

STUDY

Das IMK ist ein Institut
der Hans-Böckler-Stiftung

Nr. 54 • Januar 2017 • Hans-Böckler-Stiftung

VERGLEICHENDE EVALUATION DER KONJUNKTUR- PROGNOSEN DES IMK AN DER HANS-BÖCKLER- STIFTUNG FÜR DEN ZEITRAUM 2005-2014

Ulrich Fritsche¹, Artur Tarassow

Universität Hamburg
Korrigierte Version vom 17. Januar 2017

Kurzbeschreibung

Es werden die makroökonomischen Prognosen der acht großen deutschen Forschungsinstitute evaluiert. Die Datenbasis umfasst die publizierten Prognosen für 11 Aggregate für den Zeitraum 2005 bis 2014. Zunächst wird jede Variable separat mittels aktueller Evaluationsmethoden untersucht. Diese umfassen: Theil'sches U, Diebold-Mariano-Test, regressionsbasierte sowie nicht-parametrische Tests auf Rationalität, Bestimmung einer asymmetrischen Prognoseverlustfunktion und den dazugehörigen verallgemeinerten Test auf Rationalität, Richtungsänderungstest. Anschließend wird ein Gesamtsieger mittels des multivariaten Mahalanobis-Distanzmaßes ermittelt. Dieser Ansatz untersucht mehrere Variablen gemeinsam auf ihre empirische Kohärenz. Die quantitative Prognosegüte ist sowohl über die Institute wie über die Variablen hinweg heterogen. Für die meisten berücksichtigten Aggregate – insbesondere für Wirtschaftswachstum (BIP) und Inflationsrate – und über die meisten Institute hinweg betrachtet ist auch eine Prognoseverbesserung gegenüber einer naiven Prognose darstellbar. Allerdings ist ein gewisses Problem mit der Prognosegüte von privaten Konsumausgaben sowie der Arbeitslosenquote sichtbar. Es gibt kaum Hinweise auf Verletzung des Rationalitätskriteriums für das BIP, die Inflationsrate, die Ausrüstungsinvestitionen, die Ein- und Ausfuhren sowie die Arbeitslosenquote. Schwierigkeiten bei den Wendepunktprognosen gibt es für die Inflationsrate, die Konsumausgaben des Staates und der privaten Haushalte, die Bauinvestitionen sowie die Arbeitslosenquote. Im Hinblick auf die Form der Verlustfunktion kann für die Wachstumsrate des BIP oder die Inflationsrate keine Asymmetrie diagnostiziert werden. Bei der multivariaten Prognosegüte ergeben sich überraschend deutliche Unterschiede bezüglich der kohärenten „Stories“ zwischen den Instituten. Das IMK schneidet hier mit dem ifo Institut sowie dem IWH zusammen am besten ab. Alle Daten und Computerroutinen sind öffentlich verfügbar.

Die verwendeten Daten und Computerroutinen sind abrufbar unter (Stand Januar 2017):

http://www.boeckler.de/pdf/p_imk_study_54_2017_replikation.zip

¹ Verantwortlicher Autor: Ulrich Fritsche, Universität Hamburg, KOF ETH Zürich, GWU Resarch Program on Forecasting, Email: ulrich.fritsche@wiso.uni-hamburg.de.

Vergleichende Evaluation der Konjunkturprognosen des Instituts für Makroökonomie und Konjunkturforschung an der Hans-Böckler-Stiftung für den Zeitraum 2005-2014

Ulrich Fritsche*
Artur Tarassow

Universität Hamburg
1. Version: 20. Dezember 2016

Korrigierte Version vom 17. Januar 2017

1 Einführung

Die Diskussion über die Qualität und Bedeutung von Prognosen sozialer und ökonomischer Ereignisse ist ein klassischer Untersuchungsgegenstand theoretischer wie empirischer Wirtschaftsforschung ([Friedman, 1953](#); [Löwe, 1929](#); [Morgenstern, 1928](#)). Die Tatsache, dass Journalisten und verschiedene Kommentatoren ausführlich die Prognosefehler der Ökonomenzunft nach der Finanzkrise 2008 debattierten (siehe [Nienhaus, 2009](#), und die dort zitierten Belege), zeigt das breite und zunehmende Interesse der Öffentlichkeit an den Prognosen der Konjunkturforscher und ihren “Irrtümern”. Dies betrifft neben der journalistischen Berichterstattung weite Teile der wirtschaftspolitisch interessierten Öffentlichkeit und der Politik.

Angesichts der Tatsache, dass die Finanzkrise und die nachfolgende tiefe, weltweite Rezession — inzwischen gemeinhin als “Great Recession” bezeichnet — nicht nur zu kurzfristigen Finanzmarkturbulenzen, sondern zu einem tiefen, langanhaltenden Einbruch der Wirtschaftstätigkeit mit hoher Arbeitslosigkeit in vielen Ländern geführt hat, erstaunt dies nicht.

*Verantwortlicher Autor: Ulrich Fritsche, Universität Hamburg, KOF ETH Zürich, GWU Research Program on Forecasting, Email: ulrich.fritsche@wiso.uni-hamburg.de.

Danksagung: Wir danken dem IMK an der Hans-Böckler-Stiftung für die finanzielle Unterstützung und insbesondere Gustav Horn, Sven Schreiber, Peter Hohlfeld und Thomas Theobald für Kommentare und Verbesserungsvorschläge. Wir danken Roland Döhrn, Oliver Holtemöller und Katja Heinisch für Hinweise, die gegenüber der ersten Fassung zur Korrektur eines Fehlers im Programmcode der Studie führten und zur Korrektur zweier Datenpunkte in der Datenbank sowie einigen kleineren Änderungen in der Beschreibung der Methodik. Ein zu dieser Studie verfügbares Erratum auf der Website des IMK erläutert den Fehler im Programmcode der ersten Version und die umgesetzten Änderungen in der vorliegenden korrigierten Version. Wir danken ausdrücklich Patrick Nüß für die Forschungsunterstützung und die unermüdliche Arbeit beim Aufbau der Datenbank.

Die “Great Recession” bleibt dabei nicht ohne Auswirkungen auf die Zunft der Ökonomen und Prognostiker. Zum einen steigt der Erklärungsdruck für Fehlprognosen aus Bereichen der Öffentlichkeit und des Journalismus.¹ So kritisiert die Journalistin Lisa Nienhaus in ihrem — in polemischem Stil, aber durchaus nicht ohne akribische Recherche geschriebenen — Buch “Die Blindgänger” die systematische Ausblendung der Möglichkeit von Finanzkrisen in Standard-Prognosemodellen (Nienhaus, 2009). Aber auch Wissenschaftler aus dem Mainstream ökonomischen Denkens sprechen von einer “Schuld der Ökonomen” (Riedel, 2013). Zum anderen wird die Frage, welche Schlussfolgerungen aus den Fehlprognosen der vergangenen Jahre zu ziehen sind, innerhalb der ökonomischen Profession selbst sehr unterschiedlich beurteilt (Bezemer, 2011; Gaffney, 2011). Das heterodoxe Journal “Real World Economics” lobte einen Preis für diejenigen aus, die die Finanzkrise prognostizierten und einen ironischen “Dynamit-Preis” für diejenigen, die die theoretischen Modelle entwickelten, welche nach Meinung vieler, nicht nur heterodoxer Ökonomen zur Krise maßgeblich beitrugen. Aber auch Mainstream-Ökonomen und internationale Institutionen wie der Internationale Währungsfonds und die OECD haben eine Diskussion darüber begonnen, welche Krisenursache in der retrospektiven Perspektive unerkannt blieben und inwieweit die makroökonomischen Modelle angepasst werden müssen (Stiglitz, 2011, 2014). Die Finanzkrise führte in diesem Zusammenhang zu einer Zunahme von Publikationen, die Erweiterungen der dominanten DSGE-Modelle v.a. im Bereich der Modellierung von Unvollständigkeits auf den Finanzmärkten vorstellen (Brunnermeier u. a., 2012; De Paoli u. a., 2014) oder die Mikrofundierung bestehender Modelle kritisch diskutieren (Stiglitz, 2014).

Ein zentrales Kriterium für die Beurteilung von Theorien ist deren Prognosefähigkeit (Friedman, 1953; Morgenstern, 1928). Eine regelhafte und systematische Analyse der makroökonomischen Prognosen ist daher unumgänglich. Evaluationsstudien von Konjunkturprognosen in Deutschland sind in der Vergangenheit immer wieder mit unterschiedlichen Graden des Interesses in der wirtschaftspolitisch interessierten Öffentlichkeit diskutiert worden (siehe Döpke und Fritsche, 2006a, und insbesondere die dort zitierten Studien). In den vergangenen Jahren stieg das Interesse an dieser Art von Evaluationen wieder an. Wichtige internationale Organisationen haben ihre Prognosen evaluiert oder extern evaluieren lassen (IMF, 2014; Pain u. a., 2014). Aber auch Forschungsinstitute in Deutschland sind durch ihre Auftraggeber aufgefordert, regelmäßig ihre Prognosen z.B. im Rahmen der Gemeinschaftsdiagnose zu evaluieren oder tun dies selbst im Rahmen von Forschungsarbeiten (Döhrn und Schmidt, 2011; Henzel u. a., 2013; Nierhaus, 2014; Wollmershäuser, 2015). Allerdings beziehen sich diese Evaluationen oft entweder nur auf einzelne Aspekte, kurze Zeiträume oder einzelne Institutionen. Interessante jährliche Prognoserankings legten in den vergangenen Jahren auch Printmedien wie die *Financial Times Deutschland*, die *Süddeutsche Zeitung* oder das *Handelsblatt* vor. Allerdings sind die Daten,

¹Diese Kritik erstaunt langjährige Beobachter der Prognosedebatten nicht. Prognostikerschelte ist in Krisenzeiten immer wohlfeil. Einer der Autoren der Studie, Ulrich Fritsche, der mit seinem Ko-Autor Jörg Döpke die Qualität von Konjunkturprognosen in Deutschland ab 1970 in einer größeren Studie schon 2006 untersuchte, erinnert sich noch gut an die damalige “Prognostikerschelte”, die von Artikeln des Spiegel über polternde Kommentare Gerhard Schröders bis hin zur öffentliche Schelte von Prognostikern reichte (siehe Döpke und Fritsche, 2009; Fritsche und Döpke, 2006, und die dort zitierten Quellen). Die Bemerkungen von Bundeskanzlerin Angela Merkel in der jüngsten Debatte zur Qualität von Konjunkturprognosen schließen da nahtlos an die ihres Vorgängers im Amt an und sind angesichts der großen Prognosefehler schon fast höflich formuliert.

auf denen diese Evaluationen beruhen (v.a. für private Institutionen wie Banken oder Versicherungen), oftmals nicht verfügbar oder die Methodik dieser Evaluationen ist nicht immer nachvollziehbar. Eine Ausnahme bildet Müller-Dröge u. a. (2014), deren Methodik auch hier eingesetzt wird.

Das Ziel des vorliegenden Evaluationsprojektes kann wie folgt beschrieben werden. Zum Ersten soll die Literatur über angemessene Evaluationsmethoden kritisch reflektiert und die im vergangenen Jahrzehnt stattgefundenen methodologischen Verbesserungen soweit aufgenommen werden, dass es sich um eine Evaluation nach dem aktuellen “state of the art” handelt. Zweitens soll ein möglichst breiter Datensatz mit vielen prognostizierten Variablen betrachtet werden, und nicht wie in vielen Studien nur Wachstumsprognosen des BIP. Drittens soll ein Vektor von Prognosen über verschiedene Variablen analysiert werden, da dies der Idee einer Konjunkturprognose als “konsistenter Geschichte” (siehe auch Tietzel, 1989, für entsprechende Argumente) entspricht. Viertens hat diese Studie die relative Güte der IMK-Prognosen im Vergleich zu den wichtigsten Mitbewerbern im Auge, was die Auswahl der in die Evaluation aufzunehmenden Institutionen beeinflusst hat.² Überdies wird ein besonderer Augenmerk auf die Veränderung der Prognosegüte über den Zeitraum der Finanzkrise 2007/2008 und danach gelegt.

Die Studie ist folgendermaßen aufgebaut: Im Teil 2 wird die verwendete Methodik erläutert. Danach werden im Teil 3 die verwendeten Daten vorgestellt. Im Teil 4 werden die zentralen Ergebnisse detailliert für Wachstums- und Inflationsprognosen und zusammenfassend für alle anderen betrachteten Variablen erläutert. Im Teil 5 werden die multivariaten Evaluationsergebnisse dargestellt. Teil 6 enthält Diskussion und Schlussfolgerungen.

2 Verwendete Methoden

2.1 Prognosegüte

Um einen Überblick über die Prognosequalität zu bekommen, wurde in einem ersten Schritt die Prognosegüte beurteilt. Dazu wurde das *Verhältnis des mittleren quadrierten Fehlers zum mittleren quadrierten Fehlers einer Referenzprognose* (relative mean squared error) berechnet. In der Literatur zu Gütemaßen von Prognosen wird diese Maßzahl oftmals auch als das Theilsche U bezeichnet, wobei in der ursprünglichen Definition dieses Maßes die Referenzprognose in der Annahme keiner Veränderung bestand.

$$U = \frac{1}{T} \sum_{t=1}^{T-1} (y_{t+1} - \hat{y}_{t+1|t})^2 \times \left[\frac{1}{T} \sum_{t=1}^{T-1} (y_{t+1} - y_t)^2 \right]^{-1}$$

Im Bereich der Prognosegütebeurteilung verstecken sich jedoch einige Herausforderungen:

- Die naive Prognose bzw. die Annahme keiner Änderung gegenüber der Vor-

²In einem weitergehenden, von der Deutschen Forschungsgemeinschaft im Rahmen des Schwerpunktprogramms 1859 “Erfahrung und Erwartung: Historische Grundlagen wirtschaftlichen Handelns” geförderten Forschungsprojekt werden die Anzahl der Institutionen und der erfassten Prognosen auf der Zeitachse in die Vergangenheit deutlich erweitert.

periode ist ein geeigneter Ausgangspunkt für die Berechnung des U -Wertes nach Theil (1966). Damit können als gemeinsamer Nenner auch die Prognosen verschiedener Institutionen oder gleicher Institutionen zu verschiedenen Prognosehorizonten relativ zueinander verglichen werden. Allerdings ist die Annahme der Nichtänderung nur für den Spezialfall eines *Random Walk* auch die optimale Prognose. Unter der Annahme von Lernprozessen, der Kenntnis von Anpassungsdynamiken und mglw. autokorrelierten Schocks und unter Absehen vom Spezialfall des *Random Walk* ist eine derartige Referenzprognose kein “fairer” Benchmark. Realistischere Vergleichsprognosen können z.B. aus rekursiv oder mit rollendem Fenster geschätzten (V)ARIMA-Modellen gewonnen werden oder aus anderen — z.B. nichtlinearen — Zeitreihenmodellen, wobei die optimale Struktur und die Länge der Verzögerungen über Informationskriterien bestimmt werden können.

