

IBES DISKUSSIONSBEITRAG

Institut für Betriebswirtschaft und Volkswirtschaft

Nr. 208

MÄRZ 2015

**Der Prognostiker des Jahres
– ein Zufallsergebnis?**

**Möglichkeiten einer mehrdimensionalen
Evaluierung von Konjunkturprognosen**

Prof. Dr. Roland Döhrn

IBES

IBES DISKUSSIONSBEITRAG

Nr. 208

März 2015

Der Prognostiker des Jahres – ein Zufallsergebnis?

Möglichkeiten einer mehrdimensionalen Evaluierung von Konjunkturprognosen

Prof. Dr. Roland Döhrn (roland.doehrn@rwi-essen.de)

Impressum: Institut für Betriebswirtschaft und Volkswirtschaft (IBES)
Universität Duisburg-Essen
Universitätsstraße 12
45141 Essen
E-Mail: IBES-Diskussionsbeitrag@medman.uni-due.de

Inhaltsverzeichnis

Kurzfassung.....	4
1. Einleitung	5
2. Methodische Grundlagen	6
3. Mehrdimensionale Bewertung von Konjunkturprognosen	8
3.1 Die Daten	8
3.2 Abstandsmaße auf Basis unterschiedlicher Verlustfunktionen.....	9
3.3 Sensitivität gegenüber der Auswahl der Variablen.....	10
3.4 Korrelationen zwischen Prognosen und Prognosefehlern.....	12
4. Einordnung der Prüfmaße.....	15
4.1 Vergleich mit naiven Prognosen	15
4.2 Bewertung vor dem Hintergrund von Datenrevisionen	15
5. Schlussfolgerungen.....	18
Literatur.....	19
Anhang.....	20

Kurzfassung

Man findet in zunehmendem Maße Versuche, Prognosen nicht nur danach zu bewerten, wie hoch ihre Treffsicherheit bei der Vorhersage einer Variablen ist, sondern die Prognosegüte bei einem Satz von Variablen ins Bild nehmen. Eine solche mehrdimensionale Evaluierung von Prognosen stellt zum Teil Neuland dar. Der vorliegende Beitrag stellt einige Kennziffern vor, die auf in der multivariaten Analyse gebräuchlichen Abstandsmaßen basieren und die für eine solche Evaluierung herangezogen werden können. Dabei zeigt sich, dass die Bewertung von Prognosen stark von dem gewählten Abstandsmaß und von der Auswahl der in die Evaluierung einbezogenen Variablen abhängt. Ein besonderes Problem resultiert dabei daraus, dass die prognostizierten Variablen in der Regeln nicht unabhängig von einander sind, Dies legt nahe, die Mahalanobis-Distanz als Abstandsmaß zu verwenden an Stelle der gebräuchlicheren Eulerschen Distanz oder einer City-Block-Metric. Allerdings bestraft die Mahalanobis-Distanz solche Prognosefehler härter, die den in der Vergangenheit beobachteten Korrelationen zwischen den Variablen widersprechen. Berücksichtigt man die Unsicherheit in den Daten, die als Referenz für die Bewertung der Prognosen herangezogen werden, so zeigt sich allerdings, dass die Abstandsmaße mit einem relativ großen Unschärfebereich zu versehen sind. Daher sollten aus kleinen Unterschieden in den Distanzmaßen keine allzu weitgehenden Schlüsse gezogen werden.

Stichworte: Konjunkturprognosen, Prognoseevaluation, Mehrdimensionale Evaluierung, Eulersche Distanz, Mahalanobis-Distanz

JEL: C52, E17

I. Einleitung¹

Prognose-Rankings erfreuen sich derzeit in der deutschen Presse einer gewissen Beliebtheit. So kürte die *Financial Times Deutschland* bis zum Einstellen ihres Erscheinens einen „Prognostiker des Jahres“; dieses Ranking wird mittlerweile von der *Süddeutschen Zeitung* fortgeführt (ftd/SZ-Ranking). Das *Handelsblatt* veröffentlichte erstmals Anfang 2014 eine Auswertung zur Treffsicherheit von Prognosen, die sich sowohl methodisch als auch hinsichtlich des Kreises der erfassten Prognosen von dem ftd/SZ-Ranking unterscheidet. Gemeinsam ist beiden Ranglisten jedoch, dass sie eine mehrdimensionale Sichtweise verfolgen. Geachtet wird nicht nur auf die Genauigkeit, mit der Wirtschaftswachstum – gemessen an der Veränderungsrate des realen Bruttoinlandsprodukts – prognostiziert wurde, sondern es werden weitere Variablen in die Bewertung einbezogen. Damit soll der Qualität der Begründung der Prognosen Rechnung getragen werden.

Beim ftd/SZ-Ranking geschieht dies in einer eher qualitativen Weise (SZ 2014). In seinem Kern steht die Prognose des BIP. Ergänzend wird mit einem Ampelsystem gekennzeichnet, ob die Prognosen des privaten Konsums, der Ausrüstungsinvestitionen und der Exporte in einem Konfidenzband um die beobachteten Werte liegen. Befinden sich die Vorhersagen einer oder mehrere dieser drei Variablen außerhalb des Konfidenzbandes, so wird die Prognose bei gleicher Treffsicherheit der BIP-Prognose in der Rangfolge nach unten gesetzt.

Das Handelsblattranking berechnet hingegen eine Kennziffer, die die Prognosegenauigkeit von sieben Variablen (BIP, Private Konsumausgaben, Anlageinvestitionen, Exporte, Importe, Inflationsrate und Arbeitslosenquote) zusammenfasst. Grundlage dieser Kennziffer ist der Mittlere Absolute Prognosefehler der jeweiligen Variablen, der um deren Streuung in der Vergangenheit bereinigt wird. Dadurch wird der Tatsache Rechnung getragen, dass die prognostizierten Variablen unterschiedlich stark im Zeitablauf schwanken. So werden Fehler bei stark schwankenden Größen geringer gewichtet als solche bei wenig variierenden.

Bewegt man sich bei der Evaluierung einzelner Prognosen auf einem statistisch wie methodisch vergleichsweise sicherem Terrain, so sind solche mehrdimensionalen Bewertungen von Prognosen bisher methodisch kaum untersucht worden – ein Beispiel ist Sinclair et al. (2013). Ihr Ergebnis dürfte von zahlreichen Faktoren abhängen. Bei Evaluierung der Prognosen nur einer Variablen stellt die Wahl der Verlustfunktion den einzigen Freiheitsgrad dar. Die mehrdimensionale Bewertung hängt außerdem von der Auswahl der einbezogenen Indikatoren ab sowie davon, ob und wie Interdependenzen zwischen den ausgewählten Variablen berücksichtigt werden sollen. So ist eine Prognose, die eine zu optimistische Einschätzung des BIP liefert, zugleich aber auch die Investitionen und die Importe zu hoch einschätzt möglicherweise hinsichtlich ihrer Qualität anders zu bewerten als eine Prognose, die über-optimistisch bezüglich des BIP ist, aber zu pessimistisch bezüglich der Investitionen und der Importe. Die beiden letztgenannten Größen sind nämlich erfahrungsgemäß positiv mit dem BIP korreliert, so dass es nur konsequent wäre, mit einer hohen BIP-Schätzung eine hohe Investitions- und Importprognose zu verbinden.

Ziel des vorliegenden Beitrags ist eine systematische Auseinandersetzung mit Verfahren einer mehrdimensionalen Evaluierung von Prognosen. Jede Konjunkturprognose bietet Vorhersagen für eine

¹ Der Verfasser dankt Heinz Gebhardt, Svetlana Rujin und Torsten Schmidt für hilfreiche Anmerkungen zu früheren Fassungen dieses Beitrags.

ganze Palette ökonomischer Variablen. Will man eine Auswahl davon mit den entsprechenden Realisationen vergleichen, so bietet es sich an, hierfür die in der multivariaten Analyse gebräuchlichen Distanzmaße zu verwenden. Sie werden in Abschnitt 2 dargestellt. In Abschnitt 3 werden diese Distanzmaße auf ein Beispiel angewendet, nämlich die in die Handelsblatt-Rankings für die Jahre 2013 und 2014 einbezogenen Prognosen. Die Berechnungen zeigen, wie stark die Wahl der Verlustfunktion die Rangfolge beeinflusst und welche Rolle die Variablenauswahl bei der Bewertung von Prognosen spielt. Noch stärker verschiebt sich die Rangfolge der Prognosen, wenn man Interdependenzen zwischen den prognostizierten Variablen berücksichtigt. Im vierten Abschnitt wird die Fragestellung erweitert und nach Bewertungsmaßstäben für die berechneten Distanzmaße gesucht. Diese beschreiben nämlich nur den Abstand zwischen zwei Prognosen, sagen aber wenig darüber aus, ob er gemessen an Alternativprognosen oder gemessen an der Datenunsicherheit groß oder klein ist. Die Berechnungen zeigen, dass zumindest in den beiden hier betrachteten Jahren bereits kleine Unterschiede in den Abstandsmaßen zu großen Unterschieden in den Rängen führen, ohne dass man daraus zwangsläufig auf eine Überlegenheit der besser platzierten Prognosen schließen kann. Der anschließende Abschnitt 5 fasst die Erkenntnisse der hier vorgenommenen Analysen zusammen und liefert einige Hinweise, worauf man bei einer mehrdimensionalen Bewertung von Prognosen achten sollte.

