

# DOKUMENTATION UND ANALYSEN

## Wissenschaftliche Wahlprognosen – Alternative oder Ergänzung zu Umfragen?

Arndt Leininger\*

Prognosen bevölkern unseren Alltag, seien es Wettervorhersagen, Konjunkturprognosen oder Wetten auf Sportereignisse. Wahlen bilden hier keine Ausnahme. Viele Menschen suchen ihren Ausgang zu prognostizieren, darunter Politiker, Journalisten und individuelle Wähler, bisher jedoch noch eher selten Politikwissenschaftler. Die seriöseren Vorhersagen basierten daher bisher nahezu durchweg auf Meinungsumfragen. Sicherlich verbessern diese den Informationsgehalt der von Medienexperten geäußerten Einschätzungen und Meinungen, doch gibt es auch Defizite.

So besteht das Problem, alle Teile der Gesellschaft zu erreichen (die wachsende Zahl an Haushalten ohne Festnetzanschluss stellt eine Herausforderung für die Umfrageforschung dar), dass die Befragten falsche Antworten geben (wer räumt schon ein, dass er die NPD wählt?) oder, was noch problematischer ist, dass sie bis zum Tag der Wahl ihre Meinung ändern. Aus diesen und weiteren Gründen kann es keine perfekte Repräsentativität von Umfragedaten geben. Diese können die Wahlintentionen in der Bevölkerung notwendigerweise nur ungenau abbilden. Natürlich hat die Umfrageforschung Methoden entwickelt, systematische Abweichungen zu korrigieren. Stellt ein Meinungsforschungsinstitut beispielsweise fest, dass in den Rohdaten Männer zwischen 20 und 30 Jahren unterrepräsentiert sind gegenüber den Zensusdaten, können Befragte, die diesem Profil entsprechen, in der Aggregation der Daten mehr Gewicht erhalten, um diese Unterrepräsentation auszugleichen. Auch über einen Abgleich des erinnerten Wahlentscheids mit den Ergebnissen der letzten Bundestagswahl kann eine Gewichtung vorgenommen werden. Details darüber, welche Faktoren und Methoden dabei zum Einsatz kommen, geben die Institute allerdings nicht preis.

Solche Gewichtung kann Umfragewerte in die Nähe von Prognosen rücken, wenn es im Vorfeld von Wahlen auch darum geht, Intentionswerte zu generieren, die möglichst nahe am späteren Wahlergebnis liegen. So veröffentlicht beispielsweise die Forschungsgruppe Wahlen zur Sonntagsfrage neben der „Politischen Stimmung“ auch eine sogenannte „Projektion“. Dennoch sind Umfragen in den meisten Fällen eher als Momentaufnahmen und weniger als Vorhersagen zu betrachten, besonders wenn die Wahl noch weit entfernt ist. Meinungsforscher selbst werden nicht müde, dies zu betonen, jedoch wird ihre Warnung oft von den Medien und einer breiteren Öffentlichkeit ignoriert.

\* Ich bedanke mich ganz herzlich bei Marie Bielke, Simone Brühl, Thorsten Faas, Lukas Fesenfeld, Andreas Graefe, Thomas Gschwend, Ramona Kemper, Theresa Küntzler, Wolfgang Leininger, Simon Munzert und Carsten Reinemann sowie den Teilnehmern der 41. Annual Conference of the International Association for the Study of German Politics für sehr hilfreiche Kommentare und Hinweise. Alle möglicherweise verbleibenden Fehler sind natürlich allein dem Autor zuzuschreiben.

Zur Bundestagswahl 2013 bot sich eine zwar weiterhin bescheidene, jedoch bisher nicht dagewesene Vielfalt an wissenschaftlich fundierten Prognosen, die explizit darauf ausgerichtet waren, das Wahlergebnis mittels statistischer Methoden teilweise Monate vor der Wahl vorherzusagen. Dies nimmt der vorliegende Beitrag zum Anlass, einen Überblick über wissenschaftliche Wahlprognosen in Deutschland zu geben. Erstmals wurden diese ebenso wie Meinungsumfragen eingesetzt bei Präsidentschaftswahlen in den USA, und zwar bereits seit den 1980er Jahren. Vor allem im Vorfeld der Präsidentschaftswahl 2012 erregten sie – wie beispielsweise *Nate Silver* mit seinem Blog „FiveThirtyEight“ – einige Aufmerksamkeit, auch weil sie eine beeindruckende Genauigkeit auszeichnete. So sagten Prognosemodelle den Gewinner in fast allen oder sogar allen 50 US-Bundesstaaten korrekt voraus.<sup>1</sup> Während in den USA mittlerweile eine stattliche Anzahl von Modellen entwickelt wurde, die dort recht gut funktionieren<sup>2</sup>, steckt die Erstellung von Prognosen in Deutschland noch in den Kinderschuhen. So ist es möglich, in diesem Beitrag einen umfassenden Überblick über alle bestehenden Prognosemodelle für Bundestagswahlen zu geben.

### 1. Probleme der Meinungsforschung und das Potential von Prognosemodellen

Zum einen sind Wahlprognosen aus vielen der Gründe, die Prognosen generell interessant machen, attraktiv. Sie bedienen unsere intellektuelle Neugier: Wir wollen im Voraus wissen beziehungsweise abschätzen, wie sich zukünftige Ereignisse gestalten werden. Im Fall von Wahlen konkret: Wer wird die Wahl gewinnen und mit wieviel Stimmen? Dass diese Frage nicht nur Politikwissenschaftler interessiert, wird daran deutlich, dass Spekulationen über den möglichen Wahlausgang ein fester Bestandteil der politischen Berichterstattung vor Wahlen sind – so werden Wahlkampfereignisse mitunter mit Blick auf ihre mögliche Auswirkung auf Umfragewerte und damit auch das Wahlergebnis diskutiert. Prognosen mussten sich dabei bisher fast ausschließlich auf Einschätzungen von Experten – von politischen Insidern oder Kommentatoren – oder auf Meinungsumfragen stützen. Politische Insider besitzen per definitionem exklusives Wissen, jedoch sind ihre Vorhersagen ideologisch voreingenommen. Das heißt, sie wollen, dass ihre Seite gewinnt. Politische Kommentatoren – Journalisten, Essayisten, Think-Tanks – mögen weniger parteiisch sein, ihre auf Erfahrung basierenden Ad-hoc-Prognosen folgen jedoch keiner Systematik, so dass auch sie selbst diese wahrscheinlich nicht reproduzieren könnten.

Andere Formen der Prognose können also das Spektrum erweitern, was nicht zuletzt deshalb wichtig erscheint, weil die Fähigkeit der Meinungsforschungsinstitute, Vorhersagen für Wahlen zu treffen, die erst Wochen oder Monate später stattfinden, nicht überschätzt werden sollte. So wurde etwa bei der Landtagswahl in Niedersachsen im Januar 2013 aufgrund der Umfragen vor der Wahl der Anteil der FDP-Stimmen um nicht weniger als fünf Prozentpunkte unterschätzt – das entspricht nahezu der Hälfte des tatsächlichen Stimmen-

1 Vgl. *Mark Blumenthal*, Polls, Forecasts, and Aggregators, in: *PS: Political Science & Politics*, 47. Jg. (2014), H. 2, S. 297 – 300; *Simon Jackman*, The Predictive Power of Uniform Swing, in: ebenda, S. 317 – 321; *Michael W. Traugott*, Public Opinion Polls and Election Forecasting, in: ebenda, S. 342 – 344.

2 Siehe zum Beispiel die 13 Prognosen in der Oktoberausgabe 2012 der *PS: Political Science and Politics*.

anteils der Partei von 9,9 Prozent. Solche drastischen Abweichungen sind jederzeit möglich, jedoch nur schwer vorhersagbar. Für die letzten fünf Bundestagswahlen (1998 bis 2013) wichen die durch die Forschungsgruppe Wahlen für das ZDF erstellten Politbarometer-Projektionen des Stimmenanteils der bisherigen Koalitionsparteien zwei Monate vor der Wahl jedoch nur um durchschnittlich 2,6 (einen Monat vorher um 3,3) Prozentpunkte vom tatsächlichen Wahlergebnis ab. Die letzten Projektionen unmittelbar vor der Wahl waren mit einer durchschnittlichen Abweichung von 1,3 Prozentpunkten sogar sehr genau. Es bleibt jedoch eine teils deutliche Differenz zwischen den durch Umfragen erhobenen Momentaufnahmen und dem tatsächlichen Wahlergebnis, die in der Tendenz größer ist, je weiter die Wahl in der Zukunft liegt.<sup>3</sup> Dies liegt auch daran, dass die Gruppe der so genannten Späterschlossenen von Wahl zu Wahl gewachsen ist.<sup>4</sup> Daher publizieren die großen Umfrageinstitute mittlerweile noch bis kurz vor der Wahl neue Umfragewerte – so auch das ZDF und die Forschungsgruppe Wahlen 2013 zum ersten Mal nur noch drei Tage vor der Wahl.

