



Diskussionspapiere
Discussion Papers

Diskussionspapier Nr. 73

Wie gut sind Prognosen?
Eine Überprüfung der Vorhersagen der Gemeinschafts-
diagnose und des DIW am Beispiel des BSP

von

Thiess Büttner u. Gustav A. Horn*

22. FEB. 1995 Weltwirtschaft
Kiel

W 620 (73) ~~ni~~ sig ka

Deutsches Institut für Wirtschaftsforschung, Berlin
German Institute for Economic Research, Berlin

Die in diesem Papier vertretenen Auffassungen liegen ausschließlich in der Verantwortung des Verfassers und nicht in der des Instituts.

Opinions expressed in this paper are those of the author and do not necessarily reflect views of the Institute.

Deutsches Institut für Wirtschaftsforschung

Diskussionspapier Nr. 73

Wie gut sind Prognosen? Eine Überprüfung der Vorhersagen der Gemeinschafts- diagnose und des DIW am Beispiel des BSP

von
Thiess Büttner u. Gustav A. Horn*

- *) Thiess Büttner, Universität Konstanz
Gustav A. Horn, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

Deutsches Institut für Wirtschaftsforschung, Berlin
Königin-Luise-Str. 5, 1000 Berlin 33
Telefon: 49-30 - 82 991-0
Telefax: 49-30 - 82 991-200

1. Einleitung

Der Frage nach der Güte einer Prognose begegnen viele -insbesondere die Prognostiker selbst - mit unverhohlener Skepsis. Eine sinnvolle Überprüfung von Prognosen sei prinzipiell nicht möglich, zum einen weil sich die Annahmen z.B. über das Verhalten der wirtschaftspolitischen Instanzen, auf der die Vorhersagen basieren häufig als falsch erweisen; zum anderen sei der Informationsstand zum Zeitpunkt der Prognose über die Vergangenheit aufgrund üblicher Revisionen von Daten unsicher und lasse sich im nachhinein auch nicht mehr rekonstruieren. Die Beurteilung einer Prognose erfolge daher zwangsläufig auf einer verzerrten Basis.

Diese Argumente überzeugen allenfalls teilweise und schon gar nicht als prinzipielles Hindernis. Eine Vorhersage mit wissenschaftlichen Anspruch muß aus Prinzip, gerade um der Forderung nach Wissenschaftlichkeit zu genügen, überprüfbar sein. Die im nachhinein falsche Setzung von Annahmen ist hierfür kein Gegenargument. In einem wohldefinierten Modell, in dem exakt zwischen Annahmen über z.B. wirtschaftspolitische Setzungen und endogenen Entwicklungen unterschieden werden kann, läßt sich der Fehler, der ausschließlich von falschen Annahmen herrührt, ermitteln. Ein - primär methodisches - Problem besteht freilich darin, daß in der Regel die Prognosen nicht auf einem exakten Modell basieren, sondern eher auf intuitiven Überlegungen z.B. über Konjunkturverläufe. Dann ist allerdings eine Unterscheidung zwischen Annahmen und endogenen Entwicklungen nicht mehr möglich. Aber unabhängig davon ist zu fragen, ob die Setzung falscher Annahmen nicht ohnehin den Prognostikern anzulasten ist. Wenn es nämlich das Ziel der Vorhersage ist, ein möglichst genaues Bild der zukünftigen Entwicklung zu entwerfen, dann gehört hierzu auch die systematisch richtige Einschätzung z.B. des wirtschaftspolitischen Verhaltens.

Stichhaltiger sind die Argumente über den Informationsstand zum Zeitpunkt der Prognose. Jede Beurteilung der Vorhersagegüte muß sich um die Rekonstruktion der Datengrundlage zum Zeitpunkt der Erstellung der Prognose bemühen. Wenn eine Fehlprognose sich gleichsam erst ex post einstellt, weil die Ausgangsdaten im nachhinein revidiert wurden, darf das Ergebnis nicht als eine falsche Vorhersage gewertet werden.

Im folgenden werden in einem ersten Kapitel Kriterien für die Güte einer Prognose aufgestellt. Dabei ist nicht nur eine möglichst hohe Übereinstimmung von Vorhersage und derem Eintreffen von Bedeutung, sondern auch ob die Abweichungen unverzerrt sind und alle verfügbaren Informationen ausgenutzt werden.

Diese Kriterien werden dann auf Prognosen der Gemeinschaftsdiagnose der Arbeitsgemeinschaft wirtschaftswissenschaftlicher Forschungsinstitute (GD) und des Deutschen Instituts für Wirtschaftsforschung (DIW), das auch bei der GD mitwirkt, angewendet. Beide Institutionen veröffentlichen seit Jahren regelmäßig ihre Analysen, so daß eine längerfristig angelegte Untersuchung möglich ist.

2. Kriterien für die Prognosegüte

Bei der Beurteilung der Güte von Prognosen dienen folgende Eigenschaften üblicherweise als Kriterien:

1. Unverzerrtheit
2. Suffizienz
3. Effizienz

Eine Prognose ist unverzerrt, wenn ihr Erwartungswert mit der dann eintretenden Beobachtung übereinstimmt. In diesem Fall entsprechen sich Vorhersage und Beobachtung im Mittel. Eine Prognose ist suffizient, wenn sie alle verfügbaren Informationen voll ausschöpft. Eine Vorhersage ist effizient, wenn sie minimal um die Beobachtungen streut. Eine unverzerrte und suffiziente Prognose wird zudem als rational bezeichnet, Schätzfehler resultieren dann nur noch aus dem Unvermögen, die Realität zu erklären.

