kombinierter Prognosen sollte v.a. dann deutlich sein, wenn sich die Gewichte aus der früheren Performance der einzelnen Prognosen ableiten.

Kombinierte VAR-Prognosen können weitere Vorteile aufweisen. Erstens kann die Gewichtung der verschiedenen Prognosen – falls von der früheren Performance abhängig – Informationen über die relative Wichtigkeit der verschiedenen Modelle und Variablen enthalten. Zweitens kann eine Änderung der Streuung der einzelnen Prognosen einen frühzeitigen Hinweis auf eine Verschlechterung der Prognosegüte geben. Diese zusätzlichen Punkte werden in diesem Aufsatz nicht untersucht. Wir beschäftigen uns ausschliesslich mit der Frage, ob CVAR-Prognosen präziser sind als VAR-Prognosen.

Die in dieser Untersuchung angewandten Methoden zur Gewichtung der Prognosen werden anhand eines Beispiels erklärt, in welchem Inflationsprognosen für den Zeitpunkt t aus drei VAR-Modellen verfügbar sind: $\hat{\pi}_{VAR1,t}$, $\hat{\pi}_{VAR2,t}$ und $\hat{\pi}_{VAR3,t}$.[15] Die kombinierte Prognose $\hat{\pi}_{CVAR,t}$ ist ein gewichtetes Mittel dieser drei Einzelprognosen

(1) $\hat{\pi}_{CVAR,t} = w_0 + w_1\hat{\pi}_{VAR1,t} + w_2\hat{\pi}_{VAR2,t} + w_3\hat{\pi}_{VAR3,t}$,

wobei w_i $i = 0,\ldots 3$ die Gewichte sind.

Die erste, sehr häufige Methode zur Kombination von Prognosen besteht darin, den *einfachen Durchschnitt (simple average* oder *SA)* der einzelnen Prognosen zu berechnen. Dabei wird jeder Prognose das gleiche Gewicht beigemessen. Ferner beträgt die Summe der Gewichte eins:

(2) $w_0 = 0 \;\; w_1 = 1/3 \;\; w_2 = 1/3 \;\; w_3 = 1/3.$

Die Gewichte hängen nicht von der in der Vergangenheit beobachteten Genauigkeit der Einzelprognosen ab.

Im Gegensatz zu dieser werden die Gewichte bei den anderen Kombinationsmethoden aus der früheren Performance der einzelnen Prognosen bestimmt. Dies geschieht durch eine lineare Regression, wobei die effektive Inflationsrate als abhängige Variable und die einzelnen Out-of-sample-Prognosen als erklärende Variablen eingesetzt werden. Die Koeffizienten können mit Restriktionen geschätzt werden, so dass die Eigenschaften von Gewichten ($0 \leq w_i \leq 1$, $\sum_i w_i = 1$) teilweise oder vollständig erfüllt werden. In unserem Beispiel mit drei Prognosen ist die Regression:[16]

(3) $\pi_t = \beta_0 + \beta_1\hat{\pi}_{VAR1,t} + \beta_2\hat{\pi}_{VAR2,t} + \beta_3\hat{\pi}_{VAR3,t} + \varepsilon_t.$

Die zweite Kombinationsmethode, die wir hier anwenden, ist die *gewöhnliche Kleinstquadratemethode* (*least square method* oder *LS*). Bei der LS-Kombinationsprognose (1) dienen die geschätzten Koeffizienten der Gleichung (3) als Gewichte zur Berechnung der kombinierten Prognose:

(4) $w_0 = \hat{\beta}_0 \;\; w_1 = \hat{\beta}_1 \;\; w_2 = \hat{\beta}_2 \;\; w_3 = \hat{\beta}_3.$

Die Schätzung der Koeffizienten in Gleichung (3) erfolgt ohne Restriktionen. Man beachte, dass der Koeffizient $\hat{\beta}_0$ für nichtverzerrte Prognosen null ist.

Die dritte Kombinationsmethode ist die *Kleinstquadratemethode mit restringierter Konstante* (*constant restricted least square method* oder *CRLS*). Die Prognose wird als unverzerrt angenommen, und der konstante Term auf null restringiert:

(5) $w_0 = \hat{\beta}_0 = 0.$

Die vierte Kombinationsmethode ist die *Kleinstquadratemethode mit Gleichheitsrestriktion* (*equality restricted least square method* oder *ERLS*). Diese enthält die zusätzliche Restriktion, dass die Summe der Gewichte der Prognosen eins ergibt. Die Regression (3) wird demzufolge mit den folgenden Restriktionen geschätzt:

(6) $\hat{\beta}_0 = 0$ und $\hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3 = 1.$

Das fünfte und letzte in dieser Arbeit verwendete Verfahren ist die Kleinstquadratemethode mt Ungleichheitsrestriktion (non-negativity inequality restricted least square method oder NRLS). In NRLS-Kombinationsprognosen sind die Gewichte nicht negativ. Die Regression (3) wird mit den folgenden (nichtlinearen) Restriktionen geschätzt:

(7) $\hat{\beta}_0 = 0, \hat{\beta}_1 \geq 0, \hat{\beta}_2 \geq 0,$ und $\hat{\beta}_3 \geq 0.$

Im folgenden Abschnitt verwenden wir diese fünf Methoden zur Erstellung kombinierter Prognosen. Dabei soll geklärt werden, welche Prognosen – CVAR- oder einzelne VAR-Prognosen – die besseren sind.

14 Vgl. z.B. Granger (1989), Granger und Newbold (1973, 1986). Vgl. Jungmittag (1993) für eine Einführung ins Diversifikationsargument.

15 Vgl. Aksu und Gunter (1992).

16 Kombinierte VAR-Prognosen ergeben nicht nur eine Prognose, sondern auch zusätzliche Informationen, die für die Prognostiker und für die Durchführung der Geldpolitik von Nutzen sein können. Erstens geben kombinierte VAR-Prognosen nützliche Hinweise auf die Quelle der Prognose-Performance. Dies folgt aus Gleichung (3). Prognosen, die in dieser Regression keinen signifikanten Koeffizienten aufweisen, enthalten im Vergleich mit anderen Prognosen keine zusätzlichen Informationen (vgl. Diebold (1989) für eine Diskussion der Prognosekombination und «forecast encompassing» sowie West (2001) für einen geeigneten Test für «forecast encompassing»). Es kann also abgeleitet werden, welche Variablen gute Aussagen über die Inflation für einen bestimmten Prognosehorizont ergeben. Ein Achsenabschnitt ungleich null in der Regression weist auf eine Verzerrung der Prognosen hin (vgl. Holden und Peel, 1989). Zweitens können Strukturbrüche im Inflationsverlauf durch die Analyse der Veränderungen der geschätzten Gewichte frühzeitig identifiziert werden (vgl. Diebold und Pauly, 1986). Die wahrscheinliche Ursache des Strukturbruchs kann ebenfalls von der Veränderung der Gewichte abgeleitet werden.

