ZU WENIG WETTBEWERB? ZU STAND UND ENTWICKLUNG DER GENAUIGKEIT MAKROÖKONOMISCHER PROGNOSEN[1]

von *Ullrich Heilemann* und *Sabine Klinger*[2]

Ungeachtet der vielfältigen Anstrengungen hat sich die Genauigkeit makroökonomischer Prognosen seit den 1970er Jahren nicht verbessert. Deshalb mahnen (manche) Ökonomen „mehr Wettbewerb“ unter den Prognoseinstituten an – und hoffen, dass die allokative Effizienz marktlicher Lösungen auch bei Makroprognosen eintritt. Allerdings handelt es sich um sehr spezifische Informationsgüter: Sie ähneln öffentlichen Gütern, ihre Genauigkeit und ihre allgemeine Güte lassen sich schwer messen, aus vergangenen Ergebnissen kann nicht auf den künftigen Erfolg/Misserfolg geschlossen werden. Folgerichtig führt die wettbewerbstheoretische Analyse nicht zu klaren Indizien, dass „mehr Wettbewerb“ die Prognosegüte erhöht. Dieses noch vage Ergebnis wird von einer empirischen Untersuchung ausgewählter nationaler Prognosemärkte unterstrichen: Weder die Anzahl der Prognoseinstitute noch das Verhältnis von kommerziellen zu nicht-kommerziellen Prognostikern zeigte einen positiven Einfluss auf die Treffsicherheit. Trotz der Skepsis der Autoren darüber, dass makroökonomische Prognosen zurzeit überhaupt wesentlich verbesserbar wären, werden Vorschläge unterbreitet, wie der Prognosewettbewerb intensiviert werden könnte.

Abstract

Despite manifold efforts, accuracy of macroeconomic forecasts has not improved since the 1970s. Therefore, (some) economists demand “more competition” among forecasting institutions, hoping that allocative efficiency of market solutions holds in the field of macroeconomic forecasts, too. However, forecasts are very special information goods: they show some similarities with public goods, their accuracy as well as their general quality is hard to measure, past performance is no guide to future success. Theoretical deliberations do not reveal any clear indications for improvements via “more competition”. This result is supported by an empirical analysis of selected national markets for macroeconomic forecasts. Neither the number of forecasters nor the ratio of profit to non-profit institutions showed a positive impact on forecasting accuracy. Though the authors are sceptical that at present macroeconomic forecasts have much room for improvement, they offer some proposals on how to increase competition in this field.

Key Words: empirical macroeconomics, forecasting accuracy, competition

JEL Classification: D4, D83, E01, E3

---

[1] Vortrag im Rahmen der Jahrestagung 2005 des Wirtschaftspolitischen Ausschusses des Vereins für Socialpolitik, Marburg, 15.-17. März 2005. Für kritische Kommentare danken die Verfasser *Hans Willgerodt* und den anderen Tagungsteilnehmern sowie *Rolf Hasse*. Die Aktualisierung der Berechnungen besorgte Frau Dipl.-Vw. *Annika Blaschzik.* – Die Arbeit entstand im Rahmen des SFB 475 „Komplexitätsreduktion in multivariaten Datenstrukturen“ gefördert durch die Deutsche Forschungsgemeinschaft an der Universität Dortmund.

[2] Institut für Empirische Wirtschaftsforschung, Universität Leipzig, Marschnerstraße 31, 04109 Leipzig, heilemann@wifa.uni-leipzig.de bzw. Institut für Arbeitsmarkt- und Berufsforschung, Weddigenstraße 20-22, 90478 Nürnberg, sabine.klinger@iab.de.

A. Einleitung

Die Kenntnis der Zukunft wird in der Regel als nützlich angesehen. Von alters her erfreuen sich Prognostiker hoher Aufmerksamkeit sowie eines beträchtlichen Zuflusses an Ressourcen zur Entwicklung neuer Verfahren und Techniken. Unsere Zeit gilt mit etlicher Berechtigung als prognosebesessen – *futurehype* (*Dublin* 1989) – ob dies auch in längerfristiger Perspektive zutrifft, soll hier offen bleiben. Immerhin gelten, vom *Orakel von Delphi* abgesehen, das 16. und 17. Jahrhundert als das goldene Zeitalter der „Wahrsager“. *Gaule* zählt 1652 bereits 52 Prognoseverfahren – von der *Akromantie*, der Weissagung mit Hilfe der Luft bis zur *Tyromantie*, der Weissagung mit Hilfe von Käse (*Mackay* 1992 [1852], S. 176 ff.). Zu Überheblichkeit hat unsere Zeit übrigens wenig Anlass: nach einer Meldung der *Washington Post* vom 28. November 1995 gab die Regierung der Vereinigten Staaten im Zeitraum 1985 bis 1990 ca. 11 Million Dollar für parapsychologische „Fernaufklärung“ des sowjetischen U-Boot-Programms oder von Drogenschmuggel bzw. zur Überprüfung des Wahrheitsgehaltes dieser Methoden aus, und der Erfolg der „Seher“ wurde immerhin „unterschiedlich“ beurteilt. Bereits den Alten war freilich bewusst, dass die Methode eine, die richtige Interpretation ihrer Ergebnisse eine andere Sache ist, und sie verwendeten daher nicht wenig Scharfsinn darauf, wie die Prophezeiungen des Orakels von Delphi, aber auch die Praxis der „wissenschaftlichen“ Astrologie des 16. Jahrhunderts (*Grafton* 2001) belegen.

Die Geschichte der quantitativen Wirtschaftsprognosen ist sehr viel kürzer. Ihre Anfänge in den Vereinigten Staaten reichen in bis in die 1920er Jahre zurück. In Deutschland finden sich quantitative Vorausschauen aber erst seit dem Zweiten Weltkrieg, detaillierte, am Aufbau der Volkswirtschaftlichen Gesamtrechnung (VGR) orientierte Prognosen, wie sie heute üblich sind, sogar erst seit Anfang der 1960er Jahre. Der gestiegene Umfang und Detaillierungsgrad hat die Treffsicherheit der Prognosen erfahrungsgemäß nicht erhöht und an den notorischen Klagen über ihre Unzuverlässigkeit nichts geändert. Nach wie vor vergeht kaum ein Jahr, das dazu keinen Anlass bietet. Unzufriedenheit wird dabei nicht nur von der breiten Öffentlichkeit und den Medien geäußert, sondern vor allem von der Politik. Allerdings haben die Prognostiker auch fachintern einen schweren Stand, und nicht selten wird ihnen die Wissenschaftlichkeit abgesprochen (*Viner* 1958, S. 103ff.) oder gar die „Unmöglichkeit von Wirtschaftsprognosen“ postuliert (*Morgenstern* 1928, *Urban* 1973) – eine „Anmaßung von Unwissen“ (*Willgerodt* 2004), die über das Spannungsfeld von Möglichkeiten der und Erwartungen an die Wirtschaftsprognose hinausschießt, wie nicht zuletzt auch die folgenden Ausführungen zeigen werden. Die Mehrheit der Fachvertreter dürfte sich jedenfalls zunehmend darüber im Klaren sein, dass sich die Profession auch über Prognoseleistungen legitimieren muss, und es gibt wohl kaum DFG-Anträge von Ökonomen, in denen für den Fall einer Förderungszusage nicht eine beträchtliche Verbesserung der prognostischen Leistungsfähigkeit in Aussicht gestellt wird.

Ob die Unzufriedenheit mit den Leistungen der makroökonomischen Prognosen stets berechtigt ist, ob die Erwartungen an ihre Treffsicherheit nicht überzogen sind, wird dabei selten gefragt. Leider halten die Prognostiker meist ihre diesbezügliche Neugier im Zaum[3] und überlassen die Bestimmung der Treffsicherheit ihrer Prognosen lieber Dritten. Die Ergebnisse dieser mittlerweile sehr zahlreichen und detaillierten Untersuchungen bestätigen ausnahmslos den negativen Eindruck, der sich bei Politik und Öffentlichkeit gebildet hat. Sichtbare Resignation scheinen die Befunde allerdings bei den Prognostikern ebenso wenig ausgelöst zu haben wie erkennbare Aktivitäten zur Verbesserung der Treffsicherheit. Gelegentlich münden die Befunde bei den Prognostikern in der Forderung nach mehr Ressourcen, nach „mehr und besseren Daten“; ein Ruf nach „besseren Theorien“ ist dagegen selten zu hören. Diesen spezifischen Antworten seitens der Produzenten auf den Vorwurf ungenügender Treffsicherheit wird häufig aus der Nachfrageseite die Forderung „Mehr Wettbewerb“ zur Seite gestellt, wie jüngst im Zusammenhang mit der Kritik an den vermeintlich zu optimistischen Wachstums- und Defizitprognosen von Regierungen der EU-Mitgliedsländer (*Jonung / Larch* 2004). Neben der Generalvermutung der segensreichen Wirkungen des Gewinnstrebens auch im Wissenschaftsbereich (*Beckmann* 1987) wird als Begründung z. B. in Deutschland vor allem auf das „Kartell der Prognostiker“ und ihre staatliche Alimentierung verwiesen.[4] In der Tat werden die einflussreichsten Prognosen –

---

[3] Anders als im Lehrbuch propagiert, lautet das Motto in der Regel „forecast often and don`t keep records“.

[4] Vgl. dazu auch die – mit Blick auf die Ergebnisse der vorliegenden Arbeit – etwas blauäugigen Erwartungen von *Cassel* (2001), S. 122 ff., an die wirtschaftspolitische Beratung allgemein.

die der „Gemeinschaftsdiagnose“ (GD) bzw. der sechs großen Wirtschaftsforschungsinstitute (DIW, HWWA, ifo, IfW, IWH, RWI), des *Sachverständigenrats zur Begutachtung der gesamtwirtschaftlichen Entwicklung* (SVR), der Bundesregierung („Jahreswirtschaftsbericht“ (JWB)) sowie der EU-Kommission, der OECD und des Internationalen Währungsfonds – mehr oder weniger staatlich finanziert. Änderungen dieses Kreises wie auch innerhalb der an diesen Prognosen maßgeblich beteiligten Personen als Folge von Fehlprognosen sind nicht bekannt.[5]

Die skizzierte Sachlage gibt hinreichend Anlass, sich mit der Treffsicherheit makroökonomischer Prognosen, ihren Tendenzen und den Möglichkeiten der Verbesserung auseinanderzusetzen. Im Folgenden soll namentlich soll der Frage nachgegangen werden, inwiefern „Mehr Wettbewerb“ eine höhere Prognosegüte verspricht. Die Untersuchung beschränkt sich dabei auf das Feld der Prognosen – der übergeordnete Zusammenhang „Wettbewerb in der Wissenschaft“ (dazu z. B. *Beckmann* 1987, *Böhme* 1978) wird nur kurz angesprochen. Gewiss ist diese Beziehung auch für die Wirtschaftswissenschaften ein lohnendes Feld, pflegt doch die Politik aller Ebenen mit Hinweisen auf Pluralität und Wettbewerb eine breite Landschaft von Forschungsinstituten (*Wissenschaftsrat* 1998, S. 33 f.). Selbstredend liefert die vorliegende Untersuchung auch einen Beitrag zum Verständnis des Komplexes „Wettbewerb und Wissenschaft“. Insgesamt setzt sie mit ihrer Beschränkung auf die Frage nach der Treffsicherheit von Konjunkturprognosen und die Rolle, die der Anbieterwettbewerb dabei spielt oder spielen kann, einen sehr viel engeren Rahmen.

Im Einzelnen ist das weitere Vorgehen wie folgt: der nächste Abschnitt (B) gibt einen Überblick über die Treffsicherheit ausgewählter makroökonomischer Prognosen für die Bundesrepublik und die G7-Länder. In Abschnitt C wird die Hypothese, durch Wettbewerbserhöhung verbessere sich die Treffsicherheit von Wirtschaftsprognosen, entwickelt und aus theoretischer Sicht beleuchtet. Daran schließt sich im Abschnitt D eine Übersicht

---

[5] Suizid, wie ihn 1865 der erste Leiter des neu gegründeten, staatlichen britischen Wetterdienstes, Captain *Robert Fitzroy*, angesichts erheblicher öffentlicher Kritik an seinen prognostischen Fehlleistungen beging (*Brooks* 2004, S. 134), ist – zum Glück! – selten. Die dieser Tage vom *Moskauer* Bürgermeister geforderte harte Bestrafung meteorologischer Fehlleistungen (NZZ, Internationale Ausgabe Nr. 27 vom 25. Februar 2005) wird mit Sicherheit Ankündigung bleiben.

über die Struktur des deutschen und des internationalen Prognosemarktes sowie – darauf aufbauend – der Versuch einer empirischen Überprüfung der in C entwickelten Hypothese an. Zusammenfassung und einige Folgerungen beschließen die Ausführungen (E).

B. Treffsicherheit makroökonomischer Prognosen

Die folgenden empirischen Befunde zur Treffsicherheit makroökonomischer Prognosen oder Konjunkturprognosen, d. h. Prognosen mit einem Horizont von bis zu zwei Jahren, beschränken sich auf die Bundesrepublik Deutschland und einige ausgewählte OECD-Länder: Sie unterscheiden sich nicht wesentlich von denen für andere, vergleichbare Länder, wie sie z. B. die Arbeiten von *Pons* (2000), *Zarnowitz* (1992), *McNees* (1986) für die Vereinigten Staaten, *Kreinin* (2000), *Öller / Barot* (2000), *Koutsogeorgopoulou* (1998), *Ash / Smyth / Heravi* (1997, 1993), *Ballis* (1989) für große Industrieländer oder von *Keeremann* (2004), *Jonung / Larch* (2004) sowie *Jokinen* (1993) für die Mitgliedsländer der EU bzw. des Euro-Währungsgebiets vorgelegt haben; auf einige Unterschiede und Ausnahmen wird unten eingegangen.