Demoskopen weisen zwar Schwankungsbreiten zur Quantifizierung der Ungenauigkeit der errechneten Schätzwerte aus, diese können jedoch zukünftige Meinungsänderungen nicht berücksichtigen. So zeigen *Schnell* und *Noack* in einer Analyse von 232 Umfragen unmittelbar vor den Wahlen zwischen 1957 und 2013 auf, dass das tatsächliche Wahlergebnis nur für etwa zwei Drittel der Umfragen innerhalb der durch die Institute ausgegebenen Schwankungsbreiten der Umfragewerte lag.<sup>5</sup>

Es bleibt also festzuhalten, dass Umfragen, besonders wenn die Wahlen noch weit entfernt sind, eher die aktuelle politische Stimmung widerspiegeln und in ihren Möglichkeiten begrenzt sind, verlässliche Vorhersagen zu liefern. Wissenschaftliche Prognosemodelle sind im Kontrast dazu explizit darauf ausgelegt, den Ausgang von Wahlen Wochen oder gar Monate im Voraus vorherzusagen. Zugleich folgen sie wissenschaftlichen Prinzipien der intersubjektiven Nachvollziehbarkeit und Reproduzierbarkeit. Sie sind damit eine für die interessierte Öffentlichkeit relevante Ergänzung zu Umfragen und können so auch zu einem reflektierteren Umgang mit diesen anregen und Aufmerksamkeit für politikwissenschaftliche Forschung schaffen.

Welche *fachlichen* Gründe sprechen aber dafür, sich für Wahlprognosen zu interessieren? Ist es nicht Aufgabe der Politikwissenschaft, auf Basis handhabbarer, also vergangener, Daten Erkenntnisse zu gewinnen, statt sich an Spekulationen über die Zukunft zu beteiligen? In der Tat hängen jedoch die Retrospektive der Analyse und die Prognose eng zusammen. Politikwissenschaftliche Analysen zielen in der Regel nicht darauf, Besonderheiten einzelner Wahlen hervorzuheben, sondern wollen allgemeine Muster und auch kausale Mechanismen aufdecken. Auf dieser Basis sollten Fortschreibungen auch in die Zukunft möglich sein. Prognostische Validität ist ein Eckpfeiler der Wissenschaft, schwieriger zu erreichen als Erklärung und entsprechend nach Meinung einiger Autoren oft in den Sozialwissenschaften

- 3 Vgl. *Jochen Groß*, Die Prognose von Wahlergebnissen: Ansätze und empirische Leistungsfähigkeit, Wiesbaden 2010.
- 4 Vgl. *Rüdiger Schmitt-Beck*, Empirische Wahlforschung in Deutschland: Stand und Perspektiven zu Beginn des 21. Jahrhunderts, in: *ders.* (Hrsg.), Wählen in Deutschland, PVS-Sonderheft 45, Baden-Baden 2012, S. 2 – 39; *Thomas Plischke*, Wann Wähler entscheiden. Abläufe von Entscheidungsprozessen und der Zeitpunkt der Wahlentscheidung, Baden-Baden 2014.
- 5 Vgl. *Rainer Schnell* / *Marcel Noack*, The Accuracy of Pre-Election Polling of German General Elections, in: *Methods, Data, Analyses*, 8. Jg. (2014), H. 1, S. 5 – 24.

vernachlässigt.<sup>6</sup> Wahlprognosen sind daher vielversprechend, weil sie auf Theorien basieren und die Grundlagen der durch die Wahlforschung belegten Wahltheorien erfassen wollen. Mit ihnen lassen sich Theorien auf den denkbar härtesten Prüfstand stellen – die Vorhersage zukünftiger Wahlergebnisse.<sup>7</sup>

Wissenschaftliche Wahlprognosen, sobald einmal berechnet, werden nicht durch nachfolgende Ereignisse während des Wahlkampfes beeinflusst. Theoriebasierte Prognosen können daher eine Art erwartetes normales Wahlergebnis bereitstellen, auf dessen Grundlage eine tatsächliche Wahl beurteilt werden kann, so dass Prognosen selbst dann einen Erklärungswert haben, wenn sie in manchen Fällen daneben liegen. Sie stellen so einen Referenzrahmen dar, vor dem die Auswirkungen als besonders erachteter Ereignisse, die in jeder Kampagne auftreten, bewertet werden können. Darüber hinaus können Wahlprognosen einen eigenen Beitrag zum Forschungsstand leisten. So haben sie etwa zum Verständnis des retrospektiven und kurzfristigen Effekts wirtschaftlicher Rahmenbedingungen auf das Wahlergebnis wesentlich beigetragen.<sup>8</sup>

## 2. Prognosemodelle für Bundestagswahlen im Vergleich

Prognosemodelle verwenden statistische Methoden, um auf der Basis von vor der Wahl verfügbaren Daten deren Ausgang vorherzusagen. Dabei kann man derzeit drei Arten unterscheiden: Von besonderem Interesse sind hier so genannte strukturelle Modelle, die den Zusammenhang zwischen Wahlergebnissen und erklärenden Faktoren wie der wirtschaftlichen Lage oder der Popularität des amtierenden Kanzlers in vergangenen Wahlen ermitteln und auf dieser Basis zukünftige Wahlergebnisse vorherzusagen suchen. Strukturelle Modelle greifen auf Theorie und Empirie der Wahlforschung zurück und machen diese prädiktiv nutzbar – sie sind daher für Wahlforscher und an Wahlforschung Interessierte besonders relevant. In Deutschland zählen zu dieser Gruppe die Modelle von *Jérôme*, *Jérôme-Speziari* und *Lewis-Beck*, die 1998 die erste Prognose für eine Bundestagswahl veröffentlichten<sup>9</sup> und für die Wahl 2013 ein neues „Political Economy Model“<sup>10</sup> vorlegten, das „Kanzlermodell“ von *Norpoth* und *Gschwend*<sup>11</sup> sowie das von *Kayser* und *Leininger* entwickelte „Benchmarking Modell“<sup>12</sup>. Daneben werden für Prognosen auch „Aggregationsmodelle“ verwendet,

6 Vgl. *Alexander Rosenberg*, *Economics – Mathematical Politics or Science of Diminishing Returns?*, Chicago 1992; *Philip A. Schrodt*, *Seven Deadly Sins of Contemporary Quantitative Political Analysis*, in: *Journal of Peace Research*, 51. Jg. (2014), H. 2, S. 287 – 300.

7 Vgl. *Mark A. Kayser* / *Arndt Leininger*, *A Predictive Test of Voters' Economic Benchmarking: The 2013 German Bundestag Election*, in: *German Politics* (im Erscheinen).

8 Vgl. *William G. Mayer*, *What, If Anything, Have We Learned from Presidential Election Forecasting?*, in: *PS: Political Science & Politics*, 47. Jg. (2014), H. 2, S. 329 – 331.

9 Vgl. *Bruno Jérôme* / *Véronique Jérôme-Speziari* / *Michael S. Lewis-Beck*, *Prévisions politico – économiques et élections législatives allemandes de septembre 1998*, in: *Le Figaro* vom 10. Juli 1998.

10 Vgl. *dies.*, *A Political-Economy Forecast for the 2013 German Elections: Who to Rule with Angela Merkel?*, in: *PS: Political Science & Politics*, 46. Jg. (2013), H. 3, S. 479 – 480.

11 Vgl. *Helmut Norpoth* / *Thomas Gschwend*, *Chancellor Model Picks Merkel in 2013 German Election*, in: *PS: Political Science & Politics*, 46. Jg. (2013), H. 3, S. 481 – 482.

12 Vgl. *Mark A. Kayser* / *Arndt Leininger*, *A Benchmarking Forecast of the 2013 Bundestag Election*, Berlin 2013, im Internet auch abrufbar unter dem Titel: „Economic Growth Trumps Political

**Tabelle 1: Übersicht über die vorgestellten Prognosemodelle anhand der vorhergesagten Merkmale, der Methode und der wichtigsten theoretischen Hintergründe**

