

Haben die Demoskopien die FDP aus dem Bundestag vertrieben? Ergebnisse einer experimentellen Studie

Thorsten Faas und Sascha Huber

1. Ausgangslage: Das schwierige Verhältnis zwischen der FDP und den Demoskopien im Wahljahr 2013 und die Erwartungen der Wähler

Das Verhältnis zwischen der FDP und den Demoskopien war im Wahljahr 2013 eine Geschichte voller Missverständnisse. Am 20. Januar 2013 fand die Landtagswahl in Niedersachsen statt. Die Ausgangslage für die Liberalen war – analog zur Situation im Bund – bescheiden. In Umfragen lag die Partei bundesweit zu diesem Zeitpunkt stabil unter fünf Prozent; Parteichef *Philipp Rösler* stand kurz vor dem Rücktritt. Die Wahl in Niedersachsen – „seinem“ Heimatland – sollte über sein Schicksal entscheiden. Auch dort sah die Situation nicht gut für ihn und seine Partei aus, wie die letzten Umfragen zeigten: Am 10. Januar veröffentlichten sowohl ARD (basierend auf Daten von Infratest dimap) als auch ZDF (basierend auf Daten der Forschungsgruppe Wahlen) Zahlen, die die FDP in Niedersachsen bei genau fünf Prozent sahen. Am 12. Januar veröffentlichte der Focus (basierend auf Daten der INFO GmbH) neue Umfrageergebnisse, wonach die FDP bei 4,5 Prozent und damit sogar unterhalb der Fünf-Prozent-Sperrklausel lag. Am 17. Januar 2013 schließlich kam in SAT1 Norddeutschland (basierend auf Zahlen von GMS) die letzte frische Umfrage vor der Wahl mit der FDP bei fünf Prozent.

Allerorten wurde im Angesicht solcher Zahlen mit einem spannenden Wahlabend gerechnet. So geschah es auch, allerdings ging es dabei überhaupt nicht um die Frage, ob die FDP in den Landtag in Hannover einziehen würde oder nicht. Am Ende standen 9,9 Prozent für die Liberalen zu Buche – doppelt so viel, wie ihnen die Umfragen vorhergesagt hatten. Ein Wermutstropfen für die FDP war allerdings, dass sie trotzdem die gemeinsame Regierung mit der CDU unter Ministerpräsident *David McAllister* nicht fortsetzen konnte. Wie ganz am Ende eines langen Wahlabends feststand (und das war der Quell der Spannung), lagen SPD und Grüne hauchdünn vor Union und FDP; *Stephan Weil* würde später *David McAllister* als niedersächsischer Ministerpräsident ablösen.

Konsequenzen hatte die Wahl in Niedersachsen auch für die demoskopische Landschaft in Deutschland. Die öffentlich-rechtlichen Fernsehanstalten hatten sich über viele Jahre hinweg eine Selbstverpflichtung auferlegt¹: In der Woche unmittelbar vor einer Wahl werden keine neuen Ergebnisse von Wahlumfragen mehr veröffentlicht, damit die Wähler „in Ruhe“ ihre Entscheidung treffen können. Für den niedersächsischen Fall bedeutete dies, wie oben gesehen, dass die letzten Zahlen von ARD und ZDF jeweils zehn Tage vor der Wahl veröffentlicht wurden. Nach der Wahl aber kündigte das ZDF diese Selbstverpflichtung einseitig auf. In einer Pressemitteilung des Senders hieß es dazu: „Das ZDF veröffentlicht vor der Bundestagswahl im September sein ‚Politbarometer‘ erstmals am Donnerstag vor der Wahl. Ab 2014 wird dieses Modell grundsätzlich vor allen Bundestags-, Landtags- und Europawahlen angewendet. Die bisher geltende freiwillige Beschränkung, in der Wo-

1 Siehe dazu auch den Beitrag von *Thomas Petersen* in diesem Heft der ZParl.

che vor einer bundesweiten Wahl oder einer Landtagswahl keine Umfrageergebnisse zu veröffentlichen, wird damit aufgehoben, kündigte ZDF-Intendant Dr. *Thomas Bellut* vor dem Fernsehrat in Mainz an. Das Wählerverhalten hat sich massiv verändert. Wechselwähler machen heute einen viel größeren Teil der Wählerschaft aus, und der Wähler entscheidet sich immer später. Wir sehen uns daher in der Pflicht, den Wähler mit einem aktuellen Stimmungsbild zu informieren und ihn nicht wider besseres Wissen auf dem Stand veralteter Informationen zu lassen², erklärte *Bellut*. Eine umfassende Auswertung der Forschungsgruppe Wahlen habe ergeben, dass die bisher unveröffentlichten, kurzfristigen Umfragen vor Wahlen deutlich geringere Fehlerquoten aufwiesen als die letztveröffentlichten Daten zehn Tage vor der Wahl.³

Diese Mitteilung ist in mindestens zweierlei Hinsicht bemerkenswert: ARD und ZDF haben also immer schon auch in der letzten Woche vor einer Wahl Umfragen durchgeführt, deren Ergebnisse aber nicht mehr veröffentlicht. Zweitens möchte man seitens des ZDF mit dem neuen Vorgehen die Wähler bewusst nicht mehr „in Ruhe lassen“, sondern gezielt mit Informationen versorgen – und zwar unter der Annahme, dass diese neueren Zahlen auch besser und belastbarer seien.

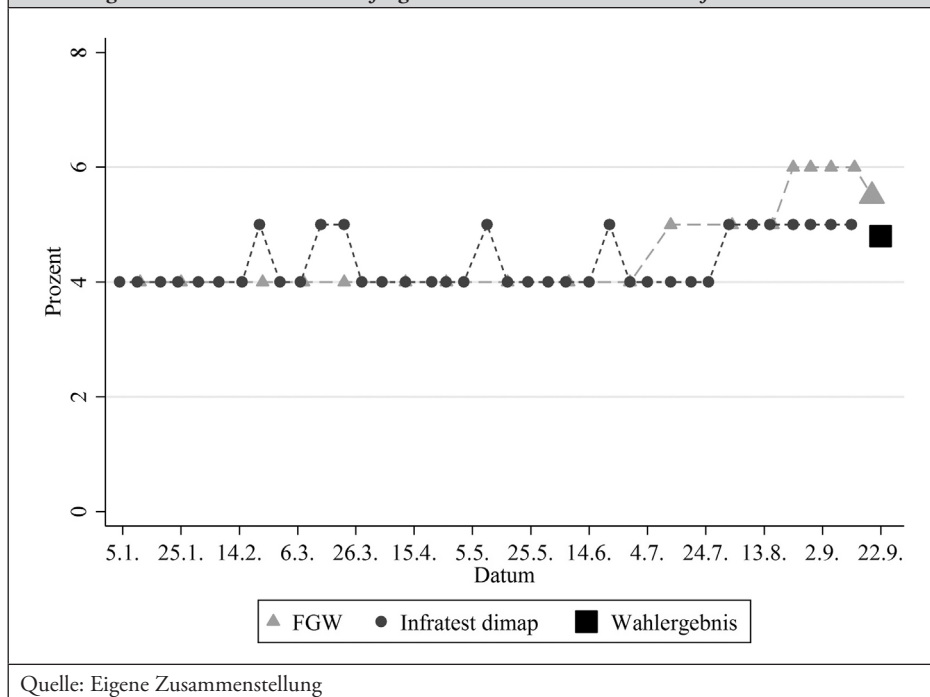
An der demoskopischen Ausgangslage für die FDP im Wahljahr 2013 änderte sich dadurch allerdings überhaupt nichts. Im Gegenteil: Abbildung 1 stellt die Werte dar, die ARD/Infratest dimap und ZDF/Forschungsgruppe Wahlen im Laufe des Wahljahres 2013 für die Partei auswiesen. Es blieb für die FDP trotz des Überraschungserfolgs in Niedersachsen eine Hängepartei: Bis in den Juli hinein stand sie zumeist bei vier Prozent, nur gelegentlich kam sie in Umfragen von Infratest dimap auf fünf Prozent. Ab August wurden die Liberalen in Umfragen dann recht stabil bei fünf Prozent gesehen, im ZDF-Politbarometer später sogar bei sechs Prozent. Auch im letzten Politbarometer vor der Wahl, das – in der Logik der oben zitierten Pressemitteilung – erstmals am Donnerstag vor der Bundestagswahl veröffentlicht wurde, lag die FDP bei 5,5 Prozent und damit noch immer oberhalb der Sperrklausel von fünf Prozent. Dabei hängen diese skizzierten Muster in keiner Weise von den gewählten Instituten ab: Unmittelbar vor der Wahl gab es praktisch keine demoskopischen Zahlen, die die FDP unterhalb der Fünf-Prozent-Hürde sahen. Vielmehr ist die Stabilität, mit der die Partei während der heißen Phase des Wahlkampfes in den letzten beiden Monaten vor der Bundestagswahl von allen Umfragen bei genau fünf Prozent oder knapp darüber geführt wurde, erstaunlich (vgl. Abbildung 1).

