

Wissenschaftliche Wahlprognosen – Alternative oder Ergänzung zu Umfragen?¹

Arndt Leininger

Hertie School of Governance

30. September 2015

Beitrag für einen Themenschwerpunkt der *Zeitschrift für Parlamentsfragen* zum Thema „Wahlen, Demoskopie und Demokratie“. PRE-PRINT.

Abstract

Vor der Bundestagswahl 2013 zogen Vorhersagen über das Ergebnis der Wahl wieder viel Aufmerksamkeit auf sich. Die seriöseren Vorhersagen basierten dabei nahezu durchweg auf Meinungsumfragen. Diese sind jedoch Momentaufnahmen, wie die Meinungsforschungsinstitute stets selbst, aber meist ungehört, betonen, und weniger Vorhersagen des Wahlergebnisses, besonders wenn die Wahl noch weit entfernt ist. Zugleich wurde zur Bundestagswahl 2013 eine bisher nicht dagewesene Vielfalt an wissenschaftlich fundierten Prognosen veröffentlicht, die explizit darauf ausgerichtet waren, das Wahlergebnis, teilweise Monate vor der Wahl, vorherzusagen. Dieser Beitrag bietet einen Überblick über diese Modelle. Ein Schwerpunkt liegt dabei auf sogenannten „strukturellen Modellen“, welche Theorie und Empirie der Wahlforschung prädiktiv nutzen. Weiterhin werden Methoden zur Aggregation von Umfragen sowie synthetische Modelle, die strukturelle Modelle mit Umfragenaggregation verbinden, angesprochen. Der Beitrag diskutiert die Leistungsfähigkeit solcher Ansätze im Kontext der Bundestagswahl 2013. Ein besonderer Fokus liegt dabei auf der Frage, welchen Beitrag wissenschaftliche Prognosemodelle zur Wahlforschung leisten können.

In the run-up to the German federal election of 2013, forecasts about who would win attracted considerable attention. Leaving casual punditry aside, the more serious forecasts were predominantly based on opinion polls. Polls offer a good snapshot in time, but when the election is distant, they predict outcomes less well – a point stressed by pollsters themselves but often ignored by the media and the wider public. However, the 2013 election also saw an unprecedented number of scientific forecasting models that explicitly aimed at forecasting the election result, in some cases months in advance. In this article I provide an overview of scientific election forecasting in Germany focusing on so-called “structural models.” These utilize theory and empirics of electoral research for their predictions. Furthermore, I cover survey aggregation methods as well as synthetic models that combine these with structural models. I discuss the performance of such approaches in the context of the general election in 2013 and address the question what, if anything, scientific election forecasting can contribute to electoral research.

¹ 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 TeilnehmerInnen der *41st 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.

1. Einleitung

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 PolitikerInnen, JournalistInnen und individuelle WählerInnen, bisher jedoch noch eher selten PolitikwissenschaftlerInnen. Die seriöseren Vorhersagen basierten daher bisher nahezu durchweg auf Meinungsumfragen. Sicherlich verbessern diese den Informationsgehalt der von MedienexpertInnen 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 gibt schon zu, 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. Umfragewerte 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 in der Gewichtung zum Einsatz kommen, werden durch die Institute allerdings nicht veröffentlicht.

Diese Gewichtung kann Umfragewerte in die Nähe von Prognosen rücken, wenn es im Vorfeld von Wahlen auch darum geht, Wahlintentionenwerte zu generieren, die möglichst nahe am Wahlergebnis liegen werden. 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. MeinungsforscherInnen selbst werden nicht müde, dies zu betonen, jedoch wird ihre Warnung oft von den Medien und einer breiteren Öffentlichkeit ignoriert.

Zugleich bot sich zur Bundestagswahl 2013 eine zwar weiterhin bescheidene, jedoch bisher nicht dagewesene Vielfalt an wissenschaftlich fundierten Prognosen, welche 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. Im Kontext dieses Themenheftes „Wahlen, Demoskopie und Demokratie“ sollen Wahlprognosen einer kritischen Betrachtung unterzogen werden.

Erstmals eingesetzt wurden wissenschaftliche Wahlprognosen ebenso wie Meinungsumfragen bei US-Präsidentschaftswahlen – dort 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.² Während in den USA mittlerweile eine stattliche Anzahl von Modellen entwickelt wurde, die dort recht gut funktionieren³, 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. Dabei sollen zunächst die Probleme bisheriger, vor allem auf Meinungsforschung basierender Vorhersagen diskutiert (2.) und die verschiedenen deutschen Prognosemodelle vorgestellt werden (3). Im Folgenden geht es dann am Beispiel der Bundestagswahl 2013 um die Frage, wie sich Modelle untereinander und mit Umfragen vergleichen lassen und wie akkurat die Vorhersagen der Prognosemodelle im Vergleich zu Umfragen sind (4.). Daran schließt sich eine kritische Diskussion der Möglichkeiten und Grenzen wissenschaftlicher Prognosemodelle an (5). Der letzte Abschnitt fasst zusammen und bietet einen Ausblick.

2. Probleme der Meinungsforschung und das Potential von Prognosemodellen

Warum sollten wir uns mit Wahlprognosen befassen? Zum einen sind Wahlprognosen aus vielen der Gründe, welche Prognosen generell interessant machen, attraktiv. Prognosen 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 PolitikwissenschaftlerInnen interessiert, wird darin 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.

² *Mark Blumenthal*. „Polls, Forecasts, and Aggregators“. In: PS: Political Science & Politics 47.02 (2014), S. 297–300; *Simon Jackman*. „The Predictive Power of Uniform Swing“. In: PS: Political Science & Politics 47.02 (2014), S. 317–321; *Michael W. Traugott*. „Public Opinion Polls and Election Forecasting“. In: PS: Political Science & Politics 47.02 (2014), S. 342–344.

³ siehe z.B. die 13 Prognosen in der Oktoberausgabe 2012 von PS: Political Science and Politics

Prognosen über den Wahlausgang mussten sich dabei jedoch bisher fast exklusiv auf Einschätzungen von ExpertInnen – von Politischen InsiderInnen oder KommentatorInnen – oder Meinungsumfragen stützen. Politische InsiderInnen besitzen per definitionem exklusives Wissen, jedoch sind ihre Prognosen ideologisch voreingenommen. Das heißt, sie wollen, dass ihre Seite gewinnt. Politische KommentatorInnen – JournalistInnen, EssayistInnen, Think-Tanks – mögen weniger parteiisch sein, ihre auf Erfahrung basierenden Ad-hoc-Prognosen folgen jedoch keiner Systematik, sodass auch sie selbst ihre Prognosen wahrscheinlich nicht reproduzieren könnten.