In dieser Arbeit wird versucht, Prognosen auf diese Kriterien hin zu überprüfen. Ob das Kriterium der Unverzerrtheit erfüllt ist, kann weitgehend eindeutig anhand der Prognosefehler erkannt werden. Die Überprüfung der Suffizienz setzt freilich eine explizite Festlegung der Menge der zum Prognosezeitpunkt verfügbaren relevanten Informationen voraus. Daher kann in der Regel nicht abschließend geklärt werden, ob dieses Kriterium tatsächlich erfüllt ist. Eine suffiziente Prognose muß jedoch notwendigerweise Informationen aus bekannten Fehlern früherer Vorhersagen ausschöpfen. Ist dies der Fall, bleiben Erkenntnisse, die in Fehlern enthalten sind, nicht ungenutzt. Eine einfache adaptive Prognose, die lediglich eine vorherige Voraussage in Richtung des eingetretenen Vorhersagefehlers korrigiert, wäre hingegen nicht suffizient. Nach diesen Überlegungen ist eine notwendige Bedingung für das Vorliegen von Suffizienz, daß die Vorhersagefehler aus der Vergangenheit nicht miteinander korreliert sind. Im folgenden wird lediglich diese notwendige Bedingung überprüft.

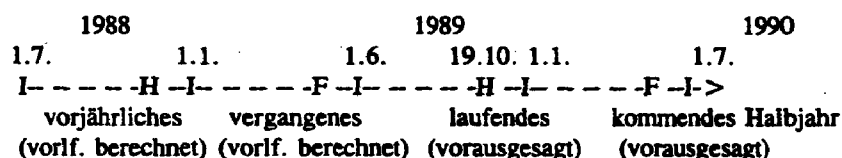
Um ein Maß für die Effizienz einer Prognose zu erhalten, ist eine Referenzprognose erforderlich. Bei makroökonomischen Vorhersagen ist es üblich, hierfür ein univariates ARIMA-Modell zu verwenden. Dies entspricht einer Prognose, die gleichsam ohne jeden ökonomischen Sachverstand allein aus den Vergangenheitswerten einer Zeitreihe ermittelt wird. Die Unterschiede in den Vorhersagefehlern von

Prognosen und Referenzprognose zeigen dann auf, ob und in welchem Ausmaß die auf ökonomischen Modellüberlegungen beruhende Prognose durch die Referenzprognose verbessert werden kann. Ist dies nicht der Fall, heißt die Prognose effizient.

3. Die BSP-Prognosen der Gemeinschaftsdiagnose

3.1 Der Mittelwert des Prognosefehlers

Im Rahmen der Gemeinschaftsdiagnose der fünf großen Wirtschaftsforschungsinstitute in der Bundesrepublik Deutschland werden zweimal jährlich Prognosen für die zukünftige Entwicklung der deutschen und der internationalen wirtschaftlichen Entwicklung aufgestellt. Eine erste Prognose erscheint jeweils im Frühjahr, eine zweite im Herbst eines jeden Jahres. Die veröffentlichten Prognosen des realen Bruttosozialprodukts durch die GD beziehen sich auf das laufende und das jeweils kommende Halbjahr. In den Herbstgutachten wird darüber hinaus sogar eine Vorhersage für das übernächste Halbjahr aufgestellt. Da eine solche Prognose jedoch nicht in den Frühjahrsgutachten veröffentlicht wird, beschränken wir uns aus Gründen der Vergleichbarkeit auf das jeweils laufende und das kommende Halbjahr. Für das jeweilige vergangene Halbjahr liegt bei der Prognose erst eine vorläufige Berechnung des BSPs vor. Diese Rahmenbedingungen lassen sich zum Beispiel für das Herbstgutachten 1989 wie folgt skizzieren:



Die laufenden Vorhersagen basieren somit auf noch unsicheren ex-post-Ergebnissen des Vorquartals und des vergangenen Halbjahres. Ein Grund für Fehlprognosen könnte somit in den zum Zeitpunkt der Prognose noch unvollständigen Informationen bestehen. Es wäre allerdings irreführend, diesen Irrtum anderen Prognosefehlern gleichzusetzen, da er unvermeidbar ist. Bei der Beurteilung der Prognosegüte sollte dem Rechnung getragen werden. Daher wird der Vorhersagefehler hier als Differenz von erster vorläufiger Berechnung und der Voraussage eines Halbjahres definiert. Die Voraussage zum Beispiel von 1989 für die erste Jahreshälfte 1990 würde dann mit der im Herbst 1990 veröffentlichten vorläufigen Berechnung für das BSP dieser Jahreshälfte konfrontiert. Zwar wird diese vorläufige Berechnung gleichfalls nachfolgend oft noch korrigiert, die Gegenüberstellung mit endgültig berechneten Werten würde aber den der Prognoseleistung zurechenbaren Vorhersagefehler überzeichnen. Wegen der lediglich graduellen Unterschiede der Niveauwerte für das BSP und wegen der unterschiedlichen Preisbasis werden die Prognosen anhand der vorhergesagten Veränderungsraten beurteilt. Sie stehen auch im Mittelpunkt der wirtschaftspolitischen Diskussion. Auf der Basis des für das vorangegangene Halbjahr vorläufig berechneten

