

Online-Umfragen: Eine geeignete Er- hebungsmethode für die Wahlforschung?

*Ein Vergleich unterschied-
licher Befragungsmodi am
Beispiel der Bundestagswahl
2009*

Online-Surveys: An appropriate survey method for electoral research?

*A comparison of different
survey modes in use at the
2009 German Federal
Election*

Ina Elisabeth Bieber und Evelyn Bytzek

Zusammenfassung

Online-Umfragen werden in den Sozialwissenschaften immer beliebter, so auch in der Wahlforschung. Zahlreiche Studien konnten jedoch zeigen, dass sich Teilnehmer/innen an Online-Umfragen nicht nur hinsichtlich des soziodemographischen Hintergrunds, sondern auch der politischen Einstellungen von Teilnehmer/innen an persönlich-mündlichen oder telefonischen Umfragen deutlich unterscheiden. Bei Analysen des Wählerverhaltens sind diese Unterschiede jedoch unproblematisch, wenn sich die Zusammenhänge zwischen den Untersuchungsmerkmalen im Rahmen der verschiedenen Befragungsmodi ähnlich gestalten und somit die Wahlentscheidung durch dieselben Faktoren erklärt werden kann. Der vorliegende Beitrag widmet sich der Betrachtung dieser Zusammenhänge, indem Modelle des Wahlverhaltens miteinander verglichen werden, die mit Daten verschiedener Umfragemodi berechnet wurden. Hierfür werden Daten der German Longitudinal Election Study

Abstract

Online surveys are becoming more and more popular in the social sciences, for example in electoral research. Plenty of studies have shown that participants who take part in online-surveys differ significantly from participants taking part in face-to-face, or telephone surveys, in terms of their socio-demographic background and political attitudes. Still, since electoral research aims primarily to explain voting behavior, online surveys are deemed useful tools, if the relationship between dependent and independent variables are similar in different survey types, for instance, if vote choice can be described by the same factors, regardless of the survey mode. This paper analyzes these relationships by comparing models of voting behavior which are based on data from different survey modes. Survey data from the German Longitudinal Election Study (GLES) is used since this project simultaneously conducted a face-to-face and telephone survey, as well as online

(GLES) herangezogen, da im Rahmen dieses Projekts im Vorfeld der Bundestagswahl 2009 eine persönlich-mündliche, eine telefonische und mehrere Online-Umfragen nahezu zeitgleich durchgeführt wurden, was ideale Bedingungen für einen Vergleich der verschiedenen Umfragemodi bietet. Diese Untersuchung trägt somit dazu bei, die Eignung von Online-Umfragen für Analysen der empirischen Wahlforschung besser einschätzen zu können.

surveys ahead of the 2009 German Federal Election. Hence, these are ideal conditions for comparing different survey modes. Consequently, this paper enables one to evaluate the use of online surveys for empirical electoral research

1 Einleitung

Lange Zeit galten die persönlich-mündliche, die schriftliche und die Telefonbefragung in der quantitativen empirischen Sozialforschung als beste Möglichkeiten, um mittels Befragungen Daten mit guter Qualität zu generieren, die eine adäquate Auskunft über die Meinungen, Einstellungen und Verhaltensweisen der Bevölkerung geben (vgl. hierzu Dillman 1978). Seit Beginn dieses Jahrhunderts erfreut sich jedoch eine weitere Befragungsmethode zunehmender Beliebtheit: In der Meinungsforschung und auch in der empirischen Sozialforschung spielen Online-Umfragen eine immer größere Rolle (vgl. Dillman 2000; Jakob et al. 2009; Pötschke 2010; Welker/Wenzel 2007; Zerback et al. 2009). Ihre Vorteile liegen klar auf der Hand: Im Vergleich zu den klassischen Umfragetypen können Online-Umfragen schnell und kostengünstig realisiert werden. Die moderne Technik macht es möglich, dass die Daten bereits wenige Tage nach der Erhebung den Forscher/innen zur Verfügung stehen und dadurch erste Forschungsergebnisse zeitnah der Forschungsgemeinschaft präsentiert werden können (Bandilla/Hauptmanns 1998; Baur/Florian 2009).

Trotz dieser Vorteile stehen Online-Umfragen häufig in der Kritik: Insbesondere die Ziehung einer Zufallsstichprobe aus einer definierten Grundgesamtheit erweist sich als problematisch (Baur/Florian 2009; Couper/Miller 2008). Dies ist jedoch die Voraussetzung für die in der Wahlforschung häufig angewandten inferenzstatistischen Verfahren (Pötschke 2010). Neben der Präzision und der Genauigkeit sollte eine Zufallsstichprobe insbesondere das Kriterium der Repräsentativität sicherstellen, was bedeutet, „dass jedes Element einer Grundgesamtheit eine angebbare, von Null verschiedene Auswahlchance hat und sie zufällig ausgewählt werden“ (Pötschke 2010:58). Ziel dabei ist es, dass die Stichprobe die Grundgesamtheit in angemessener Weise abbildet und die Forscher/innen auf der Grund-

lage der Stichprobe Aussagen über die Grundgesamtheit treffen können (Bortz/Döhring 2005:401; Kromrey 2006).

Die Grundgesamtheit in Umfragen der Wahlforschung bildet zumeist die wahlberechtigte Bevölkerung. Lange Zeit galten das persönlich-mündliche Interview und die telefonische Befragung als adäquate Methode, um Daten zu generieren, die durch ein elaboriertes Stichprobendesign ein angemessenes Abbild der wahlberechtigten Bevölkerung wiedergeben. Aufgrund zunehmend niedriger Ausschöpfungsquoten, insbesondere bei Telefonbefragungen, ist jedoch anzuzweifeln, dass dies immer gelingt. Online-Umfragen stehen jedoch im Vergleich zu diesen Umfragen vor einem Problem: Die adäquate Abbildung der wahlberechtigten Bevölkerung in Online-Befragungen ist prinzipiell nur dann möglich, wenn entweder die Online-Penetration weit fortgeschritten ist oder wenn die Rekrutierung grundsätzlich offline erfolgt, wobei Personen ohne Internetanschluss für die Zeit der Befragung mit technischen Mitteln ausgestattet werden, die die Beantwortung von Internet-Umfragen ermöglichen (ADM 2001). Die meisten Online-Umfragen basieren jedoch nicht auf einer solch aufwändigen Rekrutierung. Stattdessen werden einfachere Rekrutierungswege angewandt: Hier kann grundlegend zwischen einer *aktiven* und *passiven Auswahl* der Befragungsteilnehmer unterschieden werden (ADM 2001): Bei einer passiven Auswahl rekrutieren sich die Teilnehmer/innen selbst (sog. Selbstrekrutierung, vgl. Couper 2000). Das Untersuchungsinstitut kann folglich nicht selbst entscheiden, wer befragt werden soll. Diese Art der Stichprobenziehung genügt nach ADM nicht den wissenschaftlichen Anforderungen aufgrund potentieller Stichprobenverzerrungen und mangelnder Repräsentativität. Nur eine aktive Auswahl, bei der das Institut selbst bestimmt, wen es um eine Teilnahme bittet, erfüllt die wissenschaftlichen Standards (ADM 2001).

Um diese aktive Auswahl zu erfüllen, greifen die Erhebungsinstitute häufig auf Befragte aus sogenannten Online- bzw. Access-Panels zurück. Dies ist in der Regel ein Pool an Personen, die sich grundsätzlich bereit erklärt haben, an Online-Umfragen teilzunehmen und von den Online-Panel-Anbietern zu Befragungen eingeladen werden. Befragte werden dabei mittels einer mathematisch-statistischen Zufallsauswahl oder spezifischen, für die Untersuchung relevanten Quotierungen aus diesem Pool ausgewählt (ADM 2001).

Die Online-Panels unterscheiden sich zudem deutlich hinsichtlich der Art der Befragtenrekrutierung: Hierbei ist zwischen *Probability-based* und *Non-probability Samples* zu unterscheiden (vgl. Couper 2000, Couper/Miller 2008). Ein Probability-based Sample zeichnet sich dadurch aus, dass auf der Grundlage einer Zufallsstichprobe (bspw. Telefonstichprobe) die Befragten ausgewählt werden. Wenn eine Person über keinen PC oder keinen Internetanschluss verfügt, wird

ihr die Hardware und die Internetverbindung für den Zeitraum der Untersuchung zur Verfügung gestellt (vgl. Chang/Krosnick 2009; Couper/Coutts 2006). Obwohl diese Verfahrensweise grundsätzlich als Methode der Wahl gehandelt werden kann – Chang und Krosnick (2009) ziehen diese Methode sogar der Telefonbefragung vor – kann diese weder als kostengünstig, noch in der Bereitstellungsphase der Hardware als zeitsparend eingestuft werden (vgl. hierzu auch Bandilla et al. 2001). Demgegenüber bestehen Non-probability Samples aus einem offenen Teilnehmer/innenkreis, wobei hier selbstrekrutierte Umfragen von Umfragen mittels sogenannten Volunteer-Panels bzw. Opt-in-Panels zu unterscheiden sind: Bei Volunteer- bzw. Opt-in-Panels müssen sich die Teilnehmer/innen zunächst bei dem Panelanbieter registrieren (vgl. Bandilla/Hauptmanns 1998; Couper 2000; Yeager et al. 2011), daher ist die Zusammensetzung von Volunteer-Panels nicht zufällig. Bei selbstrekrutierten Umfragen werden die Teilnehmer/innen mittels allgemeiner Aufforderungen oder Hinweise auf eine Umfrage aufmerksam gemacht und nicht das Befragungsinstitut entscheidet, welche Personen zur Umfrage eingeladen werden (sog. passive Auswahl) (ADM 2001).

Diese Unterschiede in der Ziehung der Stichprobe und der Auswahl der Befragungsteilnehmer/innen spiegeln sich auch deutlich in Unterschieden in der Verteilung soziodemographischer Merkmale und politischer Einstellungen zwischen persönlich-mündlichen und telefonischen Befragungen auf der einen und Online-Umfragen auf der anderen Seite wider: Mitglieder in Volunteer-Panels sind zumeist jünger, männlich, höher gebildet und haben eine stärkere Affinität, Parteien des linken Spektrums zu wählen, im Vergleich zur Durchschnittsbevölkerung (vgl. Baur/Florian 2009; Couper 2000; Roster et al. 2004; Sanders et al. 2007; Schenk/Wolf 2006; Wolsing/Faas 2009).

