

## Spezialthema: Analyse der Prognosequalität für die Schweiz

### Executive Summary

Die vorliegende Studie analysiert die Qualität der Prognosen für das reale BIP-Wachstum der Schweizer Volkswirtschaft. Die Prognosedaten von sieben unterschiedlichen Instituten, darunter die BIP-Prognosen der Expertengruppe des Bundes, über eine Zeitperiode von 1985-2012 fliessen in die Studie ein. Die Qualität der Prognosen wird anhand von statistischen Messgrössen sowie der Erfüllung der Optimalitätseigenschaft von unverzerrten Prognosen überprüft. Wenig überraschend, weisen die Ergebnisse auf eine zunehmende Treffsicherheit der Prognosen mit kürzerem Prognosehorizont hin. Des Weiteren zeigt die Analyse, dass sich die Qualität der Prognosen der Expertengruppe des Bundes nicht signifikant von jener der übrigen Institute unterscheidet. Generell bestätigt sich, dass Konjunkturprognosen, die über einen Horizont von vier Quartalen (bzw. ein Jahr) hinausreichen, insbesondere diejenigen mit einem Prognosehorizont von sieben und mehr Quartalen, nur eine begrenzte Treffsicherheit besitzen. Im Weiteren ist die Eigenschaft der Unverzerrtheit der Prognosen aller Institute - unabhängig des Prognosehorizonts - erfüllt.

### Einleitung

Obwohl die Treffsicherheit von Konjunkturprognosen oftmals den Erwartungen der Anwender nicht gerecht wird, dürfte ein Verzicht auf Prognose keine sinnvolle Option darstellen. Die Wichtigkeit von Konjunkturprognosen ist in der wirtschaftspolitischen Praxis weitgehend unbestritten. Namentlich in den Bereichen der Finanzpolitik, die gesamtwirtschaftliche Prognosen für die Planung der öffentlichen Einnahmen und Ausgaben, sowie der Geldpolitik im Rahmen der Inflationsbeurteilung spielen Einschätzungen über die zukünftige Wirtschaftsentwicklung eine zentrale Rolle. Auch im Bereich der Arbeitslosenversicherung werden für die Budgetplanung Prognosen für die künftige Entwicklung der Arbeitslosenzahlen benötigt, welche vom Konjunkturausblick abhängen. Für wirtschaftspolitische Entscheidungen bilden dabei sowohl das Gesamtbild und die zugrundeliegenden Einschätzungen der Ökonomen als auch eine spezifische Zahl (Punktpgnose) zentrale Grundlagen.

Auf der anderen Seite geniessen Konjunkturprognosen in der Öffentlichkeit, insbesondere in den Medien, keinen besonders guten Ruf. Dass sich die quantitativen Prognosen im Nachhinein oft als nicht zutreffend erweisen, wird besonders bemängelt. Diese Kritik hat sich in den letzten Jahren seit der Finanzkrise weiter verstärkt. Ob die bisherigen Prognosemodelle die Komplexität und verstärkten Verflechtungen im Wirtschaftssystem - etwa zwischen Banken, Finanzmärkten und der Realwirtschaft - genug erfassen können, wird zunehmend hinterfragt.

Prognoseunsicherheit und Prognosefehler lassen sich jedoch in einer Welt voller Ungewissheiten per Definition nicht vermeiden. Stochastisch unabhängige Ereignisse prägen das Wirtschaftsgeschehen in vielen Ebenen und in allen Ländern, ob gross oder klein. Kein Prognosemodell, so raffiniert es auch sein mag, wird jemals den Anspruch an Genauigkeit erfüllen können, um unerwartete, schockartige Ereignisse vorherzusagen zu können. Prognosefehler gehören demzufolge zum Alltag der Prognostiker.

### Annahmen sind zentral

Eine Aussage über die Zukunft hängt immer von a priori gewählten „exogenen Annahmen“ ab. Diese Annahmen erlauben, von fixierten Eckwerten auszugehen. Für die Schweiz als kleine offene Volkswirtschaft sind Annahmen über das zukünftige Wachstum der Eurozone, der USA sowie von gewissen Schwellenländern zentral, um Aussagen und Prognosen für die Schweizer Wirtschaftsentwicklung zu machen. Diese exogenen Annahmen, welche sich aus Informationen der Gegenwart zusammensetzen, können sich im Nachhinein als ungenau oder gar unzutreffend erweisen. Ursachen für Prognosefehler sind vielfältig und können sich zu einem bestimmten Prognosefehler aufaddieren (Hendry und Clemens, 2001). Das Bewusstsein jedoch, dass Prognosefehler sich kaum vermeiden lassen, sollte selbstverständlich die Prognostiker nicht entmutigen, die Qualität Ihrer Aussage im Nachhinein zu prüfen und daraus mögliche Schlüsse zu ziehen, um einige (vielleicht nur bescheidene) Fortschritte zu erzielen. Die ex-post Analyse und Darstellung der Prognosequalität sollte daher als wichtiger Bestandteil der gesamten Prognoseübung wahrgenommen werden.

### Ziele der Studie

In diesem Sinne werden in dieser Studie die Prognosen zu Grössen der Volkswirtschaftlichen Gesamtrechnung (VGR) für die Schweiz analysiert, um deren Güte und Zuverlässigkeit zu beurteilen. Dazu werden Prognosedaten von sieben Instituten, welche Prognosen für die Schweizer Wirtschaft erstellen, auf deren Genauigkeit, sowie

auf das Vorhandensein von systematischen Verzerrungen überprüft. In erster Linie fokussieren wir uns auf die Prognosen für das Wachstum des Bruttoinlandproduktes (BIP) zu konstanten Preisen. Wir analysieren die Jahresprognosen für das BIP-Wachstum des laufenden sowie für das darauffolgende Jahr für den Zeitraum 1985-2012. Da jeweils für ein bestimmtes Jahr mehrere Prognosen zu unterschiedlichen Zeitpunkten zur Verfügung stehen, werden unterschiedliche Prognosehorizonte in die Analyse einbezogen.

Ziel dieser Studie ist nicht eine Rangfolge der Zuverlässigkeit der Prognosen der Institute zu ermitteln, sondern die Treffsicherheit der Prognose der Expertengruppe des Bundes in einem relativen Kontext zu betrachten. Diese Evaluation dient dazu, gegebenenfalls auf Differenzen zwischen der Prognose der Expertengruppe des Bundes und den anderen Instituten hinzuweisen. Dies wäre z.B. ein Indiz für eine notwendige Änderung/Anpassung des vom Bund angewendeten Prognoseprozesses. Nicht überraschend deuten unsere Ergebnisse darauf hin, dass die Genauigkeit der Prognosen mit kürzeren Prognosehorizonten zunimmt. Der Zeitpunkt der Veröffentlichung der Prognose ist somit entscheidend. Je später die Prognose veröffentlicht wird (und je kürzer der Zeithorizont der Prognose), desto mehr Informationen stehen zur Verfügung, welche in die Modelle einfließen und desto mehr kann eine Prognose eine gewisse (relative) Genauigkeit erreichen. Umgekehrt besitzen Prognosen über sieben oder mehr Quartale eine bescheidene Aussagekraft. Die Ergebnisse weisen darauf hin, dass die Unterschiede zwischen den Instituten marginal sind und die Qualität der Prognosen sich nicht signifikant voneinander unterscheiden.

## Literatur

Eine umfangreiche Forschungsliteratur<sup>8</sup> widmet sich der Analyse von Konjunkturprognosen. Die Anzahl der Studien, welche den Fokus nur auf Prognosen für die Schweizer Volkswirtschaft setzen, fällt jedoch bescheiden aus.

Professor Aurelio Mattei von der Universität Lausanne hat regelmässig Studien über die Qualität der Prognosen für die Schweizer Volkswirtschaft von unterschiedlichen Instituten erstellt und seine eigene Prognose in die Analyse integriert. Neben der BIP Prognose wie auch diejenige der verwendungsseitigen Komponenten wurden die Inflations- und die Arbeitslosenrate ebenfalls miteinbezogen. Seine letztveröffentlichte Studie (Mattei, 2007) stützt sich auf die Prognosen über eine Zeitperiode von 1977-2006, wobei einzig die Herbstprognosen für das darauffolgende Jahr in die Analyse einbezogen wurden. In seiner Schlussfolgerung erwähnt er die abgenommenen Prognosefehler in den letzten Jahren (2003-2006). Diese begründet er mit der relativ stabilen Wirtschaftsentwicklung während dieser Zeit.

Die Studie von Ruoss und Savioz (2002) untersucht die Qualität der BIP-Prognosen von 14 ausgewählten Instituten, welche zwischen 1981 und 2000 publiziert wurden. Die Daten unterschiedlicher Prognosehorizonte werden integriert, wodurch der Zeitpunkt der Publikation mitberücksichtigt wird. Die Ergebnisse zeigen, dass die Prognosefehler mit kürzerem Prognosehorizont abnehmen. Mittelfristige Prognosen jedoch im Vergleich zu einer naiven Prognose eine geringe Aussagekraft aufweisen.

Noormamode (2003) untersucht die Prognosen der Institute BAK Basel Economics AG (BAK), der Konjunkturforschungsstelle ETH Zürich (KOF) und der Expertengruppe des Bundes über die Zeitperiode von 1981-2000. Die Prognosegenauigkeit wird anhand von vier Hypothesen überprüft. In den Ergebnissen der Studie wird kein Institut bestimmt, welches klar überlegen ist und durchgehend bessere Prognosen publiziert. Neben den BIP-Prognosen werden die Prognosen der einzelnen verwendungsseitigen Komponenten der VGR ebenfalls analysiert.