Wertes, wird folglich eine prognostizierte Wachstumsrate für das laufende Halbjahr ermittelt. Da der Wert für das Vorjahreshalbjahr nicht im Rahmen der GD veröffentlicht wird, greifen wir auf die zum Prognosezeitpunkt vorliegenden Daten der vierteljährlichen Volkswirtschaftlichen Gesamtrechnung des DIW's zurück.¹ Sie sind mit denen bei der Prognose zugrunde gelegten Werten weitgehend identisch. Der Vorhersagefehler ist hier also definiert als Differenz von erster vorläufiger Berechnung und Voraussage der Wachstumsrate eines Halbjahres.

Für die Gemeinschaftsdiagnose liegen Daten für den Zeitraum von 1970 bis 1992 vor. Dies ergibt eine kontinuierliche Reihe von 46 Vorhersagefehlern für die Prognose des jeweils laufenden Halbjahres. Für die Vorhersagen des jeweils kommenden Halbjahres liegen 43 Beobachtungen vor. Das liegt daran, daß im Herbst 1972 und 1974 wegen des Ölpreisschocks eine Ganzjahresprognose veröffentlicht wurde. Da eine kontinuierliche Reihe für die Beurteilung benötigt wird, greifen wir hier ersatzweise auf die Jahresrate zurück. Auf diese Weise erhält man eine Reihe von 45 Vorhersagefehlern.

Unter Prognosefehler wird im folgenden der durchschnittliche quadrierte Vorhersagefehler (DQF) verstanden. Dieser kann in 3 Komponenten zerlegt werden:²

$$DQF = (m_v - m_B)^2 + (S_v - r * S_B)^2 + (1 - r^2) * S_B^2$$

Die erste Komponente ist das Quadrat der Differenz zwischen den Mittelwerten der Vorhersagen und der Beobachtungen. Sie ist Null, wenn eine Regressionsgrade zwischen Vorhersage und Beobachtung durch den Ursprung geht. Die zweite Komponente enthält Informationen über Abweichungen in der Bewegungsrichtung zwischen Vorhersage und Beobachtung. Sie ist dann Null, wenn die Regressionsgrade die Steigung 1 aufweist. Die dritte Komponente enthält Informationen über die Korrelation zwischen Vorhersagen und Beobachtungen. Sie ist Null, wenn die Korrelation total, d.h. das Bestimmtheitsmaß gleich Eins ist. Zerlegt man nun den DQF für die Vorhersage der Gemeinschaftsdiagnose in diese 3 Komponenten, so erhält man folgendes Ergebnis:

¹Siehe die DIW-Veröffentlichungen in den Vierteljahrsheften zur Wirtschaftsforschung.

²Siehe Granger/Newbold (1986).

Tabelle 1

Prognose der Gemeinschaftsdiagnose

	m_v	m_B	r	S_v	S_B
1. Halbjahr	2,45	2,39	0,90	2,12	2,23
2. Halbjahr	2,63	2,31	0,58	1,64	2,20

$m_{v(B)}$ ist der Mittelwert der Vorhersagen (Beobachtungen)

$S_{v(B)}$ Standardabweichungen von Vorhersagen (Beobachtungen)

r ist der Korrelationskoeffizient zwischen Vorhersagen und Beobachtungen

	DQF	A	B	C
1. Halbjahr	0,8771	0,0047	0,0091	0,8660
2. Halbjahr	3,4563	0,1037	0,1340	3,2185

$$A = (m_v - m_B)^2, B = (S_v - r S_B)^2, C = (1 - r^2) S_B^2$$

Die Komponentenzerlegung für die Vorhersagefehler zeigt, daß die Hauptursache für das Entstehen von Abweichungen zwischen Vorhersage und Prognose in einer mangelnden Korrelation begründet liegt. Dabei zeigen sich erhebliche Unterschiede zwischen den beiden Halbjahren. Während das Bestimmtheitsmaß für das jeweils erste Halbjahr gut 0,9 beträgt, sinkt es für das zweite Halbjahr auf weniger als 0,6 ab. Das bedeutet, daß für das zweite Halbjahr weniger als 60% der Varianz der Beobachtung durch die Vorhersage erklärt werden können. Hier zeigt sich die entscheidende Bedeutung der Informationsausstattung für die Prognose. Da für das laufende Halbjahr bereits erste Monatswerte vorliegen, ist hier die Informationsausstattung erheblich besser als für das zweite Halbjahr. Dies zeigt aber auch, daß die Vorhersagen der Gemeinschaftsdiagnose relativ "kurzsichtig" sind. Der Erklärungsgehalt läßt mit der Ausdehnung des Prognosezeitraums rasch nach. Dies läßt sich auch anhand des mittleren Prognosefehlers ($m_v - m_B$) anzeigen. Der mittlere Prognosefehler beträgt für das jeweils laufende Halbjahr 0,06 Prozentpunkte und steigt für das anschließende auf fast 0,32 Prozentpunkte.