Da die empirische Wahlforschung einerseits zeitnah die Entwicklungen in der Wählerschaft analysieren möchte, jedoch andererseits auf Daten angewiesen ist, mittels derer Aussagen über die wahlberechtigte Bevölkerung gemacht werden können, stellt sich die Frage, ob es aufgrund der Stichprobenziehung und der Befragtenrekrutierung überhaupt sinnvoll ist, Online-Umfragen zur Untersuchung des Wahlverhaltens heranzuziehen. Im Gegensatz zu den kommerziellen Umfrageinstituten besteht das Bestreben der wissenschaftlichen Wahlforschung jedoch weniger in der exakten Darstellung von Häufigkeitsverteilungen, sondern vielmehr in der Erklärung des politischen Verhaltens, insbesondere von Wahlbeteiligung und Wahlentscheidung, in Anbetracht soziodemographischer Merkmale und politischer Einstellungen: Warum und aus welchem Grund sind einige Personen nicht zur Wahl gegangen und welche Wähler/innen haben welche Parteien bzw. Politiker/innen aus

welchen Gründen gewählt? Diese Fragen stehen in der Regel im Mittelpunkt der wissenschaftlichen Wahlforschung.

Aus diesem Blickwinkel können Online-Umfragen für die empirische Wahlforschung durchaus von Nutzen sein. Die Voraussetzung hierfür ist, dass sich die Beziehungen zwischen den abhängigen und unabhängigen Variablen in Auswertungen auf Basis von Online-Umfragedaten ähnlich darstellen wie in Auswertungen auf Basis persönlich-mündlicher sowie telefonischer Umfragedaten. Ob dieser Zusammenhang tatsächlich besteht und empirisch nachgewiesen werden kann, steht im Mittelpunkt des vorliegenden Beitrags. Die Untersuchung beschränkt sich dabei auf Modelle der Wahlbeteiligung und der Wahlentscheidung, da diese die zentralen Explananda der Wahlforschung darstellen. Um die Qualität von Online-Umfragen in Hinblick auf Erklärungen der Wahlbeteiligung und Wahlentscheidung zu untersuchen, werden die Ergebnisse entsprechender Modelle auf Basis von zwei Online-Umfragen mit einer persönlich-mündlichen und einer telefonischen Befragung verglichen. Da eine Vielzahl an Faktoren neben Stichprobenziehung und Ausschöpfungsquote zu Unterschieden zwischen Umfragen führen können, vergleichen wir zusätzlich die persönlich-mündliche und die telefonische Umfrage miteinander. Bewegen sich die Abweichungen zwischen Online- und persönlich-mündlicher bzw. telefonischer Umfrage in ähnlichen Größenordnungen wie die Abweichungen zwischen persönlich-mündlicher und telefonischer Umfrage, gehen wir davon aus, dass Online-Umfragen in Hinblick auf Modelle des Wählerverhaltens vergleichbare Ergebnisse liefern wie diese beiden „klassischen“ Umfrageformen und damit einen deutlichen Nutzen für die empirische Sozialforschung haben.

Wesentlich bei solchen Vergleichen ist, dass möglichst viele Faktoren konstant gehalten werden. Die vorliegende Untersuchung hat hierbei den Vorteil, dass im Rahmen der *German Longitudinal Election Study* (GLES) zur Bundestagswahl 2009 Umfragen mit unterschiedlichen Modi, aber ähnlichen Feldzeiten und weitgehend identischen Frageformulierungen erhoben wurden (vgl. hierzu Schmitt-Beck et al. 2010a).¹ Diese Datenlage ermöglicht einen in seiner Stringenz bislang nicht dagewesenen Vergleich zwischen Online-, Telefon- und persönlich-mündlicher Befragung.

Zur Beantwortung der Frage nach der Qualität von Analysen mit Online-Umfragen wird zunächst ein kurzer Überblick über die bisherige Forschung gegeben. Im Anschluss daran werden die verwendeten Daten und Modelle detailliert

1 Minimale Unterschiede zwischen den Frageformulierungen sind notwendigen Anpassungen an den jeweiligen Erhebungsmodus geschuldet.

vorgestellt. Darauf folgt die Darstellung der Ergebnisse, die im letzten Teil dieses Beitrags zusammengefasst und diskutiert werden.

2 Forschungsstand und theoretischer Hintergrund

Die Beliebtheit von Online-Umfragen hat diese zum Gegenstand zahlreicher methodischer Forschungsarbeiten über deren Datenqualität gemacht (vgl. hierzu Fricker et al 2005; Galesic/Bosnjak 2009; Jakob et al. 2009; Malhotra 2008; Taddicken 2008; Toepoel 2008; Welker/Wenzel 2007). Bei bisherigen Untersuchungen zur Qualität von Online-Umfragen stehen verschiedene Aspekte im Mittelpunkt: So wird grundsätzlich die Frage gestellt, ob es Unterschiede in der Ausschöpfung zwischen klassischen und Online-Umfragen gibt, wie es um die Häufigkeit von Item-Nonresponse oder der Extremität von Antworten in Online-Umfragen bestellt ist, wie lange die Befragten zur Beantwortung spezifischer Items benötigen, ob unter den Interviewabbrechern spezifische Muster vorzufinden sind und inwiefern mittels Quotierung oder Gewichtungen (beispielsweise Propensity Score Weighting) die Qualität von Online-Daten verbessert werden kann (vgl. hierzu Bandilla et al. 2009; Baur/Florian 2009; Blasius/Brandt 2009; Bosnjak et al. 2001; Callegaro/Disogra 2008; Faas 2003a, 2003b; Fricker et al. 2005; Heerwegh 2009; Heerwegh/Loosveldt 2008; Jin 2011; Kaplowitz et al. 2004; Manfreda et al. 2008; Maurer/Jandura 2009; Roßmann et al. 2011; Schonlau et al. 2006; Yeager et al. 2011).

Zahlreiche Studien versuchen auch die Güte der Online-Daten dahingehend zu prüfen, ob und inwiefern sich die Randverteilungen soziodemographischer und bzw. oder politischer Einstellungsvariablen unterscheiden. Hier konnten deutliche Unterschiede in den Randverteilungen sowohl bei den soziodemographischen Merkmalen als auch bei den politischen Einstellungen zwischen Online- und klassischen Umfragen identifiziert werden (vgl. Bandilla et al. 2001; Berrens et al. 2003; Blasius/Brandt 2009; Chang/Krosnick 2009; Faas 2003a, 2003b; Faas/Schoen 2006, 2009; Yeager et al. 2011), wobei Bandilla et al. (2009) auch zeigen konnten, dass bei Befragten mit ähnlichem Bildungshintergrund auch ähnliche Verteilungen in Hinblick auf inhaltliche Themen vorliegen.

Die Replikation von derartigen Häufigkeitsverteilungen ist jedoch nicht das primäre Interesse der empirischen Wahlforschung. Wie bereits angedeutet, werden häufig Modelle des Wahlverhaltens berechnet, die Auskunft darüber geben, wann welche Wähler/innen mit welchen Einstellungen, Meinungen und Verhaltensweisen sich an Wahlen beteiligen bzw. nicht beteiligen („Wahlbeteiligung“) und für welche Partei(en) diese dann am Wahltag votieren („Wahlentscheidung“). Der

Fokus liegt hierbei folglich auf der Untersuchung der Beziehungen zwischen verschiedenen Merkmalen. Diesbezüglich herrscht in der Forschung bisher kein einheitlicher Befund vor: Während das soeben referierte Ergebnis von Bandilla et al. (2009) grundsätzlich darauf hindeutet, dass in Abhängigkeit vom Bildungshintergrund ähnliche Verteilungen bezüglich inhaltlicher Themen vorliegen und somit grundlegend eine Beziehung zwischen verschiedenen Eigenschaften der Befragten besteht, gehen Faas und Schoen (2006, 2009) davon aus, dass die Teilnehmer/innen an Online-Umfragen mit Selbstrekrutierung dem Gegenstand der Umfrage ein stärkeres Interesse entgegenbringen als Teilnehmer/innen an persönlich-mündlichen Umfragen oder Online-Umfragen mit Teilnehmer/innen aus einem Volunteer-Panel. So ist davon auszugehen, dass aufgrund des stärkeren Interesses die Befragten spezifische Vorstellungen haben und daher stärkere Zusammenhänge zwischen unterschiedlichen politischen Einstellungen bestehen als bei Teilnehmer/innen klassischer Umfragen. Dies gilt laut Faas und Schoen (2006) jedoch nur für Online-Umfragen, die auf einer Selbstrekrutierung der Befragten basieren und nicht für Stichproben aus Volunteer-Panels. Die Erwartung, dass die Beziehungen zwischen politischen Einstellungen in selbstrekrutierten Online-Umfragen stärker ausgeprägt sind als in einer persönlich-mündlichen Umfrage, konnte von Faas und Schoen (2006) bestätigt werden. Erstaunlicherweise – und entgegen ihren Erwartungen – zeigt sich ein solcher Effekt in Teilen jedoch auch bei der Online-Umfrage, deren Teilnehmer/innen aus einem Volunteer-Panel gezogen wurden. Die Autoren begründen dies mit Unterschieden in der Feldzeit beider Umfragen. Dennoch wirft dieser Befund die Frage auf, ob Befragte aus Online-Umfragen – ob nun selbstrekrutiert oder im Rahmen eines Volunteer-Panels ausgewählt – generell ein anderes Interesse am Gegenstand der Befragung haben als Teilnehmer/innen klassischer Umfragen. Im Rahmen unserer Untersuchung wäre dann zu erwarten, dass die Beziehungen zwischen Merkmalen in den Online-Umfragen ausgeprägter sind als in der telefonischen und der persönlich-mündlichen Umfrage.

Dem steht der Befund einer Forschergruppe aus Großbritannien entgegen, die Untersuchungen zu genau dieser Frage mit Daten der British Election Study 2001 und 2005 durchgeführt hat (Sanders et al. 2002, 2007). Es wurden Modelle der Wahlbeteiligung und Wahlentscheidung von persönlich-mündlichen Umfragen in direkten Vergleich zu Online-Umfragen (mit Volunteer-Panel-Rekrutierung) gesetzt. Die Forscher/innen konnten so zeigen, dass Modelle des Wahlverhaltens auf Basis unterschiedlicher Datenquellen zu vergleichbaren Ergebnissen hinsichtlich signifikanter Erklärungsfaktoren gelangen und auch die Erklärungskraft dieser Modelle beeindruckende Ähnlichkeiten aufweisen. Dennoch weisen Sanders et al. (2007) auch darauf hin, dass diese Ergebnisse möglicherweise die spezifische Situ-

ation in Großbritannien widerspiegeln und fordern die internationale Forschungsgemeinschaft auf, ähnliche modusvergleichende Studien durchzuführen. Dieser Forderung kamen Stephenson und Crête (2010) bei Provinzwahlen im kanadischen Quebec 2007 nach und haben Modelle des Wahlverhaltens mittels Daten aus Telefonbefragungen mit Online-Umfragen verglichen. Ihre Ergebnisse sind mit denen von Sanders et al. (2007) vergleichbar. Somit lässt sich bisher zusammenfassend sagen, dass keine eindeutigen Befunde vorliegen, die einen gravierenden Unterschied in den Beziehungen zwischen Merkmalen in Modellen der Wahlbeteiligung und Wahlentscheidung zwischen Online-, persönlich-mündlichen oder telefonischen Umfragen erwarten lassen.