Andere Formen des Prognose können also das Spektrum der Vorhersagen 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. Schauen wir uns zum Beispiel die Landtagswahl in Niedersachsen im Januar 2013 an, um die mit solchen Umfragen verbundenen Schwierigkeiten zu illustrieren. Aufgrund der Umfragen vor der Wahl wurde der Anteil der FDP-Stimmen um nicht weniger als 5 Prozentpunkte unterschätzt, das entspricht nahezu der Hälfte des tatsächlichen Stimmenanteils der Partei von 9,9 Prozent. Solche drastischen Abweichungen sind jederzeit möglich, jedoch nur schwer vorhersagbar. Für die letzten fünf Bundestagswahlen (1998-2013) wichen die durch die Forschungsgruppe Wahlen für das ZDF erstellten Politbarometer-Projektionen des Stimmenanteils der bisherigen Koalitionsparteien zwei Monate (ein Monat) vor der Wahl jedoch nur um durchschnittlich 2,6 (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 liegt⁴. Dies liegt auch daran, dass die Gruppe der sogenannten Spätentschlossenen von Wahl zu Wahl gewachsen ist.⁵ Daher publizieren die großen Umfrageinstitute mittlerweile noch bis kurz vor der Wahl neue Umfragewerte – auch das ZDF und die Forschungsgruppe Wahlen publizierten 2013 zum ersten Mal noch drei Tage vor der Wahl eine Umfrage.

DemoskopInnen weisen zwar Schwankungsbreiten zur Quantifizierung der Ungenauigkeit der errechneten Schätzwerte aus, diese können jedoch nicht zukünftige Meinungsänderungen 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 Umfrageinstitute ausgegebenen Schwankungsbreiten

⁴ Groß, Jochen. „Die Prognose von Wahlergebnissen: Ansätze und empirische Leistungsfähigkeit“. Wiesbaden: VS Verlag für Sozialwissenschaften, 2010.

⁵ Rüdiger Schmitt-Beck. „Empirische Wahlforschung in Deutschland: Stand und Perspektiven zu Beginn des 21. Jahrhunderts“. In: *Wählen in Deutschland*. Hrsg. von Rüdiger Schmitt-Beck. 1. Aufl. Politische Vierteljahresschrift : Sonderheft 45. Baden-Baden: Nomos, 2012, S. 2–39; Thomas Plischke. Wann Wähler entscheiden. Abläufe von Entscheidungsprozessen und der Zeitpunkt der Wahlentscheidung. Baden-Baden: Nomos, 2014.

der Umfragewerte lag.⁶

Es bleibt also festzuhalten, dass Umfragen, besonders wenn die Wahl 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. Wissenschaftliche Prognosen sind damit eine für die interessierte Öffentlichkeit relevante Ergänzung zu Umfragen, wenn es um Prognosen des Ausgangs von Wahlen geht. Sie können so auch zu einem reflektierteren Umgang mit Umfragen anregen und Aufmerksamkeit für politikwissenschaftliche Forschung schaffen.

Welche fachlichen Gründe sprechen aber dafür, sich als PolitikwissenschaftlerIn 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 Prospektive der Prognose eng zusammen. Die meisten PolitikwissenschaftlerInnen analysieren Wahlen nicht, um Besonderheiten einzelner Wahlen hervorzuheben, sondern um allgemeine Muster und auch kausale Mechanismen aufzudecken. Warum also nicht diese Forschung für Prognosen nutzen? Wenn Forschung auf Basis vorangegangener Wahlen allgemeine Muster aufgedeckt hat, sollten sich diese auch in die Zukunft fortschreiben lassen. Prognostische Validität ist ein Eckpfeiler der Wissenschaft, schwieriger zu erreichen als Erklärung, und entsprechend nach Meinung einiger AutorInnen oft in den Sozialwissenschaften vernachlässigt.⁷ Wahlprognosen sind daher vielversprechend, weil sie theoriebasiert sind und darauf abzielen, die Grundlagen der durch die Wahlforschung belegten Wahltheorien zu erfassen. Mit ihnen lassen sich Theorien auf den denkbar härtesten Prüfstand stellen – die Vorhersage zukünftiger Wahlergebnisse.⁸

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, sodass Prognosen selbst dann einen Erklärungswert haben, wenn sie in manchen Fällen daneben liegen. Wahlprognosen stellen so einen Referenzrahmen dar, vor dem die Auswirkungen als besonders erachteter Ereignisse, die in jeder Kampagne auftreten, bewertet

⁶ *Rainer Schnell und Marcel Noack.* „The Accuracy of Pre-Election Polling of German General Elections“. In: *methods, data, analyses* 8.1 (2014), S. 5–24.

⁷ *Alexander Rosenberg.* *Economics - Mathematical Politics Or Science of Diminishing Returns?* University of Chicago Press, 1992; *Philip A. Schrodt.* „Seven deadly sins of contemporary quantitative political analysis“. en. In: *Journal of Peace Research* (2013), S. 287–300.

⁸ *Mark A Kayser und Arndt Leininger.* „A Predictive Test of Voters’ Economic Benchmarking: The 2013 German Bundestag Election“. In: *Working Paper* (2015).

werden können. Darüber hinaus können Wahlprognosen einen eigenen Beitrag zum Forschungsstand leisten. So haben Wahlprognosen selbst wesentlich zum Verständnis des retrospektiven und kurzfristigen Effekts wirtschaftlicher Rahmenbedingungen auf das Wahlergebnis beigetragen.⁹