Blickt man auf die Erwartungen der Wähler bezüglich der Chancen der FDP, in den Bundestag einzuziehen, scheinen sehr viele diesen veröffentlichten Zahlen vertraut zu haben. Die Mehrheit war verhältnismäßig sicher, dass die FDP wieder im Bundestag vertreten sein würde. Nach Daten aus der Vorwahlbefragung der *German Longitudinal Election Study* gaben 52 Prozent der Wähler an, dass sie es als sicher oder zumindest sehr wahrscheinlich einschätzten, dass die FDP den Sprung über die Fünf-Prozent-Hürde schaffen würde.³ Dies ist ein bemerkenswert hoher Wert: Bei der Bundestagswahl 2009, bei der die FDP von den

2 ZDF, „Politbarometer“ künftig dichter am Wahltag: ZDF-Intendant Bellut: Reaktion auf geändertes Wählerverhalten, <http://www.presseportal.de/pm/7840/2478732> (Abruf am 15. Juli 2015).

3 Eigene Berechnungen anhand der face-to-face Vorwahlbefragung der German Longitudinal Election Study (GLES), ZA5700.

Abbildung 1: FDP-Werte in den Umfragen von ARD und ZDF im Wahljahr 2013



Umfrageinstituten vor der Wahl bei deutlich über zehn Prozent gesehen wurde, waren es mit 66 Prozent der Befragten nur geringfügig mehr.

Aufschlussreich ist dabei auch ein Vergleich mit den Erwartungen der Wähler zu den Einzugschancen der Alternative für Deutschland (AfD) bei der Bundestagswahl 2013. Hier waren sich nur vier Prozent der Wähler mehr oder weniger sicher, dass die Partei in den Bundestag einziehen würde. Für die AfD wurden von den Instituten vor der Wahl zwar konsistent geringere Umfragezahlen als für die FDP ausgewiesen; die Unterschiede waren aber eher gering. Während die FDP in der Vorwahlzeit meist bei fünf bis sechs Prozent lag, wurde die AfD von den Instituten meist bei drei bis fünf Prozent gesehen. Bedenkt man die beträchtliche Unsicherheit dieser Zahlen für kleine Parteien bei Stichproben von häufig nur etwa 1.000 bis 2.000 Befragten, war es für beide Parteien sehr unsicher, ob sie es in den Bundestag schaffen würden oder nicht. Aus statistischer Sicht ließe sich wahrscheinlich bei den meisten der durchgeführten Vorwahlumfragen der Meinungsforschungsinstitute nicht bestimmen, ob die FDP überhaupt bessere Chancen auf den Einzug in den Bundestag hatte als die AfD. Die Irrtumswahrscheinlichkeit wird häufig zu groß gewesen sein. Dennoch nahmen die Wähler die in den Umfragen ausgewiesenen Unterschiede offensichtlich sehr ernst und bildeten extrem unterschiedliche Erwartungen über die Chancen der beiden Parteien, in den Bundestag einzuziehen.

Umso überraschender war dann das Ergebnis am Wahlabend, als der Balken für die FDP letztlich bei 4,8 Prozent stehen blieb. Erstmals hatten die Liberalen damit, entgegen allen Erwartungen, den Einzug in den Deutschen Bundestag verpasst. Die Meinungsforschungsinstitute lagen damit zwar statistisch relativ nah am letztlichen Wahlergebnis – in

dessen politischer Bedeutung, nämlich dem Scheitern der FDP, aber fundamental daneben. Genau um dieses Phänomen soll es im vorliegenden Beitrag gehen, insbesondere um die Frage, ob die Demoskopen zu diesem Scheitern der FDP selbst mit ihren Veröffentlichungen einen Beitrag geleistet haben.

2. Veröffentlichte Umfragen und Wählerverhalten: Forschungsstand und Forschungsfrage

Die Frage, ob das Verhalten von Wählern – und darüber vermittelt Wahlen insgesamt – durch die Veröffentlichung von Umfragen beeinflusst werden, ist keineswegs neu. Im Public Opinion Quarterly hat *Herbert A. Simon*⁴ schon 1954 einen wichtigen Beitrag dazu verfasst. Inzwischen gibt es so viele Studien, dass sogar schon verschiedene Überblicksdarstellungen zu den möglichen Einflüssen von Umfragen auf Wahlen vorliegen, etwa von *Wolfgang Donsbach*⁵, *Frank Brettschneider*⁶, *Harald Schoen*⁷ oder *David Hopmann*⁸. Diese Beiträge verweisen auf eine beachtliche Vielfalt an theoretisch denkbaren Effekten und zugrunde liegenden Mechanismen: Wird die Wahlbeteiligung beeinflusst oder aber die Entscheidung, für oder gegen eine Partei zu stimmen? Geht dieser (mögliche) Einfluss von der absoluten Höhe der veröffentlichten Umfragewerte aus oder den Veränderungen der veröffentlichten Zahlen im Vergleich zu einer früheren Umfrage? *Hopmann*⁹ unterscheidet zudem mögliche Umfrageeffekte danach, ob sie bei den Empfängern eine affektive oder eine kognitive Reaktion auslösen. Zu den affektiven Reaktionen zählt er etwa den Wunsch, nicht isoliert zu sein oder zu den Gewinnern zu gehören (so genannter Bandwagon-Effekt). Aber auch mögliches Mitleid mit (vermeintlichen) Verlierern (und in der Folge ein Wahlverhalten zu deren Gunsten) zählen in diesen Kreis (so genannter Underdog-Effekt).

Diesen gegenüber stehen kognitive Prozesse, zu denen *Hopmann* Signalempfang, Signalverarbeitung sowie vor allem strategisch-taktisches Wahlverhalten zählt. Im ersten Fall nutzt der Empfänger einer Umfrage diese als sehr einfache Entscheidungshilfe gemäß der Logik: Ein Kandidat steht in Umfragen schlecht da, also kann er nicht gut sein. Eine „Signalverarbeitung“ dagegen impliziert, dass der Empfänger nach Erhalt einer Umfrage erneut seinen eigenen Standpunkt zu einem Kandidaten oder einer Sachfrage überdenkt.¹⁰ Unmittelbar

4 Vgl. *Herbert A. Simon*, Bandwagon and Underdog Effects and the Possibility of Election Predictions, in: *Public Opinion Quarterly*, 18. Jg. (1954), H. 3, S. 245 – 253.