I. Empirische Ergebnisse zur Treffsicherheit der makroökonomischen Prognosen

Die folgende Analyse beschränkt sich auf die gesamtwirtschaftlichen Vorausschätzungen der Gemeinschaftsdiagnose [GD] (1967 ff.), des Sachverständigenrats [SVR] (1967 ff.), der OECD (1967 ff.) im Herbst für das jeweils kommende Jahr sowie des Jahreswirtschaftsberichts der Bundesregierung [JWB] (1967 ff.) im Januar für das jeweils laufende Jahr. Die Institutionen sind allgemein bekannt; die von ihnen verwendeten Prognosemethoden interessieren im vorliegenden Zusammenhang nicht, sodass auf ihre Vorstellung verzichtet werden kann (vgl. dazu z. B. die Hinweise in *Heilemann / Stekler* 2003). Die Untersuchung beschränkt sich auf zwei wichtige wirtschaftspolitische Ziele: das Wirtschaftswachstum (Veränderungsraten des realen BIP[6] gegenüber dem Vorjahr) sowie die Inflation (Veränderungsraten des BIP-Deflators gegenüber dem Vorjahr). Neben den durchschnittlichen absoluten Fehlern (MAE) der Prognosen (prognostizierter abzüglich beobachteter Variablenwert) werden die relativen Fehler (RMSPE) sowie die Verzerrtheit (Bias) der Prognosen berechnet.[7] Darüber hinaus wird in komparativer Perspektive gefragt, wie sich die Treffsicherheit im Vergleich mit einer „naiven“ *no-change*-Prognose (*Theil*scher Ungleichheitskoeffizient) darstellt. Die Genauigkeit der Prognosen mit Blick auf die vorhergesagte Richtung der Entwicklung oder Konjunkturphase wird hier nur kurz angesprochen, desgleichen die Frage der „Rationalität“ der Prognosen (ausführlicher: *Heilemann / Stekler* 2003). Die Ergebnisse werden für den Zeitraum 1967 bis 2003 präsentiert, d. h. den aktuell längsten gemeinsamen Stützbereich. Neben den Ergebnissen für den gesamten Zeitraum werden auch solche für die Jahre 1970 bis 1979, 1980 bis 1989, 1990 bis 2003 vorgestellt. Den Ergebnissen für Deutschland werden die von der OECD (1967 ff.) für *Frankreich*, *Japan*, *Kanada*, das *Vereinigte Königreich*, die *Vereinigten Staaten* (G7) ebenfalls im Herbst für das kommende Jahr prognostizierten Werte gegenüber gestellt (bis 2001). Als „beobachtete“ Werte dienen im Falle Deutschlands die in der „Gemeinschaftsdiagnose“ der Forschungsinstitute im Frühjahr veröffentlichten Werte für das jeweilige Vorjahr. Im Falle der OECD-Prognosen wird entsprechend verfahren. Da ein Teil der Prognosen auf halbe vH-Punkte gerundete Prognosewerte ausweist, z. B. die GD bis Frühjahr 1995, wurden hier aus Vergleichbarkeitsgründen alle Prognosewerte entsprechend gerundet.

---

[6] Bis 1993: BSP, danach: BIP.

[7] Zur Berechnung der Prüfmaße vgl. die Übersicht im Anhang.

Tabelle 1

Prognosen der jährlichen Veränderungsraten des realen BIP und des BIP-Deflators für Deutschland: ausgewählte Fehlermaße

1967 bis 2003

	reales BIP				BIP-Deflator			
	GD	SVR	OECD	JWB	GD	SVR	OECD	JWB
1967 bis 2003								
MAE	1,4	1,2	1,3	1,1	0,7	0,8	0,7	0,8
RMSPE	122,5	117,2	133,8	103,0	91,3	90,7	79,5	91,5
Bias	0,2	0,2	0,2	0,1	-0,1	-0,1	-0,1	-0,2
U	0,3	0,3	0,3	0,2	0,1	0,1	0,1	0,1
UM	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
UV	0,3	0,3	0,4	0,4	0,3	0,4	0,2	0,4
UC	0,7	0,7	0,6	0,6	0,7	0,6	0,8	0,6
RMSE/σ	0,8	0,7	0,8	0,7	0,5	0,5	0,5	0,5
1970 bis 1979								
MAE	1,9	1,5	1,4	1,3	1,1	1,2	0,8	1,1
RMSPE	173,5	140,9	198,4	67,9	24,8	23,4	19,9	22,1
Bias	0,7	0,7	0,6	0,6	-0,6	-0,7	-0,4	-0,8
U	0,4	0,3	0,3	0,3	0,1	0,1	0,0	0,1
UM	0,1	0,1	0,1	0,1	0,2	0,2	0,0	0,3
UV	0,2	0,2	0,3	0,3	0,3	0,3	0,1	0,3
UC	0,7	0,7	0,6	0,6	0,5	0,5	0,9	0,4
RMSE/σ	0,9	0,8	0,9	0,7	0,5	0,5	0,5	0,5
1980 bis 1989								
MAE	0,9	0,7	1,0	0,9	0,4	0,5	0,8	0,5
RMSPE	76,2	54,7	85,7	88,7	18,0	22,2	34,2	22,1
Bias	-0,1	0,1	0,2	0,1	0,1	-0,2	-0,2	-0,2
U	0,2	0,2	0,3	0,2	0,0	0,0	0,1	0,0
UM	0,0	0,0	0,0	0,0	0,0	0,1	0,0	0,1
UV	0,0	0,1	0,1	0,0	0,3	0,1	0,1	0,1
UC	1,0	0,9	0,9	1,0	0,6	0,8	0,9	0,8
RMSE/σ	0,7	0,6	0,8	0,8	0,3	0,5	0,7	0,5
1990 bis 2003								
MAE	1,1	1,0	1,1	0,8	0,6	0,8	0,5	0,7
RMSPE	117,2	140,4	117,0	137,3	99,4	120,5	120,8	122,1
Bias	0,4	0,1	0,2	0,3	0,1	0,3	0,2	0,1
U	0,3	0,3	0,3	0,2	0,1	0,1	0,1	0,1
UM	0,1	0,0	0,0	0,1	0,0	0,1	0,1	0,0
UV	0,5	0,2	0,4	0,3	0,4	0,7	0,6	0,7
UC	0,4	0,8	0,6	0,6	0,6	0,2	0,3	0,3
RMSE/σ	0,7	0,7	0,7	0,6	0,4	0,5	0,4	0,4

Eigene Berechungen. Für Quellen, Abkürzungen und Berechnung der Fehlermaße siehe Text bzw. Anhang.

1. Deutschland[8]

Wie Tabelle 1 zeigt, ergibt sich für den gesamten Stützzeitraum bei den Wachstumsprognosen im Durchschnitt ein Mittlerer Absoluter Fehler (MAE) von 1,3 vH-Punkten, was etwa 50 vH der durchschnittlichen Veränderungsrate entspricht. Für die Inflationsrate lauten die entsprechenden Werte 0,8 vH-Punkte und 25 vH. Laut dem RMSPE, d. h. in relativer Perspektive und bei quadratischer Gewichtung der Fehler, beträgt die durchschnittliche Abweichung 120 vH für die BIP- bzw. 90 vH für die Inflationsprognosen. Im Vergleich mit den *no-change*-Szenarien zeigt sich, dass die Prognosen für beide Variablen weitaus überlegen und die Fehler zum überwiegenden Teil *zufälliger* Natur sind, d. h., der Kovarianzfehler (UC) überwiegt bei weitem den Mittelwert- (UM) und den Varianzfehler (UV). Aus kognitiver Sicht ist dies als eher lässliche Sünde anzusehen bzw. positiv zu werten. Das gute Abschneiden der Prognosen im Vergleich mit simplen extrapolativen Verfahren überrascht nicht. Deutschland verzeichnete im relevanten Zeitraum fünf Rezessionen, die wie die klaren Abwärtstendenzen bei den Veränderungen von Wachstum und Inflation von naiven Verfahren nur schwer nachzuzeichnen sind. Ähnliches gilt auch für die Mehrzahl der unten betrachteten Länder.

Insgesamt lassen die Prognosen eine leichte Tendenz zur Überschätzung erkennen. Dies zeigt sich vor allem in den Aufschwungperioden, während die Abschwungperioden unterschätzt werden. Insbesondere wurde *keine* der Rezessionen, definiert als negative Veränderungsraten gegenüber dem Vorjahr, prognostiziert.

Wie die Ergebnisse für die drei Teilzeiträume belegen, hat sich die Treffsicherheit im Zeitablauf trotz einer Zunahme der Anzahl von Prognostikern nicht verbessert (so für die

---

[8] Der Abschnitt stützt sich im Wesentlichen auf *Heilemann / Stekler* (2003), wo sich auch Aussagen zur Treffsicherheit der Tendenzaussagen finden und die Ergebnisse verschiedener Rationalitätstests wiedergegeben werden.

Vereinigten Staaten auch *Zarnowitz* (1996, S. 387 bzw. 1992)). Zwar sind die Prognosefehler in den 1980er Jahren, verglichen mit den 1970er Jahren, zurückgegangen, aber in den 1990er und den folgenden Jahren sind sie im Fall der Wachstumsprognosen, zumindest was den RMSPE betrifft, wieder angestiegen. Berücksichtigt man explizit den Rückgang des Niveaus der jeweiligen Veränderungsraten (MAE/AVG)[9] oder die Standardabweichung (RMSE/σ), so sind für die untersuchten Teilzeiträume allenfalls sehr geringe Veränderungen zu registrieren.

Ein „bester" Prognostiker lässt sich anhand dieser Ergebnisse kaum ausmachen, zumal wenn man bedenkt, dass zwischen der Prognose der GD und dem JWB praktisch fast vier Monate liegen, was insbesondere für die Prognosebasis[10] als auch für ggf. zu berücksichtigende wirtschaftspolitische Maßnahmen von Bedeutung ist. Erst recht gilt dies in kognitiver Perspektive, d. h., wenn man hier von den unterschiedlichen Annahmen oder – im Fall makroökonometrischer Modellprognosen: sonstigen Korrekturen der Modellergebnisse – abstrahiert. Die Abweichungen in den Werten der Prüfmaße sind gering und die „Ränge" wechseln je nach Betrachtungszeitraum und Variable – ein wenig überraschendes Ergebnis, da sich die Unterschiede zwischen den Prognosen der einzelnen Institutionen in engen Grenzen bewegen: der Korrelationskoeffizient (r) erreicht sowohl im Falle der Wachstums- als auch der Inflationsprognosen fast durchweg Werte zwischen 0,8 und 1 (Tabelle 2). Entsprechende Tests führen nicht zu einer Verwerfung der Null-Hypothese der Unverzerrtheit der Prognosen, machen aber deutlich, dass die Prognosefehler autokorreliert sind (*Heilemann* 2003a).

Insgesamt entsprechen die Ergebnisse den seit längerem vorliegenden internationalen Erfahrungen (für die Vereinigten Staaten oder das Vereinigte Königreich vgl. z. B. *Fildes / Stekler* 2002, *Zarnowitz* 1996, S. 385 ff., *McNees* 1986).

---

[9] Diese werden nur für die OECD-Prognosen (Tabellen 3 und 4) ausgewiesen.

[10] Immerhin beträgt der MAE der Prognosen des realen BIP-Wachstums für das laufende Jahr im Herbst beträgt im Falle der GD und des SVR jeweils etwa ¼ vH-Punkt, was via „statistischem Überhang" für die Prognose für das kommende Jahr durchaus von Bedeutung sein kann.

Tabelle 2

Korrelationen[1] zwischen den Prognosen wichtiger Institutionen für Deutschland
1967 bis 2003

		GD	SVR	OECD	JWB
GD	1967 bis 2003	.	0,9	0,8	0,9
	1970 bis 1979	.	1,0	0,8	0,8
	1980 bis 1989	.	0,9	0,9	1,0
	1990 bis 2003	.	0,9	0,9	0,9
SVR	1967 bis 2003	1,0	.	0,9	0,9
	1970 bis 1979	1,0	.	0,9	0,9
	1980 bis 1989	0,9	.	0,9	0,9
	1990 bis 2003	1,0	.	1,0	0,9
OECD	1967 bis 2003	1,0	1,0	.	0,8
	1970 bis 1979	1,0	1,0	.	0,7
	1980 bis 1989	0,8	0,9	.	0,9
	1990 bis 2003	1,0	0,9	.	0,9
JWB	1967 bis 2003	1,0	1,0	1,0	.
	1970 bis 1979	1,0	1,0	1,0	.
	1980 bis 1989	0,9	1,0	0,9	.
	1990 bis 2003	0,9	0,9	1,0	.

Eigene Berechnungen. – 1) r zwischen den Prognosen des realen BIP (rechts der Hauptdiagonalen) sowie zwischen den Prognosen des BIP-Deflators (links der Hauptdiagonalen).