3. Prognosemodelle für Bundestagswahlen im Vergleich

Prognosemodelle verwenden statistische Methoden, um auf der Basis von vor der Wahl verfügbaren Daten den Ausgang einer Wahl vorherzusagen. Dabei kann man derzeit drei Arten von Prognosemodellen unterscheiden: Von besonderem Interesse sind hier sogenannte „strukturelle Modelle“, welche den Zusammenhang zwischen Wahlergebnissen und erklärenden Faktoren wie der wirtschaftlichen Lage oder der Popularität des oder der amtierenden KanzlerIn in vergangenen Wahlen ermitteln, und auf Basis dessen zukünftige Wahlergebnisse vorherzusagen suchen. Strukturelle Modelle greifen auf Theorie und Empirie der Wahlforschung zurück und machen diese prädiktiv nutzbar – sie sind daher für WahlforscherInnen und an Wahlforschung Interessierte besonders relevant. In Deutschland zählen zu dieser Gruppe die Modelle von Jérôme, Jérôme-Speziari und Lewis-Becks, die für die Wahl 1998 die erste Wahlprognose für eine Bundestagswahl veröffentlichten¹⁰ und für die Wahl 2013 ein neues „Political Economy Model“¹¹ vorlegten, das „Kanzlermodell“ Norpoth und Gschwends¹² sowie das von Kayser und Leininger entwickelte „Benchmarking Modell.“¹³ Daneben werden für Prognosen auch „Aggregationsmodelle“ verwendet, die mehrere Umfragen kombinieren und die in Umfragen inhärenten Prognosefehler zu minimieren suchen. In Deutschland zählen zu diesen Modellen das exklusiv umfragedatenbasierte Prognosemodell von Selb und Munzert¹⁴ sowie der von Andreas Graefe¹⁵ entwickelte „PollyVote“, der sich auf verschiedenste Quellen stützt. Außerdem existieren sogenannte „synthetische Modelle“, die strukturelle Modelle mit Umfragenaggregation verbinden. Hier ist das Modell von Theresa Küntzler zu nennen,¹⁶ die das „Kanzlermodell“ mit einem eigenen Umfragemodell kombiniert.

⁹ William G. Mayer. „What, If Anything, Have We Learned from Presidential Election Forecasting?“ en. In: PS: Political Science & Politics 47.02 (2014), S. 329–331.

¹⁰ Bruno Jérôme, Véronique Jérôme-Speziari und Michael S. Lewis-Beck. „Prévisions politico - économiques et élections législatives allemandes de septembre 1998“. In: Le Figaro (Juli 1998).

¹¹ Bruno Jérôme, Véronique Jérôme-Speziari und Michael S. Lewis-Beck. „A Political-Economy Forecast for the 2013 German Elections: Who to Rule with Angela Merkel?“ In: PS: Political Science & Politics 46.03 (2013), S. 479–480.

¹² Helmut Norpoth und Thomas Gschwend. „Chancellor Model Picks Merkel in 2013 German Election“. In: PS: Political Science & Politics 46.03 (2013), S. 481–482.

¹³ Mark Kayser und Arndt Leininger. „A Benchmarking Forecast of the 2013 Bundestag Election“. In: (2013).

¹⁴ Peter Selb und Simon Munzert. Forecasting the 2013 Bundestag Election Using Data from Various Polls. SSRN Scholarly Paper ID 2313845. Rochester, NY: Social Science Research Network, 2013.

¹⁵ Andreas Graefe. „German Election Forecasting: Comparing and Combining Methods for 2013“. In: German Politics 24.2 (2015), S. 195–204.

¹⁶ Theresa Küntzler. Using Bayesian Data Combination of Fundamental Variable-Based Forecasts and Poll-Based

Tabelle 1 bietet einen ersten Überblick über die hier vorzustellenden Modelle. Diese unterscheiden sich sowohl nach den Werten, die vorhergesagt werden, der Methodik als auch der zu Grunde liegenden Theorie – die Tabelle nennt hier nur den wichtigsten theoretischen Bezugsrahmen.

Modell	Prognosewert(e)	Methode	Theorie
Strukturelle Modelle			
Jérôme et al.	Parteistimmanteile	6 Regressionsmodelle für 5 Parteien und “Sonstige”	Regierungsparteien werden für wirtschaftliche Lage verantwortlich gehalten und KanzlerkandidatInnen spielen eine herausgehobene Rolle für den Wahlscheid.
Kayser & Leininger	Koalitionsstimmanteil	1 Regressionsmodell für den Stimmanteil der Koalition	Die wirtschaftliche Lage relativ zu anderen EU-Ökonomien, besonders relevant in Wahlen nach der Finanz- und Eurokrise.
Norpoth & Gschwend	Koalitionsstimmanteil	1 Regressionsmodell für den Stimmanteil der Koalition	Die Popularität der/s AmsinhaberIn ist indikativ für die Bewertung der Arbeit der Regierung und damit den Wahlscheid.
Aggregationsmodelle & Synthetische Modelle			
Graefe	Parteistimmanteile	Mittelwertbildung über Prognosen der anderen Prognosemodelle, Umfragen, Experteneinschätzungen und Wettmärkte	Keine Prognosemethode ist systematisch besser als andere. Der Mittelwert aller Methoden kann daher dennoch akkurater sein als einzelne Methoden.
Selb & Munzert	Parteistimmanteile	1 Regressionsmodell	Es gibt instituts- und parteispezifische Fehler in den Umfragedaten. Diese werden durch ein statistisches Modell geschätzt und ausgeglichen.
Küntzler	Koalitionsstimmanteil	Mittelwertbildung aus zwei Regressionsmodellen	Berechnung eines gewichteten Mittels aus Umfragedaten und dem Kanzlermodell.

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

3.1 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 Umfragen konzipiert worden. Wer Umfragen als Prognosen nutzt, muss die Annahme treffen, dass die darin geäußerten Wahlintentionen bis zur Wahl unverändert bleiben oder sich individuelle Veränderungen im Aggregat ausgleichen. Dies ist aus den obengenannten Gründen eine nur schwerlich haltbare Annahme.

Dabei gehen strukturelle Prognosemodelle natürlich auch von einigen Annahmen aus, allerdings von weniger gewagten. Zunächst setzen solche Prognosemodelle die Annahme 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 EntwicklerInnen einer Wahlprognose zunächst eine Entscheidung treffen, welche Größe prognostiziert werden soll – das Wahlergebnis einer oder mehrerer Parteien oder der Regierungsfraktion. Eine Prognose abzugeben scheint in einem Mehrparteiensystemen wie Deutschland nämlich deutlich schwerer als in einem Zweiparteiensystem wie den USA, in deren Kontext Wahlprognosen für Präsidentschaftswahlen zunächst entwickelt wurden.¹⁷¹⁸

Die meisten Menschen interessieren sich für Prognosen, weil sie wissen möchten, wer die Wahl gewinnen wird. Wer die Wahl gewinnt, hängt jedoch in einem Mehrparteiensystem auch von den nach der Wahl stattfindenden Koalitionsverhandlungen ab. Daher ist es sinnvoller, nicht die GewinnerInnen und VerliererInnen zu betrachten, sondern sich auf die Stimmenanteile der Parteien zu konzentrieren. Ein Fokus auf den Stimmenanteil der Koalition erlaubt es vorherzusagen, ob dieselbe Regierung an der Macht 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,¹⁹ Kayser und Leininger²⁰ und Küntzler,²¹ letztere da auf dem „Kanzlermodell“ aufbauend, entscheiden sich für Prognosen des Koalitionsstimmanteils. Die Alternative dazu ist, die Stimmenanteile für alle im Parlament vertretenen Parteien zu prognostizieren, wie es Jérôme und Ko-Autoren²² tun. Ebenso prognostizieren die Modelle Selb und Munzerts²³ sowie Graefes²⁴ die Stimmenanteile aller fünf im 17. Deutschen Bundestag vertretenen Parteien und den „Sonstigen.“

Die Prognose der Stimmanteile aller Parteien bietet mehr, aber nicht unbedingt bessere

¹⁷ Michael S. Lewis-Beck und Tom W. Rice. *Forecasting elections*. Washington, D.C.: CQ Press, 1992.