<sup>8</sup> Siehe Dovern und Weisser (2011), Ager, Kappler und Osterloh (2009), Osterloh (2006), Gultekin und Lahiri (2007), Koutsogeorgopoulou (2000), Spiwox, Scheier, Hein (2011), Kappler (2005).

**Daten: BIP-Prognosen**

Die Daten der Institute, welche Prognosen für das Bruttoinlandsprodukt (BIP) zu konstanten Preisen der Schweizer Volkswirtschaft erstellen, fliessen in unsere Analyse ein. Namentlich sind dies die Basler Arbeitsgruppe für Konjunkturforschung BAK Basel Economics AG (BAK), das Forschungsinstitut der ETH Zürich (KOF), die Crédit Suisse AG (CS)<sup>9</sup>, die OECD (Organisation for Economic Co-operation and Development), die Créa (Institut Créa de macroéconomie appliqué de l'Université de Lausanne), die Expertengruppe des Bundes (Prog.Bund) und die UBS AG<sup>10</sup>. Grundsätzlich fliessen alle Prognosen der Jahre 1985 bis 2012<sup>11</sup> in die Analyse ein, welche im Herbst (September bis November) publiziert wurden und sich auf das darauffolgende Jahr beziehen. Die Prognosen der OECD werden innerhalb des World Economic Outlook veröffentlicht, welche im Dezember publiziert werden.

Zudem sind die Prognosedaten der Institute BAK, KOF, der Expertengruppe des Bundes (Prog.Bund) und der OECD für weitere Prognosehorizonte erhältlich und werden in unsere Analyse integriert. Diese beinhalten die Prognosen, welche im Frühling (März bis Mai<sup>12</sup>) publiziert wurden und sich auf das aktuelle und auf das darauffolgende Jahr beziehen sowie die Herbstprognosen (September bis November), welche sich auf das aktuelle Jahr beziehen. Diese zusätzlichen Prognosehorizonte können für die anderen Institute (CS, UBS, CREA) aufgrund der begrenzten Verfügbarkeit der Daten nicht in diese Studie integriert werden.

Die Institute werden in den Ergebnissen unserer Berechnungen anonymisiert und mit der Bezeichnung Institut 1 bis 7 aufgeführt.

**Vier Prognosehorizonte**

In unserer Studie beziehen wir vier Prognosehorizonte ein. Dies sind die Frühlings- und Herbstprognosen für das laufende wie auch für das nächste Jahr, siehe Tabelle 5.

**Tabelle 5: Unterscheidung und Benennung der vier Prognosehorizonte**

	Jahr der Erstellung der Prognose	Jahr zu prognostizieren	Quartale, die "implizit" prognostiziert werden	Bezeichnung der Prognose im Text
Frühlingsprognosen	t	t+1	7	<i>PF(7)</i>
Herbstprognosen	t	t+1	5	<i>PH(5)</i>
Frühlingsprognosen	t	t	3	<i>PF(3)</i>
Herbstprognosen	t	t	1	<i>PH(1)</i>

Die Zahl in der Klammer bei der Benennung der Prognosehorizonte mit PF(7), PH(5), PF(3) oder PH(1) widerspiegelt die Anzahl der zu prognostizierenden Quartale bis zum Ende des Prognosejahres. Dementsprechend werden die Einschätzungen der Ökonomen im Beispiel von PH(5) fünf Quartale im Voraus bis zum Abschluss des zu prognostizierenden Jahres veröffentlicht. PH verweist auf eine Herbstprognose hin, PF bezieht sich auf eine Frühlingsprognose.

**Referenzreihen**

Für die Wahl der Referenzreihe zur Überprüfung der Prognosegenauigkeit folgen wir der Standardliteratur<sup>13</sup> und verwenden die erste provisorische Schätzung der jährlichen reale BIP-Wachstumsrate basierend auf der vierteljährlichen Volkswirtschaftlichen Gesamtrechnung (VGR), veröffentlicht durch das Staatssekretariat für Wirtschaft (SECO) (wird in der vorliegenden Studie einfachheitshalber VGR-BIP-1 bezeichnet). Diese ersten provisorischen Jahreswerte werden jeweils im März des darauffolgenden Jahres vom Staatssekretariat für Wirtschaft (SECO) publiziert<sup>14</sup>. Als zusätzlicher Vergleich werden in einem zweiten Schritt die ersten Werte der jährlichen VGR, welche jeweils im September des darauffolgenden Jahres vom Bundesamt für Statistik (BFS) publiziert werden, in die Analyse einbezogen (VGR-BIP-2).

<sup>9</sup> Vor 1997: Schweizerische Kreditanstalt

<sup>10</sup> Vor 1998: Schweizerische Bankgesellschaft (SBG)

<sup>11</sup> Die Berechnungen beziehen sich auf die Jahre 1985-2012. Für die grafische Betrachtung werden die Werte für 2013 einbezogen.

<sup>12</sup> Die Prognosen der OECD innerhalb des Economic Outlook werden jeweils im Juni bzw. Dezember publiziert.

<sup>13</sup> Vgl. Clements, Joutz und Stekler (2007), Ruoss und Savioz (2002).

<sup>14</sup> 1985-1986: publiziert durch Bundesamt für Statistik (BFS), 1987-1998: Bundesamt für Konjunkturfragen, ab 1999: Staatssekretariat für Wirtschaft (SECO).

**Tabelle 6: Unterscheidung und Benennung der Referenzreihen**

	Quelle	Bezeichnung der Referenzreihe im Text
jährliche BIP-Wachstumsrate zu konstanten Preisen des Vorjahres; basierend auf der <i>vierteljährlichen</i>	SECO	<i>VGR-BIP-1</i>
jährliche BIP-Wachstumsrate zu konstanten Preisen des Vorjahres; basierend auf den ersten Ergebnissen der <i>jährlichen</i> Volkswirtschaftlichen Gesamtrechnung	BFS	<i>VGR-BIP-2</i>

Die prognostizierten BIP-Wachstumsraten für unterschiedliche Zeithorizonte der Institute BAK, KOF und die Prog.Bund entnehmen wir aus Noormamode (2002). Die Hauptquelle der Daten der restlichen Institute für den Prognosehorizont PH(5) stellt die Datenbank von Mattei (2007) dar. Die fehlenden Daten bis zum heutigen Zeitpunkt wurden hauptsächlich über Internetseiten und Archive oder auch durch persönlichen Kontakt mit den Instituten ergänzt.<sup>15</sup>

Während die Prognosedaten für PH(3) und PH(5) über die gesamte Zeitperiode (1985-2012) berücksichtigt werden, fällt das Sample aufgrund der beschränkten Verfügbarkeit der Daten für PF(3) (1992-2012 (ohne 1997)) und PH(7) (1994-2012 (ohne 1997)) kürzer aus.

Die Daten der Referenzreihen (reale BIP-Wachstumsrate) - sowohl VGR-BIP-1 und VGR-BIP-2 – wurden aus den internen Archiven entnommen. Historische Daten bis in die 80er Jahre konnten mithilfe der Datenbank des OECD Economic Outlook<sup>16</sup> zusammengestellt werden. Tabelle 13 und Tabelle 14 im Appendix zeigen die deskriptive Statistik der Daten, welche in unsere Studie einfließen.

#### Grafische Darstellung

In Abbildung 33 bis Abbildung 34 sind die Prognosedaten gegliedert nach Prognosehorizonten dargestellt. Dabei entsprechen in den Grafiken die Y-Achse der jeweiligen Prognose und die X-Achse dem ersten verfügbaren Jahreswert (VGR-BIP-1). Fällt der Punkt auf die 45 Grad Linie, stimmen die Prognosen und der erste SECO-Jahreswert genau überein.

Je kürzer der Prognosehorizont, desto genauer die Prognose

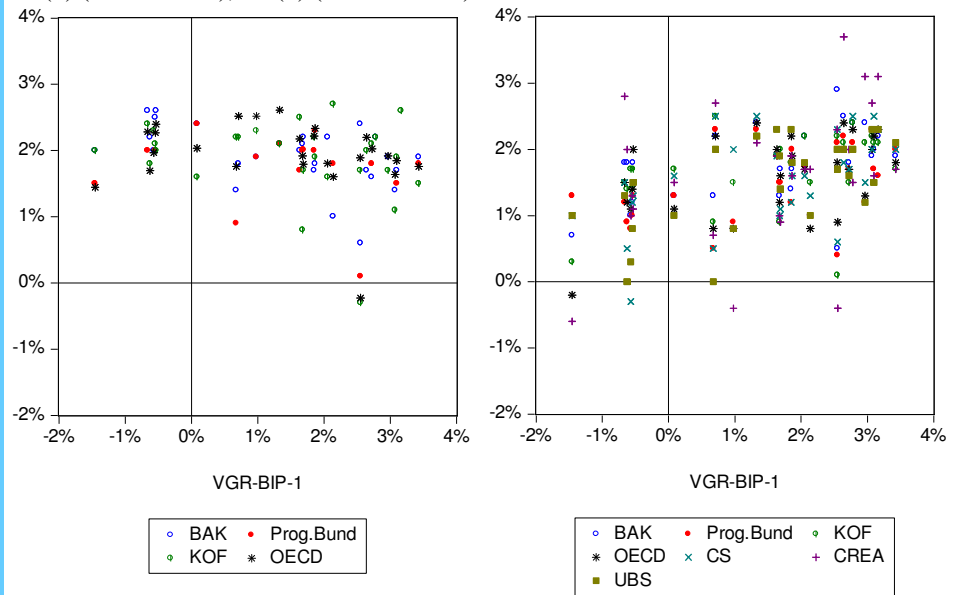
Die Grafiken verdeutlichen eine abnehmende Differenz zwischen den Prognosen und der ersten BIP-Jahresschätzung des SECO mit kürzerem Prognosehorizont. Dies wird durch das „Näherrücken“ an die 45 Grad Linie verdeutlicht. Mit kürzerem Prognosehorizont nimmt somit die Genauigkeit der Prognose sichtbar zu. Ausserdem lässt sich feststellen, dass wenige Prognosen über einen mittelfristigen Horizont von fünf oder sieben Quartalen mit einem negativen Vorzeichen (d.h. eine Rezession) versehen waren (siehe Abbildung 33). Dies deutet darauf hin, dass Rezessionen über einen mittelfristigen Horizont von sieben Quartalen kaum prognostiziert wurden.