Unsere Untersuchung schließt direkt an die Arbeiten von Sanders et al. (2002, 2007) und Stephenson und Crête (2010) an. Wir gehen jedoch über diese Untersuchungen hinaus, indem wir Online-Umfragen nicht nur mit einer persönlich-mündlichen oder einer telefonischen Befragung vergleichen, sondern mit beidem. Zusätzlich vergleichen wir auch die persönlich-mündliche mit der telefonischen Befragung, da wir auch zwischen diesen Umfragemodi leichte Unterschiede erwarten. Dieses Vorgehen ermöglicht uns daher eine bessere Einschätzung dazu, wie gravierend potentielle Unterschiede zwischen klassischen und Online-Umfragen sind. Bevor wir zu den Ergebnissen dieser Untersuchung kommen, stellen wir im folgenden Abschnitt zunächst die verwendeten Daten und Methoden vor.

3 Daten und Methoden

Im Rahmen dieser Untersuchung werden Daten der *German Longitudinal Election Study* (GLES) zur Bundestagswahl 2009 verwendet. Es stehen zwei Online-Umfragen, eine persönlich-mündliche und eine Telefonbefragung zur Verfügung: Die persönlich-mündliche Befragung wurde als computer-assisted personal interview (CAPI) durchgeführt und fand zwischen dem 12. August und dem 26. September 2009 statt, insgesamt wurden 2.173 Personen befragt.² Es handelt sich um eine klassische Querschnittsbefragung mit einer dreistufigen Zufallsstichprobe und

2 Die Ausschöpfungsquote der persönlich-mündlichen Umfrage beträgt 52,4 %, basierend auf der um stichprobenneutrale Ausfälle bereinigten Bruttostichprobe (Studienbeschreibung zu GLES 2009 Vor- und Nachwahl-Querschnitt (Kumulation), ZA5302, v6.0.0). Alle verwendeten Daten und Methodenberichte sind öffentlich zugänglich und können nach einer Registrierung bei <http://www.gesis.org/wahlen/gles/> heruntergeladen werden. Die ZA-Nummern der Umfragen lauten ZA5300 (persönlich-mündliche Querschnittsbefragung), ZA5303 (telefonische Rolling Cross-Section-Umfrage), ZA5338 (Online-Befragung im August 2009) und ZA5339 (Online-Befragung im September 2009).

einem Oversampling von ostdeutschen Befragten (das Verhältnis von ost- zu westdeutschen Befragten beträgt 1:2). Es wurde daher ein entsprechendes Gewicht eingesetzt, das ein Designgewicht zur Korrektur dieses Oversamplings und ein Transformationsgewicht zum Transformieren der Haushalts- in eine Personenstichprobe kombiniert. Da diese Befragung die Basis für ein dreiwelliges Panel darstellt, das sich über die beiden kommenden Bundestagswahlen erstreckt, wurden auch Personen unter 18 Jahren befragt. Um vergleichbare Daten zu erhalten, werden diese Fälle aus dem Datensatz eliminiert, wodurch 2.144 Fälle untersucht werden.

Die telefonische Umfrage (computer-assisted telephone interview, CATI) startete am 29. Juli 2009; vor der Wahl wurden insgesamt 6.008 Fälle realisiert.³ Die beiden Online-Umfragen waren vom 24. August bis zum 1. September 2009 (Online 1, 1.039 Befragte) und vom 18. bis zum 27. September 2009 (Online 2, 1.010 Befragte) im Feld und decken damit nur einen Teil des Zeitraums der persönlichen und der telefonischen Umfrage ab.⁴ Es handelt sich um Quotenstichproben anhand der Merkmale Alter (18-29 Jahre: 25 %, 30-39 Jahre: 20 %, 40-49 Jahre: 25 %, 50-59 Jahre: 15 %, 60 Jahre und älter: 15 %), Geschlecht (weiblich: 50 %, männlich: 50 %) und Bildung (niedrig: 35 %, mittel: 40 %, hoch: 25 %) aus dem von der Respondi AG betriebenen Online-Access-Panel.⁵

Um einschätzen zu können, ob die Verwendung von Anpassungsgewichten in den folgenden Analysen sinnvoll ist, betrachten wir die Abweichungen der vier Umfragen bei den Merkmalen Geschlecht, Alter und Bildung von offiziellen Daten des Statistischen Bundesamtes (Statistisches Bundesamt 2009). Hierfür wird für jede Ausprägung eines Merkmals (in Bezug auf Geschlecht also für männlich/weiblich) die Differenz zwischen den Anteilen dieser Ausprägung in der Bevölkerung

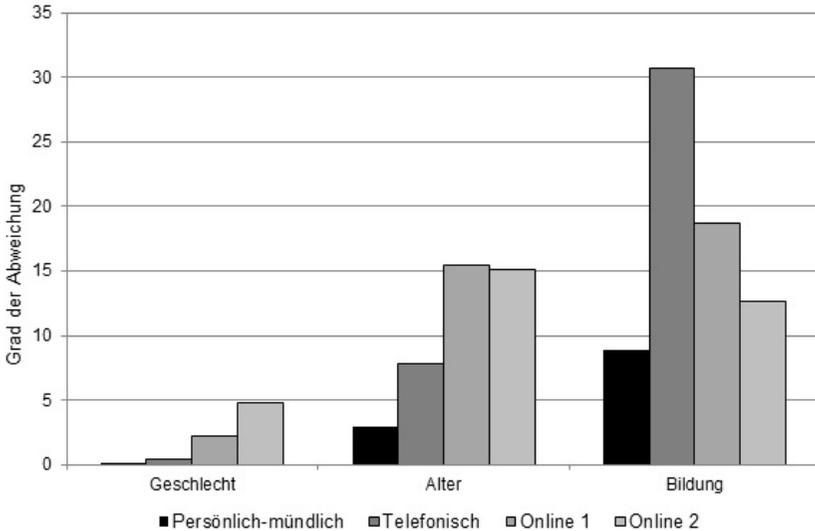
3 Die Ausschöpfungsquote der telefonischen Umfrage beträgt 20,0 %, basierend auf der um stichprobenneutrale Ausfälle bereinigten Bruttostichprobe (Schmitt-Beck et al. 2010b).

4 Bei beiden Online-Umfragen machen wir nur von den Fällen Gebrauch, bei denen wir davon ausgehen können, dass die Befragten die Umfrage ernstgenommen haben. Sogenannte Zeitunterschreiter, die weniger als 60 % der durchschnittlichen Bearbeitungsdauer für die vollständige Bearbeitung der Umfrage benötigten, wurden von der Analyse ausgeschlossen.

5 Die Panelisten für dieses Online-Access-Panel (Volunteer Panel) werden vorwiegend über Respondi-eigene Online-Meinungsportale (z. B. <http://www.sozioland.de>) rekrutiert. Darüber hinaus werden auch Onsite-Befragungen, Suchmaschinen und telefonische Rekrutierung über Partnerinstitute aus der Marktforschung zur Rekrutierung genutzt. Die Quotierung erfolgte in Anlehnung an den Mikrozensus und die Internetpopulation (Studienbeschreibung zu GLES 2009 Langfrist-Online-Tracking, T5, ZA5338, Version 3.0.0). Im Kontext von Online-Umfragen spricht man aufgrund fehlender Zufälligkeit bei der Auswahl der Panelmitglieder von „participation rate“, die den Anteil an nutzbaren Interviews an der Gesamtzahl an Einladungen zu der Umfrage angibt (AAPOR 2010). Diese liegt in der ersten Online-Befragung bei 35,8 % (eigene Berechnung auf Basis der Studienbeschreibung zu GLES 2009 Langfrist-Online-Tracking T5, ZA5338, v3.0.0) und in der zweiten Online-Befragung bei 34,8 % (eigene Berechnung auf Basis der Studienbeschreibung zu GLES 2009 Langfrist-Online-Tracking T6, ZA5339, v2.0.0).

laut Statistischem Jahrbuch und den Anteilen in der Umfrage ermittelt. Diese Differenz wird quadriert, über die Ausprägungen eines Merkmals hinweg summiert und radiziert. Abbildung 1 trägt diese Abweichungen ab.⁶

Abbildung 1 Abweichung der demographischen Merkmale von offiziellen Daten (in Prozentpunkten)



Während in Hinblick auf das Geschlecht der Befragten bei keiner der vier Umfragen große Abweichungen von den offiziellen Daten zu verzeichnen sind, weichen die beiden Online-Umfragen sowohl beim Alter (nach Klassen) sowie der Bildung deutlich von den Daten des Statistischen Bundesamtes ab.⁷ Bei der telefonischen und persönlich-mündlichen Befragung sind die Abweichungen in Hinblick auf das Merkmal Alter gering, die telefonische Befragung sticht jedoch beim Merkmal Bildung mit einer sehr großen Abweichung hervor. Diese Abweichungen legen die Verwendung von verschiedenen Anpassungsgewichten nahe: Bei der telefonischen Umfrage wird im Folgenden in Kombination mit einem Transformationsgewicht ein Gewicht verwendet, das die überproportionale Zahl höher gebildeter Befrag-

6 Die Daten der persönlich-mündlichen Befragung (CAPI) wurden – wie erwähnt – in Hinblick auf das Oversampling ostdeutscher Befragter und in Hinblick auf die Transformation einer Haushalts- in eine Personenstichprobe gewichtet. Letzteres wurde auch bei der telefonischen Befragung (CATI) angewandt.

7 Unterschiede in den Abweichungen zwischen den beiden Online-Umfragen kommen durch Unterschiede bei der Quotenerfüllung zustande.

ter korrigiert. Zudem werden die beiden Online-Umfragen mittels eines Anpassungsgewichts in Anlehnung an den Mikrozensus gewichtet, um Alter, Bildung, Geschlecht und regionale Verteilung an die Allgemeinbevölkerung anzugleichen. Die Gewichtung dient dem Zweck, im Hinblick auf die soziodemographische Struktur möglichst ähnliche Daten zu erhalten.⁸

Zur Feststellung der Nützlichkeit von Online-Umfragen in der an Kausalmodellen interessierten Wahlforschung stellen wir drei Vergleiche an: Zunächst vergleichen wir die beiden Online-Umfragen je mit der persönlich-mündlichen und der telefonischen Umfrage. Hierfür werden die beiden Online-Umfragen zu einem Datensatz zusammengefasst, da das grundsätzlich gleiche Design der beiden Online-Umfragen dies erlaubt. Außerdem vergleichen wir die persönlich-mündliche und die telefonische Umfrage miteinander. Der Grund hierfür ist, dass nicht pauschal davon ausgegangen werden kann, dass keinerlei Abweichungen zwischen Online- und klassischen Umfragen zu finden sind. Diese Abweichungen können auf eine Vielzahl an Faktoren zurückzuführen sein, da die Rahmenbedingungen von Umfragen nicht perfekt konstant gehalten werden können. Dies trifft auch auf den Vergleich zwischen persönlich-mündlicher und telefonischer Umfrage zu. Diese Art des Vergleichs hilft uns folglich bei der Einschätzung, wie gravierend mögliche Abweichungen zwischen Online- und klassischen Umfragen sind.