3.2 Unverzerrtheit

Die Unverzerrtheit der Prognose für das jeweils laufende Halbjahr kann einfach anhand einer linearen Regression des Vorhersagefehlers auf eine Konstante überprüft werden³. Wenn die Konstante signifikant von Null verschieden ist, muß die Hypothese der Unverzerrtheit verworfen werden. Das zu schätzende Modell ist folglich:

$$VF(t) = c + e(t)$$

$e(t)$ sind unabhängig und identisch verteilte Störvarianten.

Als Ergebnis erhält man:

Koeffizient	Standardfehler	t-Signifikanz
c -0.0687	0.1395	-0.4926
N:46	SEE:0.9461	DW: 1.8050

Die H_0 Hypothese, daß der Koeffizient c sich nicht von Null unterscheidet, also Unverzerrtheit vorliegt, kann somit auf einem Signifikanzniveau von 0.10 % nicht verworfen werden. Die Prognose gilt somit als unverzerrt.

Bei der Überprüfung der Voraussage für das jeweils kommende Halbjahr muß berücksichtigt werden, daß es sich hier um eine zweistufige Prognose handelt. Der Fehler der zuletzt getroffenen Vorhersagen ist zum Zeitpunkt der neuen nicht bekannt. Dies hat zur Folge, daß sich auch unvorhergesehene Störungen in der laufenden Periode auf den Fehler der Prognose für das kommende Halbjahr auswirken. Auf diese Weise pflanzen sich Vorhersagefehler von dem früheren Prognosezeitraum auf den späteren fort. Bezeichnet u_t das Störglied der Regression des Vorhersagefehlers auf eine Konstante $VF_t = c + u_t$, dann sind bei der zweistufigen Prognose die u_t ein gleitender Durchschnitt (Moving Average Prozeß erster Ordnung) der unabhängig und identisch verteilten e_t in der Form $u_t = e_t + be_{t-1}$. Deswegen wurde die Gleichung $VF_t = c + e_t + be_{t-1}$ geschätzt.

³Zur Methodik der Überprüfung von Vorhersagen auf Unverzerrtheit siehe Holdin/Peel (1990).

Die Schätzung liefert folgendes Resultat:

	Koeffizient	Standardfehler	t-Signifikanz
c	-0.3307	0.2618	-1.2620
b	0.3362	0.1446	2.3246
	N:45	See: 1.7561	DW: 1.9133

Die Ergebnisse führen zu der Schlußfolgerung, daß auch für die zweistufige Prognose die Hypothese der Unverzerrtheit nicht verworfen werden kann. Diese Resultate implizieren, daß die Vorhersagen bei der Gemeinschaftsdiagnose keine systematischen Abweichungen von den später tatsächlich zu beobachtenden Werten für das BSP enthalten. Dies bedeutet auch, daß die relativ schwache Korrelation für die zweistufige Prognose nicht auf systematischen Fehlern, wie z.B. einer fortwährenden Unterschätzung der Wachstumsrate, beruht.

3.3. Zur Suffizienz der Prognose

Ob eine Prognose suffizient ist, hängt davon ab, ob die Vorhersagefehler miteinander korreliert sind, also nicht alle verfügbaren Informationen ausgenutzt werden. Diese Überprüfung kann für die einstufige Prognose des laufenden Jahres mittels der Durbin/Watson-Prüfgröße (DW) durchgeführt werden. Da die Konstante bei der obigen Schätzung nicht signifikant war, erfolgt die Berechnung direkt durch die Regression des Vorhersagefehlers auf die Störterme; es ergibt sich ein DW von 1.7953, sodaß die Hypothese der Suffizienz nicht verworfen werden kann. Mit der Ljung/Box-Prüfgröße⁴ kann außerdem überprüft werden, ob zugleich auch keine Autokorrelationen höherer Ordnung vorliegen. Als Prüfgröße ergab sich für 24, 20 und 16 Lags mit 15.20, 14.11 und 7.39. Die jeweiligen kritischen Werte lauten für ein Signifikanzniveau von 10% [15.659;...], [12.443;...] und [9.312;...]. Die H₀-Hypothese fehlender Autokorrelation kann also nur für eine maximale Verzögerung von 20 verworfen werden. Zusammengekommen lassen diese Ergebnisse nicht auf fehlende Suffizienz schließen. Da wir ebenso Unverzerrtheit fanden, ist die Prognose für das laufende Halbjahr also als bedingt rational zu beurteilen - bedingt, weil nur eine notwendige Bedingung zur Suffizienz überprüft werden konnte.

⁴Die Prüfgröße $Q = n \cdot (n+2) \cdot (r_1^2/(n-1) + \dots + r_i^2/(n-i) + \dots + r_m^2/(n-m))$ ist CHI-Quadrat verteilt mit m-p-q Freiheitsgraden.

Wobei n die Zahl der Beobachtungen, r_i der i-te Korrelationskoeffizient der Residuen des betrachteten ARMA(p,q) Modells und m die Zahl der Verzögerungen (Lags) ist, die berücksichtigt werden.