- Der einfache Vergleich von Prognosegütemaßen (wie RMSE) beantwortet nicht die Frage, ob der Abstand zwischen Prognosemaßen im Sinne einer Signifikanzbewertung hinreichend groß ist, um die Prognosegüte zweier Institutionen als auf Basis der Stichproben unterschiedlich bewerten zu können. In der Regel werden in der Literatur bilaterale Prognosevergleiche über die Differenzen der RMSE-Werte mit dem Verfahren von Diebold und Mariano (1995) durchgeführt. Zusammen mit der Information aus dem Theilschen U (Prognoseverbesserung bei $U < 1$, Prognoseverschlechterung bei $U > 1$) kann mit dem Diebold-Mariano-Test nicht nur die Signifikanz der Abweichung des Theilschen U vom Wert 1 beurteilt, sondern auch eine signifikante Prognoseverbesserung diagnostiziert werden. Konkret: Mit $d_{12} = L(e_1) - L(e_2)$ als Vektor von t Verlustdifferentialen zwischen zwei Prognosen, kann unter der Null gleicher Prognosequalität (mit dem Erwartungswert $E(d_{12}) = 0$) mit einer asymptotisch z -verteilten Statistik getestet werden. D.h. es gilt:

$$DM = \frac{\bar{d}_{12}}{\hat{\sigma}_{12}} \sim N(0, 1)$$

wobei \bar{d}_{12} den Stichprobendurchschnitt der Verlustdifferentialie bezeichnet und $\hat{\sigma}_{12}$ einen konsistenten Schätzer der Standardabweichung für \bar{d}_{12} darstellt. Da jedoch die Prognosefehler autokorreliert sein könnten, empfiehlt es sich, das Verlustdifferential d_{12} auf eine Konstante mittels OLS zu regressieren und robuste HAC Standardfehler anzuwenden

$$d_{12t} = \beta_0 + u_t \tag{1}$$

wobei u_t gleich *i.i.d.*-verteilte Störterme sind. Unter der Nullhypothese gleicher Prognosequalität gilt das $\beta_0 = 0$. Die t -Statistik des β_0 -Koeffizienten ist zugleich die Teststatistik. Harvey u. a. (1997) schlagen ergänzend eine Korrektur für kleine Stichproben vor, die wir verwenden.

- Prognoseevaluationen auf Basis der o.g. Methoden vergleichen die durchschnittliche Prognosegüte. Wie Dovern und Janssen (2015); Fildes und Stekler (2002); Stekler (1972) gezeigt haben, passieren die entscheidenden Prognosefehler beim “Verpassen” der konjunkturellen Wendepunkte. Um die Güte von Wendepunktvorhersagen zu testen (z.B. anhand der Akzeleration und Dezele-

ration von Wachstumsraten des Bruttoinlandsproduktes), haben [Diebold und Lopez \(1996\)](#) einen Test entwickelt, um den Informationsgehalt von Prognosen bezüglich der Wendepunkte zu beurteilen. Dazu wird auf eine einfache Kontingenztabelle zurückgegriffen, bei der die beiden Zustände “auf” und “ab” mit “i” und “j” als Subskripte bezeichnet werden und die klassifizierten Prognosen in den Zeilen und die Realisationen in den Spalten abgetragen werden.

| | Realisation i | Realisation j | Zeilensumme |
|--------------|---------------|---------------|-------------|
| Prognose i | O_{ii} | O_{ij} | $O_{i.}$ |
| Prognose j | O_{ji} | O_{jj} | $O_{.j}$ |
| Spaltensumme | $O_{.i}$ | $O_{.j}$ | O |

Das Subskript “ii” in den Zellen steht dann für ein prognostiziertes Event (hier Wachstumsbeschleunigung, also “auf”), das auch stattfand, bzw. “ij” für eine prognostizierte Wachstumsbeschleunigung, die nicht stattfand. Analoges gilt für “jj” und “ji” in Bezug auf Wachstumsverringerungen. Der Informationsgehalt einer Sequenz von Prognosen kann über das Maß $I = \frac{O_{ii}}{O_{ii}+O_{ji}} + \frac{O_{jj}}{O_{jj}+O_{ij}}$ zusammengefasst werden. In einem reinen Zufallsprozess, in der ein Münzwurf den Prognostiker leitet, wären $O_{ii} \approx O_{ji}$ sowie $O_{jj} \approx O_{ij}$, und damit $I \rightarrow 1$. Bei einer perfekten Prognose ergibt sich: $O_{ji} = O_{ij} = 0$ und $I \rightarrow 2$. Das bedeutet, dass jeder Wert $1 < I \leq 2$ für sich genommen, einen positiven Informationsgehalt (Prognoseverbesserung gegenüber dem Referenz Münzwurf) indiziert. Der konsistente Schätzer für die erwartete Zellbelegung unter der Null keiner Informationsverbesserung ergibt sich als $\hat{E}_{ij} = O_{i.}O_{.j}/O$. Als Teststatistik wird dann üblicherweise das Maß $C = \sum_{i,j=1}^2 (O_{ij} - \hat{E}_{ij})^2 / \hat{E}_{ij} \sim \chi_1^2$ verwendet, dass unter der Null einer χ^2 -Verteilung mit einem Freiheitsgrad folgt ([Döpke und Fritsche, 2006a](#)). In der vorliegenden Untersuchung wird C für verschiedene Institutionen mit verschiedenen Prognosehorizonten und für unterschiedliche Variablen berechnet, um die Güte der Wendepunktprognosen zu bestimmen.

2.2 Rationalität und Effizienz

Jede Prognoseevaluation muss sich mit den Aspekten “Rationalität” und “Effizienz” auseinandersetzen. Sie bilden gewissermaßen den Lackmustest, an dem die “Wissenschaftlichkeit” von Prognosen — in Abgrenzung zur Kaffeesatzleserei oder zu “Prophezeiungen” — festgemacht wird. [Mincer und Zarnowitz \(1969\)](#) haben in ihrer zentralen Studie mit Verweis auf [Theil \(1966, S. 10ff\)](#) argumentiert: “An economic forecast may be called ‘scientific’ if it is formulated as a verifiable prediction by means of an explicitly stated method which can be reproduced and checked.” ([Mincer und Zarnowitz, 1969, S. 3](#)). Nicht modellbasierte Prognosen — und dabei handelt es sich in vielen Fällen ([European Central Bank, 2014](#)) — könnten mit einem solch strikten Anspruch jedoch nicht falsifiziert werden. Allerdings können Kriterien aufgestellt werden, nach denen eine Prognose als rational oder effizient beurteilt werden kann.

- Prognosen sollten im Erwartungswert unverzerrt sein, also keine systematische Verzerrung aufweisen. Um dies zu testen, schlagen [Mincer und Zarnowitz](#)

(1969) vor, die Gleichung

$$y_{t+1} = \beta_0 + \beta_1 \hat{y}_{t+1|t} + u_{t+1}$$

zu schätzen und die gemeinsame Hypothese

$$H_0 : \begin{cases} \beta_0 = 0 \\ \beta_1 = 1 \end{cases} \quad (2)$$

zu testen.

- Für den Effizienztest beziehen wir uns auf den Schätzansatz in [Kirchgässner \(1984\)](#), wobei in dem Aufsatz argumentiert wird, dass für eine optimale Prognose der Prognosefehler mit jeder ex ante bekannten Information unkorreliert sein sollte. Das impliziert als Spezialfall Orthogonalität zu den verzögerten Prognosefehlern. Das dahinterstehende Konzept wird i.d.R. als “schwache Effizienz” bezeichnet.³ Als gemeinsamen Test auf Unverzerrtheit und Effizienz schätzen wir daher

$$y_{t+1} = \beta_0 + \beta_1 \hat{y}_{t+1|t} + \beta_2 e_{t|t-1} + u_{t+1}$$

und testen die verbundene Hypothese:

$$H_0 : \begin{cases} \beta_0 = 0 \\ \beta_1 = 1 \\ \beta_2 = 0 \end{cases} \quad (3)$$

mit einem F-Test.

- In der Literatur findet sich eine noch strengere Fassung des Rationalitätskonzeptes, die gelegentlich als “starke (Form der) Rationalität” bezeichnet wird. In dieser Interpretation wird nicht nur mit korrektem Erwartungswert und unter Berücksichtigung der in der Vergangenheit gemachten eigenen Fehler prognostiziert, sondern es werden alle öffentlich verfügbaren Informationen in die Prognose vollständig und korrekt aufgenommen. Wenn im Umkehrschluss öffentlich verfügbare Informationen existieren — z.B. über Zinssätze, Ölpreise, etc. — die in systematischer Weise die künftigen Prognosefehler erklären, so kann eine solche Prognose nicht rational im strengen Sinne sein. Den Arbeiten von [Holden und Peel \(1990\)](#); [Kirchgässner und Savioz \(2001\)](#) und [Döpke und Fritsche \(2006a\)](#) folgend, wird die Gleichung

$$y_{t+1} = \beta_0 + \beta_1 \hat{y}_{t+1|t} + \beta_2 X_t + u_{t+1}$$

als gemeinsamer Test auf Rationalität und Effizienz geschätzt, wobei X_t ein Vektor öffentlich verfügbarer Informationen zum Zeitpunkt der Prognoseerstellung ist. Die Nullhypothese lautet

³[Nordhaus \(1987\)](#) argumentiert überdies, dass unter der Annahme schwacher Effizienz der Prognosefehler im Zeitpunkt t unkorreliert mit allen Prognoserevisionen bis zum Zeitpunkt t sein sollte und dass die Prognoserevision im Zeitpunkt t ebenfalls mit allen vergangenen Prognoserevisionen sein sollte.

$$H_0 : \begin{cases} \beta_0 = 0 \\ \beta_1 = 1 \\ \beta_2 = 0 \end{cases} \quad (4)$$

und wird über einen F-Test geprüft. [Döpke und Fritsche \(2006a\)](#) verwenden als Kontrollvariable folgende Variablen: kurzfristiger Zinssatz, realer effektiver Wechselkurs, Jahresveränderungsraten des Ölpreises und der US-amerikanischen Industrieproduktion. In dieser Studie wird der reale kurzfristige 3-Monatszins verwendet.

- Die regressionsbasierten Ansätze können durch nicht-parametrische Ansätze ergänzt werden. Zum Testen auf das Vorliegen von Unverzerrtheit schlagen [Dufour \(1981\)](#) und [Campbell und Ghysels \(1995, 1997\)](#) folgende Teststrategie vor: Man ordnet die absoluten Prognosefehler der Größe nach und vergibt die Ränge. Unter der Nullhypothese, dass die Prognosefehler unabhängig voneinander sind mit dem Median von Null und der zusätzlichen Annahme, dass die Prognosefehler symmetrisch um die Null verteilt sind, folgt die Teststatistik der Verteilung des bekannten Wilcoxon-Vorzeichen-Rang-Tests ([Wilcoxon u. a., 1970](#)). Für weitere Details siehe [Campbell und Ghysels \(1995, S. 556\)](#). Die Nullhypothese kann implizit formuliert werden als:

$$H_0^{CG1} : \text{median}(e_{t|t-1}) = 0 \quad (5)$$

Den Argumenten in [Campbell und Ghysels \(1995, 1997\)](#) folgend, hat dieser Test eine gute *Power* in kleinen Stichproben. Als Test auf Effizienz wird ebenfalls auf den Vorschlag in [Campbell und Ghysels \(1995, 1997\)](#) zurückgegriffen. Hierbei wird auf das Fehlen von Autokorrelation (erster Ordnung) in den Prognosefehlern getestet, indem das Produkt zweier aufeinanderfolgender Prognosefehler daraufhin untersucht wird, ob es einen Median von Null aufweist, d.h.:

$$H_0^{CG2} : \text{median}(e_{t|t-1} \times e_{t-1|t-2}) = 0 . \quad (6)$$

- [Batchelor und Peel \(1998\)](#); [Elliott u. a. \(2005\)](#); [Patton und Timmermann \(2007\)](#) diskutieren die traditionellen Unverzerrtheits- und Effizienztests äußerst kritisch und argumentieren, dass beobachtbare Prognosefehler durchaus im Einklang mit der Hypothese rationaler Erwartungen stehen können, auch wenn die üblichen Tests eine Verletzung dieser Hypothese nahelegen. Der Grund liegt in der impliziten Annahme einer symmetrischen Verlustfunktion, die den meisten Ansätzen zugrunde liegt.⁴ Um auf das Vorliegen möglicher Asymmetrien in der Verlustfunktion zu testen, ist der Ansatz von [Elliott u. a. \(2005\)](#) ein guter Ausgangspunkt. Darin wird folgende verallgemeinerte Verlustfunktion vorgeschlagen, die den Symmetriefall als Spezialfall enthält:

$$L(p, \alpha, \Theta) = [\alpha + (1 - 2\alpha) \times \mathbf{I}(y_{t+1} - \hat{y}_{t+1} < 0) \times |y_{t+1} - \hat{y}_{t+1}|^p] . \quad (7)$$

\mathbf{I} bezeichnet dabei eine Indikatorfunktion. In dieser Verlustfunktion bildet der

⁴Ein entsprechendes Argument war schon von [Granger \(1969\)](#) formuliert worden.

Parameter p die (gesetzte) Annahme über die (zu schätzende) Form der Verlustfunktion ab. Konkret wird unter der Setzung von $p = 1$ eine linear approximierte Verlustfunktion (sog. *lin-lin*) und unter $p = 2$ eine quadratisch approximierte Verlustfunktion (sog. *quad-quad*) geschätzt. Darüber hinaus enthält die Verlustfunktion den Parameter α , der frei geschätzt werden kann. Werte von $\alpha = 0.5$ bilden den Symmetriefall ab, Werte von $\alpha > 0.5$ zeigen überoptimistische Prognosen an, und $\alpha < 0.5$ indiziert überpessimistische Prognosen. [Elliott u. a. \(2005\)](#) zeigen, dass sich in diesem Ansatz Kriterien für Prognoseoptimalität ergeben, die sich als Momentenbedingungen eines Instrumentenvariablen-schätzers (hier: GMM, also Generalisierte Momentenmethode) darstellen lassen.⁵ Der Test auf Rationalität beruht dann auf einem Sargan-Hansen J-Test auf überidentifizierende Restriktionen in einem GMM-Ansatz ([Hansen, 1982](#); [Sargan, 1958](#)).

2.3 Mahalanobis-Distanz als multivariates Evaluationswerkzeug

Die Mahalanobis-Distanz (MD) ist ein statistisches Werkzeug der multivariaten Analyse. Dieses Maß misst die Distanz zwischen den Prognosen eines Instituts und den realisierten Werten, wobei der zu untersuchende Vektor mehrere Variablen umfassen kann. Die MD wird in der folgenden quadratischen Form ausgedrückt

$$D^2 = (\mathbf{F} - \mathbf{A})' \boldsymbol{\Sigma}^{-1} (\mathbf{F} - \mathbf{A}) \quad (8)$$

wobei \mathbf{F} der Vektor der durchschnittlichen Prognosen, \mathbf{A} ein Vektor der dazugehörigen durchschnittlichen Realisationen und $\boldsymbol{\Sigma}$ die positiv-definite Kovarianzmatrix der historischen Realisationen sind. Anlehnend an [Müller-Dröge u. a. \(2014\)](#) legen wir für die Kovarianzmatrix jeweils die letzten 10 bzw. 20 Jahre der historischen Realisierungen der relevanten Variablen zugrunde. Das MD-Maß nimmt den Wert Null an, wenn alle Prognosen genau den Realisationen entsprechen. Dieses multivariate Maß berücksichtigt explizit die Korrelationen zwischen den Variablen sowie deren realisierte Varianzen. Für den Fall, dass $\boldsymbol{\Sigma} = \mathbf{I}$ eine Identitätsmatrix ist, reduziert sich das Maß D^2 auf das bekannte quadratische Euklidische Distanzmaß. Für weitere Details sei [McLachlan \(1999\)](#) empfohlen.

Da die Kovarianzmatrix $\boldsymbol{\Sigma}$ durchaus durch 'Ausreißer' negativ beeinflusst sein kann, was wiederum das Gesamtergebnis berührt, wurde folgende Robustheitsprüfung vorgenommen. Basierend auf den historischen Realisationen der letzten 20 Jahre der berücksichtigten Zeitreihen, werden mittels Mahalanobis-Distanzmaß die beiden Beobachtungen ermittelt, für welche das Maß am größten ist. Die Matrix der um diese beiden Beobachtungen bereinigten Zeitreihen wird dann als Grundlage für die Berechnung der Kovarianzmatrix $\boldsymbol{\Sigma}$ genutzt. Somit werden eventuelle Ausreißer außer Acht gelassen, was als eine Art pseudo-robuste Schätzung von $\boldsymbol{\Sigma}$ gelten kann.