Um diese drei paarweisen Vergleiche anzustellen orientieren wir uns an dem Vorgehen von Sanders et al. (2002, 2007) und berechnen Modelle der Wahlbeteiligung und Wahlentscheidung zunächst getrennt mit den Daten der jeweiligen Umfrage. Hieran lässt sich sehen, ob trotz unterschiedlicher Datengrundlagen die gleichen Variablen signifikant Einfluss auf die abhängigen Variablen nehmen. Im nächsten Schritt werden je zwei Datensätze (also für Vergleich 1 die persönlich-mündliche und die Online-Umfragen, für Vergleich 2 die telefonische und die Online-Umfragen und für Vergleich 3 die persönlich-mündliche und die telefonische Umfrage) zu einem Datensatz zusammengeführt und die Modelle für Wahlbeteiligung und Wahlentscheidung mit diesen gepoolten Datensätzen berechnet. In diese Modelle werden für jede unabhängige Variable Interaktionseffekte zwischen unabhängiger Variable und Datenbasis eingeführt. Zur Bildung eines Interak-

8 Da die Verwendung von Gewichten bekanntermaßen Einfluss auf die Ergebnisse nehmen kann, haben wir alle Modelle sowohl mit als auch ohne Einschalten der oben genannten Anpassungsgewichte berechnet (die jeweiligen Transformations- und Designgewichte wurden generell verwendet) und konnten nur in sechs von insgesamt 60 Fällen einen Unterschied in der Signifikanz der Moduseffekte feststellen. Diese Unterschiede beziehen sich zudem nur auf die Wirkung der soziodemographischen Merkmale auf Wahlbeteiligung und Wahlentscheidung, also auf eben jene Merkmale, die auch Gewichtungsfaktoren sind. Wir gehen daher von einem zu vernachlässigbaren Effekt der Gewichtung auf die Ergebnisse unserer Analyse hinsichtlich unserer substantiellen unabhängigen Variablen aus.

tionseffekts konstruieren wir jeweils eine Variable, die den Wert 1 für Befragte aus dem einen Umfragetyp und den Wert 0 für Befragte aus dem anderen Umfragetyp annimmt. Diese Variable multiplizieren wir beispielsweise mit der unabhängigen Variablen Parteiidentifikation und bilden so den Interaktionseffekt zwischen dieser unabhängigen Variablen und der Datenbasis.⁹ Ist einer dieser Interaktionseffekte signifikant, bedeutet dies, dass der Effekt der darauf bezogenen unabhängigen Variablen auf die abhängige Variable zwischen den beiden Datenquellen unterschiedlich ist. Ist ein Interaktionseffekt nicht signifikant, gibt es in Hinblick auf die Wirkung dieser unabhängigen Variablen keine signifikanten Unterschiede zwischen den beiden Datenquellen (Sanders et al. 2007: 271). Da uns jedoch bewusst ist, dass Interaktionseffekte in nicht-linearen Regressionen unter bestimmten Voraussetzungen statistische Signifikanz verlieren können (Best/Wolf 2010; Ai/Norton 2003), wenden wir für nicht-signifikante Interaktionseffekte folgendes an den Vorschlag in Best und Wolf (2010) angelehntes Verfahren an: Wir berechnen mit den gepoolten Daten zunächst ein Grundmodell ohne Interaktionseffekt und darauf aufbauend für jeden nicht-signifikanten Interaktionseffekt ein Modell, das diesen zusätzlich enthält. Diese beiden Modelle vergleichen wir in Hinblick auf die Werte des Aikake-Schwarz-Informationskriteriums (AIC) sowie des Bayesianischen Informationskriteriums (BIC) miteinander. Geben beide Informationskriterien einen besseren Modellfit für das Modell mit Interaktionseffekt gegenüber dem Grundmodell an, kann dies als ein deutliches Indiz für einen nicht unwesentlichen Einfluss des Erhebungsmodus auf die Wirkung der unabhängigen auf die abhängige Variable gelten. Deutet lediglich das AIC, das weniger stark für eine Erhöhung der Zahl an Schätzparametern bestraft als das BIC, auf eine Verbesserung des Modellfits durch den Interaktionseffekt hin, kann dies bestenfalls als schwacher Hinweis in Richtung eines Einflusses des Erhebungsmodus interpretiert werden.

In Hinblick auf unsere Analysestrategie ist uns bewusst, dass wir die Ergebnisse nicht daraufhin interpretieren können, welche Datengrundlage das „wahre“ Ergebnis liefert. Hierfür fehlen uns Theorien zur Relation der Effektstärken. Die Erwartung ist jedoch, dass die Unterschiede in den Effektstärken zwischen persönlich-mündlicher und telefonischer Befragung gering sein sollten. Unser Vor-

9 Im Falle des Vergleichs zwischen persönlich-mündlicher und telefonischer Umfrage werden die unabhängigen Variablen mit einer dichotomen Variable multipliziert, die den Wert 1 für Fälle aus der telefonischen Befragung und den Wert 0 für Fälle aus der persönlich-mündlichen Befragung annimmt. Im Falle des Vergleichs zwischen persönlich-mündlicher und Online-Umfragen nimmt diese Variable den Wert 1 an, wenn es sich um Befragte aus der persönlich-mündlichen Umfrage handelt, andernfalls den Wert 0. Beim Vergleich zwischen telefonischer und Online-Umfragen nimmt die Variable den Wert 1 für Befragte aus den Online-Umfragen an, andernfalls 0.

gehen kann daher nicht darüber Auskunft geben, welche Datenquelle das „wahre“ Ergebnis liefert, aber es kann zeigen, ob und inwieweit die Ergebnisse voneinander abweichen.

Da im vorliegenden Beitrag die Qualität von Analysen der Wahlbeteiligung und der Wahlentscheidung mit Daten aus Online-Umfragen untersucht werden soll, richtet sich die Wahl der statistischen Analysemodelle nach den Charakteristika der beiden abhängigen Variablen. Die abhängige Variable im Modell der Wahlbeteiligung ist dichotom codiert mit dem Wert 1 für Wähler/innen und dem Wert 0 für Nichtwähler/innen und geht aus einer sechsstufigen Frage nach der Wahrscheinlichkeit der Wahlbeteiligung hervor.¹⁰ Daher werden logistische Regressionen berechnet. Als unabhängige Variablen werden Geschlecht, Alter (in Klassen) und Bildung als soziodemographische Kontrollvariablen herangezogen. Substantielle unabhängige Variablen sind das Politische Interesse, das Vorliegen einer Parteiidentifikation und der Grad an Unterschieden, der zwischen den Parteien wahrgenommen wird (gemessen anhand der Abweichung zwischen den unterschiedlichen Parteienskalometern bei einem Befragten: umso höher der Wert, umso größer sind die Unterschiede, die zwischen den Parteien wahrgenommen werden, vgl. Johann 2009). Politisches Interesse und Parteiidentifikation werden dichotom gemessen mit dem Wert 1 für das Vorliegen Politischen Interesses bzw. einer Parteiidentifikation und dem Wert 0 andernfalls. Von den drei substantiellen Variablen wird jeweils ein positiver Effekt auf die individuelle Wahlbeteiligung erwartet.

Da bei der Wahlentscheidung zwischen diskreten Alternativen, den Parteien, ausgewählt werden kann, berechnen wir multinomiale Logit-Modelle mit den Entscheidungsalternativen CDU/CSU, SPD, FDP, Bündnis 90/Die Grünen und die Linke. Hierfür haben wir einen „gestapelten“ Datensatz gebildet, in dem jeder Befragte mit fünf Zeilen vertreten ist, für jede Entscheidungsalternative eine. Jede Zeile gibt folglich in Hinblick auf die abhängige Variable an, ob der Befragte für (1) oder gegen diese Partei (0) gestimmt hat. Dies erlaubt uns, für jede unabhängige Variable nur jeweils einen (Interaktions-) Effekt zu berechnen. So gibt beispielsweise der Effekt der Parteiidentifikation auf die Wahlentscheidung die Wirkung dieser Variable im Mittel über alle fünf Alternativen hinweg an. Hierdurch können wir die Modellergebnisse trotz der hohen Zahl an unabhängigen Variablen (insbesondere

10 Die Frage, wie wahrscheinlich es ist, dass ein Befragter zur Wahl gehen wird, hat die folgenden Ausprägungen: (1) Bestimmt, (2) wahrscheinlich, (3) vielleicht, (4) wahrscheinlich nicht, (5) bestimmt nicht, (6) habe bereits Briefwahl gemacht. Die Ausprägungen 1, 2 und 6 wurden zum Wert 1 der dichotomen Wahlbeteiligungsvariable zusammengefasst, die Ausprägungen 3, 4 und 5 zum Wert 0. Da unseres Erachtens die Antwort „vielleicht“ nicht auf eine feste Absicht zur Wahl zu gehen schließen lässt, stellt dies für uns die bestmögliche Unterscheidung zwischen Wählern und Nichtwählern dar.

in den Modellen mit Interaktionseffekten) und Alternativen kompakt darstellen. Da es sich bei unseren Berechnungen um einfache Modelle mit gut abgesicherten substantiellen unabhängigen Variablen handelt, halten wir diese Herangehensweise für angemessen. Als soziodemographische Kontrollvariablen werden wiederum Geschlecht, Alter (in Klassen) und Bildung herangezogen. Substantielle Variablen sind die Identifikation mit der jeweiligen Partei, ob man der Partei die Lösung des derzeit wichtigsten Problems zutraut und die Bewertung der Spitzenkandidat/innen der Partei (basierend auf der Skalometereinstufung, bei der Union wurde der Mittelwert der Spitzenkandidaten von CDU und CSU herangezogen).¹¹ Von allen drei substantiellen Variablen wird ein positiver Effekt auf die Wahrscheinlichkeit, die jeweilige Partei zu wählen, erwartet.

4 Ergebnisse

4.1 Wahlbeteiligung

Bevor die Ergebnisse des Modells zur Erklärung der Wahlbeteiligung vorgestellt werden, wird zunächst kurz auf Unterschiede in der Wahlbeteiligung zwischen den verschiedenen Umfragen im Vergleich zur tatsächlichen Wahlbeteiligung bei der Bundestagswahl 2009 eingegangen: Die Wahlbeteiligung lag bei der Bundestagswahl bei 70,8 %. Mit jeweils knapp 90 % überschätzen die telefonische und die beiden Online-Umfragen die Wahlbeteiligung (trotz Anpassungsgewichtung) deutlich, während dies bei der persönlich-mündlichen Umfrage mit knapp 75 % weniger stark der Fall ist.