2. Die G7-Länder

Die Ergebnisse für die G7-Länder sind denen für Deutschland sehr ähnlich. Zwar unterscheiden sie sich bezüglich der absoluten und der relativen Fehler beträchtlich: Über den gesamten Stützzeitraum liegt bei den Wachstumsprognosen Frankreich an der Spitze der Treffsicherheitsskala und Japan rangiert am Ende (Tabelle 3). Bereinigt man die Ergebnisse um die unterschiedlichen Niveaus der Veränderungsraten des (realen) BIP bzw. der Inflation (MAE/AVG) oder um die unterschiedlichen Standardabweichungen, so verschwinden die Fehlerunterschiede fast vollständig. Bei den Inflationsraten liegen die Prognosen

Tabelle 3

Jährliche Wachstumsprognosen der OECD: ausgewählte Fehlermaße

1967 bis 2001

	Kanada	Frankreich	Deutschland	Italien	Japan	UK	US	G7	OECD
1967 bis 2001									
MAE	1.4	0.9	1.4	1.4	1.7	1.1	1.1	0.6	0.8
MAE/AVG	0.4	0.3	0.5	0.5	0.4	0.5	0.4	0.1	0.1
RMSPE	124.9	91.8	347.5	155.7	391.1	583.2	216.0	19.7	16.7
RMSE/σ	0.4	0.4	0.4	0.4	0.2	0.4	0.3	0.1	0.2
Bias	0.4	0.1	0.1	0.3	0.1	0.1	0.0	-0.1	0.3
U	0.2	0.1	0.3	0.3	0.2	0.3	0.1	0.0	0.0
UM	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1
UV	0.3	0.1	0.3	0.0	0.2	0.1	0.1	0.2	0.0
UC	0.7	0.9	0.7	1.0	0.8	0.9	0.9	0.9	0.8
1970 bis 1979									
MAE	1.2	1.1	1.5	2.7	2.3	1.3	1.0	1.3	1.4
MAE/AVG	0.3	0.3	0.4	0.8	0.3	0.5	0.3	0.2	0.2
RMSPE	157.2	112.8	233.2	178.7	165.4	1078.3	120.2	15.8	24.9
RMSE/σ	0.5	0.4	0.3	0.4	0.3	0.6	0.2	0.5	0.4
Bias	0.6	0.2	0.5	0.4	0.5	0.7	0.7	-1.3	-1.1
U	0.1	0.2	0.3	0.6	0.2	0.5	0.2	0.0	0.1
UM	0.2	0.0	0.0	0.0	0.0	0.2	0.2	0.5	0.3
UV	0.3	0.3	0.6	0.0	0.6	0.1	0.2	0.2	0.0
UC	0.5	0.7	0.4	1.0	0.4	0.9	0.6	0.2	0.7
1980 bis 1989									
MAE	1.6	0.7	1.1	0.9	1.0	1.0	1.0	0.5	0.7
MAE/AVG	0.6	0.4	0.6	0.5	0.3	0.8	0.5	0.1	0.1
RMSPE	51.2	87.1	84.8	191.8	28.1	42.2	170.3	12.3	14.4
RMSE/σ	0.3	0.9	0.6	0.4	0.9	0.3	0.2	0.5	0.5
Bias	-0.2	-0.2	-0.1	-0.2	-0.2	-0.8	-0.6	0.2	0.4
U	0.3	0.2	0.3	0.2	0.1	0.2	0.1	0.0	0.0
UM	0.0	0.0	0.0	0.0	0.0	0.4	0.2	0.1	0.3
UV	0.5	0.0	0.1	0.2	0.1	0.1	0.2	0.0	0.1
UC	0.5	1.0	0.9	0.8	0.8	0.5	0.6	1.0	0.6
1990 bis 2001									
MAE	1.5	0.9	1.0	0.9	1.4	0.9	1.4	0.8	0.9
MAE/AVG	0.5	0.4	0.5	0.4	0.6	0.4	0.6	0.4	0.4
RMSPE	149.9	87.0	123.7	143.7	687.3	143.7	315.9	119.6	119.1
RMSE/σ	0.5	0.8	0.7	1.1	0.4	0.5	0.6	1.3	1.3
Bias	0.7	0.4	0.2	0.6	0.8	0.4	-0.1	-0.1	0.2
U	0.3	0.2	0.3	0.3	0.5	0.3	0.3	0.2	0.2
UM	0.2	0.1	0.0	0.4	0.2	0.1	0.0	0.0	0.0
UV	0.3	0.3	0.3	0.1	0.3	0.3	0.4	0.2	0.2
UC	0.5	0.6	0.7	0.5	0.5	0.6	0.6	0.8	0.8

Eigene Berechnungen. Für Quellen und Berechnung der Fehlermaße siehe Text bzw. Anhang.

Tabelle 4

Jährliche Inflationsprognosen der OECD: ausgewählte Fehlermaße

1967 bis 2001

	Kanada	Frankreich	Deutschland	Italien	Japan	UK	US	G7	OECD
1967 bis 2001									
MAE	1.3	0.8	0.7	1.6	1.7	1.6	0.9	0.8	0.9
MAE/AVG	0.3	0.1	0.2	0.2	0.5	0.2	0.2	0.3	0.3
RMSPE	91.6	48.5	82.1	28.3	176.7	28.5	24.6	138.0	661.0
RMSE/σ	0.1	0.1	0.2	0.1	0.2	0.1	0.2	0.4	0.4
BIAS	-0.1	-0.2	-0.2	-1.0	0.2	-0.6	0.0	-0.3	0.1
U	0.1	0.0	0.1	0.0	0.3	0.7	0.1	0.1	0.1
UM	0.0	0.0	0.0	0.2	0.0	0.1	0.0	0.1	0.0
UV	0.2	0.1	0.3	0.0	0.2	0.2	0.0	0.0	0.2
UC	0.8	0.9	0.7	0.8	0.8	0.8	1.0	0.9	0.8
1970 bis 1979									
MAE	1.8	1.2	0.9	2.7	3.6	3.4	1.5	0.4	1.1
MAE/AVG	0.3	0.1	0.2	0.2	0.5	0.3	0.3	0.1	0.3
RMSPE	25.3	15.1	18.9	23.9	61.5	32.4	24.3	82.2	1218.5
RMSE/σ	0.3	0.2	0.4	0.2	0.2	0.1	0.4	0.1	0.3
BIAS	-1.4	-0.9	-0.6	-2.1	-1.1	-2.7	-0.8	-0.1	0.5
U	0.1	0.0	0.1	0.1	0.3	0.1	0.1	0.0	0.2
UM	0.2	0.3	0.2	0.4	0.0	0.4	0.2	0.1	0.1
UV	0.0	0.0	0.2	0.0	0.2	0.1	0.0	0.1	0.3
UC	0.8	0.7	0.6	0.6	0.8	0.5	0.8	0.8	0.6
1980 bis 1989									
MAE	1.0	0.5	0.6	1.4	1.3	1.4	0.9	1.0	0.7
MAE/AVG	0.2	0.1	0.2	0.1	0.5	0.2	0.2	0.4	0.3
RMSPE	28.4	10.4	25.3	17.7	230.6	21.3	26.8	189.1	165.2
RMSE/σ	0.1	0.0	0.6	0.1	1.5	0.1	0.2	0.5	0.5
BIAS	0.1	0.0	-0.3	-0.9	1.0	0.1	0.9	-0.7	-0.4
U	0.0	0.0	0.1	0.0	0.8	0.0	0.0	0.1	0.1
UM	0.0	0.0	0.2	0.3	0.4	0.0	0.6	0.3	0.2
UV	0.0	0.2	0.0	0.0	0.2	0.0	0.0	0.2	0.3
UC	1.0	0.8	0.8	0.7	0.4	1.0	0.4	0.5	0.5
1990 bis 2001									
MAE	1.2	0.6	0.5	0.8	0.5	0.5	0.4	0.4	0.5
MAE/AVG	0.5	0.3	0.2	0.2	0.5	0.1	0.1	0.2	0.2
RMSPE	158.4	82.9	48.9	21.4	207.0	23.3	23.2	25.6	15.6
RMSE/σ	1.4	0.9	0.2	0.3	0.6	0.2	0.5	0.7	1.3
BIAS	1.2	0.6	0.2	-0.2	0.4	0.1	0.2	0.3	-0.3
U	0.8	0.1	0.0	0.1	0.2	0.0	0.0	0.1	0.0
UM	0.8	0.8	0.1	0.0	0.5	0.0	0.2	0.2	0.2
UV	0.0	0.0	0.3	0.1	0.0	0.0	0.2	0.1	0.0
UC	0.2	0.2	0.6	0.9	0.5	1.0	0.7	0.6	0.8

Eigene Berechnungen. Für Quellen und Berechnung der Fehlermaße siehe Text bzw. Anhang.

für Deutschland vorn, während die für Japan wiederum die geringste Treffsicherheit aufweisen (Tabelle 4). Auch bei den Inflationsprognosen ebnet die Standardisierung die Unterschiede sehr ein, lediglich im Falle der Prognosen für Japan ändert sich nichts. Bei den Wachstumsprognosen sind mit Ausnahme Kanadas nur geringe (positive) Verzerrungen der Prognosen zu verzeichnen, bei den Inflationsprognosen überwiegen die negativen Verzerrungen, die allerdings mit Ausnahme *Italiens* und des Vereinigten Königreichs ebenfalls gering sind. Auch in komparativer Perspektive fällt sowohl bezüglich der Wachstums- als auch der Inflationsprognosen das Urteil sehr positiv aus, wiederum u. a. auch deshalb, weil die Fehler überwiegend als „zufällig" zu klassifizieren sind.

Im Zeitablauf ändert sich dieses Bild in ähnlicher Weise, wie oben bei den Prognosen für Deutschland festgestellt: Die 1980er Jahre bringen leichte Verbesserungen, die 1990er Jahre einen „Rückfall", wobei freilich auch in diesem Fall die standardisierten Maße ein recht konstantes Bild von der Treffsicherheit liefern. In der Rangfolge der Treffsicherheit sind lediglich bezüglich Japan und den Vereinigten Staaten Verbesserungen bzw. Verschlechterungen festzustellen. Eine durchgehende Verschlechterung ist bezüglich des Bias zu beobachten.

Auch bei den OECD-Prognosen für die G7-Länder zeigt sich – durchaus überraschend – eine vergleichsweise hohe Korrelation der Prognosefehler sowohl im Falle der Wachstums- als auch der Inflationsprognosen (Tabelle 5). Bemerkenswerterweise nimmt allerdings der Fehlerzusammenhang nicht zu, wie aufgrund der wachsenden internationalen Verflechtung der Gütermärkte und der Finanzmärkte erwartet werden könnte und die bei der OECD besondere Berücksichtigung finden sollte. Die Rationalitätstests konstatieren wie im Fall Deutschlands, dass die Null-Hypothese der Unverzerrtheit nicht widerlegt werden kann und die Fehler autokorreliert sind.[11]

---

[11] Inwiefern dies u. a. auch daher rührt, dass die Prognosen – anders als die Rationalitätstests – nicht auf beobachteten, sondern auf geschätzten Werten für das Vorjahr aufbauen, muss hier offen bleiben.

Tabelle 5

Korrelation zwischen den Fehlern der OECD-Wachstums- bzw. Inflationsprognosen

1967 bis 2001

	Kanada	Frank-reich	Deutsch-land	Italien	Japan	UK	US
Kanada							
1967-2001	-	0.6	0.5	0.5	0.4	0.6	1.0
1970-1979	-	0.8	0.1	0.4	0.4	0.6	0.9
1980-1989	-	-0.2	0.6	0.6	0.0	0.7	0.9
1990-2001	-	0.0	-0.4	-0.2	-0.5	0.8	0.9
Frankreich							
1967-2001	0.9	-	0.7	0.9	0.7	0.8	0.9
1970-1979	1.0	-	0.2	0.8	0.7	0.9	0.9
1980-1989	0.9	-	0.5	1.0	0.2	0.7	0.9
1990-2001	0.8	-	0.9	1.0	0.4	0.9	0.9
Deutschland							
1967-2001	0.6	0.6	-	0.6	0.6	0.6	0.7
1970-1979	0.3	0.1	-	0.0	0.6	0.1	0.4
1980-1989	1.0	0.9	-	1.0	0.2	0.9	0.9
1990-2001	0.5	0.9	-	0.9	0.6	0.8	0.6
Italien							
1967-2001	0.9	0.7	0.6	-	0.7	0.4	0.9
1970-1979	0.8	0.7	0.4	-	0.8	0.6	0.8
1980-1989	0.9	0.5	0.9	-	0.2	0.7	0.9
1990-2001	0.7	0.6	0.7	-	0.4	0.0	0.9
Japan							
1967-2001	0.8	0.8	0.8	0.7	-	0.8	0.8
1970-1979	0.8	0.7	0.8	0.5	-	0.6	0.8
1980-1989	0.9	0.9	1.0	0.9	-	0.9	1.0
1990-2001	0.7	0.8	0.8	0.8	-	0.8	0.8
UK							
1967-2001	0.9	0.4	0.4	0.8	0.2	-	0.9
1970-1979	0.9	0.8	0.4	0.8	0.6	-	0.8
1980-1989	0.8	0.2	0.6	0.7	-0.1	-	0.9
1990-2001	0.9	0.1	-0.1	0.8	-0.3	-	0.9
US							
1967-2001	0.7	0.3	0.3	0.2	0.3	0.5	-
1970-1979	0.5	0.4	0.1	0.0	0.3	0.2	-
1980-1989	0.9	-0.3	0.3	0.4	-0.1	0.8	-
1990-2001	0.8	0.1	-0.2	0.1	-0.2	0.8	-

Eigene Berechnungen. – 1) r zwischen den Prognosen des realen BIP (rechts der Hauptdiagonalen) sowie zwischen den Prognosen des BIP-Deflators (links der Hauptdiagonalen).

II. Probleme der Bestimmung der Treffsicherheit makroökonomischer Prognosen

Die Bestimmung der Treffsicherheit makroökonomischer Prognosen, ihr Vergleich und erst recht die Bewertung ihrer Ergebnisse gestalten sich weitaus schwieriger, als von der Öffentlichkeit, aber vermutlich auch von vielen Fachvertretern erwartet. Nicht alle Untersuchungen machen dies deutlich. Zur Illustration dieser Schwierigkeiten bietet sich das folgende allgemeine Schema („wissenschaftlicher“) Prognosen an, wie es der makroökonometrischen Modellprognose, im Kern aber allen Verfahren zugrunde liegt, die mit expliziter Angabe der erklärenden Variablen arbeiten, also auch der so genannten Iterativen VGR-Prognose (vgl. dazu z. B. *Tichy* 1994, S. 206 ff.) bzw. dem so genannten *informal GDP-model* (*Zarnowitz* 1996, S. 385 ff.).[12]

(1) $y_t = f\left(y_{t-k}; x_{t-i}; \beta; e_t\right)$,

mit:

$f = (f_1, \ldots, f_N)$: Vektor der N funktionalen Beziehungen;

y_{t-k} : Vektor der (zeitverzögerten) endogenen Variablen, $k = 0, \ldots, m$;

x_{t-i} : Vektor der (zeitverzögerten) exogenen Variablen, $i = 0, \ldots, n$;

$\beta = (\beta_1, \ldots, \beta_N)$: Matrix der geschätzten strukturellen Parameter;

$e_t = (e_{1t}, \ldots, e_{Nt})$: Vektor der Störterme (mit den üblichen Annahmen).