4 Daten, Zeitreiheneigenschaften und VAR-Modelle

Um den Rechen- und Programmierungsaufwand in Grenzen zu halten, beschränken wir uns auf die folgenden fünf Variablen: den Konsumentenpreisindex P, das Geldaggregat $M3$, das Total inländischer Bankkredite C, das reale Bruttoinlandprodukt (BIP) Q und den langfristigen Zinssatz für Schweizer Franken R. Gemäss den Haupttheorien des Transmissionsmechanismus in der Geldpolitik sind Geld, Kredit, Wirtschaftsaktivität und der langfristige Zinssatz wichtige Determinanten im Inflationsprozess.[17]

Die Untersuchungsperiode umfasst den Zeitraum vom ersten Quartal 1974 bis zum dritten Quartal 2000.[18] Tabelle 1 zeigt den erweiterten Dickey-Fuller-Einheitswurzeltest für die fünf Variablen. Die Logarithmen aller fünf Variablen werden in diesem Aufsatz als I(1) angenommen.[19] Nur VAR-Modelle mit stationären Variablen werden berücksichtigt. Deshalb gehen alle Variablen als erste Differenzen in die Modelle ein. Für den Zweck dieses Aufsatzes verzichten wir auf Spezifikationen in Bezug auf Kointegration und Vektorfehlerkorrektur.

Die betrachteten VAR-Modelle enthalten eine Konstante und vier Lags. Es werden keine Trend- oder Saisondummies berücksichtigt.[20] Alle VAR-Modelle müssen im Minimum die Inflationsrate π_t beinhalten. Das kleinste VAR-Modell umfasst deshalb nur π_t, das grösste alle fünf Variablen. Innerhalb dieser Anordnung können 16 verschiedene VAR-Modelle spezifiziert werden: 1 Modell mit 1 Variablen, 4 Modelle mit 2 Variablen, 6 Modelle mit 3 Variablen, 4 Modelle mit 4 Variablen und 1 Modell mit 5 Variablen.

Kombinierte Prognosen werden nur aus Prognosen von VAR-Modellen mit der gleichen Anzahl Variablen erstellt. Mit dieser Restriktion können 79 verschiedene kombinierte Prognosen spezifiziert werden: 11 kombinierte Prognosen aus Modellen mit 2 Variablen, 57 aus Modellen mit 3 Variablen und 11 aus Modellen mit 4 Variablen (vgl. Tabelle 4).

5 Out-of-Sample-Prognosen

In diesem Abschnitt untersuchen wir die Möglichkeit, die Prognosegüte durch die Kombination verschiedener Prognosen zu verbessern. Wir vergleichen die Ergebnisse von CVAR-Prognosen mit den individuellen VAR-Prognosen. Sowohl das Durchschnittsergebnis von VAR- und CVAR-Prognosen als auch das Ergebnis der besten VAR- und CVAR-Prognosen werden verglichen. Wir konzentrieren uns dabei auf die Prognose der jährlichen Inflationsrate

(8) $\pi_t = 100{*}\log(P_t/P_{t-4})$

für Prognosehorizonte von ein, zwei und drei Jahren. Dies sind die für die Geldpolitik relevantesten Horizonte.

Um die Prognosen zu beurteilen, stützen wir uns auf die Wurzel des durchschnittlichen quadratischen Prognosefehlers (*root mean square error* oder *RMSE*)

(9) $RMSE = \sqrt{1/T\sum_{t=1}^{T}(\pi_t - \hat{\pi}_t)^2}$,

wobei π_t die realisierte und $\hat{\pi}_t$ die für den Zeitpunkt t prognostizierte Inflation widerspiegelt. Die Differenz $\pi_t - \hat{\pi}_t$ zeigt den Prognosefehler an, und T steht für die Anzahl Prognosen. Der MSE ($MSE = 1/T\sum_{t=1}^{T}(\pi_t - \hat{\pi}_t)^2$) ist ein Mass für den durchschnittlichen quadratischen Prognosefehler. Durch Quadrierung der Prognosefehler werden grosse Prognosefehler proportional stärker gewichtet als kleine. Durch Wurzelziehen erhält man den RMSE, welcher die gleiche Dimension hat wie die Inflationsrate π_t. Sind die Prognosen vollkommen, ist der RMSE null.

Die Analyse zieht nur Out-of-sample-Prognosen in Betracht. Wie u.a. von Bernanke (1990) sowie Thoma und Gray (1994) aufgezeigt, hängt die Nützlichkeit eines Prognosemodells letztlich von seiner Fähigkeit ab, Out-of-sample-Prognosen zu erstellen. Erstellt ein Modell hervorragende In-sample-Prognosen, muss dies bei Out-of-sample-Prognosen nicht automatisch auch der Fall sein.

Die Prognosen werden anhand von rollenden Regressionsanalysen *(rolling regression models)* errechnet: Eine rollende Schätzung der VAR-Modelle ergibt eine Reihe von einzelnen Out-of-sample-Prognosen für verschiedene Prognosehorizonte $k = 4, 8, 12$. Die Prognosen für den Horizont von k Quartalen berechnen sich wie folgt: Zuerst wird das VAR-Modell unter Einbezug von Beobachtungen zwischen dem Zeitpunkt s und dem Zeitpunkt $s - 29$ geschätzt, wobei s

17 Die Studie von Jordan (1999a) hat gezeigt, dass Kreditaggregate gute Eigenschaften für die Voraussage von Inflation aufweisen. Für die Bedeutung von Geldaggregaten vgl. Baltensperger, Jordan und Savioz (2001) sowie Kirchgässner und Savioz (2001).
18 Man beachte, dass in letzter Zeit die Inflationsvariabilität sehr klein und das Inflationsniveau tief waren. Die neusten Daten sind deshalb wenig informativ bezüglich Performancetests von Modellen zur Prognose von (hoher) Inflation.
19 Vgl. Miller, Clemen und Winkler (1992).
20 Die Variablen sind saisonbereinigt.