Modell	Prognosewert(e)	Methode	Theoretische Annahmen
Strukturelle Modelle			
<i>Jérôme u.a.</i>	Parteistimmen- anteile	6 Regressionsmodelle für 5 Parteien und „Sonstige“	Regierungsparteien werden für die wirtschaftliche Lage verantwortlich gehalten, und Kanzlerkandidaten spielen eine herausgehobene Rolle für den Wahlentscheid.
<i>Kayser / Leininger</i>	Koalitionsstimmen- anteil	1 Regressionsmodell für den Stimmenanteil der Koalition	Die wirtschaftliche Lage ist relativ zu anderen EU-Ökonomien, besonders relevant in Wahlen nach der Finanz- und Eurokrise.
<i>Norpoth / Gschwend</i>	Koalitionsstimmen- anteil	1 Regressionsmodell für den Stimmenanteil der Koalition	Die Popularität der Amtsinhaber ist indikativ für die Bewertung der Arbeit der Regierung und damit für den Wahlentscheid.
Aggregationsmodelle und Synthetische Modelle			
<i>Graefe</i>	Parteistimmen- anteile	Mittelwertbildung über Prognosen der anderen Prognosemodelle, Um- fragen, Experteneinschät- zungen und Wettermärkte	Keine Prognosemethode ist systema- tisch besser als andere. Der Mittel- wert aller Methoden kann daher dennoch akkurater sein als einzelne Methoden.
<i>Selb / Munzert</i>	Parteistimmen- anteile	1 Regressionsmodell	Es gibt instituts- und parteispezifi- sche Fehler in den Umfragedaten. Diese werden durch ein statistisches Modell geschätzt und ausgeglichen.
<i>Küntzler</i>	Koalitionsstimmen- anteil	Mittelwertbildung aus zwei Regressions- modellen	Berechnung eines gewichteten Mittels aus Umfragedaten und dem Kanzlermodell
Quelle: Eigene Zusammenstellung.			

die mehrere Umfragen kombinieren und die in Umfragen inhärenten Prognosefehler zu minimieren versuchen. Hierzu gehören in Deutschland das exklusiv umfragedatenbasierte Prognosemodell von *Selb* und *Munzert*<sup>13</sup> sowie der von *Graefe*<sup>14</sup> entwickelte „PollyVote“, der sich auf verschiedenste Quellen stützt. Außerdem existieren „synthetische Modelle“, die strukturelle Modelle mit Umfragenaggregation verbinden. *Theresa Küntzler*<sup>15</sup> beispielsweise kombiniert das „Kanzlermodell“ mit einem eigenen Umfragemodell.

Tabelle 1 bietet einen ersten Überblick über die hier vorzustellenden Modelle. Diese unterscheiden sich sowohl nach den Werten, die vorhergesagt werden, der Methodik als

Campaigning: CDU/FDP Coalition Set to Win another Election“, 6. August 2015, <https://www.hertie-school.org/blog/mark-kayser/> (Abruf am 27. November 2015).

13 Vgl. *Peter Selb / Simon Munzert*, Forecasting the 2013 Bundestag Election Using Data from Various Polls. SSRN Scholarly Paper ID 2313845, Rochester 2013.

14 Vgl. *Andreas Graefe*, German Election Forecasting: Comparing and Combining Methods for 2013, in: *German Politics*, 24. Jg. (2015), H. 2, S. 195 – 204.

15 Vgl. *Theresa Küntzler*, Using Bayesian Data Combination of Fundamental Variable-Based Forecasts and Poll-Based Forecasts to Predict the 2013 German Election, SSRN Scholarly Paper ID 2602955, Rochester 2014.

auch den zu Grunde liegenden theoretischen Annahmen (die Tabelle nennt hier nur den wichtigsten theoretischen Bezugsrahmen).

## 2.1. Strukturelle Prognosemodelle

Strukturelle Modelle waren die ersten und sind immer noch die am weitesten verbreiteten wissenschaftlichen Prognosemodelle. Anders als Umfragen sind sie explizit als Vorhersage und daher auch als Ergänzung zu diesen konzipiert worden. Wer Umfragen als Prognosen nutzt, muss die Annahme treffen, dass die darin geäußerten Intentionen bis zur Wahl unverändert bleiben oder sich individuelle Veränderungen im Aggregat ausgleichen. Dies ist aus den oben genannten Gründen eine nur schwerlich haltbare Annahme.

Dabei gehen strukturelle Prognosemodelle natürlich auch von einigen Annahmen aus, allerdings von weniger gewagten. Zunächst setzen sie voraus, dass keine Wahl wirklich einzigartig ist, und zweitens, dass die Bedeutung verschiedener die Wahlentscheidung beeinflussender Faktoren über die Zeit hinweg konstant bleibt. Der konkrete Aufbau eines strukturellen Modells leitet sich aus Theorie und Empirie der Wahlforschung ab.

In Mehrparteiensystemen müssen Entwickler einer Wahlprognose zunächst die Entscheidung treffen, welche Größe prognostiziert werden soll – das Wahlergebnis einer oder mehrerer Parteien oder der die Regierung tragenden Fraktion. Eine Prognose abzugeben scheint in einem Mehrparteiensystem wie Deutschland nämlich deutlich schwerer als in einem Zweiparteiensystem wie den USA, in deren Kontext Wahlprognosen für Präsidentschaftswahlen zunächst entwickelt wurden.<sup>16</sup>

Die meisten Menschen interessieren sich für Prognosen, weil sie wissen möchten, wer die Wahl gewinnen wird. Dies hängt jedoch in einem Mehrparteiensystem auch von den nach der Wahl stattfindenden Koalitionsverhandlungen ab. Daher ist es sinnvoller, nicht die Gewinner und Verlierer zu betrachten, sondern sich auf die Stimmenanteile der Parteien zu konzentrieren. Ein Fokus auf den Stimmenanteil der Koalition erlaubt es vorherzusagen, ob diese an der Regierung bleiben wird. Da ein guter Teil der Theoriebildung über Wahlen auf der Unterscheidung zwischen Regierung und Opposition beruht, lässt sich eine breite politikwissenschaftliche Literatur für die Konstruktion solcher Modelle nutzen.

*Norpoth* und *Gschwend*<sup>17</sup>, *Kayser* und *Leininger*<sup>18</sup> und *Küntzler*<sup>19</sup> (da auf dem „Kanzlermodell“ aufbauend), entscheiden sich für Prognosen des Koalitionsstimmenanteils. Die Alternative dazu ist, die Stimmenanteile für alle im Parlament vertretenen Parteien zu prognostizieren, wie es *Jérôme* und Ko-Autoren<sup>20</sup> tun. Ebenso prognostizieren die Modelle

16 Vgl. *Michael S. Lewis-Beck* / *Tom W. Rice*, *Forecasting Elections*, Washington D.C. 1992. Prognosemodelle der US-Präsidentschaftswahlen sagen üblicherweise den Stimmenanteil des Kandidaten der Partei des Amtsinhabers voraus, relativ zu den für eine der beiden großen Parteien abgegebenen Stimmen. Der prognostizierte Stimmenanteil für die andere Partei ist dann einfach die Differenz zu 100 Prozent.

17 Vgl. *Thomas Gschwend* / *Helmut Norpoth*, Soll und Haben: Die deutsche Wählerschaft rechnet mit den Parteien ab, in: *Markus Klein* / *Wolfgang Jagodzinski* / *Ekkehard Mochmann* / *Dieter Ohr* (Hrsg.), *50 Jahre Empirische Wahlforschung in Deutschland. Entwicklung, Befunde, Perspektiven*, Daten, Wiesbaden 2000, S. 389 – 409.

18 Vgl. *Mark A. Kayser* / *Arndt Leininger*, a.a.O. (Fn. 12).

19 Vgl. *Theresa Küntzler*, a.a.O. (Fn. 15).

20 Vgl. *Bruno Jérôme* / *Véronique Jérôme-Speziari* / *Michael S. Lewis-Beck*, a.a.O. (Fn. 10).



Selbs und Munzerts<sup>21</sup> sowie Graefes<sup>22</sup> die Stimmenanteile aller fünf im 17. Deutschen Bundestag vertretenen Parteien und den „Sonstigen“.

Die Prognose der Stimmenanteile aller Parteien bietet mehr, aber nicht unbedingt bessere Informationen. Sie können leicht addiert werden, um den Koalitionsstimmenanteil und so die Wahrscheinlichkeit, dass die Regierung im Amt bleiben kann, zu berechnen. Weiterhin erlauben sie in Wahlsystemen wie dem deutschen, das eine Fünf-Prozent-Hürde vorsieht, vorherzusagen, ob einzelne Parteien im Parlament vertreten sein werden. Jedoch scheinen die Prognosen für einzelne Parteien weniger akkurat zu sein als Prognosen über den Koalitionsstimmenanteil. Erstere bieten somit mehr Informationen, beinhalten jedoch das Risiko größerer Ungenauigkeit. Zudem sind zur Konstruktion von Modellen, die den Stimmenanteil einzelner Parteien vorhersagen, mehr Ad-hoc-Annahmen zu treffen.

Strukturelle Modelle in ihrer einfachsten Form basieren auf einem simplen statistischen Modell, einem so genannten Regressionsmodell, das die Stimmenanteile der Regierungsparteien oder einzelner Parteien mit erklärenden Variablen, zum Beispiel dem Wirtschaftswachstum im Vorfeld der Wahl, in Zusammenhang setzt. Dieses Modell wird auf einer Zeitreihe vergangener Wahlen geschätzt. Das Ergebnis der statistischen Schätzung ist eine Gleichung, die den Zusammenhang zwischen Wahlergebnissen und den inkludierten Erklärungsfaktoren widerspiegelt. Diese erlauben es, den zu erwartenden Stimmenanteil in einer bevorstehenden Wahl zu schätzen, indem man aktuelle Werte für die erklärenden Variablen in die Gleichung einsetzt.