5 Vgl. *Wolfgang Donsbach*, Die Rolle der Demoskopie in der Wahlkampf-Kommunikation. Empirische und normative Aspekte der Hypothese über den Einfluss der Meinungsforschung auf die Wählermeinung, in: *ZfP*, 31. Jg. (1984), H. 4, S. 388 – 409.

6 Vgl. *Frank Brettschneider*, Demoskopie im Wahlkampf – Leitstern oder Irrlicht?, in: *Markus Klein / Wolfgang Jagodzinski / Ekkehard Mochmann / Dieter Ohr* (Hrsg.), 50 Jahre empirische Wahlforschung in Deutschland, Opladen 2000, S. 477 – 505.

7 Vgl. *Harald Schoen*, Wirkung von Wahlprognosen auf Wahlen, in: *Thomas Berg* (Hrsg.), *Moderner Wahlkampf*, Opladen 2002, S. 171 – 191.

8 Vgl. *David Hopmann*, Vom emotionalen Underdog zur bewussten Strategie: Wie Meinungsumfragen die Parteipräferenzen der Wähler beeinflussen, in: *Thorsten Faas / Kai Arzheimer / Sigrid Roßteutscher* (Hrsg.), *Information – Wahrnehmung – Emotion: Politische Psychologie in der Wahl- und Einstellungsforschung*, Wiesbaden 2010, S. 51 – 69.

9 Vgl. *David Hopmann*, a.a.O. (Fn. 8).

10 Vgl. *Diana C. Mutz*, *Impersonal Influence: How Perceptions of Mass Collectives Affect Political Attitudes*, Cambridge 1998; *dies.*, *Impersonal Influence. Effects of Representations of Public Opinion on Political Attitudes*, in: *Political Behavior*, 14. Jg. (1992), H. 1, S. 89 – 122.

auf das Wahlverhalten bezogen sind schließlich mögliche strategisch-taktische Überlegungen, die Wähler in Folge von Umfragen anstellen. Dabei wollen sie ihren eigenen Einfluss auf das Wahlergebnis (und seine Folgen) optimieren und so ihre Stimme bestmöglich einsetzen.

Trotz oder vielleicht gerade wegen dieser Vielfalt an theoretischen Argumenten zu möglichen Auswirkungen von veröffentlichten Meinungsumfragen gilt bis heute, was *Alexander Gallus* schon 2003 bilanzierte: „Auf die Frage, ob und wie Umfrageergebnisse und insbesondere Wahlprognosen das Wahlverhalten der Bevölkerung beeinflussen, fehlen bis heute gleich eindeutige wie empirisch gesicherte Antworten.“¹¹ An einigen Stellen mag dies theoretisch begründet sein – einige der formulierten Hypothesen stehen in direktem Gegensatz zueinander. Aber auch was die Empirie betrifft, finden sich immer wieder gegensätzliche Befunde. So postuliert die Bandwagon-Hypothese, dass erfolgreich erscheinende Parteien wegen ihres Erfolges zusätzliche Unterstützung in der Wählerschaft erfahren – dies wird an einigen Stellen auch empirisch bestätigt.¹² Allerdings konnte im Kontext der Bundestagswahl 2005 ebenso gezeigt werden, dass die SPD, die zunächst noch in Umfragen wie der sichere Verlierer der von ihr selbst vorgezogenen Neuwahl aussah, gegen Ende hin aufholen konnte und dass diese Aufholjagd auch in Teilen durch demoskopische Veröffentlichungen befördert wurde¹³, was der Bandwagon-These widerspricht und theoretisch eher die Underdog-Hypothese stützt.

Die empirische Befundlage ist also uneinheitlich, obwohl doch die skizzierten Thesen hohe Plausibilität aufweisen. Wie kann das sein? Möglich ist zunächst, dass die – teils gegenläufigen – Hypothesen jeweils nur in bestimmten Teilen der Wählerschaft zutreffen. In der Summe würden sich die Effekte dann gegenseitig aufheben. Denkbar erscheint aber auch, dass die individuellen Kalküle der Wähler mit den jeweiligen kontextuellen Bedingungen interagieren und eben je nach Kontext zu verschiedenen Reaktionen führen. Das

- 11 *Alexander Gallus*, Wahl als „Demoskopiedemokratie“? Überlegungen zur Meinungsforschung und ihren Wirkungen aus Anlass der Bundestagswahl 2002, in: *Eckhard Jesse* (Hrsg.), Bilanz der Bundestagswahl 2002, Wiesbaden 2003, S. 128.
- 12 Vgl. *Herbert A. Simon*, a.a.O. (Fn. 4); *Rüdiger Schmitt-Beck*, Mass Media, the Electorate, and the Bandwagon. A Study of Communication Effects on Vote Choice in Germany, in: *International Journal of Public Opinion Research*, 8. Jg. (1996), H. 3, S. 266 – 291; *Sascha Huber* / *Thomas Gschwend* / *Michael Meffert* / *Franz-Urban Pappi*, Erwartungsbildung über den Wahlausgang und ihr Einfluss auf die Wahlentscheidung, in: *Oskar W. Gabriel* / *Jürgen W. Falter* / *Bernhard Wessels* (Hrsg.), Wahlen und Wähler. Analysen aus Anlass der Bundestagswahl 2005, Wiesbaden 2009, S. 562 – 584.
- 13 Vgl. *Thorsten Faas* / *Christian Mackenrodt* / *Rüdiger Schmitt-Beck*, Polls that Mattered: Effects of Media Polls on Voters' Coalition Expectations and Party Preferences in the 2005 German Parliamentary Election, in: *International Journal of Public Opinion Research*, 20. Jg. (2008), H. 3, S. 299 – 325; *Thorsten Faas* / *Rüdiger Schmitt-Beck*, Wahrnehmung und Wirkungen politischer Meinungsumfragen. Eine Exploration zur Bundestagswahl 2005, in: *Frank Brettschneider* / *Oskar Niedermayer* / *Bernhard Wessels* (Hrsg.), Die Bundestagswahl 2005. Analysen des Wahlkampfes und der Wahlergebnisse, Wiesbaden 2007, S. 233 – 267; *Hanna Hoffmann* / *Markus Klein*, Wirkungen von veröffentlichten Wahlumfragen auf die Koalitionserwartung, die Wahlbeteiligung und die Wahlentscheidung bei der Bundestagswahl 2009: Eine Mehrebenenanalyse auf der Grundlage der Rolling Cross-Section-Befragung der German Longitudinal Election Study (GLES), in: *Thorsten Faas* / *Kai Arzheimer* / *Sigrid Roßteutscher* / *Bernhard Wessels* (Hrsg.), Koalitionen, Kandidaten, Kommunikation: Analysen zur Bundestagswahl 2009, Wiesbaden 2013, S. 221 – 246.

sollte gerade für kognitive Effekte von Umfragen gelten, denn hier versuchen Bürger ja gerade, ihr Verhalten in Reaktion auf äußere Umstände zu optimieren.