die Periode ist, nach welcher die erste Prognose beginnt.[21] Die geschätzten Koeffizienten werden sodann zur Berechnung der Prognose für die Zeit $s+k$ verwendet. Für die Prognose des Zeitraums $s+k$ wird nur die zur Zeit s verfügbare Information verwendet. Danach wird das Sample um eine Periode erweitert, und die Gleichung wird nochmals mit Daten für den Zeitraum $s+1$ bis $s-29$ geschätzt. Die neu geschätzten Koeffizienten werden dann für die Berechnung der Prognose für den Zeitraum $s+1+k$ verwendet. Dieses Vorgehen wird bis zum aktuellen Rand fortgesetzt. Das Sample wird jedoch konstant gehalten, sobald 50 Beobachtungen erreicht worden sind. Mit dieser Technik können 73 einzelne Einjahresprognosen von 1982:3 bis 2000:3, 69 Prognosen mit einem Zweijahreshorizont von 1983:3 bis 2000:3 und 65 Prognosen mit einem Dreijahreshorizont von 1984:3 bis 2000:3 berechnet werden.

Kombinierte Prognosen können anhand von Gewichtungen erstellt werden, die mit den fünf oben vorgestellten Methoden errechnet worden sind. Zur Ermittlung dieser Gewichtungen werden rollende Regressionstechniken angewendet. Die Regression der Gleichung (3) wird mit den ersten 30 einzelnen Prognosen durchgeführt, um die erste kombinierte Prognose für den Zeitpunkt $s+2k+29$ zu erhalten, wobei $s+k$ dem Quartal entspricht, für welches die erste einzelne Prognose mit dem Horizont k verfügbar ist. Sodann erfolgt die Regression mit den ersten 31 einzelnen Prognosen, um die kombinierte Prognose für den Zeitpunkt $s+2k+30$ zu erstellen. Diese Vorgehensweise wird bis zum aktuellen Rand wiederholt. Die Anzahl der in der Regression verwendeten Einzelprognosen bleibt nach Erreichen von 50 wieder konstant. Dank diesem Vorgehen erhalten wir 39 kombinierte Prognosen mit einem Einjahreshorizont von 1991:3 bis 2000:3, 31 kombinierte Prognosen mit einem Zweijahreshorizont von 1993:3 bis 2000:3 und 23 kombinierte Prognosen mit einem Dreijahreshorizont von 1995:3 bis 2000:3.[22]

Für die Periode von 1991:1 bis 2000:3 betrug die durchschnittliche effektive Inflationsrate 1,9% bei einer Maximalrate von 6,1% und einem *root mean square* ($RMS = 1/T\sum_{t=1}^{T}\sqrt{\pi_t^2/T}$) von 2,62%. Für die Periode von 1993:1 bis 2000:3 lag die Inflation durchschnittlich bei 1,2% und erreichte ein Maximum von 3,4% sowie einen RMS von 1,54%. In der Periode von 1995:1 bis 2000:3 war die durchschnittliche Inflation lediglich 0,9% und immer unter 2%. Der RMS betrug 1,09%. Die Prognose-Performance wird mit Vorteil über eine Periode beurteilt, in der die Inflation gewissen Schwankungen unterworfen war. Deshalb verzichten wir auf die Wiedergabe der Ergebnisse von Prognosen mit verschiedenen Zeithorizonten anhand eines gemeinsamen Sample.

Die Veränderung sowohl der Volatilität als auch des Inflationsniveaus erschwert die Einschätzung der Prognose-Performance der Modelle zwischen den verschiedenen Prognoseperioden und -horizonten. Um die Prognosegüte zwischen verschiedenen Perioden zu vergleichen, bietet sich u.a. Theil's Ungleichheitskoeffizient (U) an:[23]

$$(10)\quad U = \frac{RMSE}{RMS} = \frac{\sqrt{1/T\sum_{t=1}^{T}(\pi_t - \hat{\pi}_t)^2}}{1/T\sum_{t=1}^{T}\sqrt{\pi_t^2/T}}\,.$$

Das Theil'sche U vergleicht den RMSE der Inflationsprognosen mit dem RMS der tatsächlichen Inflation. Die Prognosefehler werden so relativ zum Inflationsniveau skaliert, da absolute Prognosefehler in Perioden mit tiefer Inflation tendenziell kleiner sind als in Perioden mit hoher Inflation. Ausserdem erlaubt das Theil'sche U eine Beurteilung der Performance eines Modells im Vergleich zu einer einfachen Prognose stabilen Preisniveaus (d.h. Nullinflation). U ist gleich 1, wenn das Modell dieselbe Voraussagekraft hat wie die einfache Prognose. Ist U kleiner (grösser) als 1, liefert das Modell genauere (weniger genaue) Prognosewerte als die einfache Prognose. Man sollte sich jedoch der Tatsache bewusst sein, dass eine einfache Prognose eines konstanten Preisniveaus unter gewissen Umständen sehr gut sein kann. Ein U höher als 1 würde demnach nicht unbedingt auf eine schlechte Prognosefähigkeit hindeuten. Dies trifft insbesondere für die Periode von 1995:1 bis 2000:3 zu, während der eine Prognose eines konstanten Preisniveaus durchaus akzeptabel gewesen wäre. Während dieser beinahe sechsjährigen Periode stieg das Preisniveau nur um 4,3% bzw. 0,75% pro Jahr. Für Prognosen von 1995:1 bis 2000:3 mit einem Dreijahres-Horizont bedeutet deshalb ein U grösser als 1 nicht unbedingt eine schlechte Performance.

Tabelle 2 zeigt die Resultate der Out-of-sample Performance einzelner VAR-Prognosen.[24] Wir legen durchschnittliche Resultate für verschiedene Gruppen von VAR-Modellen vor. Wie die RMSE- und U-Masszahlen zeigen, verschlechtert sich im Allgemeinen die Prognosequalität, je länger der Prognosehorizont wird. Die Reihe mit einer Variablen entspricht einem AR(4)-Modell der Veränderung des Preisniveaus. Dieses Modell dient als Benchmark. Die anderen VAR-Prognosen übertreffen diesen Benchmark im Durch-

21 Es wird somit keine VAR-Prognose mit weniger als dreissig Beobachtungen geschätzt.

22 Wird der Prognosehorizont um ein Jahr erweitert, verkleinert sich das Sample für die Evaluation der kombinierten Prognosen um zwei Jahre (vgl. Tabelle 2 ff.). Das erste Jahr geht verloren, da durch die Verlängerung des Prognosehorizonts bei gegebenem Datensatz weniger Prognosen errechnet werden können. Ebenso geht das zweite Jahr verloren, weil bei gegebener Anzahl Prognosen weniger kombinierte Prognosen errechnet werden können.