Im Mittelpunkt der folgenden Ausführungen steht der methodische Aspekt. Auf eine Kommentierung der Prognosegenauigkeit einzelner Prognostiker wird daher weitgehend verzichtet, zumal eine Bewertung von deren Fähigkeiten auf der Grundlage von nur zwei analysierten Prognosen ohnehin nicht möglich sein dürfte.

2. Methodische Grundlagen

In formaler Hinsicht stellt sich das Problem des Vergleichs eines Satzes prognostizierter Variablen mit den entsprechenden realisierten Werten wie folgt dar: P ist ein Vektor von n prognostizierten Variablen p_i und R ist ein Vektor von den n Realisationen r_i .

$$(1) \quad P = \{p_1, p_2, p_3 \dots p_n\}$$

$$(2) \quad R = \{r_1, r_2, r_3 \dots r_n\}$$

Diese Darstellung legt nahe, die (mehrdimensionale) Treffsicherheit der Prognosen anhand der in der multivariaten Analyse gebräuchlichen Distanzmaße D vorzunehmen. Wie diese Distanz aussieht, hängt von der gewählten Verlustfunktion V ab.

$$(3) \quad D = V(R - P)$$

Der einfachste Fall eines Distanzmaßes wäre die City-Block-Metrik (Backhaus et al. 2011, 544). In ihr werden die absoluten Differenzen der jeweiligen Wertepaare aufaddiert, was eine lineare Verlustfunktion impliziert.

$$(4) \quad D_{CB} = \sum_i |r_i - p_i|$$

Bei der Evaluierung einer einzelnen prognostizierten Größe entspricht dieses Abstandsmaß dem mittleren absoluten Prognosefehler. Bei multivariater Betrachtung hängt die Größe der Distanz aber hier wie auch von der Dimensionierung der einbezogenen Variablen ab. Im Niveau höhere Variablen beeinflussen die Distanz stärker als Variablen, die durch kleine Zahlen gekennzeichnet sind. Um zu ver-

hindern, dass D wesentlich durch die Dimensionierung geprägt wird, ist es notwendig, mit standardisierten Variablen zu arbeiten.

$$(5) \quad \check{r}_i = \frac{r_i - \mu_{r,i}}{\sigma_{r,i}} \text{ bzw. } \check{p}_i = \frac{p_i - \mu_{p,i}}{\sigma_{p,i}}$$

Dabei sind $\mu_{r,i}$ bzw. $\mu_{p,i}$ die Mittelwerte der Realisationen bzw. der Prognosen und $\sigma_{r,i}$ bzw. $\sigma_{p,i}$ die entsprechenden Standardabweichungen. Da für unverzerrte und effiziente Prognosen aber gilt

$$(6) \quad \mu_{r,i} = \mu_{p,i} = \mu_i \text{ bzw. } \sigma_{r,i} = \sigma_{p,i} = \sigma_i,$$

erhält man als standardisierte City-Block Metric

$$(7) \quad \check{D}_{CB} = \sum_i \frac{|r_i - p_i|}{\sigma_i}.$$

Ein solches Abstandsmaß liegt in modifizierter Form dem Handelsblatt-Ranking zu Grunde. Die Modifikation resultiert daraus, dass die einzelnen Prognosen mit unterschiedlichem Gewicht in das Abstandsmaß eingehen. Um ein Übergewicht des Außenhandels in der Bewertung der Prognosen zu verhindern, werden die Prognosen für die Ausfuhren und für die Einfuhren jeweils mit dem Gewicht 0,5 versehen.

Gebräuchlich ist die Euklidische Distanz D_{EU} .

$$(8) \quad D_{EU} = \sqrt{\sum_i (r_i - p_i)^2}$$

bzw. in standardisierter Form

$$(9) \quad \check{D}_{EU} = \sqrt{\sum_i \frac{(r_i - p_i)^2}{\sigma_i^2}}.$$

Die Euklidische Distanz impliziert eine quadratische Verlustfunktion. Sie korrespondiert mit der Wurzel aus dem mittleren quadratischen Prognosefehler bei univariater Betrachtung.

Die Euklidische Distanz führt allerdings dann zu einer verzerrten Darstellung des Abstandes zwischen beiden Vektoren, wenn die r_i und entsprechend auch die p_i miteinander korreliert sind (Sinclair et al. 2013). In diesem Fall sollte man als Abstandsmaß die Mahalanobis-Distanz D_{MA} wählen. Diese ist definiert als

$$(10) \quad D_{MA} = (R - P)'W^{-1}(R - P).$$

Dabei ist W die Varianz-Kovarianz-Matrix der betrachteten Variablen, wobei erneut unterstellt wird, dass die Prognosen unverzerrt und effizient sind.

Diese Prüfmaße sollen im Folgenden zur Evaluierung der Prognosen herangezogen werden. Da dabei jeweils nur die Prognosen für ein Jahr betrachtet werden, müssen σ_i in (7) und (9) bzw. W in (10) aus den in der Vergangenheit beobachteten Werten der p_i abgeleitet werden.

3. Mehrdimensionale Bewertung von Konjunkturprognosen

3.1 Die Daten

Die folgenden Rechnungen basieren auf den Daten, die in das Prognoseranking des Handelsblatts eingehen. Für 2014 werden 26 Prognosen betrachtet, die in den letzten Wochen des Jahres 2013 erstellt wurden (Tabelle 1). Die älteste ausgewählte Prognose dürfte die der Gemeinschaftsdiagnose vom Herbst sein, die bereits Mitte Oktober 2013 abgeschlossen und die daher auf Grundlage einer noch deutlich unsichereren Datenbasis erstellt wurde als die anderen hier dargestellten.² Auch die Prognosen des Sachverständigenrates wurden noch vor Veröffentlichung der Volkswirtschafts-

Tabelle 1

Konjunkturprognosen ausgewählter Institutionen für das Jahr 2014 Veränderungen gegenüber dem Vorjahr in %, soweit nicht anders vermerkt

Institution	Kürzel	BIP	Privater Konsum	Anlageinvestitionen	Export	Import	Verbraucherpreise	Arbeitslosenquote in % ¹
Berliner Sparkasse	BES	1,7	0,9	4,1	4,8	5,2	1,2	6,8
Bundesbank	BUBA	1,7	1,4	4,3	3,9	5,1	1,3	6,8
Bayern LB	BLB	1,5	1,1	3,9	4,6	5,4	1,6	6,8
Feri	FERI	1,7	1,3	3,6	4,6	4,8	1,5	6,5
HSBC	HSBC	1,7	1,3	2,8	3,0	3,1	1,7	6,7
Deutsche Bank	DB	1,5	0,9	4,0	5,6	6,1	1,5	6,7
MM Warburg	MMW	2,3	1,2	4,2	5,0	4,2	1,5	6,8
IW	IW	1,5	1,0	2,5	3,5	3,5	2,0	6,5
IMK	IMK	1,2	1,4	4,2	4,5	6,1	1,5	7,0
RWI Essen	RWI	1,5	1,3	4,8	4,8	5,7	1,7	6,8
IWH Halle	IWH	1,8	1,6	4,5	4,0	5,0	1,5	6,5
ING	ING	1,7	0,5	4,4	3,7	3,2	1,6	6,9
DIW Berlin	DIW	1,6	1,3	4,7	5,3	6,1	1,6	7,0
Ifo Institut	IFO	1,9	1,5	4,5	5,9	6,5	1,5	6,8
Allianz	ALL	2	1,6	4,2	4,9	4,9	1,6	6,8
IKB	IKB	2	1,6	4,3	5,2	5,3	1,6	6,8
Bankenverband	BV	1,9	1,4	4,5	5,5	5,9	1,7	6,8
IfW Kiel	IFW	1,7	1,4	4,3	6,0	6,7	1,7	6,9
Commerzbank	COM	1,7	1,3	5,0	5,2	6,9	1,7	7,0
HWWI	HWWI	1,7	1,4	4,5	5,4	5,9	1,8	6,4
Sachverständigenrat	SVR	1,6	1,4	4,9	5,2	6,3	1,9	6,8
Helaba	HLB	2	1,3	4,0	6,0	6,0	2,0	6,7
Deka	DEKA	1,7	1,6	4,9	4,9	5,8	1,8	6,8
Gemeinschaftsdiagnose	GD	1,8	1,4	5,2	5,0	6,2	1,9	6,8
DZ Bank	DZ	2,3	1,4	5,5	7,4	7,8	2,1	6,8
Kiel Economics	KE	2,8	2	5,1	3,2	4,3	2,3	6,2
nachrichtlich								
Ist ²		1,5	1,1	3,1	3,7	3,3	0,9	6,7

Quelle: Handelsblatt (2015). – ¹ in % der gesamten Erwerbspersonen. – ² Erste Jahresergebnisse des Statistischen Bundesamtes.