Eine Vorstellung davon, wie gut ein Modell tatsächlich ist, erhält man nur durch Prognosen zu Wahlen außerhalb der Stichprobe.<sup>23</sup> Die Bundestagswahl von 2013 stellte in dieser Hinsicht einen wichtigen Test dar. Nach einer Wahl lässt sich die Genauigkeit einer Vorhersage einfach bestimmen. Doch muss man nicht bis zur nächsten Wahl warten, um zu prüfen, ob ein Modell geeignet ist, gute Prognosen zu generieren.

Viele Autoren tun dies, indem sie eine Wahl aus der Stichprobe nicht berücksichtigen, das Modell in den übrigen Wahlen neu schätzen und mit den Werten der nicht berücksichtigten Beobachtungen eine Prognose für diese erstellen. Dies tun sie für einige Wahlen, vergleichen die Prognose mit dem tatsächlichen Ergebnis und berechnen einen durchschnittlichen Fehler, um die Genauigkeit des Modells zu beurteilen. Der geschätzte Zusammenhang zwischen Wahlergebnis und Erklärungsfaktoren sollte dabei über alle Schätzungen hinweg stabil bleiben, das heißt nicht durch einzelne Wahlen stark beeinflusst sein. Nur wenn der geschätzte Zusammenhang über die Zeit hinweg stabil ist, kann davon ausgegangen werden, dass dieser wahrscheinlich auch in Zukunft fortbestehen wird.

Als Autoren eines strukturellen Modells sind zunächst *Bruno Jérôme, Véronique Jérôme-Speziari* und *Michael Lewis-Beck* zu nennen. Sie haben bereits viele Vorhersagen für Wahlen in Frankreich, den USA und anderen Ländern erstellt und konstruierten auch das erste Prognosemodell für eine deutsche Bundestagswahl.<sup>24</sup> Zentrale Komponente ihres „Political Economy“-Modells, das die Abwahl *Helmut Kohls* 1998 erfolgreich prognostizierte, war die

21 Vgl. *Peter Selb / Simon Munzert*, a.a.O. (Fn. 13).

22 Vgl. *Andreas Graefe*, a.a.O. (Fn. 14).

23 Der Begriff der Stichprobe bezeichnet hier die Menge der vergangenen Wahlen über die das Modell geschätzt wird. Diese beinhaltet notwendigerweise nicht die bevorstehende Wahl, kann jedoch sonst aus der vollständigen Population aller vergangenen Bundestagswahlen bestehen.

24 Vgl. *Bruno Jérôme / Véronique Jérôme-Speziari / Michael S. Lewis-Beck*, a.a.O. (Fn. 9).

Arbeitslosenrate in den letzten zwei Quartalen vor der Wahl. Ein um die Wahlintentionen für den kleineren Koalitionspartner erweitertes Modell sagte 2009 einer Koalition aus CDU/CSU und FDP ein Ergebnis von 51,2 Prozent voraus – tatsächlich vereinigte die im Anschluss an diese Wahl auch gebildete Koalition einen Stimmenanteil von nur 41,9 Prozent auf sich, dieser war jedoch ausreichend für eine parlamentarische Mehrheit.

2013 stellte das Autorenteam ein neues „Political Economy“-Modell vor.<sup>25</sup> Dieses besteht aus für jede der fünf großen Parteien sowie für die Kategorie „Sonstige“ jeweils separaten Schätzgleichungen. Ihre Variablenauswahl beruht auf der Wahlforschung und schließt eine ökonomische Einflussgröße wie die Arbeitslosenrate ein, aber auch politische Variablen wie die Popularität der Kanzlerin.

Das „Kanzlermodell“ war zwar nicht das erste Prognosemodell für eine Bundestagswahl, seine Urheber *Helmut Norpoth* und *Thomas Gschwend* können dennoch als Wegbereiter für Wahlprognosen in Deutschland gelten.<sup>26</sup> Ihre letzte Prognose vor der Wahl hat eine Vorlaufzeit von einem Monat, das Modell erlaubt aber Vorhersagen zu früheren Zeitpunkten mit vorläufigen Daten. Seit den Bundestagswahlen 2002 erhielten ihre Prognosen relativ viel Aufmerksamkeit und haben sehr gute Ergebnisse geliefert. Im Jahr 2002 sagten sie den Stimmenanteil der amtierenden rot-grünen Regierung sogar exakt voraus. Ihr Modell stützt sich auf drei Faktoren: die Popularität des amtierenden Kanzlers (Prozentsatz der Menschen, die diesen dem Herausforderer vorziehen), den langfristigen Wählerrückhalt (durchschnittlicher Wahlerfolg der Koalitionsparteien bei den drei vorangegangenen Wahlen) und die Anzahl der Wahlperioden, während derer die Partei des Kanzlers an der Regierung war.<sup>27</sup>

Für die Bundestagswahl 2013 stellten *Kayser* und *Leininger* erstmals ihr „Benchmarking-Modell“ vor. Sie prognostizierten ebenfalls den Stimmenanteil der Koalitionsregierung. Die namensgebende Variable ihres Modells ist das Wachstum in Deutschland im Vergleich zu den drei nächstwichtigsten Volkswirtschaften in Europa: Großbritannien, Frankreich und Italien. Indem sie die Abweichung der deutschen Wachstumsraten vom durchschnittlichen britischen, französischen und italienischen Wachstum benutzen, nehmen sie implizit an, dass die Wähler den Zustand ihrer Volkswirtschaft im Vergleich zu dem anderer Länder beurteilen, da die Wahrnehmung der wirtschaftlichen Lage vor allem durch Medien vermittelt sei.<sup>28</sup> Dieses Maß relativer Wirtschaftsleistung könnte für die Prognose der Wahl 2013 von besonderer Bedeutung gewesen sein, da das Wachstum in Deutschland zwar verhalten war, aber im Vergleich mit anderen europäischen Ländern, die an den Folgen der Finanz- und Eurokrise litten, immer noch vergleichsweise hoch ausfiel. *Kayser* und *Leininger* nehmen außerdem als Prädiktoren den Stimmenanteil der derzeit amtierenden Parteien in der vor-

25 Vgl. *dies.*, a.a.O. (Fn. 10).

26 Vgl. *Thomas Gschwend* / *Helmut Norpoth*, a.a.O. (Fn. 17).

27 Vgl. *dies.*, „Wenn am nächsten Sonntag ...“: Ein Prognosemodell für Bundestagswahlen, in: *Hans-Dieter Klingemann* / *Max Kaase* (Hrsg.), *Wahlen und Wähler: Analysen aus Anlass der Bundestagswahl 1998*, Wiesbaden 2001, S. 473 – 499.

28 Belege für ein solches „Benchmarking“ über Grenzen hinweg finden sich bei *Kayser* und *Peress*, die das Phänomen im Wesentlichen so erklären, dass die Medien positiver über die Wirtschaft berichten, wenn sie im Vergleich zu anderen Ländern besser abschneidet (vgl. *Mark A. Kayser* / *Michael Peress*, *Benchmarking Across Borders: Electoral Accountability and the Necessity of Comparison*, in: *American Political Science Review*, 106. Jg. (2012), H. 3, S. 661 – 684).



herigen Wahl hinzu, den Anteil der Befragten, die angaben, eine Bindung an eine der Parteien zu haben, sowie die Anzahl von Wahlperioden, die eine Regierung im Amt war.

Alle hier vorgestellten Modelle konzentrieren sich in Anlehnung an das „Michigan-Modell“ der Wahlentscheidung auf eine Kombination aus lang-, mittel- und kurzfristigen Faktoren.<sup>29</sup> Dabei kommt allerdings keines der Modelle ohne Umfragedaten aus. Bei *Norpoth* und *Gschwend* ist es die Kanzlerpopularität, bei *Jérôme*, *Jérôme-Speziari* und *Lewis-Beck* sind es die Kanzlerpräferenz sowie Wahlintentionen für die Grünen, Die Linke und die sonstigen Parteien. *Kayser* und *Leininger* setzen auf die Politbarometer-Frage zur Parteiidentifikation. Auch deshalb ist es wenig sinnvoll, Prognosen als Alternative zu Umfragen zu betrachten. Für die im folgenden Abschnitt dargestellten Modelle sind Umfragedaten eine noch wichtigere, in einem Fall sogar die einzige Datenquelle.