⁵Für die Prozedur werden Instrumente benötigt. [Elliott u. a. \(2005\)](#) schlagen vor, (i) eine Konstante, (ii) eine Konstante und den verzögerten Prognosefehler, (iii) eine Konstante und die verzögerte Prognose sowie (iv) eine Konstante, verzögerte Prognosefehler und verzögerte Prognose zu wählen.

2.4 Friedman-Test auf Rangstabilität

Als weiteren Test wenden wir eine nicht-parametrische Prozedur an, die auf [Friedman \(1937, 1940\)](#) zurückgeht. Diese erlaubt es, die Prognoseperformance eines jeden Instituts einzuschätzen. Unter der Nullhypothese des Tests ist die Gruppe der Institute homogen, d.h. jedes Institut hat dieselbe Wahrscheinlichkeit, zu jeder Periode dieselbe Platzierung wie alle anderen Institute zu erreichen. Dies impliziert, dass die Platzierung eines Instituts im Durchschnitt nicht vom mittleren Rang abweichen sollte.

Für den Test werden die Platzierungen (Ränge) der einzelnen Institute für jedes Jahr notiert. Die resultierende Matrix ist eine $T \times k$ Matrix, wobei T die Anz. der Jahre (2005 bis 2014) und k die Anz. der Institute beschreibt. R_j steht dabei für die Rangsumme des j -ten Instituts über alle T Beobachtungen. Die Friedman-Statistik lautet

$$F_R = \frac{12}{Tk(k+1)} \sum_{j=1}^k R_j^2 - 3T(k+1) \sim \chi_{k-1}^2. \quad (9)$$

Die Teststatistik folgt für $T \geq 5$ einer χ^2 -Verteilung mit $k - 1$ Freiheitsgraden.

3 Daten

Die Datengrundlage für die Studie bilden alle für uns verfügbaren Prognosen folgender Institute zwischen den Jahren 2005 und 2014: das DIW in Berlin, das ifo Institut in München, das ifw in Kiel, das IMK in Düsseldorf, das RWI in Essen, das IWH in Halle, das HWWI in Hamburg sowie die Gemeinschaftsdiagnose (GD). Unsere Datenbank protokolliert alle eventuellen Einträge auf monatlicher Frequenz.

Tabelle 1 gibt einen Überblick über die von uns berücksichtigten Aggregate. Der Prognosewert der Institute für das Kalenderjahr t bezieht sich immer auf den prognostizierten durchschnittlichen Jahreswert (Jahreswachstumsrate).

| Variable | Transformation |
|-----------------------------------|---------------------|
| Reales Bruttoinlandsprodukt | Jahreswachstumsrate |
| Reale priv. Konsumausgaben | Jahreswachstumsrate |
| Reale öffentl. Konsumausgaben | Jahreswachstumsrate |
| Reale Ausrüstungsinvestitionen | Jahreswachstumsrate |
| Reale Bauinvestitionen | Jahreswachstumsrate |
| Reales sonst. Anlageinvestitionen | Jahreswachstumsrate |
| Reale Einfuhren | Jahreswachstumsrate |
| Reale Ausfuhren | Jahreswachstumsrate |
| BIP Deflator | Jahreswachstumsrate |
| Verbraucherpreisindex | Jahreswachstumsrate |
| Arbeitslosenquote | — |

Tabelle 1: Prognostizierte Aggregate aller Institute und jeweilige Transformation.

Alle durchschnittlichen Jahreswerte der Realisierungen der in Tabelle 1 aufgelisteten Aggregate wurden der Datenbank des Statistischen Bundesamtes entnommen.

Für die Tests auf strenge Rationalität (siehe Gl. 4) wird im Folgenden als zusätzliche Kontrollvariable der kurzfristige reale (basierend auf BIP-Deflator) EURIBOR 3-Monatszins verwendet (Datenquelle: Bundesbank).

In dieser Studie werden unterschiedliche Prognosehorizonte erfasst (siehe Tabelle 2). Aufgrund des hohen Informationsumfangs beschränken wir uns jedoch bei der Präsentation der Ergebnisse auf die in der 2. Jahreshälfte des aktuellen Kalenderjahres (t) abgegebenen Prognosen für das folgende Kalenderjahr ($t + 1$).⁶

| Prognose abgegeben | Prognose für das | |
|--------------------|------------------|----------------------|
| | Kalenderjahr t | Kalenderjahr $t + 1$ |
| 1. Halbjahr in t | x | x |
| 2. Halbjahr in t | x | x |

Tabelle 2: Illustration der untersuchten Prognoseevaluationszeiträume.

Dies bedeutet, dass für die folgenden Prognoseevaluationstests *alle* in der 2. Jahreshälfte des aktuell laufenden Kalenderjahres publizierten Prognosewerte eines Instituts für das kommende Kalenderjahr berücksichtigt werden, um eine genügend große Stichprobe zu erhalten. Eventuelle Abweichungen von dieser Regel werden bekanntgegeben. Die Prognosen eines Instituts für ein beliebiges Aggregat werden nur berücksichtigt, wenn für den relevanten Prognosehorizont mindestens 10 Beobachtungen für den Zeitraum 2005 bis 2014 existieren.

Es sei darüber hinaus angemerkt, dass die für den Diebold-Mariano Test (siehe Gl. 1) relevante *naive* Vergleichsprognose in dieser Auswertung immer die Vorjahresrealisation der zu untersuchenden Größe ist.

4 Grundlegende Ergebnisse für das Wachstum des realen Bruttoinlandsprodukts und die Inflation

Aufgrund der massiven Informationsdichte werden im Folgenden zunächst lediglich die Ergebnisse für die zentralen Variablen Wachstum und Inflation im Detail vorgestellt. Diese Auswahl schließt an die bisher vorliegende Literatur an (Döpke und Fritsche, 2006a).

Die Ergebnisse der unterschiedlichen Statistiken und Tests werden in grafischer Form dargestellt. Der Fokus liegt hierbei auf den Maßen der Prognosegüte, den Tests auf Effizienz und Rationalität unter der Annahme symmetrischer als auch asymmetrischer Verlustfunktionen. Zum Schluss folgen die Testergebnisse für die Prognose von Richtungsänderungen.

⁶Diese Prognosen des zweiten Halbjahres eines Kalenderjahres für das kommenden Kalenderjahr halten wir für einen sehr relevanten Prognosehorizont, in dem die Unterschiede in der Prognosequalität zum Tragen kommen. Bei kürzeren Prognosehorizonten werden, wie schon in der Vergangenheit gezeigt, die Unterschiede in den Prognosen deutlich kleiner (Döpke und Fritsche, 2006b). Alle anderen Ergebnisse sind auf Nachfrage verfügbar und werden in naher Zukunft über einen Online-Appendix zur Verfügung gestellt.

4.1 Wachstum des realen Bruttoinlandsprodukts

Abbildung 1 zeigt die Ergebnisse für die Prognose des realen BIP-Wachstums für alle acht Institute. In Bezug auf das Theilsche U liefern das ifo, ifw und HWWI die präzisesten Prognosen (siehe Abb. 1(a)). Die Ergebnisse des Diebold-Mariano Tests weisen darauf hin, dass jedes der Institute signifikant (im Folgenden wird eine 10% Irrtumswahrscheinlichkeit als “Signifikanzgrenze” angenommen) verschieden von der naiven Prognose prognostiziert.

In Bezug auf die Tests von Mincer-Zarnowitz (MZ) und Kirchgässner (EFF) auf Effizienz bzw. Rationalität unter der Annahme einer symmetrischen Verlustfunktion (siehe Abb. 1(b)), kann die jeweilige Nullhypothese lediglich für das DIW, ifw, RWI und die GD nicht abgelehnt werden. Für das IMK sowie HWWI kann zumindest der MZ-Test auf Rationalität nicht abgelehnt werden. Legt man jedoch den nicht-parametrischen CG1 Test zugrunde, dann kann die Null der Rationalität für keines der Institute verworfen werden. Nichtsdestotrotz lassen sich Hinweise auf Autokorrelation 1. Ordnung in den Prognosen finden: So z.B. für das RWI und das HWWI. Zusammengefasst weisen mindestens 3 der 4 Tests für alle Institute bis auf das ifo, IWH sowie HWWI auf keine Probleme bei der Rationalität hin.

In Hinblick auf den EKT-Test auf Symmetrie der Verlustfunktion (siehe Gl. (7)), kann die Nullhypothese für keines der Institute abgelehnt werden (siehe Abb. 1(c)). Das selbe gilt für die Tests auf Rationalität unter der Annahme von Symmetrie bzw. unter der Anwendung des geschätzten Symmetrie-Parameters $\hat{\alpha}$, wobei letztere Hypothese zumindest zum 10%-Niveau für das DIW verworfen werden kann. Alles in allem bestätigen die Ergebnisse auf Rationalität unter möglicher Asymmetrie die vorherigen Ergebnisse. Mögliche Hinweise auf Probleme bei der Rationalität gibt es demnach für das DIW, das ifo, das IWH und das HWWI.

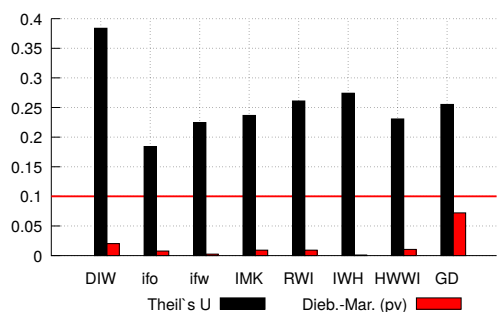
Mit Blick auf die Richtungsprognose (siehe Abb. 1(d)) kann gesagt werden, dass die Nullhypothese einer sich nicht vom Münzwurf unterscheidenden Richtungsprognose für alle Institute zum 10% Niveau abgelehnt werden kann, auch wenn die Irrtumswahrscheinlichkeiten für das DIW und GD bei ca. 8 % bzw. 6% liegen.⁷

4.2 Inflation basierend auf dem BIP-Deflator

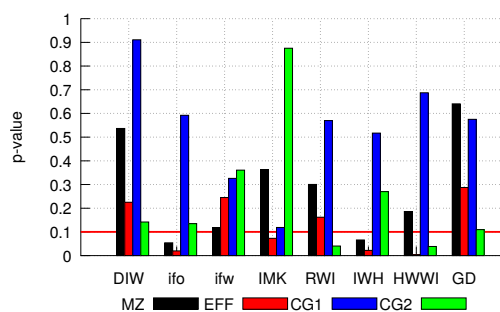
Wie bereits zuvor für die Prognose der BIP-Wachstumsrate, können wir die Nullhypothese, dass die Prognose für die Inflation im Durchschnitt nicht signifikant verschieden von der naiven Prognose ist, bis für das ifw (die Irrtumswahrscheinlichkeit liegt bei knapp unter 10% für das DIW) für alle übrigen Institute ablehnen (siehe Abb. 2(a)). Mit Bezug auf das Theilsche U liefern die präzisesten Prognosen die GD, das IMK, gefolgt vom HWWI.

Die parametrischen Tests auf Rationalität und strikte Effizienz können, wie in Abb. 2(b) ersichtlich, jeweils nur für das ifw zum 10% Signifikanzniveau abgelehnt werden. Unter Zuhilfenahme der beiden nicht-parametrischen Tests, finden wir lediglich für das RWI, HWWI und die GD übereinstimmende Hinweise auf keinerlei Probleme bzgl. der Rationalität. Zumindest für drei der vier Tests kann die Null der Rationalität jeweils nicht für das DIW abgelehnt werden. Die Ergebnisse sind weniger eindeutig für das ifo, ifw, IMK sowie IWH.

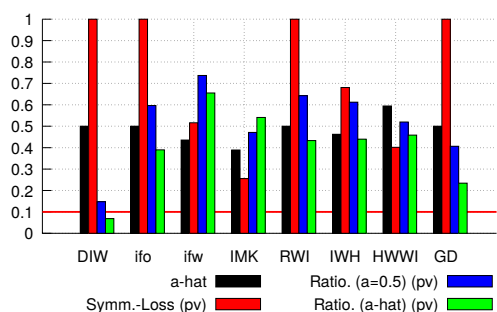
⁷Für alle in dieser Studie präsentierten Ergebnisse finden wir Evidenz dafür, dass der jeweilige Informationsgehalt der Prognosen des Diebold-Lopez (1996) Tests gleich $I = \frac{O_{ii}}{O_{ii}+O_{ji}} + \frac{O_{jj}}{O_{jj}+O_{ij}} > 1$ ist.



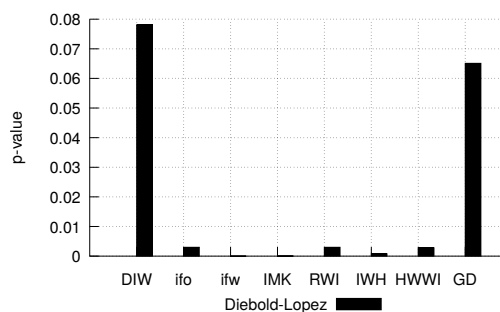
(a) Prognosegüte



(b) Tests auf Rationalität und Effizienz



(c) Test auf Symmetrie and Rationalität unter (A-)Symmetrie



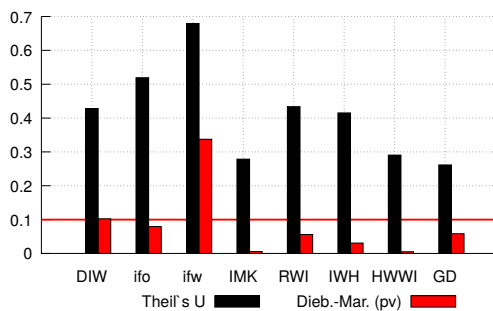
(d) Test der Richtungsprognose

NOTE: *Dieb.-Mar* bezeichnet die Irrtumswahrscheinlichkeit des Diebold-Mariano Tests mit der Nullhypothese, dass die Prognose nicht unterscheidbar von einer naiven Prognose ist (siehe Gl. (1)) unter der Annahme einer quadratischen (U-shaped) Verlustfunktion. *MZ* steht für die Bootstrap (999 Iterationen) Irrtumswahrscheinlichkeit des Mincer-Zarnowitz Tests auf Rationalität beschrieben in Gl. (2). *EFF* bezeichnet die Bootstrap (999 Iterationen) Irrtumswahrscheinlichkeit des erweiterten Tests von Kirchgässner (siehe Gl. (3)) auf strikte Rationalität. *CG1* und *CG2* bezeichnen die Irrtumswahrscheinlichkeiten des nicht-parametrischen Tests darauf, dass die Prognosefehler um Null zentriert (Gl. (5)) sind bzw. die Prognosefehler nicht autokorreliert 1. Ordnung sind (Gl. (6)). *a-hat*, *Sym-Loss*, *Ratio(a=0.5)* und *Ratio(a-hat)* bezeichnen den geschätzten Symmetrie-Parameter \hat{a} , die Irrtumswahrscheinlichkeit des Tests auf Symmetrie, die Irrtumswahrscheinlichkeit auf Rationalität unter der Annahme von Symmetrie bzw. unter Verwendung des Punktschätzers von \hat{a} (siehe dazu Elliott u. a., 2005, S. 1113). Unterstellt wird eine Lin-Lin Verlustfunktion sowie als Instrumente der verzögerte Prognosewert und der kurzfristige Realzins. *Diebold-Lopez* steht für die Irrtumswahrscheinlichkeit der Null, dass die Richtungsprognose nicht besser als ein Münzwurf ist.