2 Döhrn/Schmidt (2011) zeigen, dass die Treffsicherheit einer Prognose mit der Verringerung des zeitlichen Prognosehorizonts zunimmt. Ein fairer Vergleich von Prognosen sollte solchen Unterschieden im Informationsstand Rechnung tragen.

lichen Gesamtrechnungen für das dritte Quartal 2013 erstellt, das bereits wichtige Informationen für die Prognose des kommenden Jahres enthält. Es fehlen in der Zusammenstellung die Prognosen internationaler Organisationen. Es wäre teilweise auch schwierig, sie in einem solchen Vergleich einzubeziehen, da sie Verbraucherpreise und Arbeitslosenquoten oft in abweichender Abgrenzung prognostizieren. Dies gilt auch für das Jahr 2013, für das in von 2014 abweichender Zusammensetzung 25 Prognosen betrachtet wurden (Tabelle AI im Anhang). Die Bewertung der Prognosen erfolgt anhand der Ergebnisse für sieben Variablen: Jeweils die Vorjahresveränderungen des Bruttoinlandsprodukts, der privaten Konsumausgaben, der Anlageinvestitionen, der Exporte und der Importe, wobei – wie erwähnt – die beiden letztgenannten nur mit dem Gewicht 0,5 in die Bewertung eingehen. Hinzu kommen die Inflationsrate (gemessen am Verbraucherpreisindex) und die Arbeitslosenquote in der in Deutschland gebräuchlichen Definition.

Für die Berechnungen werden auch die die in der Vergangenheit beobachteten Werte der prognostizierten Variablen benötigt, um die Varianzen und Kovarianzen zu berechnen. Das Handelsblatt gibt an, diese aus den Daten der vergangenen 20 Jahre abzuleiten. Um vergleichbare Ergebnisse zu erhalten, werden auch hier für die Bewertung der Prognose 2014 die beobachteten Werte der Jahre 1994 bis 2013 verwendet und im Falle der Prognose 2013 die der Jahre 1993 bis 2012.

3.2 Abstandsmaße auf Basis unterschiedlicher Verlustfunktionen

In einem ersten Untersuchungsschritt wurde versucht, die Ergebnisse des Handelsblatts zu reproduzieren. Dies gelang nicht ganz; es bleibt eine kleine Differenz zwischen den Berechnungen hier und denen, die das Handelsblatt veröffentlicht hat. Ursache hierfür kann nur sein, dass die zur Normierung verwendeten Streuungen der prognostizierten Variablen in der Vergangenheit voneinander abweichen.³ Die Unterschiede sind aber gering und haben keinen Einfluss auf die Rangfolge der Prognosen (Tabelle 2).

Gewichtet man die Prognosen für alle Variablen gleich (Spalte D_{CB}), so sind die Unterschiede zum Handelsblatt-Ranking zwar gering, aber es kommt bereits zu einigen Verschiebungen in der Rangfolge. So schiebt sich die HSBC auf den dritten Rang vor, und ING verbessert sich um drei Ränge. Gravierender werden die Änderungen, wenn man eine quadratische Verlustfunktion wählt und daher die Euklidische Distanz berechnet. Zwar ändert sich nichts an den beiden Spitzenplätzen und an den beiden letztplatzierten Prognosen. Das gesamte Mittelfeld wird aber durcheinandergewirbelt; kaum eine Prognose behält ihren Rang. Am gravierendsten ändert sich die Position des IWV, das von Rang 8 im ursprünglichen Handelsblatt-Ranking auf Rang 21 zurückfällt.

3 Da die Unterschiede zwischen beiden Rechnungen mit den Fehlern der Inflationsprognose hoch korreliert sind, liegt der Verdacht nahe, dass kleine Unterschiede in den für die Vergangenheit verwendeten Inflationsraten ausschlaggebend sind.

Tabelle 2

Maße der Treffsicherheit der betrachteten Konjunkturprognosen für 2014

Kürzel ¹	Methode des Handelsblatts				City-Block Metric		Euklidische Distanz	
	Handelsblatt		Eigene Rechnungen		D _{CB}	Rang	D _{EU}	Rang
	D	Rang	D	Rang				
BES	1,39	1	1,38	1	1,64	1	0,72	1
BUBA	1,63	2	1,62	2	1,80	2	0,86	2
BLB	1,69	3	1,66	3	1,93	4	1,23	7
FERI	1,78	4	1,75	4	1,97	5	1,07	3
HSBC	1,80	5	1,76	5	1,84	3	1,33	9
DB	1,85	6	1,83	6	2,24	8	1,19	5
MMW	2,01	7	1,98	7	2,17	6	1,12	4
IW	2,26	8	2,20	8	2,24	7	1,80	21
IMK	2,27	9	2,24	9	2,56	11	1,22	6
RWI	2,34	10	2,31	10	2,62	12	1,46	15
IWH	2,35	11	2,32	11	2,51	10	1,23	8
ING	2,38	12	2,35	12	2,36	9	1,36	12
DIW	2,41	13	2,38	13	2,77	14	1,37	13
IFO	2,51	14	2,49	14	2,97	16	1,34	10
ALL	2,54	15	2,51	15	2,75	13	1,36	11
IKB	2,62	16	2,59	16	2,90	15	1,39	14
BV	2,65	17	2,61	17	3,00	17	1,51	16
IFW	2,69	18	2,65	18	3,15	18	1,56	17
COM	2,76	19	2,72	19	3,18	20	1,58	18
HWWI	2,84	20	2,80	20	3,18	19	1,65	19
SVR	2,95	21	2,90	22	3,31	22	1,82	22
HLB	2,95	22	2,90	21	3,34	23	1,93	24
DEKA	2,98	23	2,94	23	3,27	21	1,70	20
GD	3,09	24	3,04	24	3,42	24	1,83	23
DZ	4,08	25	4,02	25	4,74	25	2,33	25
KE	4,90	26	4,83	26	4,96	26	2,63	26

Quelle: Handelsblatt, eigene Berechnungen. – ¹Zu den Kürzeln siehe Tabelle 1.

Die Ergebnisse für das Jahr 2013 sind im Anhang dargestellt (Tabelle A2). Auch hier ließen sich die im Handelsblatt veröffentlichten Ergebnisse nicht exakt reproduzieren. Zwei Prognosen tauschen nach der Rechnung hier sogar ihren Rang.⁴ Ansonsten gilt auch hier, dass sich durch eine Gleichgewichtung aller Variablen wie auch durch den Übergang zu einer quadratischen Verlustfunktion das Ranking der Prognostiker ändert. Allerdings sind die Auswirkungen im Vergleich zu 2014 etwas geringer.

3.3 Sensitivität gegenüber der Auswahl der Variablen

Welche Variablen in die Evaluation der Prognosegüte eingehen, ist letztlich eine subjektive Entscheidung. In das Handelsblatt-Ranking gehen Kennziffern des Wachstums, der Inflation und des Arbeitsmarktes ein, ergänzt durch Verwendungskomponenten des BIP, um zu berücksichtigen, ob die Struk-

4 Das Gütemaß für das Jahr 2013 wurde im Handelsblatt nur mit einer Nachkommastelle ausgewiesen. Die hier festgestellten Abweichungen bewegen sich allesamt im Rahmen der Rundungsdifferenzen. Der erwähnte Rangtausch ist bedingt durch einen Unterschied in der dritten Nachkommastelle. Zu beachten ist, dass es im September 2014 zu einer großen Revision der Volkswirtschaftlichen Gesamtrechnungen kam. Daher basierte die Bewertung der Prognosen des Jahres 2013 auf einem anderen Datenstand als die Evaluation für 2014.

tur der gesamtwirtschaftlichen Expansion zutreffend vorhergesagt wurde. Warum dabei der Anstieg der Verbraucherpreise als Kennziffer ausgewählt wurde und nicht beispielsweise der den inländischen Preisauftrieb besser charakterisierende BIP-Deflator, oder warum die Arbeitslosenquote und nicht die Zunahme der Beschäftigung hier erscheinen, lässt sich allenfalls mit Praktikabilitätsgründen rechtfertigen: Die ausgewählten Kennziffern dürften zu jenen gehören, die man in allen Prognosen findet.