Die Überschätzung in den Online-Umfragen trotz fehlendem Interviewerkontakts, der häufiger zu sozial erwünschten Antworten führt, könnte durch ein höheres Politisches Interesse der Befragten begründet sein. Daher ist das Politische Interesse auch eine der unabhängigen Variablen in unserem Modell der Wahlentscheidung. Ob in Hinblick auf Erklärungen der Wahlbeteiligung Unterschiede zwischen den Umfragen vorliegen, kann nur die Berechnung der oben vorgestellten Modelle klären. Tabelle 1 gibt zunächst die Ergebnisse der nach Befragungsmodus getrennten Modelle wieder.

11 Bei der Wahl der substantiellen unabhängigen Variablen haben wir gängige Erklärungsfaktoren für das jeweils zu erklärende Verhalten herangezogen (für die Wahlbeteiligung vgl. z. B. Johann 2009, für die Wahlentscheidung Kellermann 2007), zu denen die notwendigen Informationen in allen verwendeten Datensätzen zur Verfügung stehen.

Tabelle 1 Ergebnisse des Wahlbeteiligungsmodells mit Daten aus unterschiedlichen Befragungsmodi

	Persönlich-mündlich	Telefonisch	Online
Geschlecht	0,06 (0,13)	0,06 (0,09)	0,12 (0,18)
Alter: 30-39	0,69 (0,21)***	0,25 (0,14)	0,03 (0,28)
Alter: 40-49	0,70 (0,19)***	0,74 (0,13)***	0,46 (0,27)
Alter: 50-59	1,13 (0,21)***	0,58 (0,14)***	0,45 (0,29)
Alter: 60 und älter (Referenzkategorie Alter: 18-29)	1,03 (0,19)***	1,26 (0,16)***	0,65 (0,27)*
Bildung: Mittel	0,12 (0,15)	0,53 (0,10)***	0,26 (0,20)
Bildung: Hoch (Referenzkategorie Bildung: Niedrig)	0,73 (0,20)***	1,23 (0,15)***	0,66 (0,26)*
Politisches Interesse	2,40 (0,29)***	1,39 (0,15)***	1,46 (0,25)***
Parteiidentifikation	1,14 (0,13)***	0,75 (0,09)***	1,25 (0,18)***
Parteiendifferential	2,97 (0,41)***	1,12 (0,26)***	2,04 (0,43)***
Konstante	-1,82 (0,28)***	-0,19 (0,19)	-0,51 (0,38)
Log-likelihood	-789,60	-1627,22	-481,70
Pseudo-R ²	0,23	0,14	0,22
AIC	1601,20	3276,45	985,40
BIC	1662,39	3349,25	1044,96
N	1926	5535	1660

Anmerkung: Die Zellen enthalten Koeffizienten logistischer Regression (Standardfehler in Klammern), Signifikanzniveaus: *** $p < 0,001$, ** $p < 0,01$, * $p < 0,05$.

Betrachtet man zunächst die Ergebnisse getrennt nach Befragungsmodus, ist festzustellen, dass die drei substantiellen Variablen Politisches Interesse, Parteiidentifikation und Parteiendifferential in allen drei Umfragen einen signifikanten Effekt auf die Wahlbeteiligung ausüben.¹² Zwischen persönlich-mündlicher und Telefonumfrage bestehen zusätzlich nur wenige Unterschiede hinsichtlich signifikanter Effekte soziodemographischer Merkmale auf die Wahlbeteiligung.

Lediglich in Hinblick auf ein mittleres Bildungsniveau zeigt das Modell, das auf den Daten der Telefonumfrage basiert, einen signifikanten Effekt, das auf den Daten des persönlich-mündlichen Interviews basierende Modell jedoch nicht. Der

12 Beim Vergleich der persönlich-mündlichen und der Online-Umfragen mit der telefonischen Umfrage ergibt sich das Problem, dass die Fallzahlen der Telefonbefragung nahezu dreimal so hoch sind wie die der anderen beiden Umfragen. Da die Fallzahl einen Einfluss auf die Signifikanz von Koeffizienten haben kann, haben wir bei der Telefonumfrage als Test die Fälle jedes zweiten Erhebungstags aus der Analyse herausgenommen. Die Ergebnisse dieses Tests geben keine Anhaltspunkte dafür, dass die Signifikanz der Effekte in den Modellen der Telefonbefragung von der Fallzahl abhängt.

umgekehrte Fall liegt beim Effekt der Altersgruppe der 30- bis 39-Jährigen vor. Dagegen lassen sich bei den Modellen auf Basis der Online-Umfragen kaum signifikante Effekte der soziodemographischen Merkmale auf die Wahlbeteiligung feststellen. Hier liegt folglich ein deutlicher Unterschied zur persönlich-mündlichen und der telefonischen Umfrage vor.

Doch welches Ergebnis bringt nun der paarweise Vergleich der drei Umfragemodi anhand von Interaktionseffekten? Ist einer dieser Interaktionseffekte signifikant, bedeutet dies, dass der Effekt der darauf bezogenen unabhängigen Variablen auf die abhängige Variable zwischen den beiden Datenquellen unterschiedlich ist (Sanders et al. 2007: 271). Gibt es nun signifikante Unterschiede zwischen der Wirkung der unabhängigen Variablen in Abhängigkeit vom Befragungsmodus? Tabelle 2 zeigt die Ergebnisse dieser Berechnungen. Hierbei beschränken wir uns auf die Darstellung der Interaktionseffekte (mit Ausnahme des Modus-Haupteffekts), da die Haupteffekte dieser Regressionen identisch mit den Effekten in den nach Befragungsmodus getrennten Modellen in Tabelle 1 sind.

Überraschenderweise zeigen sich hier deutliche Unterschiede zwischen der persönlich-mündlichen und der telefonischen Umfrage in Hinblick auf die drei substantiellen unabhängigen Variablen, wie sich anhand der signifikanten Interaktionseffekte zeigen lässt. In diesem Ergebnis spiegeln sich die deutlichen Unterschiede in den Effektstärken wider, die in der telefonischen Umfrage durchweg schwächer ausgeprägt sind als in der persönlich-mündlichen. Dagegen sind die Effekte des Bildungsniveaus in der persönlich-mündlichen Umfrage schwächer, was sich wiederum in signifikanten Interaktionseffekten für dieses soziodemographische Merkmal niederschlägt. Im Gegensatz dazu lassen sich weder für den Vergleich zwischen persönlich-mündlicher noch telefonischer Umfrage mit den Online-Umfragen signifikante Interaktionseffekte für die soziodemographischen Merkmale finden. Auch die Unterschiede zwischen den Wirkungen der substantiellen unabhängigen Variablen scheinen weniger stark ausgeprägt zu sein: Beim Vergleich zwischen persönlich-mündlicher und den Online-Umfragen zeigt sich ein signifikanter Unterschied in der Wirkung des Politischen Interesses auf die Wahlbeteiligung, beim Vergleich zwischen telefonischer und den Online-Umfragen liegt hingegen ein signifikanter Unterschied bei der Wirkung des Vorhandenseins einer Parteiidentifikation vor. Bei beiden Vergleichen deutet zudem der oben beschriebene Test nicht signifikanter Interaktionseffekte anhand von Informationskriterien schwach in die Richtung, dass es auch Unterschiede in Hinblick auf die Wirkung des Parteiendifferentials zwischen persönlich-mündlicher bzw. telefonischer und Online-Umfragen gibt (zu Ergebnissen dieser Tests siehe Tabelle A1 im Anhang).

Tabelle 2 Ergebnisse des Wahlbeteiligungsmodells, Vergleiche zwischen den Befragungsmodi

	Vergleich 1: Persönlich- mündlich vs. Online	Vergleich 2: Telefonisch vs. Online	Vergleich 3: Persönlich-mündlich vs. Telefonisch
Geschlecht	-0,06 (0,22)	0,06 (0,20)	0,00 (0,16)
Alter: 30-39	0,66 (0,35)	-0,21 (0,31)	0,45 (0,26)*
Alter: 40-49	0,24 (0,34)	-0,28 (0,31)	-0,03 (0,24)
Alter: 50-59	0,68 (0,36)	-0,13 (0,32)	0,55 (0,26)*
Alter: 60 und älter (Referenzkategorie Alter: 18-29)	0,38 (0,33)	-0,61 (0,31)	-0,23 (0,26)
Bildung: Mittel	-0,13 (0,25)	-0,27 (0,22)	-0,40 (0,18)*
Bildung: Hoch (Referenzkategorie Bildung: Niedrig)	0,07 (0,33)	-0,57 (0,30)	-0,50 (0,25)*
Politisches Interesse	0,94 (0,39)*	0,07 (0,29)	1,01 (0,33)**
Parteiidentifikation	-0,12 (0,22)	0,50 (0,20)*	0,38 (0,16)*
Parteiendifferential	0,93 (0,59)+	0,92 (0,50)+	1,85 (0,49)***
Modus	-1,31 (0,47)**	-0,32 (0,43)	-1,63 (0,34)***
Konstante	-0,51 (0,38)	-0,19 (0,19)	-0,19 (0,19)
Log-likelihood	-1271,30	-2108,93	-2416,82
Pseudo-R ²	0,24	0,16	0,19
AIC	2586,60	4261,85	4877,65
BIC	2722,67	4413,24	5029,83
N	3586	7195	7461

Anmerkung: Die Zellen enthalten Koeffizienten logistischer Regression (Standardfehler in Klammern), abgebildet sind lediglich die Koeffizienten der Interaktionseffekte, Signifikanzniveaus: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + Verbesserung des Modellfit durch Interaktionseffekt (nur bei AIC).

Wie lässt sich dieses Ergebnis vor dem Hintergrund einer nicht unwesentlichen Zahl an signifikanten Effekten soziodemographischer Merkmale in den Modellen auf Basis der persönlich-mündlichen und der telefonischen Umfrage, die in den Online-Umfragen-Modellen nicht signifikant sind, erklären (vgl. Tabelle 1)? Die Interaktionseffekte beziehen sich lediglich auf Unterschiede in der Höhe der Effekte, ob diese Effekte im Zusammenspiel mit dem Standardfehler konventionelle Signifikanzniveaus erreichen, spiegeln sie nicht wider.