¹⁸ Prognosemodelle der US-Präsidentschaftswahlen sagen üblicherweise den Stimmanteil 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%.

¹⁹ Thomas Gschwend und Helmut Norpoth. „Soll und Haben: Die deutsche Wählerschaft rechnet mit den Parteien ab“. In: *50 Jahre Empirische Wahlforschung in Deutschland. Entwicklung, Befunde, Perspektiven, Daten*. Hrsg. von Markus Klein u. a. Wiesbaden: Westdeutscher Verlag, 2000, S. 389–409.

²⁰ Mark Kayser und Arndt Leininger, „A Benchmarking Forecast of the 2013 Bundestag Election“.

²¹ Theresa Küntzler, *Using Bayesian Data Combination of Fundamental Variable-Based Forecasts and Poll-Based Forecasts to Predict the 2013 German Election*.

²² Bruno Jérôme, Véronique Jérôme-Speziari und Michael S. Lewis-Beck, „A Political-Economy Forecast for the 2013 German Elections“.

²³ Peter Selb und Simon Munzert, *Forecasting the 2013 Bundestag Election Using Data from Various Polls*.

²⁴ Andreas Graefe, „Combined Forecasts of the 2013 German Election“.

Informationen. Sie können leicht addiert werden, um den Koalitionsstimmanteil und so die Wahrscheinlichkeit, dass die Regierung im Amt bleiben kann, zu berechnen. Weiterhin erlauben sie in Wahlsystemen wie dem deutschen, das eine 5%-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 Koalitionsstimmanteil. Die Prognose des Stimmanteils aller relevanter Parteien bietet somit mehr Informationen, beinhaltet jedoch das Risiko ungenauerer Prognosen. Zudem sind mehr Ad-hoc-Annahmen zur Konstruktion von Modellen, die den Stimmanteil einzelner Parteien vorhersagen, zu treffen.

Strukturelle Modelle in ihrer einfachsten Form basieren auf einem einfachen statistischen Modell, einem sogenannten Regressionsmodell, welches die Stimmanteile der Regierungsparteien oder einzelner Parteien mit erklärenden Variablen, z.B. 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.²⁵ Die Wahl von 2013 stellte in dieser Hinsicht einen wichtigen Test dar. Nach einer Wahl lässt sich die Genauigkeit einer Prognose einfach bestimmen. Doch muss man nicht bis zur nächsten Wahl warten, um zu prüfen, ob ein Modell geeignet ist, gute Prognosen zu generieren.

Viele AutorInnen 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 Prognosefehler, 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

²⁵ 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.

Zukunft fortbestehen wird.

Nachdem die Grundprinzipien struktureller Wahlprognosemodelle dargelegt wurden, schließt sich nun die Vorstellung der Modelle an, die diese Methodik auf Bundestagswahlen anwenden. Hier ist zunächst das Autorenteam Bruno Jérôme, Véronique Jérôme-Speziari und Michael Lewis-Beck zu nennen, das bereits viele Prognosen für Wahlen in Frankreich, den USA und anderen Ländern erstellt hat. Sie konstruierten auch das erste Prognosemodell für eine deutsche Bundestagswahl.²⁶ Zentrale Komponente ihres „Political Economy“ Modells, das die Abwahl Helmut Kohls 1998 erfolgreich vorhersagte, war die Arbeitslosenrate in den letzten zwei Quartalen vor der Wahl. Ein um die Wahlintentionen für den kleineren Koalitionsparteien erweitertes Modell sagte 2009 einer Koalition aus CDU/CSU und FDP ein Ergebnis von 51,2% voraus – tatsächlich vereinigte die im Anschluss an diese Wahl auch gebildete Koalition einen Stimmanteil von nur 41,9% auf sich, dieser war jedoch ausreichend für eine parlamentarische Mehrheit.

2013 stellten sie ein neues „Political Economy Model“ vor.²⁷ Dieses besteht aus für jede der fünf großen Parteien sowie der 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.²⁸ Ihre finale Prognose hat eine Vorlaufzeit von einem Monat, das Modell erlaubt aber Prognosen zu früheren Zeitpunkten mit vorläufigen Daten. Seit den Bundestagswahlen 2002 erhielten ihre Prognosen jeweils relativ viel Aufmerksamkeit und haben bei der Vorhersage vergangener Wahlen 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 der amtierenden Kanzlerin (Prozentsatz der Menschen, die die amtierende Kanzlerin dem Herausforderer vorziehen), den langfristigen WählerInnenrückhalt (durchschnittlicher Wahlerfolg der Koalitionsparteien bei den drei vorangegangenen Wahlen) und die Anzahl der Legislaturperioden, während derer die Partei der Kanzlerin an der Regierung war.²⁹

Für die Bundestagswahl 2013 stellten Kayser und Leininger erstmals ihr „Benchmarking Modell“ vor. Sie prognostizierten ebenfalls den Stimmanteil der Koalitionsregierung. Die namensgebende

²⁶ Bruno Jérôme, Véronique Jérôme-Speziari und Michael S. Lewis-Beck, „Prévisions politico - économiques et élections législatives allemandes de septembre 1998“.

²⁷ Bruno Jérôme, Véronique Jérôme-Speziari und Michael S. Lewis-Beck, „A Political-Economy Forecast for the 2013 German Elections“.

²⁸ Thomas Gschwend und Helmut Norpoth, „Soll und Haben: Die deutsche Wählerschaft rechnet mit den Parteien ab“.