Tabelle 3

**Maße der Treffsicherheit bei alternativen Variablensätze
Korrelation mit den Prüfmaßen mit vollem Variablensatz**

Prüfmaß ohne ...	2013				2014			
	City-Block-Metric		Euklidischer Abstand		City-Block-Metric		Euklidischer Abstand	
	D _{CB}	Rang ¹	D _{EU}	Rang ¹	D _{CB}	Rang ¹	D _{EU}	Rang ¹
BIP	0,988	0,948	0,999	0,998	0,988	0,974	0,999	0,999
Konsum	0,924	0,927	0,932	0,892	0,966	0,954	0,991	0,977
Investitionen	0,955	0,910	0,973	0,966	0,992	0,952	0,999	0,988
Export	0,991	0,928	0,956	0,939	0,996	0,958	0,999	0,997
Import	0,991	0,931	0,962	0,923	0,991	0,937	0,993	0,989
Inflation	0,918	0,894	0,936	0,915	0,884	0,799	0,661	0,636
Arbeitsl.	0,989	0,933	0,999	0,996	0,996	0,973	1,000	0,997

Eigene Berechnungen. – ¹ Spearman'sche Rangkorrelation.

Um zu überprüfen, wie sensitiv die ermittelten Rangfolgen auf die Auswahl der Indikatoren reagieren, wurde hier jeweils ein Indikator aus dem Datensatz ausgeschlossen und anschließend ermittelt, wie stark sich die aufgrund der verbleibenden sechs Indikatoren ermittelte Rangfolge von der unter Verwendung aller sieben Indikatoren berechneten abweicht. Diese Unterschiede wurden für das absolute Abstandsmaß D_{CB} und für die Euklidische Distanz D_{EU} ermittelt. Tabelle 3 zeigt sowohl die Korrelation der Abstandsmaße selbst als auch die der auf deren Basis ermittelten Ränge.

Ein erstes Ergebnis erscheint auf den ersten Blick überraschend: Aus dem hier verwendeten Variablensatz hätte man zumindest in den beiden vergangenen Jahren das Bruttoinlandsprodukt, also eine an sich zentrale Größe jeder Konjunkturprognose, ausschließen können, ohne dass dies nennenswerte Auswirkungen auf die Abstandsmaße und auch auf die Rangfolge der Prognosen gehabt hätte. Diese gilt insbesondere bei Wahl einer quadratischen Verlustfunktion. Plausibel wird dieses Ergebnis allerdings, wenn man sich die Korrelation zwischen den Prognosefehlern der Variablen in der Vergangenheit anschaut. Tabelle 4 stellt dies beispielhaft für die Gemeinschaftsdiagnose vom Herbst, die Prognose des Sachverständigenrates sowie die Dezember-Prognosen des RWI und des ifo Instituts dar;⁵ andere Prognosen zeigen allerdings kein grundsätzlich abweichendes Bild. Die Prognosefehler bei BIP, Anlageinvestitionen, Export und Import sind hoch korreliert, so dass man kaum eine Information bezüglich der Prognosegüte verliert, wenn man eine dieser Variablen weglässt. Auf die sich darauf ergebenden Probleme wird noch einzugehen sein.

5 Die Daten zur Berechnung von Tabelle 4 wurden der Prognosedatenbank des RWI entnommen. In ihr sind keine Prognosen der Arbeitslosenquoten erfasst.

Tabelle 4

**Interkorrelation der Prognosefehler ausgewählter Prognosen
1991-2014**

	BIP	Konsum	Investitionen	Export	Import	Inflation
<i>Gemeinschaftsdiagnose vom Herbst</i>						
BIP	1,000	0,357	0,879	0,882	0,825	0,512
Konsum		1,000	0,220	0,173	0,359	-0,204
Investitionen			1,000	0,741	0,833	0,429
Export				1,000	0,875	0,550
Import					1,000	0,336
Inflation						1,000
<i>Sachverständigenrat</i>						
BIP	1,000	0,436	0,865	0,889	0,838	0,334
Konsum		1,000	0,295	0,253	0,483	-0,249
Investitionen			1,000	0,728	0,825	0,289
Export				1,000	0,853	0,466
Import					1,000	0,158
Inflation						1,000
<i>RWI, Dezember-Prognose</i>						
BIP	1,000	0,277	0,872	0,820	0,789	0,217
Konsum		1,000	0,223	-0,065	0,246	-0,269
Investitionen			1,000	0,641	0,775	0,186
Export				1,000	0,823	0,326
Import					1,000	0,063
Inflation						1,000
<i>ifo Institut, Dezemberprognose</i>						
BIP	1,000	0,328	0,810	0,824	0,783	0,157
Konsum		1,000	0,119	0,086	0,368	-0,357
Investitionen			1,000	0,684	0,804	0,123
Export				1,000	0,838	0,410
Import					1,000	0,123
Inflation						1,000

Eigene Berechnungen.

Kaum einen Einfluss auf die Rangfolge übt die Arbeitslosenquote aus, und zwar in beiden betrachteten Jahren und bei beiden Verlustfunktionen. Dies ist insofern nicht überraschend, als die Arbeitslosenquote die einzige Bestandsgröße im Datensatz ist und sie sich deshalb im Zeitablauf nur langsam ändert. Sie lässt sich daher leichter prognostizieren als Veränderungsraten, weshalb die Unterschiede in den Prognosen gering sind. Die einflussreichste Variable ist die Inflationsrate. Schließt man sie aus dem zur Evaluierung herangezogenen Datensatz aus, erhält man deutlich abweichende Prüfmaße und Rangfolgen. Dabei sind die Unterschiede bei quadratischer Verlustfunktion ausgeprägter als bei linearer, und der Effekt ist 2014 deutlicher als 2013. Im Jahr 2014 konnte lediglich das IWV einen größeren Fehler in der Inflationsprognose durch sehr geringe Fehler in anderen Prognosen ausgleichen. Ansonsten entspricht die Rangfolge der Prognosen anhand des Gesamtindex fast genauer jener Rangfolge, die sich ergeben würde, würde man die Prognosen allein aufgrund des Fehlers bei der Inflationsprognose anordnen.

3.4 Korrelationen zwischen Prognosen und Prognosefehlern

Die in Tabelle 4 dargestellten Korrelationen zwischen den Prognosefehlern machen deutlich, dass die Prognosen von BIP, Investitionen, Import und Export nicht unabhängig voneinander sind, so dass eine ähnliche Information mehrfach in die Evaluation eingebracht wird, wenn man alle vier Variablen berücksichtigt. Eine hohe Korrelation besteht nicht nur zwischen den Prognosefehlern, sondern auch zwischen den evaluierten Variablen selbst (Tabelle 5), was die Prognostiker in ihren Prognosen allem

Anschein nach widerspiegeln.⁶ Daher dürfte die Euklidische Distanz als Maß des Abstandes zwischen dem Vektor der beobachteten und der prognostizierten Variablen verzerrt sein und man müsste stattdessen die Mahalanobis-Distanz anwenden. Da hinter dieser auch eine quadratische Verlustfunktion steht, soll sie hier mit der Euklidischen Distanz verglichen werden (Tabelle 6).

Tabelle 5

Interkorrelation der prognostizierten Variablen in der Vergangenheit 1992-2013

	BIP	Konsum	Investitionen	Export	Import	Inflation	Arbeitslosenquote
BIP	1,000	0,487	0,892	0,910	0,847	0,450	0,099
Konsum		1,000	0,559	0,343	0,508	0,091	0,051
Investitionen			1,000	0,736	0,856	0,309	0,064
Export				1,000	0,881	0,346	0,326
Import					1,000	0,144	0,351
Inflation						1,000	-0,203
Arbeitslosenquote							1,000

Eigene Berechnungen

Die Rangfolge der Prognosen ändert sich bei Betrachtung der Mahalanobis-Distanz zum Teil beträchtlich. Zwar behält die Berliner Sparkasse ihren ersten Rang. Das IMK fällt aber beispielsweise von Rang 6 auf Rang 15 zurück, während sich die Allianz-Versicherung von Rang 11 auf Rang 5 verbessert. Diese Verschiebungen ergeben sich daraus, dass die Mahalanobis-Distanz nicht nur die Abweichung der Prognose vom realisierten Wert berücksichtigt, sondern auch die Zusammenhänge zwischen den realisierten Werten der Vergangenheit. Auch für die Mahalanobis-Distanz gilt zwar, dass die Distanzen umso geringer sind, je dichter die prognostizierten Werte bei den Realisationen liegen. Sie berücksichtigt darüber hinaus aber auch die logische Konsistenz der Prognosen: Weisen die Fehler bei in der Vergangenheit eng miteinander korrelierten Variablen das der Korrelation entsprechende Vorzeichen auf, wird dies von der Mahalanobis-Distanz weniger hart „bestraft“ als abweichende Vorzeichen bei solchen Variablen. Die Folgen dieser Logik soll hier am Beispiel der beiden angesprochenen Prognosen verdeutlicht werden, deren Ränge sich zwischen beiden Abstandsmaßen besonders stark unterscheiden.