## 2.2. Aggregations- und synthetische Modelle

Während die bisher dargestellten Modelle sich explizit auf Erkenntnisse der Wahlforschung beziehen, fußen die im Folgenden dargestellten Aggregations- und synthetischen Modelle primär auf statistischer Theorie. Ohne sie wäre ein umfassender Überblick über wissenschaftliche Wahlprognosen jedoch nicht vollständig. Diese Modelle arbeiten nicht mit den Rohdaten der Umfragen, sondern fassen die von den Instituten berechneten Aggregatwerte zu einer Prognose zusammen. *Selb* und *Munzert* wählen dabei einen den strukturellen Modellen ähnlichen methodischen Ansatz, der den Zusammenhang zwischen Umfragedaten und Ergebnissen vorangegangener Wahlen extrapoliert. *Graefe* beschränkt sich auf die Umfragedaten der bevorstehenden Wahl, bezieht jedoch auch die zuvor vorgestellten Prognosemodelle, Experteneinschätzungen und Prognosemärkte<sup>30</sup> ein und kann so tagesaktuelle Vorhersagen liefern. *Küntzler* konstruiert ein synthetisches Modell, das ein gewichtetes Mittel aus dem „Kanzlermodell“ und einem eigenen umfragedatenbasierten Modell als Prognose berechnet.

Der Ansatz *Selbs* und *Munzerts* unterscheidet sich dadurch von den bisher vorgestellten Modellen, dass er ausschließlich auf Umfragedaten basiert. Durch die Schätzung des Zusammenhangs zwischen den Umfrageergebnissen vor der Wahl und den tatsächlichen Resultaten hoffen sie, etwaige bei Umfragen vor der Wahl auftretende partei- oder institutspezifische Verzerrungen zu eliminieren und so zu einer genaueren Prognose zu gelangen als einzelne Umfragen. Interessanterweise fanden sie heraus, dass sie die zutreffendste Prognose auf Basis von Umfragen erhalten, die acht Monate vor der Wahl durchgeführt wurden. Die Genauigkeit nimmt dabei im Kampagnenverlauf zunächst ab und steigt erst wieder, wenn Umfrageergebnisse herangezogen werden, die näher an der Wahl liegen.

29 Vgl. *Harald Schoen* / *Cornelia Weins*, Der sozialpsychologische Ansatz zur Erklärung von Wahlverhalten, in: *Jürgen W. Falter* / *Harald Schoen* (Hrsg.), Handbuch Wahlforschung, Wiesbaden 2005, S. 187 – 242.

30 Prognosemärkte sind virtuelle Marktplätze, auf denen Teilnehmer auf die großen Parteien lautende Anteilsscheine, die in einer Phantasiewährung notiert sind, handeln. Die endgültige Wertsetzung der Anteilsscheine erfolgt mit dem Wahlergebnis, das den Auszahlungswert bestimmt. Über den Marktmechanismus soll dabei das gesammelte Wissen der Teilnehmer für eine Wahlprognose genutzt werden, die dem momentanen Kurs der virtuellen Anteilsscheine entspricht.

Eine weitere, simple, jedoch nicht minder effektive Methode der Umfrageaggregation ist der von *Andreas Graefe* entwickelte „PollyVote“.<sup>31</sup> Dieser Ansatz kombiniert vier Kategorien von Prognosen – Umfragen, Prognosemärkte, Expertenurteile und die im vorangegangenen Abschnitt vorgestellten Modelle. Er wurde von *Graefe* bereits seit 2004 für die Prognose von US-Präsidentchaftswahlen eingesetzt.<sup>32</sup> Die Idee hinter dem PollyVote ist, dass keine einzelne Methode durchweg besser als andere Methoden ist. *Graefe* berechnet bewusst nur einen einfachen Durchschnitt der Prognosen in allen vier Kategorien, da kompliziertere Kombinationsverfahren keine Vorteile hinsichtlich der Genauigkeit böten.<sup>33</sup> So zeigt er in einer Analyse über sechs Wahlen von 1992 bis 2012 auf, dass die Methode, die für eine Wahl die genaueste Prognose lieferte, in der darauf folgenden Wahl oftmals die schlechteste war.<sup>34</sup>

*Küntzler* präsentiert, erstmals in Deutschland, ein aus einer Kombination eines strukturellen Modells und eines umfragebasierten Modells bestehendes dynamisches Prognosemodell. Dieses erlaubt es, mit jeder neu verfügbaren Umfrage die Prognose zu aktualisieren. Die erste Komponente ist eine Replikation des „Kanzlermodells“. Die zweite Komponente ist ein statistisches Modell, das Wahlergebnisse auf Basis von Umfragewerten vorhersagt. Um zu einer finalen Prognose zu gelangen, kalkuliert *Küntzler* ein mit der jeweiligen Unsicherheit gewichtetes Mittel aus den beiden Einzelvorhersagen. Je näher die Wahl rückt, desto mehr Gewicht erhält die umfragebasierte Prognose in der Kombination.

### 3. Die Vorhersagen von Prognosemodellen und Umfragen bei der Bundestagswahl 2013

Die Frage, welche Prognosemodelle bei der Bundestagswahl 2013 am besten abschnitten, ist nicht einfach zu beantworten, denn die genaueste Vorhersage ist nicht notwendigerweise auch die beste. Denn was ist eine Prognose wert, die exakt zutrifft, aber nur einen Tag vor der Wahl entstand? Eine Prognose, die beispielsweise zwei Prozentpunkte vom Wahlergebnis abweicht, aber schon acht Wochen vor der Wahl veröffentlicht wurde, kann durchaus als bessere bezeichnet werden. Hier kann es individuell unterschiedliche Meinungen geben, wie sich die richtige Balance zwischen der Vorlaufzeit und der Genauigkeit darstellt. Daher wird hier auf ein Ranking der verschiedenen Prognosen verzichtet und stattdessen nur eine Zusammenfassung präsentiert (vgl. Tabelle 2).

Das „Kanzlermodell“, das „Benchmarking Modell“ sowie *Küntzlers* synthetisches Modell sagten nur den kombinierten Stimmenanteil der Parteien in der amtierenden Koalition (CDU/CSU und FDP) voraus. Diese Modelle waren daher nicht in der Lage, das Ausscheiden der FDP aus dem Bundestag vorherzusehen. Modelle, die die individuellen Stimmenanteile der Parteien prognostizieren, haben hier im Prinzip einen Vorteil, jedoch erwarteten alle Modelle – ebenso wie die Institute Allensbach, Forsa und die Forschungsgruppe Wah-

31 Vgl. *Andreas Graefe*, a.a.O. (Fn. 22).

32 Vgl. *ders.* / *J. Scott Armstrong* / *Randall J. Jones* / *Alfred G. Cuzan*, Combining Forecasts: An Application to Elections, in: *International Journal of Forecasting*, 30. Jg. (2014), H. 1, S. 43 – 54.

33 Vgl. *Andreas Graefe* / *Helmut Kuechenhoff* / *Veronika Stierle* / *Bernhard Riedl*, Limitations of Ensemble Bayesian Model Averaging for Forecasting Social Science Problems, in: *International Journal of Forecasting*, 31. Jg. (2015), H. 3, S. 943 – 951.

34 Vgl. *Andreas Graefe* / *J. Scott Armstrong* / *Randall J. Jones* / *Alfred G. Cuzan*, a.a.O. (Fn. 32).

**Tabelle 2: Das offizielle Endergebnis und die Abweichungen der Prognosen und letzten Umfragen vor der Wahl vom tatsächlichen Wahlergebnis**

		Strukturelle Modelle					Synthetische Modelle		Umfragen		
	Endgültiges Ergebnis	<i>Jérôme u.a.</i>	<i>Norpoth / Gschwend</i>	<i>Kayser / Leininger</i>	<i>Selb / Munzert</i>		<i>Graefe</i>	<i>Küntzler</i>	Forschungsgruppe Wah-len (19.9.)	Forsa (20.9.)	Allensbach (20.9.)
Genauigkeit (MAE)	–	1,82	4,9	0,75	2,82		1,15	0,98	0,84	0,89	0,89
Vorlaufzeit (Monate)	–	5	5	2	8		2	1	3 Tage	3 Tage	3 Tage
Koalition	46,3	0,7	4,9	0,75	-2,8		1,5	0,98	-0,8	-1,3	-1,3
CDU/CSU	41,5	-0,5			3,4		-2,5		-1,5	-1,5	-2,0
FDP	4,8	1,2			0,6		1,0		-0,7	0,2	0,7
SPD	25,7	2,3			2,5		0,5		1,3	0,3	1,3
Grüne	8,4	1,6			5,1		1,9		0,6	1,6	0,6
Die Linke	8,6	0,4			-0,9		-0,2		0,1	0,4	0,4
Sonstige	10,9	-4,9			-4,4		-0,8		-0,9	-0,9	-0,9
Quelle: Eigene Zusammenstellung.											

len –, dass die FDP mindestens fünf Prozent erhalten würde. Das tatsächliche Ergebnis lag jedoch innerhalb der Fehlermargen sowohl der Umfragen als auch der Prognosemodelle.

Natürlich birgt ein Modell, das mehrere Stimmenanteile voraussagt, ein größeres Fehlerisiko. Um einen fairen Vergleich der Genauigkeit zu bieten, werden daher auch die mittleren absoluten Fehler<sup>35</sup> über alle Prognosen berechnet (vgl. Tabelle 2). Das Modell *Selbs* und *Munzerts* hat die größte Vorlaufzeit, aber einen durchschnittlichen Prognosefehler von 2,82 Prozentpunkten. Das „Benchmarking Modell“ hat den kleinsten Fehler mit 0,75 Prozentpunkten, aber eine Vorlaufzeit von lediglich zwei Monaten und sagt zudem nur den Koalitionsstimmenanteil voraus.