Die Schwierigkeiten der Bestimmung der Treffsicherheit beginnen bei der Wahl der Referenzgröße, d. h., was ist als „beobachteter Wert“ dem jeweils prognostizierten Variablenwert (y_t) gegenüberzustellen. Der erste veröffentlichte Wert der amtlichen Statistik unterscheidet sich z. B. beim realen BIP (Veränderungsrate gegenüber dem Vorjahr) vom „endgültigen“ um 1 vH-Punkt vom „endgültigen“ (z. B. *Heilemann* 2002). Überwiegend wird – wie hier – der erste Wert als Maßstab verwendet; es gibt auch für den endgültigen Wert gute Gründe, die hier aber nicht weiter interessieren. Damit ist auch das generelle Problem

---

[12] Zeitreihenansätze, also Verfahren, die ohne Angabe ökonomisch erklärender Variablen arbeiten, spielen in den hier betrachteten Prognosen eine allenfalls komplementäre Rolle.

der Daten- und Informationsbasis (y_{t-i}, x_{t-i}) angesprochen: zu unterschiedlichen Zeitpunkten ist auch der Informationsstand der Prognostiker bezüglich der amtlichen Statistik (z. B. der Werte der VGR), der einzelnen vorlaufenden, gleich- oder nachlaufenden Indikatoren oder sonstiger Informationen wie z. B. über die zu erwartende Entwicklung der Tariflöhne oder der Ölpreise, unterschiedlich. Diese Kenntnisse müssen nicht zwangsläufig die Treffsicherheit erhöhen, aber sie erschweren die Beurteilung der Leistungsfähigkeit der Prognostiker. Im vorliegenden Fall liegt die Vermutung nahe, dass z. B. die vergleichsweise hohe Treffsicherheit des JWB auch mit dessen Informationsvorsprung angesichts seiner späten Veröffentlichung/Erstellung in Verbindung zu bringen ist, umgekehrt im Fall der OECD-Prognose (für Deutschland), die einem zeitaufwändigen internen Abstimmungsprozess unterliegt (vgl. dazu z. B. *Pons* 2000).

Erhebliche Schwierigkeiten mindestens der Vergleichbarkeit, erst recht aber mit Bezug auf die Bestimmung der Leistungsfähigkeit der Prognostiker, ergeben sich mit Blick auf die Annahmen der Prognose (x_{t-i}). Je nach Annahmen bezüglich wichtiger Determinanten der makroökonomischen Entwicklung – Stichworte: weltwirtschaftliche Entwicklung, finanzpolitische Interventionen, Ölpreis, Tariflohnentwicklung – können sich Prognosen bei sonst gleichen oder ähnlichen Hypothesen erheblich unterscheiden. Bei Prognosen, die nicht mit ökonometrischen Modellen erstellt werden, ist es im Allgemeinen sehr schwierig, den Einfluss der einzelnen Annahmen zu ermitteln. Dass Prognosen mit „falschen“ Annahmen denen mit „richtigen“ überlegen sind, ist z. B. bei amerikanischen Prognosen keine Seltenheit (*Heilemann* 2002). – Das im Zusammenhang mit Prognosebeurteilungen vor allem von den Prognostikern häufig bemühte Argument des „*self fulfilling*“ oder „*self destroying*“ – sie bildeten bereits den Kern der *Morgenstern*schen These von der Unmöglichkeit von Wirtschaftsprognose (*Morgenstern* 1928) – sollte ungeachtet der logischen Auseinandersetzung mit dem Problem (*Grunberg / Modigliani* 1954) bis auf weiteres in den Fußnoten verbleiben; die empirische Relevanz ist bescheiden, wenn man vom Fall der Börsenkrise 1987 absieht (*Heilemann* 1990).[13]

---

[13] Inwiefern in den Vereinigten Staaten die erfolgreiche makroökonomische Rezessionsbekämpfung 2001 ein neuerliches Beispiel liefert, bleibt allerdings zu prüfen (vgl. dazu auch die Hinweise in Fußnote 27).

Davon unabhängig stellt sich aber die Frage, ob und wie ggf. die Beurteilung der Treffsicherheit erfolgen sollte. Nicht alle Prognoseadressaten werden Anhänger des Konzeptes der *induktiven Hypothesenwahrscheinlichkeit* (*Wild* 1974, S. 134 ff.) sein und lediglich Wert auf empirisch und theoretisch gut fundierte Prognosen – also die richtigen ß und niedrige e_t – legen und in ihren jeweiligen Prognosenutzenfunktionen die Frage der richtigen Annahmen gering gewichten.

III. Externe und interne Bedeutung der Treffsicherheit

Die Argumente der Prognosenutzenfunktion lassen sich allerdings noch grundsätzlicher relativieren: eine hohe Treffsicherheit der Prognosen muss nicht zwangsläufig oder jedenfalls nicht immer eine hohe Wertschätzung bei den Prognoseadressaten besitzen. Die Veröffentlichung gesamtwirtschaftlicher Prognosen kann z. B. bei der Wirtschaftspolitik unerwünschten Handlungsdruck auslösen. Ein bekanntes Beispiel stellt die dilatorische Rezeption der Abschwungprognosen seitens der Bundesregierung im Vorfeld der Bundestagswahl im September 2002 dar, die andernfalls finanzpolitische Reaktionen (Nachtragshaushalt) erfordert hätten, die ihr angesichts des bevorstehenden Wahlkampfs als wenig opportun erschienen.[14] Eine ähnlich differenzierte und ggf. sich wandelnde Einstellung zur Treffsicherheit von (gesamtwirtschaftlichen) Prognosen ist auch für andere Träger der Wirtschaftspolitik, z. B. die Notenbank, und für Unternehmen – namentlich auch für börsennotierte – vor dem Hintergrund ihrer Berichtspflicht auszumachen. Zwar ist grundsätzlich vorstellbar, dass Prognosen eine unterschiedliche *interne* und *externe* Nützlichkeit beigemessen und je nach situativer Bedeutung das Argument „Treffsicherheit“ unterschiedlich gewichtet wird. Es erscheint jedoch mehr als zweifelhaft, dass die Prognoseadressaten die Verwendung unterschiedlicher Prognosen der Öffentlichkeit glaubhaft vermitteln können.[15] Alles in allem führt das zu vermutende unterschiedliche Interesse an der

---

[14] Erst recht in kognitiver Sicht ist die Treffsicherheit als Prüfkriterium keineswegs unangefochten, wie die Diskussion um *Friedman*s (1953) “*positive economics*“ zeigte.

[15] Vgl. dazu z. B. die Ausführungen im sog. „Lügenausschuss“ des Deutschen Bundestages (2003, S. 48ff., S. 134ff.).

Treffsicherheit oder „Wahrheit“ zu der ernüchternden Erkenntnis: *less is (sometimes) more*![16]

C. Mit „Mehr Wettbewerb“ zu besseren Prognosen? Das Gut „Prognose“ in wettbewerbstheoretischer Sicht

Die allokative Effizienz von Wettbewerb wird von Ökonomen in der Regel nicht in Zweifel gezogen, die wettbewerbliche Marktorganisation verspricht eine optimale Versorgung. Im vorliegenden Zusammenhang bedeutet dies, dass Prognoseproduzenten je nach Aufwand ein preislich und qualitativ differenziertes Angebot bereitstellen, das von den Prognosekonsumenten entsprechend ihrer Prognosenutzenfunktion und ihrem Budget in Anspruch genommen wird. Sind im Falle – makroökonomischer bzw. kurzfristiger – Wirtschaftsprognosen die Voraussetzungen der allokativen Effizienz einer marktwirtschaftlichen Lösung des Prognoseproblems gegeben? Oder handelt es sich bei „Prognosen“ um ein Gut, für das die marktwirtschaftliche Lösung ineffiziente Ergebnisse generiert? Weist der Prognosemarkt, wie z. B. der Markt für Bildungs- oder für Gesundheitsleistungen, Unvollkommenheiten auf, die einer effizienten Marktlösung entgegenstehen?

An der Offenheit des Prognosemarktes bestehen kaum Zweifel – zur Prognoseerstellung bedarf es im Prinzip weder einer Lizenz noch einer Zertifizierung.[17] Allerdings tangiert die Frage der Finanzierung das Offenheitskriterium. Wenn staatlich subventionierte Prognosen

---

[16] Aufschlussreich sind in diesem Zusammenhang einige jüngere Fälle in der amerikanischen Rechtssprechung, in denen Unternehmen der pharmazeutischen Industrie im Falle von Schadensersatzforderungen wesentlich härter bestraft wurden, wenn sie intern die Risiken von Medikamenten ausführlich untersucht hatten, während Unternehmen, die darauf verzichtet hatten, wesentlich günstiger wegkamen. „While the kind of weighing of risk and benefit may be medically rational, in the legal arena it’s poison.” (*Surowiecki* 2005, S. 38). Ähnliche Reaktionen waren in den Vereinigten Staaten übrigens bereits im Zusammenhang mit dem „Freedom of Information Act“ seitens der staatlichen Bürokratie zu beobachten gewesen (Guida 1989).

[17] Das war nicht immer so, wie *Fögen* (1997) in einer Untersuchung zur Entstehung des kaiserlichen Wissensmonopols im Römischen Reich – unter maßgeblicher Beteiligung christlicher Autoren – im dritten und vierten nachchristlichen Jahrhundert zeigt. – Neuerdings zeichnen sich übrigens Tendenzen der Zertifizierung im Bereich der *surveys* ab, so z. B. in Frankreichs *Conseil National de l’Information Statistique* (vgl. http://www.cnis.fr/). Ähnliche Bemühungen der Qualitätssicherung finden sich auch seitens der OECD.

der Öffentlichkeit regelmäßig, in hoher Qualität und für sie kostenlos zur Verfügung gestellt und damit zum quasi-öffentlichen Gut „erklärt“ werden, ist die Hürde für den Markteintritt neuer Anbieter angesichts der hohen Produktionskosten sehr hoch.[18]

Die Wettbewerbstheorie benennt Ausnahmebereiche, in denen entweder der Markt als Koordinations- und Ausschlussmechanismus versagt oder der Wettbewerb seine Funktion der ökonomischen Auslese zu eigenen Ungunsten („über“) erfüllt (vgl. z. B. *Bartling* 1983, *Olten* 1995, S. 70 ff. oder *Wolf* 1997, S. 17 ff.). Inwiefern derartige Ausnahmebereiche als Folge von Skalen- und Netzwerkeffekten, Externalitäten, öffentlichen Gütern und hohen Transaktionskosten für den Prognosemarkt zutreffen, ist in der Literatur bislang eher stiefmütterlich behandelt worden. Mit Bezug auf Informationsgüter (siehe *Klodt* 2003, S. 111 ff., *Bössmann* 1988, S. 199, *Arrow* 1962), unter die Prognosen zu subsumieren sind, ergibt sich folgendes Bild:

Das Argument wettbewerbsbehindernder *economies of scale* (*Bössmann* 1988, S. 199, *Daub* 1987, S. 80) und Netzwerkeffekte scheint für gesamtwirtschaftliche Prognosen nur mit Einschränkungen stichhaltig zu sein. Im Fall von Prognosen z. B. auf der Grundlage makroökonometrischer Modelle[19] besteht zunächst die Tendenz zum natürlichen Monopol, da Erstellung, Tests und Pflege eines solchen Modells mit hohem fixen Aufwand verbunden sind, wohingegen seine Anwendung geringe zusätzliche Kosten verursacht. Die Entwicklung fast aller großen Modelle war übrigens in erheblichem Umfang mit öffentlichen Mitteln gefördert worden – auch in den Vereinigten Staaten (*Heilemann* 1980). Zudem können die – überwiegend schriftlich oder elektronisch veröffentlichten – Modelle „leicht“ durch Veränderung weniger Modellgleichungen modifiziert werden. Die technologischen Voraussetzungen sprechen demnach nicht einmal im Fall modellbasierter Prognosen (oder,

---

[18] Zwar schreibt die EU die wettbewerbliche Vergabe von Dienstleistungsaufträgen durch EU-weite Öffentliche Ausschreibung vor (siehe Verdingungsordnung für Leistungen VOL (A), Abschn. 2, §§ 2, 3) – dies wird jedoch z. B. im Fall der Gemeinschaftsdiagnose aus einer Reihe von Gründen (noch) nicht praktiziert. Siehe auch Fußnote 30.

[19] Für andere Verfahren – wie z. B. die Iterative VGR-Prognose – gilt dies entsprechend dem jeweiligen Gewicht der einzelnen Schritte der Prognoseerstellung – in ähnlicher Weise.

dank leistungsfähiger Informationstechnologie, nicht mehr) gegen eine wettbewerbliche Lösung – wie ja auch die Praxis lehrt.