Das IMK hatte für das Jahr 2014 eine vergleichsweise pessimistische Prognose des BIP (1,2%). Die Prognose der Investitionen (4,2%) war gemessen daran aber recht optimistisch. Die Bundesbank prognostizierte beispielsweise bei einer um 0,5 Prozentpunkte stärkeren Expansion des BIP eine fast gleich starke Zunahme der Investitionen. Auch die Exportprognose war, wenngleich in langfristiger Perspektive nicht allzu hoch, gemessen an den realisierten Werten zu optimistisch, und die Importprognose gehört zu den höheren unter den betrachteten. Diese Konstellation aus Unterschätzung des BIP und Überschätzung der drei anderen Größen bestraft die Mahalanobis-Distanz. Hätte das IMK einen Zuwachs des BIP um 1,8% prognostiziert, die tatsächliche Entwicklung also um 0,3 Prozentpunkte über- statt unterschätzt, wäre die Euklidische Distanz unverändert geblieben, die Mahalanobis-Distanz aber auf 4,38 gefallen.

6 Vergleicht man die Korrelationskoeffizienten in den Tabellen 4 und 5, so zeigt sich ein bemerkenswerter Unterschied: Die Prognosefehler von Konsum und Inflation sind bei allen Prognosen negativ korreliert, während sich in der Vergangenheit keine Korrelation zwischen beiden Größen zeigt. Dies weist darauf hin, dass die Konsumprognosen wesentlich auf einer Prognose der nominalen Einkommen basieren, aus denen je nach Inflation mehr oder weniger starke Zuwächse der Realeinkommen und des realen Konsums abgeleitet werden.

Tabelle 6

Euklidische Distanz und Mahalanobis-Distanz im Vergleich, 2014

Kürzel ¹	Euklidische Distanz		Mahalanobis Distanz	
	D _{eu}	Rang	D _{MA}	Rang
BES	0,72	1	1,37	1
BUBA	0,86	2	1,81	3
BLB	1,23	7	4,46	12
FERI	1,07	3	2,55	4
HSBC	1,33	9	4,34	10
DB	1,19	5	5,11	13
MMW	1,12	4	1,55	2
IW	1,80	21	7,09	21
IMK	1,22	6	6,51	15
RWI	1,46	15	6,52	16
IWH	1,23	8	2,74	6
ING	1,36	12	3,48	8
DIW	1,37	13	5,39	14
IFO	1,34	10	3,83	9
ALL	1,36	11	2,69	5
IKB	1,39	14	2,96	7
BV	1,51	16	4,45	11
IFW	1,56	17	6,69	18
COM	1,58	18	7,00	20
HWWI	1,65	19	6,74	19
SVR	1,82	22	8,93	24
HLB	1,93	24	7,35	22
DEKA	1,70	20	6,63	17
GD	1,83	23	7,54	23
DZ	2,33	25	10,04	25
KE	2,63	26	12,99	26

Quelle: Handelsblatt, eigene Berechnungen. – ¹Zu den Kürzeln siehe Tabelle 1.

Die Allianz hatte mit einem BIP-Wachstum um 2,0% eine der höchsten Prognosen unter den betrachteten Institutionen. Dies ging mit der Vorhersage einer Entwicklung von Exporten, Importen und Investitionen einher, die ebenfalls jeweils zu den höheren unter den hier analysierten gehört. Man stelle sich vor, die Allianz hätte eine Zunahme der Investitionen um lediglich 2% prognostiziert, also deutlich unterschätzt. Die Euklidische Distanz hätte dies nicht verändert; die Mahalanobis-Distanz würde aber auf 8,25 springen, weil einer Überschätzung von BIP, Export und Import eine Unterschätzung der mit den drei Variablen in der Vergangenheit hoch und positiv korrelierten Investitionen gegenüberstünde.

Deutlichere Verschiebungen als für 2014 ergeben sich in der Rangfolge der Prognosen, wenn man für 2013 die Mahalanobis-Distanzen als Kriterium heranzieht (Tabelle A2). Da die Abstände zwischen den Prognosen in diesem Jahr generell geringer waren, führt die Berücksichtigung der Kovarianzen der prognostizierten Variablen u.a. dazu, dass die Prognose von Kiel Economics von Rang 19 auf 3 nach vorn springt und die von MM Warburg vom zweiten auf den elften Rang zurückfällt.

Alles in allem stellt die Mahalanobis-Distanz ein zwar nicht unbedingt einfach zu interpretierendes Maß zur multidimensionalen Bewertung von Prognosen dar, sie liefert gleichwohl wichtige Erkenntnisse, die in den anderen Abstandsmaßen nicht enthalten sind.

4. Einordnung der Prüfmaße

Die bisher diskutierten Abstandsmaße beschreiben zwar Unterschiede zwischen Prognose und Realisation. Sie sagen aber nichts darüber aus, wie diese zu bewerten sind, ob und vor dem Hintergrund welcher Vergleichsmaßstäbe sie beispielsweise also als groß anzusehen sind. So zeigt sich gerade in den beiden hier betrachteten Jahren, dass sich die Treffsicherheit vieler Prognosen nur wenig unterscheidet, so dass bereits kleine Unterschiede in den Abstandsmaßen zu großen Unterschieden in den Rängen führen können. Im Folgenden sollen zwei Vergleichsmaßstäbe herangezogen werden: Zum einen – ähnlich wie bei univariaten Prognoseevaluations – ein Vergleich mit „naiven“ Prognosen, zum anderen eine Beurteilung der Prognosen vor dem Hintergrund der ohnehin vorhandenen Ungenauigkeiten in den Daten.

4.1 Vergleich mit naiven Prognosen

Bei der eindimensionalen Evaluation von Prognosen werden häufig Prüfmaße verwendet, die es erlauben, den Prognosefehler einzuordnen. Ein solches Prüfmaß ist der Theil'sche Ungleichheitskoeffizient U , der den Quadratischen Prognosefehler in Relation setzt zum Fehler einer Naiven Prognose. Dabei findet man in der Literatur zwei Varianten (Döhrn 2014: 157). Die erste unterstellt einen *random walk* in den Niveaus, was prognostizierte Veränderungsrate von 0 impliziert. Die zweite unterstellt einen einfachen autoregressiven Prozess erster Ordnung und geht davon aus, dass die Prognose der zuletzt beobachteten Veränderungsrate entspricht.⁷ Eine Mindestanforderung an eine ernst zu nehmende Prognose ist, dass U kleiner als 1 ist.

Tabelle 7

Maße der Treffsicherheit „naiver“ Konjunkturprognosen für 2013 und 2014

Kürzel ¹	Methode des Handelsblatts				City-Block-Metric		Euklidische Distanz	
	Handelsblatt		Eigene Rechnungen		D_{CB}	Rang	D_{EU}	Rang
	D	Rang	D	Rang				
2013								
Naiv I ¹	-	-	3,32	26	3,46	25	3,23	26
Naiv II ²	-	-	1,48	9	1,71	5	0,78	6
2014								
Naiv I ¹	-	-	4,88	28	5,49	28	2,34	27
Naiv II ²	-	-	3,22	25	3,68	25	1,59	19

Eigene Berechnungen – ¹Veränderungsrate gegenüber Vorjahr = 0, Arbeitslosenquote entspricht der des Vorjahres. – ²Prognosen entsprechen den zu Beginn des Jahres 2014 beobachteten Werten für 2014.

Tabelle 7 zeigt, dass im Jahr 2014 alle Prognosen mit einer Ausnahme bei Zugrundelegung der City-Block-Metric die erste Variante der naiven Prognosen schlagen. Die zweite Variante lässt auf Grundlage der Handelsblatt-Berechnung und der City-Block-Metric zwei der im Ranking betrachteten Prognosen hinter sich, auf Grundlage der Euklidischen Distanz sogar 8 Prognosen. Für 2013 fällt die Bilanz für die Prognostiker ungünstiger aus, obwohl die Prognosefehler gemessen an den diversen Abstandsmaßen generell kleiner sind als 2014: Der einfache autoregressive Prozess wäre – je nach gewähltem Abstandsmaß – auf Ränge zwischen 5 und 9 unter den Prognosen gekommen.

4.2 Bewertung vor dem Hintergrund von Datenrevisionen

⁷ Hier wurden die zu Beginn von 2014 beobachteten Werte für 2013 als Referenz gewählt, die aufgrund der zwischenzeitlichen Revision der VGR von den heute berichteten abweichen.