Dass der jeweilige Kontext eine wichtige Rolle spielen kann, wird schon daraus ersichtlich, dass bestimmte Thesen zum möglichen Einfluss von Umfragen auf Wählerverhalten zunächst explizit für bestimmte Kontexte formuliert worden sind. So besagt das *wasted-vote*-Argument als eine Variante strategisch-taktischen Wahlverhaltens, dass Wähler nur jenen Kandidaten (oder Parteien) ihre Stimme geben, die eine Chance auf Erfolg haben. Dieser Zusammenhang wurde ursprünglich für Mehrheitswahlsysteme amerikanischer oder britischer Provenienz formuliert. Wenn Wähler also wirklich das Ziel haben „to make their votes count“¹⁴, so sollten sie nur für solche Kandidaten (oder Parteien) votieren, die eine Chance haben, als erste ins Ziel zu kommen. Um dies aber überhaupt einschätzen zu können, müssen sie entsprechende Erwartungen bilden, wofür wiederum Umfragen eine zentrale Rolle spielen.¹⁵ Erst auf diesen Erwartungen aufbauend können sie versuchen, eine für sie optimale Entscheidung zu treffen.¹⁶ Ob sie also ihrer eigentlich am stärksten präferierten Partei ihre Stimme geben können oder nicht, hängt davon ab, ob diese Partei überhaupt eine Chance auf den ersten Platz hat. Ist dies nicht der Fall, müssen sie nach Alternativen suchen.

Ursprünglich für den Kontext von Mehrheitswahlsystemen konzipiert, können solche strategisch-taktischen Überlegungen auch in Verhältniswahlsystemen eine Rolle spielen.¹⁷ *Harald Schoen*¹⁸ diskutiert die so genannte *Fallbeil-Hypothese*, derzufolge Parteien bei Bundestagswahlen dann an Zustimmung verlieren, wenn sie unterhalb der Fünf-Prozent-Hürde gesehen werden. Eine Stimme für sie wäre dann nämlich ganz im Sinne des *wasted-vote*-Arguments vergeudet, da sie bei der Vergabe von Sitzen im Deutschen Bundestag nicht berücksichtigt würde. Andere argumentieren gleichzeitig, dass Wähler in Mehrparteiensystemen nicht nur an der parteipolitischen Zusammensetzung des Parlaments, sondern vor allem auch an der Zusammensetzung der Regierung interessiert sind.¹⁹ Sie werden deshalb versuchen, ihre Stimme möglichst so einzusetzen, dass sie den größtmöglichen Einfluss auf die Bildung einer bestimmten Koalition hat. Das erfolgt beim so genannten *Leihstimmen-Wählen*: Ein Wähler vergibt hierbei seine Stimme nicht an die von ihm eigentlich bevorzugte Partei, sondern an deren potenziellen – kleineren – Partner in der gewünschten Koalition, wenn dessen Einzug ins Parlament gefährdet ist. Würde der Bürger seine eigent-

14 Gary W. Cox, *Making Votes Count. Strategic Coordination in the World's Electoral Systems*, Cambridge 1997.

15 Vgl. Thorsten Faas / Rüdiger Schmitt-Beck, a.a.O. (Fn. 13); vgl. Galen A. Irwin / Joop M. van Holsteyn, *According to the Polls: The Influence of Opinion Polls on Expectations*, in: *Public Opinion Quarterly*, 66. Jg. (2002), H. 1, S. 92 – 104.

16 Vgl. R. Michael Alvarez / Jonathan Nagler, *A New Approach for Modelling Strategic Voting in Multiparty Elections*, in: *British Journal of Political Science*, 30. Jg. (2000), H. 1, S. 57 – 75; Andre Blais / Thomas Gschwend, *Strategic Desertion across Elections, Parties, and Voters*, in: Russell J. Dalton / Christopher J. Anderson (Hrsg.), *Citizens, Context and Choice: How the Political Context Shapes Electoral Behavior*, Oxford 2011, S. 176 – 195.

17 Vgl. Andre Blais / Elisabeth Gidengil / Neil Nevitte, *Do Polls Influence the Vote?*, in: Henry E. Brady / Richard Johnston (Hrsg.), *Capturing Campaign Effects*, Ann Arbor 2006, S. 263 – 279.

18 Vgl. Harald Schoen, *Split-Ticket Voting in German Federal Elections, 1953-1990: An Example of Sophisticated Balloting?*, in: *Electoral Studies*, 18. Jg. (1999), H. 4, S. 473 – 496.

19 Vgl. Franz U. Pappi / Paul W. Thurner, *Electoral Behavior in a Two-Vote System: Incentives for Ticket Splitting in German Bundestag Elections*, in: *European Journal of Political Research*, 41. Jg. (2002), H. 2, S. 207 – 232.

lich bevorzugte Partei wählen, so liefe er Gefahr, dass der mögliche Koalitionspartner an der Fünf-Prozent-Hürde scheitert, das gegnerische Lager eine Mehrheit der Sitze erhält und damit zukünftig die Regierung stellt. Mit einer Stimme für den kleineren Partner kann zumindest versucht werden, dessen Einzug ins Parlament zu ermöglichen und so der gewünschten Regierungskoalition zur Mehrheit zu verhelfen. Entgegen der Fallbeil-Hypothese würden also hier Parteien, die in Umfragen nahe an der Fünf-Prozent-Hürde liegen, eher davon profitieren.

Welche dieser beiden gegenläufigen Annahmen zutrifft, entscheidet also letztlich darüber, ob eine Partei Schaden oder Nutzen aus der Tatsache zieht, dass sie mit einer Sperrklausel ringt. Um diese konkurrierenden Thesen prüfen zu können, braucht es aber in jedem Fall einen elektoralen Kontext, in dem eine Partei an der Sperrklausel scheitern *könnte*. Das ist keinesfalls immer und automatisch bei jeder Wahl gegeben, wie das Beispiel der Bundestagswahl 2009 zeigt. Damals wurden die „kleinen Parteien“ FDP, Grüne und Linkspartei in allen Umfragen jeweils weit oberhalb der Sperrklausel gesehen. 2013 aber war dies nicht der Fall – wie eingangs gezeigt, bestand das Wahljahr für die FDP vielmehr aus einem permanenten Kampf mit der Sperrklausel. Damit ist genau dieses Wahljahr auch sehr gut geeignet, um die verschiedenen theoretischen Erwartungen (und damit die Rolle von Umfragen) in einem realweltlichen Kontext zu überprüfen.

Als globale Erwartung kann formuliert werden, dass die FDP bei Gültigkeit der wasted-vote-Hypothese Schaden nehmen sollte, wenn sie in Umfragen unterhalb der Fünf-Prozent-Hürde rangiert. Liegt sie dagegen bei fünf Prozent (oder sogar darüber), sollte ihr dies eher nützen: Zweifel am Einzug der Liberalen in den Bundestag sollten dann zunehmend schwinden, so dass eine strategische Abkehr von der FDP aus Angst vor der Verschwendung der eigenen Stimme zunehmend abwegig erscheint. Umgekehrt würde es sich verhalten, wenn die Leihstimmen-Hypothese zutreffen würde: Demnach sollte die FDP gerade dann profitieren, wenn der Einzug der Partei in den Bundestag gefährdet ist. Je weiter sie von der Fünf-Prozent-Hürde entfernt liegt, desto weniger einschlägig sollten solche Überlegungen werden, da der Einzug entweder als sicher oder nicht möglich erscheint.

Über diese globalen Hypothesen hinaus lassen sich noch spezifischere Erwartungen ableiten, welche Personengruppen die Träger der jeweiligen Effekte sind: Leihstimmen sollten von Personen kommen, die eine starke Präferenz für eine Koalition aus Union und FDP haben, aber als Partei eigentlich die Union bevorzugen. Das wasted-vote-Argument dagegen sollte auf Personen zutreffen, die eine Präferenz für die FDP haben, aber gegebenenfalls in der bundesdeutschen Parteienlandschaft auch eine Alternative finden (die mit großer Sicherheit in den Bundestag einziehen kann).