23 Das Theil'sche U wird hier wie in Greene (2000, S. 310) definiert.

24 Für ähnliche Ergebnisse vgl. Jordan (1999b).

schnitt. Die Prognosen der VAR-Modelle mit vier Variablen ($n = 4$) schneiden durchschnittlich bei einem einjährigen Prognosehorizont am besten ab. Für die zweijährigen Prognosen schneiden die VAR-Modelle mit drei Variablen ($n = 3$) durchschnittlich am besten ab. Der Unterschied zu den VAR-Modellen mit vier Variablen ($n = 4$) ist aber sehr gering. Für Dreijahresprognosen erzielen die VAR-Modelle mit drei Variablen ($n = 3$) durchschnittlich die besten Resultate, gefolgt von den bivariaten VAR-Modellen ($n = 2$) und den VAR-Modellen mit vier Variablen ($n = 4$). Die Performance des einfachen VAR-Modells mit fünf Variablen ($n = 5$) ist nicht besonders gut. Dies mag auf die kleine Anzahl von Freiheitsgraden in der Schätzung dieses VAR-Modells zurückzuführen sein. Die Resultate zeigen eindeutig, dass der Einbezug zusätzlicher Variablen die Performance verbessert, zumindest was Prognosen auf ein und zwei Jahre hinaus anbetrifft.

5.1 Vergleich mit der durchschnittlichen Performance der individuellen VAR-Prognosen

Wie ist die Performance von CVAR-Prognosen im Vergleich mit VAR-Prognosen? Im Folgenden vergleichen wir die Performance der CVAR-Prognosen mit der *durchschnittlichen Performance* der VAR-Prognosen, wie sie in der ersten Zeile der Tabelle 2 aufgeführt wird. Tabelle 3 zeigt den durchschnittlichen RMSE sowie die U-Masszahl der 79 CVAR-Prognosen für die verschiedenen Kombinationsmethoden. Aus der Tabelle geht auch die aus der Kombination von Prognosen erreichte Verbesserung im Vergleich zum Durchschnitt aller einzelnen VAR-Prognosen hervor. Die einfache Durchschnittsmethode (SA) erzielt für alle Prognosehorizonte ein gutes Resultat. Für den Einjahreshorizont ist die Errechnung des einfachen Durchschnitts sogar die beste Methode zur Gewichtung der Einzelprognosen. Ein Vorteil dieser Methode ist, dass die Gewichte nicht geschätzt werden müssen. Die Kleinstquadratemethode (LS) schneidet bei allen Prognosehorizonten sehr schlecht ab. Die Anwendung der Kleinstquadratemethode mit restringierter Konstante (CRLS) verbessert die Performance für den Prognosehorizont von zwei und drei Jahren erheblich.[25] Was den Dreijahreshorizont anbelangt, ist CRLS die beste Methode. Mit der Gleichheitsrestriktion (ERLS), gemäss welcher die Summe der Gewichte 1 betragen muss, verbessert sich die Performance nur für den Einjahres-Prognosehorizont.[26] Die Methode, welche die Gewichte der Prognosen auf nicht negative Werte beschränkt (NRLS), erzielt bei Zwei- und Dreijahresprognosen gute Ergebnisse. Im Zweijahreshorizont schliesst sie sogar am besten ab.

Die Anwendung von konstanten Gewichtungen (SA) eignet sich insbesondere für kurze Prognosehorizonte. Für längerfristige Prognosen scheint es jedoch wichtig, dass sich die Gewichte mit der Zeit verändern können, indem sie für jede Periode neu geschätzt werden. Wie jedoch die schlechte Performance der Kleinstquadratemethode (LS) zeigt, sollten der Schätzung von Gewichten Restriktionen auferlegt werden. Während die ERLS-Methode nicht die richtige Restriktion für die Schätzung der Gewichte anzuwenden scheint, schneiden die CRLS- und die NRLS-Methode diesbezüglich gut ab. Wenn die einzelnen VAR-Prognosen sehr ähnlich, aber nicht identisch sind, ist CRLS numerisch immer noch machbar, NRLS möglicherweise nicht.[27]

Für die nachfolgende Analyse konzentrieren wir uns auf die für jeden Prognosehorizont besten Methoden. Für den Einjahreshorizont verwenden wir SA. Für den Zweijahreshorizont haben wir die Wahl zwischen CRLS und NRLS. Die Resultate sind sehr ähnlich. Um mögliche numerische Probleme zu vermeiden, entscheiden wir uns für CRLS anstatt NRLS. CRLS ist auch für den Dreijahreshorizont die beste Methode.

Tabelle 5 zeigt die Resultate für verschiedene Untergruppen der CVAR-Prognosen. Die Ergebnisse der durchschnittlichen VAR- und CVAR-Prognosen sind in der Tabelle im Sinne eines Benchmark ebenfalls aufgeführt. Wir bilden zwei verschiedene Untergruppen mit 79 kombinierten Prognosen. Zuerst werden Untergruppen für alle kombinierten Prognosen gebildet, die aus derselben Anzahl von VAR-Prognosen berechnet sind. Die Gruppe $m = 2$ besteht beispielsweise aus allen 27 kombinierten Prognosen, die aus zwei Einzelprognosen gebildet wurden. Diese Einzelprognosen können aus VAR-Modellen mit 2, 3, oder 4 Variablen stammen. Danach werden Untergruppen mit allen kombinierten Prognosen gebildet, die aus VAR-Prognosen mit der gleichen Anzahl Variablen errechnet wurden. Die Untergruppe ($n = 2$) besteht zum Beispiel aus allen 11 kombinierten Prognosen, welche aus Prognosen mit einzelnen VAR mit 2 Variablen berechnet sind. Die kombinierten Prognosen können sich aus 2, 3, oder vier Einzelprognosen zusammensetzen. Die Untergruppen sind in Tabelle 4 dargestellt.

