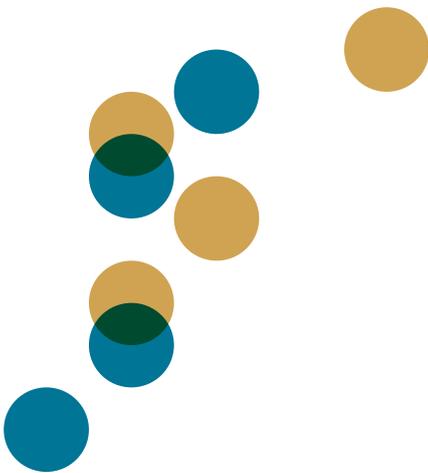


Working Paper

Kampagnendynamik bei der Bundestagswahl 2013

Die Rolling Cross-Section-Studie im
Rahmen der „German Longitudinal
Election Study“ 2013

Julia Partheymüller
Rüdiger Schmitt-Beck
Christian Hoops



*mannheimer zentrum
für europäische sozialforschung*

UNIVERSITÄT
MANNHEIM

Julia Partheymüller
Rüdiger Schmitt-Beck
Christian Hoops

Kampagnendynamik bei der Bundestagswahl 2013

Die Rolling Cross-Section-Studie im
Rahmen der „German Longitudinal
Election Study“ 2013

Arbeitspapiere – Working Papers
Nr. 154, 2013

Mannheimer Zentrum für Europäische Sozialforschung

Partheymüller, Julia:

Kampagnendynamik bei der Bundestagswahl 2013: Die Rolling Cross-Section-Studie im Rahmen der „German Longitudinal Election Study“ 2013 / Julia Partheymüller; Rüdiger Schmitt-Beck; Christian Hoops – Mannheim: 2013 (Arbeitspapiere - Mannheimer Zentrum für Europäische Sozialforschung; 154) ISSN 1437-8574

Not available in book shops.

Token fee: € 3,00

Purchase: Mannheimer Zentrum für Europäische Sozialforschung (MZES)

D – 68131 Mannheim

www.mzes.uni-mannheim.de

Editorial Note:

Dipl.-Pol. Julia Partheymüller ist wissenschaftliche Mitarbeiterin im Projekt „German Longitudinal Election Study (GLES)“ am Mannheimer Zentrum für Europäische Sozialforschung (MZES), Universität Mannheim.

Prof. Dr. Rüdiger Schmitt-Beck ist Professor für Politikwissenschaft – Politische Soziologie an der Universität Mannheim – und Co-Projektleiter der „German Longitudinal Election Study (GLES)“.

Dipl.-Stat. Christian Hoops ist als Research Analyst bei Ipsos Public Affairs tätig. Seine Forschungsschwerpunkte liegen seither im Bereich der quantitativen Methoden, insbesondere Sampling und Gravitationsanalysen.

Abstract

Im Rahmen der „German Longitudinal Election Study“ 2013 wurde der dritte Rolling Cross-Section Survey zu einer Bundestagswahl in Deutschland realisiert. Der Rolling Cross-Section-Survey (RCS) ist eine der Panel-Befragung komplementäre Methode zur Analyse dynamischer Prozesse der Meinungsbildung und des Meinungswandels. Die Grundidee des RCS-Designs besteht darin, die Durchführung einer auf einer Zufallsstichprobe basierenden Querschnittsbefragung in strikt kontrollierter Weise so über einen vorab definierten Zeitraum zu spreizen, dass die Befragten jedes einzelnen Tages jeweils in sich eine Zufallsstichprobe aus der Grundgesamtheit konstituieren. Solche Umfragen ermöglichen in höchst flexibler Weise Analysen dynamischer Phänomene der öffentlichen Meinung. Der Beitrag erörtert detailliert die idealtypischen Eigenschaften von RCS-Studien und diskutiert Probleme ihrer praktischen Umsetzung am Beispiel der RCS-Studie, die im Rahmen der German Longitudinal Election Study (GLES) im Vorfeld der Bundestagswahl 2013 durchgeführt wurde. Gleichzeitig dokumentiert er wesentliche methodische Merkmale dieser Studie, um sie für Nutzer leichter erschließbar zu machen. Analysen der Stichprobenstruktur nach Replikatalter und der Verteilungsstabilität soziodemographischer Merkmale im Zeitverlauf geben Auskunft über die Qualität der gewonnenen Daten. Das inhaltliche Analysepotential von RCS-Daten wird anhand der Entwicklung von Perzeptionen, Erwartungen und Präferenzen der Wähler im Verlauf des Wahlkampfes skizziert.