Die BIP-Prognosen im Frühling, welche sich auf das nächste Jahr beziehen (PH(7)) liegen grösstenteils zwischen 1.5% und 2.5% (ca. mit einem Anteil von 80%), was im Bereich des langfristigen Durchschnitts der BIP-Wachstumsraten ist (1.8%). Allerdings unterliegen die effektiven BIP-Jahreswerte stärkeren Schwankungen und liegen somit oftmals ausserhalb dieser Bandbreite. Dass die BIP-Wachstumsrate längerfristig zurück zum Trendwachstumspfad (langfristiger Durchschnitt) findet, ist somit weit verbreitet unter den Prognostikern bzw. in den Prognosemodellen. Eine Untersuchung auf Verzerrung – ständiges Überschätzen oder Unterschätzen des Wachstums – wird daher später durchgeführt.

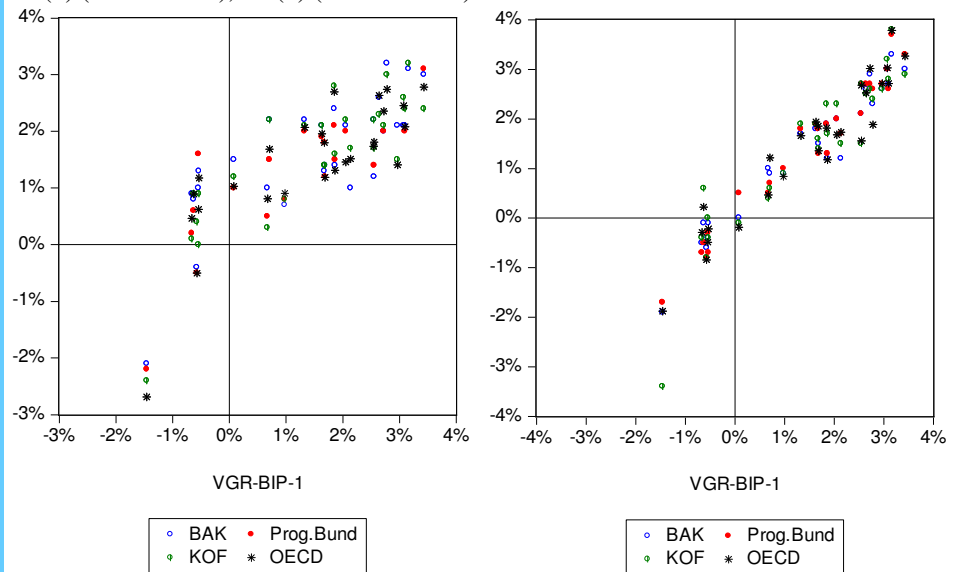
<sup>15</sup> Hiermit möchten wir Herrn Mattei ganz herzlich danken für die zur Verfügungsstellung seiner Datenbank. Weiter, möchten wir uns bei Herrn Alexis Körber (BAK Basel Economics AG) wie auch Herrn Claude Maurer (Credit Suisse AG) für die Zustellung der Daten bedanken.

<sup>16</sup> Die historischen Daten vom OECD Economic Outlook sind verfügbar unter: <http://stats.oecd.org/>

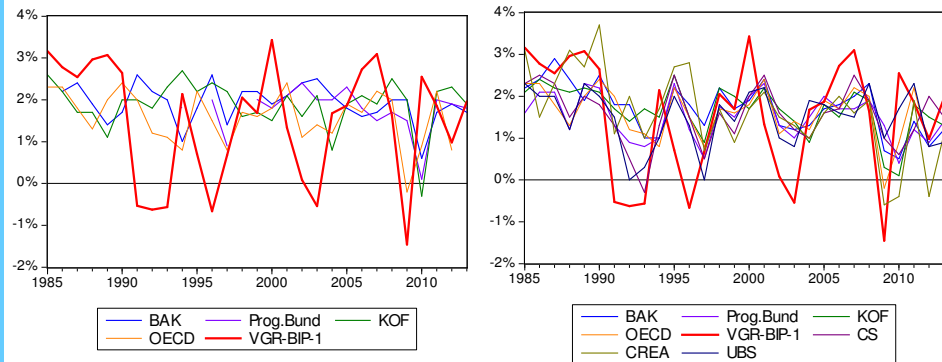
**Abbildung 33: Prognosen (y-Achse) vs. VGR-BIP-1 (x-Achse), PF(7) (linke Grafik), PH(5) (rechte Grafik)**



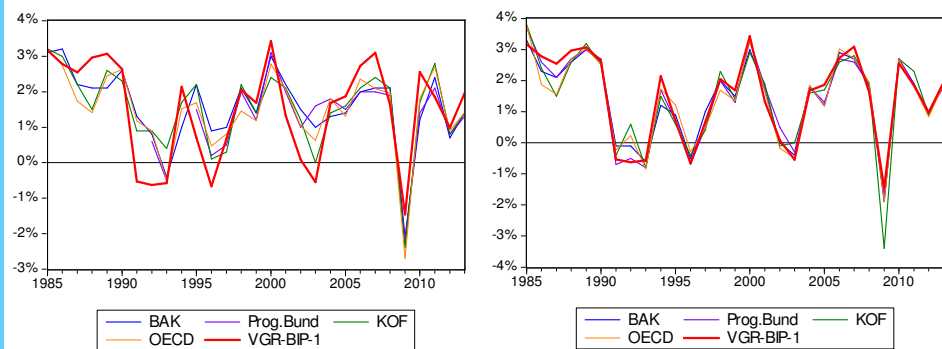
**Abbildung 34: Prognosen (y-Achse) vs. VGR-BIP-1 (x-Achse), PF(3) (linke Grafik), PH(1) (rechte Grafik)**



**Abbildung 35: Prognosen vs. VGR-BIP-1**  
PF(7) (linke Grafik), PH(5) (rechte Grafik)



**Abbildung 36: Prognosen vs. VGR-BIP-1**  
PF(3) (linke Grafik), PH(1) (rechte Grafik)



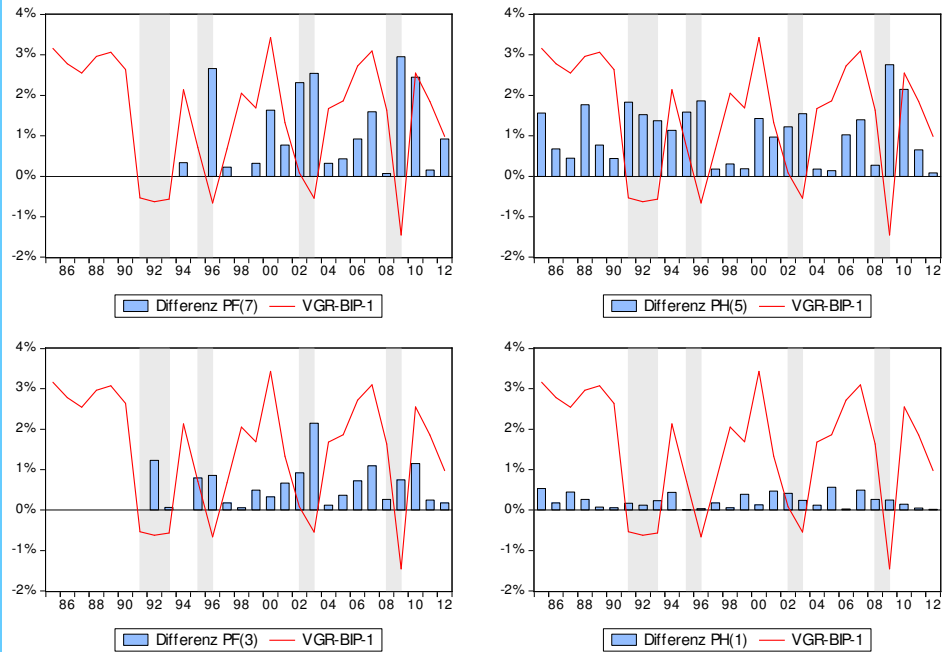
In Abbildung 35 und Abbildung 36 sind die jeweiligen Prognosen und die realisierten Werte (VGR-BIP-1) im Zeitverlauf gegliedert nach Prognosehorizonten dargestellt. Die Abnahme der Differenz zwischen den Prognosen und dem realisierten Wert und eine vermehrte Übereinstimmung zwischen den Prognosen und den BIP-Jahreswerten mit kürzeren Zeithorizonten wird in diesen Abbildungen deutlich. Tiefe Rezessionen, wie die Rezession Anfangs der 90er Jahre oder die Finanzkrise 2008/2009, wurden erst mit kürzeren Prognosehorizonten erkannt. Weiter ist zu erkennen, dass die Nachholeffekte in 2004 oder in 2010 in den Prognosehorizonten PF(7) und PH(5) teilweise unterschätzt wurden.