²⁹ Thomas Gschwend und Helmut Norpoth. „Wenn am nächsten Sonntag ...“: Ein Prognosemodell für Bundestagswahlen“. In: *Wahlen und Wähler: Analysen aus Anlass der Bundestagswahl 1998*. Hrsg. von Hans-Dieter Klingemann und Max Kaase. Wiesbaden: Westdeutscher Verlag, 2001, S. 473–499.

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 WählerInnen den Zustand ihrer Volkswirtschaft im Vergleich zu dem der Volkswirtschaften anderer Länder beurteilen, da die Wahrnehmung der wirtschaftlichen Lage vor allem durch Medien vermittelt seien.³⁰ 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 war. Weiterhin inkludieren sie als Prädiktoren den Stimmenanteil der derzeit amtierenden Parteien in der vorherigen Wahl, den Anteil der Befragten, die angeben, eine Bindung an eine der Parteien zu haben, sowie die Anzahl von Legislaturperioden, 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.³¹ 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, denn keines der hier vorgestellten Prognosemodelle kommt ohne Umfragedaten aus. Für die im folgenden Abschnitt dargestellten Modelle sind Umfragedaten eine noch wichtigere, in einem Fall sogar die einzige Datenquelle.

3.2 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

³⁰ 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 (*Mark Andreas Kayser und Michael Peress*. „Benchmarking across Borders: Electoral Accountability and the Necessity of Comparison“. In: *American Political Science Review* 106.03 [2012], S. 661–684).

³¹ *Harald Schoen und Cornelia Weins*, „Der sozialpsychologische Ansatz zur Erklärung von Wahlverhalten.“ In: *Handbuch Wahlforschung*. Hrsg. von Jürgen W. Falter und Harald Schoen. Wiesbaden: VS Verlag für Sozialwissenschaften, 2005, S. 187-242.

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 ein und kann so tagesaktuelle Vorhersagen liefern. Küntzler konstruiert ein synthetisches Modell, das ein gewichtetes Mittel aus der Prognose des „Kanzlermodells“ und einem eigenen umfragedatenbasierten Modell als Prognose berechnet.

Der Ansatz Selb 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 institutsspezifische Verzerrungen zu eliminieren und so zu einer genaueren Prognose zu gelangen, als einzelne Umfragen. Interessanterweise gelangen sie zu dem Ergebnis, 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.

Eine weitere, simplere, jedoch nicht minder effektive Methode der Umfrageaggregation ist der von Andreas Graefe entwickelte „PollyVote“.³² Dieser Ansatz kombiniert vier Kategorien von Prognosen – Umfragen, Prognosemärkte, ExpertInnenurteile und die im vorangegangenen Abschnitt vorgestellten Modelle. Er wurde von Graefe bereits seit 2004 für die Prognose von US-Präsidentenwahlen eingesetzt.³³ 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 der Prognose böten.³⁴ 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.³⁵

Küntzler präsentiert, erstmals in Deutschland, ein aus einer Kombination eines strukturellen Modells und eines umfragebasierten Modells bestehendes dynamisches Prognosemodell. Dieses

³² *Andreas Graefe*, „Combined Forecasts of the 2013 German Election“.

³³ *Andreas Graefe* u. a. „Combining forecasts: An application to elections“. In: *International Journal of Forecasting* 30.1 (2014), S. 43–54.

³⁴ *Andreas Graefe* u. a. „Limitations of Ensemble Bayesian Model Averaging for Forecasting Social Science Problems“. In: *International Journal of Forecasting* forthcoming (2015).

³⁵ *Andreas Graefe* u. a., „Combining forecasts“.

erlaubt es, mit jeder neu verfügbaren Umfrage die Prognose zu aktualisieren. Die erste Komponente ihres Modells 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 Prognoseunsicherheit gewichtetes Mittel aus den beiden Einzelprognosen. Je näher die Wahl rückt, desto mehr Gewicht erhält die umfragebasierte Prognose in der kombinierten Prognose.

4. Die Vorhersagen von Prognosemodellen und Umfragen bei der Bundestagswahl 2013

Nachdem die Logik der verschiedenen Prognosemodelle vorgestellt wurden, stellt sich nun die Frage, welche Prognosemodelle bei der Bundestagswahl 2013 am besten abschnitten. Dies ist jedoch keine einfache Frage, denn die genaueste Vorhersage ist nicht notwendigerweise auch die beste Vorhersage. Denn was ist eine Prognose wert, die exakt zutrifft, aber nur einen Tag vor der Wahl entstand? Eine Prognose, die, sagen wir, zwei Prozentpunkte vom Wahlergebnis abweicht, aber schon acht Wochen vor der Wahl veröffentlicht wurde, kann durchaus als bessere Prognose bezeichnet werden. Hier kann es individuell unterschiedliche Meinungen geben, wie sich die richtige Balance zwischen der Vorlaufzeit und der Genauigkeit darstellt. Daher verzichte ich auf ein Ranking der verschiedenen Prognosen, sondern präsentiere nur eine Zusammenfassung der Prognosen (siehe 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 vorherzusagen. Modelle, die die individuellen Stimmanteile der Parteien prognostizieren, haben hier im Prinzip einen Vorteil, jedoch sahen alle Modelle die FDP bei über 5% - ebenso wie die Institute Allensbach, forsa und die Forschungsgruppe Wahlen die FDP bei jeweils mindestens 5% sahen. Das tatsächliche Ergebnis der FDP lag jedoch innerhalb der Fehlermargen sowohl der Umfragen als aus der Prognosemodelle.

Natürlich hat ein Modell, das mehrere Stimmanteile voraussagt, mehr Potential, daneben zu liegen. Um einen fairen Vergleich der Genauigkeit der Modelle zu bieten, berechne ich daher auch die mittleren absoluten Prognosefehler³⁶ über alle Prognosen (Tabelle 2). Das Modell Selb und Munzerts hat die größte Vorlaufzeit aber einen durchschnittlichen Prognosefehler von 2,82 Prozentpunkten. Das „Benchmarking Modell“ hat den kleinsten Prognosefehler mit 0,75

³⁶ Für Modelle, die nur den Koalitionsstimmanteil prognostizieren, ist dieser die absolute Abweichung vom tatsächlichen kombinierten Stimmanteil der CDU/CSU und FDP.

Prozentpunkten, aber auch nur eine Vorlaufzeit von lediglich zwei Monaten und sagt zudem nur den Koalitionsstimmanteil voraus.