Contents

1. Einleitung	7
2. Der Rolling Cross-Section Survey als Instrument dynamischer Analyse	8
3. Die RCS-Studie zur Bundestagswahl 2013	11
3.1 Grundkonzept	11
3.2 Feldverlauf	12
3.3 Bearbeitung der Replikate	14
4. Zur Qualität der RCS-Daten	17
4.1 Strukturgleichheit der Tagesstichproben	17
4.2 Soziodemographische Strukturgleichheit	21
5. Analysepotenzial	22
6. Fazit	27
Literatur	29

1. Einleitung

Eines der großen Themen der Forschung über öffentliche Meinung und politisches Verhalten ist seit mehreren Jahrzehnten die Beobachtung, dass die Orientierungen der Bürger an Stabilität verlieren. Ersichtlich wird dies beispielsweise bei Wahlen. So steigt seit Jahrzehnten der Anteil der Wähler, die von Wahl zu Wahl ihre Parteipräferenzen ändern. Ein anderer Indikator ist die wachsende Bedeutung der sogenannten Spätentscheider – jener Personen also, die sich erst während des Wahlkampfes, oft sogar erst am Wahltag selbst, endgültig auf eine Partei festlegen. Umfasste der Anteil dieser Gruppe vor der Bundestagswahl 1983 bei keinem Wahlgang mehr als fünf bis zehn Prozent der Wähler, so wuchs er in den nachfolgenden beiden Jahrzehnten allmählich an, um Ende der 1990er Jahre eine Größenordnung von 15 bis 20 Prozent zu erreichen. Hierzu trug zwar die schwächere parteipolitische Verwurzelung der Bürger aus den neuen Bundesländern durchaus bei. Doch auch im Westen der Republik ist ein deutlicher Trend zu verzögerten Wahlentscheidungen unübersehbar (Schmitt-Beck/Partheymüller 2012). Eine ganz neue Qualität erreichte dieses Phänomen freilich bei der Bundestagswahl 2005, bei welcher der Anteil der Spätentscheider sprunghaft auf über 40 Prozent anstieg – eine Größenordnung, die auch bei der jüngsten Bundestagswahl am 22. September 2013 erneut erreicht wurde.

Wenn so viele Wähler erst kurz vor Wahlen entscheiden, wie sie stimmen werden, dann gewinnen Wahlkämpfe erheblich an Bedeutung. Anders als in früheren Jahrzehnten ist daher heute die Analyse von Wahlkämpfen und ihren Konsequenzen ein unverzichtbarer Bestandteil wahlsoziologischer Forschung (Schoen 2005, Schmitt-Beck 2007). Hierzu bedarf es allerdings neuer methodischer Instrumente. Wahlkämpfe sind hochgradig dynamische Phänomene. Um sie angemessen zu untersuchen, sind Verfahren der Datenerhebung und –analyse vonnöten, welche diese Dynamik einzufangen vermögen. Im vorliegenden Beitrag wird ein solches Verfahren vorgestellt – das so genannte „*Rolling Cross-Section*“-Design (im Folgenden: RCS). Das zentrale Ziel von RCS-Studien besteht darin, dynamische Prozesse der Kristallisierung und Veränderung von Einstellungen und Verhaltensorientierungen abzubilden und zu analysieren. Dazu ist die Durchführung einer auf einer Zufallsstichprobe basierenden Querschnittsbefragung in strikt kontrollierter Weise so über einen vorab definierten Zeitraum zu spreizen, dass nicht nur alle Befragten zusammen, sondern auch die Befragten jedes einzelnen Tages jeweils in sich eine Zufallsstichprobe aus der Grundgesamtheit konstituieren.