Abbildung 37 stellt die absoluten Prognosefehler (blaue Balken) der Expertengruppe des Bundes im zeitlichen Verlauf für die einzelnen Prognosehorizonte dar. Zusätzlich sind Jahre, in denen sich die Schweizer Wirtschaft in einer Rezession befand (VGR-BIP-1 ist negativ) grau schraffiert.<sup>17</sup> Diese grau markierten Flächen deuten auf eine Rezession und somit auf Phasen unsicherer Konjunkturentwicklung hin.

<sup>17</sup> In der gängigen Literatur wird der Begriff Rezession verwendet, falls zwei darauffolgende Quartale ein negatives BIP-Wachstum verzeichnen, siehe Meichle et al. (2011). Wir stützen uns auf die Jahresdaten und schraffieren diejenigen Jahre mit negativen BIP-Jahreswachstumsraten.



**Abbildung 37:** Differenz zwischen der Prognose der Expertengruppe des Bundes (Prog.Bund) und dem realisierten Wert (VGR-BIP-1), nach Prognosehorizonten<sup>18</sup>



Prognosefehler in der Finanz- und Wirtschaftskrise

Mit Blick auf die letzten 30 Jahre sind über einen mittelfristigen Prognosehorizont (PF(7), PH(5)) die Prognosefehler für die Jahre mit relativ stabiler Konjunkturentwicklung geringer im Vergleich zu Jahren mit unsicherer Wirtschaftsentwicklung (siehe grau schraffierte Fläche in Abbildung 37). Grundsätzlich wurden Rezessionen über einen mittelfristigen Prognosehorizont über ein Jahr hinaus (PF(7), PH(5)) selten prognostiziert (siehe Abbildung 35). Die Prognosefehler für die Rezessionsjahre in den 90er Jahren, 2003 sowie 2009 für diese Prognosehorizonte sowie umgekehrt für die darauffolgenden Jahre (unterschätzter Aufholeffekt) sind somit deutlich höher als der Durchschnitt aller Prognosefehler. Dies widerspiegelt die Erfahrung, dass Konjunkturprognosen zumeist nicht in der Lage sind, abrupte wirtschaftliche Wendepunkte - sowohl negativ in der Krise als auch positiv bei einsetzender Erholung - mit deutlichem Vorlauf vorherzusehen. Ausgeprägt war dies in der Finanz- und Wirtschaftskrise festzustellen, welche sich im Herbst 2008 in einem schlagartigen Einbruch vieler Wirtschaftsindikatoren manifestierte und erst anschliessend zu stark negativen Prognosekorrekturen für das folgende Jahr 2009 führte. Weiter lässt sich feststellen, dass je nach Tiefe der Rezession die Prognosefehler ausgeprägter sind, womit der absolute Prognosefehler für das Jahr 2009 für die Prognosehorizonte PF(7) wie auch PH(5) höher zu Buche schlägt als z.B. für 2003. Diese ausgeprägten Prognosefehler für Rezessionsjahre sowie für die Folgejahre sind jedoch mit kürzeren Zeithorizonten kaum sichtbar.

**Genauigkeit der Prognosen: Überblick**

Die bisher durchgeführte grafische Betrachtung der Daten gibt einige Hinweise bezüglich der Prognosequalität wider. Eine gründliche Analyse der Genauigkeit der Prognosen (forecast accuracy) erfordert jedoch die Berechnung statistischer Messgrössen, sogenannter Prognosefehlermasse. Wir berechnen in der vorliegenden Studie die in der Literatur gängigen Messgrössen wie den durchschnittlichen absoluten Prognosefehler (*MAE, Mean Absolute Error*) und die Wurzel des Durchschnittes der quadrierten Prognosefehlern (*RMSE, Root Mean Squared Error*). Der *MAE* berechnet die durchschnittliche absolute Abweichung der Prognosen von den tatsächlich realisierten BIP-Wachstumsraten. Diese Kennzahl gibt an, wie hoch der absolute Prognosefehler eines Instituts im historischen Durchschnitt für die jeweiligen Prognosehorizonte ist. Der *RMSE* ist ein weiteres sehr verbreitetes Fehlermass. Dieser gewichtet aufgrund der Bildung des quadrierten Fehlers grössere Abweichungen stärker, womit grobe Fehlprognosen der Institute ein höheres Gewicht erhalten.

<sup>18</sup> Die Differenz entspricht einer absoluten Differenz. Da die Prognosen sich auf das prozentuale Wachstum des BIP in konstanten Preisen beziehen, ist die Differenz in Prozentpunkten.

**Statistische Messgrössen:**  
**MAE**

Schliesslich wird der Theil's Ungleichheitskoeffizient berechnet. Dieser erlaubt uns die Prognosen der Institute mit einer naiven Prognose zu vergleichen.

Der Mean Absolute Error (*MAE*) entspricht der durchschnittlichen Differenz zwischen der Prognose und dem realisierten BIP. Der *MAE* wird für die einzelnen Institute wie auch die unterschiedlichen Prognosehorizonte berechnet. Als Referenzgrösse des realisierten BIP-Wachstums wird im ersten Schritt der Jahreswert VGR-BIP-1 herangezogen (siehe linker Abschnitt in Tabelle 1), in einem zweiten Schritt werden die ersten Werte des jährlichen VGR (VGR-BIP-2) als Referenzreihe gebraucht (siehe rechter Abschnitt in Tabelle 7).

$$MAE = \frac{1}{n} \sum_{t=1}^n |a_t - f_t| \quad (1)$$

Mit *n*: Anzahl Beobachtungen, *a<sub>t</sub>*: realisierte BIP-Wachstumsreihe (VGR-BIP-1 bzw. VGR-BIP-2), *f<sub>t</sub>*: Prognosen der jeweiligen Institute, *t*: das zu prognostizierende Jahr.

**Tabelle 7: MAE für unterschiedliche Prognosehorizonte, Zeitraum 1985-2012**  
 VGR-BIP-1 als Referenzwert (links), VGR-BIP-2 als Referenzwerte (rechts)

	PH(1)	PF(3)	PH(5)	PF(7)		PH(1)	PF(3)	PH(5)	PF(7)
Institut 1	0.35	0.68	1.05	1.33	Institut 1	0.37	0.70	1.03	1.31
Institut 2	0.27	0.79	1.05	1.28	Institut 2	0.27	0.78	1.09	1.26
Institut 3	0.23	0.63	1.05	1.21	Institut 3	0.35	0.69	1.08	1.28
Institut 4	0.41	0.70	1.08	1.34	Institut 4	0.41	0.76	1.12	1.41
Institut 5			1.03		Institut 5			1.06	
Institut 6			1.09		Institut 6			1.19	
Institut 7			0.95		Institut 7			0.96	
Durchschnitt	0.32	0.70	1.04	1.29	Durchschnitt	0.35	0.73	1.08	1.32

Grundsätzlich lässt sich feststellen, dass mit kürzerem Zeithorizont der durchschnittliche absolute Prognosefehler deutlich abnimmt. Im historischen Vergleich über die Zeitperiode 1985-2012 waren die Prognosen mit einem Prognosehorizont über ein Jahr hinaus im Durchschnitt 1%-Punkt vom realisierten Wert entfernt. Der durchschnittliche Fehler sinkt mit kürzerem Prognosehorizont deutlich. Herbstprognosen für das laufende Jahr waren im Durchschnitt um 0.3 bzw. 0.4 Prozentpunkte von den realisierten Werten entfernt.

Zwischen den Instituten lassen sich kleine Unterschiede der durchschnittlichen Prognosefehler feststellen. Im Beispiel von PH(1) weist Institut 3 den kleinsten MAE aus, falls die Prognosen mit den VGR-BIP-1-Werten verglichen werden. Im Falle von der VGR-BIP-2 als Referenzreihe zeigt Institut 2 den kleinsten MAE auf. Mit Blick auf Tabelle 7 lässt sich somit kein Institut mit dem kleinsten durchschnittlichen Fehler über alle Zeithorizonte identifizieren. Ob für die einzelnen Prognosehorizonte die Unterschiede zwischen den Instituten signifikant sind wird weiter unten in der vorliegenden Studie untersucht.

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (a_t - f_t)^2} \quad (2)$$

**RMSE**

Der *RMSE* (durchschnittlicher quadrierter Prognosefehler) ist eine weitere Messgrösse, um die Genauigkeit einer Prognose („forecast accuracy“) zu berechnen. Durch das Quadrieren der Prognosefehler erhalten grössere Abweichungen eine stärkere Gewichtung im Vergleich zum *MAE*. In den untenstehenden Tabellen wird der *RMSE* für verschiedene Prognosehorizonte berechnet. Analog zum *MAE* nimmt mit sinkendem Prognosehorizont der quadrierte Prognosefehler ab.