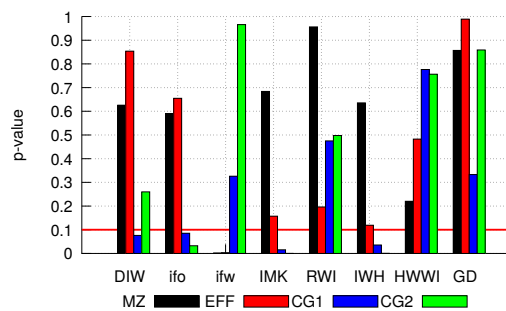
Abbildung 1: Zusammenfassung der Prognoseevaluationsergebnisse für das *reale BIP Wachstum*. Berücksichtigt werden alle Prognosen eines Instituts getätigt im zweiten Halbjahr des aktuellen Kalenderjahres für das kommende Kalenderjahr. Periode: 2005–2014.

Jedoch scheint für einige der Institute die Symmetrieannahme nicht gerechtfertigt zu sein. In Abb. 2(c) wird deutlich, dass die Null einer symmetrischen Verlustfunktion für das DIW (hier knapp zum 10% Niveau), IMK, IWH sowie HWWI jeweils abgelehnt werden kann. Diese Institute tendieren dazu, die Inflationsrate zu unterschätzen, da der Punktschätzer des Symmetrieparameters $\hat{a} < 0.5$ ist, was gleichbedeutend mit einer Verlustfunktion ist, welche die Überschätzung der Inflation härter als die Unterschätzung bestraft. Der Test auf Rationalität unter der Annahme von Symmetrie kann lediglich für das IMK sowie HWWI abgelehnt werden. Berücksichtigt man den geschätzten Symmetrie-Parameter \hat{a} und somit die Möglichkeit einer asymmetrischen Verlustfunktion, kann die Rationalitätsannahme nur für das RWI sowie das HWWI abgelehnt werden.

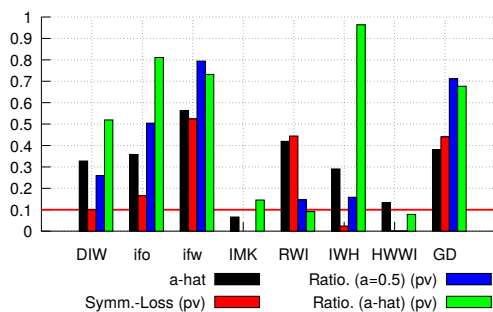
Wir finden des Weiteren, dass die Richtungsprognose für die Inflationsrate aller Institute signifikant verschieden als ein Münzwurf ausfällt (siehe Abb. 2(d)).



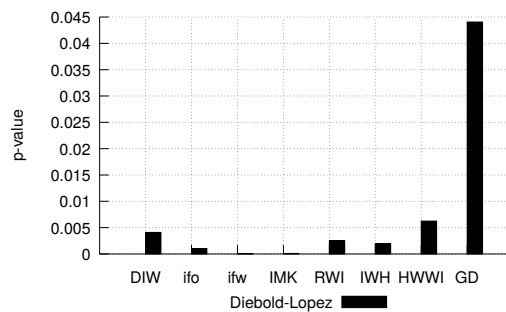
(a) Prognosegüte



(b) Test auf Rationalität und Effizienz



(c) Test auf Symmetrie und Rationalität unter (A-)Symmetrie



(d) Test der Richtungsprognose

NOTE: Für Erklärungen siehe Abb. 1.

Abbildung 2: Zusammenfassung der Prognoseevaluationsergebnisse für die *Inflationsrate* basierend auf dem BIP-Deflator. Berücksichtigt werden alle Prognosen getätigt im zweiten Halbjahr des aktuellen Kalenderjahres für das kommende Kalenderjahr. Periode: 2005–2014.

4.3 Inflation basierend auf dem Konsumentenpreisindex

Ähnlich im Vergleich zur Prognose der Inflationsrate basierend auf dem BIP-Deflator finden wir Hinweise darauf, dass die Prognose für den Verbraucherpreisindex signi-

signifikant verschieden von der naiven Prognose ausfällt (siehe Abb. 3(a)). Nimmt man das Theilsche U als Kriterium der Prognosegüte, so liefert das IMK gefolgt vom ifo, ifw und dem HWWI die präzisesten Prognosen.

Wie in Abb. 3(b) ersichtlich, kann für das ifw sowie das IWH die Nullhypothese des parametrischen Tests auf strenge Rationalität (wenn auch knapp) abgelehnt werden. Interessanterweise wird der nicht-parametrische CG1 Test auf Rationalität nur für das RWI sowie das ifo (nahe der 10% Signifikanzgrenze) abgelehnt; nicht aber für die übrigen Institute. Blickt man jedoch auf eventl. Autokorrelation in den Prognosefehlern mittels des CG2 Tests, so gibt es Hinweise auf Probleme der Rationalität bei allen Instituten.

Anders als im Falle des BIP-Deflators, kann die Symmetrieannahme der Verlustfunktion für keines der Institute zum 10% Signifikanzniveau abgelehnt werden, wie in Abb. 3(c) sichtbar. Unter der Annahme einer symmetrischen Verlustfunktion finden wir lediglich für das ifw sowie IWH Hinweise darauf, dass die Rationalitätsannahme abgelehnt werden kann — was wiederum die vorherigen Ergebnisse bestätigt. Die Irrtumswahrscheinlichkeit für DIW, ifo sowie auch IMK liegen nur geringfügig über dem 10% Niveau. Alles in allem sind die Ergebnisse bzgl. der Effizienz sowie Rationalität bei Inflationsprognosen nicht eindeutig, weisen aber zumindest auf Autokorrelation der Prognosefehler hin.

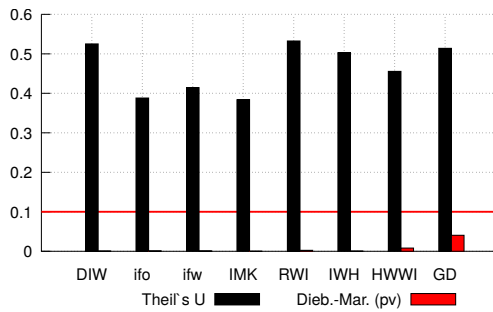
Im Unterschied zur Inflationsrate basierend auf dem BIP-Preisindex findet sich kein Institut, dessen Richtungsprognose (zum 10% Niveau) signifikant unterschiedlich zum Münzwurf ausfällt (siehe Abb. 3(d)). Allerdings liegen die Irrtumswahrscheinlichkeiten für das ifw, IMK, IWH sowie HWWI nur knapp über dem gewählten Signifikanzniveau.

4.4 Diskussion der Ergebnisse für weitere prognostizierte Variablen

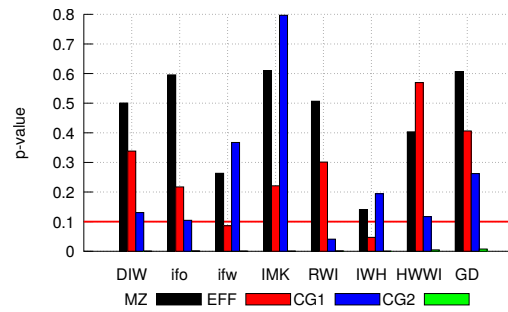
An dieser Stelle soll nur kurz auf die Ergebnisse für weitere Teilkomponenten des BIP sowie andere relevante makroökonomische Variablen eingegangen werden. Die dazugehörigen Abbildungen finden sich im Appendix ab Abb. 5.

In Bezug auf das Wachstum der privaten realen Konsumausgaben tendiert die Prognose für keines der Institute dazu, signifikant verschieden von der naiven Prognose abzuschneiden. Eine Verbesserung gegenüber der naiven Prognose anhand des Theilschen U s finden wir nur für das ifo, RWI und IWH. Lediglich für das DIW, ifo, IMK, RWI, IWH, HWWI und die GD gibt es Hinweise auf rationale Prognosen dieses Aggregats; entweder unter Anwendung nicht-parametrischer Tests bzw. unter Berücksichtigung einer asymmetrischen Verlustfunktion. Überdies scheinen das DIW, ifo und ifw bezüglich des privaten Konsums eine asymmetrische Verlustfunktion zu haben, mit der Eigenschaft, überoptimistische Prognosen zu präferieren. Auch finden wir für das DIW, das HWWI und die GD Schwächen in der Richtungsprognose für diese Größe.

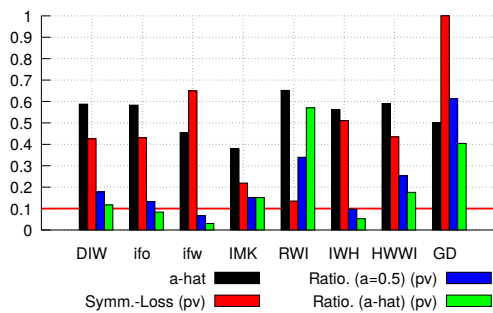
Im Gegensatz dazu finden wir, dass die Prognosen der öffentlichen Konsumausgaben sich signifikant von der naiven Prognose unterscheiden — dies gilt für alle Institute. Das Theilsche U liegt dabei für alle Institute unter 1. Wir finden starke Evidenz für asymmetrische Verlustfunktionen mit deutlich überpessimistischen Prognosen für das Wachstums der öffentlichen Konsumausgaben für alle Institute. Berücksichtigt man diese asymmetrische Verlustfunktion, finden wir für das ifw, IMK



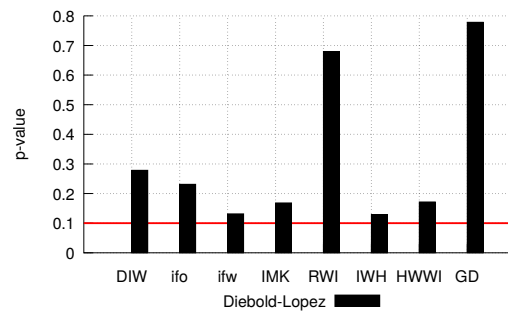
(a) Prognosegüte



(b) Test auf Rationalität und Effizienz



(c) Test auf Symmetrie und Rationalität unter (A-)Symmetrie



(d) Test der Richtungsprognose

NOTE: Für Erklärungen siehe Abb. 1.

Abbildung 3: Zusammenfassung der Prognoseevaluationsergebnisse für die *Inflationsrate* basierend auf dem Verbraucherpreisindex. Berücksichtigt werden alle Prognosen getätigt im zweiten Halbjahr des aktuellen Kalenderjahres für das kommende Kalenderjahr. Periode: 2005–2014.

und IWH Hinweise auf nicht-rationale Prognosen. Weiterhin zeigen drei von acht Institute, namentlich das DIW, ifw und RWI, Schwächen in der Richtungsprognose für dieses Teil-Aggregat.

Für die Ausrüstungsinvestitionen liefern das ifo, das IMK, das RWI und die GD die präzisesten Prognosen auf Basis des Theilschen U . Im Durchschnitt unterscheiden sich die Prognosen aller Institute dieses Aggregats signifikant von der naiven Prognose. Mit Blick auf die sonstigen Investitionsausgaben finden wir die geringsten Werte des Theilschen U für das IWH, die GD, das IMK und das HWWI. Lediglich für das DIW und die ifw gibt es keine ausreichenden Hinweise darauf, dass die Prognose signifikant verschieden von der naiven Prognose ist. Mit Blick auf die Bauinvestitionen gibt das IMK die genaueste Prognose im Durchschnitt ab, wobei die übrigen Institute ähnlich präzise prognostizieren. Wie bereits bei den Ausrüstungsinvestitionen finden wir, dass die Prognosen aller Institute jeweils signifikant besser als die naive Prognose ausfällt.

Bezüglich der Ausrüstungsinvestitionen gibt es geringe Hinweise gegen Rationalität für die untersuchten Institute. Mit Ausnahme des DIW gibt es keine Hinweise auf eine asymmetrische Verlustfunktion. Legt man für alle Institute die nicht- sowie parametrischen Tests zugrunde, so weisen mindestens 3 von 4 Tests auf Rationalität der Ausrüstungsinvestitionsprognosen für das ifo, ifw, RWI, IWH, HWWI sowie die GD hin. Für das DIW spricht nichts gegen die Annahme einer rationalen Prognose unter einer asymmetrischen Verlustfunktion.

Für die sonstigen sowie Bauinvestitionen finden wir starke Hinweise auf eine asymmetrische Verlustfunktion für die meisten Institute. Während für die sonstigen Investitionen die Null auf Rationalität unter einer asymmetrischen Verlustfunktion lediglich für das IWH abgelehnt werden kann, finden wir bzgl. der Bauinvestitionen Schwächen bei der Rationalität der Prognosen für das ifo, RWI, IWH sowie HWWI. Des Weiteren gibt es für das HWWI und die GD Hinweise auf Probleme bei der Richtungsprognose der Bauinvestitionen.

Mit Hinblick auf die außenwirtschaftliche Seite kommen wir zu folgenden Ergebnissen: Das ifo, ifw, IMK, HWWI und die GD geben die präzisesten Prognosen für das Exportwachstum ab, wenngleich diese für jedes der acht Institute signifikant verschieden von der naiven Prognose ist. Unter der Annahme einer symmetrischen Verlustfunktion gibt es, basierend auf den parametrischen und/oder nicht-parametrischen Tests, Hinweise auf Nicht-Rationalität allein für das RWI. Für das HWWI finden wir Evidenz einer asymmetrischen Verlustfunktion (wenngleich auch die Irrtumswahrscheinlichkeit des DIW knapp über 10% liegt). Insgesamt scheinen die meisten Institute keine Probleme bzgl. der Rationalität als auch der Richtungsprognose für dieses Aggregat zu haben.

Das ifw, das ifo und die GD geben die präzisesten Prognosen bzgl. des Importwachstums ab. Auch für dieses Aggregat fallen die Prognosen der Institute signifikant verschieden von der naiven Prognose aus. Unter der Annahme einer symmetrischen Verlustfunktion gibt es nur schwache Evidenz gegen Rationalität und Effizienz mit Ausnahme des DIW und zum Teil der GD. Jedoch gibt es, mit Ausnahme der GD, des IMK und des RWI, deutliche Hinweise für eine asymmetrische Verlustfunktion unter den Instituten. Legt man die Ergebnisse des Tests auf eine asymmetrische Verlustfunktion zugrunde, dann lassen sich für das DIW und das IWH Probleme bei der Rationalität finden. Durchweg scheinen die Institute keine Probleme bei der Richtungsprognose des Wachstums der Exporte zu haben.

Ein interessantes Bild ergibt sich bzgl. der Arbeitslosenquote. Das ifo, das HWWI, das ifw und die GD geben auf Basis des Theilschen U die präzisesten Prognosen ab. Nichtsdestotrotz finden wir nur für das ifo, ifw, IMK, RWI und HWWI, dass die Prognose systematisch verschieden von einer naiven Prognose ist. Die Nullhypothese einer symmetrischen Verlustfunktion kann nur für das DIW, IMK und das HWWI abgelehnt werden. Für die übrigen Institute weisen die nicht- und/oder parametrischen Tests (auf Rationalität) in der Regel auf Rationalität hin. Für das DIW, IMK sowie auch HWWI kann jeweils die Null der Rationalität unter einer asymmetrischen Verlustfunktion nicht zum 10% Niveau abgelehnt werden, wobei der Signifikanzwert des HWWI nur knapp darüber ist. Lediglich für das ifo (Irrtumswahrscheinlichkeit knapp bei 10%), ifw, und IMK finden wir Evidenz für eine Richtungsprognose, die systematisch besser als ein Münzwurf ausfällt. Die übrigen Institute zeigen diesbezüglich deutliche Schwächen.

5 Multivariate Analyse mittels Mahalanobis-Distanz

In diesem Abschnitt werden die Ergebnisse für die multivariate Analyse vorgestellt. Der Vektor der berücksichtigten Variablen umfasst alle die in Tabelle 1 aufgelisteten Größen. Die Kovarianzen der tatsächlichen Realisierungen werden jeweils basierend auf den historischen Daten der letzten 10 bzw. 20 Jahre kalkuliert (siehe Müller-Dröge u. a., 2014). Zur Prüfung der Robustheit der Ergebnisse werden zur Berechnung der Kovarianzmatrix die beiden Beobachtungen mit den größten Ausreißern aussortiert, wie in Kap. 2.3 beschrieben.