Die Ergebnisse zeigen deutlich: Je mehr Prognosen kombiniert werden, desto besser ist die Pro-

25 Vgl. auch Granger und Ramanathan (1984) betreffend die CRLS-Methode.

26 Vgl. auch Clemen (1986) betreffend die ERLS-Methode.

27 Wenn die neue Regression keine numerischen Resultate ergibt, werden die Gewichte der vorhergehenden Regression verwendet.

gnose-Performance. Die Verbesserung ist in der zunehmenden Anzahl kombinierter Prognosen monoton. Werden 6 Prognosen ($m = 6$) kombiniert, verringert sich der RMSE verglichen mit dem Durchschnitt der individuellen VAR-Prognosen um über 20%. Diese Verbesserung könnte entweder das Ergebnis eines Diversifikationseffekts (eine grössere Anzahl Prognosen) oder eines Informationseffekts (eine grössere Anzahl Variablen) sein. Die Resultate zeigen auch, dass kombinierte Prognosen aus VAR-Modellen mit 3 oder 4 Variablen ($n = 3$, $n = 4$) durchschnittlich eine markantere Verbesserung des RMS erzielen als kombinierte Prognosen aus bivariaten VAR-Prognosen ($n = 2$). Die Resultate für den Zweijahreshorizont sind etwas befremdend. Dort sind kombinierte Prognosen aus Modellen mit 4 Variablen ($n = 4$) nicht sehr befriedigend.

5.2 Vergleich mit den besten VAR-Prognosen

Die Ergebnisse der Analyse zeigen bis jetzt, dass das Kombinieren von Prognosen mit verschiedenen Modellen im Allgemeinen die Güte der Prognosen verbessern kann. Voraussetzung ist die Wahl einer angemessenen Kombinationsmethode. Dieses Resultat basierte auf einem Vergleich der durchschnittlichen Performance von CVAR-Prognosen mit derjenigen von VAR-Prognosen. Es stellen sich zwei Fragen: Gibt es erstens einzelne VAR-Prognosen, welche den Durchschnitt der kombinierten Prognosen erheblich übertreffen? Ist es zweitens möglich herauszufinden, welche Kombinationen besser abschneiden und ob es von Vorteil wäre, sich auf bestimmte Kombinationen zu konzentrieren? Zur Beantwortung dieser Fragen nehmen wir die Performance der besten VAR- und CVAR-Prognosen etwas genauer unter die Lupe.

Tabelle 6 zeigt die Ergebnisse der drei besten VAR-Prognosen für jeden einzelnen Prognosehorizont. Ein Vergleich mit Tabelle 3 zeigt, dass die besten VAR-Prognosen die durchschnittlichen CVAR-Prognosen beim einjährigen Prognosehorizont übertreffen (alle Gewichtungsmethoden). Bei der zweijährigen Prognose übertreffen nur die zwei besten VAR-Prognosen die durchschnittlichen CVAR-Prognosen, welche mit der CRLS-Methode erstellt wurden. Die beste VAR-Prognose ist jedoch nicht besser als die durchschnittlichen CVAR-Prognosen, welche mit der NRLS-Methode erstellt wurden. Beim dreijährigen Prognosehorizont schneidet die beste VAR-Prognose schlechter ab als der Durchschnitt der CVAR-Prognosen, die entweder mit der CRLS- oder mit der NRLS-Methode erstellt wurden. Man beachte ferner auch, dass die besten VAR-Modelle nicht alle Variablen beinhalten, insbesondere bei langfristigen Prognosehorizonten. Die Ergebnisse in Tabelle 6 zeigen auf, dass das Kombinieren von Prognosen besonders für langfristige Prognosen wichtig ist.

Tabelle 7 vergleicht die besten VAR- mit den drei besten CVAR-Prognosen für jeden Prognosehorizont. Wird die Methode des einfachen Durchschnitts (SA) angewendet, übertreffen die besten CVAR-Prognosen die besten VAR-Prognosen beim Einjahreshorizont um 9%. Für längere Prognosehorizonte verbessert sich der RMSE der besten CVAR-Prognose (CRLS) allerdings um über 30%. Dies ist eine erhebliche Verbesserung der Prognosequalität. Beim Prognosehorizont von zwei (drei) Jahren erreichen 59 (49) von 79 möglichen CVAR-Prognosen tiefere RMSE und U-Werte als die beste VAR-Prognose.

Ist es möglich, die beste Kombination zu bestimmen? Interessanterweise beinhalten die drei besten Modelle für jeden der Prognosehorizonte ausnahmslos sämtliche fünf in dieser Analyse betrachteten Variablen. Man beachte ferner auch das folgende interessante Resultat: Für den einjährigen Prognosehorizont stützen sich nur drei von fünfzehn möglichen Kombinationen (Anzahl zweier Prognosen aus trivariaten VAR-Modellen) auf Informationen aus allen Variablen. Die beste CVAR-Prognose ist eine dieser drei Kombinationen. Dies könnte nicht nur auf einen «Diversifikationsvorteil», sondern auch auf einen «Informationsvorteil» von kombinierten Prognosen gegenüber einzelnen Prognosen hindeuten. Ein praktischer Vorschlag wäre, VAR-Prognosen so zu kombinieren, dass eine grosse Anzahl von Variablen in der CVAR-Prognose berücksichtigt wird.

6 Schlussbemerkungen

Unbedingte Inflationsprognosen sind für die Durchführung der Geldpolitik von grosser Bedeutung. VAR-Modelle eignen sich gut für die Erstellung solcher unbedingter Prognosen. Da jedoch bei makroökonomischen Studien in der Regel nur kurze Zeitreihen verfügbar sind, haben VAR-Modelle den Nachteil, dass sie sich auf einige wenige Variablen beschränken müssen. Wir haben einen Modellansatz zur Berechnung unbedingter Prognosen entwickelt, der das Problem einer begrenzten Anzahl Variablen überwindet. Die Methode macht sich die Eigenschaften kombinierter Prognosen zu Nutze und besteht aus zwei Schritten: Zuerst wird eine grosse Anzahl VAR-Prognosen berechnet, die von unterschiedlich spezifizierten Modellen stammen und verschiedene Variablen beinhalten. In einem zweiten Schritt werden die Gewichte für die Kombination der Prognosen aufgrund ihrer Güte in der Vergangenheit bestimmt. Dann werden die unbedingten kombinierten Prognosen mittels dieser Gewichte berechnet. Mit diesem Ansatz lassen sich «Echtzeit»-Informationen über die Prognosegüte und mögliche Strukturbrüche aus Gewichtsveränderungen ableiten. Ausserdem können die Gewichte zeigen, welche Gruppe von Variablen Informationen über die zukünftige Inflation für einen bestimmten Prognosehorizont enthalten.