Bisherige Anwendungen finden sich vor allem im Bereich der Wahlkampfforschung – auch wenn das RCS-Design für *jede* Form kurzfristigen Wandels der öffentlichen Meinung einen attraktiven Analyseansatz darstellt. Das erste, allerdings noch in wöchentlichem Turnus realisierte RCS-Pilotprojekt wurde 1984 im Rahmen der American National Election Studies (ANES) durchgeführt. Der eigentliche Durchbruch des RCS-Designs setzte die Umstellung auf tägliche Interviews voraus, welche erstmals im Rahmen der Canadian Election Study (CES) 1988 vorgenommen wurde. Seither hat das RCS-Design weltweite Verbreitung gefunden. Es wird heute bei etlichen nationalen Wahlstudien eingesetzt, etwa seit 1988 kontinuierlich in Kanada (CES), in jüngerer Zeit aber auch in Neuseeland (New Zealand Election Study NZES), Großbritannien (British Election Study BES), Italien (Italian National Election Study ITANES), Schweiz (Swiss Electoral Studies SELECTS), Österreich (Austrian National Election Study AUTNES) und den USA (National Annenberg Election Study NAES). Zur Europawahl 2009 wurde sogar ein international vergleichendes RCS-Projekt realisiert.

In Deutschland wurde die erste RCS-Umfrage anlässlich der Bundestagswahl 2005 durchgeführt. Im Rahmen der „German Longitudinal Election Study“ (GLES)¹ wurden bei den Bundestagswahlen 2009

1 Die GLES-Studie (vgl. www.gles.eu) wird als bislang größtes Projekt der deutschen Wahlforschung im Rahmen der Langfristförderung der Deutschen Forschungsgemeinschaft (DFG) unter der Leitung von Hans Rattinger (Universität Mannheim), Sigrid Roßteut-

(Schmitt-Beck/Faas/Wolsing 2010) und 2013 erneut solche Studien realisiert. Der vorliegende Beitrag schließt an die Dokumentationen der beiden früheren Studien an (Schmitt-Beck/Faas/Wolsing 2010, Schmitt-Beck/Faas/Holst 2006) und skizziert zunächst in idealtypischer Perspektive die Eigenheiten des RCS-Designs. Sodann wird erläutert, wie das Design in der RCS-Bundestagswahlstudie 2013 umgesetzt wurde. Hierbei wird auch auf ein neues Verfahren der Quantifizierung der Strukturgleichheit mit einem Effizienzmaß eingegangen. Der Beitrag schließt mit einigen exemplarischen Analysen, welche Rückschlüsse auf die Datenqualität ermöglichen und anhand exemplarischer Befunde zur Bundestagswahl 2013 die besonderen Vorzüge von RCS-Daten verdeutlichen.

2. Der Rolling Cross-Section Survey als Instrument dynamischer Analyse

Wahlkämpfe sind hochgradig dynamische Kommunikationsphänomene, die sich als Sequenzen von Ereignissen darstellen lassen. Diese werden von Parteien und Kandidaten, aber auch von den Medien oder anderen politischen Akteuren (z.B. Verbänden) initiiert. Überdies muss stets mit der Möglichkeit gerechnet werden, dass unvorhersehbare Ereignisse, wie z.B. Naturkatastrophen oder Geschehnisse der internationalen Politik, auf das Wahlkampfgeschehen Einfluss nehmen. Studien zur Analyse von Wahlkampfwirkungen müssen daher längsschnittlich angelegt sein und diese kurzfristige Dynamik abbilden können.