Effiziente Marktlösungen setzen des Weiteren voraus, dass der Markt (transparent) sämtliche Informationen über die Marktbesetzung sowie die qualitativen und quantitativen Eigenschaften des Gutes internalisiert und über den Preis widerspiegelt. Die Qualität einer Prognose hat viele Dimensionen. Sie reichen von ihrer „rechtzeitigen“ Verfügbarkeit bzw. Zugänglichkeit über ihren Informationsgehalt (Umfang, inhaltliche und zeitliche Disaggregation, Prognosehorizont), die theoretische und empirische Fundierung, den Präzisionsgrad, die Bedingtheit bis hin zu ihrer Treffsicherheit. Mit Ausnahme der Treffsicherheit und der theoretischen und empirischen Fundierung lassen sich diese Dimensionen recht leicht bestimmen, und die Mehrzahl der Konjunkturprognosen unterscheiden sich in dieser Hinsicht wenig. Was die Identifikation der theoretischen und empirischen Fundierung der Prognosen angeht, so geben hierüber die Prognostiker selten Auskunft, und Dritte können erst recht darüber vielfach nur mutmaßen und auch dies oft nur nach sehr aufwändigen Untersuchungen. Besonders große Probleme ergeben sich hinsichtlich der Bestimmung der Treffsicherheit. Dies gilt nicht nur – wie oben ausgeführt – für die Treffsicherheit der jeweiligen Prognoseinstitution in der Vergangenheit, sondern auch hinsichtlich der künftig zu erwartenden Genauigkeit. Im Falle gesamtwirtschaftlicher Prognosen kommt hinzu, dass es sich bei den hier diskutierten Prognosen um konjunktive, aus den Teilaggregaten errechnete Prognosen handelt, sowohl was die einzelnen wirtschaftspolitischen Ziele als auch was die konstituierenden Nachfragekomponenten angeht. Ferner ist zu beachten, dass die zu erwartende Treffsicherheit weder aus der bisherigen Treffsicherheit noch – wie vielleicht vermutet werden könnte – aus den verwendeten Methoden abgeleitet werden kann. Jedenfalls lautet so ein in mehreren Studien erhärteter Befund (*Zarnowitz* 1996, S. 519 ff., *McNees* 1986). Kurz, bei Prognosen wäre eine Preisdifferenzierung nach Qualitätsmerkmalen außerordentlich problembehaftet.

Abgesehen davon vermindert sich auch die Aussagekraft des Preises als Knappheitsindikator, wenn nicht die gesamte effektive Nachfrage am Markt artikuliert wird. Externe Effekte verhindern jedoch die vollständige Ausschließbarkeit von Nutzern des Gutes „Prognose“, die nicht als Nachfrager am Markt aufgetreten sind. Die Durchsetzung von Urheber- und

Eigentumsrechten ist in diesem Fall schwierig. Mindestens durch die Anwendung einer solchen Information erhalten andere teilweise Kenntnis von ihr (*Bössmann* 1988, S. 199, *Daub* 1987, S. 96), selbst wenn man berücksichtigt, dass die Handlungen der Akteure in der Regel nicht ausschließlich von der internen Kenntnis über die künftige Wirtschaftsentwicklung, sondern auch von anderen Einflüssen bestimmt werden. Die Verbreitung, das Nach-Außen-Dringen von Informationen/Prognoseergebnissen wird jedoch zusätzlich von den modernen Informations- und Kommunikationstechnologien erleichtert.

Ausschließbarkeit von der Nutzung einer Prognose ist somit zwar vom Prinzip her möglich, jedoch schwierig umzusetzen und einzuhalten (siehe z. B. *Klodt* 2003, S. 111). Einigkeit besteht dahingehend, dass mit der Nutzung kein Verbrauch verbunden ist[20], die Information somit vielen Nachfragern ohne zusätzliche Kosten zur Verfügung steht. Der Wegfall des Rivalitätskriteriums in Verbindung mit der schwierigen Umsetzung der Ausschließbarkeit führt die Prognose nah an die Eigenschaften eines öffentlichen Gutes heran (so z. B. *Daub* 1987, S. 96, *Klodt* 2003, S. 111 f., ähnlich *Wolf* 1997, S. 60, für wissenschaftliche Forschung allgemein). Dies spricht – wenn auch nicht zwingend – für die kostenlose bzw. staatlich alimentierte Bereitstellung von (wissenschaftlichen) Prognoseergebnissen. Insbesondere rechtfertigt es die Zweifel daran, dass „Mehr Wettbewerb“ *per se* die Treffsicherheit von Prognosen steigert.

Besondere Schwierigkeiten für die Verwirklichung von Wettbewerb auf dem Prognosemarkt resultieren nicht nur aus den Eigenschaften der Prognose als „quasi-öffentliches“ Gut, sondern auch aus dem Charakteristikum von Informationen, Erfahrungsgüter zu sein (*Klodt* 2003, S. 117 ff., *Arrow* 1962). Selbst wenn man vereinfachend die Mehrdimensionalität der Prognosegüte außer Acht ließe und sich allein auf einen numerischen Prognosewert konzentrierte, sind die Qualität und der Nutzwert dieser Information für den Konsumenten nicht im Vorhinein absehbar. Würde ihm jedoch vorab Einblick gewährt, gelänge er kostenlos in den Genuss der Information. Damit ist die Bereitstellung von Informa-

---

[20] Die Wertminderung der Information tritt im Zeitverlauf ein, nicht durch Nutzung. Sie ist für die Abgrenzung eines öffentlichen Gutes jedoch nicht relevant.

tion nicht nur mit Risiken bei der Generierung verbunden, sondern auch mit einer unvollständigen Abschöpfbarkeit der Erträge. Auf diesem „Informationsparadoxon“ fußt *Arrow*s Unterinvestitionshypothese (*Arrow* 1962)[21], der zufolge z. B angesichts der ökonomischen und gesellschaftlichen Schäden von Rezessionen ihrem „rechtzeitigen“ Erkennen zu wenig Ressourcen gewidmet werden. Lediglich in längerfristigen Marktbeziehungen stelle privatwirtschaftlich organisierte Vertragsforschung eine einigermaßen wettbewerbskonforme Lösung dar, weil dann der Auftraggeber aus der Qualität in der Vergangenheit Rückschlüsse auf das aktuell zu erwerbende Gut ziehen könne. Wie erwähnt, ist dies bei Prognosen indessen nicht möglich. Vermutlich ist dies auch ein Grund dafür, dass selbst in primär privatwirtschaftlich organisierten Prognosemärkten überwiegend auf ein Jahr begrenzte Verträge geschlossen werden.

Angesichts der Unmöglichkeit, aufgrund der bisherigen auf die künftige Prognosegenauigkeit zu schließen, sowie der schwierigen Zuordenbarkeit des Ergebnisses zu seinem „Urheber“ kann der Erwerb und die Nutzung des Gutes „Prognose“ von relativ hohen Such- und Kontrollkosten begleitet sein, was ebenfalls eher für eine unternehmens- oder regierungsinterne Forschung als für eine (externe) wettbewerbliche Lösung spricht. Wird dennoch die Marktlösung bevorzugt, gewinnt in Anbetracht der Schwierigkeiten bei der Beurteilung des Gutes die Reputation der Prognostiker für die Nachfrage an Bedeutung. In diesem Zusammenhang wäre unter Prognostikern, deren beruflicher Erfolg von Reputation abhängt, ein indirekter Anreizmechanismus zu Gunsten „besserer“ Prognosen zu vermuten. Wie gezeigt, hebt sich jedoch kein Anbieter mit durchgängig erfolgreicheren Prognosen von den Konkurrenten ab.

Die Anreiz- und Kontrollmechanismen der Marktwirtschaft, das Fundament des Wettbewerbsarguments, sind offenbar auf dem Prognosemarkt nur sehr eingeschränkt funktionstüchtig: Hohe Such- und Kontrollkosten angesichts eines diffizilen, facettenreichen Gutes, das zudem noch sehr weitgehend Züge eines öffentlichen Gutes trägt, lassen *a priori* in

---

[21] Anders bekanntlich *Hirshleifer* (1971), der mit Verweis auf die privaten erwerbswirtschaftlichen Interessen diese Möglichkeit nicht sieht.

mehr marktlichem Wettbewerb kein Verbesserungspotenzial erkennen. Dessen ungeachtet soll im Folgenden der Frage nachgegangen werden, inwiefern es bei der Prognoseerstellung nicht bereits Wettbewerb herrscht oder wo er sich ggf. installieren ließe.

D. Prognosemarkt und Wettbewerb – eine Bestandsaufnahme für Deutschland und die G7

I. Deutschland

Ein gleichermaßen umfassendes wie detailliertes Bild vom Prognosemarkt, wie es z. B. *Daub* (1987) für Kanada in den 1980er Jahren gezeichnet hat, liegt für Deutschland nicht vor. Sollte dies versucht werden, wäre ein erheblicher Aufwand in erster Linie in Form von Befragungen zu leisten, denn über die deutsche „Prognoseindustrie“ liegen weder Daten der Amtlichen Statistik vor, noch lassen sich aus Bilanzveröffentlichungen Hinweise auf Kosten-/Absatzstrukturen, Unternehmensstrategien usw.[22] gewinnen. Dies vor allem deshalb, weil es hierzulande in diesem Segment des Dienstleistungsbereichs wenig Unternehmen mit eigener Rechtspersönlichkeit gibt und die übrigen Unternehmen darüber nicht publizieren. Vieles spricht dafür, dass sich ein entsprechendes Bild für Deutschland in seinen Grundzügen, vor allem was die Kostenseite angeht, kaum von dem Daubs unterscheiden würde. Insofern dürfte auch dessen Ergebnis, dass angesichts der von ihm errechneten Profitabilität von Mitte der siebziger bis Mitte der achtziger Jahre die kanadische Prognoseindustrie – letztlich auch mit Blick auf die Treffsicherheit – als effizient anzusehen ist (*Daub* 1987, S. 100 ff.), übertragbar sein. Ein weiteres wichtiges Ergebnis, nicht unabhängig von dem zuletzt genannten Befund, dürfte sich ebenfalls bestätigen: die Treffsicherheit der Prognosen spielt, verglichen mit anderen Absatzparametern wie Breite des Angebotes, Transportabilität (: problemlose elektronische Verfügbarkeit, Weiterverarbeitungs- und Präsentationsmöglichkeiten), generelles Renommee des Produzenten, preisliche Wettbe-

---

[22] Die wenigen Studien zu nichtdeutschen Märkten für Prognosen bzw. makroökonometrische Modelle nutzen derartige Angaben, um die Märkte nach dem Schema *structure – conduct – performance* zu analysieren (*Daub* 1987, S. 88 ff., *Smith* 1994, S. 71 ff.)

werbsfähigkeit usw., eine geringe Rolle. Daub verweist in diesem Zusammenhang übrigens ebenfalls auf die Probleme der Bestimmung der Prognosegenauigkeit. Die 1990er Jahre dürften die Struktur des Prognosemarktes weltweit durch das Vordringen der Internetnutzung als primärem Präsentations- und Vertriebskanal insofern erheblich verändert haben, als die Gebühren für das Nutzen der Datenbanken (*time-sharing*) und damit eine wichtige Einnahmequelle der Industrie weggefallen ist. Für Deutschland war diese Entwicklung aus den genannten Gründen allerdings ohne größere Bedeutung.

Der deutsche Markt für (kurzfristige) makroökonomische Prognose ist stark von staatlichen Einflüssen geprägt. Auf der Angebotsseite ist gegenwärtig von ca. 50 Anbietern auszugehen. Die Expansion der Finanzmärkte hat diese Zahl vor allem in den letzten Jahrzehnten deutlich ansteigen lassen. Hervorgehobene Beachtung bei Politik oder Medien finden davon weniger als ein Drittel. Neben der eingangs erwähnten „Gemeinschaftsdiagnose“ bzw. den sechs großen Wirtschaftsforschungsinstituten (DIW, HWWA, ifo, IfW, IWH, RWI), dem Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung und der Bundesregierung mit ihrem „Jahreswirtschaftsbericht“ sind dies die EU-Kommission sowie die internationalen Organisationen, d. h. OECD, IWF und Weltbank.[23] Mit Blick auf die Bedeutung wäre also von einem weiten Oligopol mit polypolistischer Tendenz zu sprechen. Von den ca. zehn einflussreichsten ist kein Institut privatwirtschaftlich ausgerichtet oder privat finanziert. Rein kommerziell ausgerichtete Prognoseunternehmen, wie z. B. die nationalen Ableger der in den Vereinigten Staaten ansässigen Unternehmen DRI-McGraw-Hill, WEFA usw., sind nicht in Deutschland tätig. Der Aufwand der öffentlichen Hand belief sich für die „Gemeinschaftsdiagnose“ im Jahr 2003 auf 1,3 Mill. €, für den Sachverständigenrat auf 1,7 Mill. €, wovon freilich weniger als 3 vH auf die Erstellung der gesamtwirtschaftlichen Prognose entfallen sein dürfte.

---

[23] Die namhaften Institutionen erweisen sich hinsichtlich Prognosegenauigkeit allerdings gegenüber kleineren Anbietern, Abteilungen von größeren Unternehmen oder von Banken als unterlegen (*Blix u.a.* 2001, S. 45 ff.). Inwiefern deren möglicherweise höhere Abhängigkeit vom „Prognoseerfolg“ dabei eine Rolle spielt, muss allerdings offen bleiben.

Auch als „natürliche“ Folge der staatlichen, steuerfinanzierten Alimentierung werden die Prognosen der Öffentlichkeit und Interessierten in der Regel kostenlos zur Verfügung gestellt (*Kuhlen* 2001), seit einigen Jahren auch in elektronischer Form. Die wesentlichen Teile der numerischen Prognosen finden in der Berichterstattung der Medien große Beachtung.