Eine andere Form der Relativierung der Ergebnisse erhält man, wenn man berücksichtigt, dass es sich bei den hier als Vergleich herangezogenen ersten Daten der Volkswirtschaftlichen Gesamtrechnungen lediglich um Schätzungen handelt, die im Laufe der an den Zeitpunkt der Veröffentlichung jeweils anschließenden drei Jahre noch mehrfach revidiert werden. Im Durchschnitt wurde die Veränderungsrate des realen BIP zwischen 1993 und 2010, dem derzeit letzten Jahr, für das endgültige Werte vorliegen, um knapp 0,5 Prozentpunkte verändert. In ähnlichem Maße änderte sich auch die Veränderungsrate der privaten Konsumausgaben. Noch größer sind die Änderungen bei Investitionen und Außenhandel. Kaum revisionsanfällig sind hingegen die Inflationsraten. Hier ergeben sich im Laufe der Zeit nur vereinzelt kleine Veränderungen des Jahresergebnisses, etwa wenn das Basisjahr für die Berechnung des Verbraucherpreisindex wechselt. Die Arbeitslosenquote wird zwar in dem Maße revidiert, wie sich die Angaben zum Nenner, also zur Erwerbstätigkeit ändern; allerdings sind die Änderungen hier äußerst gering.

Um die Aussagekraft der verwendeten Abstandsmaße zu evaluieren, bietet es sich an, die Prognosefehler nicht in Relation zu den vorliegenden ersten Veröffentlichungen zu berechnen, sondern in Relation zu den derzeit noch unbekanntem endgültigen Daten. Diese endgültigen Daten werden hier generiert, indem die vorläufigen Daten um in der Vergangenheit beobachtete Revisionen korrigiert werden.⁸ Da die Revisionen der einzelnen Verwendungskomponenten der Volkswirtschaftlichen Gesamtrechnungen nicht unabhängig voneinander sind, ist es hier nicht möglich, wie man bei einem *bootstrapping* vorgehen würde, den einzelnen Variablen unterschiedliche Korrekturen zuzulassen. Vielmehr werden zu den Vektoren der vorläufigen Werte die Vektoren der Revisionen der verschiedenen Datenjahrgänge addiert. Betrachtet werden hier die Revisionen der Werte für die Jahre 1993 bis 2010, so dass 18 Rechnungen durchgeführt werden können.⁹

Tabelle 8 zeigt die Ergebnisse dieses Experiments am Beispiel der Euklidischen Distanzen. Die Ergebnisse für die Handelsblatt-Methode und die City-Block-Metric sind im Anhang dargestellt. Die Rechnungen verdeutlichen, dass nicht nur die Abstandsmaße, sondern auch die Platzierungen der Prognosen erheblich schwanken können. Stets im oberen Bereich liegt die Prognose der Bundesbank, die man in allen Fällen nachträglicher Korrekturen auf dem ersten oder zweiten Platz findet. Die nach dem Handelsblatt-Ranking führende Berliner Sparkasse erreicht auch hier häufig den ersten Rang, findet sich in einem Fall aber auch auf dem 15. Rang wieder. Es gibt aber ein breites Mittelfeld von Prognosen, bei denen das Euklidische Abstandsmaß zwischen gut 1 und rund 2 schwankt, je nach nachträglicher Korrektur der Volkswirtschaftlichen Gesamtrechnungen, und die sich – je nachdem – auf Rängen zwischen 3 und 15 wiederfinden. Berechnet man z.B. um die Prognose des DIW, die mit einem Euklidischen Abstandsmaß von 1,37 in der Mitte der Rangfolge liegt, mit Hilfe der Standardabweichung der unter Verwendung der verschiedenen Revisionen generierten Abstandsmaße ein 68%-

8 Der Revisionszyklus der Daten der VGR erstreckt sich über etwa dreieinhalb Jahre. Die ersten Angaben für das Jahr t werden im ersten Quartal des Jahres $t+1$ veröffentlicht. Danach werden fortlaufend weitere im Laufe der Zeit bekanntwerdende Daten der amtlichen Statistik in das Rechenwerk eingearbeitet. Letztmalig werden Veränderungen im August des Jahres $t+3$ vorgenommen. Danach können sich die Angaben zwar auch noch ändern, dann aber aufgrund konzeptioneller Änderungen bei der Erstellung der VGR. Um die Revisionen aufgrund neuer Daten von konzeptionellen Änderungen zu trennen, werden hier die nach dreieinhalb Jahren vorliegenden Daten als endgültige Werte angesehen. Revisionen aufgrund neuer Daten und aufgrund konzeptioneller Änderungen vermischen sich allerdings in den Jahren, in denen eine Revision der VGR in den Dreieinhalb-Jahreszeitraum fällt.

9 Dieses Vorgehen setzt voraus, dass das Ausmaß der Revisionen nicht mit der Höhe der Veränderungsraten der betrachteten Variablen korreliert, was hier der Fall ist.

Intervall, so umfasst dieses den Bereich zwischen 1,18 und 1,56. In diesem Bereich liegen 12 der betrachteten 26 Prognosen. Eine signifikante Abstufung zwischen diesen Prognosen hinsichtlich ihrer Treffsicherheit lässt sich – gegeben diese Unschärfe – aufgrund der vorliegenden Daten kaum vornehmen. Eindeutiger ist das Bild wieder am Ende der Rangfolge: Keine der in den vergangenen Jahren registrierten Revisionen würde die Position der dort platzierten Prognosen nennenswert verbessern, und in das 68%-Intervall des Euklidischen Abstandsmaßes dieser Prognosen fällt auch keine andere der hier aufgeführten.

Tabelle 8

Maße der Treffsicherheit der betrachteten Konjunkturprognosen für 2014 unter Berücksichtigung nachträglicher Revisionen

Euklidische Distanz

Kürzel ¹	Niedrigster Wert		Höchster Wert		Standard- Abweichung	Mittelwert	
	DEU	Rang	DEU	Rang		DEU	Rang
BES	0,59	1	2,09	15	0,355	0,99	3,0
BUBA	0,74	1	1,62	2	0,213	0,98	1,4
BLB	1,19	5	2,16	18	0,230	1,38	9,5
FERI	1,02	3	1,88	9	0,193	1,22	4,8
HSBC	1,35	8	2,15	20	0,184	1,55	14,7
DB	1,05	3	2,28	21	0,296	1,34	8,7
MMW	1,01	3	1,95	10	0,207	1,28	6,9
IW	1,82	20	2,71	25	0,214	2,01	23,3
IMK	1,07	3	1,84	9	0,184	1,26	5,6
RWI	1,33	11	2,06	16	0,176	1,51	13,6
IWH	1,06	1	1,59	10	0,164	1,26	5,4
ING	1,30	6	2,74	26	0,345	1,68	17,4
DIW	1,19	7	1,98	12	0,188	1,40	10,2
IFO	1,09	3	1,71	12	0,172	1,31	6,8
ALL	1,18	3	1,68	16	0,144	1,39	9,8
IKB	1,18	4	1,69	15	0,146	1,40	10,4
BV	1,37	10	1,97	17	0,153	1,53	14,4
IFW	1,38	12	2,01	18	0,154	1,56	15,9
COM	1,35	13	2,09	19	0,182	1,59	16,3
HWWI	1,52	14	2,09	20	0,141	1,67	18,2
SVR	1,64	20	2,21	22	0,142	1,82	21,3
HLB	1,86	22	2,40	24	0,130	1,97	23,3
DEKA	1,52	8	1,93	21	0,138	1,68	18,1
GD	1,64	19	2,20	23	0,148	1,83	21,4
DZ	2,06	24	2,54	25	0,144	2,27	24,7
KE	2,35	23	2,87	26	0,156	2,60	25,8

Quelle: Handelsblatt, eigene Berechnungen. – ¹Zu den Kürzeln siehe Tabelle 1.

Wie wichtig es ist, diese Unschärfen in den Daten zu berücksichtigen, zeigen auch die Erfahrungen des Jahres 2014. Die hier als Vergleichsmaßstab herangezogenen, im Januar veröffentlichten ersten Jahresergebnisse wiesen eine Zunahme des BIP um 1,5% aus. Die im Februar publizierten ersten vollständigen Volkswirtschaftlichen Gesamtrechnungen hingegen beziffern des BIP-Wachstum auf 1,6%.

5. Schlussfolgerungen

Prognoserankings finden derzeit in Deutschland erhöhte Aufmerksamkeit. Der Wunsch ist durchaus nachvollziehbar, dass man diese nicht auf die Treffsicherheit reduzieren möchte, mit der eine einzelne volkswirtschaftlichen Kennziffer prognostiziert wurde, sondern dass man ein Gütemaß verwenden will, das mehrere Dimensionen der Prognose abdeckt. Allerdings bewegt man sich bei einer solchen mehrdimensionalen Prognoseevaluation auf einem bisher wenig erkundeten Terrain. Der vorliegende Beitrag versucht, dieses auszuleuchten. Die dabei gewonnenen Erkenntnisse lassen sich wie folgt zusammenfassen:

Erstens hängt die Beurteilung von Prognosen wesentlich von der gewählten Verlustfunktion ab. Ob man alle Fehler gleich gewichtet, oder größere Fehler mit höheren Gewichten versieht, ist entscheidend für die Rangfolge der Prognosen. Dies gilt allerdings insbesondere für das „Mittelfeld“ der betrachteten Prognosen, während sowohl die Plätze an der Spitze wie auch die am Ende – zumindest in den hier betrachteten Fällen – recht robust sind im Hinblick auf die Wahl der Verlustfunktion.