Einer der bedeutendsten methodischen Fortschritte in diesem Bereich war die Entwicklung des *Panel-Designs* mehrfacher Wiederholungsbefragungen derselben Personen durch Paul Lazarsfeld und seine Mitarbeiter Mitte des vergangenen Jahrhunderts (Lazarsfeld/Berelson/Gaudet 1944). Panel-Designs haben den großen Vorteil, Vorgänge des Wandels von Einstellungen, Vorstellungen und Verhaltensorientierungen auf der individuellen Ebene nachvollziehbar zu machen. Allerdings weisen sie auch einige gravierende Nachteile auf. Dazu zählen Konditionierungseffekte (d.h. Lerneffekte bei wiederholten Befragungen zu denselben Gegenständen), aber auch die unvermeidliche Panel-Mortalität und die damit verbundene Selbstselektion der zur erneuten Befragung bereiten Personen, die typischerweise mit studienrelevanten Merkmalen verknüpft ist. Für Wahlkampfforscher ist darüber hinaus die zeitliche „Grobkörnigkeit“ solcher Erhebungen ein Problem. Aus logistischen Gründen ist es schwierig, zahlreiche Panel-Wellen in sehr dichter Folge zu realisieren. Je mehr Zeit aber zwischen den einzelnen Panel-Wellen verstreicht, desto schwieriger ist es, Panel-Effekte eindeutig einem bestimmten Ereignis zuzuschreiben, weil die Zahl der potenziell für solche Effekte verantwortlichen Ereignisse mit der Zeit steigt. Außerdem weist das Instrument hinsichtlich möglicher Untersuchungsgegenstände eine gravierende Einschränkung auf. Der typische Untersuchungsplan, um Wirkungen eines Ereignisses mittels Panel-Daten zu analysieren, ist das Pre-Post-Design, d. h. die Durchführung von Befragungen vor und nach einem mutmaßlich einflussreichen Ereignis. Das ist aber nur möglich, wenn vor Studienbeginn schon klar ist, welches Ereignis untersucht werden soll. Nicht vorhersehbare Vorkommnisse sind mit diesem Instrument ebenso wenig analysierbar wie Ereignisse, die zwar im Prinzip antizipierbar sein mögen, für die aber mangels geeigneter Theorien keine Erwartungen politischer Effekte existieren. Für die in einer solchen Situation erforderlichen explorativen Analysen sind Panel-Designs ungeeignet (Bartels 2006).

Mit dem *Rolling Cross-Section-Design* wurde eine Methode zur Analyse zeitlichen Wandels von Einstellungen und anderen Orientierungen entwickelt und konzeptionell perfektioniert, welche diese Probleme nicht aufweist (Johnston 2001, Romer et al. 2006). Wie bereits erwähnt, besteht die Grundidee solcher

scher (Universität Frankfurt), Rüdiger Schmitt-Beck (Universität Mannheim) und Bernhard Weißels (WZB Berlin) in Kooperation mit der Deutschen Gesellschaft für Wahlforschung (DGfW) durchgeführt und untersucht die Bundestagswahlen 2009, 2013 und 2017 (Schmitt-Beck et al. 2010). Alle in diesem Projekt erhobenen Daten stehen Interessierten für eigene Analysen zur freien Verfügung und werden über GESIS verfügbar gemacht (vgl. www.gesis.org/gles). Die Datensätze der RCS-Studien 2005, 2009 und 2013 tragen die Studiennummern ZA4991, ZA5303, ZA5703.

Erhebungen darin, eine Querschnittsbefragung kontrolliert über einen vorab definierten Zeitraum zu spreizen, so dass *auch die Befragten jedes einzelnen Tages* jeweils in sich eine Zufallsstichprobe aus der Grundgesamtheit konstituieren. Herkömmliche Umfragen können dies nicht leisten. Dort nämlich werden stets bestimmte Zielpersonen eher zu Beginn einer mehr oder weniger lange dauernden Feldperiode, andere dagegen erst zu einem späteren Zeitpunkt befragt. Das ist unter anderem eine Folge der unterschiedlichen Erreichbarkeit dieser Befragungspersonen. Manche von ihnen können durch Umfrageinstitute leicht kontaktiert und interviewt werden, weil sie sich oft zuhause aufhalten und über ihre Zeit relativ frei disponieren können. Andere sind hingegen schwer erreichbar, weil sie aufgrund ihrer Berufstätigkeit und vielleicht auch aktiver Freizeitgestaltung selten in ihrer Wohnung anzutreffen sind, und noch schwieriger wird es, wenn solche Personen alleine leben. Dann nämlich fällt die Nichterreichbarkeit von Zielpersonen sogar mit der Nichterreichbarkeit des gesamten Haushaltes zusammen, und es besteht nicht einmal die Möglichkeit, Befragungstermine für die Zielperson mit anderen Haushaltsmitgliedern zu vereinbaren. Das RCS-Design trägt diesen praktischen Rahmenbedingungen der Umfrageforschung Rechnung und erlaubt es, trotz dieser Widrigkeiten Stichproben so zu ziehen, dass sie für jeden einzelnen Befragungstag eine eigenständige Zufallsstichprobe konstituieren.