Die Nachfrageseite des deutschen Prognosemarktes bleibt zu großen Teilen im Dunkeln. Das Interesse der Politik und der Wirtschaftspolitik einschließlich der großen Verbände liegt auf der Hand. Es wird durch die Beauftragung der großen Forschungsinstitute mit der Erstellung der Gemeinschaftsdiagnose, durch die Finanzierung des Sachverständigenrats und die Erstellung des Jahreswirtschaftsberichtes belegt. Unbestritten ist auch das Interesse der Öffentlichkeit bzw. das der Medien.[24] Sehr viel schwerer abzuschätzen ist die Nachfrage des Privaten Sektors. Setzte man in Analogie zu den in den Vereinigten Staaten in den 1970er Jahren beobachteten Relationen (*Heilemann* 1980, S. 62) ein Nachfragepotential entsprechend dem Vorhandensein einer volkswirtschaftlichen Abteilung voraus, so dürfte es heute noch etwa 50 Unternehmen umfassen. Berücksichtigt man ferner, dass bezüglich Informationsgehalt, theoretischer und empirischer Begründung sowie „Treffsicherheit“ ein vergleichsweise hochwertiges Angebot – neuerdings in elektronischer Form – für die Nutzer kostenlos zur Verfügung steht, so scheint das kommerziell nutzbare Nachfragepotential gegen Null zu tendieren. Zumindest wird mit dieser Begründung der hierzulande geringe Markterfolg amerikanischer Anbieter erklärt.

Auch über das Wettbewerbsverhalten der Prognoseanbieter liegen verständlicherweise wenig Informationen vor. Im Wesentlichen dürfte ungeachtet der staatlichen Alimentierung der in der Öffentlichkeit meist beachteten Prognostiker das Gleiche gelten, was *Zarnowitz* (1992) für die Vereinigten Staaten festgestellt hat (siehe unten). Bemerkenswert ist die außerordentlich hohe Korrelation der Fehler der Wachstums- und Inflationsprognosen (siehe für Deutschland nochmals die Tabelle 2). Über den gesamten Untersuchungszeitraum wie

---

[24] Die wichtigsten Ergebnisse beispielsweise der GD finden trotz Sperrfristen meist einige Tage vor Veröffentlichung ihren Weg in die Medien.

über alle Teilperioden erreicht sie Werte um 0,9 und mehr! Es sei dahingestellt, ob dies als Indiz mangelnder Wettbewerbsintensität oder als risikoaverse Vermeidung einer Abseitsposition auf einem fast freien Markt (*Smith* 1994, S. 75) zu begreifen ist.

II. International

Untersuchungen über die Struktur anderer nationaler Prognosemärkte sind selten: der unreflektierten Übernahme einer Analyse der makroökonometrischen Modellindustrie (*Smith* 1994, S.71 f.) sind Grenzen durch Eigenschaften des Prognose- vs. Modellgutes gesetzt. Allerdings scheint eine oligopolistische Besetzung der Angebotsseite gegenwärtig typisch für den Prognosemarkt zu sein. Der Untersuchung von *Blix u.a.* (2001) zufolge erscheint das deutsche Angebot, gemessen an der Zahl der Anbieter, als reichlich, lediglich übertroffen von den Vereinigten Staaten[25]; gemessen an den Einwohnerzahlen oder dem BIP/Kopf relativiert sich diese Aussage beträchtlich.

Zarnowitz (1992, S. 131 ff.) sieht den US-amerikanischen Prognosemarkt als Wettbewerbsindustrie, die keine erstbeste Lösung des Prognoseproblems zu Tage fördert. Zwar besteht ebenfalls eine polypolistische, teils oligopolistische Konkurrenz zwischen vielen, unterschiedlich großen Anbietern (nach *Zarnowitz* 1996, S. 387, etwa 200-250); und die Marktzutrittsbarrieren sind gering. Gerade im Zeitalter der *high tech*-Information und -Kommunikation, der schnellen softwarebasierten Rechenlösungen und der Datenubiquität schrumpft jedoch das Diversifizierungspotenzial jedes einzelnen Marktteilnehmers. Im Wesentlichen stehen Allen ähnliche Ressourcen zur Verfügung, mit denen sie die gleiche Aufgabe zu lösen trachten. Unterschiede zwischen den Prognostikern ergeben sich lediglich hinsichtlich des Methodenmix, wobei Vor- und Nachteile durch die Wahl verschiedener Methoden aufgewogen werden, ohne dass ein durchgängig bester Mix gefunden wird. Der vermeintliche Wettbewerbsvorteil durch einen geringen Zentralisierungsgrad wird also

---

[25] Allgemein zu den Märkten für gesamtwirtschaftliche Prognosen in den Vereinigten Staaten und dem modellbasierten Prognoseangebot: *Sherden* (1998), S. 56 ff. Zur Entstehung der Modell-Industrie vgl. auch z. B. *Heilemann* (1980).

Tabelle 6

Prognosegenauigkeit und Niveau und Struktur des Prognoseangebots in ausgewählten Ländern,
1991 bis 2000

Land1)	BIP-Prognosen			Inflationsprognosen		
	Anzahl der Prognostiker	durchschnittlicher RMSE	Varianz der RMSE	Anzahl der Prognostiker	durchschnittlicher RMSE	Varianz der RMSE
Vereinigten Staaten	34	1,2	0,04	34	0,5	0,02
kommerziell	31	1,2	0,04	31	0,5	0,01
nicht-kommerziell	3	1,2	0,00	3	0,7	0,02
statistisch	4	1,8	0,78	4	0,6	0,05
Deutschland	30	1,2	0,06	29	0,7	0,01
kommerziell	23	1,2	0,06	23	0,7	0,01
nicht-kommerziell	7	1,1	0,03	6	0,6	0,01
statistisch	4	1,5	0,26	4	09	0,11
Frankreich	22	1,0	0,03	21	0,5	0,01
kommerziell	18	1,0	0,03	18	0,5	0,01
nicht-kommerziell	4	1,0	0,01	3	0,4	0,00
statistisch	4	1,2	0,07	4	0,6	0,02
Japan	20	1,7	0,06	18	0,4	0,01
kommerziell	18	1,6	0,06	17	0,4	0,01
nicht-kommerziell	2	1,8	0,01	1	0,5	0,00
statistisch	4	2,1	0,26	4	0,7	0,11
Italien	15	0,9	0,01	14	0,7	0,01
kommerziell	11	0,9	0,02	11	0,7	0,01
nicht-kommerziell	4	0,9	0,00	3	0,7	0,00
statistisch	4	1,1	0,06	4	0,8	0,12
Schweden	12	1,0	0,03	11	0,9	0,06
kommerziell	10	1,0	0,03	10	0,9	0,06
nicht-kommerziell	2	1,2	0,00	1	1,2	0,00
statistisch	4	1,5	0,86	4	1,6	0,81
Korrelation mit Anzahl der Prognostiker (r)	1,0	0,2	0,5	1,0	-0,5	-0,5
kommerziell	1,0	0,4	0,5	1,0	-0,5	-0,4
nicht-kommerziell	1,0	-0,5	0,9	1,0	-0,4	0,4

Eigene Berechnungen nach Angaben in *Blix u.a.* (2001), S. 56 ff. – 1) absteigend geordnet nach der Anzahl der Prognostiker ohne rein statistische Verfahren (Mittelwert, Median, Modus, *random walk*).

kompensiert oder zumindest relativiert durch die Ähnlichkeit der Voraussetzungen, mit denen jedes Institut seine Prognosen erstellt. Und in der Tat stechen die Vereinigten Staaten keineswegs durch besondere Treffsicherheit hervor sondern liegen eher im Mittelfeld, wie Tabellen 3 und 4 zeigen.

Verglichen mit den USA, ist in Frankreich verstärkt eine in den öffentlichen Sektor integrierte Prognoselandschaft zu erkennen. In enger Verknüpfung zum Staat und staatlich inspirierten Untersuchungszielen steht nicht nur das reine Prognoseprodukt, sondern auch sein Einsatz in der Politikanalyse und -beratung auf der Agenda (*Smith* 1994, S. 71). Was die Besetzungshäufigkeit der Angebotsseite anbelangt, liegt Frankreich mit reichlich 20 Prognostikern im Mittelfeld (Tabelle 6). Die Prognosen über die französische Volkswirtschaft haben sich als einzige längerfristig als vergleichsweise treffsicher erwiesen und haben sich sogar im Zeitablauf leicht verbessert (Tabellen 3, 4). Dieser Befund ist gerade im Vergleich von Frankreich und Deutschland zu den USA ein Indiz dafür, dass weder „Masse“ noch Wettbewerbsintensität *per se* hohe Treffsicherheit gewährleisten. Die kenntnisreiche, Erfahrung und Intuition verbindende Mischung von Daten, Methoden, Theorien und letzten Endes auch Glück scheint bislang ohne weiteres von „Mehr Wettbewerb“ profitieren zu können. Insbesondere nicht in vergleichsweise kleinen Märkten, die Beispiele Frankreich oder Schweden illustrieren. Für *Frankreich*s prognostischen Erfolg muss nach anderen Ursachen gesucht werden – zumal angesichts der Tatsache, dass das über ein ähnlich großes Angebot verfügende *Japan*, was die Prognosegenauigkeit anbelangt, auf hinteren Rängen landet.

III. Empirische Ergebnisse

Die Untersuchung des Einflusses des Wettbewerbs auf die Treffsicherheit der Prognosen erfordert dessen Operationalisierung. Explizite Informationen über dessen Intensität und die Entwicklung im Zeitablauf existieren nicht. Hilfsweise kann – wie oben bereits geschehen – auf die Anzahl der Anbieter und ihre Entwicklung im Zeitverlauf zurückgegriffen werden.

Unmittelbare Hinweise auf den möglichen Zusammenhang zwischen der Zahl der Anbieter und der Prognosegenauigkeit der Wachstums- und der Inflationsprognosen gibt eine Gegenüberstellung der Anzahl der Prognostiker, der durchschnittlichen RMSE und deren Varianz für die Vereinigten Staaten, Deutschland, Frankreich, Japan, Italien und Schweden

(Tabelle 6). Empirische Basis sind die Tabellen in *Blix u.a.* (2001, S. 56 ff.).[26] Zunächst werden tendenziell die Ergebnisse der Tabellen 3 und 4 (siehe oben) bestätigt. Erneut zeigte sich die etwas höhere Treffsicherheit für Frankreich, aber auch für Italien. Alles in allem bewegen sich die Fehler jedoch in einer schmalen Bandbreite – wie aufgrund der oben präsentierten Korrelationsergebnisse der Prognosefehler nicht anders zu vermuten. Die Varianz der RMSE spiegelt beim Vergleich der Länder auch die Volatilität der jeweiligen wirtschaftlichen Entwicklung wider. Untergliedert man die Prognostiker nach ihrer kommerziellen Orientierung, lässt sich keine eindeutige Aussage zu Gunsten oder Ungunsten von mehr privatem Wettbewerb ableiten: Die Fehlermaße liegen nah beieinander; eine Tendenz, dass kommerzielle Organisationen „bessere“ Prognosen erstellten als staatlich finanzierte oder umgekehrt, ist nicht erkennbar.

Die Korrelationsmaße für den Zusammenhang zwischen der Anzahl der Anbieter auf dem Prognosemarkt (ohne „statistische Verfahren“, d. h. „Prognosen ohne Erklärung“) und der Prognoseungenauigkeit zeigen für die Wachstumsprognosen erstaunlicherweise (abgesehen von den *non-profit*-Organisationen) positive Vorzeichen. Trotz der angebrachten Vorsicht in der Interpretation angesichts einer knappen Datenbasis und geringer Werte des Koeffizienten stützt dieses Ergebnis eher die Vermutung, dass Prognosen nicht ohne weiteres durch zusätzliche Anbieter auf dem Markt verbessert werden. Bei den Inflationsprognosen ist der Zusammenhang zwischen der Anzahl der Prognostiker und den Fehlermaßen etwas stärker ausgeprägt. Zudem spricht das (meist) negative Vorzeichen dafür, dass eine hohe Besetzung der Angebotsseite tendenziell mit geringeren Prognosefehlern einhergeht.

Diese Ergebnisse werden bestätigt, wenn man unterstellt – wofür vieles spricht –, dass sich die Anzahl der Anbieter seit den Sechzigerjahren eher erhöht als vermindert hat. Die Prognosegenauigkeit (RMSE/σ, RMSE/AVG) hat jedenfalls weder für Deutschland (Tabelle 1) noch für die G7-Länder (Tabellen 3, 4) in den einzelnen Teilzeiträumen zugenommen.

---

[26] Die Auflistung in *Blix u.a.* (2001) ist wegen der speziellen Fragestellung der Untersuchung nicht erschöpfend; so fehlt in der Liste der französischen Prognose-Institutionen beispielsweise das INSEE.

E. Zusammenfassung und Schlussfolgerungen

Ausgangspunkt der Untersuchung war die in der Öffentlichkeit, bei Politik, aber auch in der Wissenschaft weit verbreitete Unzufriedenheit mit der Treffsicherheit oder Genauigkeit makroökonomischer Prognosen. Neben den traditionellen Forderungen aus dem Kreis der Prognostiker, vor allem auch der Wissenschaft, nach mehr und besseren Daten sowie Methoden – seltener nach besseren Theorien –, wird seit langem auch die Forderung nach „Mehr Wettbewerb" zwischen den Prognostikern bzw. nach Sanktionen im Falle prognostischer Fehlleistungen erhoben.

Die Untersuchung der Treffsicherheit ausgewählter „prominenter" gesamtwirtschaftlicher Wachstums- und Inflationsprognosen für Deutschland (Gemeinschaftsdiagnose, Sachverständigenrat, OECD, Jahreswirtschaftsbericht) bzw. die G7-Länder hat gezeigt, dass die Fehler vergleichsweise hoch sind, insbesondere, wenn man sie z. B. an den Erfordernissen stabilisierungspolitischer Maßnahmen hinsichtlich Genauigkeit, Rechtzeitigkeit, Sicherheit usw. misst. In den letzten 40 Jahren hat sich die Treffsicherheit zudem praktisch nicht verbessert. Das erste Ergebnis bestätigt zahlreiche frühere Befunde. *De facto* gilt dies auch für den zweiten Befund. Dass aus kognitiver wie aus handlungsorientierter Sicht diese Ergebnisse außerordentlich ernüchternd sind, liegt auf der Hand – die Analyse der aggregativen Struktur der Prognosen würde diesen Eindruck übrigens noch drastisch verstärken. Es erklärt den Unmut vor allem auch der Politik, wenngleich stabilisierungspolitische Maßnahmen in den letzten Jahren bei uns ohnehin nicht auf der wirtschaftspolitischen Agenda standen.[27].