Zweitens sollte man bei der Auswahl der in eine solche mehrdimensionale Prognoseevaluation eingehenden Variablen beachten, dass viele Daten der Volkswirtschaftlichen Gesamtrechnungen wie auch deren Prognosen stark miteinander korreliert sind. So besteht in Deutschland ein deutlicher konjunktureller Gleichlauf des Bruttoinlandsprodukts sowohl mit den Investitionen als auch mit den Exporten und Importen. Nimmt man alle vier Variablen in eine Prognoseevaluation auf, läuft das implizit auf eine stärkere Gewichtung der BIP-Prognose hinaus. Um eine solche „Übergewichtung“ zu vermeiden, sollte man Variablen verwenden, die unterschiedliche Aspekte der Prognose abdecken. Im vorliegenden Fall sind dies die Wirtschaftsleistung – gemessen am Bruttoinlandsprodukt oder an dessen Verwendungskomponenten – und die Inflation. Die Arbeitslosigkeit wäre ein dritter solcher Aspekt, jedoch leistet die Arbeitslosenquote als im Konjunkturverlauf nur wenig veränderliche Bestandgröße hier keinen nennenswerten Beitrag zur Differenzierung der Gütemaße der einzelnen Prognosen. Interessant wäre, wenn man ferner eine verteilungsseitige Größe, etwa die für die Finanzplanung des Staates wichtigen Bruttolöhne und -gehälter, einbeziehen würde. Aber diese Größe dürfte nicht in allen betrachteten Prognosen ausgewiesen werden.

Drittens: Will man auch miteinander stark korrelierte Variablen in die Evaluation einbeziehen, so sollte man statt der Euklidischen Distanz die Mahalanobis Distanz verwenden, in deren Berechnung die Kovarianzen der Variablen einbezogen werden. Sie liefert von der Euklidischen Distanz, die Unabhängigkeit der Variablen unterstellt, deutlich abweichende Ergebnisse, ist aber unter Umständen nicht einfach zu interpretieren. Sie gewichtet nämlich Abweichungen, die von dem in der Vergangenheit beobachteten Muster des Zusammenhangs zwischen den Variablen abweichen stärker als Abweichungen, die diesem Muster entsprechen.

Viertens sollte man vorsichtig sein, aus Rangdifferenzen allzu weitreichende Schlüsse auf das Leistungsvermögen der einzelnen Prognostiker zu ziehen. Vielmehr sollte man stets die Abstandsmaße selbst im Auge behalten. Oft liegen diese nämlich sehr nahe beieinander. Dies gilt insbesondere, wenn man berücksichtigt, dass die Vergleichsmaßstäbe für deren Berechnung – die zuletzt beobachteten realisierten Werte – selbst Schätzungen sind, die später noch beträchtlich revidiert werden können. Zieht man frühere Datenrevisionen heran, um Unsicherheitsmargen für die Abstandsmaße zu ermitteln, so unterscheiden gerade die im „Mittelfeld“ rangierenden Prognosen oft nicht signifikant von einander.

Alles in allem zeigt die vorstehende Analyse, dass eine Bewertung von Prognosen anhand mehrerer Merkmale Bewertungsprobleme nach sich zieht, die bei der eindimensionalen Evaluation von Prognosen nicht auftreten. Die Distanzmaße können je nach Auswahl der einbezogenen Variablen und der Berechnungsart zu erheblichen Unterschieden führen. Beruhigend ist zwar, dass es zumindest mit den hier verwendeten Verfahren und in den hier betrachteten Jahren nicht gelungen ist, durch Veränderungen des Berechnungsverfahrens eine der weniger treffsicheren Prognosen zum Prognostiker des Jahres zu küren. Aber für die Medaillenvergabe können Stellgrößen wie Variablenauswahl und Distanzmaß durchaus entscheidend sein. Eine weitere Stellgröße, deren Folgen hier nicht vertieft wurden, wäre eine unterschiedliche Gewichtung der Fehler bei den einzelnen Variablen. Diese spielt bei der Handelsblatt-Methode am Rande eine Rolle, da die Fehler bei Ausfuhr und bei Einfuhr jeweils nur mit dem halben Gewicht in das Gütemaß eingehen, um ein Übergewicht des Außenhandels zu verhindern. Man könnte mit der gleichen Logik überlegen, Fehler bei der Prognose BIP als zentrale Kennziffer der Wirtschaftsleistungen höher zu gewichten.¹⁰ Insofern besteht reichlich Spielraum, Prognoserankings zu beeinflussen.

Bei alledem dürfen zwei Aspekte nicht übersehen werden: Erstens bewerten die hier vorgestellten Kennziffern Prognosen allein aus der Sicht ihrer Nutzer. Den Produzenten geben sie jedoch keine Hinweise, wie sie ihre Prognose verbessern können. Dies kann nur eine Evaluation der Prognose Variable für Variable leisten. Zweitens adressieren auch solche mehrdimensionale Verfahren nicht die Frage, ob die Prognosen konsistent aus ihrem jeweiligen Bedingungsrahmen abgeleitet wurden, also z.B. aus den Annahmen zu Wechselkursen oder Rohstoffpreisen. Denn bei Evaluationen darf nicht übersehen werden, dass jede Prognosen eine bedingte Aussage darstellt.

Literatur

Backhaus, K. et al. (2011), *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung*. 13. Auflage. Berlin: Springer.

Döhrn, R. (2014), *Konjunkturdiagnose und -prognose – eine anwendungsorientierte Einführung*. Berlin: Springer-Gabler.

Döhrn, R. und C. M. Schmidt (2011), *Information or Institution? – On the determinants of forecast accuracy*. *Jahrbücher für Nationalökonomie und Statistik*. 231: 9-27.

Fricke, T. (2014), *Noch mehr Prognostiker des Jahres? – Aufklärungsversuch*. <http://neuewirtschaftswunder.de/2014/03/26/noch-mehr-prognostiker-des-jahres-aufklarungsversuch/>. Zugegriffen am 12. Februar 2015.

Handelsblatt (2014), *Weidmann ist der beste Hellseher*. Handelsblatt vom 17.6.2014. <http://www.handelsblatt.com/politik/konjunktur/nachrichten/prognose-ranking-weidmann-ist-der-beste-hellseher/9584946.html>. Zugegriffen am 12. Februar 2015.

Handelsblatt (2015), *Der Sieger: Berliner Sparkasse*. Handelsblatt vom 16. Januar 2015. S. 10

Sinclair, T, H. Stekler und W. Carnow (2013). *A new approach for evaluating economic forecasts*. *Economic Bulletin*, 32 (3), S. 2331-2342.

SZ – Süddeutsche Zeitung (2014), *Prognostiker des Jahres 2014. Das Jahr der Pessimisten*. Süddeutsche Zeitung vom 19. Dezember 2014, S. 18.

¹⁰ Diese zentrale Rolle des BIP ist eine wesentliche Motivation des ftd/SZ-Verfahrens. Vgl. Fricke 2014.