Bei der zweistufigen Prognose wird zunächst auf Autokorrelation zweiter Ordnung geprüft, da ihre Signifikanz ein Hinweis auf fehlende Suffizienz wäre. Da die Voraussage unverzerrt ist, kann auch hier die Konstante in der Schätzung vernachlässigt werden. Das zu schätzende Modell ist:

$VF(t) - a \cdot VF(t-2) = e(t) + b \cdot e(t-1)$. Das Ergebnis der Schätzung lautet:

	Koeffizient	Standardfehler	t-Signifikanz
a	0.0401	0.1504	0.2665
b	0.3720	0.1489	2.4982
	N: 43 ⁴	DW: 1.9476	SEE: 1.8079

Während sich die Schätzung in Bezug auf den Koeffizienten des Moving-Average-Prozeß kaum von der ersten Schätzung unterscheidet, ist der zusätzlich in die Schätzgleichung aufgenommene autoregressive Prozeß nicht signifikant von Null verschieden. Er trägt somit nichts zur Erklärung des Vorhersagefehlers bei.

Analog zur Untersuchung der einstufigen Prognose kann auch für diese letzte Schätzung ein Ljung/Box-Test durchgeführt werden, ob die Residuen des Moving-Average-Prozesses nicht autokorreliert sind. Es ergab sich bei 24, 20 und 16 Lags für Q: 11.52, 10.56 und 8.81 (Ablehnungsbereich für ein Niveau von 10% [14.848;...], [11.651;...] und [8.547;...]) sodaß die Hypothese fehlender Autokorrelation auf einem Niveau von 10% nur im Falle von 16 Lags nicht gesichert ist. Die Resultate lassen somit darauf schließen, daß wohl alle verfügbaren Informationen aus Vorhersagefehlern ausgenutzt werden. Insgesamt kann damit auch die Prognose des jeweils kommenden Halbjahres als bedingt rational eingestuft werden.

3.4 Effizienz

Basierend auf Daten für das Bruttosozialprodukt in Preisen von 1985 wurde zunächst eine Referenzprognose mittels eines ARIMA-Modells erstellt. Die Identifikation des adäquaten Modells erfolgte auf der Basis der Quartalswerte der vierteljährlichen volkswirtschaftlichen Gesamtrechnung des DIW über den gesamten Zeitraum von 1960 bis 1992.⁵ Die Referenzprognose läßt sich nach dem Verfahren von Box/Jenkins auf der Basis einer sukzessiven Schätzung der Parameter des identifizierten Modells ermitteln. Dieses Vorgehen unterstellt, daß die Struktur der Zeitreihe im Zeitverlauf qualitativ konstant ist, so daß die Identifikation eindeutig ist. Diese Untersuchung führte zur Answahl eines ARIMA(4,1,0) Modells:

⁵vgl. DIW: lange Reihen der vierteljährlichen volkswirtschaftlichen Gesamtrechnung für die Bundesrepublik Deutschland Januar 1993 (für 1960-1980) und April 1993 (für 1980-1992)

$$(1 - a_1B^1) (1 - a_3B^3) (1 - a_4B^4) (1 - B^4)y_t = c + e_t^6$$

Für den Zeitraum von 1970-1992 wurden Vorhersagen für die jeweils kommenden vier Quartale auf der Basis von Quartalswerten der vorangegangenen zehn Jahre mit Hilfe dieses ARIMA-Modells errechnet. Die jeweils letzten vier Quartalswerte einer Schätzung wurden durch die zum Prognosezeitpunkt verfügbaren vorläufigen Berechnungen ersetzt, da die zu beurteilenden Prognosen ja ebenfalls auf der Basis vorläufiger Werte errechnet wurden. Hier wurden, wie schon bei der Ermittlung der Raten, die zum Prognosezeitpunkt vorliegenden Daten aus den Vierteljahrsheften zur Wirtschaftsforschung des DIW verwendet. Um den Unterschied in der Preisbasis zu korrigieren, berechneten wir mittels der vorläufigen Werte Zuwachsraten auf die Vorjahreswerte. Aus den Quartalswerten lassen sich dann Halbjahreswerte errechnen und mit den vorläufigen Werten des jeweiligen Vorjahrs in vorausgesagte Wachstumsraten transformieren. Der Vorhersagefehler der Referenzprognose wurde wie oben durch Differenz mit den ersten vorläufigen Berechnungen der Raten errechnet.

Diese Referenzprognose hat schlechtere Eigenschaften als die Prognosen der GD: Der durchschnittliche quadrierte Vorhersagefehler ist für die Prognosen beider Halbjahre größer, nämlich 2.1115 bzw. 6.3363. die Korrelation zwischen beobachteter und vorhergesagter Wachstumsrate ist deutlich geringer.⁷ Überprüft man auch die Referenzprognose auf Unverzerrtheit, so erhält man folgendes Ergebnis für das Modell:

$$RVF(t) = c + e(t)$$

Koeffizient	Standardfehler	t-Signifikanz
c: -0.0786	0.2163	-0.3635
N: 46	SEE: 1.4670	DW: 2.0348

Für die zweistufige Referenzprognose ergibt sich für die Regressionsgleichung:

$$RVF(t) = c + e(t) + b \cdot e(t-1)$$

⁶B ist der Lagoperator $B^k X_t = X_{t-k}$.