Um die Beobachtung der öffentlichen Meinung im Zeitverlauf noch weiter zu standardisieren, wurde darüber hinaus in den meisten bisherigen RCS-Studien versucht, jeden Tag ungefähr gleich viele Interviews zu realisieren. Das ist jedoch nicht zwingend; die täglichen Fallzahlen können auch, wenn dies für das Untersuchungsziel angemessen erscheint, in designkompatibler Weise variiert werden. Zumal für konstantere Tagesfallzahlen auch stärker variierende Replikatsgrößen notwendig sind, die wiederum eine hohe Strukturgleichheit gefährden können. Von zentraler Bedeutung ist hingegen die Zufälligkeit und damit gleichzeitig auch Repräsentativität der täglich realisierten Interviews. Die für die inferenzstatistische Absicherung von Befunden zwingend notwendige Zufallsauswahl von Befragten kommt bei RCS-Studien auf doppelte Weise zum Tragen: Erstens im Hinblick auf die Chance eines Mitglieds der Grundgesamtheit, überhaupt in die Studie einbezogen zu werden, zweitens im Hinblick auf den Zeitpunkt innerhalb des Erhebungszeitraumes, an dem eine ausgewählte Befragungsperson interviewt wird. Zentrale Parameter, um diese kennzeichnenden Merkmale von RCS-Studien zu gewährleisten, sind die Stichprobenziehung und die Feldadministration:

Die Stichprobenziehung (präziser: die Ziehung der Haushaltsstichprobe) erfolgt bei einer RCS-Studie in zwei Stufen. Zunächst wird eine normale Zufallsstichprobe gezogen. Diese wird dann jedoch nicht wie üblich sofort komplett oder in unsystematischer – oftmals rein feldlogistischen Erwägungen folgender – Weise über einen gewissen Zeitraum freigegeben. Vielmehr wird sie in einem zweiten Schritt *zufällig* in Teilstichproben (so genannte *Replikate*) unterteilt. Jedes Replikat stellt ebenfalls eine zufällige Stichprobe aus der Grundgesamtheit dar. Diese Teilstichproben werden dann nach einem festgelegten „Fahrplan“ jeweils an einem bestimmten, zufällig ausgewählten Tag für die Befragung freigegeben.

Nach der Freigabe wird für jedes Replikat eine Bearbeitungsroutine in Gang gesetzt, die einem strengen Protokoll folgt. Diese hat das Ziel, durch intensive Feldarbeit jede Teilstichprobe möglichst gut auszuschöpfen. Dies ist notwendig, weil in der Feldpraxis nicht alle Mitglieder eines Replikats umgehend am Tag der Freigabe interviewt werden können. Das wird nur für leicht erreichbare (und befragungsbereite) Personen der Fall sein, nicht jedoch für schwerer erreichbare Personen. Mit diesen kommen Interviews typischerweise erst nach mehreren vergeblichen Kontaktversuchen zustande. Essentiell für das RCS-Design ist nun, dass dieses Protokoll für alle Replikate in einheitlicher Weise abgearbeitet wird. Für jedes Replikat gelten also identische Kontaktierungsregeln; weder der Start- oder Wochentag noch sonstige Gesichtspunkte spielen dabei eine Rolle. Diese Regeln beinhalten einen genau festgelegten, für alle Replikate gleich langen, mehrtägigen Zeitraum, während dem die Kontaktdaten eines Replikats aktiv bleiben und Kontaktversuche unternommen werden, für die im Hinblick auf Anzahl und Zeitpunkte eine festgelegte Routine einzuhalten ist. Würden diese Abläufe geändert, z.B. um Tagesstichproben exakt gleicher Größe