5.1 Ergebnisse basierend auf der Gesamtstichprobe

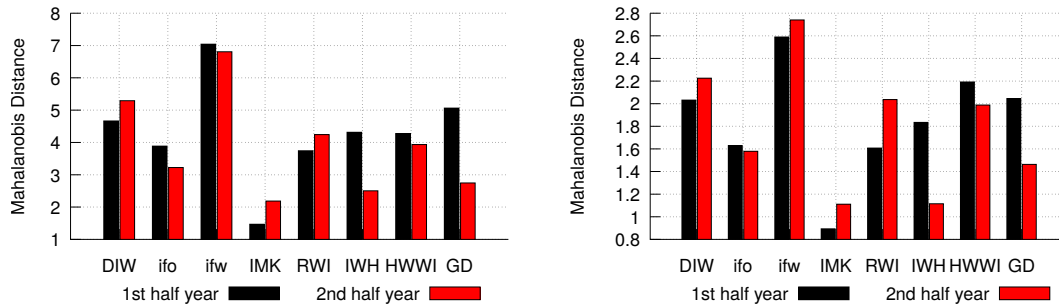
In diesem Abschnitt werden die Ergebnisse der multivariaten Analyse basierend auf der gesamten (pooled) Stichprobe von 2005 bis 2014 vorgestellt. Für die folgende Analyse berücksichtigen wir jeweils alle abgegebenen Prognosen der ersten Hälfte des Kalenderjahres t für das Folgejahr $t + 1$ bzw. im zweiten Fall alle Prognosen der zweiten Jahreshälfte in t für $t + 1$.

Die Wahl des zugrundeliegenden Fensters zur Berechnung der gemeinsamen Kovarianz hat keine qualitativen Auswirkungen auf die Ergebnisse, wie in Abb. 4 deutlich wird. Die Resultate zeigen, dass das IMK, gefolgt vom RWI, ifo und dem IWH, die präziseste Gesamtprognose für das kommende Kalenderjahr in der ersten Jahreshälfte des aktuellen Kalenderjahres abgab. Auf den hinteren Plätzen lassen sich das ifw, die GD sowie das DIW finden. Legt man jedoch die Prognose der 2. Kalenderjahreshälfte zugrunde, dann geben das IMK, gefolgt vom IWH, der GD und dem ifo die präzisesten Gesamtprognosen ab. Interessant ist, dass es für drei (das DIW, IMK und RWI) der acht Institute zu einer leichten Erhöhung der Mahalanobis-Distanz im Laufe des Kalenderjahres kommt.

Es sollte erwähnt werden, dass diese Ergebnisse sich von der robusten Variante kaum unterscheiden, wie in Abb. 13 im Appendix ersichtlich.

5.2 Ergebnisse für die jeweiligen Jahre

Um einen Eindruck über die zeitvariierende Gesamtprognoseperformance der Institute zu erhalten, berechnen wir im Folgenden die multivariate Mahalanobis-Distanz



(a) Basierend auf 10-jähriger Kovarianz

(b) Basierend auf 20-jähriger Kovarianz

NOTE: Der berücksichtigte Vektor an Variablen umfasst: Wachstum des realen BIP, Wachstum der privaten sowie öffentlichen Konsumausgaben, Wachstum der Ausrüstungs-, Bau- und sonstigen Investitionen, Wachstum der Ex- und Importe, der Arbeitslosenquote sowie der Preisinflation basierend auf dem BIP- bzw. VPI-Deflator.

Abbildung 4: Multivariate Mahalanobis-Distanz. Berücksichtigt werden alle Prognosen getätigt im ersten bzw. zweiten Halbjahr des aktuellen Kalenderjahres für das kommende Kalenderjahr. Periode: 2005–2014.

für jedes Jahr unserer Stichprobe von 2005 bis 2014. Berücksichtigt werden für jedes Kalenderjahr t nun alle abgegebenen Prognosen für das kommende Kalenderjahr — also *nicht* nur die publizierten Prognosen der 2. Jahreshälfte des aktuellen Kalenderjahres. Der Vektor der Prognosevariablen bleibt unverändert. Die zugrundeliegende Kovarianzmatrix basiert auf den realisierten Werten der letzten 20 Jahre.

In der Tabelle 3 werden die Platzierungen der Institute gem. der geschätzten multivariaten Mahalanobis-Distanz für jedes Jahr (rollierendes Fenster von einem Jahr) ausgewiesen. Beispielhaft weist nach diesen Ergebnissen das IWH Institut die genaueste Prognose für das Jahr 2006 im Jahr 2005 aus. Zusammenfassend lässt sich feststellen, dass das IMK in dem Zeitraum von 2005 bis 2014 fünf Mal den ersten Platz einnimmt. Das IWH und HWWI konnten jeweils zwei Mal den ersten Platz einnehmen und das DIW ein Mal. Zu berücksichtigen ist auch, dass das ifo vier Mal bzw. das IMK jeweils zwei Mal den zweiten Platz belegen konnten. Obgleich das HWWI zwei Mal den ersten Platz belegt, rangiert es in vier Jahren auf dem siebten Platz.

Die beste Krisenprognose für das Jahr 2009 gab dementsprechend das IMK ab, gefolgt vom ifo, dem IWH und der GD. Während das HWWI, gefolgt vom ifo Institut, die zügige Erholung der Konjunktur im Jahre 2009 für das Jahr 2010 wiederum in der “Story” weitgehend präzise voraussagte, lag das IMK auf dem siebten Platz in diesem Jahr.

Tabelle 4 gibt die durchschnittlichen Ränge und dazugehörigen Standardfehler aus. Es gibt über den Zeitraum von 10 Jahren doch systematische Unterschiede. Das IMK, das ifo und das IWH haben die besten Durchschnittsränge, allerdings ist die Varianz beim IMK deutlich höher. Auf ein Mittelfeld von RWI und DIW mit durchschnittlichen Rängen und kleiner bis mittlerer Standardabweichung folgt ein Schlussfeld mit GD, HWWI und ifw. Berücksichtigt man eventuelle Ausreißer bei der Berechnung der Kovarianz, dann belegen weiterhin das Trio aus IMK, ifo und IWH die Spitzenplätze, wie anhand der Tabellen 8 sowie 9 im Appendix deutlich wird. Jedoch rangiert nun das ifo auf dem ersten Platz, gefolgt vom IMK und dem IWH.

| Institut | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |
|----------|------|------|------|------|------|------|------|------|------|------|
| DIW | 7 | 3 | 3 | 8 | 8 | 8 | 1 | 2 | 4 | 3 |
| IFO | 2 | 4 | 5 | 2 | 2 | 4 | 6 | 4 | 3 | 2 |
| IFW | 5 | 8 | 6 | 5 | 3 | 6 | 8 | 8 | 8 | 8 |
| IMK | 8 | 1 | 8 | 1 | 7 | 2 | 2 | 1 | 1 | 1 |
| RWI | 3 | 7 | 7 | 7 | 4 | 3 | 3 | 5 | 2 | 5 |
| IWH | 1 | 2 | 4 | 3 | 5 | 1 | 4 | 6 | 5 | 4 |
| HWWI | 6 | 6 | 1 | 6 | 1 | 7 | 7 | 3 | 7 | 7 |
| GD | 4 | 5 | 2 | 4 | 6 | 5 | 5 | 7 | 6 | 6 |

NOTE: Die Spalten bezeichnen das Kalenderjahr der getätigten Prognose für das kommende Kalenderjahr, z.B. 2005 bezieht sich auf die Prognose aus 2005 für das Jahr 2006. Der Vektor der berücksichtigten Variablen umfasst: BIP, priv. Konsum, öffentl. Konsum, Ausrüstungs-, Bau- und sonst. Anlageinvestitionen, Ausfuhren, Einfuhren (alle in realen Größen und Wachstumsraten), Inflation (BIP), Inflation (VPI) sowie die Arbeitslosenquote. Die zugrundeliegende Kovarianzmatrix basiert auf den realisierten Werten der letzten 20 Jahre.

Tabelle 3: Jährlicher Rang der Institute basierend auf der Mahalanobis-Distanz für jedes Kalenderjahr.

| Institut | Durchschn.Rang | Std.-Fehler |
|----------|----------------|-------------|
| DIW | 4.7 | 2.6 |
| IFO | 3.4 | 1.4 |
| IFW | 6.5 | 1.7 |
| IMK | 3.2 | 3.0 |
| RWI | 4.6 | 1.8 |
| IWH | 3.5 | 1.6 |
| HWWI | 5.1 | 2.3 |
| GD | 5.0 | 1.3 |

Tabelle 4: Durchschnittliche Ränge und Standardabweichungen der Platzierungen aus Tabelle 3.

Das Ergebnis des Friedman Tests auf Ranghomogenität zwischen allen Instituten ist in Tabelle 5 ausgewiesen. Der Test basiert auf den ermittelten Prognoserängen der vorherigen Tabelle 3. Unter der Nullhypothese gibt es keine systematischen Differenzen in den Platzierungen zwischen den Instituten. Die Irrtumswahrscheinlichkeit die Nullhypothese abzulehnen, liegt knapp unter 5%, was auf Unterschiede in den Prognoserängen zwischen den Instituten hinweist.⁸ Dieses Ergebnis bestätigt sich auch dann, wenn für eventl. Ausreißer bei der Berechnung der Kovarianz kontrolliert wird (siehe Tabelle 10 im Appendix).

In Tabelle 6 replizieren wir die Ergebnisse aus Tabelle 5. Jedoch testen wir nun für alle Kombinationen zwischen zwei Instituten, ob deren Prognoseränge sich systematisch voneinander unterscheiden. So finden wir keine Evidenz dafür, dass es zwischen den Top-3 Instituten, dem IMK, ifo und IWH, jeweils signifikante Unterschiede in den Prognoserängen gibt. Jedoch kann z.B. die Null, dass es keine systematischen

⁸Wir haben den Test auch für diejenigen Prognosen durchgeführt, die sich auf das laufende Kalenderjahr beziehen und können dort ebenfalls mind. zum 10%-Niveau ablehnen.

| Teststat. | Prob. |
|-----------|-------|
| 14.267 | 0.047 |

Tabelle 5: Friedman Test auf keine systematischen Differenzen zwischen allen Instituten in den Prognoserängen gem. der Mahalanobis-Distanz. Periode: 2005–2014.

Platzierungsunterschiede zwischen dem ifo und der GD gibt, zum 10% Niveau verworfen werden. Ähnliches gilt z.B. für das IWH und die GD. Die Null, dass es keine systematischen Platzierungsdifferenzen zwischen der GD und dem IMK gibt, kann nicht abgelehnt werden. Im Appendix (siehe Tabelle 11 im Appendix) präsentieren wir die Ergebnisse basierend auf der um eventuelle Ausreißer korrigierten Kovarianzmatrix bei der Berechnung der Mahalanobis Distanz. Interessanterweise finden wir nun weitere Hinweise drauf, dass es durchaus systematische Unterschiede in den Prognoserängen zwischen den Instituten gibt.

| Institut | DIW | IFO | IFW | IMK | RWI | IWH | HWWI |
|-------------|-------|-------|-------|-------|-------|-------|-------|
| DIW | | | | | | | |
| IFO | 0.527 | | | | | | |
| IFW | 0.527 | 0.002 | | | | | |
| IMK | 0.206 | 0.206 | 0.206 | | | | |
| RWI | 1.000 | 0.206 | 0.206 | 0.206 | | | |
| IWH | 1.000 | 1.000 | 0.011 | 0.527 | 0.527 | | |
| HWWI | 1.000 | 0.206 | 0.206 | 0.206 | 1.000 | 0.206 | |
| GD | 1.000 | 0.058 | 0.011 | 0.206 | 0.206 | 0.011 | 0.206 |

Tabelle 6: Ablehnungswahrscheinlichkeiten des Friedman Test auf keine systematischen Differenzen in den Prognoserängen gem. der Mahalanobis-Distanz zwischen jeweils zwei Instituten. Periode: 2005–2014.

6 Zusammenfassung

Eine Zusammenfassung der recht uneinheitlichen Resultate scheint schwierig, dennoch schält sich bei näherer Betrachtung ein Gesamtbild heraus. Um die relevanten Ergebnisse dieser Studie zusammenzufassen, haben wir die folgende Tabelle 7 erstellt. Dort finden sich für alle berücksichtigten Aggregate Hinweise zur Güte, Rationalität, Symmetrie sowie Richtungsprognose. Diese Zusammenfassung gilt für die Ergebnisse der Prognosen, die in der zweiten Hälfte des laufenden Kalenderjahres für das kommende Kalenderjahr abgegeben wurden.⁹

In der Beurteilung der *Güte* ist das Fazit sowohl hinsichtlich der Resultate über die betrachteten Institute wie über die Variablen durchaus heterogen. Für die Mehr-

⁹Zu beachten ist, dass wir hier exemplarisch die Herbst-Winter-Prognosen der betrachteten Institute für das jeweilige kommende Kalenderjahr zusammengefasst haben. Resultate zu anderen Horizonten und aggregiert nach Institut über verschiedene Horizonte sind ebenfalls möglich und auf Anfrage verfügbar.

heit der Variablen — insbesondere die vielbeachteten Größen Wirtschaftswachstum (gemessen am BIP) und Inflationsrate — sind die Ergebnisse durchaus als gut zu bezeichnen. Für die meisten berücksichtigten Aggregate und über die meisten Institute hinweg betrachtet ist auch eine deutliche Prognoseverbesserung — gemessen am Theilschen U — gegenüber einer naiven Prognose darstellbar und in dem meisten Fällen auch statistisch signifikant. Über den betrachteten Zeitraum wird allerdings ein gewisses Problem mit der Güte der Prognosen von privaten Konsumausgaben sichtbar, da die Ergebnisse auf keine statistisch signifikante Abweichung von der naiven Prognose hindeuten. Darüber hinaus zeigen das DIW, das IWH sowie die GD Schwächen bei der Prognosegüte der Arbeitslosenquote, da auch in diesen Fällen die Prognose statistisch nicht signifikant von der naiven Prognose ausfällt.

Die Ergebnisse der *Rationalitätstests* — und auch das ist ein in der oben diskutierten Literatur schon mehrfach beschriebenes Ergebnis — sind uneinheitlich: Während es kaum Hinweise auf Verletzung des Rationalitätskriteriums für das BIP, die Inflationsrate, die sehr volatilen Ausrüstungsinvestitionen, die Ein- und Ausfuhren sowie Arbeitslosenquote gibt, deutet das Ergebnis für die übrigen Variablen auf Abweichungen vom Kriterium der Rationalität hin und die Notwendigkeit, hier die Qualität der Prognose weiter zu verbessern.

Wendepunktprognosen sind auch nach den Ergebnissen der hier vorgelegten Untersuchung weiterhin als ein schwieriges Feld zu bezeichnen. In einigen Fällen kann — trotz eines positiven Informationsgehaltes der Prognose — die Nullhypothese eines Münzwurfes nicht abgelehnt werden; dies betrifft insbesondere die Inflationsrate basierend auf dem Verbraucherpreisindex, die Konsumausgaben des Staates und der privaten Haushalte, die Bauinvestitionen sowie die Arbeitslosenquote. Bei den meisten anderen Aggregaten ist die Wendepunktprognose jedoch im statistischen Sinne deutlich besser als die naive Alternative des Münzwurfes.