In Hinblick auf das Wahlbeteiligungsmodell sind die Ergebnisse folglich gemischt: Positiv ist hervorzuheben, dass die substantiellen unabhängigen Variablen Politisches Interesse, Parteiidentifikation und Parteiendifferential bei der Berechnung mit unterschiedlichen Datenquellen immer signifikante Effekte auf die Wahlbeteiligung zeigen würden, die grundsätzliche Interpretation der Ergeb-

nisse folglich sehr ähnlich wäre. Zwischen der persönlich-mündlichen und der telefonischen Umfrage liegen hierbei jedoch deutliche Unterschiede in den Effektstärken vor, die sich in signifikanten Interaktionseffekten im direkten Vergleichsmodell niederschlagen. Dies ist weniger stark für den Vergleich zwischen den Online-Umfragen und der persönlich-mündlichen sowie der telefonischen Umfrage der Fall. Bevor auf Basis dieser Befunde eine Einschätzung der Nützlichkeit von Online-Umfragen für Untersuchungen des Wahlverhaltens gegeben wird, werden im nächsten Abschnitt zunächst die Ergebnisse der Modelle zur Wahlentscheidung vorgestellt.

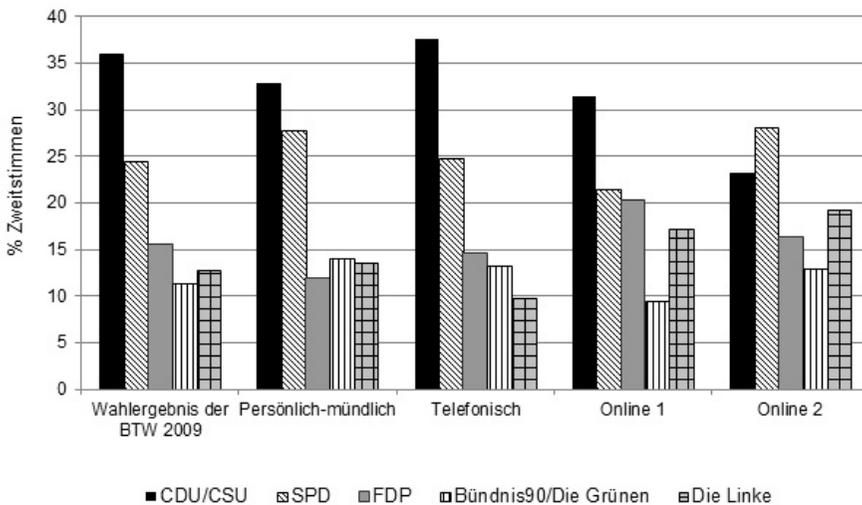
4.2 Wahlentscheidung

Vor der Betrachtung der Ergebnisse des Beziehungsmodells erfolgt ein Vergleich der vier Umfragen mit der Verteilung der Zweitstimmen bei der Bundestagswahl 2009 (Abbildung 2): Während die Verteilungen der Wahlabsicht bei der telefonischen und der persönlich-mündlichen Umfrage auf den ersten Blick dem realen Wahlergebnis sehr ähneln, gibt es im Detail doch einige Unterschiede: Die persönlich-mündliche Umfrage überschätzt den Anteil von SPD sowie Bündnis 90/Die Grünen und unterschätzt den Anteil von CDU/CSU und FDP. Damit kann man sie als leicht linkslastig bezeichnen. Die telefonische Umfrage ist bei SPD und FDP sehr nah am realen Wahlergebnis, überschätzt aber den Anteil der CDU/CSU, und zwar in erster Linie zu Lasten der Linken. Im Vergleich zu den beiden Online-Umfragen sind diese Abweichungen jedoch gering: Während die im August durchgeführte Online-Umfrage (Online 1) den Anteil der FDP und der Linken massiv zu Lasten von CDU/CSU und SPD überschätzt, aber in der Relation der Parteistärken immerhin noch mit dem realen Ergebnis zu vergleichen ist, ist selbst dies bei der direkt vor der Wahl durchgeführten Online-Umfrage (Online 2) nicht der Fall. Hier wäre die SPD die stärkste Partei, gefolgt von der CDU/CSU und der Linken.

Anhand dieser Ergebnisse bestätigt sich der schon mehrmals gezeigte Befund, dass Online-Umfragen für die Darstellung von Verteilungen politischer Einstellungen nur schwer herangezogen werden können. Doch wie sieht es mit den Beziehungen zwischen politischen Einstellungen und Wahlabsicht aus? Tabelle 3 zeigt die Ergebnisse des Wahlentscheidungsmodells auf Basis unterschiedlicher Datenquellen. Betrachtet man zunächst wieder diese Ergebnisse, zeigt sich wiederum, dass die drei substantiellen Variablen Parteiidentifikation, Problemlösungskompetenz und Kandidatenbewertung in allen drei Umfragen signifikante Effekte auf die Wahlentscheidung ausüben. Dagegen haben die soziodemographischen Merkmale nur in wenigen Ausnahmen einen signifikanten Effekt, beispielsweise in

Hinblick auf ein hohes Bildungsniveau in der telefonischen Umfrage. Bei Betrachtung signifikanter Effekte zeigen sich folglich nur geringfügige Unterschiede zwischen den drei Umfragemodi.

Abbildung 2 Wahlabsicht und reales Wahlergebnis



Wie sieht es nun mit den Interaktionseffekten im direkten paarweisen Vergleich aus? Tabelle 4 zeigt die Ergebnisse dieser Berechnungen, wiederum mit Beschränkung auf die Darstellung der Interaktionseffekte (mit Ausnahme des Modus-Haupteffekts), da die Haupteffekte dieser Regressionen identisch mit den Effekten in den nach Erhebungsmodus getrennten Modellen in Tabelle 3 sind.

Auch hier lassen sich so gut wie keine Unterschiede in Hinblick auf die Wirkung der soziodemographischen Merkmale auf die Wahlentscheidung festzustellen, lediglich in Hinblick auf die Wirkung der Bildung geben die Tests der nicht-signifikanten Interaktionseffekte anhand der Informationskriterien einen schwachen Hinweis auf einen Unterschied zwischen telefonischer und Online-Umfragen einerseits und persönlich-mündlicher und telefonischer Umfrage andererseits (vgl. Tabelle A2 im Anhang). Erstaunlicherweise lassen sich die größten Unterschiede in Hinblick auf die substantiellen unabhängigen Variablen wiederum zwischen persönlich-mündlicher und telefonischer Umfrage finden: Sowohl bei der Wirkung der Parteiidentifikation als auch bei der Problemlösungskompetenz zeigen sich signifikante Unterschiede zwischen den beiden Datenquellen. Beim Vergleich zwischen persönlich-mündlicher und Online-Umfragen lässt sich nur ein signifi-

kanter Unterschied in der Wirkung der Parteiidentifikation auf die Wahlentscheidung nachweisen. Dagegen gibt es einen signifikanten Unterschied in der Wirkung der Problemlösungskompetenz zwischen telefonischer und Online-Umfragen sowie einen schwachen Hinweis (basierend auf der Betrachtung der Informationskriterien in einem Modell mit und ohne den Interaktionseffekt zur Kandidatenbewertung, siehe Tabelle A2 im Anhang) darauf, dass es auch Unterschiede in Hinblick auf die Wirkung der Kandidatenbewertung zwischen den beiden Umfragemodi geben könnte.

Tabelle 3 Ergebnisse des Wahlentscheidungsmodells mit Daten aus unterschiedlichen Befragungsmodi

	Persönlich-mündlich	Telefonisch	Online
Geschlecht	-0,05 (0,09)	-0,12 (0,05)*	-0,11 (0,08)
Alter: 30-39	0,19 (0,17)	0,10 (0,09)	0,09 (0,17)
Alter: 40-49	0,22 (0,16)	0,11 (0,08)	0,17 (0,15)
Alter: 50-59	0,31 (0,16)	0,18 (0,08)*	0,26 (0,16)
Alter: 60 und älter (Referenzkategorie Alter: 18-29)	0,12 (0,15)	0,08 (0,08)	0,10 (0,14)
Bildung: Mittel	0,18 (0,11)	0,05 (0,06)	0,17 (0,09)
Bildung: Hoch (Referenzkategorie Bildung: Niedrig)	0,07 (0,12)	0,32 (0,07)***	0,14 (0,10)
Parteiidentifikation	2,56 (0,10)***	2,29 (0,05)***	2,13 (0,09)***
Problemlösungskompetenz	1,73 (0,10)***	1,23 (0,05)***	1,89 (0,09)***
Kandidatenbewertung	3,66 (0,23)***	3,45 (0,13)***	3,61 (0,18)***
Konstante	-5,21 (0,24)***	-5,14 (0,14)***	-4,88 (0,23)***
Log-likelihood	-1877,91	-5936,47	-2226,04
Pseudo-R ²	0,49	0,37	0,42
AIC	3777,83	11894,94	4474,08
BIC	3855,80	11984,54	4552,35
N (Beobachtungen)	8855	25485	9095
N (Befragte)	1771	5097	1819

Anmerkung: Die Zellen enthalten Koeffizienten multinomialer Logit-Modelle (Standardfehler in Klammern), Signifikanzniveaus: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Insgesamt schneiden die Online-Umfragen folglich im direkten Vergleich zu persönlich-mündlicher und telefonischer Umfrage nicht schlecht ab. Die Unterschiede in den Effektstärken sind teilweise sogar geringer als zwischen persönlich-mündlicher und telefonischer Befragung. Darüber hinaus würde man bei einer Interpreta-

tion der Signifikanzen hinsichtlich der substantiellen unabhängigen Variablen auf Basis aller drei Umfragen zu demselben Ergebnis kommen.

Tabelle 4 Ergebnisse des Wahlentscheidungsmodells, Vergleiche zwischen den Befragungsmodi

	Vergleich 1: Persönlich- mündlich vs. Online	Vergleich 2: Telefonisch vs. Online	Vergleich 3: Persönlich-mündlich vs. Telefonisch
Geschlecht	0,06 (0,12)	0,01 (0,09)	-0,07 (0,10)
Alter: 30-39	0,10 (0,24)	-0,00 (0,19)	0,10 (0,19)
Alter: 40-49	0,05 (0,22)	0,07 (0,17)	0,11 (0,18)
Alter: 50-59	0,04 (0,23)	0,08 (0,18)	0,13 (0,18)
Alter: 60 und älter (Referenzkategorie Alter: 18-29)	0,03 (0,20)	0,02 (0,16)	0,05 (0,17)
Bildung: Mittel	0,01 (0,14)	0,12 (0,11)+	0,12 (0,12)+
Bildung: Hoch (Referenzkategorie Bildung: Niedrig)	-0,07 (0,16)	-0,18 (0,12)+	-0,25 (0,14)+
Parteiidentifikation	0,43 (0,13)**	-0,16 (0,10)	0,28 (0,11)*
Problemlösungskompetenz	-0,15 (0,13)	0,66 (0,11)***	0,50 (0,11)***
Kandidatenbewertung	0,05 (0,29)	0,16 (0,23)+	0,21 (0,27)+
Modus	-0,32 (0,33)	0,26 (0,27)	-0,06 (0,28)
Konstante	-4,88 (0,23)***	-5,14 (0,14)***	-5,14 (0,14)***
Log-likelihood	-4103,95	-8162,51	-7814,38
Pseudo-R ²	0,45	0,39	0,41
AIC	8251,91	16369,02	15672,76
BIC	8423,41	16554,94	15858,53
N (Beobachtungen)	17950	34580	34340
N (Befragte)	3590	6916	6868

Anmerkung: Die Zellen enthalten Koeffizienten multinomialer Logit-Modelle (Standardfehler in Klammern), abgebildet sind lediglich die Koeffizienten der Interaktionseffekte, Signifikanzniveaus: *** $p < 0,001$, ** $p < 0,01$, * $p < 0,05$, + Verbesserung des Modellfit durch Interaktionseffekt (nur bei AIC),

5 Zusammenfassung und Schlussfolgerung

Was lässt sich auf Basis dieser Ergebnisse über den Nutzen von Online-Umfragen für Modelle zu Beziehungen zwischen politischen Einstellungen und politischem Verhalten in der Wahlforschung sagen? Zunächst wurde ein unlängst bekanntes Phänomen bestätigt, nämlich dass Online-Umfragen in den Randverteilungen von

soziodemographischen Merkmalen und politischen Einstellungen zum Teil gravierend von „klassischen“ Umfragen abweichen (vgl. hierzu auch Bandilla et al. 2001; Berrens et al. 2003; Blasius/Brandt 2009; Faas 2003a, 2003b; Faas/Schoen 2006, 2009, Yeager et al 2011). Für die Darstellung deskriptiver Ergebnisse sind Online-Umfragen folglich nur schwer heranzuziehen und das trotz Quotenstichprobe und Anpassungsgewichtung.