Ob von „Mehr Wettbewerb" im Prognosebereich – sofern dies für möglich gehalten wird, was mit guten Gründen bezweifelt werden kann – eine Erhöhung der Treffsicherheit zu

---

[27] Die sehr erfolgreiche makroökonomische Stabilisierungspolitik der Vereinigten Staaten seit den 1990er Jahren stellt in dieser Hinsicht eine Ausnahme dar. Sie war wesentlich dem durchaus risikobehafteten *forbearing* des Vorsitzenden der Federal Reserve Bank, *Alan Greenspan*, und einem „glücklichen" Zusammentreffen von Steuersenkungen und 11. September zu danken (*Heilemann* (2003b neuerdings *Council of Economic Advisers* (2005), S. 62 ff.). Die Ergebnisse in Tabelle 3 deuten übrigens daraufhin, dass der Politik ihre Aufgabe keineswegs durch besonders hohe Treffsicherheit der Konjunkturprognosen erleichtert worden wäre – im Gegenteil!

erwarten ist, erscheint aus mehreren Gründen als mindestens zweifelhaft. Immerhin bewegen sich die Treffsicherheit und ihre Tendenzen national und international im Grunde auf sehr ähnlichen Niveaus. Dies deutet auf eine hohe Prognoseeffizienz hin.[28] Außerdem zeigte sich, dass die Anzahl der Anbieter oder ihre Zugehörigkeit zum öffentlichen oder privaten Sektor keine relevanten Unterschiede hinsichtlich ihrer Treffsicherheit erkennen lässt. Hinsichtlich der Nachfrage könnte dabei in Verbindung mit dem erreichten Treffsicherheitsniveau eine Rolle spielen, dass für sie andere Bestandteile des Angebots wichtiger sind.

Die vorgelegten Befunde sprechen nicht dagegen, über Schritte in Richtung einer Stärkung wettbewerblicher Elemente des Prognosemarktes nachzudenken. An erster Stelle wäre an die Erhöhung der Transparenz bezüglich der prognostischen Leistungsfähigkeit der Prognostiker zu denken. Dies könnte dadurch geschehen, dass bei konsekutiven Prognosen, wie z. B. im Fall der GD oder des SVR, „neue“ Prognosen sich mit der Treffsicherheit der vorangegangenen detailliert auseinandersetzen und sie im Lichte ihrer längerfristigen Treffsicherheitsbilanz diskutieren. Damit hat z. B. das RWI im Sommer 1996 begonnen (*Rheinisch-Westfälisches Institut für Wirtschaftsforschung* 1996, S. 36), die GD unternimmt, wenn auch weniger detailliert, ähnliche Anstrengungen seit dem Frühjahr 2000 (*Arbeitsgemeinschaft* 2000). Der SVR hat sich von derlei Bemühungen bislang nicht inspirieren lassen. Die Aufgabe könnte prinzipiell auch nach dem Vorbild des *Macroeconomic Modelling Bureau*s – es analysierte von 1983 bis 1998 u. a. auch die Treffsicherheit der wichtigsten makroökonometrischen Modellprognosen für das Vereinigte Königreich (vgl. z. B. *Wallis u. a.* 1986) – von Dritten übernommen werden. Inwiefern auch bei der Präsentation der Prognose die erfahrungsgemäßen Fehlermargen, z. B. nach Maßgabe eines Prognosefächers erfolgt, bedarf angesichts der dabei offen zutage tretenden Unsicherheiten der sorgfältigen Prüfung.[29]

---

[28] Zu einer skeptischen Einschätzung bezüglich weiterer Verbesserungsmöglichkeiten der gesamtwirtschaftlichen Prognose vor dem Hintergrund der Leistungsfähigkeit der makroökonomischen Theorie vgl. *Heilemann* (2004).

[29] Für die Prognose der GD vom Herbst 2004 für das Jahr 2005 ergab sich bei 50 vH (ex post) Sicherheit immerhin eine Bandbreite von 2,4 bis 0,6 vH Wirtschaftswachstum (*Heilemann / Blaschzik* 2004, zum Vergleich dazu auch *Wallis* 2003).

Ein anderer Schritt wäre – neben der Aufgabe der Institution „Gemeinschaftsdiagnose“ – die (häufigere) Anwendung des Instituts der Öffentlichen Ausschreibung der staatlichen Aufträge zur Erstellung gesamtwirtschaftlicher Prognosen. Im Fall der GD geschieht dies allerdings bereits. Die Ausschreibung erfolgt allerdings, mit plausiblen Gründen und vielleicht auch wegen eines bezüglich der Leistungsfähigkeit der Anbieter vermuteten positiven *home bias*, auf nationaler Ebene.[30]

Ob die radikalen Lösungen z. B. von *Jonung / Larch* (2004) in Anlehnung an die „Bestrafung“ der für die Geldpolitik Verantwortlichen im Falle von Zielverfehlungen in die richtige Richtung weisen, erscheint mehr als ungewiss, wenn man den gegenwärtigen Stand der Treffsicherheit als *state of the art* makroökonomischer Prognosegenauigkeit akzeptiert, von den Problemen der Fehlerbestimmung ganz abgesehen. Auf jeden Fall würde sie wohl erhebliche Entgeltverbesserungen angesichts des gestiegenen Berufsrisikos der Prognostiker bedeuten müssen.

Unabhängig von den hier vorgelegten numerischen Ergebnissen bezüglich der Treffsicherheit und der Beurteilung der wettbewerblichen Verfassung des Prognosemarktes stellt sich die Frage, auf welchen Wegen konkret „Mehr Wettbewerb“ die Treffsicherheit erhöhen sollte. Ein wettbewerbsbasiertes, aber von den üblichen drei Säulen der empirischen Wirtschaftsforschung, Daten – Theorien – Methoden, abweichendes Prognoseverfahren wird in den Vereinigten Staaten seit kurzem durch *prediction markets* entwickelt (z. B. *Wolfers / Zitzewitz* 2004). Unter der von den Finanzmärkten übernommenen Prämisse effizienter, Information perfekt verarbeitender Märkte lassen sich durch den freien Handel von Wertpapieren Aussagen über die Eintrittswahrscheinlichkeit, den Mittelwert oder Median spezieller Ereignisse in der Zukunft ableiten. Die Treffsicherheit der aus dem Handel solcher

---

[30] Zur Umsetzung der EU-Richtlinie vgl. VOL (A), Abschn. 2, §§ 2, 3. Ausnahmebereiche, für die eine Beschränkte Ausschreibung akzeptiert wird, ergeben sich nach §3, Nr. 3a, dann „wenn die Leistung nach ihrer Eigenart nur von einem beschränkten Kreis von Unternehmen in geeigneter Weise ausgeführt werden kann, besonders wenn außergewöhnliche Fachkunde oder Leistungsfähigkeit oder Zuverlässigkeit erforderlich ist“. Ähnliches schreibt § 3, Nr. 4 für die Freihändige Vergabe von Dienstleistungsaufträgen vor. Zusätzlich spielt hier eine Rolle, ob besonders hohe Entwicklungsleistungen, als welche z. B. die Erstellung eines makroökonometrischen Modells angesehen werden könnte, den Verzicht auf die Öffentliche Ausschreibung rechtfertigen.

Wertpapiere abgeleiteten Prognosen, z. B. über Wahlergebnisse, steht nicht hinter der herkömmlichen Verfahren, z. B. Umfragen, zurück. Erste Versuche werden auch hinsichtlich der Prognose ökonomischer Variabler wie Beschäftigung, Industrieproduktion oder Einzelhandelsumsätzen unternommen. Ein Urteil, ob sich ein *prediction market* als Prognoseinstrument für Indikatoren bewährt und ob seine Anwendung auch auf die mehrdimensionalen gesamtwirtschaftliche Prognosen ausgeweitet werden kann, steht indes noch aus.

Für den Fall traditioneller „wissenschaftlicher Prognosen“ ist davon auszugehen, dass die Prognosegenauigkeit das Ergebnis von „Theorie“, „Daten“ und „Methoden“ bzw. ihrer mehr oder weniger geschickten Kombination und anderer subjektiver Elemente ist. Die beiden zuletzt genannten sollen hier außer Betracht bleiben.[31] Die wissenschaftliche Theorieentwicklung und -bildung erfolgt im wissenschaftlichen System, bei dem üblicherweise „Wettbewerb“ unterstellt wird. Ob hier Intensivierungen bzw. Verbesserungen möglich sind, z. B. durch stärkere finanzielle Anreize, erscheint als offen, denn schon bislang waren diese ja – Stichwort: Finanzmärkte – gegeben, ohne bislang zu erkennbaren Fortschritten geführt zu haben. Hinzu kommt, dass im Fall von Prognosen ein spezifischer „Fortschritt“ gefordert ist: es geht nicht allein um eine bessere (*ex post*-)Erklärung, sondern um eine bessere Prognose (*ex ante*-Erklärung). Mit anderen Worten: das *Explanans* muss leichter zu erklären sein als das *Explanandum*, d. h. die „rechte Seite“ einer Gleichung (vgl. dazu B. oben (1)) leichter/besser als die linke. Dies ist allerdings eine Forderung, der sich kognitive Interessen selten, wenn überhaupt verpflichtet fühlen. Das Problem stellt sich naturgemäß im Fall mehrdimensionaler, konjunktiver Prognosen wie es Konjunkturprognosen in potenzierter Form sind. *Comte*’s *„savoir pour prévoir, prévoir pour prévenir“* ist – bestenfalls – ein langfristiges Ziel kognitiven Interesses. – Die Datenbasis der gesamtwirtschaftlichen Prognosen wird im Allgemeinen von der Amtlichen Statistik erstellt. Ohne Frage besteht auch hier ein gewisser Wettbewerb. Er findet jedoch in Europa nur indirekt auf der nationalen Ebene statt, primär auf der transnationalen, d. h. zwischen den Vereinigten Staaten und der EU und betrifft – bei durchaus unterschiedlichen qualitativen Standards

---

[31] Dass gerade für die Kombination der Inputs oder für subjektive Faktoren wie Erfahrung, Fingerspitzengefühl durch mehr Wettbewerb weitere Konzentrationsvorteile erschließbar seien, ist zweifelhaft (siehe oben, *Zarnowitz* (1992).

– die Veröffentlichungsfristen.[32] Sieht man von der expandierenden Produktion von Indikatoren ab, so scheint die Produktion „eigener“ Daten mit der Ausweitung des Datenangebotes der Amtlichen Statistik abgenommen zu haben und sich auf die Bereitstellung komplementärer Informationen zu konzentrieren. Bessere Daten, aus denen bessere Prognosen ableitbar wären, scheinen sich die einzelnen Institute von eigener Produktion nicht zu versprechen. – Ein vergleichsweise scharfer Wettbewerb herrscht vor allem auf dem methodischen Gebiet (analog zu *Daub* 1987, S. 15 ff.). In der Tat drängt sich dieser Eindruck auf, wenn man sich zum einen die Vielzahl von Methodenvergleichen und zum andern die hohe Visibilität und das vermutlich hohe Prestige „moderner“ Methoden bei den Prognosekonsumenten vor Augen führt. Dennoch konnte sich in der Vielzahl der Methodenvergleiche kein Verfahren als durchgängig bestes durchsetzen.[33] Faktisch bewegt sich die Verwendung moderner Methoden z. B. in gesamtwirtschaftlichen Prognosen in engen Grenzen, nicht zuletzt weil sich auch innerhalb der Prognoseinstitutionen ein zunehmender Methodenpluralismus oder -eklektizismus ausbreitet und gleichzeitig die Grenzen zwischen den Methoden verschwinden. An Wettbewerb fehlt es jedenfalls bei den Methoden der gesamtwirtschaftlichen Prognose nicht. Ob er stets in die richtige Richtung geht, mag man angesichts der großen Beliebtheit von „Prognose ohne Erklärung“, also Zeitreihenverfahren, bezweifeln, aber das steht auf einem anderen Blatt.

---

[32] Im Kern handelt es sich dabei um Reaktionen auf Forderungen der Finanzmärkte, bei denen Qualitätsaspekte der Statistik möglicherweise weniger interessieren als Aktualität und auch drastische Revisionen (Beispiel Vereinigte Staaten) durchaus in Kauf genommen werden. Vgl. Europäische Kommission (Hrsg.) 2005, z. B. S. 21ff.

[33] In den drei M-*competitions* hoben sich komplexe Modelle nicht positiv von einfachen Verfahren ab. Zudem variierte die Rangfolge je nach Prüfmaß, Prognosehorizont und Methodenbündel. Siehe – auch zu einer Zusammenfassung bisheriger Methodenwettbewerbe im Prognosebereich – *Makridakis / Hibon* (2000).

Literatur

Arbeitsgemeinschaft deutscher wirtschaftswissenschaftlicher Forschungsinstitute e.V. (1967 ff.): Die Lage der Weltwirtschaft und der deutschen Wirtschaft im Herbst 1966 ff. Essen u. a. O.: Arbeitsgemeinschaft.

Arrow, K. (1962): Economic welfare and the allocation of resources for invention. In: National Bureau of Economic Research (ed.), The rate and direction of inventive activity: economic and social factors. Princeton, N.J. : Princeton Univ. Press, S. 609-626.