Der limitierende Faktor bei der Vorlaufzeit ist meist die umfragebasierte Variable, so dass das Modell erst geschätzt werden kann, wenn die entsprechenden Umfragedaten zur Verfügung stehen. *Graefes* und *Küntzlers* Modelle haben keine klar definierte Vorlaufzeit, da ihre Prognosen kontinuierlich auf Basis neuer Daten aktualisiert werden – bei *Graefe* ab einem Zeitraum von 58 Tagen und bei *Küntzler* ab 30 Tage vor der Wahl. Zum Vergleich mit anderen Modellen wird daher der mittlere Prognosefehler über den gesamten Zeitraum herangezogen. Dieser fällt bei den synthetischen Modellen kleiner aus als bei den strukturellen Modellen und liegt zwischen diesen und den letzten Umfragen kurz vor der Wahl (vgl. Tabelle 2).

Strukturelle Modelle produzieren meist eine einzige Prognose zum ersten Zeitpunkt, zu dem alle Daten verfügbar sind. Jedoch lassen sich mit ihrer Hilfe auf Basis vorläufiger Da-

<sup>35</sup> Für Modelle, die nur den Koalitionsstimmenanteil prognostizieren, ist dieser die absolute Abweichung vom tatsächlichen kombinierten Stimmenanteil der CDU/CSU und FDP.

ten auch früher Prognosen erstellen. So berechneten *Norpoth* und *Gschwend* 2013 mittels aktueller Daten nicht nur eine, sondern drei leicht unterschiedliche Vorhersagen.<sup>36</sup> Diese sind in Abbildung 1 dargestellt. Alle basieren auf demselben Modell, variieren jedoch leicht, da sie jedes Mal aktualisierte Umfragedaten (für die Kanzlerpopularität) nutzen. Strukturelle Modelle bieten auch die Möglichkeit, Szenarien zu berechnen. Anstatt zu warten, bis neue Daten verfügbar werden, können für einen Bereich wahrscheinlicher Werte der Prädiktoren, zum Beispiel dem (wiederum prognostizierten) Wirtschaftswachstum, Vorhersagen gemacht werden.

Nur wenn die Wahl näher rückt, können wir mehr Vertrauen in die prognostische Validität von Wahlabsichtsumfragen setzen, aber auch dann bieten Prognosemodelle noch vergleichbare oder gar bessere Vorhersagen (vgl. Tabelle 2). Das gilt besonders, wenn man die Prognosen mit den zum Zeitpunkt ihrer Veröffentlichung verfügbaren Umfragen vergleicht (vgl. Abbildung 1).

Hier ist zu erwähnen, dass die Meinungsforschungsinstitute 2013, anders als in vorherigen Wahlen, schon früh überraschend genaue Prognosen lieferten, wenn man die Umfrageergebnisse zur Wahlintention als solche interpretiert. Die hinter dem „Politbarometer“ des ZDF stehende Forschungsgruppe Wahlen beispielsweise veröffentlichte seit Mai 2013 Werte für die Stimmenanteile der Koalitionsparteien, die sich zwischen 45 und 47 Prozent bewegten. Die Umfrageergebnisse im Vorfeld zukünftiger Wahlen müssen jedoch nicht so stabil bleiben, da sie durch kurzfristige Ereignisse beeinflusst werden können.<sup>37</sup>

Eine Erklärung für diese durchaus verblüffende Leistung wissenschaftlicher Prognosemodelle ist, dass die Wähler im Verlauf einer Kampagne Informationen über die so genannten fundamentals (insbesondere die wirtschaftliche Lage) erhalten und diese an Wichtigkeit für ihre Wahlabsicht zunehmen.<sup>38</sup> Als Vorhersagen interpretierte Umfragen weit vor einer Wahl können diesen Effekt nicht berücksichtigen.

#### 4. Probleme von Prognosemodellen

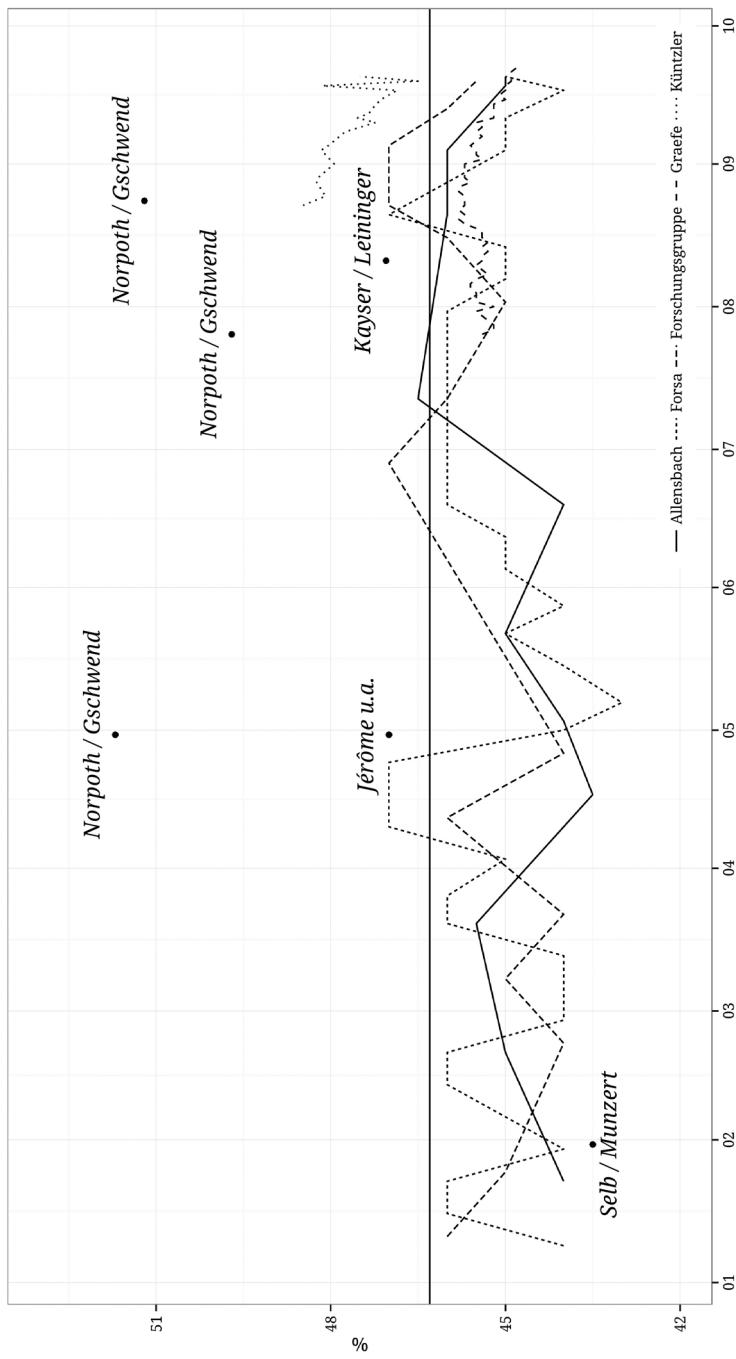
Modellbasierte Prognosen scheinen also zum Zeitpunkt ihrer Veröffentlichung in ihrer prognostischen Güte den zum selben Zeitpunkt publizierten Umfragen in nichts nachzustehen. Jedoch stützt sich diese Feststellung nur auf die letzten drei Bundestagswahlen. Erst für die Wahl 2013 stand eine nennenswerte Zahl unterschiedlicher Prognosen zur Verfü-

36 Vgl. *Helmut Norpoth* / *Thomas Gschwend*, a.a.O. (Fn. 11). Siehe auch *dies.*, Kanzlermodell sagt Wiederwahl von Merkel voraus, in: Zeit Online, Blog Zweitstimme vom 26. Juli 2013, <http://blog.zeit.de/zweitstimme/2013/07/26/kanzlermodell-sagt-wiederwahl-von-merkel-voraus/>; *dies.*, Mach's noch einmal, Mutti, in: Zeit Online, Blog Zweitstimme vom 24. August 2013, <http://blog.zeit.de/zweitstimme/2013/08/24/machs-noch-einmal-mutti/> (Abruf am 22. Oktober 2015).

37 Man denke nur an die Atomkatastrophe in Fukushima, in deren Folge Bündnis 90/Die Grünen kurzzeitig im Sommer 2011 in der Sonntagsfrage bei über 20 Prozent lagen. Auch jenseits solcher singulärer Vorkommnisse sind größere „Prognosefehler“ in Umfragen keine Seltenheit. 2009 beispielsweise, als die amtierende Regierung aus einer Großen Koalition zwischen CDU/CSU und SPD bestand, überschätzten Umfragen den Stimmenanteil von CDU/CSU und SPD noch kurz vor der Wahl um 4,2 Prozentpunkte.