zu erreichen, würden in Abhängigkeit der Erreichbarkeit nicht mehr alle Mitglieder der Grundgesamtheit aus der Nettostichprobe mit nahezu derselben Wahrscheinlichkeit in die jeweilige Tagesstichprobe inkludiert und die Verteilung des Replikalters würde variieren. Damit könnte die für das RCS-Design zentrale Annahme der völligen strukturellen Gleichartigkeit aller Replikate nicht mehr aufrechterhalten werden.

Die dichte zeitliche Taktung von RCS-Studien auf Tagesbasis und die Notwendigkeit strikter Kontrolle über den Befragungszeitpunkt präjudiziert den Erhebungsmodus. Die Implementation des RCS-Designs wurde erst durch das Aufkommen der Methode der computergestützten telefonischen Befragung Anfang der 1980er Jahre ermöglicht. Durch eine zentralisierte Stichprobensteuerung und Feldkontrolle war es möglich die Clearingperioden der täglichen Stichproben kontrolliert so zur Überlappung zu bringen, dass an jedem einzelnen Tag sowohl leicht erreichbare als auch schwer erreichbare Personen gleichermaßen befragt werden können. Ein derart kontrolliertes Vorgehen bei der Feldarbeit ist im Rahmen von dezentral organisierten persönlich-mündlichen Befragungen logistisch kaum durchführbar. In den letzten Jahren wurden auch einige RCS-Studien online realisiert (Johnston 2008, Faas/Blumenberg 2012), doch müssen bei diesem Modus Abstriche im Hinblick auf das Ziel der Repräsentativität der Stichproben für die Grundgesamtheit der wahlberechtigten Bevölkerung akzeptiert werden. Bei genauer Umsetzung der beschriebenen methodischen Anforderungen erhält man eine Umfrage, die alle Eigenschaften einer normalen Querschnittsstudie auf der Grundlage einer Zufallsstichprobe besitzt und auch als solche analysiert werden kann. Zugleich aber kann sie unter zeitlichen Aspekten in beliebiger und völlig flexibler Weise in ebenfalls zufällige Unterstichproben für beliebige Zeitpunkte oder -perioden zerlegt werden. Durch diese Eigenschaft von RCS-Daten werden höchst differenzierte Analysen des dynamischen Geschehens im Verlauf von Wahlkämpfen möglich. Die Replikate, die bei der Erhebung von RCS-Daten zentrale Bedeutung haben, treten bei diesen Analysen nicht mehr in Erscheinung. Die Auswertungen orientieren sich vielmehr an den Kalendertagen, an denen die Interviews realisiert werden. Auf ihrer Grundlage können auch „synthetische Querschnittsstichproben“ gebildet werden, die alle Interviews beinhalten, die innerhalb eines vom Forscher definierten Zeitraums durchgeführt wurden.

Die Logik des RCS-Designs gewährleistet, dass die für jeden Erhebungstag zur Verfügung stehenden Interviews jeweils identische Mischungen von Befragten aus mehreren Replikaten und damit von leicht, aber auch schwer erreichbaren Personen beinhalten. Mit Blick auf die Stichprobe insgesamt ist der Befragungstag innerhalb der Feldzeit damit nicht mehr systematisch, sondern zufällig bedingt. Die an den verschiedenen Erhebungstagen realisierten Interviews unterscheiden sich durch nichts außer dem Datum ihrer Durchführung bzw. – präziser ausgedrückt – nur hinsichtlich des situativen Kontextes im Hinblick auf das Wahlkampfgeschehen, das diesem Datum entspricht. Alle Änderungen von Wahrnehmungen, Einstellungen, Präferenzen oder Verhaltensorientierungen, die im Zeitverlauf identifiziert werden, können daher nur zwei Quellen haben: den zufälligen Stichprobenfehler und realen Wandel unter dem Eindruck des Kampagnengeschehens. Diese können durch geeignete statistische Verfahren separiert werden (Johnston/Brady 2002, Brady/Johnston/Sides 2006, Romer et al. 2006)