⁷Wenn wie oben der Mittelwert des Prognosefehlers in seine Komponenten zerlegt wird, zeigt sich, daß erneut die mangelnde Korrelation den Hauptanteil ausmacht:

Koeffizient	Standardfehler	t-Signifikanz
c: -0.2366	0.3359	-0.7044
b: 0.7939	0.0908	8.7453
N: 45	SEE: 2.2535	DW: 2.4158

Die Konstante ist somit in beiden Regressionen nicht gegen Null gesichert. Die Vorhersagen sind folglich unverzerrt.

Bildet man eine kombinierte Vorhersage, $B(t)$, aus der Prognose durch die Gemeinschaftsdiagnose, V , und der Referenzprognose, RV , so kann deren Effizienz ermittelt werden, indem man eine Gewichtung der beiden Prognosen so errechnet, daß der Vorhersagefehler ex post minimiert wird. Die optimale Gewichtung k der kombinierten Vorhersage $k \cdot V + (1-k) \cdot RV$, wird durch Anwendung der Methode der kleinsten Quadrate auf die Gleichung: $B(t) = k \cdot V(t) + (1-k) \cdot RV(t)$ bestimmt, wobei B der tatsächliche Beobachtungswert ist. Einfacher noch ist die Regression: $(B(t) - RV(t)) = k \cdot (V(t) - RV(t)) + e(t)$

Die Schätzung ergab:

	Koeff.	Std.fhl.	t-Sign.	t(k=1)	SEE	R ²	N
1. Halbj.	1.0532	0.1341	7.855	0.3967	0.9576	0.5828	46
2. Halbj.	0.9543	0.1578	6.0497	-0.2896	1.8910	0.4552	45

Für beide Prognosen kann die Hypothese, daß k sich nicht von eins unterscheidet nicht verworfen werden. Die relative Effizienz der GD-Prognose kann also in beiden Fällen in der Nähe von 100% eingeschätzt werden. Das heißt, die Berücksichtigung der ARIMA-Prognose kann die GD-Prognose nicht verbessern. Das höhere Bestimmtheitsmaß der einstufigen Prognose spiegelt den Umstand wider, daß die kurzfristige Prognose größere Erklärungskraft hat.

Die reale Sozialproduktsprognose der GD für das jeweils laufende und kommende Halbjahr ist somit unverzerrt und bedingt suffizient, genügt also notwendigen Bedingungen der Rationalität. Gegenüber einer ARIMA-Referenzprognose ist die Prognose effizient. Ähnliche Ergebnisse fand auch Kirchgässner (1983) in seiner umfassenderen Beurteilung der Prognoseleistung der GD.

4. Die Prognose des realen Sozialprodukts durch das DIW

4.1 Der Mittelwert des Prognosefehlers

Das DIW veröffentlicht seit Anfang der 70er Jahre Vorausschätzungen zur Jahreswende für das jeweils vergangene, das nächste und übernächste Halbjahr in den "Grundlinien der Wirtschaftsentwicklung". Seit 1982 werden auch zur Jahresmitte Vorausschätzungen für das jeweils vergangene und das jeweils kommende Halbjahr in den "Tendenzen der Wirtschaftsentwicklung" herausgegeben. Die beiden Veröffentlichungen bilden zusammengekommen eine ähnliche Reihe von Vorhersagen wie die Frühjahrs- und Herbstgutachten der GD. Es kann also eine Schätzung für das jeweils vergangene und eine Prognose für das jeweils kommende Halbjahr untersucht werden.

Der Fehler für die Schätzung des vergangenen Halbjahres wird wie oben als Differenz von erster vorläufiger Berechnung und der Schätzung der Wachstumsrate definiert. Die Voraussagen aus den "Grundlinien" und "Tendenzen" wurden also mit den im nächsten respektive übernächsten Frühjahrs- und Herbstgutachten ermittelten ersten vorläufig berechneten Wachstumsraten der Halbjahre konfrontiert. Entsprechend wurde der Prognosefehler für die Rate des kommenden Halbjahres gebildet.

Die Stichprobe umfaßt den Zeitraum von 1970 bis 1992. Für die Jahre 1971-1973 und 1975-1992 sind "Grundlinien" zur Jahreswende veröffentlicht worden, und in den Jahren 1982 bis 1992 "Tendenzen" in der Jahresmitte. Es werden 33 Vorausschätzungen für das jeweils vergangene und 32 für das jeweils kommende Halbjahr betrachtet. Die jeweiligen kontinuierlichen Prognosereihen haben dann 23 respektive 22 Beobachtungen.

Tabelle 2

Prognosen des DIW

	m_v	m_B	r	S_v	S_B
1. Halbjahr	2,00	2,21	0,97	1,83	1,82
2. Halbjahr	2,08	2,37	0,71	1,80	2,23

Komponenten des Prognosefehlers

	DF	A	B	C
1. Halbjahr	0,24	0,04	0,00	0,20
2. Halbjahr	2,66	0,08	0,05	2,53

Die Ergebnisse entsprechen der Erwartung, daß auch hier die geringe Korrelation den Prognosefehler dominiert. Verglichen mit den Sozialproduktsprognosen der GD treten aber geringere Fehler auf. Dies beruht im wesentlichen auf einer besseren Informationsgrundlage, da inzwischen ein Vierteljahr vergangen ist.