38 Vgl. *Andrew Gelman* / *Gary King*, Why Are American Presidential Election Campaign Polls so Variable When Votes Are so Predictable?, in: *British Journal of Political Science*, 23. Jg. (1993), H. 4, S. 409 – 451.

Abbildung 1: Zeitreihe der Wahlprognosen für die Summe der von Koalitionsparteien CDU/CSU und FDP zu erwartenden Stimmenanteile (Januar 2013 bis zur Wahl im September 2013)



Anmerkung: Modellbasierte Prognosen sind zum frühestmöglichen Zeitpunkt der Prognose eingetragen. Die Prognosen Graefes und Künzlers sowie die durch Allensbach, Forsa und Forschungsgruppe Wahlen erhobenen Umfragewerte sind, um deren zeitliche Veränderungen darzustellen, als Linien abgetragen. Die horizontale Linie stellt das tatsächliche Wahlergebnis dar.  
Quelle: Eigene Darstellung.

gung – fünf von sechs der in diesem Beitrag besprochenen Modelle kamen zum ersten Mal zum Einsatz. Ob Prognosemodelle im komplexen deutschen Mehrparteiensystem verlässliche Vorhersagen liefern können, lässt sich erst mit größerer Sicherheit nach weiteren Wahlen sagen. Abgesehen von mangelnden Erfahrungswerten gibt es natürlich weitere Gründe, die eine gesunde wissenschaftliche Skepsis gegenüber Prognosemodellen nahelegen.

Dieser Beitrag leitete ein mit einer Kritik an der Interpretation von Umfragewerten als Wahlprognosen. Solch eine Interpretation setzt entweder voraus, dass sich Wahlabsichten im Verlauf der Kampagne nicht ändern oder dass sich dieses und weitere Probleme durch Gewichtung der Rohdaten beheben lassen. Diese Annahmen sind meist nur schwer haltbar. Jedoch basieren natürlich auch Prognosemodelle auf Annahmen, die ebenso in Frage gestellt werden können. Regressionsmodelle sind in Deutschland wie auch etwa in den USA das vorherrschende Mittel zur Erstellung von Wahlprognosen. Kritiker bemängeln, dass solche Modelle zwar für Situationen mit vielen Datenpunkten und wenigen Variablen geeignet sind, nicht jedoch für Situationen mit wenigen Datenpunkten und vielen potenziellen Variablen, wie sie im Falle der Wahlprognose typisch sind.<sup>39</sup> Die Annahme, dass die in einem statistischen Modell geschätzten Zusammenhänge so auch in Zukunft stabil sein werden, darf durchaus als gewagt bezeichnet werden.

Urheber struktureller Modelle mögen dem entgegenhalten, dass die statistischen Modelle natürlich nicht für sich stehen, sondern durch Theorien der Wahlforschung motiviert sind. Diese Theorien, die durch empirische Forschung gestützt werden, begründen, dass es sich um dauerhaft geltende Zusammenhänge handelt. Hier ist jedoch anzumerken, dass das Gros dieser Theorien und empirischen Anwendungen auf der Individualebene angesiedelt ist. Ebenso wie beim Schluss von Zusammenhängen auf der Aggregatebene auf solche der Individualebene<sup>40</sup> höchste Vorsicht angezeigt ist, ist dies auch umgekehrt der Fall – gerade wenn es darum geht, Vorhersagen zu treffen. Es bedarf überzeugender Begründungen, warum und wie Zusammenhänge auf der Individualebene auf die Aggregatebene übertragen werden können.

Urheber von Umfrageaggregationen und synthetischen Modellen gehen mit dem Problem geringer Fallzahlen anders um. Sie setzen auf Daten aus mehreren Quellen, um ihre Datenbasis zu erweitern. Die politikwissenschaftlich-theoretische Komponente verliert in diesen Modellen an Bedeutung.

Es ist jedoch gerade der Bezug auf politikwissenschaftliche Theorien, der strukturelle Modelle auch dann noch nützlich macht, wenn sie „mal daneben liegen“. So wichen *Norpoths* und *Gschwends* Prognosen des Stimmenanteils der Koalitionsparteien CDU/CSU und FDP bei der Bundestagswahl 2013 um fast fünf Prozentpunkte vom tatsächlichen Ergebnis ab. Die Autoren führen dies auf den unvorhergesehenen Erfolg der Alternative für Deutschland sowie auf die außergewöhnliche Popularität Kanzlerin *Merkels* zurück. *Angela*

39 *Andreas Graefe* beispielsweise rät daher zu einem einfacheren Verfahren der Mittelwertbildung das allen Prädiktoren gleiches Gewicht zumisst. Dies erlaubt die Inklusion beliebig vieler Prädiktoren. Vgl. *J. Scott Armstrong*, Illusions in Regression Analysis, in: *International Journal of Forecasting*, 28. Jg. (2012), H. 3, S. 689 – 694; *ders.* / *Kesten C. Green* / *Andreas Graefe*, Golden Rule of Forecasting: Be Conservative, in: *Journal of Business Research*, 68. Jg. (2015), H. 8, S. 1717 – 1731; *Andreas Graefe*, Improving Forecasts Using Equally Weighted Predictors, in: ebenda, S. 1792 – 1799.

40 Hierbei handelt es sich um einen ökologischen Fehlschluss.



*Merkel* war 2013 weit über das eigene Lager hinaus beliebt, weshalb das Maß der Kanzlerpopularität 2013 weniger Aussagekraft als Prädiktor für Wahlintentionen gehabt habe.<sup>41</sup>

Eine ähnliche Situation – das Aufkommen des Wahlbündnisses zwischen PDS und WASG im Vorfeld der Wahl 2005 – veranlasste *Norpoth* und *Gschwend* damals, eine einmalige Korrektur ihres Modells vorzunehmen und die Kanzlerpopularität um den zu erwartenden Stimmenanteil der PDS/WASG zu bereinigen. Während sie solche Anpassungen als „praktische Ausnahmeregelungen, die die Logik des Modells nicht in Frage stellen“<sup>42</sup>, beschreiben, ist *Markus Klein* im gleichen Heft der Politischen Vierteljahresschrift deutlich kritischer.<sup>43</sup> Er argumentiert, *Norpoths* und *Gschwends* Prognose könne nicht als Erfolg betrachtet werden, da das Modell augenscheinlich nicht in der Lage sei, allgemeingültige Vorhersagen zu produzieren. Die Anpassung sei zudem aus theoretischer und empirischer Sicht nicht überzeugend. Dessen ungeachtet können, wie eingangs schon dargestellt, Prognosemodelle ein erwartetes ‚normales‘ Wahlergebnis bereitstellen, vor dessen Hintergrund die Wahl beurteilt werden kann; und so können sie auch zur Theoriebildung beitragen.

Wie beispielsweise die Wahlprognosen für 2005 und 2013 illustrieren, scheint es, dass in einem fraktionalisierten Parteiensystem die Popularität des Regierungschefs weniger Vorhersagekraft für den Wahlentscheid besitzt. Diese Hypothese, abgeleitet aus einem Prognosemodell, ließe sich in einem klassischeren Forschungsdesign systematisch testen. Ebenso lassen sich am „erfolgreichen Scheitern“ der Prognose an der Linkspartei 2005 Hypothesen bilden. Das Aufkommen neuer Parteien ist jedoch ein Problem, das Wahlprognosemodelle in den Griff bekommen müssen, wenn sie zuverlässige Prognosen in Mehrparteiensystemen generieren sollen.

Ein Problem mit Vorhersagen wie auch mit Umfragen tritt auf, wenn diese ohne die entsprechenden Schwankungsbreiten kommuniziert werden, und somit die ihnen inhärente Unsicherheit nicht deutlich gemacht wird. Prognosemodelle sollten keinen falschen Eindruck von Genauigkeit erwecken, indem sie auf die Angabe von Schwankungsbreiten verzichten. Prognosewerte sind ebenso wie Umfragen mit Unsicherheit behaftet, die sich jedoch mit den zu Grunde liegenden statistischen Methoden quantifizieren lässt. So kann eine Schwankungsbreite des Prognosewertes analog zu den aus Umfragen bekannten berechnet werden, innerhalb dessen sich das tatsächliche Wahlergebnis höchstwahrscheinlich bewegt. Daraus lässt sich auch die Wahrscheinlichkeit ableiten, dass die amtierende Regierung im Amt verbleibt. Die Darstellung der Wahlprognosen als Wahrscheinlichkeiten ist nützlich, weil sie die Aufmerksamkeit nicht nur auf die Punktschätzung der voraussichtlichen Sieger, sondern auch auf die den Prognosen inhärente Unsicherheit lenkt.<sup>44</sup> Diese sollten daher unbedingt analog zu Umfragen Maße zur Quantifizierung der ihnen inhärenten Unsicherheit ausweisen.