**Tabelle 8: RMSE für unterschiedliche Prognosehorizonte, Zeitraum 1985-2012**  
 VGR-BIP-1 als Referenzwert (links), VGR-BIP-1 als Referenzreihe (rechts)

	PH(1)	PF(3)	PH(5)	PF(7)		PH(1)	PF(3)	PH(5)	PF(7)
Institut 1	0.43	0.79	1.24	1.64	Institut 1	0.49	0.80	1.24	1.68
Institut 2	0.35	0.93	1.29	1.65	Institut 2	0.39	0.96	1.32	1.67
Institut 3	0.29	0.81	1.26	1.57	Institut 3	0.46	0.88	1.36	1.61
Institut 4	0.58	0.80	1.32	1.68	Institut 4	0.55	0.91	1.38	1.73
Institut 5			1.21		Institut 5			1.29	
Institut 6			1.41		Institut 6			1.44	
Institut 7			1.12		Institut 7			1.21	
Durchschnitt	0.41	0.83	1.26	1.63	Durchschnitt	0.47	0.89	1.32	1.67



Theil's U

Die bisher aufgeführten statistischen Masse ermöglichen, die Prognosefehler der Institute zu quantifizieren und miteinander zu vergleichen. Eine grundsätzliche Beurteilung, ob eine Prognoseleistung „gut“ oder „schlecht“ ist, kann basierend auf diesen Statistiken jedoch nur bedingt durchgeführt werden. Aus diesem Grund wird ein weiteres Mass (Theil's Ungleichheitskoeffizient) im folgenden Abschnitt berechnet. Der Theil's Ungleichheitskoeffizient (*Theil's U*) erlaubt die Prognosen der Institute mit einer sogenannten „naiven Prognose“ zu vergleichen. Als naive Prognosen können Ergebnisse eines vereinfachten Prognoseverfahrens berücksichtigt werden. Der *Theil's U* kann einen Hinweis geben, ob die Prognosen im Vergleich zu einer naiven Prognose hinreichend gut, oder klar verbesserungsfähig sind.

Der *Theil's U* berechnet sich aus der Division des *RMSE* der Prognosen der Institute durch den *RMSE* des naiven Prognoseverfahrens. Der kritische Wert beim *Theil's U* ist somit der Wert 1. Liegt der Index zwischen null und eins ( $0 < U < 1$ ), sind die durchgeführten Prognosen der Institute „genauer“ im Vergleich zu einer naiven Prognose. Mit *Theil's U*=1, entspricht der *RMSE* der naiven Prognose demjenigen der Institute. Demensprechend schneiden beide gleich (gut/schlecht) ab. Übersteigt der Index den Wert 1, ist der durchschnittliche quadrierte Prognosefehler der naiven Prognosen kleiner als derjenige des jeweiligen Instituts. Dies wäre ein Indiz, dass die Prognosen der Institute ein Verbesserungspotenzial hätten.

Naive Prognose 1:  
Trendwachstum

Ein Beispiel für ein naives Prognoseverfahren ist das sogenannte Trendwachstum, welches wir als Durchschnitt der BIP-Wachstumsraten (in unserer Studie die VGR-BIP-1) der letzten 10 Jahre definieren (*Theil's U* „naive Prognose 1“). Die Ergebnisse sind aus Tabelle 9 zu entnehmen.

**Tabelle 9: *Theil's U* „naive Prognose 1“: Durchschnitt der letzten 10 Jahre, Zeitraum 1985-2012<sup>19</sup>, VGR-BIP-1 als Referenzreihe (links), VGR-BIP-2 als Referenzreihe (rechts)**

	<i>PH(1)</i>	<i>PF(3)</i>	<i>PH(5)</i>	<i>PF(7)</i>		<i>PH(1)</i>	<i>PF(3)</i>	<i>PH(5)</i>	<i>PF(7)</i>
<b>Institut 1</b>	0.30	0.53	0.87	1.19	<b>Institut 1</b>	0.35	0.53	0.87	1.27
<b>Institut 2</b>	0.24	0.59	0.90	1.20	<b>Institut 2</b>	0.27	0.64	0.93	1.29
<b>Institut 3</b>	0.20	0.65	0.89	1.12	<b>Institut 3</b>	0.32	0.67	0.95	1.22
<b>Institut 4</b>	0.40	0.45	0.92	1.26	<b>Institut 4</b>	0.38	0.49	0.97	1.35
<b>Institut 5</b>			0.85		<b>Institut 5</b>			0.90	
<b>Institut 6</b>			0.99		<b>Institut 6</b>			1.01	
<b>Institut 7</b>			0.78		<b>Institut 7</b>			0.85	
<b>Durchschnitt</b>	0.29	0.55	0.89	1.19	<b>Durchschnitt</b>	0.33	0.59	0.93	1.28

Die Ergebnisse für den *Theil's U* in Tabelle 9, zeigen einen zunehmenden Wert des Indexes mit längerem Prognosehorizont. Für den Prognosehorizont *PF(7)* übersteigt der *Theil's U* den kritischen Wert von 1, sowohl im Vergleich zum realisierten Wert VGR-BIP-1 wie auch zu VGR-BIP-2. Der *RMSE* der Prognosen der Institute übersteigt somit den *RMSE* der naiven Prognose für *PF(7)*. Ein Trendwachstum über die letzten 10 Jahre scheint für den Prognosehorizont *PF(7)* mindestens gleich gut zu sein wie die Prognosen der Institute. Mit kürzeren Prognosehorizonten nimmt die relative Prognosegüte der Institute im Vergleich zur „naiven Prognose 1“ zu.

<sup>19</sup> Aufgrund der Durchschnittsbildung über die letzten 10 Jahre beinhaltet das eigentliche Sample die Zeitperiode 1995-2012.

Naive Prognose 2:  
ARIMA-Prognosen in  
Real-time

Bei kurzfristigen Prognosen für das laufende Jahr (PF(3), PH(1)) sind bereits BIP-Quartalswerte für das zu prognostizierende Jahr verfügbar. Diese stellen eine wichtige Informationsquelle für eine Prognose für das laufende Jahr dar. Aufgrund des Trendwachstums mittels Jahreswachstumsraten können diese in der „naiven Prognose 1“ nicht berücksichtigt werden. Daher liegt eine gewisse „Benachteiligung“ der „naiven Prognosen 1“ bei kurzfristigen Zeithorizonten vor. In einem zweiten Schritt werden aus diesem Grund die bereits vorhandenen Quartalsdaten für die „naive Prognose 2“ berücksichtigt. Hierzu wird eine einfache ARIMA Prognose als „naive Prognose 2“ modelliert und mit den Prognosen der Institute verglichen (*Theil's U* „naive Prognose 2“). Eine ARIMA (Autoregressive integrated moving average) Prognose stellt ein einfaches Prognoseverfahren dar, wobei anhand der vergangenen vierteljährlichen BIP-Werte die BIP-Reihe zu konstanten Preisen in die Zukunft fortgeschrieben wird. Die bereits veröffentlichten BIP-Quartalsdaten für das laufende Jahr können dadurch berücksichtigt werden. Die ARIMA-Fortschreibung (ARIMA-Extrapolation) ist ein mechanisches Prognoseverfahren, welches einzig auf der Dynamik der vergangenen BIP-Werte basiert. Exogene Annahmen werden dabei nicht berücksichtigt (z.B. Annahmen über Wechselkurse, Welt-BIP, Zinsen etc.).

Die Wahl des ARIMA Modells für die Fortschreibung (ARIMA-Extrapolation) wird basierend auf den statistischen Kriterien durchgeführt und fixiert. Auf die Wahl der ARIMA-Modells wird weiter nicht eingegangen.

Mithilfe eines Real-time Datensatzes (Indergand und Leist, 2014) kann diese Übung in einem Real-time Kontext durchgeführt werden. Dies bedeutet, dass die zum Zeitpunkt der Prognose verfügbare BIP-Quartalsreihe (ohne nachträgliche Revisionen) in die ARIMA Prognose einbezogen werden können. Wir gehen somit von der gleichen Informationsbasis (betreffend BIP-Reihe zu konstanten Preisen) aus, welche den Prognostikern zum Zeitpunkt der Prognose zur Verfügung stand.

Aufgrund der begrenzten Verfügbarkeit der Daten im Real-time Datensatz werden in diesem Abschnitt die Prognosen über einen Zeithorizont von 2003-2012 betrachtet. Die Ergebnisse sind aus Tabelle 10 zu entnehmen. Die Berechnung des *Theil's U* wurde nur für die Institute 1 bis 4 durchgeführt.

**Tabelle 10: *Theil's U* „naive Prognose 2“: ARIMA-Extrapolation,**  
Zeitraum 2003-2012, VGR-BIP-1 als Referenzreihe (links), VGR-BIP-2 als Referenzreihe (rechts)

	PH(1)	PF(3)	PH(5)	PF(7)		PH(1)	PF(3)	PH(5)	PF(7)
Institut 1	0.81	0.65	0.51	0.79	Institut 1	0.71	0.58	0.57	0.76
Institut 2	0.76	0.68	0.66	0.76	Institut 2	0.71	0.68	0.69	0.75
Institut 3	0.69	0.62	0.77	0.75	Institut 3	0.78	0.62	0.78	0.73
Institut 4	1.66	0.61	0.69	0.86	Institut 4	1.24	0.58	0.71	0.83
Institut 5					Institut 5				
Institut 6					Institut 6				
Institut 7					Institut 7				
Durchschnitt	0.98	0.64	0.66	0.79	Durchschnitt	0.86	0.62	0.69	0.77

Der *Theil's U* liegt für den Grossteil der Institute über alle Prognosehorizonte um den Wert 0.7. Dies weist auf eine höhere Prognosegüte der Prognosen der Institute hin im Vergleich zur naiven Prognose 2. Nur im Falle von Institut 4 für den Prognosehorizont PH(1) übersteigt der *Theil's U* den kritischen Wert von 1.

Aufgrund der unterschiedlich berücksichtigten Zeitperiode sind Tabelle 9 und Tabelle 10 zwar nicht direkt miteinander vergleichbar, jedoch lassen sich einzelne Feststellungen daraus ableiten. Im Gegensatz zum *Theil's U* („naive Prognose 1“) nimmt die relative Genauigkeit der Prognosen der Institute im Vergleich zur „naiven Prognose 2“ mit kürzerem Prognosehorizont nicht zu, sondern bleibt praktisch konstant.