4.2 Unverzerrtheit

Der Test auf Unverzerrtheit führte für die einstufige Prognose zu folgenden Ergebnissen:

Regression: $VF(t) = c + e(t)$				
Koeff.	Std.fehler	t-Sign.	SEE	N
c: 0.2093	0.0754	2.777	0.4330	33

Die Konstante ist signifikant von Null verschieden auf einem Niveau von 2%. Es gibt also einen Hinweis auf eine leichte Verzerrtheit in Höhe von einem Fünftel Prozent-Punkt.

Bei der Prognose für das zweite Halbjahr ist wiederum zu unterstellen, daß der Berichtswert über das vergangene Halbjahr nicht vorliegt, und deswegen ein Moving-Average-Prozeß 1.Ordnung angenommen werden muß. Da für diese Hypothese eine kontinuierliche Reihe von Beobachtungen vorliegen muß, stehen nur 22 Werte zur Verfügung. Regression: $VF(t) = c + e(t) + b \cdot e(t-1)$

Koeffizient	Standardfehler	t-Signifikanz
c: 0.3836	0.2542	1.5090
b: 0.1909	0.2269	0.8416
N: 22	DW: 1.8736	SEE: 1.1907

Bei der Prognose des 2. Halbjahres zeigt sich keine signifikante Verzerrung. Der Moving-Average-Prozeß ist nicht signifikant; damit wird die leichte Verfälschung der einstufigen Prognose im Zeitablauf wieder ausgeglichen.

4.3 Suffizienz

Suffizienz bei der Schätzung für das vergangene Halbjahres kann nur vorliegen, wenn der Fehler nicht autokorreliert ist. Die Überprüfung erfolgt erneut nur anhand der kontinuierlichen Reihe. Unter

Beibehaltung der signifikanten Konstanten wurde die Signifikanz eines Autokorrelationsterms überprüft:

Regression: $VF(t) - a \cdot VF(t-1) = c + e(t)$

Koeffizient	Standardfehler	t-Signifikanz
c: 0.2326	0.0554	4.1973
a: -0.4720	0.1977	-2.3870
N: 22	DW: 1.7956	SEE: 0.3826

Die Konstante bleibt weiterhin signifikant. Für die Residuen der Regression auf die Konstante zeigt sich auch beim Ljung/Box-Test, daß Autokorrelation vorliegt. Die Prüfgröße zeigt für die Zahl von 20, 16 und 12 Lags Werte von 33.86, 28.51 und 23.03 welche alle deutlich über dem kritischen Wert bei 10% liegen. Somit gibt es Hinweise, daß die Schätzung des jeweils vergangenen Halbjahres nicht den Bedingungen der Rationalität genügt. Dieses Ergebnis wird dadurch relativiert, daß der DQF äußerst gering ist. Ferner spielt der Wert für das jeweils vergangene Halbjahr nur eine unbedeutende Rolle für die DIW-Prognosen.

Die Überprüfung der Prognose für das jeweils zweite Halbjahr auf Suffizienz wird unter Berücksichtigung des unterstellten Moving-Average-Prozesses durchgeführt. Zunächst soll überprüft werden, ob Autokorrelation mit dem letzten zum Prognosezeitpunkt bekannten Fehler vorliegt.

Regression: $VF(t) - a \cdot VF(t-2) = e(t) - b \cdot e(t-1)$

Koeffizient	Standardfehler	t-Signifikanz
a: 0.2980	1.1917	1.5543
v: 0.4856	0.2216	2.1915
N: 20	DW: 1.661	SEE: 1.0380

Wegen fehlender Signifikanz der Konstante wurde die Ljung/Box-Prüfgröße der Residuen für die Gleichung $VF(t) = e(t) + b \cdot e(t-1)$ berechnet. Es ergaben sich für die Lags 10, 12, 14 und 16 Werte für die Prüfgröße von 3.58, 6.54, 7.21 und 7.87. Für die Lags 12 und 14 muß die Hypothese fehlender Autokorrelation verworfen werden, bei 10 und 16 Lags kann sie nicht verworfen werden. Es ergeben sich also nur schwache Hinweise auf fehlende Autokorrelation, die auch angesichts der mit der Modellspezifizierung schwankenden Signifikanz des Moving-Average Prozesses nicht den Schluß auf mangelnde Suffizienz zulassen.

4.4 Effizienz

Die Referenzprognosen wurden anhand des gleichen Verfahrens wie im vorigen Kapitel ermittelt. Für die jeweils letzten vier Quartalswerte, auf die sich die Voraussage stützt, wurden wiederum die zum damaligen Zeitpunkt vorliegenden vorläufigen Werte verwendet.

Der mittlere quadrierte Prognosefehler des DIW ist erwartungsgemäß geringer als bei der vorherigen Referenzprognose. Jedoch ist der Unterschied zwischen den Prognosen größer als bei der DD. Die Korrelation zwischen Vorhersage und Beobachtung fällt deutlich ab. Erneut bestimmt die mangelnde Korrelation im wesentlichen die Höhe des DQF. Die entsprechende Komponente des Vorhersagefehlers weist die Werte 0.4208 und 4.6611 auf. Der mittlere quadrierte Vorhersagefehler ist mit 0.5083 für das vergangene und 5.2662 für das kommende Halbjahr erwartungsgemäß geringer. Das Ergebnis zeigt, daß durch das Arima-Modell relativ verläßlich nur zwei Quartalswerte vorausgesagt werden können.