Der limitierende Faktor bei der Vorlaufzeit ist meist die umfragebasierte Variable des Modells, sodass 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 30 Tagen vor der Wahl. Zum Vergleich mit anderen Modellen wird daher der mittlere Prognosefehler über den gesamten Zeitraum herangezogen. Der durchschnittliche Prognosefehler der synthetischen Modelle fällt dabei kleiner aus als der durchschnittliche Fehler der strukturellen Modelle und liegt zwischen diesen und den finalen Umfragen kurz vor der Wahl (Tabelle 2).

	Strukturelle Modelle				Synthetische Modelle		Umfragen			
	Endgültiges Ergebnis	Jérôme et al	Norpoth & Gschwend	Kayser & Leininger	Selb & Munzert	Graefe	Küntzler	Forschungsgruppe Wahlen, 19.09.	Forsa, 20.09.	Allensbach, 20.09.
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
FDP	4,8	1,2			0,6	1		-0,7	0,2	0,7
SPD	25,7	2,3			2,5	0,5		1,3	0,3	1,3
Bündnis 90 / Die Grünen	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

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

Strukturelle Modelle produzieren meist eine einzige Prognose zum ersten Zeitpunkt, zu dem alle Daten verfügbar sind. Jedoch lassen sich auch mittels struktureller Modelle auf Basis vorläufiger Daten frühere Prognosen errechnen. So berechneten Norpoth und Gschwend 2013 mittels aktueller Daten nicht nur eine, sondern drei leicht unterschiedliche Prognosen.^{37,38,39} Diese sind in Abbildung 2 dargestellt. Alle Prognosen basieren auf dem gleichen 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, z.B. dem (widerum prognostizierten) Wirtschaftswachstum, Vorhersagen berechnet werden.

Nur wenn die Wahl näher rückt, können wir mehr Vertrauen in die prognostische Validität von

³⁷ Helmut Norpoth und Thomas Gschwend, „Chancellor Model Picks Merkel in 2013 German Election“.

³⁸ Zeit Blog. Zweitstimme: Kanzlermodell sagt Wiederwahl von Merkel voraus. Juli 2013.

³⁹ Zeit Blog. Zweitstimme: Mach's noch einmal, Mutti. Aug. 2013.

Wahlabsichtsumfragen setzen, aber auch dann bieten Prognosemodelle noch vergleichbare oder gar bessere Vorhersagen (Tabelle 2). Das gilt besonders, vergleicht man die Prognosen mit den zum Zeitpunkt ihrer Veröffentlichung verfügbaren Umfragen (siehe Abbildung 1).

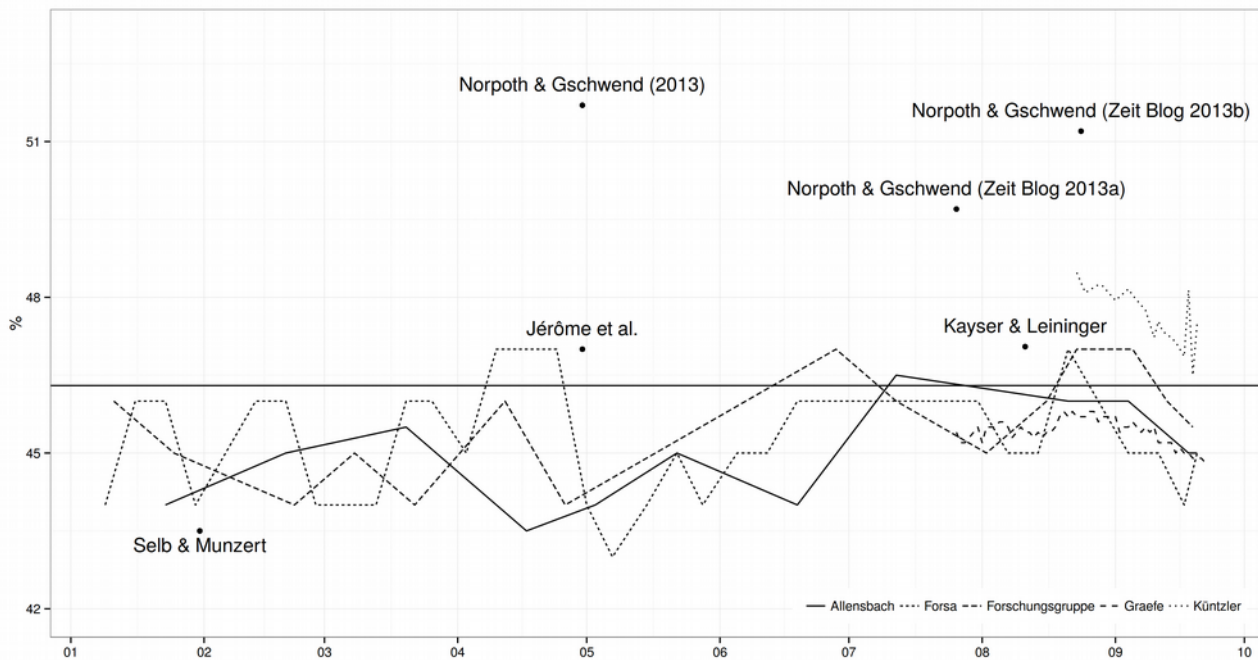


Abbildung 1: Zeitreihe der Wahlprognosen für die Summe der von Koalitionsparteien CDU/CSU und FDP zu erwartenden Stimmanteile – von Januar 2013 bis zur Wahl im September 2013. Modellbasierte Prognosen sind zum frühestmöglichen Zeitpunkt der Prognose eingetragen. Die Prognosen Graefes und Kuntzlers sowie die durch Allensbach, Forsa und Forschungsgruppe Wahlen erhobenen Umfragemerte sind, um deren zeitliche Veränderungen darzustellen, als Linien abgetragen. Die horizontale Linie stellt das tatsächliche Wahlergebnis dar.

Hier ist zu erwähnen, dass die Meinungsforschungsinstitute 2013, anders als in vorherigen Wahlen, schon früh überraschend genaue Prognosen lieferten, wenn man die Umfragemerte zur Wahlintention als solche interpretiert. Die hinter dem „Politbarometer“ des ZDF stehende Forschungsgruppe Wahlen beispielsweise veröffentlichte seit Mai 2013 Werte für die Stimmanteile der Koalitionsparteien, die sich zwischen 45 und 47 Prozent bewegten. Die Umfrageergebnisse im Vorfeld zukünftiger Wahlen müssen jedoch nicht so stabil bleiben, da Umfragemerte durch kurzfristige Ereignisse beeinflusst werden können.⁴⁰

Eine Erklärung für diese durchaus verblüffende Leistung wissenschaftlicher Prognosemodelle ist, dass die Wählerinnen und Wähler im Verlauf einer Kampagne Informationen über die sogenannten ‚fundamentals‘ (insbesondere die wirtschaftliche Lage) gewinnen und diese an Wichtigkeit für ihre Wahlabsicht gewinnen.⁴¹ Als Prognosen interpretierte Umfragen weit vor einer Wahl können diesen

⁴⁰ Man denke nur an die Atomkatastrophe in Fukushima, in deren Folge Bündnis 90/Die Grünen kurzzeitig im Sommer des Jahres 2011 in der Sonntagsfrage bei über 20% lagen. Auch jenseits solcher singulärer Vorkommnisse sind größere 'Prognosefehler' in Umfragen keine Seltenheit. Im Jahr 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.