In der ARIMA-Extrapolation wird die relativ kurzfristige vergangene Dynamik fortgeschrieben. Im Gegensatz dazu wird in der „naiven Prognose 1“ das durchschnittliche BIP-Wachstum der letzten zehn Jahren berücksichtigt. Wir gehen einen Schritt weiter und vergleichen die naiven Prognosen miteinander. Aus den Ergebnissen kann man darauf schliessen, dass die relative Genauigkeit der naiven Prognose 2 für kurzfristige Prognosehorizonte (insbesondere für PH(1)) im Vergleich zu der naiven Prognose 1 höher ist. Im Gegensatz dazu weist die naive Prognose 1 über einen Prognosehorizont über ein Jahr hinaus einen kleineren RMSE auf. Ob die Differenzen zwischen den

Diebold Mariano Test  
Statistik

naiven Prognosen 1 und 2 signifikant sind und ob sich die Resultate aufgrund der unterschiedlichen Zeitperioden ergeben, wurde weiter nicht untersucht.

Die bisherigen Messstatistiken geben einen ersten Hinweis über die Grösse und Unterschiede der Prognosefehler der einzelnen Institute. Die Frage, ob Prognosefehler von gewissen Instituten oder der naiven Prognosen *signifikant* kleiner (grösser) sind im Vergleich zu den Prognosen der Expertengruppe des Bundes, wird im folgenden Abschnitt anhand von paarweisen Signifikanztests (Diebold und Mariano, 1995)<sup>20</sup> untersucht. Der Diebold Mariano Test beurteilt, ob die Unterschiede der Prognosefehler zwischen den Instituten signifikant sind und ob ein Institut somit bessere Prognosen erstellt. Wir vergleichen in paarweisen Signifikanztests die Prognosefehler der Expertengruppe des Bundes (Prog.Bund) mit denjenigen der übrigen Institute wie auch mit den naiven Prognosen für alle Prognosehorizonte mit der Referenzreihe VGR-BIP-1. Dabei wird die Differenz zwischen den (quadrierten) Prognosefehlern der Expertengruppe des Bundes (Prog.Bund) und dem jeweiligen Institut auf eine Konstante regressiert. Die Ergebnisse des *Diebold Mariano Test* (paarweiser Signifikanztest der quadrierten Prognosefehler) mit der VGR-BIP-1 als Referenzreihe sind in Tabelle 11 aufgelistet.

Die Koeffizienten in Tabelle 11 deuten darauf hin, dass die Nullhypothese (die quadrierten Prognosefehler der Institute im Vergleich zur Prognose der Expertengruppe sind gleich) auf einem 5% Signifikanzniveau für kein Institut und Prognosezeithorizont verworfen werden kann. Somit lassen sich keine signifikanten Differenzen zwischen der Prognose der Expertengruppe und derjenigen der übrigen Institute feststellen.

Wird der *Diebold Mariano Test* für den Vergleich mit der „naiven Prognose 1“ angewendet, kann die Nullhypothese (keine Unterschiede zwischen den Prognosefehlern) auf dem 5%-Signifikanzniveau für die Prognosehorizonte PH(1) und PF(3) verworfen werden. Die Koeffizienten weisen sowohl für PH(1) wie auch für PF(3) ein negatives Vorzeichen auf, und weisen auf einen kleineren durchschnittlichen quadrierten Prognosefehler der Prog.Bund hin. Diese Ergebnisse bestätigen, dass die Prognosen der Expertengruppe des Bundes (Prog.Bund) im Vergleich zu einer naiven Prognose (Durchschnitt der letzten 10 Jahre) im Falle von Prognosen für das laufende Jahr eine signifikant höhere Treffsicherheit aufweisen.

Im Falle der Frühlingsprognosen für das nächste Jahr (PF(7)) bestätigt der Koeffizient die bisherigen Ergebnisse aus den Berechnungen des *Theil's U* („naive Prognose 1“), wo der kritische Wert von 1 überschritten wird. Die quadrierten Prognosefehler der Prog.Bund sind im Durchschnitt zwar grösser als die Fehler der naiven Prognose 1, die Höhe der Standardfehler weist jedoch darauf hin, dass eine grosse Streuung vorherrscht und dass die naive Prognose 1 nicht systematisch besser ist.

Autoregressiver Charakter  
der Konjunktur

Beim paarweisen Signifikanztest mit der „naiven Prognose 2“ kann die Nullhypothese für keinen Prognosehorizont abgelehnt werden. Dass für kein Prognosehorizont eine bestimmte Signifikanz erreicht werden kann hängt sicherlich mit der geringen Anzahl Beobachtungen zusammen (10 Beobachtungen). Mit Blick auf die (negativen) Koeffizienten lässt sich eine abnehmende Differenz zwischen den Prognosen der Expertengruppe und der ARIMA-Extrapolation mit kürzerem Prognosehorizont feststellen (der absolute Wert des Koeffizienten wird kleiner). Daraus lässt sich schliessen, dass ARIMA-Extrapolationen bei einem kurzfristigen Prognosehorizont (insbesondere die Herbstprognose für das laufende Jahr) über den Zeithorizont von 2003-2012 praktisch gleich gut waren wie die Prognosen der Expertengruppe des Bundes. Dass eine ARIMA-Prognose kurzfristig ähnlich abschneiden kann man durch den „autoregressiven“ Charakter der Konjunktur erklärt werden. Die Zeitreihe des BIP weist eine gewisse „Autokorrelation“ aus. Dies bedeutet, dass die jüngste Vergangenheit im Konjunkturverlauf die gegenwärtige Konjunkturelle Entwicklung beeinflusst, insbesondere in relativ stabilen Konjunkturperioden.

<sup>20</sup> Die Nullhypothese des Tests von Diebold und Mariano (1995) ist, dass der Durchschnitt der quadrierten Prognosefehler zweier Institute (in unserer Studie im Vergleich zur Prognose von Prog.Bund) gleich ist, und somit die Treffsicherheit der Prognosen beider Institute gleich ist. Kann die Nullhypothese auf einem gewissen Signifikanzniveau verworfen werden, ist die Treffsicherheit eines Instituts signifikant „besser“.

Paarweise Signifikanztests

**Tabelle 11: Diebold-Mariano paarweise Tests** der Differenz der quadrierten Prognosefehler der Prognose der Expertengruppe des Bundes (Prog.Bund) vs. übrige Institute, Zeitraum 1985/1992/1994-2012 bzw. 2003-2012, VGR-BIP-1 als Referenzreihe

	<i>PH(1)</i>	<i>PF(3)</i>	<i>PH(5)</i>	<i>PF(7)</i>
Institut 1	-0.106 (0.053)	0.014 (0.179)	0.067 (0.349)	-0.342 (0.215)
Institut 2	-0.040 (0.031)	-0.195 (0.171)	-0.065 (0.289)	-0.511 (0.293)
Institut 3	-0.249 (0.122)	-0.012 (0.250)	-0.140 (0.300)	-0.538 (0.398)
Institut 4			0.137 (0.188)	
Institut 5			-0.391 (0.426)	
Institut 6			0.347 (0.182)	
naive Prognose 1	-1.96*** (0.461)	-1.39* (0.544)	-0.434 (0.336)	0.456 (0.585)
naive Prognose 2	-0.090 (0.039)	-0.321 (0.458)	-1.157 (0.681)	-1.260 (1.170)

Bemerkungen: Standardfehler in Klammern. \*\*\*/\*\*/\* Signifikanzniveau auf dem 5%/2%/1% - Niveau. Robuste Standardfehler nach Newey West (1987).

Synthese der ersten Ergebnisse

Die Ergebnisse zur Berechnung der Genauigkeit der Prognosen legen dar, dass die Höhe der Prognosefehler mit kürzeren Prognosehorizonten abnehmen. Diese Ergebnisse sind nicht überraschend und bestätigen die wissenschaftliche Literatur. Mit kürzerem Prognosehorizont steht erstens ein grösserer Pool von Daten und Informationen zur Verfügung, welche in die Prognosemodelle einfließen und somit die Treffsicherheit der Prognosen erhöhen. Zweitens ist die Wahrscheinlichkeit für das Eintreten von unvorhergesehenen Schocks (z.B. eine unerwartete, deutliche Veränderung der Weltkonjunktur) mit kürzeren Zeithorizonten beschränkter. Anhand der statistischen Messgrößen lässt sich kein Institut identifizieren, welches systematisch bessere Prognosen als die anderen liefern kann. Die Prognosefehler der Expertengruppe und der übrigen Institute unterscheiden sich nicht signifikant voneinander.

Der Vergleich mit naiven Prognosen erlaubt uns die Prognosen der Institute den Ergebnissen eines vereinfachten Prognoseverfahrens gegenüberzustellen. Dabei zeigt sich, dass die Aussagekraft der mittelfristigen Prognosen der Institute, d.h. über ein Jahr hinaus, verglichen mit einer naiven Prognose eher beschränkt ist. Die mittelfristigen Prognosen der Institute hatten im historischen Vergleich von 1985-2012 eine ähnliche Treffsicherheit wie eine naive Prognose anhand des Durchschnittswachstums über die letzten 10 Jahre. Eine weitere naive Prognose durch eine ARIMA-Extrapolation der bisherigen BIP-Reihe scheint für kurzfristige Prognosen für das laufende Jahr eine ähnliche Treffsicherheit zu erreichen wie die Prognosen der Expertengruppe des Bundes. Dies deutet darauf hin, dass der Einbezug der bereits publizierten BIP-Quartalswerte für das laufende Jahr eine relevante Information für die Wirtschaftsprognosen darstellt.