3. Ein Experiment

Es gibt unterschiedliche Möglichkeiten, Wirkungszusammenhängen zwischen Umfragen und dem Ausgang von Wahlen empirisch auf den Grund zu gehen. Die Arbeiten von *Frank Brettschneider*²⁰ etwa stützen sich auf Selbsteinschätzungen von Befragten und geben damit

20 Vgl. *Frank Brettschneider*, Der taktische und rationale Wähler. Über den Einfluß von Wahlumfragen auf das Wählerverhalten bei den Bundestagswahlen 1983 bis 1990, in: PVS, 33. Jg. (1992), H. 1, S. 55 – 72.

die subjektive Perspektive von Personen wieder, ob sie bei ihrer Entscheidung Umfragen berücksichtigen oder nicht. Allerdings ist diese Art der Erfassung mit dem Risiko von Fehlwahrnehmungen behaftet und könnte folglich zu fehlerhaften Einschätzungen von demoskopischen Effekten führen.²¹ Insbesondere besteht die Gefahr, dass Wähler in Wunschdenken verfallen, wenn sie nach ihren Erwartungen über den Wahlausgang gefragt werden und folglich die Chancen der eigenen Partei deutlich besser einschätzen.²² Entsprechend finden sich auch zahlreiche Arbeiten, die Alternativen nutzen und etwa der Logik von Zusammenhangsanalysen folgen. So wird gefragt, welche Konsequenzen (berichtete) Kontakte mit demoskopischen Informationen bei Wählern nach sich ziehen. Dazu wird überprüft, ob Menschen, die von Kontakten mit demoskopischen Zahlen berichten, ein anderes Wahlverhalten an den Tag legen als Menschen ohne solche Kontakte.²³ Eine verfeinerte Variante solcher Verfahren, die zusätzlich die zeitliche Dynamik von Umfragen im Vorfeld von Wahlen integriert, stellt die Verwendung von rollierenden Querschnittsstichproben dar: Eine Bevölkerungsumfrage wird dabei kontrolliert über einen längeren Zeitraum hin gestreckt, um so einerseits Veränderungen in der Bevölkerungsmeinung nachzeichnen zu können, andererseits aber auch kontextuelle Variationen (etwa sich verändernde demoskopische Zahlen) direkt mit diesen Veränderungen in Verbindung bringen zu können.²⁴

Diese Verfahren ermöglichen, wie auch die anderen Beiträge in diesem Heft zeigen, tiefgehende Einblicke in die Wirkungsweise von Umfragen. Gleichwohl sind sie mit einem Problem behaftet: Die Zuwendung zu Umfragen (und in der Folge deren Wahrnehmung samt zugehöriger Effekte) könnten von bestimmten Drittvariablen abhängen, die sowohl die Wahrnehmung dieser Umfragen als auch etwaige Muster des Wählerverhaltens beeinflussen. In der Folge könnte ein enger Zusammenhang zwischen Wahrnehmung und Wählerverhalten bestehen, der aber wegen der mangelnden Drittvariablenkontrolle nicht in einem kausalen Sinne interpretiert werden dürfte. Solche Probleme löst eine experimentelle Herangehensweise, die auch in der einschlägigen Forschung schon Tradition hat.²⁵

- 21 Vgl. *Jack Vowles*, Did the Polls Influence the Vote? A Case Study of the 1999 New Zealand General Election, in: *Political Science*, 54. Jg. (2002), H. 1, S. 67 – 78.
- 22 Vgl. *Michael Meffert / Sascha Huber / Thomas Gschwend / Franz Urban Pappi*, More than Wishful Thinking: Causes and Consequences of Voters' Electoral Expectations about Parties and Coalitions, in: *Electoral Studies*, 30. Jg. (2011), H. 4, S. 804 – 815.
- 23 Vgl. *Rüdiger Schmitt-Beck*, a.a.O. (Fn. 12); siehe auch den Beitrag von *Thomas Zerback* und *Carsten Reinemann* in diesem Heft der ZParl.
- 24 Siehe auch den Beitrag von *Hanna Hoffmann* in diesem Heft der ZParl; *dies. / Markus Klein*, a.a.O. (Fn. 13); *Thorsten Faas / Christian Mackenrodt / Rüdiger Schmitt-Beck*, a.a.O. (Fn. 21); *André Blais / Elisabeth Gidengil / Neil Nevitte*, a.a.O. (Fn. 17); *André Blais / Jack Vowles / Kees Aarts*, Does the Impact of Polls Vary Across Electoral Systems?, Konferenzpapier für das 2002 Annual Meeting of the American Political Science Association, Boston 2002; zum Verfahren einer RCS-Studie siehe *Rüdiger Schmitt-Beck / Thorsten Faas / Christian Holst*, Der Rolling Cross-Section Survey – ein Instrument zur Analyse dynamischer Prozesse der Einstellungsentwicklung. Bericht zur ersten deutschen RCS-Studie anlässlich der Bundestagswahl 2005, in: *ZUMA-Nachrichten*, 58. Jg. (2006), S. 13 – 49.
- 25 Vgl. *Gregor Daschmann*, Vox Pop & Polls: The Impact of Poll Results and Voter Statements in the Media on the Perception of a Climate of Opinion, in: *International Journal of Public Opinion Research*, 12. Jg. (2000), H. 2, S. 160 – 181; *Diana C. Mutz*, Impersonal Influence. Effects of Representations of Public Opinion on Political Attitudes, a.a.O. (Fn. 10); *Jürgen Maier / Frank Brettschneider*, Wirkungen von Umfrageberichterstattung auf Wählerverhalten: Ein Online-Experiment zu den Landtagswahlen in Baden-Württemberg 2006, Rheinland-Pfalz 2006 und Hessen

Entscheidend bei dieser Methode ist, dass man einen bestimmten Faktor, von dem erwartet wird, dass er eine Wirkung hat, bewusst kontrolliert und manipuliert. Experimentalforscher sprechen von einem „Treatment“, das gesetzt wird. Allerdings bedarf es zusätzlich noch einer Vergleichsmöglichkeit, um evaluieren zu können, ob dieses gesetzte Treatment tatsächlich den vermuteten Effekt hat. Dazu greift man auf die Unterscheidung zwischen mindestens einer Experimental- und einer Kontrollgruppe zurück: Während die Experimentalgruppe dem Treatment bewusst ausgesetzt wird, bleibt genau dies in der Kontrollgruppe aus. Um belastbar sagen zu können, dass das eingesetzte Treatment tatsächlich gewirkt hat, muss sichergestellt sein, dass sich Experimental- und Kontrollgruppe in nichts unterscheiden – außer dem eingesetzten Treatment. Indem Befragte zufällig verschiedenen Gruppen zugewiesen werden, wird gewährleistet, dass es keine systematischen Unterschiede zwischen den Gruppen gibt. Sind diese nämlich hinreichend groß, kann man tatsächlich davon ausgehen, dass sich infolge der zufälligen Verteilung die Gruppen in nichts unterscheiden – außer dem eingesetzten Treatment. Damit wird eine belastbare Evaluation der Wirkung möglich. Beobachtbare Unterschiede zwischen den Gruppen können dann tatsächlich auf das Treatment – und nichts anderes – zurückgeführt werden. Den Nachweis von Kausalität kann kein Forschungsdesign besser erbringen als das Experiment.