41 Vgl. *Helmut Norpoth* / *Thomas Gschwend*, A Near Miss for the Chancellor Model, in: EU Political Economy Bulletin, Nr. 17 (2014), S. 4 – 8.

42 *Thomas Gschwend* / *Helmut Norpoth*, Prognosemodell auf dem Prüfstand: Die Bundestagswahl 2005, in: PVS, 46. Jg. (2005), H. 4, S. 682 – 688, S. 684.

43 Vgl. *Markus Klein*, Die Zauberformel. Über das erfolgreiche Scheitern des Prognosemodells von Gschwend und Norpoth bei der Bundestagswahl 2005, in: PVS, 46. Jg. (2005), H. 4, S. 689 – 691.

44 Vgl. *Andrew Leigh* / *Justin Wolfers*, Competing Approaches to Forecasting Elections: Economic Models, Opinion Polling and Prediction Markets, in: Economic Record, 82. Jg. (2006), H. 258, S. 325 – 340, S. 328.

Darüber hinaus haben Modelle, die den Koalitionsstimmenanteil vorhersagen, ein weiteres fundamentales Problem. Sie waren nicht in der Lage, 2013 das Ausscheiden der FDP aus dem Bundestag vorherzusehen. Modelle, die die individuellen Stimmenanteile der Parteien prognostizieren, haben hier im Prinzip einen Vorteil, da sie Prognosen für alle Parteien liefern; jedoch erwarteten alle Modelle, dass die FDP die Fünf-Prozent-Hürde überspringen würde. Ebenso sahen sämtliche Umfragen die FDP bei mindestens fünf Prozent.

Ein weiterer Aspekt muss Erwähnung finden, der wiederum Prognosen und Umfragen gleichermaßen trifft: Die Veröffentlichung von Umfragedaten noch unmittelbar vor der Wahl 2013 war umstritten, da die Ergebnisse das Wahlverhalten der Bürger beeinflussen könnten. Die Wirkung von Wahlumfragen auf das Wahlverhalten ist daher auch ein Thema in der Wahlforschung, wie *Hanna Hoffmann* in ihrem Beitrag in diesem Heft der ZParl darstellt. Für Wahlprognosen ist von einer solchen Wirkung zumindest momentan nicht auszugehen – dafür finden sie noch zu wenig öffentliche Beachtung. Bei Prognosemodellen, die explizit für sich in Anspruch nehmen, den Ausgang einer Wahl exakt vorherzusagen, scheint dieses Thema jedoch besonders relevant. So stellt sich die Frage, ob manche Wähler der Wahlurne fernbleiben werden, wenn das Ergebnis schon bekannt ist. Wie auch für Umfragen gilt hier, dass Prognosen Informationen bereitstellen, die Wähler in ihre Wahlentscheidung einbeziehen können oder auch nicht. In der Tat könnten Vorhersagen, sofern sie einen knappen Wahlausgang prognostizieren, die Wahlbeteiligung auch steigern. Eine Voraussetzung dafür wäre, dass über die in Prognosen inhärente Unsicherheit besser informiert und sie entsprechend stärker beachtet wird. Hier gibt es wie auch bei Umfragen noch Verbesserungspotential in der Kommunikation der Umfrage- beziehungsweise Prognosewerte.

Nicht nur aufgrund begrenzter Erfahrungswerte, sondern natürlich auch angesichts der angeführten Kritikpunkte sind Prognosemodelle keine Alternative zu Umfragen; sie stellen aber eine sinnvolle Ergänzung zur klassischen Sonntagsfrage dar.

### 5. Fazit: Prognosemodelle als sinnvolle Ergänzung von Umfragen

Das Wahlergebnis ist schon deutlich vor seinem Eintreten von erheblichem öffentlichen Interesse, jedoch musste sich die interessierte Öffentlichkeit bisher primär auf Expertenmeinungen und Bevölkerungsumfragen stützen, wollte sie sich vor der Wahl eine Meinung über den möglichen Ausgang bilden. Mit der Bundestagswahl 2013 gab es erstmals eine nennenswerte Anzahl wissenschaftlicher Prognosemodelle in Ergänzung zu Umfragen und Experteneinschätzungen.

Festzuhalten ist, dass, anders als die meisten Umfragen, Prognosemodelle explizit als Vorhersagen intendiert sind. Die hier vorgestellten Modelle unterscheiden sich von einfachen Meinungsumfragen, da sie sich auf Basis der Analyse vorangegangener Wahlen und auf teils weit vor der Wahl zur Verfügung stehende Daten stützen. Wiederum anders als Expertenmeinungen folgen sie wissenschaftlichen Prinzipien der intersubjektiven Nachvollziehbarkeit und Reproduzierbarkeit. Diese Modelle haben sich in ihrer Genauigkeit überwiegend genauso gut, manchmal besser als traditionelle Meinungsumfragen bewährt.

Dies kann im Hinblick auf zukünftige Prognosen im Allgemeinen und Voraussagen für deutsche Bundestagswahlen im Besonderen optimistisch stimmen – mehr jedoch nicht.

Während in den USA aufgrund langjähriger Erfahrungen mit einer größeren Zahl an Modellen die Zuverlässigkeit von Prognosemodellen als relativ gesichert gelten kann, ist dies in Deutschland noch nicht der Fall. Die Zukunft wird zeigen, ob in einem komplexeren Mehrparteiensystem wie dem deutschen, im Unterscheid zum US-amerikanischen Zweiparteiensystem, auf Dauer verlässliche Prognosen möglich sind.

Gleichzeitig verändert sich das Feld der Wahlprognosen auch in Deutschland rasch. Ältere strukturelle Modelle, in diesem Beitrag besonders hervorgehoben, verlieren an Boden gegenüber synthetischen Modellen, die Umfragen und strukturelle Modelle, manchmal auch Experteneinschätzungen oder Prognosemärkte, kombinieren. Synthetische Modelle finden sowohl in den USA<sup>45</sup> als auch in Europa<sup>46</sup> vermehrt Verbreitung.

Angesichts dieser Entwicklung ist die Frage berechtigt, ob strukturelle Prognosen mit ihren offensichtlichen Mängeln mehr sein können als Inputs für synthetische Modelle, besonders da sich letztere wie beispielsweise der PollyVote als akkurater als ihre einzelnen Komponenten erweisen. Für einen Zweck bleiben jedoch Strukturmodelle im Vorteil: das prädiktive Testen von Theorien des Wahlverhaltens. Umfragewerte, ob aggregiert oder für sich genommen, versuchen nicht zu verstehen, warum die Wähler sich verhalten, wie sie es tun, sondern sie messen nur Anteilswerte in der Bevölkerung. Zumindest ist dies der Fall für die hier vorgestellten Aggregationsmodelle. Bei Umfragen mögen theoretische Überlegungen hinter den Gewichtungen stehen; diese bleiben jedoch im Verborgenen, da die Gewichtungsmethoden nicht publik gemacht werden.

Strukturelle Modelle sind explizit theoriebasiert und zielen darauf ab, Theorie und Empirie der Wahlforschung prädiktiv nutzbar zu machen. Sie werden durch spezifische Entwicklungen während des Wahlkampfes weniger beeinträchtigt als Meinungsumfragen. Natürlich laufen Prognosemodelle dadurch auch Gefahr, durch idiosynkratische Ereignisse kurz vor der Wahl ad absurdum geführt zu werden. Theoriebasierte Prognosen können aber, positiv betrachtet, eine Art erwartetes normales Wahlergebnis bereitstellen, auf dessen Grundlage die tatsächliche Wahl beurteilt werden kann. So haben Prognosen selbst dann einen Erklärungswert, wenn sie in manchen Fällen irren. All dies spricht dafür, dass sich die Politikwissenschaft stärker mit ihnen beschäftigt.

Wahlprognosen, wie sie hier präsentiert wurden, stellen damit eine sinnvolle Ergänzung, jedoch sicher nicht eine Alternative im Sinne eines Entweder-oder zu den traditionellen Ansätzen der Expertenmeinungen und Umfragen dar. Es ist anzunehmen, dass in Zukunft noch weitere Prognosemodelle hinzukommen werden, denn mit jeder neuen Wahl erweitert sich die Datenbasis für Vorhersagen, für Modelle und damit auch für prädiktive Tests elektoraler Theorien.

45 Vgl. *Drew A. Linzer*, Dynamic Bayesian Forecasting of Presidential Elections in the States, in: *Journal of the American Statistical Association*, 108. Jg. (2013), H. 501, S. 124 – 134; *Robert S. Erikson / Christopher Wlezien*, Forecasting US Presidential Elections Using Economic and Noneconomic Fundamentals, in: *PS: Political Science & Politics*, 47. Jg. (2014), H. 2, S. 313 – 316.

46 Vgl. *Andreas Graefe*, a.a.O. (Fn. 22); *Michael S. Lewis-Beck / Ruth Dassonneville*, Forecasting Elections in Europe: Synthetic Models, in: *Research & Politics*, 2. Jg. (2015), H. 1, S. 1 – 11.