Unsere Resultate sind konsistent mit den Ergebnissen aus der bisherigen Literatur und für weitere Länder ebenfalls zutreffend. Laut Ruoss und Savioz (2002) widerspiegeln mittelfristige Prognosen über ein Jahr hinaus nicht die Bewegung des Konjunkturzyklus, sondern geben bestenfalls das angenommene Trendwachstum wider. Ähnliche Ergebnisse liefert auch die Studie von Gultekin und Lahiri (2007), welche die Prognosegenauigkeit von *Consensus Economics Inc.* Prognosen für 18 Länder untersucht. Sie kommen zum Schluss, dass Prognosen über einen Horizont von 18 Monaten (Erstellung der Prognose 18 Monate vor Abschluss des Jahres  $t$  keinen Mehrwert, welches über das Trendwachstum hinausgeht, aufweisen.

**Sind Prognosen systematisch verzerrt?**

Ein weiteres Mass um die Qualität der Prognosen zu analysieren, ist die Erfüllung der Eigenschaft der Unverzerrtheit. Eine Verzerrung der Prognosen liegt vor, wenn Institute systematisch zu hoch oder zu tief prognostizieren. Falls eine Verzerrung vorliegt, sind die Prognosen nicht optimal und könnten basierend der Verzerrung aus der Vergangenheit verbessert werden. Die Darstellungen in Abbildung 33 bis Abbildung 34 können einen ersten Hinweis auf eine allfällige Verzerrung liefern.

Eine systematische Verzerrung von Prognosen wird in der Literatur (Clements, Joutz und Stekler, 2007) anhand der erweiterten Mincer-Zarnowitz-Regression überprüft, wobei die Prognosefehler der jeweiligen Institute auf eine Konstante  $\tau$  und einen Fehlerterm  $v_{th}$  regressiert werden. Der Prognosefehler wird definiert als:

$$e_{th} = f_{th} - a_t \tag{3}$$

Mit  $e_{th}$  als Prognosefehler,  $a_t$  der realisierten BIP-Wachstumsrate (VGR-BIP-1) und  $f_{th}$  entspricht der Prognosen der jeweiligen Institute.

Der Prognosefehler wird auf eine Konstante  $\tau$  und einen Fehlerterm  $v_{th}$  regressiert werden.  $h$  kennzeichnet den jeweiligen Planungshorizont.

$$e_{th} = \tau + v_{th} \tag{4}$$

$\tau$  erfasst die Summe von möglichen Verzerrungen und  $v_{th}$  als Fehlerterm fasst die Shocks (Überraschungen) welche zwischen  $t-h$  und  $t$  eingetreten sind und somit in der Prognose nicht berücksichtigt werden konnten. Die OLS Regression untersucht die Nullhypothese  $\tau = 0$ . Kann die Nullhypothese verworfen werden, liegt eine systematische Verzerrung der Prognosen vor und die Prognosen der Institute sind nicht optimal.

Ein t-Test wird als Teststatistik angewandt, wobei die Eigenschaft der normalverteilten Prognosefehler erfüllt sein muss.<sup>21</sup> Des Weiteren, dürfen die Fehlerterme nicht autokorreliert sein. Da die Prognosen mit einem Planungshorizont von PF(7) und PH(5) sich auf das nächste Jahr beziehen und somit die Prognose für das nächste Jahr (jeweils im Herbst) vor der eigentlichen Veröffentlichung der BIP-Schätzung für das aktuelle Jahr (jeweils im März) erfolgt, liegt eine Überlappung vor. In diesen Fällen kann Autokorrelation erster Ordnung des Fehlerterms  $v_{th}$  vorliegen. Ein externer Schock, welcher während dieser Überlappungsperiode auftritt erscheint sowohl im Prognosefehler für den Zeitpunkt  $t$  wie auch für Zeitpunkt  $t-1$ . In diesem Sinne werden robuste Standardfehler nach Newey und West (1987) ausgewiesen.

**Tabelle 12: Test auf Unverzerrtheit, erweiterte Mincer-Zarnowitz-Regression**  
Zeitraum 1985-2012 bzw. 1992/1994-2012, VGR-BIP-1 als Referenzreihe

	<b>PF(7)</b>	<b>PH(5)</b>	<b>PF(3)</b>	<b>PH(1)</b>
Institut 1	0.571 (0.363)	0.132 (0.279)	0.078 (0.198)	-0.072 (0.081)
Institut 2	0.505 (0.384)	0.250 (0.272)	0.151 (0.196)	-0.063 (0.064)
Institut 3	0.400 (0.403)	0.020 (0.264)	0.128 (0.262)	-0.057 (0.054)
Institut 4	0.505 (0.376)	0.214 (0.257)	0.061 (0.183)	-0.086 (0.102)
Institut 5		0.171 (0.336)		
Institut 6		0.064 (0.254)		
Institut 7		0.025 (0.233)		

Bemerkungen: Standardfehler in Klammern.  
\*/\*/\*\* Signifikanzniveau auf dem 5%/2%/1% Niveau. (keine signifikanten Resultate in der Tabelle)  
Robuste Standardfehler nach Newey West (1987).

<sup>21</sup> Wurde anhand der Jarque-Bera Statistik überprüft. Die Prognosefehler sind mit Ausnahme einer einzigen Reihe normalverteilt. In dieser Reihe wurde die Normalverteilung aufgrund einer Beobachtung (Outlier) nicht erfüllt.



Die Koeffizienten in Tabelle 12 sind auf dem 5%-Niveau nicht signifikant und weisen darauf hin, dass im Rahmen unserer Studie keine signifikante Verzerrung der Prognosen der Institute vorliegt.<sup>22</sup> Interessant ist jedoch, dass die Konstante  $\tau$  (eine mögliche Verzerrung) für den Prognosehorizont PF(7) für alle Institute positiv ist (um 0.5) und mit kürzeren Prognosehorizonten abnimmt (für PH(1) sogar leicht negativ). Dies kann dadurch begründet werden, dass Rezessionen für die Zeitperiode 1994-2012 über einen mittelfristigen Horizont über ein Jahr hinaus selten prognostiziert wurden, und somit im Durchschnitt eine Überschätzung des BIP-Wachstums für diese Jahre vorliegt.

### Schlussfolgerungen

Je kürzer der Prognosehorizont, desto treffsicherer die Prognose

Basierend auf den Ergebnissen der vorliegenden Studie lassen sich einige Schlüsse über die Qualität der BIP-Prognosen für die Schweizer Volkswirtschaft ziehen.

Unsere Ergebnisse bestätigen die intuitive Erwartung, dass mit kürzerem Prognosehorizont die Prognosegenauigkeit zunimmt. Der Zeitpunkt der Erstellung der Prognose ist demzufolge entscheidend für die Genauigkeit der Prognosen. BIP-Prognosen mit einem Horizont von einem Jahr (bzw. vier Quartalen) sind im Durchschnitt rund einen Prozentpunkt vom realisierten Wert entfernt. Mit kürzerem Prognosehorizont steht erstens ein grösserer Pool von Daten und Informationen für die Prognose zur Verfügung und zweitens ist die Wahrscheinlichkeit, dass unvorhergesehene Schocks eintreten (z.B. eine unerwartete, deutliche Veränderung der Weltkonjunktur) beschränkter als bei einem längeren Prognosehorizont. Eine grosse Herausforderung bei Konjunkturprognosen stellt der Einbezug und objektive Beurteilung immer neuer Informationen dar, insbesondere solcher, welche einen provisorischen Charakter aufweisen und im Nachhinein (manchmal grundlegend) revidiert werden.

Aussagekraft mittelfristiger Prognosen ist limitiert

Eine eher ernüchternde Erkenntnis ist die im Durchschnitt (gilt somit nicht für alle Jahre) beschränkte Aussagekraft mittelfristiger Prognosen (über ein Jahr hinaus) für den Zeitraum von 1985-2012. Die Prognosegüte mittelfristiger Prognosen ist im historischen Vergleich ähnlich gut wie eine naive Prognose anhand eines Trendwachstums über die letzten 10 Jahre. Diese Feststellung wurde bereits in Ruoss und Savioz (2002) erbracht und in der vorliegenden Studie bestätigt. Laut Ruoss und Savioz (2002) geben Prognosen über einen langen Horizont nicht den Konjunkturverlauf, sondern bestenfalls das Trendwachstum wieder. Gultekin und Lahiri (2007), welche die Prognosegenauigkeit von *Consensus Economics Inc.* Prognosen für 18 Länder untersuchen, kommen zu einem ähnlichen Ergebnis. Hieraus den Schluss zu ziehen, in Zukunft ganz auf mittel- und langfristige Prognosen zu verzichten, ist jedoch nicht angebracht und auch nicht praktikabel. In unterschiedlichen Bereichen wie zum Beispiel der Budgetplanung des Staates sind Wirtschaftsprognosen über ein Jahr hinaus erforderlich. Dass Wirtschaftsprognosen im Allgemeinen gewissen Annahmen unterliegen, muss der Anwender berücksichtigen sowie die Tatsache, dass ein längerer Prognosehorizont mit gesteigerter Unsicherheit verbunden ist.

Keine signifikanten Unterschiede zwischen den Instituten

Signifikante Unterschiede der Prognosequalität zwischen den Instituten lassen sich in der vorliegenden Studie nicht feststellen. Die Prognosefehler der Expertengruppe des Bundes sind somit vergleichbar mit denjenigen der übrigen Institute. Das heisst, dass die Konjunkturprognosen des Bundes im Mittel weder schlechter noch besser als die der anderen Institute sind. Gleichwohl bleibt es eine permanente Aufgabe, die angewandten Prognoseprozesse und -modelle kritisch zu evaluieren und gegebenenfalls anzupassen. Prognosemodelle müssen laufend überarbeitet werden und sollten die jüngsten Erkenntnisse der Forschung miteinbeziehen. Im Weiteren ist die Optimalitätseigenschaft von unverzerrten Prognosen für alle Institute und Prognosehorizonte erfüllt.