Übertragen auf die vorliegende Forschungsfrage heißt das: Um die Effekte verschiedener Umfrageergebnisse auf das beabsichtigte Wahlverhalten von Personen untersuchen zu können, nutzt man verschiedene Experimentalgruppen, die man bestimmten, aber jeweils unterschiedlichen Umfrageinformationen aussetzt. Anschließend kann durch den Vergleich der Gruppen geprüft werden, ob sich das Wahlverhalten zwischen ihnen unterscheidet. Genau dieser Logik folgt auch unsere experimentelle Studie, die Teil einer Online-Umfrage zur Bundestagswahl 2013 – www.wahlumfrage2013.de – war.

Die Onlineumfrage war vom 30. August bis zum 22. September 2013 im Feld; insgesamt nahmen 3.285 Personen an der Befragung teil, die meisten davon in den letzten beiden Wochen vor der Wahl. Sie wurden über Mailinglisten und in sozialen und persönlichen Netzwerken rekrutiert; sie waren im Mittel 39,2 Jahre alt, 38 Prozent waren weiblich, 72 Prozent hatten als Bildungsabschluss Abitur. Natürlich ist diese Stichprobe damit nicht repräsentativ für die wahlberechtigte Bevölkerung in Deutschland. Für eine experimentelle Studie ist Repräsentativität aber von nachrangiger Bedeutung. Im Vordergrund steht vielmehr die interne Validität: Kann der Nachweis eines kausalen Einflusses der eingesetzten Treatments erfolgreich erbracht werden? Dafür ist entscheidend, dass die zufällige Verteilung von Probanden auf die verschiedenen experimentellen Bedingungen gelingt, was in der vorliegenden Studie auch tatsächlich der Fall war.

Die Befragten sollten sich vorstellen, dass sie am Tag der Bundestagswahl noch nicht gewählt haben und dann auf Twitter einen Tweet sehen, der Vorabergebnisse aus Wahltagsbefragungen zum Ausgang der Wahl beinhaltet. Zweierlei Gründe sprachen für dieses Szenario aus Twitter und Wahltagsbefragungen, um mögliche Reaktionen auf Zahlen aus Wahlumfragen zu prüfen: Erstens ist es realistisch, da es bei früheren Bundestagswahlen am Wahltag genau solche Vorabinformationen auf genau diesem Wege gab. Zweitens sind diese Wahltagsbefragungen, die die Grundlage für die 18-Uhr-Prognosen von ARD und ZDF am Wahltag darstellen, durch ein extrem hohes Maß an Präzision gekennzeichnet. Sie do-

2008, in: *Nikolaus Jakob / Harald Schoen / Thomas Zerback* (Hrsg.), *Sozialforschung im Internet. Methodologie und Praxis der Online-Befragung*, Wiesbaden 2009, S. 321 – 337.

kumentieren tatsächliches Wahlverhalten, so dass diese Informationen nicht nur sehr belastbar sind, sondern auch als tatsächliche Prognosen des Wahlausgangs interpretiert werden können. Anders formuliert: Die Erwartungen der Probanden an den Wahlausgang sollten in dieser Form direkt beeinflusst werden und in der Folge sollten sich strategisch-taktisch motivierte Reaktionen darauf sehr deutlich ablesen lassen.

Im Zuge der Befragung wurde den Probanden ein Tweet mit solchen Informationen präsentiert, der optisch einem tatsächlichen Tweet exakt entspricht. Experimenteller Logik folgend variierten diese Tweets in ihrem Informationsgehalt zwischen verschiedenen Gruppen. Dabei blieben aber alle präsentierten Umfrageergebnisse in einem für die Bundestagswahl 2013 realistischen Rahmen, waren also glaubwürdig im Kontext dieser Wahl. Angesichts des erwarteten knappen Ausgangs der Bundestagswahl genügten kleinere Veränderungen dieser Tweets, um erhebliche machtpolitische Konsequenzen nach sich zu ziehen. Dies galt gerade auch für den (Nicht-)Einzug der FDP und damit verbunden verschiedene mögliche Mehrheitsverhältnisse und Koalitionsvarianten im zukünftigen Bundestag. Insgesamt wurden die knapp 3.300 Befragten auf acht verschiedene Gruppen verteilt. Für die Zwecke des vorliegenden Papiers ist dabei entscheidend, dass der FDP in diesen verschiedenen Bedingungen jeweils unterschiedliche Prozentwerte vorhergesagt wurden, nämlich vier, fünf, sechs oder acht Prozent.²⁶ Um die Fallzahl für die differenzierten Analysen zu erhöhen, wurden die beiden Bedingungen, in denen die FDP bei über fünf Prozent (also entweder bei sechs oder bei acht Prozent) gesehen wurde, zusammengefasst. Es werden also drei Gruppen verglichen, in denen die FDP entweder bei vier, fünf oder aber mehr als fünf Prozent lag.

Anschließend wurden die Befragten gebeten anzugeben, wie sie im Lichte dieser gerade gesehenen Information ihre beiden Stimmen am Wahltag vergeben würden. Sind trotz der zufälligen Verteilung der Probanden auf die Gruppen Unterschiede bei dem hier berichteten Wahlverhalten zu verzeichnen, ist dies ein klarer Hinweis auf eine Reaktion der Probanden auf den präsentierten Tweet. Da im Rahmen dieses Beitrags vor allem das Verhältnis zwischen FDP und Demoskopie interessiert, wird im Folgenden hauptsächlich der Stimmenanteil der FDP in den verschiedenen experimentellen Gruppen betrachtet. So können die globalen Hypothesen getestet werden. Zur Prüfung der spezifischeren Hypothesen und damit der Frage, *wer* auf die präsentierten Tweets reagiert, bedurfte es zusätzlicher Informationen über die Parteipräferenzen der Probanden, um zu identifizieren, wer eine (Erst-)Präferenz für die Union beziehungsweise für die FDP hatte. Dazu dienten Skalometerfragen, in deren Rahmen die Befragten die einzelnen Parteien des deutschen Parteiensystems auf Skalen von -5 bis +5 bewerten sollten. Eine Präferenz für die Union liegt dann vor, wenn ein Befragter den eindeutig höchsten Skalometerwert für die CDU oder CSU vergibt. Umgekehrt liegt eine Präferenz für die FDP vor, wenn ein Befragter eindeutig den höchsten Skalometerwert für die FDP vergibt.

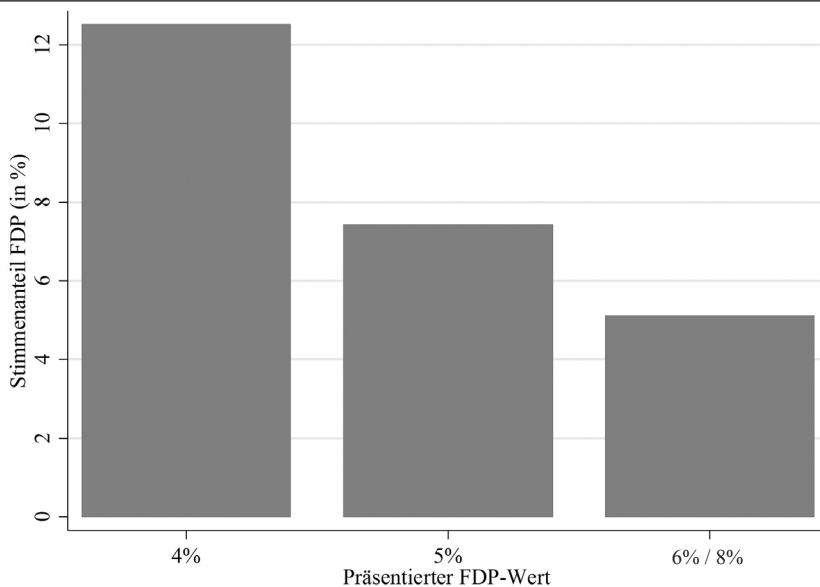
26 Die verschiedenen vorhergesagten Werte für die FDP wurden dabei zum Teil in mehreren Experimentalbedingungen verwendet: Auf vier Prozent wurde die FDP in einer Gruppe gesetzt, auf fünf Prozent in vier Gruppen, auf sechs Prozent wiederum in einer Gruppe und auf acht Prozent in zwei Gruppen. Die Mehrfachverwendung ergibt sich dabei aus der Tatsache heraus, dass in diesen Gruppen Werte für andere Parteien auch systematisch verändert wurden, die aber für den hiesigen Kontext keine Rolle spielen.