Die Referenzprognosen sind erneut unverzerrt. Die Regression für die Schätzung des vergangenen Halbjahres ergaben für die gesamte (diskontinuierliche) Stichprobe:

Regression $VF(t) = c + e(t)$

Koeffizient	Standardfehler	t-Signifikanz	N	SEE
c: 0.1287	0.1240	1.3080	33	0.7121

Der Koeffizient ist auf einem Signifikanzniveau von 10% nicht von Null unterschieden.

Die Prognose des zweiten Halbjahres kann nur für die kontinuierliche Stichprobe überprüft werden:

Regression $VF(t) = c + e(t) + b \cdot e(t-1)$

Koeffizient	Standardfehler	t-Signifikanz
c: 0.0700	0.3464	0.2022
b: 0.3599	0.2071	1.7377
N: 22	DW: 1.7746	SEE: 1.6247

Der unterstellte Moving-Average Prozeß ist noch signifikant auf einem Niveau von 10%, wogegen die Konstante sich nicht signifikant von Null unterscheidet.

Die Bestimmung der relativen Effizienz erfolgt wie oben durch Regression des Vorhersagefehlers der Referenzprognose auf die Differenz der Vorhersagen. Für die Vorausschätzung des vergangenen Halbjahres ergab sich: Regression: $(B(t)-RV(t)) = k \cdot (V(t)-RV(t)) + e(t)$

	Koeff.	Std.fehl.	t-Sign.	t(k=1)	SEE	R ²	N
1. Halbj.	0.7468	0.1015	7.3571	-2.4946	0.4413	0.6159	33
2. Halbj.	0.8825	0.1571	5.6168	-0.7476	0.6414	0.5017	32

Der Faktor k ist bei der Vorausschätzung des vergangenen Halbjahres mit Signifikanz von 10% gegen den Wert 1 gesichert. Hier kann darauf geschlossen werden, daß die kombinierte Vorhersagen der isolierten Prognose überlegen ist.

Für die Prognose des zweiten Halbjahres zeigt sich kein solcher Hinweis. Hier ist die Hypothese, der Faktor k habe den Wert eins, nicht zu verwerfen.

Insgesamt läßt sich feststellen, daß die Prognosen des realen Sozialprodukts des DIW ist für das zweite Halbjahr unverzerrt und bedingt suffizient sind also Bedingungen der Rationalität erfüllen. Die Effizienz gegenüber einer ARIMA-Prognose ist gesichert. Die Voraussage des jeweils ersten Halbjahres ist nicht unverzerrt und nicht suffizient, genügt also den Bedingungen der Rationalität nicht. Sie ist auch nicht voll-effizient. Es überrascht, daß die für den späteren Zeitraum gültige Vorhersage die Berechnungen der Rationalität besser erfüllt als jene für den früheren. Der sehr geringe Vorhersagefehler relativiert die Bedeutung der Abweichung allerdings sehr stark.

5. Fazit

Die Ergebnisse zeigen, daß sowohl die Prognosen der Gemeinschaftsdiagnose als auch des DIW für das Bruttosozialprodukt für einen Zeitraum bis zu einem halben Jahr relativ verläßlich sind. Längerfristige Vorhersagen sind freilich mit großen Fehlern behaftet. Dies führt zu der Schlußfolgerung, daß die angewandten Verfahren "kurzsichtig" sind. Da wirtschaftspolitische Entscheidungen nur in den seltensten Fällen binnen eins halben Jahres wirksam sein dürften, mindert dieser Befund die Bedeutung, die den Prognosen für aktuelle wirtschaftspolitische Entscheidungen zukommen kann, beträchtlich. Um diese Schwäche zu beleben, sind methodische Änderungen bei der Vorhersagetechnik wohl unvermeidlich. Hierbei könnte der verstärkte Einsatz ökonometrischer Modelle für die Erstellung einer Prognose mit längerer

Reichweite sinnvoll sein.

Für die weitere Forschungsarbeit ist es außerdem sinnvoll zu wissen, in welcher Weise die einzelnen Komponenten des BSP die Prognosegüte beeinflussen. Sind sie auch unverzerrt ? Oder heben sich gegenläufige Verzerrungen einzelner Komponenten wechselseitig auf? Welche Komponenten verursachen die größten Vorhersagefehler ? Eine Antwort auf diese Fragen kann nur mittels einer Analyse der einzelnen Sozialproduktkomponenten, die der oben angewandte Methodik folgt, gefunden werden.

Literatur

Fuller, W. A.: Introduction to Statistical Time Series, New York, 1976

Granger, C.W.J.; Newbold, P.: Forecasting Economic Time Series, 2.nd edition, San Diego et. al., 1986

Holden, K; Peel, D.A.: On Testing for Unbiasedness and Efficiency of Forecasts, The Manchester School, Vol.LVIII, No.2, June 1990, S. 120-127

Kirchgässner, G.: Wie gut sind die Prognosen der Arbeitsgemeinschaft wirtschaftswissenschaftlicher Forschungsinstitute in der Bundesrepublik Deutschland, Weltwirtschaftliches Archiv, Bd. 120, 1983, S. 279-300