Die Erkenntnis dieses Aufsatzes ist, dass die mit dem entwickelten Ansatz berechneten kombinierten Inflationsprognosen im Durchschnitt die beste VAR-Prognose übertreffen. Dies ist insbesondere bei langfristigen Prognosen der Fall. Die Überlegenheit kombinierter Prognosen lässt sich auf drei Gründe zurückführen: Erstens sind die Prognosefehler diversifiziert. Zweitens stützen sich kombinierte Prognosen nicht schwergewichtig auf die Spezifikation eines einzelnen VAR-Modells und reagieren deshalb unter Umständen weniger anfällig auf Spezifikationsfehler. Drittens beruhen kombinierte VAR-Prognosen im Gegensatz zu einzelnen VAR-Prognosen üblicherweise auf mehr Informationen.

Die vorliegende Studie ist mit einigen Mängeln behaftet. Um den Rechen- und Programmierungsaufwand in Grenzen zu halten, haben wir uns in zweierlei Hinsicht eingeschränkt. Einerseits haben wir nur fünf Variablen verwendet. In der Literatur zum monetären Transmissionsmechanismus wird die Bedeutung weiterer Variablen wie z. B. Wechselkurs, Importpreise, verschiedene Zinssätze etc. hervorgehoben. Andererseits haben wir nicht untersucht, ob die Resultate auch dann zutreffen, wenn andere Zeitreiheneigenschaften für die Daten angenommen werden. Dies würde die Anzahl der zu berücksichtigenden VAR- und CVAR-Prognosen drastisch ansteigen lassen, da VAR mit Niveauvariablen und Fehlerkorrektur-Modelle mit einbezogen werden müssten. Da eine ineffiziente (und unverzerrte) Prognose immer verbessert werden kann, wenn man sie mit einer anderen Prognose kombiniert, erwarten wir, dass die Prognosegüte durch eine Lockerung dieser beiden Restriktionen in einem gewissen Ausmass verbessert werden könnte. Die interessante Frage in dieser Hinsicht ist jedoch, ob diese Verbesserung so gross wäre, wie die in dieser Studie aufgezeigte. Schliesslich soll noch eine grundsätzliche Beschränkung der vorgestellten Methode erwähnt werden: Kombinierte VAR-Prognosen eignen sich nur dann, wenn unbedingte Prognosen benötigt werden. Bei strukturellen Simulationen (Impuls-Response, Varianzdekomposition, bedingte Prognosen) muss auf individuelle VAR-Prognosen abgestellt werden.

Zusammenfassend macht dieser Aufsatz deutlich, dass der Beizug vieler «kleiner» VAR-Modelle eine bessere Strategie zur Erstellung einer unbedingten Prognose sein kann als die Verwendung eines einzigen VAR-Modells. Gemäss unseren empirischen Ergebnissen kann dies insbesondere für langfristige Inflationsprognosen der Fall sein. Mit Bezug auf das Anfangszitat von Shakespeare ziehen wir folgende Schlussfolgerung: Wenn man nicht sagen kann, welches Korn spriessen wird, so ist es ratsam, nicht bloss ein einziges Korn zu wählen, sondern gleich das ganze Saatgut zu streuen. Wenn es um Prognosen und nicht um Analysen geht, zeigen dieser Aufsatz und die umfangreiche Literatur zu kombinierten Prognosen, dass es besser ist, mehr als ein einziges Modell zu verwenden. Die Inflationsprognosen der SNB basieren denn auch auf mehreren Modellen. Die Ergebnisse dieser Studie tragen dazu bei, diesen pluralistischen Ansatz zu untermauern.

7 Literaturverzeichnis

Aksu C. und Gunter, S.I. 1992. An empirical analysis of the accuracy of SA, OLS, ERLS, and NRLS combination forecasts. *International Journal of Forecasting* 8: 27–43.

Baltensperger, E., Fischer A.M. und Jordan T.J. 2002. Abstaining from inflation targets, Understanding SNB rhetoric in the inflation targeting debate. Arbeitspapier, Schweizerische Nationalbank.

Baltensperger, E., Jordan T.J. und Savioz M.R. 2001. The demand for M_3 and inflation forecasts: An empirical analysis for Switzerland. *Weltwirtschaftliches Archiv/Review of World Economics* 137(2): 244–272.

Bernanke, B.S. 1990. On the predictive power of interest rates and interest rate spreads. *New England Economic Review* November-December: 51–61.

Clemen, R.T. 1986. Linear constraints and the efficiency of combined forecasts. *Journal of Forecasting* 5: 31–38.

Clemen, R.T. 1989. Combining forecasts: A review and annotated bibliography, *International Journal of Forecasting* 8: 559–83.

Clemen, R.T. und Winkler, R.L. 1986. Combined economic forecasts. *Journal of Business and Economic Statistics* 4: 39–46.

Diebold, F.-X. 1989. Forecast combination and encompassing: Reconciling two divergent literatures. Board of Governors of the Federal Reserve System, Finance and Economics Discussion Series.

Diebold, F.-X. und Pauly, P. 1986. Structural change and the combination of forecasts. Boards of Governors of the Federal Reserve System, Special Studies Section, Discussion Paper 201.

Granger, C.W.J. 1986. *Forecasting Economic Time Series*. London: Academic Press.

Granger, C.W.J. 1989. *Forecasting in Business and Economics*. London: Academic Press.

Granger, C.W.J. und Newbold, P. 1973. Some comments on the evaluation of economic forecasts. *Applied Economics* 5: 35–47.

Granger, C.W.J. und Newbold, P. 1986. *Forecasting Economic Time Series*. Zweite Auflage, London: Academic Press.

Granger, C.W.J. und Ramanathan, R. 1984. Improved methods of combining forecasts. *Journal of Forecasting* 14(3): 367–79.

Greene, W.H. 2000. *Econometric analysis*. Vierte Auflage, London: Prentice-Hall.