Prognosefehler während der Finanzkrise

Unvorhergesehene Schocks, wie die Insolvenz der Bank Lehman Brothers (15. September 2008) oder das unerwartete Ausmass des Kollaps des Welthandels Ende 2008/Anfangs 2009, führten zu einer grundlegenden Neubeurteilung des gesamten Prognoseumfelds. Diese unerwarteten Schocks und die erst dadurch ausgelöste starke Anpassung der exogenen Annahmen (schwere weltwirtschaftliche Rezession) waren massgebliche „Ursachen“ für das grosse Ausmass der Prognosefehler (der Prognosen, welche vor Eintreten dieser Schocks erstellt wurden) in dieser turbulenten Periode. Dass die Prognosefehler während instabiler Wirtschaftsphasen besonders gross sind, ist allerdings kein neues Phänomen seit der Finanzkrise, sondern war in der Vergan-

<sup>22</sup> Die beschränkte Anzahl Beobachtungen sei hier erneut erwähnt.



Weitere Forschungsfragen	<p>genheit regelmässig festzustellen. In Rezessionsphasen ist die Prognosetreffsicherheit im Allgemeinen tiefer als in eher ruhigen Wirtschaftsphasen, in denen weniger unerwartete Schocks auftreten. Somit lassen sich Prognosefehler insbesondere in turbulenten Zeiten über einen längeren Prognosehorizont praktisch nicht vermeiden.</p> <p>Die vorliegende Studie könnte für die Zukunft noch um weitere Aspekte ergänzt/erweitert werden. Interessante und wichtige Aufschlüsse für die Prognoseprozesse der Expertengruppe des Bundes könnte eine vertiefte Analyse der unterschiedlichen Ursachen für die Prognosefehler bringen (falsche exogene Annahmen, schlechte „Prognosemodelle“ etc.). Prognosen basieren fast immer (grösstenteils) auf den in der Vergangenheit empirisch ermittelten Zusammenhängen zwischen den Variablen. Strukturbrüche, eine veränderte Beziehung zwischen den Variablen, unerwartete stochastische Änderungen stellen mögliche Ursachen für Prognosefehler dar. Ein vertieftes Verständnis dieser verschiedenen Fehlerquellen wäre wünschenswert.</p>
Weitere Qualitätsaspekte einer Prognose	<p>Nicht zuletzt sind Prognosen der Expertengruppe des Bundes (wie auch der übrigen Institute) nicht ein rein mechanisches Verfahren, welches sich nur auf die Ergebnisse der Prognosemodelle stützen. Die Einschätzungen der Experten wie auch die Schlussfolgerungen des Austausches zwischen den Ökonomen fliessen direkt oder indirekt in die Ergebnisse der Prognosen ein. Damit lässt sich die Qualität nicht allein durch eine quantitative statistische Analyse untersuchen. Die Kommunikation der zentralen Annahmen, Einschätzungen sowie der Risiken und die dazu gesetzten Nuancen sind ebenso wichtige Bestandteile einer Prognose. Diese wichtigen Aspekte einer Wirtschaftsprognose lassen sich jedoch nur beschränkt quantitativ messen.</p>

**Bibliografie**

- Ager, P., Kappler, M., & Osterloh, S. (2009). The Accuracy and Efficiency of the Consensus Forecasts: A further application and extension of the pooled approach. *International Journal of Forecasting*, 25(1), S. 167-181.
- Clements, M. P. (1997). Evaluating the Rationality of Fixed-event Forecasts. *Journal of Forecasting*, 16(4), S. 225-239.
- Clements, M. P., Joutz, F., & Stekler, H. (2007). An Evaluation of the Forecasts of the Federal Reserve: A Pooled Approach. *Journal of Applied Econometrics*, 22(1), S. 121-136.
- Diebold, F. X., & Mariano, R. S. (1995). Comparing Predictive Accuracy. *Journal of Business and Economic Statistics*, 13(3), S. 253-265.
- Dovern, J., & Weisser, J. (2011). Accuracy, Unbiasedness and Efficiency of Professional Macro-economic Forecasts: An empirical comparison for the G7. *International Journal of Forecasting*, 27(2), S. 452-465.
- Gultekin, I., & Lahiri, K. (2007). How far ahead can we forecast? Evidence from cross-country surveys. *International Journal of Forecasting*, 23(2), S. 167-187.
- Hendry, D. F., & Clements, M. P. (2001). Economic Forecasting: Some Lessons From Recent Research, *Discussion Paper Series*, N. 78, University of Oxford, Department of Economics.
- Holden, K., & Peel, D. (1990). On Testing for Unbiasedness and Efficiency of Forecasts. *The Manchester School*, 58(2), S. 120-127.
- Indergand, R., & Leist, S. (2014). A real-time dataset for Switzerland, mimeo.
- Kappler, M. (2005). Wie genau sind die Konjunkturprognosen der Institute für Deutschland? *ZEW Discussion Paper*, Nr. 06-004, ZEW - Zentrum für Europäische Wirtschaftsforschung.
- Koutsogeorgopoulou, V. (2000). A Post-Mortem on Economic Projections. *OECD Economics Department Working Papers*, Nr. 274, OECD Publishing.
- Mattei, A. (2007). La Qualité des Prévisions Economiques Suisses. Université de Lausanne.
- Meichle, M., Ranaldo, A., & Zanetti, A. (2011). Do financial variables help predict the state of the business cycle in small open economies? Evidence from Switzerland. *Financial Markets and Portfolio Management*, 25(4), S. 435-453.
- Newey, W. K., & West, K. D. (1987). A simple positive semi-definite heteroskedasticity and autocorrelation-consistent covariance matrix. *Econometrica*, 55(3), S. 703-708.
- Noormamode, S. (2002). Analyse des prévisions des principales variables macroéconomiques. *WP Discussion Paper*, Nr. 15, Staatssekretariat für Wirtschaft (SECO).
- Osterloh, S. (2006). Accuracy and properties of German business cycle forecasts. *ZEW Discussion Papers*, Nr: 06-87, Zentrum für Europäische Wirtschaftsforschung (ZEW).
- Pain, N. et al. (2014), OECD Forecasts During and After the Financial Crisis: A Post Mortem. *OECD Economics Department Working Papers*, No. 1107, OECD Publishing.
- Ruoss, E., & Savioz, M. (2002). Wie gut sind BIP-Prognosen? Eine Untersuchung für die Schweiz. *Quartalsheft*, Nr. 3/2002, Schweizerische Nationalbank.
- Spiwojs, M., Scheier, J., & Hein, O. (2011). Zur Beurteilung von Konjunkturprognosen : eine Auswertung von Prognosen zur Entwicklung des BIP, der Industrieproduktion und der privaten Konsumausgaben in zwölf Industrienationen. *Sofia-Diskussionsbeiträge zur Institutionenanalyse*, Nr. 11-01, Sonderforschungsgruppe Institutionenanalyse (Sofia).

## Appendix

**Tabelle 13: Deskriptive Statistik der Prognosen für das reale BIP-Wachstum der jeweiligen Institute (in %), für PH(1)/PH(5) Sample von 1985-2012, für PF(3): 1992-2012 (ohne 1997), für PF(7): 1994-2012 (ohne 1997)**

	<i>PH(1)</i>	<i>PF(3)</i>	<i>PH(5)</i>	<i>PF(7)</i>
<b>Institut 1</b>				
mean	1.41	1.34	1.72	1.85
median	1.75	1.40	1.80	1.90
minimum	-1.90	-2.10	0.50	0.60
maximum	3.30	3.00	2.90	2.60
st.dev.	1.34	1.14	1.60	0.52
observations	28	20	28	17
<b>Institut 2</b>				
mean	1.42	1.23	1.50	1.75
median	1.80	1.50	1.50	1.90
minimum	-1.70	-2.20	0.40	0.10
maximum	3.70	3.10	2.30	2.40
st.dev.	1.43	1.16	0.55	0.55
observations	28	20	28	17
<b>Institut 3</b>				
mean	1.40	1.24	1.59	1.90
median	1.70	1.40	1.70	2.03
minimum	-1.90	-2.70	-0.15	-0.23
maximum	3.80	2.80	2.40	2.61
st.dev.	1.39	1.22	0.61	0.64
observations	28	20	28	17
<b>Institut 4</b>				
mean	1.39	1.29	1.69	1.88
median	1.65	1.50	1.75	2.10
minimum	-3.40	-2.40	0.10	-0.30
maximum	3.80	2.80	2.50	2.70
st.dev.	1.53	1.20	0.58	0.73
observations	28	20	28	17
<b>Institut 5</b>				
mean			1.65	
median			1.70	
minimum			-0.60	
maximum			3.70	
st.dev.			1.04	
observations			28	
<b>Institut 6</b>				
mean			1.50	
median			1.65	
minimum			0.00	
maximum			2.30	
st.dev.			0.68	
observations			28	

	<i>PH(1)</i>	<i>PF(3)</i>	<i>PH(5)</i>	<i>PF(7)</i>
<b>Institut 7</b>				
mean			1.54	
median			1.60	
minimum			-0.30	
maximum			2.50	
st.dev.			0.70	
observations			28	

**Tabelle 14: Deskriptive Statistik der realisierten BIP-Wachstumsraten (in %),**  
Sample: 1985-2012.

	<b>VGR-BIP-1</b>	<b>VGR-BIP-2</b>
mean	1.47	1.51
median	1.77	1.85
minimum	-1.45	-1.91
maximum	3.43	4.00
st.dev.	1.43	1.46
observations	28	28