Im Hinblick auf die *Symmetrie* der Verlustfunktion ist zu konstatieren, dass es für die in der Öffentlichkeit populären Prognosegrößen wie Wachstumsrate des BIP oder Inflationsrate (Konsumentenpreisindex) keine asymmetrische Verlustfunktion als Regelfall diagnostiziert werden kann. Interessant sind hier die Ergebnisse bei einigen Teilaggregaten des Konsums, bei den Investitionen, bei den Importen sowie bei den Arbeitslosenzahlen, die darauf hindeuten, dass es hier entweder systematische Fehleinschätzungen der konjunkturellen Triebkräfte durch verschiedene Institute über mehrere Jahre gab oder tatsächlich eine Asymmetrie in der Gewichtung von Über- und Unterschätzungen existierte. Für das IMK speziell lässt sich hier tendenziell eine Überschätzung der sonstigen sowie Bauinvestitionstätigkeit und der Arbeitslosenquote diagnostizieren. Auch tendiert das IMK dazu, die Inflationsrate (basierend auf dem BIP-Preisindex) sowie die öffentlichen Konsumausgaben zu unterschätzen. Andere Institute weisen systematische Über- oder Unterschätzungen bei anderen Variablen auf, so z.B. DIW, ifo, und ifw beim privaten Konsum; in der Regel alle Institute bei den Bauten sowie DIW, ifo, ifw, IWH, HWWI bei den Einfuhren. Hier besteht Bedarf, die Ursachen dieser systematischen Abweichung von der symmetrischen Verlustfunktionen herauszufinden.

Bezüglich der *multivariaten Prognosegüte* ergibt sich, dass es überraschend deutliche Unterschiede im Prognoseerfolg bezüglich der kohärenten “Stories” gab — für sich genommen ein dominantes Kriterium für die Güte von Prognosen (Tietzel, 1989) — und auch über einen längeren Zeitraum Unterschiede bei der Rangfolge der Institute statistisch signifikant zu finden sind. Das IMK schneidet hier mit dem

ifo Institut sowie dem IWH zusammen am besten ab und kann für sich genommen die meisten “Siegesplätze” bei der Rangfolge der Prognosegüte in einem multivariaten Rahmen verbuchen. Allerdings weist das IMK auch die höchste Varianz in den Rängen auf. Anders gewendet könnte das Ergebnis auch dahingehend interpretiert werden, dass die IMK-Prognosen in ihrer “Story” deutlich von der anderer (durchschnittlich erfolgreicher) Institute abweichen, aber damit das IMK auch in 5 von 10 Fällen “Prognosesieger” geworden ist. Dies passt im Übrigen zum Ergebnis von [Döpke und Fritsche \(2006a\)](#), wo argumentiert wurde, dass vom Konsens abweichende Prognosen nicht unbedingt schlechter abschneiden als konsensuale Prognosen, sondern im Gegenteil oftmals besser.

| | BIP | Infl. (BIP) | Infl. (VPI) | priv. Konsum | öff. Konsum | Ausrüstungen | Bauten | Sonst. Anlagen | Ein- u. Ausfuhr | Arbeitslosenquote |
|---|--|---|---|--|---|---|--|---|---|---|
| Güte | Prognosen deutlich besser als naive Prognose, Verbesserung signifikant | Prognosen in der Regel deutlich besser als naive Prognose; keine statistische Signifikanz einzig für das ifw | Theilsches $U < 1$: Prognose besser als naive Prognose; statistische Signifikanz | Prognose schlechter als naive Prognose ($U > 1$) für das DIW, ifw, IMK, HWWI und GD; jedoch keine statistische Signifikanz | Theilsches $U < 1$: Prognose besser als naive Prognose; statistische Signifikanz | Prognose deutlich besser als naive Prognose; geringste Werte des U für ifo, RWI und IMK | Theilsches $U < 1$: Prognose deutlich besser als naive Prognose; geringste Werte des U für IMK und HWWI | $U < 1$: Prognose besser als naive Prognose; jedoch keine statistische Signifikanz für DIW und ifw, geringste Werte des U für IWH und GD | Prognose besser als naive Prognose; statistisch signifikant; GD, IFO, IFW und HWWI mit genauesten Prognosen | Prognose besser als naive Prognose; statistisch nicht signifikant für DIW, IWH und GD |
| Rationalität (symmetrische Verlustfunktion) | Kaum Abweichungen von Rationalitätshypothese | Parametrische Tests deuten kaum auf Verletzung der Rationalitätshypothese hin, nicht-parametrische Test jedoch schon für das IFO, IMK und IWH | Parametrische Tests deuten kaum auf Verletzung der Rationalitätshypothese hin, nicht-parametrischer Test auf Autokorrelation jedoch schon | Rationalitätsprobleme bei allen Institutionen bis auf GD; Autokorrelierte Prognosefehler bei IFO, IFW, IMK und IWH | Parametrische Tests deuten auf keine Abweichungen von der Rationalität hin, nicht-parametrische jedoch schon | Kaum Verletzung der Rationalität | Kaum Verletzung der Rationalität | Deutliche Abweichungen von Rationalitätshypothese bis auf IWH, HWWI und GD | Kaum Verletzungen der Rationalitätsannahme; Rationalitätsprobleme des RWI bei Exporten und des DIW bei Importen | Kaum Abweichungen von Rationalitätshypothese |
| Symmetrie und Rationalität (asymmetrische Verlustfunktion) | Kann nicht abgelehnt werden | Muss für DIW, IMK, IWH, HWWI abgelehnt werden (tendenziell Unterschätzung); Rationalitätsprobleme bei RWI und HWWI | Kann nicht abgelehnt werden; Rationalitätsprobleme bei IFW und IWH | Muss für DIW, IFO und IFW abgelehnt werden (systematische Überschätzung); jedoch Hinweise auf Rationalität bei DIW und IFO | Starke Evidenz für asymmetrische Verlustfunktion (Unterschätzung); Probleme mit Rationalität bei IFW, IMK und IWH | Nur für DIW Asymmetrie (Überschätzung) aber keine Evidenz gegen Rationalität | Starke Evidenz für asymmetrische Verlustfunktion (Überschätzung); bei DIW, IFW, IMK und GD keine Hinweise gegen Rationalität | Tendenziell Überschätzung, in der Regel rationale Prognose bei angekommener asymmetrischer Verlustfunktion bis auf IWH | Systematische Überschätzung des Importwachstums bei DIW, IFO, IFW, IWH, HWWI; Systematische Überschätzung der Exporte beim HWWI | Systematische Überschätzung bei DIW und IMK, Systematische Unterschätzung bei HWWI; Rationalität bei Asymmetrie |
| Richtungsprognose | Ablehnung der Münzwurf-Hypothese | Ablehnung der Münzwurf-Hypothese | keine Ablehnung der Münzwurf-Hypothese | Keine Ablehnung der Münzwurf-Hypothese für DIW, HWWI, GD | Keine Ablehnung der Münzwurf-Hypothese für DIW, IFW und RWI | Ablehnung der Münzwurf-Hypothese | Keine Ablehnung der Münzwurf-Hypothese für HWWI und GD | Ablehnung der Münzwurf-Hypothese | Ablehnung der Münzwurf-Hypothese | Keine Ablehnung der Münzwurf-Hypothese für DIW, IFO, RWI, IWH, HWWI, GD |

Tabelle 7: Zusammenfassung der Ergebnisse der Studie

Literatur

- [Batchelor und Peel 1998] BATCHELOR, R. A. ; PEEL, D. A.: Rationality Testing under Asymmetric Loss. In: *Economics Letters* 61 (1998), S. 49–54
- [Bezemer 2011] BEZEMER, Dirk J.: No One Saw This Coming: Understanding Financial Crisis Through Accounting Models. In: *Accounting, Organizations and Society* 35 (2011), Nr. 7, S. 676–688
- [Brunnermeier u. a. 2012] BRUNNERMEIER, Markus K. ; EISENBACH, Thomas M. ; SANNIKOV, Yuliy: Macroeconomics with financial frictions: A survey / National Bureau of Economic Research. 2012. – Forschungsbericht
- [Campbell und Ghysels 1995] CAMPBELL, B. ; GHYSELS, E.: Federal Budget Projections: A Nonparametric Assessment of Bias and Efficiency. In: *The Review of Economics and Statistics* 77 (1995), Nr. 1, S. 17 – 31
- [Campbell und Ghysels 1997] CAMPBELL, Bryan ; GHYSELS, Eric: An Empirical Analysis of the Canadian Budget Process. In: *The Canadian Journal of Economics* 30 (1997), August, Nr. 3, S. 553–576
- [De Paoli u. a. 2014] DE PAOLI, Bianca ; SBORDONE, Argia ; TAMBALOTTI, Andrea: *A Bird's Eye View of the FRBNY DSGE Model*. 2014. – URL <http://libertystreeteconomics.newyorkfed.org/2014/09/a-birds-eye-view-of-the-frbny-dsge-model.html#.VCKoURaqKDY>
- [Diebold und Lopez 1996] DIEBOLD, Francis X. ; LOPEZ, Jose A.: Forecast Evaluation and Combination. In: MADDALA, G. S. (Hrsg.) ; RAO, C. R. (Hrsg.): *Handbook of Statistics* Bd. 14:Statistical Methods in Finance. 1996, S. 241–268
- [Diebold und Mariano 1995] DIEBOLD, Francis X. ; MARIANO, Roberto S.: Comparing Predictive Accuracy. In: *Journal of Business and Economic Statistics* 20 (1995), January, Nr. 1, S. 134–144
- [Döhrn und Schmidt 2011] DÖHRN, Roland ; SCHMIDT, Christoph M.: Information or Institution? On the Determinants of Forecast Accuracy. In: *Jahrbücher für Nationalökonomie und Statistik (Journal of Economics and Statistics)* 231 (2011), Nr. 1, S. 9–27
- [Döpke und Fritsche 2006a] DÖPKE, Jörg ; FRITSCHKE, Ulrich: Growth and inflation forecasts for Germany: A panel-based assessment of accuracy and efficiency. In: *Empirical economics : a journal of the Institute for Advanced Studies, Vienna, Austria* 31 (2006), Nr. 3, S. 777–798
- [Döpke und Fritsche 2006b] DÖPKE, Jörg ; FRITSCHKE, Ulrich: When do forecasters disagree? An assessment of German growth and inflation forecast dispersion. In: *International Journal of Forecasting* 22 (2006), Nr. 1, S. 125–135
- [Döpke und Fritsche 2009] DÖPKE, Jörg ; FRITSCHKE, Ulrich: Zehn Gebote zum Umgang mit Konjunkturprognosen: Ten commandments on the use of business cycle forecasts. In: *Zeitschrift für Politikberatung* 2 (2009), Nr. 1, S. 26–42

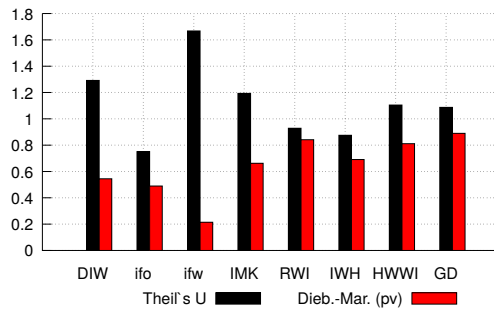
- [Dovern und Jannsen 2015] DOVERN, Jonas ; JANNSEN, Nils: Systematic Errors in Growth Expectations over the Business Cycle / Kiel Institute for the World Economy. February 2015 (1989). – Kiel Working Paper
- [Dufour 1981] DUFOUR, J.: Rank Tests for Serial Dependence. In: *Journal of Time Series Analysis* 2 (1981), Nr. 3, S. 117–128
- [Elliott u. a. 2005] ELLIOTT, Graham ; TIMMERMANN, Allan ; KOMUNJER, Ivana: Estimation and testing of forecast rationality under flexible loss. In: *The Review of Economic Studies* 72 (2005), Nr. 4, S. 1107–1125
- [European Central Bank 2014] EUROPEAN CENTRAL BANK: Fifteen Years of the ECB Survey of Professional Forecasters. In: *ECB Monthly Bulletin* 3 (2014), S. 55–67
- [Fildes und Stekler 2002] FILDES, R. ; STEKLER, H.O: The State of Macroeconomic Forecasting. In: *Journal of Macroeconomics* 24 (2002), Nr. 4, S. 435–468
- [Friedman 1937] FRIEDMAN, Milton: The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance. In: *Journal of the American Statistical Association* (1937), Nr. 32, S. 675–701
- [Friedman 1940] FRIEDMAN, Milton: A Comparison of Alternative Tests of Significance for the Problem of m Rankings. In: *Annals of the Mathematical Statistics* (1940), Nr. 11, S. 86–97
- [Friedman 1953] FRIEDMAN, Milton: The methodology of positive economics. In: CHICAGO PRESS, University of (Hrsg.): *Essays in positive economics*. University of Chicago Press, 1953, S. 30–43
- [Fritsche und Döpke 2006] FRITSCH, Ulrich ; DÖPKE, Jörg: Treffgenauigkeit, Rationalität und Streuung von Konjunkturprognosen für Deutschland. In: *Vierteljahrshefte zur Wirtschaftsforschung* 75 (2006), Nr. 2, S. 34–53
- [Gaffney 2011] GAFFNEY, Mason: An Award for Calling the Crash. In: *Econ Journal Watch* 8 (2011), Nr. 2, S. 185–192
- [Granger 1969] GRANGER, Clive W. J.: Prediction with a Generalized Cost of Error Function. In: *Operational Research Quarterly* 20 (1969), S. 199–207
- [Hansen 1982] HANSEN, Lars P.: Large Sample Properties of Generalized Method of Moments Estimators. In: *Econometrica* 50 (1982), Nr. 4, S. 1029–1054
- [Harvey u. a. 1997] HARVEY, David ; LEYBOURNE, Stephen ; NEWBOLD, Paul: Testing the equality of prediction mean squared errors. In: *International Journal of Forecasting* 13 (1997), June, Nr. 2, S. 281–291
- [Henzel u. a. 2013] HENZEL, Steffen ; NIERHAUS, Wolfgang ; WOLLMERSHÄUSER, Timo: Evaluation der ifo Konjunkturprognosen. In: *ifo Schnelldienst* 67 (2013), Nr. 17, S. 43–45
- [Holden und Peel 1990] HOLDEN, Ken ; PEEL, David A.: On testing for unbiasedness and efficiency of forecasts. In: *The Manchester School* 58 (1990), Nr. 2, S. 120–127

- [IMF 2014] IMF: *IMF Forecasts: Process, Quality, and Country Perspectives*. International Monetary Fund, 2014. – This report was prepared by an IEO team led by Hans Genberg
- [Kirchgässner und Savioz 2001] KIRCHGÄSSNER, Gebhard ; SAVIOZ, Marcel R.: Monetary Policy and Forecasts for Real GDP Growth: An Empirical Investigation for the Federal Republic of Germany. In: *German Economic Review* 2 (2001), Nr. 4, S. 339–365
- [Kirchgässner 1984] KIRCHGÄSSNER, Gebhard: Sind die Erwartungen der Wirtschaftssubjekte rational? Eine empirische Untersuchung für die Bundesrepublik Deutschland. In: *Weltwirtschaftliches Archiv* 120 (1984), S. 279–300
- [Löwe 1929] LÖWE, Adolph: Besprechung von Oskar Morgenstern, Wirtschaftsprognose. In: *Zeitschrift für die gesamte Staatswissenschaft* 87 (1929), S. 419–423
- [McLachlan 1999] MCLACHLAN, G. J.: Mahalanobis distance. In: *Resonance* 4 (1999), Nr. 6, S. 20–26. – URL <http://dx.doi.org/10.1007/BF02834632>. – ISSN 0973-712X
- [Mincer und Zarnowitz 1969] MINCER, Jacob A. ; ZARNOWITZ, Victor: *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*. Kap. The Evaluation of Economic Forecasts, S. 3–46, NBER, 1969
- [Morgenstern 1928] MORGENSTERN, Oskar: *Wirtschaftsprognose : eine Untersuchung ihrer Voraussetzungen und Möglichkeiten*. Wien : Springer, 1928
- [Müller-Dröge u. a. 2014] MÜLLER-DRÖGE, Hans C. ; SINCLAIR, Tara M. ; STEKLER, Herman O.: *Evaluating Forecasts of a Vector of Variables: a German Forecasting Competition*. 2014. – URL https://cama.crawford.anu.edu.au/sites/default/files/publication/cama_crawford_anu_edu_au/2014-07/55_2014_muller-droge_sinclair_stekleer.pdf
- [Nienhaus 2009] NIENHAUS, Lisa: *Die Blindgänger: Warum die Ökonomen auch künftige Krisen nicht erkennen werden*. Frankfurt a.M. and New York : Campus-Verl., 2009. – ISBN 3593390795
- [Nierhaus 2014] NIERHAUS, Wolfgang: Wirtschaftskonjunktur 2014: Prognose und Wirklichkeit. In: *ifo Schnelldienst* 68 (2014), Nr. 2, S. 43–49
- [Nordhaus 1987] NORDHAUS, W. D.: Forecasting Efficiency: Concepts and Applications. In: *The Review of Economics and Statistics* 69 (1987), Nr. 4, S. 667–674
- [Pain u. a. 2014] PAIN, Nigel ; LEWIS, Christine ; DANG, Thai-Thanh ; JIN, Yosuke ; RICHARDSON, Pete: OECD Forecasts During and After the Financial Crisis: A Post Mortem / Organisation for Economic Co-operation and Development. URL <http://dx.doi.org/10.1787/5jz7311qw1s1-en>, 2014 (1107). – OECD Economics Department Working Papers
- [Patton und Timmermann 2007] PATTON, Andrew J. ; TIMMERMANN, Allan: Testing Forecast Optimality Under Unknown Loss. In: *Journal of the American Statistical Association* 102 (2007), S. 1172–1184. – URL <http://ideas.repec.org/a/bs/jnlasa/v102y2007mdecemberp1172-1184.html>