Holden, K. und Peel, D.A. 1986. An empirical investigation of combinations of economic forecasts. *Journal of Forecasting* 5: 229–42.

Holden, K. und Peel, D.A. 1989. Unbiasedness, efficiency and the combination of economic forecasts. *Journal of Forecasting* 8: 175–882.

Jordan, T.J. 1999a. The information content of bank credit to forecast output and inflation: The case of Switzerland. Arbeitspapier, Schweizerische Nationalbank.

Jordan, T.J. 1999b. Inflationsprognosen mit VAR-Systemen. Arbeitspapier, Schweizerische Nationalbank.

Jordan, T.J. und Peytrignet, M. 2001. Die Inflationsprognosen der Schweizerischen Nationalbank, La prévision d'inflation de la Banque nationale suisse. *Schweizerische Nationalbank Quartalsheft* 2: 54–61.

Jordan, T.J., Kugler, P., Lenz, C. und Savioz, M.R. 2002. Inflationsprognosen mit vektorautoregressiven Modellen, Prévisions d'inflation par des modèles vectoriels autorégressifs. *Schweizerische Nationalbank Quartalsheft* 1: 40–66.

Jungmittag, A. 1993. Die Kombination von Prognosen: Ein Überblick mit Anwendungen, *Jahrbücher für Nationalökonomie und Statistik* 212(1–2): 58–72.

Kirchgässner, K. und Savioz, M.R. 1997. Empirische Forschung in den Wirtschaftswissenschaften: Ein Überblick. *Homo oeconomicus* 16: 507–538.

Kirchgässner, K., und Savioz, M.R. 2001. Monetary policy and forecasts for real GDP growth: An empirical investigation for the Federal Republic of Germany. *German Economic Review* 2: 339–365.

Kugler, P. und Jordan, T.J. 2000. Structural vector autoregressions and the analysis of monetary policy interventions: The Swiss case. Arbeitspapier, University of Basel.

Miller, C.M., Clemen, R.T. und Winkler, R.L. 1992. The effect of nonstationary on combined forecasts. *International Journal of Forecasting* 7(4): 515–529.

Ruoss, E., und Savioz, M.R. 2002. How accurate are GDP forecasts – a survey for Switzerland. *Schweizerische Nationalbank Quartalsheft* 3: 43–63.

Sims, C. 1980. Macroeconomics and reality. *Econometrica* 48: 1–49.

Stalder, P. 2001. Ein ökonometrisches Makromodell für die Schweiz, Un modèle macroéconomique pour la Suisse. *Schweizerische Nationalbank Quartalsheft* 2: 63–89.

Thoma, M.A. und Gray, J.A. 1994. On leading indicators: Getting it straight. 1994 Federal Reserve Bank of Dallas *Texas conference on monetary economics Paper No. 4*.

West, K.D. 2001. Tests for forecast encompassing when forecasts depend on estimated regression parameters. *Journal of Business & Economic Statistics* 19(1): 29–33.

Winkler, R.L. 1989. Combining forecasts: A philosophical basis and some current issues. *International Journal of Forecasting* 5: 605–609.

Anhang: Tabellen

Erweiterter Dickey-Fuller-Einheitswurzeltest 1974:1 – 2000:3 Tabelle 1

Variable	k	r	t
ΔP	5	0.699	−2.699(*)
ΔQ	10	0.058	−3.17*
ΔM	4	0.570	−2.845(*)
ΔC	0	0.833	−2.905*
ΔR	0	0.412	−6.736**

* zeigt, dass die Nullhypothese einer Einheitswurzel beim Signifikanzniveau von 5% verworfen wird. ** und (*) zeigen eine Verwerfung beim Signifikanzniveau von 1% und 10%. k ist die Anzahl verzögerter (endogener) Variablen, welche in die erweiterte Dickey-Fuller-Testgleichung eingehen. k ist der Lag zwischen 0 und 10 mit dem kleinsten Wert des AIC-Kriteriums. r ist die geschätzte Einheitswurzel und t die Teststatistik. Es werden die kritischen Werte von MacKinnon benutzt.

Individuelle Out-of-Sample-VAR-Prognosen
Durchschnittliche Resultate der 16 VAR-Prognosen Tabelle 2

Anzahl der in den VAR enthaltenen Variablen	Einjahresprognosen 1991:1 – 2000:3 39 Prognosen		Zweijahresprognosen 1993:1 – 2000:3 31 Prognosen		Dreijahresprognosen 1995:1 – 2000:3 23 Prognosen	
	RMSE	Theil's U	RMSE	Theil's U	RMSE	Theil's U
Alle VAR (16)[1]	*0.892*	*0.340*	*1.259*	*0.817*	*1.418*	*1.302*
n = 1 (1)	1.198	0.457	1.365	0.886	1.464	1.345
n = 2 (4)	0.966	0.369	1.252	0.812	1.417	1.301
n = 3 (6)	0.860	0.328	**1.242**	**0.806**	**1.402**	**1.287**
n = 4 (4)	**0.801**	**0.306**	1.245	0.808	1.425	1.309
n = 5 (1)	0.842	0.321	1.340	0.869	1.445	1.328

Anmerkung:
Die beste Statistik ist für jeden Prognosehorizont fett dargestellt. n ist die Anzahl der Variablen in einem VAR. Alle VAR sind von der Ordnung 4. Für n = 1 werden zum Beispiel die Prognosen mit einem AR(4) Modell berechnet. Für n = 2 (n = 3) werden die Prognosen mit einem bivariaten (trivariaten) VAR(4)-Modell berechnet, usw. Das Sample zur Schätzung der VAR beginnt mit 30 Beobachtungen und wird erweitert, bis es 50 Beobachtungen erreicht.

1 Die Anzahl VAR mit n Variablen wird in Klammern angegeben. «n = 3 (6)» heisst zum Beispiel, dass sechs Prognosen mit sechs trivariaten VAR(4) errechnet wurden. Das in der Tabelle angegebene Resultat entspricht dem Durchschnitt der sechs Prognosen.