Ash, J. C. K. / *Smyth*, D. J. / *Heravi*, S. M. (1993): The accuracy of OECD forecasts for Canada and the United States. North American Journal of Economics & Finance 4 (2), S. 179-210.

Ash, J. C. K. / *Smyth*, D. J. / *Heravi*, S. M. (1997): The accuracy of OECD forecasts for Japan. Pacific Economic Review 2 (1), S. 25-44.

Ballis, B. (1989): A post-mortem on OECD short term projections from 1982 to 1987. In: OECD, Department of Economics and Statistics (ed.), Working Paper 65. Paris: OECD.

Bartling, H. (1983): Wettbewerbliche Ausnahmebereiche. Rechtfertigung und Identifizierung. In: *Feldsieper*, M.; *Groß*, R. (Hrsg.), Wirtschaftspolitik in weltoffener Wirtschaft. Berlin: Duncker & Humblot, S. 325-346.

Beckmann, M. J. (1987): Wissenschaft und Wettbewerb: Versuch einer ökonomischen Theorie der Wissenschaft. In: *Korte*, B. (Hrsg.), Analysen und Prognosen wirtschaftlicher Entwicklungen. Bonn: Bouvier, S. 58-73.

Blix, M. / *Wadefjord*, J. / *Wienecke*, U. / *Ådahl*, M. (2001): How good is the forecasting performance of major institutions? Sveriges Riksbank Economic Review (3), S. 38-68.

Böhme, G. (1978): Autonomisierung und Finalisierung. In: Böhme, G. u. a. (Hrsg.), Starnberger Studien I – Die gesellschaftliche Orientierung des wissenschaftlichen Fortschritts. Frankfurt a. M.: edition suhrkamp, S. 69-130.

Bössmann, E. (1988[1982]): Information. In: Albers, W. u. a. (Hrsg.): Handwörterbuch der Wirtschaftswissenschaften, Bd. 4. Ungekürzte Studienausgabe. Stuttgart/New York: Fischer, Tübingen: Mohr (Siebeck), Göttingen/Zürich: Vandenhoeck & Ruprecht, S. 184-200.

Brooks, M. (2004): Extreme measures – the dark visions and bright ideas of Francis Galton. New York: Bloomsbury.

Bundesregierung (1967 ff.): Jahreswirtschaftsbericht der Bundesregierung. Bonn: Hegner. Nach 1994, Köln: Bundesanzeiger.

Cassel, S. (2001): Politikberatung und Politikerberatung – Eine institutionenökonomische Analyse der wissenschaftlichen Beratung der Wirtschaftspolitik. Bern: Paul Haupt.

Conseil National de L'Information Statistique (o.J.), http://www.cnis.fr/.

Council of Economic Advisers (2005): Annual Report. Economic Report of the President. Washington, DC: United States Government Printing Office.

Daub, M. (1987): Canadian economic forecasting: In a world where all's unsure. Kingston, Montreal: McGill-Queen's University Press.

Deutscher Bundestag (Hrsg.) (2003): Beschlussempfehlung und Bericht des 1. Untersuchungsausschusses nach Artikel 44 des Grundgesetzes. Drucksache 15/2100 v. 24.11.2003. Köln: Bundesanzeiger.

Dicke, H. / *Glisman*, H. G. (2002): Konjunkturprognosen und wissenschaftlich-technischer Fortschritt. Wirtschaftsdienst 82 (3), S. 167-169.

Dublin, M. (1989): Futurehype – the tyranny of prophecy. Markham, Ontario, Canada: Penguin Books.

Europäische Kommission (Hrsg.) (2005), Konjunkturindikatoren: Prioritätsbewertung, Qualität, Revisionen und Aktualität. Ergebnisse der 90. DGINS-Konferenz, Paris, 13.-14. September 2004. Luxembourg: Europäische Kommission.

Fildes, R. / *Stekler*, H. O. (2002): The state of macroeconomic forecasting. Journal of Macroeconomics 24 (4), S. 435-468.

Fögen, M. T. (1997): Die Enteignung der Wahrsager – Studien zum kaiserlichen Wissensmonopol in der Spätantike. Frankfurt a. M.: Suhrkamp.

Friedman, M. (1953): The methodology of positive economics. In: ders., Essays in positive economics. Chicago: University of Chicago Press, S. 3-43.

Grafton, A. (2001): Cardano`s cosmos: the worlds and works of a renaissance astrologer. Cambridge, MA: Harvard University Press.

Grunberg, E. / *Modigliani*, F. (1954): The predictability of social events. Journal of Political Economy 62, S. 465-478.

Guida, R. A. (1989): The costs of free information. Public Interest 97, S. 87-95.

Heilemann, U. (1980): Zur Praxis ökonometrischer Modelle in den USA. Jahrbücher für Nationalökonomie und Statistik 195 (1), S. 61-76.

Heilemann, U. (1990): Die Prognosen für 1988 im Rückblick: Der Börsenkrach vom 19. Oktober 1987 und seine gesamtwirtschaftlichen Folgen. Kredit und Kapital, 23. Jg. (1990), S. 174-195.

Heilemann, U. (2002): Increasing the transparency of macroeconomic forecasts: a report from the trenches. International Journal of Forecasting 18, S. 85-105.

Heilemann, U. (2003a): Have macroeconomic forecasts improved? The case of Germany 1967 to 2001. Essen, mimeo.

Heilemann, U. (2003b): Wahlen, Wirtschaftspolitik und Glück. In: Schäfer, W. (Hrsg.), Konjunktur, Wachstum und Wirtschaftspolitik im Zeichen der New Economy. Schriften des Vereins für Sozialpolitik 293. Berlin: Duncker & Humblot, S. 69-106.

Heilemann, U. (2004): Besser geht`s nicht – Genauigkeitsgrenzen von Konjunkturprognosen. Jahrbücher für Nationalökonomie und Statistik 224 (1+2), S. 51-64.

Heilemann, U. / *Blaschzik*, A. (2004): Que sera, sera? Zur Genauigkeit von Konjunkturprognosen. Universitätsjournal – Mitteilungen und Berichte für die Angehörigen und Freunde der Universität Leipzig 14, S. 8-9.

Heilemann, U. / *Stekler*, H. (2003): Have macroeconomic forecasts improved? The case of Germany 1967 to 2001. SFB 475 Universität Dortmund – Technical Report 31/2003. Dortmund.

Hirshleifer, J. (1971): The private and social value of information and the reward to inventive activity. American Economic Review 61 (4), S. 561-574.

Jokinen, H. (1993): Economic forecasting in the European Commission – a post mortem on the accuracy of the DGII economic forecasts. Ministry of Finance Economics Department (ed.), Discussion Paper, 39. Helsinki.

Jonung, L. / *Larch*, M. (2004): Improving fiscal policy in the EU. The case for independent forecasts. In: European Commission (ed.), Economic Paper, 210. Brüssel: EU.

Keereman, F. (1999): The track record of the Commission forecasts. In: European Commission (ed.), Economic Paper, 137. Brüssel: EU.

Kirchgässner, G. (1993): Testing weak rationality of forecasts with different time horizons. Journal of Forecasting 12, S. 541-558.

Klodt, H. (2003): Wettbewerbsstrategien für Informationsgüter. In: Schäfer, W. (Hrsg.), Konjunktur, Wachstum und Wirtschaftspolitik im Zeichen der New Economy. Schriften des Vereins für Socialpolitik, Bd. 293. Berlin: Duncker & Humblot, S. 107-123.

Koutsogeorgopoulou, V. (1998): A post-mortem on economic outlook projections. OECD Economics Department (ed.), Working Papers, 274. Paris: OECD.

Kreinin, M. (2000): Accuracy of OECD and IMF projections. Journal of Policy Modeling 22, S. 161-79.

Kuhlen, R. (2001): Wert der Information: Ware oder öffentliches Gut? Vortrag beim Hearing zur Umsetzung der Eu-Richtlinie 2001/29/EG des Europäischen Parlaments zur „Harmonisierung bestimmter Aspekte des Urheberrechts…“ in das deutsche Urheberrecht. Als Internetdokument am 8.3.2005 unter http://www.nethics.net/nethics/de/themen/urheberrecht/wert_der_information_ware_oder_oeffentliches_gut.html.

Mackay, C. (1992 [1852]): Zeichen und Wunder – Aus den Annalen des Wahns. Frankfurt a. M.: Eichborn.

Makridakis, S. / *Hibon*, M. (2000): The M3-Competition: results, conclusions and implications. International Journal of Forecasting 16 (4), S. 451-476.

McNees, S. K. (1986): Forecasting accuracy of alternative techniques: a comparison of US macroeconomic forcasts. Journal of Business and Economic Statistics 4 (1) S. 5-15.

Morgenstern, O. (1928): Wirtschaftsprognose – Eine Untersuchung ihrer Voraussetzungen und Möglichkeiten. Wien: Julius Springer.

OECD (1967 ff.): Economic outlook 1 ff. Paris: OECD.

Öller, L.-E. / *Barot*, B. (2000): The accuracy of European growth and inflation forecasts. International Journal of Forecasting 16 (3), S. 293-315.

Olten, R. (1995): Wettbewerbstheorie und Wettbewerbspolitik. München, Wien: Oldenbourg.

Pons, J. (2000): The accuracy of IMF and OECD forecasts for G 7 countries. Journal of Forecasting 19, S. 53-63.

Rheinisch-Westfälisches Institut für Wirtschaftsforschung (1996): RWI-Konjunkturberichte, Bd. 47, Berlin: Duncker & Humblot.

Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung (1967 ff.): Jahresgutachten 1967/68 ff.. Stuttgart, Mainz: Kohlhammer. Ab 1989: Stuttgart: Metzler-Poeschel.

Sherden, W. A. (1998): The fortune sellers – the big business of buying and selling predictions. New York: Wiley.

Smith, R. (1994): The macromodelling industry. In: Hall, S. (ed.), Applied economic forecasting techniques. New York: Harvester Wheatsheeaf, S. 68-88.

Surowiecki, J. (2005): Don’t do the math. The New Yorker, January 24 & 31, S. 38.

Tichy, G. (1994): Konjunktur – Stilisierte Fakten, Theorie, Prognose. Berlin: Springer.

Urban, P. (1973): Zur wissenschaftstheoretischen Problematik zeitraumüberwindender Prognosen. Untersuchungen des Instituts für Wirtschaftspolitik an der Universität zu Köln, 28. Köln: Inst. f. Wirtschaftspolitik an d. Univ. zu Köln.

Viner, J. (1958): The long view and the short: Studies in economic theory and policy. Glencoe, IL: Free Press.

Wallis, K. F. (2003): Chi-squared tests of interval and density forecasts, and the Bank of England's fan charts. International Journal of Forecasting 19 (2), S. 165-175.

Wallis, K. F. / *Andrews*, M. J. / *Fisher*, P. G. / *Longbottom*, J. A. / *Whitley*, J. D. (1986): Models of the UK Economy. A Third Review by the ESRC Macroeconomic Modelling Bureau. Oxford: Oxford University Press.

Wild, J. (1974): Grundlagen der Unternehmensplanung. rororo studium, 26. Reinbek bei Hamburg: Rowohlt.

Willgerodt, H. (2004): Die Anmaßung von Unwissen. ORDO Jahrbuch für die Ordnung von Wirtschaft und Gesellschaft 55, S. 25-35.

Wissenschaftsrat (1998): Stellungsnahme zu den Wirtschaftsforschungsinstituten der Blauen Liste in den alten Ländern. Drucksache 3320/98. Berlin.

Wolf, C. (1997): Markets or governments. Choosing between imperfect alternatives. 2. Aufl. Cambridge, MA.; London, U.K.: MIT Press.

Wolfers, J. / *Zitzewitz*, E. (2004): Prediction markets. Journal of Economic Perspectives 18 (2), S. 107-126.

Zarnowitz, V. (1992): Has macro-forecasting failed? in: Cato Journal 12 (1), S. 129-160.

Zarnowitz, V. (1996): Business cycles – theory, history, indicators, and forecasting. Chicago: NBER.

Anhang

Verwendete Prognosefehlermaße

Bias

$$\mathrm{BIAS} = \frac{1}{T}\sum_{t=1}^{T}\left(p_t - a_t\right)$$

mit:
p: Prognosewert
a: Beobachtungswert
T: Anzahl der Beobachtungen

Durchschnittlicher absoluter Prognosefehler (Mean-absolute-error)

$$\mathrm{MAE} = \frac{1}{T}\sum_{t=1}^{T}\left|p_t - a_t\right|$$

Durchschnittlicher Prognosefehler (Root-mean-square-error)

$$\mathrm{RMSE} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(p_t - a_t)^2}$$

Durchschnittlicher relativer Prognosefehler (Root-mean-square-percentage-error)

$$\mathrm{RMSPE} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\frac{p_t - a_t}{a_t}\right)^2}\cdot 100$$

*Theil*scher Ungleichheitskoeffizient

$$\mathrm{U} = \sqrt{\frac{\sum_{t=1}^{T}\left(p_t - a_t\right)^2}{\sum_{t=1}^{T}a_t^2}}$$

Mittelwertfehler

$$\mathrm{UM} = \frac{\left(\bar{p} - \bar{a}\right)^2}{\frac{1}{T}\sum_{t=1}^{T}\left(p_t - a_t\right)^2}$$

Varianzfehler

$$\mathrm{UV} = \frac{\left(s_p - s_a\right)^2}{\frac{1}{T}\sum_{t=1}^{T}\left(p_t - a_t\right)^2}$$

mit:

s_p, s_a: Standardabweichungen

Kovarianzfehler

$$\mathrm{UC} = \frac{2\left(1 - r\right)s_p s_a}{\frac{1}{T}\sum_{t=1}^{T}\left(p_t - a_t\right)^2}$$

mit:

r: Korrelationskoeffizient zwischen p und a