Anhang

Tabelle A1

Konjunkturprognosen ausgewählter Institutionen für das Jahr 2013

Veränderungen gegenüber dem Vorjahr in %, soweit nicht anders vermerkt

Institution	Kürzel	BIP	Privater Konsum	Anlageinvestitionen	Export	Import	Verbraucherpreise	Arbeitslosenquote in % ¹
Bundesbank	BUBA	0,4	1,0	-0,1	1,9	3,0	1,5	7,2
MM Warburg	MMW	0,6	0,7	-0,8	4,2	1,6	1,5	7,0
Landesbank Berlin	LBB	0,5	1,0	0,7	2,2	2,9	1,6	6,8
Ifo	IFO	0,7	0,7	0,7	3,0	3,3	1,6	6,9
Industriekreditbank	IKB	0,8	0,9	-0,6	4,3	4,0	1,9	6,9
Deka-Bank	DEKA	0,7	1,1	-0,3	3,3	3,3	1,9	6,9
IMK	IMK	0,8	0,7	0,0	3,5	3,6	1,7	7,0
DZ Bank	DZ	0,4	0,9	0,1	3,0	3,8	2,1	7,1
HSBC	HSBC	0,6	1,0	0,5	2,9	4,1	2,0	7,0
Deutsche Bank	DB	0,3	0,6	1,1	3,2	4,2	1,7	7,0
HWWI	HWWI	0,5	1,0	0,7	3,5	4,4	1,9	6,6
RWI Essen	RWI	0,3	0,2	-1,1	3,0	3,2	1,7	7,0
Postbank	POST	0,6	1,0	0,4	3,7	4,3	2,0	7,1
Commerzbank	COM	0,5	1,3	0,1	2,8	4,1	1,9	7,1
IfW Kiel	IFW	0,3	0,6	0,6	2,9	3,9	2,0	7,0
ING	ING	0,8	0,4	1,2	0,6	0,4	2,0	6,9
DIW Berlin	DIW	0,9	1,1	0,9	4,2	4,6	1,8	7,0
Sachverständigenrat	SVR	0,8	0,8	1,4	3,8	4,2	2,0	6,9
Kiel Economics	KE	1,0	1,4	1,0	3,6	4,1	1,6	6,8
UBS	UBS	0,8	0,9	1,7	2,6	3,9	2,1	7,2
IWH Halle	IWH	0,7	0,4	0,3	4,0	3,7	2,0	6,7
IW	IW	0,75	0,5	1,5	4,0	4,0	1,75	6,5
Gemeinschaftsdiagnose	GD	,01	1,1	1,9	3,8	4,6	2,1	6,8
Feri	FERI	1,2	1,2	1,9	4,1	4,1	2,0	6,6
Helaba	HLB	1,1	1,2	2,6	5,5	5,0	2,0	7,0
nachrichtlich								
Ist ²		0,4	0,9	-0,7	0,8	0,9	1,5	6,9

Quelle: Handelsblatt (2014). – ¹ in % der gesamten Erwerbspersonen. – ² Erste Jahresergebnisse des Statistischen Bundesamtes.

Tabelle A2

Maße der Treffsicherheit der betrachteten Konjunkturprognosen für 2013

Kürzel ¹	Methode des Handelsblatts				City-Block-Metric		Euklidische Distanz		Mahalanobis-Distanz	
	Handelsblatt		Eigene Rechnungen		D _{CB}	Rang	D _{EU}	Rang	D _{MA}	Rang
	D	Rang	D	Rang						
BUBA	0,7	1	0,73	1	1,01	1	0,50	1	1,18	1
MMW	0,8	2	0,76	2	1,08	2	0,60	3	4,60	11
LBB	0,9	3	0,93	3	1,22	3	0,55	2	1,91	4
IFO	1,2	4	1,18	4	1,57	4	0,70	4	1,64	2
IKB	1,2	5	1,22	5	1,77	6	0,92	9	4,65	12
DEKA	1,3	6	1,33	6	1,74	5	0,79	6	2,82	7
IMK	1,3	7	1,33	7	1,79	7	0,77	5	2,35	6
DZ	1,4	8	1,42	8	1,85	8	0,95	11	5,82	20
HSBC	1,6	9	1,55	9	2,01	10	0,92	8	4,75	15
DB	1,6	10	1,57	10	2,05	12	0,91	7	6,84	23
HWWI	1,7	11	1,65	12	2,18	13	0,96	13	6,87	24
RWI	1,7	12	1,65	11	2,02	11	1,03	16	5,12	17
POST	1,7	13	1,68	13	2,22	16	1,00	15	5,32	18
COM	1,7	14	1,73	14	2,18	14	0,96	12	4,75	16
IFW	1,8	15	1,75	15	2,18	15	0,96	14	6,78	22
ING	1,8	16	1,82	16	1,88	9	0,94	10	1,97	5
DIW	1,8	17	1,83	17	2,43	19	1,04	17	3,41	9
SVR	1,9	18	1,85	18	2,38	17	1,07	19	4,71	14
KE	2,0	19	1,94	19	2,45	20	1,06	18	1,72	3
UBS	2,0	20	2,00	20	2,41	18	1,08	20	3,53	10
IWH	2,2	21	2,16	21	2,66	21	1,12	22	5,39	19
IW	2,2	22	2,20	22	2,73	22	1,09	21	6,03	21
GD	2,4	23	2,39	23	2,96	23	1,25	24	4,70	13
FERI	2,6	24	2,61	24	3,16	24	1,24	23	3,15	8
HLB	2,8	25	2,76	25	3,50	25	1,46	25	9,32	25

Quelle: Handelsblatt, eigene Berechnungen. – ¹Zu den Kürzeln siehe Tabelle A1.

Tabelle A3

Maße der Treffsicherheit der betrachteten Konjunkturprognosen für 2014 unter Berücksichtigung nachträglicher Revisionen

Handelsblatt-Methode

Kürzel ¹	Niedrigster Wert		Höchster Wert		Standard- Abweichung	Mittelwert	
	DEU	Rang	DEU	Rang		DEU	Rang
BES	1,02	1	3,02	7	0,741	1,84	3,1
BUBA	1,18	1	2,64	5	0,743	1,84	1,9
BLB	1,76	2	3,62	20	0,743	2,34	9,3
FERI	1,51	3	3,21	10	0,745	2,16	5,7
HSBC	1,88	3	3,86	23	0,746	2,57	13,3
DB	1,44	2	3,68	21	0,747	2,32	10,4
MMW	1,43	2	3,24	17	0,748	2,30	8,6
IW	2,28	8	4,89	26	0,748	3,19	20,0
IMK	1,78	2	3,40	13	0,749	2,30	7,4
RWI	1,77	3	3,35	15	0,750	2,49	10,8
IWH	1,76	2	3,06	15	0,751	2,30	7,3
ING	2,27	5	4,28	25	0,751	3,04	19,6
DIW	1,78	4	3,41	16	0,752	2,51	11,0
IFO	1,82	3	3,09	14	0,753	2,37	8,3
ALL	1,76	3	3,11	20	0,754	2,43	10,1
IKB	1,75	2	3,19	19	0,754	2,46	10,8
BV	1,94	8	3,42	17	0,755	2,59	13,1
IFW	2,17	11	3,50	19	0,755	2,71	16,0
COM	1,99	8	3,76	21	0,755	2,81	17,2
HWWI	2,24	15	3,54	20	0,754	2,81	18,0
SVR	2,14	14	3,71	23	0,754	2,91	19,5
HLB	2,22	17	3,93	24	0,753	3,07	21,6
DEKA	2,19	8	3,54	23	0,752	2,80	17,3
GD	2,14	13	3,86	24	0,751	2,98	20,5
DZ	2,92	23	4,83	26	0,751	3,83	24,5
KE	3,35	23	5,61	26	0,749	4,60	25,7

Quelle: Handelsblatt, eigene Berechnungen. – ¹Zu den Kürzeln siehe Tabelle 1.

Tabelle A4

Maße der Treffsicherheit der betrachteten Konjunkturprognosen für 2014 unter Berücksichtigung nachträglicher Revisionen

City-Block-Metric

Kürzel ¹	Niedrigster Wert		Höchster Wert		Standard- Abweichung	Mittelwert	
	DEU	Rang	DEU	Rang		DEU	Rang
BES	1,18	1	3,17	7	0,559	1,99	3,2
BUBA	1,34	1	2,77	6	0,427	2,00	2,3
BLB	1,85	2	3,72	19	0,528	2,50	9,1
FERI	1,66	1	3,25	10	0,430	2,28	5,3
HSBC	2,02	2	4,16	23	0,541	2,83	13,7
DB	1,64	2	3,87	21	0,626	2,56	10,7
MMW	1,55	2	3,47	19	0,491	2,41	8,3
IW	2,48	7	5,11	26	0,700	3,38	19,1
IMK	1,83	1	3,57	13	0,476	2,50	7,6
RWI	1,83	5	3,70	15	0,474	2,67	10,4
IWH	1,88	2	3,27	14	0,408	2,45	7,6
ING	2,46	4	4,51	24	0,612	3,23	18,7
DIW	1,81	2	3,85	16	0,505	2,73	11,3
IFO	2,03	2	3,60	16	0,476	2,67	10,2
ALL	1,85	1	3,29	20	0,419	2,55	9,1
IKB	1,85	2	3,44	18	0,431	2,62	10,3
BV	2,14	8	3,86	17	0,479	2,81	13,1
IFW	2,32	11	4,01	20	0,469	3,04	17,7
COM	2,08	8	4,27	22	0,509	3,09	18,1
HWWI	2,29	12	3,93	20	0,466	3,02	16,9
SVR	2,18	12	4,16	23	0,490	3,14	19,4
HLB	2,60	17	4,42	24	0,469	3,33	22,1
DEKA	2,29	8	3,81	22	0,440	2,98	16,3
GD	2,15	11	4,28	24	0,512	3,20	20,3
DZ	3,33	23	5,60	26	0,597	4,37	24,9
KE	3,54	22	5,82	26	0,652	4,79	25,4

Quelle: Handelsblatt, eigene Berechnungen. – ¹Zu den Kürzeln siehe Tabelle 1.

IBES



ISSN-Nr. 2192-5208 (Print)
ISSN-Nr. 2192-5216 (Online)