Anders sieht es in Hinblick auf die Ergebnisse von Modellen zu Wahlbeteiligung und Wahlabsicht aus: Bezüglich der Wahlbeteiligung kann zunächst festgestellt werden, dass in nach Datenquelle getrennten Modellen die Ergebnisse in Hinblick auf die Signifikanz der substantiellen unabhängigen Variablen gleich sind. Bei Betrachtung der Interaktionseffekte treten jedoch deutliche Unterschiede zwischen persönlich-mündlicher und telefonischer Umfrage in Hinblick auf die Effektstärken der substantiellen Variablen auf. Vergleicht man demgegenüber die Online-Umfragen mit der persönlich-mündlichen und der telefonischen Befragung, sind die Unterschiede in den Effektstärken als deutlich geringer einzuschätzen. Ein ähnliches Bild liefern die Modelle zur Wahlentscheidung der Befragten: Die drei substantiellen Variablen haben in allen drei Modellen einen signifikanten Effekt auf die Wahlabsicht. Und auch hier sind die Unterschiede zwischen den Online-Umfragen und der persönlich-mündlichen bzw. der telefonischen Umfrage nicht höher als zwischen den beiden letztgenannten.

Unser Fazit lautet daher, dass die Berechnung von Modellen der Wahlbeteiligung und der Wahlentscheidung unabhängig von der Datenbasis sehr ähnliche Ergebnisse liefert. Neben den vielfältigen weiteren Nutzungsmöglichkeiten von Online-Umfragen (z. B. für die Untersuchung geschlossener Netzwerke oder für experimentelle Studien, vgl. Bandilla/Hauptmanns 1998), stellen sie folglich auch für die Untersuchung von Einflussbeziehungen im Bereich des Wahlverhaltens eine nützliche Datenquelle dar.

Dieses Ergebnis gewinnt dadurch an Gewicht, dass die Online-Umfragen sowohl mit einer persönlich-mündlichen als auch mit einer telefonischen Umfrage verglichen und zusätzlich die beiden letztgenannten Umfragemodi ebenfalls einem Vergleich unterzogen wurden. Die Analyse profitiert hierbei davon, dass sowohl die Erhebungszeiträume als auch die Fragenformate weitgehend konstant gehalten werden konnten, die vorliegende Untersuchung folglich stringenter ist als bisherige Studien. Zusammen mit den bisherigen Befunden, die mit unseren in Einklang stehen, ergibt sich somit ein starkes Argument für die Nutzbarkeit von Online-Umfragen. Dennoch bleiben etliche Fragen offen: Beispielsweise haben wir (auch aufgrund fehlender Daten) grundständige Modelle der Wahlbeteiligung und der Wahlentscheidung berechnet. Die Frage lautet daher, ob die Ergebnisse bei ver-

feinerten Modellen aufrechterhalten werden können. Darüber hinaus haben wir alle Befragten in unsere Modelle einbezogen. Neuere Untersuchungen der Wahlforschung versuchen jedoch Unterschiede in den Einflussstärken unterschiedlicher Faktoren zwischen Wählergruppen zu finden (beispielsweise in Hinblick auf Politisches Wissen). Auch hier ist zu fragen, ob Online-Umfragen durch die doch deutlich unterschiedliche Zusammensetzung der Befragten auch für solche Untersuchungen nützlich sind. Folglich ist weitere Forschung zur Qualität von Online-Umfragen notwendig. Die Ergebnisse bislang zeigen jedoch in eine positive Richtung, sodass Online-Umfragen aufgrund ihrer vielfältigen Vorteile das Potential haben, zu einer der wichtigsten Datenerhebungsformen der empirischen Sozialforschung zu werden.

Literatur

- American Association of Public Opinion Research, 2010: AAPOR Report on Online Panels. Online verfügbar unter: http://www.aapor.org/AM/Template.cfm?Section=AAPOR_Committee_and_Task_Force_Reports&Template=/CM/ContentDisplay.cfm&ContentID=2223, (letzter Zugriff: 16.12.2012).
- Ai, C. und E. C. Norton, 2003: Interaction Terms in Logit and Probit Models. In: *Economics Letters* 80, 123-129.
- Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute (ADM), 2001: Standards zur Qualitätssicherung für Online-Befragungen. Online verfügbar unter: http://www.adm-ev.de/fileadmin/user_upload/PDFS/Onlinestandards_D.PDF (letzter Zugriff: 30.5.2011).
- Bandilla, W. und P. Hauptmanns, 1998: Internetbasierte Umfragen als Datenerhebungstechnik für die empirische Sozialforschung? In: *ZUMA-Nachrichten* 43, 36-53.
- Bandilla, W., M. Bosnjak und P. Altdorfer, 2001: Effekte des Erhebungsverfahrens? Ein Vergleich zwischen einer Web-basierten und einer schriftlichen Befragung zum ISSP-Modul Umwelt. In: *ZUMA-Nachrichten* 49, 7-28.
- Bandilla, W., L. Kaczmarek, M. Blohm und W. Neubarth, 2009: Coverage- und Nonresponse-Effekte bei Online-Bevölkerungsumfragen. In: Jakob, Nikolaus, Schoen, Harald, Zerback, Thomas (Hg.): *Sozialforschung im Internet. Methodologie und Praxis der Online-Befragung*, Wiesbaden: VS Verlag für Sozialwissenschaften, 129-145.
- Baur, N. und M. J. Florian, 2009: Stichprobenprobleme bei Online-Umfragen. In: Jakob, Nikolaus, Schoen, Harald, Zerback, Thomas (Hg.): *Sozialforschung im Internet. Methodologie und Praxis der Online-Befragung*, Wiesbaden: VS-Verlag für Sozialwissenschaften, 109-128.
- Berrens, R. P., A. K. Bohara, H. Jenkins-Smith, C. Silva und D. L. Weimer, 2003: The Advent of Internet Surveys for Political Research: A Comparison of Telephone and Internet Samples. In: *Political Analysis* 11(1), 1-22.
- Best, H. und C. Wolf, 2010: Logistische Regression. In: Wolf, C. und H. Best (Hg.): *Handbuch der sozialwissenschaftlichen Datenanalyse*, Wiesbaden: VS Verlag für Sozialwissenschaften, 827-854.
- Blasius, J. und M. Brandt, 2009: Repräsentativität in Online-Befragungen. In: Weichbold, M., J. Bacher und C. Wolf (Hg.): *Umfrageforschung. Herausforderungen und Grenzen*, Wiesbaden: VS-Verlag für Sozialwissenschaften, 157-177.
- Bosnjak, M., T. L. Tuten und W. Bandilla, 2001: Participation in Web Surveys. A Typology. In: *Zuma-Nachrichten* 48, 7-17.

- Bortz, J. und N. Döring, 2005: *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. Heidelberg: Springer Medizin Verlag.
- Callegaro, M. und C. Disogra, 2008: Computing Response Metrics for Online Panels. In: *Public Opinion Quarterly* 72 (5), 1008-1032.
- Chang, L. und J. A. Kroshnick, 2009: National Surveys Via RDD Telephone Interviewing Versus the Internet. Comparing Sample Representativeness and Response Quality. In: *Public Opinion Quarterly* 73 (4), 641-678.
- Couper, M. P., 2000: Web Surveys: A Review of Issues and Approaches. In: *Public Opinion Quarterly* 64 (4), 464-494.
- Couper, M. und E. Coultts, 2006: Online-Befragung. Problem und Chancen verschiedener Online-Erhebungen. In: Diekmann, A. (Hg.): *Methoden der Sozialforschung*, Sonderheft 44 der KZfSS, Wiesbaden: VS Verlag für Sozialwissenschaften, 217-243.
- Couper, M. P. und P. V. Miller, 2008: Web Survey Methods. In: *Public Opinion Quarterly* 72(5), 831-835.
- Dillman, D. A., 1978: *Mail and Telephone Surveys: The Total Design Method*. New York: Wiley.
- Dillman, D. A., 2000: *Mail and Internet Surveys: The Tailored Design Method*. New York: Wiley.
- Faas, T., 2003a: Offline rekrutierte Access Panels: Königsweg der Online-Forschung? In: *ZUMA-Nachrichten* 53, 58-76.
- Faas, T., 2003b: Umfragen im Umfeld der Bundestagswahl 2002: Offline und Online im Vergleich. In: *ZA-Information* 52, 120-136.
- Faas, T. und H. Schoen, 2006: Putting a Questionnaire on the Web is not Enough – A Comparison of Online and Offline Surveys Conducted in the Context of the German Federal Election 2002. In: *Journal of Official Statistics* 22 (2), 177-190.
- Faas, T. und H. Schoen, 2009: Nur eine Frage der Zeit? Eine Analyse zweier Online-Umfragen zu den Bundestagswahlen 2002 und 2005. In: Schoen, H., H. Rattinger, und O. W. Gabriel (Hg.): *Vom Interview zur Analyse. Methodische Aspekte der Einstellungs- und Wahlforschung*, Baden-Baden: Nomos, 343-360.
- Fricker, S., M. Galesic, R. Tourangeau, und T. Yan, 2005: An Experimental Comparison of Web and Telephone Surveys. In: *Public Opinion Quarterly* 69 (3), 370-392.
- Galesic, M. und M. Bosnjak, 2009: Effects of Questionnaire Length on Participation and Indicators of Response Quality in a Web Survey. In: *Public Opinion Quarterly* 73 (2), 349-360.
- Heerwegh, D., 2009: Mode Differences Between Face-to-Face and Web Surveys: An Experimental Investigation of Data Quality and Social Desirability Effects. In: *International Journal of Public Opinion Research* 21 (1), 111-121.
- Heerwegh, D. und G. Loosveldt, 2008: Face-to-Face versus Web Surveying in a High-Internet-Coverage Population. Differences in Response Quality. In: *Public Opinion Quarterly* 72 (5), 836-846.
- Jackob, N., H. Schoen und T. Zerback (Hg.), 2008: *Sozialforschung im Internet. Methodologie und Praxis der Online-Befragung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Jin, L., 2011: Improving response rates in web surveys with default setting: The effects of default on web survey participation and permission. In: *International Journal of Market Research* 53 (1), 75-94.
- Kaplowitz, M. D., T. D. Hadlock und R. Levine, 2004: A Comparison of Web and Mail Survey Response Rates. In: *Public Opinion Quarterly* 68 (1), 94-101.
- Kellermann, C., 2007: „Trends and Constellations“: Klassische Bestimmungsfaktoren des Wahlverhaltens bei den Bundestagswahlen 1990-2005. In: Rattinger, H., O. W. Gabriel, und J. W. Falter (Hg.): *Der gesamtdeutsche Wähler. Stabilität und Wandel des Wählerverhaltens im wiedervereinigten Deutschland*, Baden-Baden: Nomos, 297-327.
- Kromrey, H., 2006: *Empirische Sozialforschung. Modelle und Methoden der standardisierten Datenerhebung und Auswertung*. Stuttgart: Lucius & Lucius Verlag.