Kombinierte Out-of-Sample-VAR-Prognosen
Durchschnittliches Resultat für die 79 kombinierten Prognosen

Tabelle 3

Methode	Einjahresprognosen 1991:1 – 2000:3 39 Prognosen			Zweijahresprognosen 1993:1 – 2000:3 31 Prognosen			Dreijahresprognosen 1995:1 – 2000:3 23 Prognosen		
	RMSE	*U*	%	RMSE	*U*	%	RMSE	*U*	%
VAR-Prognose	*0.892*	*0.340*	*100%*	*1.259*	*0.817*	*100%*	*1.418*	*1.302*	*100%*
SA	**0.749**	**0.286**	−15,9%	1.167	0.757	−7,3%	1.374	1.262	−3,1%
LS	1.069	0.408	20%	1.714	1.112	36,1%	2.528	2.322	78,3%
CRLS	0.942	0.360	5,9%	1.142	0.741	−9,3%	**1.261**	**1.158**	−11,1%
ERLS	0.868	0.331	−2,6%	1.302	0.844	3,3%	1.552	1.426	9,5%
NRLS	0.904	0.345	1,5%	**1.088**	**0.706**	−13,6%	1.268	1.165	−10,5%

Anmerkung:
VAR(4) und Kombinationen werden anhand von 50 Beobachtungen geschätzt.

Untergruppen kombinierter VAR-Prognosen

Tabelle 4

Anzahl Variablen (n) im VAR	Anzahl Prognosen, die kombiniert wurden (m):				
	$m = 2$	$m = 3$	$m = 4$	$m = 5$	$m = 6$
$n = 2$	6	4	1	–	–
$n = 3$	15	20	15	6	1
$n = 4$	6	4	1	–	–

Anmerkung:
n ist die Anzahl Variablen in den VAR. *m* ist die Anzahl der zu einer kombinierten Prognose zusammengesetzter Einzelprognosen.

Kombinierte Out-of-sample-VAR-Prognosen (Untergruppen)

Durchschnittliches Resultat für verschiedene Untergruppen der 79 kombinierten Prognosen — Tabelle 5

	Einjahresprognosen 1991:1 – 2000:3 (39 Prognosen) SA			Zweijahresprognosen 1993:1 – 2000:3 (31 Prognosen) CRLS			Dreijahresprognosen 1995:1 – 2000:3 (23 Prognosen) CRLS		
	RMSE	U	Gewinn	RMSE	U	Gewinn	RMSE	U	Gewinn
VAR-Prognose	*0.892*	*0.340*	*100%*	*1.259*	*0.817*	*100%*	*1.418*	*1.302*	*100%*
CVAR-Prognose	*0.749*	*0.286*	*−15,9%*	*1.142*	*0.741*	*−9,3%*	*1.261*	*1.158*	*−11,1%*
Untergruppen gebildet aus der Anzahl zu kombinierten VAR-Prognosen zusammengesetzter Einzelprognosen									
m = 2 (27)	0.782	0.298	−12,4%	1.208	0.784	−4,0%	1.327	1.219	−6,4%
m = 3 (28)	0.745	0.284	−16,5%	1.179	0.765	−6,4%	1.264	1.161	−10,8%
m = 4 (17)	0.721	0.275	−19,1%	1.080	0.701	−14,2%	1.201	1.104	−15,2%
m = 5 (6)	0.705	0.269	−20,9%	0.914	0.593	−27,4%	1.147	1.054	−19,0%
m = 6 (1)	**0.698**	**0.266**	−21,8%	**0.783**	**0.508**	−37,9%	**1.100**	**1.011**	−22,4%
Untergruppen gebildet aus der Anzahl der ins VAR-Modell einfliessenden Variablen									
n = 2 (11)	0.869	0.332	−2,4%	1.389	0.901	−10,3%	1.285	1.181	−9,3%
n = 3 (57)	0.734	0.280	−17,6%	**1.073**	**0.696**	−14,9%	1.258	1.156	−11,2%
n = 4 (11)	**0.706**	**0.270**	−20,6%	1.256	0.815	−0,2%	**1.251**	**1.149**	−11,8%

Anmerkung:
m ist die Anzahl der zu einer kombinierten Prognose zusammengesetzter Einzelprognosen.

Anmerkung:
n ist die Anzahl Variablen im VAR-Modell.

Die besten individuellen Prognosen

Tabelle 6

Rang	VAR	RMSE	Theil's U
	Einjahresprognosen: 1991:1 – 2000:3 (39 Prognosen)		
1	*P, M, C, R*	0.716	0.273
2	*P, M, C, Q*	0.732	0.280
3	*P, C, R*	0.742	0.283
	Zweijahresprognosen: 1993:1 – 2000:3 (31 Prognosen)		
1	*P, M, C, Q*	1.088	0.706
2	*P, M, C*	1.128	0.732
3	*P, M, Q*	1.152	0.748
	Dreijahresprognosen: 1995:1 – 2000:3 (23 Prognosen)		
1	*P, C, R*	1.318	1.210
2	*P, C, Q, R*	1.351	1.241
3	*P, C*	1.358	1.247

Anmerkung:
VAR(4) und Kombinationen
werden anhand von
50 Beobachtungen geschätzt.

Die besten kombinierten VAR-Prognosen

Tabelle 7

Rang	CVAR	RMSE	Theil's U	Gewinn
	Einjahresprognosen (SA)			
	Beste VAR-Prognose	*0.716*	*0.273*	*100%*
1	*P,M,C + P,Q,R*	0.650	0.248	−9,2%
2	*P,M,C + P,C,R + P,Q,R*	0.658	0.251	−8,1%
3	*P,M,C,Q + P,M,C,R + P,M,Q,R*	0.662	0.252	−7,7%
	Zweijahresprognosen (CRLS)			
	Beste VAR-Prognose	*1.088*	*0.817*	*100%*
1	*P,M,C + P,M,Q + P,M,R + P,C,Q + P,C,R + P,Q,R*	0.783	0.508	−38,8%
2	*P,M,C + P,M,R + P,C,Q + P,C,R + P,Q,R*	0.787	0.511	−37,5%
3	*P,M,C + P,M,R +P,C,R + P,Q,R*	0.817	0.530	−35,1%
	Dreijahresprognosen (CRLS)			
	Beste VAR-Prognose	*1.318*	*1.210*	*100%*
1	*P,M,C + P,C,R + P,Q,R*	0.921	0.846	−30,1%
2	*P,M,C + P,C,R + P,Q,R + P,M,R*	0.951	0.874	−27,8%
3	*P,M,C + P,C,R + P,Q,R + P,M,Q*	0.952	0.874	−27,8%

Anmerkung:
VAR(4) und Kombinationen
werden anhand von
50 Beobachtungen geschätzt.