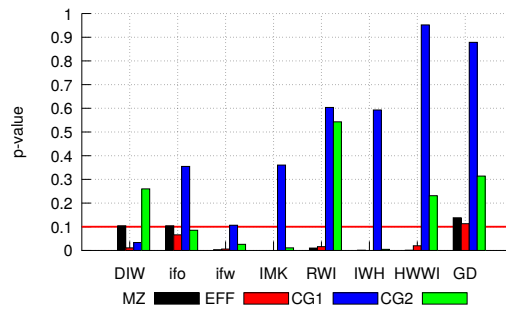
- [Riedel 2013] RIEDEL, Frank: *Die Schuld der Ökonomen: Was Mathematik und Ökonomie zur Krise beitrugen*. Berlin : Econ, 2013. – ISBN 343020156X
- [Sargan 1958] SARGAN, John D.: The Estimation of Economic Relationships using Instrumental Variables. In: *Econometrica* 26 (1958), Nr. 3, S. 393–415
- [Stekler 1972] STEKLER, Herman O.: An Analysis of Turning Point Forecasts. In: *American Economic Review* 62 (1972), Nr. 4, S. 724–729
- [Stiglitz 2011] STIGLITZ, Joseph E.: Rethinking Macroeconomics What Failed, and How to Repair It. In: *Journal of the European Economic Association* 9 (2011), August, Nr. 4, S. 591–645
- [Stiglitz 2014] STIGLITZ, Joseph E.: Reconstructing Macroeconomic Theory to Manage Economic Policy / NBER. 2014 (20517). – NBER Discussion Paper
- [Theil 1966] THEIL, Henri: *Applied Economic Forecasting*. Chicago : Rand McNally, 1966
- [Tietzel 1989] TIETZEL, Manfred: Prognoselogik oder: warum Prognostiker irren dürfen: On the logic of economic forecasting. In: *Jahrbücher für Nationalökonomie und Statistik* 206 (1989), Nr. 6, S. 546–562
- [Wilcoxon u. a. 1970] WILCOXON, Frank S. ; KATTI, S. K. ; WILCOX, Roberta: Critical values and probability levels for the Wilcoxon rank sum test and the Wilcoxon signed rank test. In: HARTER, H. L. (Hrsg.) ; OWEN, D. B. (Hrsg.): *Selected Tables in Mathematical Statistics*. Markham, 1970 (Vol. I), Kap. Critical values and probability levels for the Wilcoxon rank sum test and the Wilcoxon signed rank test
- [Wollmershäuser 2015] WOLLMERSHÄUSER, Timo: Evaluation der ifo Konjunkturprognosen – ein Vergleich mit den Prognosen von Consensus Economics. In: *ifo Schnelldienst* 68 (2015), Nr. 22, S. 26–18

Appendix

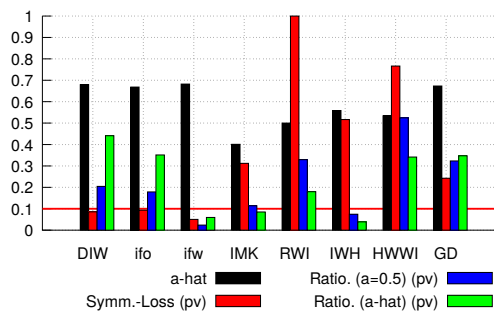
Abbildungen



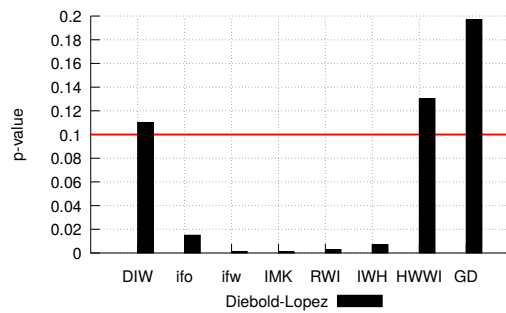
(a) Prognosegüten



(b) Test auf Rationalität und Effizienz



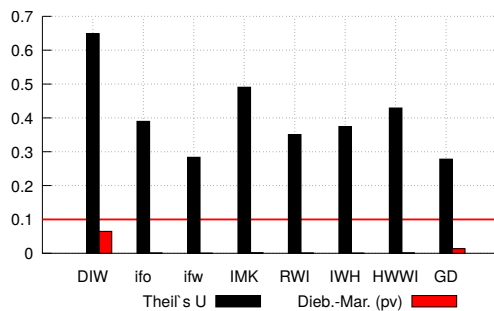
(c) Test auf Symmetrie und Rationalität unter (A-)Symmetrie



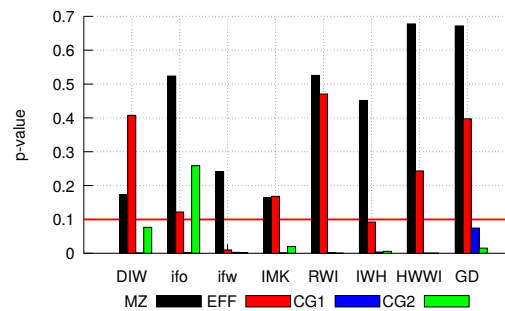
(d) Test der Richtungsprognose

NOTE: Für Erklärungen siehe Abb. 1.

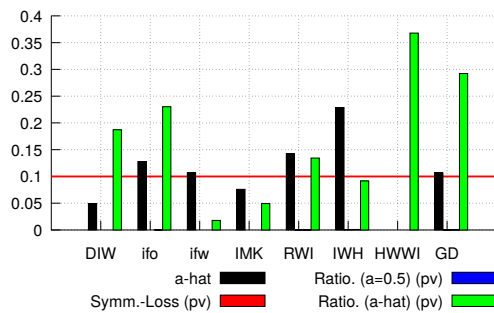
Abbildung 5: Zusammenfassung der Prognoseevaluationsergebnisse für das *Wachstum der priv. realen Konsumausgaben*. Berücksichtigt werden alle Prognosen getätigt im zweiten Halbjahr des aktuellen Kalenderjahres für das kommende Kalenderjahr. Periode: 2005–2014.



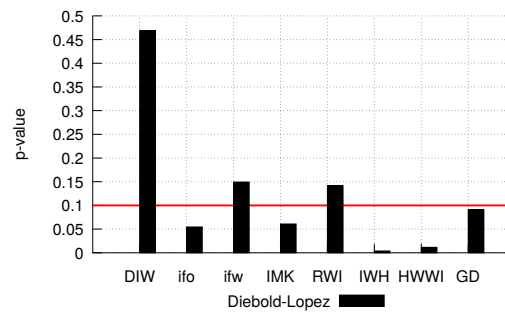
(a) Prognosegüten



(b) Test auf Rationalität und Effizienz



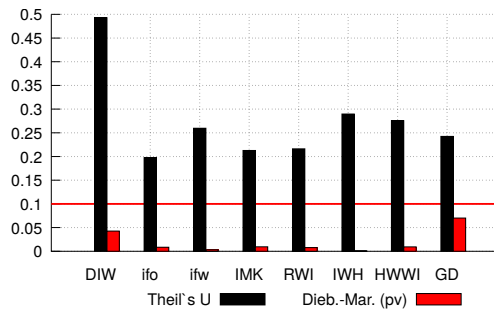
(c) Test auf Symmetrie and Rationalität unter (A-)Symmetrie



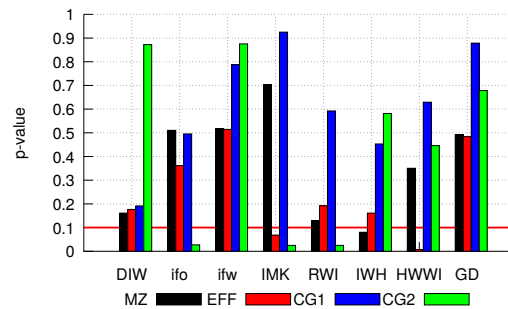
(d) Test der Richtungsprognose

NOTE: Für Erklärungen siehe Abb. 1.

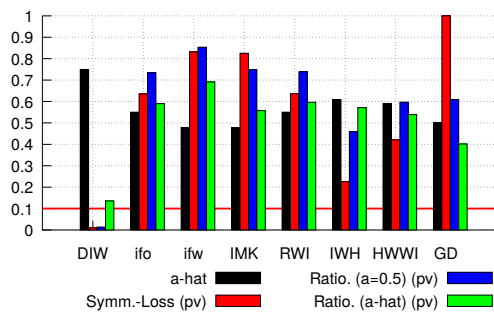
Abbildung 6: Zusammenfassung der Prognoseevaluationsergebnisse für das *Wachstum der öffentl. realen Konsumausgaben*. Berücksichtigt werden alle Prognosen getätigt im zweiten Halbjahr des aktuellen Kalenderjahres für das kommende Kalenderjahr. Periode: 2005–2014.



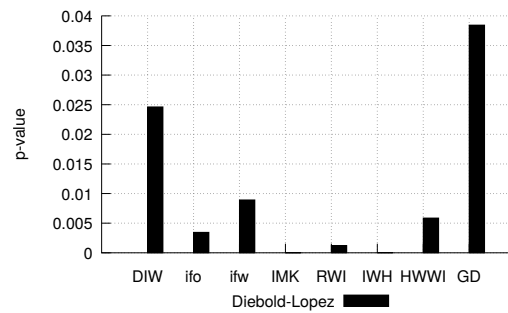
(a) Prognosegüten



(b) Test auf Rationalität und Effizienz



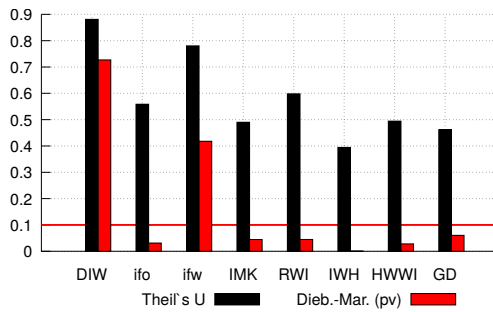
(c) Test auf Symmetrie und Rationalität unter (A-)Symmetrie



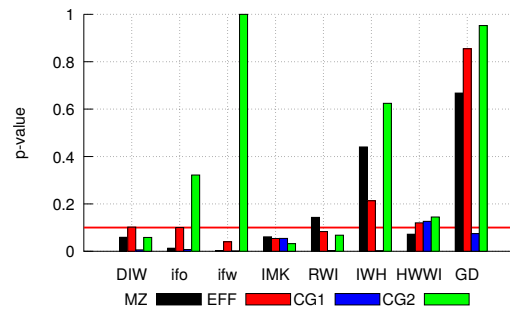
(d) Test der Richtungsprognose

NOTE: Für Erklärungen siehe Abb. 1.

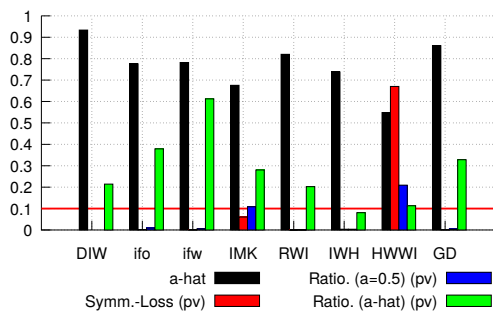
Abbildung 7: Zusammenfassung der Prognoseevaluationsergebnisse für das *Wachstum der realen Ausrüstungsinvestitionen*. Berücksichtigt werden alle Prognosen getätigt im zweiten Halbjahr des aktuellen Kalenderjahres für das kommende Kalenderjahr. Periode: 2005–2014.



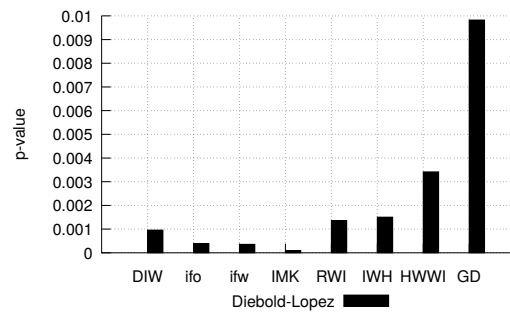
(a) Prognosegüten



(b) Test auf Rationalität und Effizienz



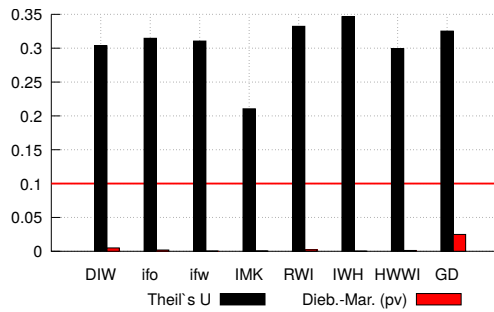
(c) Test auf Symmetrie und Rationalität unter (A-)Symmetrie



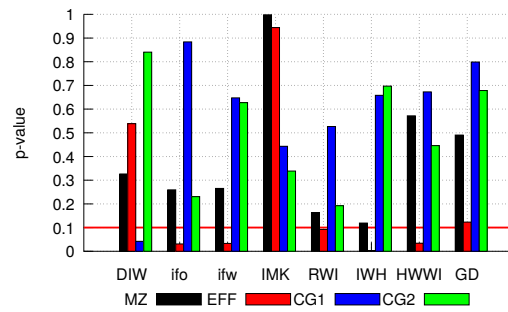
(d) Test der Richtungsprognose

NOTE: Für Erklärungen siehe Abb. 1.

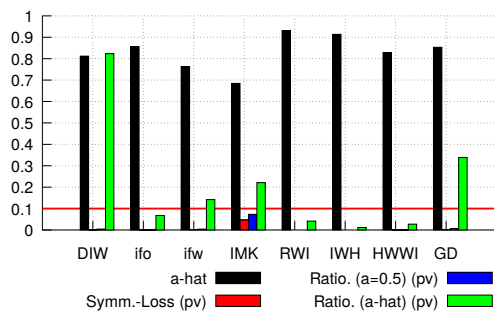
Abbildung 8: Zusammenfassung der Prognoseevaluationsergebnisse für das *Wachstum der sonstigen realen Investitionen*. Berücksichtigt werden alle Prognosen getätigt im zweiten Halbjahr des aktuellen Kalenderjahres für das kommende Kalenderjahr. Periode: 2005–2014.