4. Empirische Befunde

Sollten sich beim Vergleich der drei Gruppen (mit den verschiedenen Informationen zum jeweiligen Stand der FDP) auch unterschiedliche Anteile für die FDP in den Antworten der Befragten ergeben, würde das für einen kausalen Effekt der präsentierten Informationen auf die Angaben zum intendierten Verhalten der Befragten schließen lassen. Abbildung 2 zeigt die entsprechenden Ergebnisse – und es werden unmittelbar erhebliche Unterschiede deutlich: In der Gruppe, in der die FDP laut getwitterter Wahltagsbefragung nur bei vier Prozent liegt, geben 12,5 Prozent der Probanden an, den Liberalen ihre Zweitstimme geben zu wollen. Wird die FDP in dem gezeigten Tweet dagegen bei fünf Prozent gesehen, wollen sich nur 7,4 Prozent für die FDP entscheiden. Liegt die FDP laut Tweet sogar bei über fünf Prozent, sinkt der Anteil nochmals auf 5,1 Prozent.²⁷

Festzuhalten bleibt also zunächst, dass deutliche Unterschiede existieren. Deren Muster gibt zudem Auskunft darüber, welche der beiden Hypothesen – wasted-vote oder Leihstimmen – zutrifft. Auch diesbezüglich ist das Muster eindeutig und spricht für die Leihstimmen- und gegen die wasted-vote-Hypothese: Je schlechter die FDP laut Umfrage dasteht, desto

Abbildung 2: Wahlabsichten zu Gunsten der FDP in Abhängigkeit von den präsentierten Umfrageinformationen



Quelle: Eigene Berechnungen auf der Basis von www.wahlumfrage2013.de (n = 3.211).

²⁷ Dass die Werte in allen drei Gruppen oberhalb des Wertes liegen, den die FDP am Wahltag erzielte, stellt dabei kein Problem dar: Bei dem hier durchgeführten Experiment, das auf einer nicht-repräsentativen Umfrage basiert, geht es ja nicht darum, den Stimmenanteil der FDP exakt vorherzusagen, sondern darum zu prüfen, ob sich Unterschiede zwischen den Gruppen ergeben – und wie man sieht, ist das deutlich der Fall.

besser schneidet die Partei in den jeweiligen Gruppen ab. Diese Unterschiede erfüllen auch das Kriterium statistischer Signifikanz (vgl. Tabelle 1). Zwischen allen möglichen Paaren von Gruppen besteht ein signifikanter Unterschied bezüglich des jeweiligen FDP-Anteils.

Wie die Ergebnisse bezüglich unserer spezifischen Hypothesen in den Teilgruppen ausfallen, zeigt Tabelle 1. Die Leihstimmen-Hypothese sollte in besonderem Maße bei Anhängern der Union zu finden sein, die ihrem kleinen Koalitionspartner unter die Arme greifen möchten. Tatsächlich tritt das entsprechende Muster in dieser Gruppe besonders deutlich zu Tage: Wird die FDP mit vier Prozent ausgewiesen, wollen über ein Viertel der Befragten in der Gruppe, die eine Erstpräferenz für die Union angeben, der FDP ihre Stimme geben. Wird die FDP dagegen bei fünf oder sogar bei sechs oder mehr Prozent gesehen, geht der Stimmenanteil für die Liberalen deutlich zurück: Sie liegen dann nur noch bei 11,1 beziehungsweise 8,4 Prozent. Der Unterschied zwischen diesen beiden Gruppen erweist sich dabei auch nicht mehr als statistisch signifikant. Die trennende Linie verläuft hier also tatsächlich zwischen vier Prozent auf der einen Seite und mindestens fünf Prozent auf der anderen Seite, was – man bedenke die demoskopische Situation vor der Bundestagswahl – von entscheidender Bedeutung ist. Diese Unterschiede erweisen sich zudem immer als statistisch hoch signifikant.

Für die wasted-vote-Hypothese hatten schon die globalen Annahmen keine entsprechenden empirischen Hinweise geliefert. Auch dies wird bei der differenzierteren Betrachtung in einzelnen Teilgruppen nochmals untermauert. Bei Befragten, die eine Erstpräferenz für die FDP haben, würde man nämlich – bei Gültigkeit der Hypothese – erwarten, dass ein Wert unter fünf Prozent zu einer Abkehr von der Partei führt. Dafür gibt es aber keinerlei Hinweise: Die resultierenden Werte von 92,3 Prozent (in der Vier-Prozent-Bedingung), 86,0 Prozent (in der Fünf-Prozent-Bedingung) sowie 81,6 Prozent (in der Bedingung mit sechs oder mehr Prozent) widersprechen ihr vielmehr. Dabei erweist sich keiner der Unterschiede als statistisch signifikant (was allerdings auch an der geringen Zahl von Fällen in einigen Gruppen liegt). Alles in allem jedenfalls unterstützen die Befunde der experimentellen Studie eindrucksvoll die Leihstimmen-Hypothese und in keiner Weise die wasted-vote-Hypothese.

Tabelle 1: Wahlabsichten zu Gunsten der FDP in Abhängigkeit von den präsentierten Umfrageinformationen und Präferenzen der Befragten (in Prozent)

	4 Prozent	5 Prozent	6 / 8 Prozent	Signifikanz der Unterschiede
Alle Befragten	12,5 (n = 415)	7,4 (n = 1.626)	5,1 (n = 1.170)	*** / *** / *
Befragte mit CDU-Präferenz	26,2 (n = 61)	11,1 (n = 234)	8,4 (n = 179)	** / *** / n.s.
Befragte mit FDP-Präferenz	92,3 (n = 13)	86,0 (n = 50)	81,6 (n = 38)	n.s. / n.s. / n.s.

Anmerkung: Die ausgewiesenen Signifikanzen beziehen sich jeweils auf den Paarvergleich von zwei Gruppen, zunächst zwischen Gruppen vier/fünf Prozent, dann vier/sechs beziehungsweise acht Prozent und schließlich fünf/sechs beziehungsweise acht Prozent.

Quelle: Eigene Berechnungen, www.wahlumfrage2013.de (n = 3.211).

5. *Fazit: (Fehlende) Leihstimmen und Demoskopien haben den Ausgang der Wahl 2013 beeinflusst*

Das schwierige Verhältnis zwischen der FDP und der Demoskopie im Wahljahr 2013 gab Anlass zu fragen, ob ein Wechselverhältnis zwischen dem Abschneiden der Partei (und damit ihrem historischen Ausscheiden aus dem Deutschen Bundestag) und den Veröffentlichungen der demoskopischen Institute bestand. Nach der *wasted-vote*-Hypothese sollten der FDP Umfragewerte unter fünf Prozent schaden, weil damit ein Vergeuden der eigenen Stimme (bedingt durch die Wahl einer Partei, die letztlich den Einzug ins Parlament verpasst) droht. Diametral entgegengesetzt postuliert die Leihstimmen-Hypothese, dass Werte unterhalb von fünf Prozent den Liberalen eher nützen, weil sie ein Signal an Anhänger der Union senden, gegebenenfalls ihren gewünschten Koalitionspartner zu stützen.