- Johann, D., 2009: Eine Betrachtung der Wahlbeteiligung bei der Bundestagswahl 2005 auf Basis von Rational-Choice-Konzepten. In: Kühnel, S., O. Niedermayer, und B. Westle (Hg.): *Wähler in Deutschland. Sozialer und politischer Wandel, Gender und Wahlverhalten*, Wiesbaden: VS Verlag für Sozialwissenschaften, 424–445.
- Malhotra, N., 2008: Completion Time and Response Order Effects in Web Surveys. In: *Public Opinion Quarterly* 72 (5), 914–934.
- Manfreda, K. L., M. Bosnjak, J. Berzelak, I. Haas und V. Vehovar, 2008: Web surveys versus other survey modes. A meta-analysis comparing response rates. In: *International Journal of Market Research* 50, 79–104.
- Maurer, M. und O. Jandura, 2009: Masse statt Klasse? Einige kritische Anmerkungen zu Repräsentativität und Validität von Online-Befragungen. In: Jakob, N., H. Schoen und T. Zerback (Hg.): *Sozialforschung im Internet. Methodologie und Praxis der Online-Befragung*, Wiesbaden: VS Verlag für Sozialwissenschaften, 129–145.
- Pötschke, M., 2010: Datengewinnung und Datenaufbereitung. In: Wolf, C. und H. Best (Hg.): *Handbuch der sozialwissenschaftlichen Datenanalyse*, Wiesbaden: VS Verlag für Sozialwissenschaften, 41–64.
- Roßmann, J., J. E. Blumenstiel und M. Steinbrecher, 2011: *Why do respondents drop-out from online surveys? Results from follow-up surveys in the German Longitudinal Election Study (GLES)*. Konferenz der European Survey Research Association (ESRA). Lausanne, 18.–22.07. 2011.
- Roster, C. A., R. D. Rogers, G. Albaum und D. Klein, 2004: A Comparison of Response Characteristics From Web and Telephone surveys. In: *International Journal of Market Research* 46 (3), 359–373.
- Sanders, D., H. Clarke, M. Stewart, P. Whiteley und J. Twyman, 2002: *The 2001 British Election Study Internet Poll: a Methodological Experiment*. Konferenzpapier für das Annual Meeting der American Political Science Association, Boston, 27.–31. August 2002.
- Sanders, D., H. Clarke, D. Harold, D., M. C. Stewart und P. Whiteley, 2007: Does Mode Matter for Modeling Political Choice? Evidence from the 2005 British Election Study. In: *Political Analysis* 15 (3), 257–285.
- Schenk, M. und M. Wolf, 2006: Die digitale Spaltung der Gesellschaft. Zur politikorientierten Nutzung des Internets und der traditionellen Medien in den sozialen Milieus. In: Imhof, K., R. Blum, H. Bonfadelli und O. Jarren (Hg.): *Demokratie in der Mediengesellschaft*, Wiesbaden: VS-Verlag für Sozialforschung, 239–260.
- Schmitt-Beck, R., H. Rattinger, S. Roßteutscher und B. Weßels, 2010a: Die deutsche Wahlforschung und die German Longitudinal Election Study (GLES). In: Faulbaum, F. und C. Wolf (Hg.): *Gesellschaftliche Entwicklungen im Spiegel der empirischen Sozialforschung*, Wiesbaden: VS Verlag für Sozialwissenschaften, 141–172.
- Schmitt-Beck, R., T. Faas und A. Wolsing, 2010b: Kampagnendynamik bei der Bundestagswahl 2009: die Rolling Cross-Section-Studie im Rahmen der „German Longitudinal Election Study“ 2009. In: *MZES Arbeitspapier* 134. Mannheim: Universität Mannheim.
- Schonlau, M., A. van Soest, A. Kapteyn und M. P. Couper, 2006: Selection Bias in Web surveys and the Use of Propensity Scores. In: *RAND Working Papers* WR-279, April 2006.
- Statistisches Bundesamt, 2009: *Statistisches Jahrbuch 2009 für die Bundesrepublik Deutschland*. Wiesbaden: Statistisches Bundesamt.
- Stephenson, L. B. und J. Crête, 2010: Studying Political Behavior: A Comparison of Internet and Telephone Surveys. In: *International Journal of Public Opinion Research* 23 (1), 24–55.
- Taddicken, M., 2008: *Methodeneffekte bei Web-Befragungen. Einschränkungen der Datengüte durch ein „reduziertes Kommunikationsmedium“?* Köln: Herbert von Halem Verlag.
- Toepoel, V., M. Das, und A. van Soest, 2008: Effects of Design in Web Surveys: Comparing Trained and Fresh Respondents. In: *Public Opinion Quarterly* 72 (5), 985–1007.
- Welker, M. und O. Wenzel, 2007: *Online-Forschung 2007. Grundlagen und Fallstudien*. Köln: Herbert von Halem Verlag.

- Wolsing, A. und T. Faas, 2009: Können offene Online-Umfragen dynamische Prozesse der Einstellungsentwicklung erfassen? In: Jakob, N., H. Schoen und T. Zerback (Hg.): *Sozialforschung im Internet: Methodologie und Praxis der Online-Befragung*, Wiesbaden: VS-Verlag: 213-232.
- Yeager, D. S., J. A. Krosnick, L. Chang, H. S. Javitz, M. S. Levendusky, A. Simpser und R. Wang, 2011: Comparing the Accuracy of RDD Telephone Surveys and Internet Surveys Conducted with Propability and Non-probability samples. In: *Public Opinion Quarterly* 75(4), 709-747.
- Zerback, T., H. Schoen, N. Jakob und S. Schlereth, 2009: Zehn Jahre Sozialforschung mit dem Internet. Eine Analyse zur Nutzung von Online-Befragungen in den Sozialwissenschaften. In: Jakob, N., H. Schoen und T. Zerback (Hg.): *Sozialforschung im Internet: Methodologie und Praxis der Online-Befragung*, Wiesbaden: VS-Verlag für Sozialwissenschaften, 15-31.

Anschriften der Autorinnen Ina Elisabeth Bieber
Johann Wolfgang Goethe-Universität
Frankfurt am Main
Fachbereich 03 Gesellschaftswissenschaften
Institut für Gesellschafts- und Politikanalyse
Robert-Mayer-Straße 5
60054 Frankfurt am Main
E-Mail: bieber@soz.uni-frankfurt.de

Evelyn Bytzek
Universität Koblenz-Landau
Institut für Sozialwissenschaften
Abteilung Politikwissenschaft
Kaufhausgasse 9
76829 Landau in der Pfalz
E-Mail: bytzek@uni-landau.de

Anhang

Tabelle A1 Ergebnisse des Informationskriteriumtests nicht-signifikanter Interaktionseffekte im Wahlbeteiligungsmodell

	Vergleich 1: Persönlich-mündlich vs. Online		Vergleich 2: Telefonisch vs. Online		Vergleich 3: Persönlich-mündlich vs. Telefonisch	
	AIC	BIC	AIC	BIC	AIC	BIC
	Grundmodell	2582,03	2656,25	4260,89	4343,47	4913,10
Politisches Interesse	-	-	4262,75	4352,21	-	-
Parteiidentifikation	2584,03	2664,43	-	-	-	-
Parteiendifferential	2581,28	2661,68	4256,98	4346,44	-	-
Geschlecht	2583,51	2663,91	4262,89	4352,34	4914,97	5004,91
Alter: 30-39	2582,67	2663,07	4262,88	4352,34	4912,77	5002,70
Alter: 40-49	2583,58	2663,98	4262,89	4352,34	4913,77	5003,70
Alter: 50-59	2582,56	2662,94	4261,82	4351,27	-	-
Alter: 60 und älter	2584,00	2664,40	4260,91	4350,36	4913,79	5003,72
Bildung: Mittel	2583,81	2664,22	4262,50	4351,96	-	-
Bildung: Hoch	2583,94	2664,35	4261,83	4351,29	-	-

Tabelle A2 Ergebnisse des Informationskriteriumtests nicht-signifikanter Interaktionseffekte im Wahlentscheidungsmodell

	Vergleich 1: Persönlich-mündlich vs. Online		Vergleich 2: Telefonisch vs. Online		Vergleich 3: Persönlich-mündlich vs. Telefonisch	
	AIC	BIC	AIC	BIC	AIC	BIC
	Grundmodell	8243,71	8337,25	16400,26	16501,68	15703,31
Parteiidentifikation	-	-	16402,07	16511,93	-	-
Problemlösungskompetenz	8245,51	8346,85	-	-	-	-
Kandidatenbewertung	8245,57	8346,91	16397,34	16507,21	15699,92	15809,69
Geschlecht	8245,25	8346,59	16402,23	16512,10	15704,04	15813,81
Alter: 30-39	8245,60	8346,94	16402,13	16512,00	15705,29	15815,07
Alter: 40-49	8245,71	8347,05	16402,24	16512,10	15705,11	15814,88
Alter: 50-59	8245,70	8347,04	16401,87	16511,73	15705,13	15814,90
Alter: 60 und älter	8245,71	8347,05	16402,25	16512,11	15705,28	15815,06
Bildung: Mittel	8245,65	8346,99	16397,94	16507,80	15700,15	15809,92
Bildung: Hoch	8245,38	8346,72	16396,49	16506,35	15696,45	15806,23