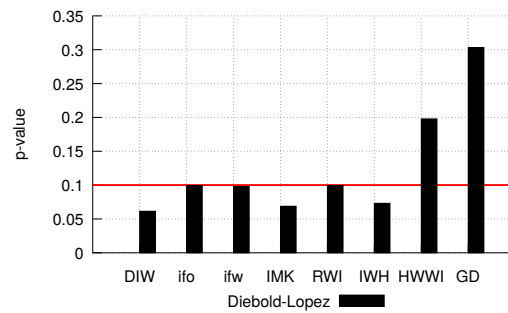
(a) Prognosegüten



(b) Test auf Rationalität und Effizienz



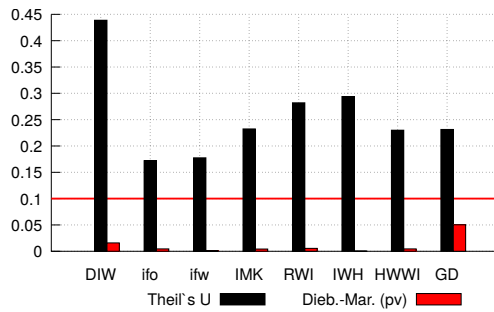
(c) Test auf Symmetrie und Rationalität unter (A-)Symmetrie



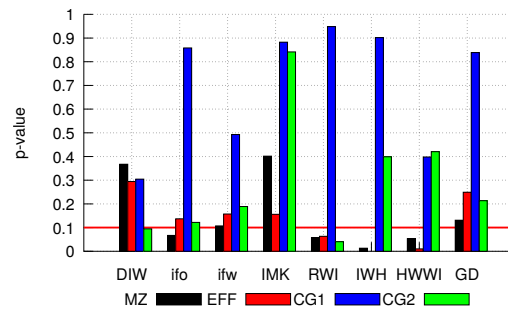
(d) Test der Richtungsprognose

NOTE: Für Erklärungen siehe Abb. 1.

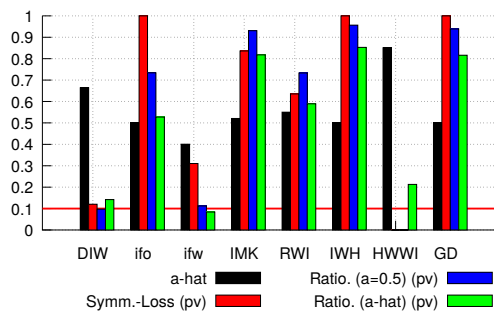
Abbildung 9: Zusammenfassung der Prognoseevaluationsergebnisse für das *Wachstum der realen Bauinvestitionen*. Berücksichtigt werden alle Prognosen getätigt im zweiten Halbjahr des aktuellen Kalenderjahres für das kommende Kalenderjahr. Periode: 2005–2014.



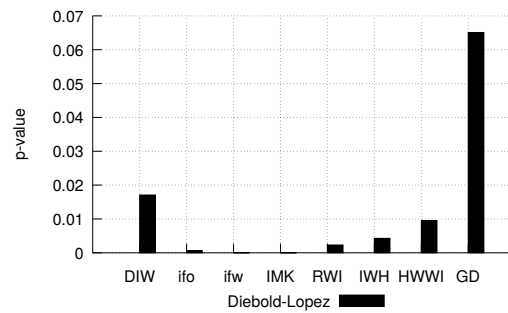
(a) Prognosegüten



(b) Test auf Rationalität und Effizienz



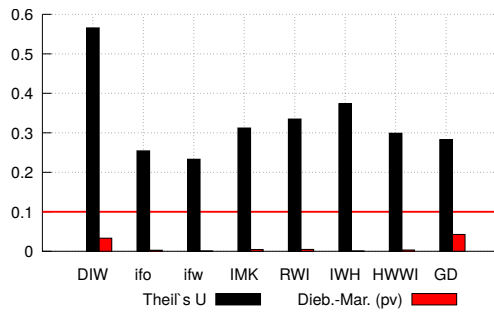
(c) Test auf Symmetrie und Rationalität unter (A-)Symmetrie



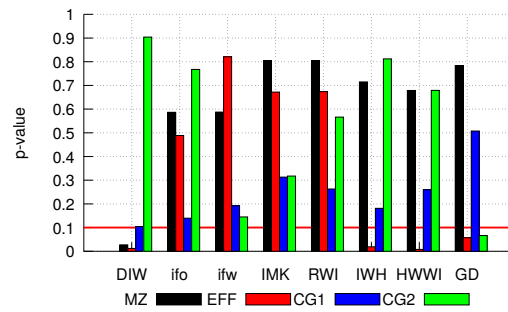
(d) Test der Richtungsprognose

NOTE: Für Erklärungen siehe Abb. 1.

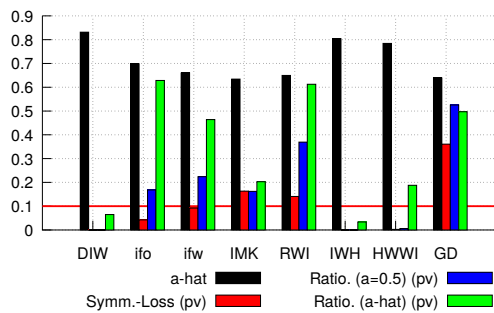
Abbildung 10: Zusammenfassung der Prognoseevaluationsergebnisse für das *Wachstum der realen Exporte*. Berücksichtigt werden alle Prognosen getätigt im zweiten Halbjahr des aktuellen Kalenderjahres für das kommende Kalenderjahr. Periode: 2005–2014.



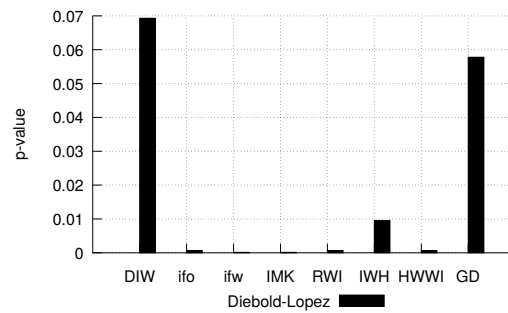
(a) Prognosegüten



(b) Test auf Rationalität und Effizienz



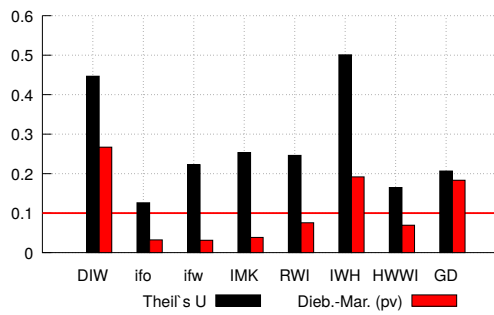
(c) Test auf Symmetrie und Rationalität unter (A-)Symmetrie



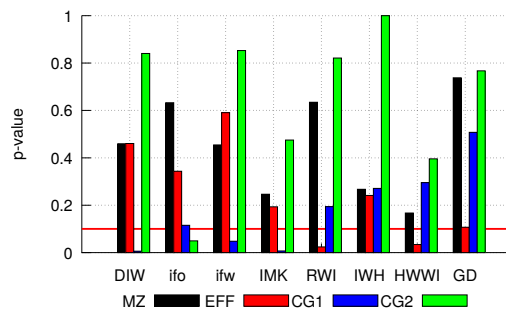
(d) Test der Richtungsprognose

NOTE: Für Erklärungen siehe Abb. 1.

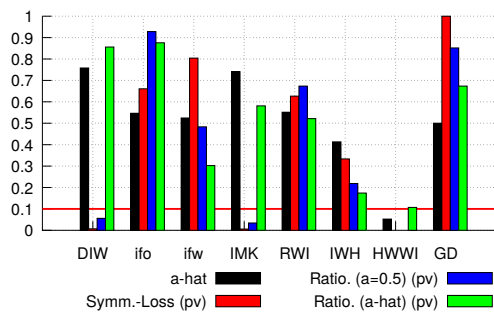
Abbildung 11: Zusammenfassung der Prognoseevaluationsergebnisse für das *Wachstum der realen Importe*. Berücksichtigt werden alle Prognosen getätigt im zweiten Halbjahr des aktuellen Kalenderjahres für das kommende Kalenderjahr. Periode: 2005–2014.



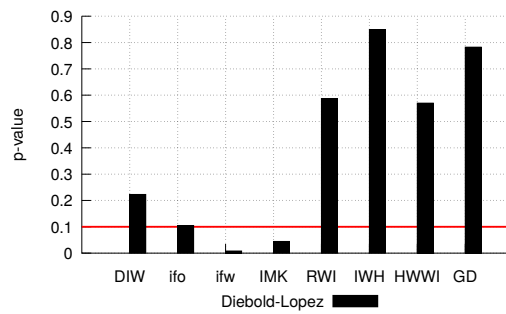
(a) Prognosegüten



(b) Test auf Rationalität und Effizienz



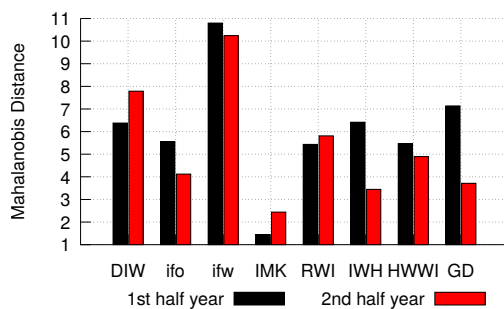
(c) Test auf Symmetrie und Rationalität unter (A-)Symmetrie



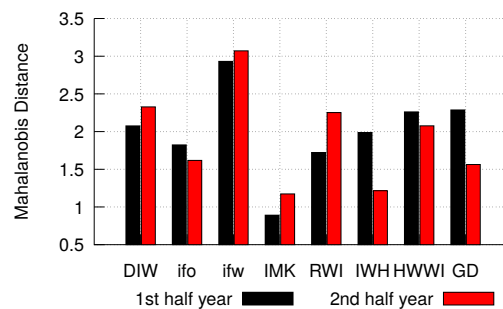
(d) Test der Richtungsprognose

NOTE: Für Erklärungen siehe Abb. 1.

Abbildung 12: Zusammenfassung der Prognoseevaluationsergebnisse für die *Arbeitslosenquote*. Berücksichtigt werden alle Prognosen getätigt im zweiten Halbjahr des aktuellen Kalenderjahres für das kommende Kalenderjahr. Periode: 2005–2014.



(a) Basierend auf 10-jähriger Kovarianz



(b) Basierend auf 20-jähriger Kovarianz

NOTE: Der berücksichtigte Vektor an Variablen umfasst: Wachstum des realen BIP, Wachstum der privaten sowie öffentlichen Konsumausgaben, Wachstum der Ausrüstungs-, Bau- und sonstigen Investitionen, Wachstum der Ex- und Importe, der Arbeitslosenquote sowie der Preisinflation basierend auf dem BIP- bzw. VPI-Deflator.

Abbildung 13: Multivariate Mahalanobis-Distanz. Berücksichtigt werden alle Prognosen getätigt im im ersten bzw. zweiten Halbjahr des aktuellen Kalenderjahres für das kommende Kalenderjahr. Periode: 2005–2014.

Tabellen

| Institut | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |
|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| DIW | 6 | 4 | 3 | 8 | 8 | 8 | 1 | 2 | 4 | 3 |
| IFO | 1 | 2 | 5 | 2 | 2 | 3 | 6 | 4 | 3 | 2 |
| IFW | 4 | 8 | 6 | 5 | 3 | 6 | 8 | 8 | 8 | 8 |
| IMK | 8 | 1 | 8 | 1 | 7 | 2 | 3 | 1 | 1 | 1 |
| RWI | 3 | 7 | 7 | 7 | 4 | 4 | 2 | 5 | 2 | 5 |
| IWH | 2 | 3 | 4 | 3 | 5 | 1 | 4 | 6 | 5 | 4 |
| HWWI | 7 | 6 | 1 | 6 | 1 | 7 | 7 | 3 | 7 | 7 |
| GD | 5 | 5 | 2 | 4 | 6 | 5 | 5 | 7 | 6 | 6 |

NOTE: Die Spalten bezeichnen das Kalenderjahr der getätigten Prognose für das kommende Kalenderjahr, z.B. 2005 bezieht sich auf die Prognose aus 2005 für das Jahr 2006. Der Vektor der berücksichtigten Variablen umfasst: BIP, priv. Konsum, öffentl. Konsum, Ausrüstungs-, Bau- und sonst. Anlageinvestitionen, Ausfuhren, Einfuhren (alle in realen Größen und Wachstumsraten), Inflation (BIP), Inflation (VPI) sowie die Arbeitslosenquote. Die zugrundeliegende Kovarianzmatrix basiert auf den realisierten Werten der letzten 20 Jahre.

Tabelle 8: Jährlicher Rang der Institute basierend auf der robusten Mahalanobis-Distanz für jedes Kalenderjahr.

| Institut | Durchschn.Rang | Std.-Fehler |
|----------|----------------|-------------|
| DIW | 4.7 | 2.5 |
| IFO | 3.0 | 1.5 |
| IFW | 6.4 | 1.8 |
| IMK | 3.3 | 2.9 |
| RWI | 4.6 | 1.9 |
| IWH | 3.7 | 1.4 |
| HWWI | 5.2 | 2.4 |
| GD | 5.1 | 1.3 |

NOTE: Die 2 stärksten Ausreisser wurden bei der Berechnung der Kovarianz zuvor eliminiert.

Tabelle 9: Durchschnittliche Ränge und Standardabweichungen der Platzierungen aus Tabelle 8. Basierend auf einer pseudo-robusten Kovarianzschätzung.

| Teststat. | Prob. |
|------------------|--------------|
| 14.733 | 0.040 |

NOTE: Die 2 staerksten Ausreisser wurden bei der Berechnung der Kovarianz zuvor eliminiert.

Tabelle 10: Friedman Test auf keine systematischen Differenzen zwischen allen Instituten in den Prognoserängen gem. der Mahalanobis-Distanz. Periode: 2005–2014.

| Institut | DIW | IFO | IFW | IMK | RWI | IWH | HWWI |
|-----------------|------------|------------|------------|------------|------------|------------|-------------|
| DIW | | | | | | | |
| IFO | 0.206 | | | | | | |
| IFW | 0.527 | 0.002 | | | | | |
| IMK | 0.206 | 0.206 | 0.206 | | | | |
| RWI | 1.000 | 0.058 | 0.206 | 0.527 | | | |
| IWH | 1.000 | 0.206 | 0.011 | 0.527 | 0.527 | | |
| HWWI | 0.527 | 0.206 | 0.206 | 0.206 | 1.000 | 0.206 | |
| GD | 1.000 | 0.058 | 0.058 | 0.206 | 0.206 | 0.011 | 0.206 |

NOTE: Die 2 staerksten Ausreisser wurden bei der Berechnung der Kovarianz zuvor eliminiert.

Tabelle 11: Ablehnungswahrscheinlichkeiten des Friedman Test auf keine systematischen Differenzen in den Prognoserängen gem. der Mahalanobis-Distanz zwischen jeweils zwei Instituten. Periode: 2005–2014.

Impressum

Publisher: Hans-Böckler-Stiftung, Hans-Böckler-Str. 39, 40476 Düsseldorf, Germany
Phone: +49-211-7778-331, IMK@boeckler.de, <http://www.imk-boeckler.de>

IMK Study is an online publication series available at:
http://www.boeckler.de/imk_5023.htm

ISSN: 1861-2180

The views expressed in this paper do not necessarily reflect those of the IMK or the Hans-Böckler-Foundation.

All rights reserved. Reproduction for educational and non-commercial purposes is permitted provided that the source is acknowledged.