Das hier vorgestellte Experiment, mit dem diese gegensätzlichen Thesen geprüft wurden, zeigte, dass die FDP dann besonders viele Stimmen von den Probanden erhielt, wenn die Partei in der vermeintlichen Wahltagsbefragung bei vier Prozent gesehen wurde. Je höher der präsentierte Anteil stieg, desto geringer wurde der Anteil der Probanden, die der FDP ihre Stimme geben wollten. Dieses Muster war besonders stark bei Personen mit einer (Erst-)Präferenz für die Union ausgeprägt. All dies spricht für die Gültigkeit der Leihstimmen- und gegen die *wasted-vote*-Hypothese.

Angesichts der tatsächlichen Umfragewerte, die die demoskopischen Institute der FDP kurz vor der Bundestagswahl 2013 zuschrieben (und die allesamt bei über fünf Prozent lagen), lässt sich folgern, dass damit ein Signal an die Wähler gesendet wurde, dass ein Verleihen von Stimmen zu Gunsten der FDP nicht nötig sei. Am Wahltag um 18 Uhr erwies sich dies aber als trügerisch. Hätten die Institute dagegen kurz vor der Wahl Werte unter fünf Prozent für die Liberalen veröffentlicht, wäre der Einzug der FDP in den Bundestag unseren Ergebnissen zufolge wahrscheinlicher geworden.

Wie eingangs beschrieben, glaubte die Mehrheit der Wähler nicht daran, dass die FDP scheitern könnte. Zum Teil mag das an früheren Erfahrungen gelegen haben, denn sie hatte es schließlich bei früheren Bundestagswahlen am Ende immer geschafft, in den Bundestag einzuziehen. Zum Teil kann es aber auch an den veröffentlichten Umfragen gelegen haben, in denen die FDP in den Wochen vor der Wahl bei allen Instituten fast durchgängig bei genau fünf Prozent oder knapp darüber geführt wurde. Die mit diesen Werten einhergehende Unsicherheit wurde von vielen Wählern offensichtlich falsch eingeschätzt. Aus Sicht der FDP hätte es wahrscheinlich geholfen, wenn die Institute sie nicht ständig oberhalb der Fünf-Prozent-Hürde ausgewiesen hätten. Angesichts einer normalen Streuung von Umfragen mit nur 1.000 bis 2.000 Befragten ist es sehr unwahrscheinlich, dass die Partei bei den Rohdaten der Institute immer bei genau fünf Prozent oder knapp darüber lag. Unabhängig von den ausgewiesenen Zahlen wäre es sicherlich hilfreich gewesen, zumindest die Unsicherheit der Umfragewerte in der Öffentlichkeit stärker deutlich zu machen.

Vor diesem Hintergrund scheint auch die Pressemitteilung des ZDF in ihrer Erwartung fragwürdig: Die dort formulierte Annahme, dass Umfragen kurz vor der Wahl genauer sind als ältere, mag zutreffen – trotzdem sind sie aber immer noch mit Unsicherheit behaftet. Bei einem ausgewiesenen Wert von fünf Prozent ist ein Einzug der FDP aus statistischer Sicht nur minimal wahrscheinlicher als ein Nicht-Einzug. Der Kontext der Wahl 2013 könnte darüber hinaus dazu geführt haben, dass Wähler besonders heftig auf eine Umfrage reagierten, gerade weil sie so kurz vor der Wahl veröffentlicht wurde und damit besonders

belastbar und genau erschien. An das vermeintlich klare Signal, Leihstimmen seien nicht nötig, haben sich offenkundig gerade die Anhänger der Union gehalten. Letztlich hat schon *Herbert A. Simon* 1954 auf diesen Punkt aufmerksam gemacht und eine pointierte Frage formuliert, die sich Demoskopien stellen sollten: Welchen Wert müssten demoskopische Institute eigentlich veröffentlichen, damit sie genau die Reaktionen auslösen, die jene Werte hervorbringen, die die demoskopischen Institute zuvor veröffentlicht haben?²⁸

Zu fragen bleibt auch, ob die hier präsentierten Befunde auf andere Parteien übertragbar sind oder nicht vielmehr die Rolle einer Partei im Parteiensystem einen Einfluss darauf hat, wie sich Umfragen potenziell auf das Wählerverhalten auswirken. Zu denken wäre im Kontext der Bundestagswahl 2013 an die AfD, denn auch diese lag in Umfragen am Ende bei vier Prozent und erzielte 4,7 Prozent der Zweitstimmen. Im Unterschied zur FDP aber stand dieser neuen Partei keine Koalitionsoption offen – insofern scheidet die Leihstimmen-Hypothese schon allein aus diesem Grund aus. Zu prüfen wäre hier, ob stattdessen das wasted-vote-Argument für solche neuen Parteien greift.

Und schließlich bleibt ungeklärt, warum die Wahlergebnisse für die FDP in Niedersachsen und dem Bund so unterschiedlich ausgefallen sind, obwohl die demoskopischen Vorhersagen gar nicht so verschieden waren. Man mag argumentieren, dass die Wahlumfragen für die FDP im Bund deutlicher über der Fünf-Prozent-Hürde lagen und damit weniger Unsicherheit über ihre Chancen signalisiert haben als in Niedersachsen. Dieses Muster passt gut zu den Ergebnissen des Experiments. Eine weitere mögliche Erklärung liegt in den jeweiligen Wahlkampfstrategien des Koalitionspartners. *David McAllister* hatte im Vorfeld der Wahl in Niedersachsen – wohl wissend, dass dies seine einzige Chance sein würde, Ministerpräsident zu bleiben – zumindest zwischen den Zeilen angedeutet, dass Leihstimmen von Unionsanhängern für die FDP sinnvoll sein könnten. In starkem Kontrast dazu hat die Union bei der Bundestagswahl ganz bewusst das Werben der FDP um Leihstimmen konterkariert. Anders als in Niedersachsen nahm sie die Zweitstimmenkampagne der FDP nicht mehr einfach hin, sondern bekämpfte sie ganz gezielt, indem sie um beide Stimmen warb und die Zweitstimme gar als „Merkelstimme“ bezeichnete. So berichtete die *Süddeutsche Zeitung* am 27. September 2013: „Also ruft *Philipp Rösler* bei *Angela Merkel* an und wirbt um Verständnis. ... Dieses Telefonat, so ist zu hören, hat es tatsächlich gegeben. Fünf Tage vor der Bundestagswahl ... Und wie reagiert *Angela Merkel*? Sie erklärt, dass die FDP sich gar keine Sorgen machen müsse, sie liege in den Umfragen stabil über sechs Prozent.“²⁹ Es sei dahin gestellt, ob die Kanzlerin an die Genauigkeit der Umfragen wirklich glaubte und die FDP lediglich nicht allzu stark sehen wollte oder ob sie mit ihrer Wahlkampfstrategie ein Scheitern der FDP an der Fünf-Prozent-Hürde bewusst in Kauf nahm. Die veröffentlichten Umfragewerte für die FDP wurden auf jeden Fall auch auf der Ebene der politischen Eliten als Argument genutzt, um Unions-Anhänger zu überzeugen, dass die FDP auch ohne Leihstimmen den Einzug schaffen würde. Das war schließlich nicht der Fall, möglicherweise auch weil viele Wähler diesen Argumenten Glauben schenkten.

28 Vgl. *Herbert A. Simon*, a.a.O. (Fn. 4), S. 252.

29 *Stefan Braun*, Wenn aus Partnern Gegner werden, in: *SZ* vom 27. September 2013, S. 6.