⁴¹ Andrew Gelman und Gary King. „Why Are American Presidential Election Campaign Polls So Variable When Votes

Effekt nicht berücksichtigen.

5. 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 Verfügung – fünf von sechs in diesem Artikel besprochenen Modelle kamen 2013 zum ersten Mal zum Einsatz. Ob Prognosemodelle im komplexen deutschen Mehrparteiensystem verlässliche Prognosen 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. Diese Kritikpunkte sollen hier aufgezeigt werden, um ein besseres Bild der Möglichkeiten und Grenzen von Prognosemodellen zu erhalten.

Dieser Beitrag leitete ein mit einer Kritik an der Interpretation von Umfragewerten als Wahlprognosen. Solch eine Interpretation setzt entweder voraus, dass sich Wahlintentionen im Verlaufe der Wahlkampagne 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 in anderen Kontexten wie den USA das vorherrschende Mittel zur Erstellung von Wahlprognosen. KritikerInnen 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.⁴² Die Annahme, dass die in einem statistischen Modell geschätzten Zusammenhänge so auch in Zukunft stabil sein werden, darf durchaus auch als gewagt betrachtet werden.

UrheberInnen struktureller Modelle mögen dem entgegenhalten, dass die statistischen Modelle natürlich nicht für sich stehen, sondern durch Theorien der Wahlforschung motiviert sind. Diese Theorien, welche durch empirische Forschung gestützt werden begründen, dass es sich um dauerhaft geltende Zusammenhänge handle. Hier ist jedoch anzumerken, dass das Gros dieser Theorien und empirischer Anwendungen auf der Individualebene angesiedelt ist. Ebenso wie beim

Are So Predictable?“ In: *British Journal of Political Science* 23.04 (1993), S. 409–451.

⁴² Andreas Graefe beispielsweise rät daher zu einem einfacheren Verfahren der Mittelwertbildung das allen Prädiktoren gleiches Gewicht zumisst. Dies erlaubt die Inklusion beliebig vieler Prädiktoren. Siehe: *J. Scott Armstrong*. „Illusions in regression analysis“. en. In: *International Journal of Forecasting* 28.3 (2012), S. 689–694; *J. Scott Armstrong, Kesten C. Green und Andreas Graefe*. „Golden rule of forecasting: Be conservative“. en. In: *Journal of Business Research* (2015); *Andreas Graefe*. „Improving forecasts using equally weighted predictors“. en. In: *Journal of Business Research* (2015).

Schluss von Zusammenhängen auf der Aggregatebene auf Zusammenhänge auf der Individualebene⁴³ 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 Quelle, um ihre Datenbasis zu erweitern. Die politikwissenschaftlich-theoretische Komponente verliert in diesen Modellen an Bedeutung.

Es ist jedoch der Bezug auf politikwissenschaftliche Theorien, der strukturelle Modelle auch dann noch nützlich macht, wenn sie ‚mal daneben liegen.‘ So wichen Norpoth und Geschwends Prognosen des Stimmanteils 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. Merkel war 2013 weit über das eigene Lager hinaus populär, weshalb das Maß der Kanzlerpopularität 2013 weniger Aussagekraft als Prädiktor für Wahlintentionen gehabt habe.⁴⁴

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 Stimmanteil der PSDS/WASG zu bereinigen. Während Norpoth und Gschwend solche Anpassungen als „praktische Ausnahmeregelungen, die die Logik des Modells nicht in Frage stellen“⁴⁵ beschreiben, ist Markus Klein in der gleichen Ausgabe der Politischen Vierteljahresschrift deutlich kritischer.⁴⁶ Letzterer argumentiert, Norpoth 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. Davon unbenommen können, wie eingangs schon dargestellt, Prognosemodelle ein erwartetes ‚normales‘ Wahlergebnis bereitstellen vor dessen Hintergrund die Wahl beurteilt werden kann. So können Wahlprognosemodelle 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 prädiktiv für den

⁴³ Ökologischer Fehlschluss

⁴⁴ *Helmut Norpoth und Thomas Gschwend.* „A Near Miss for the Chancellor Model“. In: *EUSA: EU Political Economy Bulletin* 17 (2014), S. 4–8.

⁴⁵ *Thomas Gschwend und Helmut Norpoth.* „Prognosemodell auf dem Prüfstand: Die Bundestagswahl 2005“. In: *Politische Vierteljahresschrift* 46.4 (2005), S. 682–688, S. 684.

⁴⁶ *Markus Klein.* „Die Zauberformel. Über das erfolgreiche Scheitern des Prognosemodells von Gschwend und Norpoth bei der Bundestagswahl 2005“. In: *Politische Vierteljahresschrift* 46.4 (2005), S. 689–691.

Wahlentscheid ist. Diese Hypothese, abgeleitet aus einem Prognosemodell, ließe sich in einem klassischeren Forschungsdesign systematisch testen. Ebenso können sich am „erfolgreichen Scheitern“ der Prognose an der Linkspartei 2005 Hypothesen bilden lassen. Das Aufkommen neuer Parteien ist jedoch ein Problem, das Wahlprognosemodelle in den Griff bekommen müssen, sollen sie zuverlässige Prognosen in Mehrparteiensystem generieren.

Ein Problem mit Prognosen wie auch mit Umfragen tritt auf, wenn diese ohne die entsprechenden Schwankungsbreiten kommuniziert werden, und somit die ihnen inhärente Unsicherheit nicht kommuniziert 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 auf Basis der zu Grunde liegenden statistischen Methoden quantifizieren lässt. So kann analog zu den aus Umfragen bekannten Schwankungsbreiten eine Schwankungsbreite des Prognosewertes 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 der Prognose inhärente Unsicherheit lenkt.⁴⁷ Prognosen sollten daher unbedingt analog zu Umfragen, Maße zur Quantifizierung der der Prognose inhärenten Unsicherheit ausweisen.

Darüber hinaus haben Modelle, die den Koalitionsstimmanteil vorhersagen, ein weiteres fundamentales Problem. Diese Modelle waren nicht in der Lage, 2013 das Ausscheiden der FDP aus dem Bundestag vorherzusagen. Modelle, die die individuellen Stimmanteile der Parteien prognostizieren, haben hier im Prinzip einen Vorteil, da sie Prognosen für alle Parteien liefern, jedoch sahen alle Modelle die FDP bei über 5 %. Ebenso sahen sämtliche Umfragen die FDP bei mindestens fünf Prozent. Das tatsächliche Ergebnis der FDP lag jedoch sowohl für die Modelle als auch Umfragen innerhalb der Fehlermarge.

Zu guter letzt soll noch ein Aspekt Erwähnung finden, der wiederum Prognosen und Umfragen gleichermaßen trifft. Die Veröffentlichung von Umfragedaten noch unmittelbar vor der Wahl 2013 war umstritten, da die Umfrageergebnisse das Wahlverhalten der BürgerInnen beeinflussen könnte. Die Wirkung von Wahlumfragen auf das Wahlverhalten ist daher auch ein Thema in der Wahlforschung wie Hanna Hoffmann in ihrem Beitrag in diesem Band darstellt. Für Wahlprognosen ist von einer solchen Wirkung zumindest momentan nicht auszugehen – dafür finden sie schlicht noch zu wenig Beachtung. Bei Prognosemodellen, die explizit für sich in Anspruch nehmen, den Ausgang einer Wahl exakt vorherzusagen, scheint diese Frage jedoch besonders relevant. So stellt

⁴⁷ *Andrew Leigh und Justin Wolfers.* „Competing Approaches to Forecasting Elections: Economic Models, Opinion Polling and Prediction Markets*“. In: *Economic Record* 82.258 (2006), S. 325–340, S. 6.

sich die Frage, ob manche WählerInnen der Wahlurne fernbleiben werden, wenn das Ergebnis doch schon bekannt ist. Wie auch für Umfragen gilt hier, dass Prognosen Informationen bereitstellen, die WählerInnen in ihre Wahlentscheidung einbeziehen können oder auch nicht. In der Tat könnten Wahlprognosen, sofern sie einen knappen Wahlausgang prognostizieren, die Wahlbeteiligung im Gegenteil auch steigern. Eine Voraussetzung dafür wäre, dass die in Prognosen inhärente Unsicherheit stärker kommuniziert und beachtet wird. Hier gibt es wie auch bei Umfragen noch Verbesserungspotential in der Kommunikation der Umfrage- bzw. Prognosewerte.

Nicht nur aufgrund begrenzter Erfahrungswerte, sondern natürlich auch angesichts der hier aufgeführten Kritikpunkte sind Prognosemodelle keine Alternative zu Umfragen. Eine damit einhergehende Sichtweise eines Entweder-Oder ist jedoch auch gar nicht angebracht. Dennoch lässt sich aufgrund der bisherigen Erfahrungen mit Prognosemodellen in Deutschland sagen, dass diese eine sinnvolle Ergänzung zur klassischen Sonntagsfrage darstellen.

6. Abschließende Bemerkungen

Der Wahlausgang ist schon deutlich vor seinem Eintreten von erheblichem öffentlichen Interesse, jedoch musste sich die interessierte Öffentlichkeit bisher primär auf ExpertInnenmeinungen und Meinungsumfragen stützen, wollte sie sich vor der Wahl eine Erwartung über den möglichen Ausgang bilden. Mit der Bundestagswahl 2013 gab es jedoch erstmal eine nennenswerte Anzahl wissenschaftlicher Prognosemodelle in Ergänzung zu Umfragen und ExpertInneneinschätzungen. Dies nahm dieser Beitrag zum Anlass, diese Modelle vorzustellen und die Frage zu stellen, welche Rolle diesen neben etablierten Umfragen spielen können. Sind Prognosemodelle eine sinnvolle Ergänzung oder gar Alternative zu Umfragen?

Hier war zunächst festzustellen, 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 ExpertInnenmeinungen folgen sie wissenschaftlichen Prinzipien der intersubjektiven Nachvollziehbarkeit und Reproduzierbarkeit. Diese Modelle haben sich in ihrer Genauigkeit überwiegend genauso gut, manchmal besser, bewährt wie traditionelle Meinungsumfragen.

Dies kann im Hinblick auf zukünftige Prognosen im Allgemeinen und Prognosen für deutsche Bundestagswahlen im Besonderen optimistisch stimmen – mehr jedoch nicht. Während in den USA aufgrund langjähriger Erfahrung mit einer größeren Zahl an Modellen die Reliabilität von Prognosemodellen als relativ gesichert gelten kann, ist dies in Deutschland noch nicht der Fall. Hier gibt es erst seit der letzten Wahl eine nennenswerte Zahl an Prognosemodellen. 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 Boden an synthetische Modelle, die Umfragen und strukturelle Modelle, manchmal auch ExpertInneneneinschätzungen oder Prognosemärkte, kombinieren. Synthetische Modelle finden sowohl in den USA⁴⁸ als auch in Europa⁴⁹ 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 synthetische Modelle 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. Für Umfragen mögen mögliche 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. Solche Prognosemodelle werden durch spezifische Ereignisse 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 daher 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 daneben liegen. All dies spricht für eine stärkere Beschäftigung der Politikwissenschaft mit Prognosen.

Wahlprognosen, wie sie hier vorgestellt wurden, stellen damit eine sinnvolle Ergänzung, jedoch sicher nicht eine Alternative im Sinne eines entweder oder, zu den traditionellen Ansätzen der ExpertInnenmeinungen 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 und Prognosemodelle und damit auch prädiktive Tests elektoraler Theorien.

⁴⁸ *Drew A. Linzer*. „Dynamic Bayesian Forecasting of Presidential Elections in the States“. In: *Journal of the American Statistical Association* 108.501 (2013), S. 124–134; *Robert S. Erikson* und *Christopher Wlezien*. „Forecasting US Presidential Elections Using Economic and Noneconomic Fundamentals“. In: *PS: Political Science & Politics* 47.02 (2014), S. 313–316.

⁴⁹ *Andreas Graefe*, „Combined Forecasts of the 2013 German Election“; *Michael S. Lewis-Beck* und *Ruth Dassonneville*. „Forecasting elections in Europe: Synthetic models“. en. In: *Research & Politics* 2.1 (2015).