Durch RCS-Erhebungen gewonnene Daten können in äußerst flexibler und vielfältiger Weise analysiert werden. Die einfachste Auswertungsvariante besteht darin, Trends der öffentlichen Meinung im Wahlkampfverlauf zu ermitteln (vgl. z.B. Schmitt-Beck 2009, Schmitt-Beck/Faas 2006). Man kann die Daten aber auch auf der Basis einzelner Erhebungstage analysieren und damit punktgenaue Verbindungen zwischen bestimmten Wahlkampfeignissen und korrespondierenden Änderungen von Wahrnehmungen oder Einstellungen herstellen (vgl. z.B. Schmitt-Beck/Tenscher 2008). Ebenso leicht ist es möglich, die Daten nach Formalkriterien zu gruppieren, beispielsweise wochenweise, um zu höheren Fallzahlen und damit höherer Teststärke zu gelangen. Schnittpunkte können aber auch in beliebiger Weise nach inhaltlichen Kriterien gesetzt werden, etwa in Form eines Vergleiches der Interviews unmittelbar vor und unmittelbar nach einem bestimmten Ereignis im Wahlkampf. Ein besonderer Vorzug ist dabei, dass dies keineswegs nur für antizipierbare Ereignisse wie z.B. Parteitage oder TV-Duelle gilt, sondern für jedes

beliebige Ereignis, einschließlich solcher, die im Vorfeld überhaupt nicht vorhergesehen werden können. Eine weitere interessante Analysevariante besteht in der Anreicherung von RCS-Datensätzen mit externen Daten, die über das Erhebungsdatum als Schlüssel problemlos zugespielt werden können. Beispiele sind Daten zur Medienbewertung politischer Akteure aus Inhaltsanalysen der tagesaktuellen Berichterstattung (Bevan/Krewel 2013) oder aktuelle demoskopische Daten aus Medienumfragen (Faas/Schmitt-Beck 2007), deren Effekte auf Wahrnehmungen und Einstellungen der Wähler sich so auf elegante Weise ermitteln lassen. Bei ausreichenden täglichen Fallzahlen ist es sogar möglich zu analysieren, wie und aufgrund welcher Ereignisse sich die Stärke der Effekte bestimmter Prädiktoren auf das Wählerverhalten im Verlauf von Wahlkämpfen ändert (Johnston/Hagen/Hall Jamieson 2004). Auch hinsichtlich der einsetzbaren statistischen Auswertungsverfahren bieten RCS-Daten flexible Möglichkeiten. Sie können sowohl auf der Individualebene als auch – durch Verfahren der Zeitreihenanalyse – auf der Aggregatebene analysiert werden. Für Aggregatanalysen von RCS-Daten bieten sich überdies in besonderer Weise graphische Verfahren der Datenanalyse an (Romer et al. 2006).

Bei all diesen Vorteilen darf nicht übersehen werden, dass RCS-Umfragen im Vergleich zu Panel-Befragungen auch einen Nachteil aufweisen: Sie erlauben zunächst keine Aussagen über individuellen Wandel, sondern nur Aggregataussagen über Prozesse des Wandels in der gesamten Wählerschaft oder in Wählergruppen. Von daher bietet es sich an, RCS-Studien durch eine Panel-Komponente anzureichern, welche genau dies ermöglicht. Panel- und RCS-Studien können als komplementär gelten und ergänzen einander. Die Kombination einer RCS-Erhebung als Vorwahlbefragung mit einer als Nachwahlbefragung konzipierten zweiten Panel-Welle hat sich als ideales Verfahren erwiesen, um die Kampagnendynamik bei Wahlkämpfen zu analysieren (Johnston 2001, Schmitt-Beck/Faas/Holst 2006). Aus diesem Grund wurde entschieden in die „German Longitudinal Election Study“ (GLES), die bislang umfangreichste und komplexeste deutsche Wahlstudie, neben zahlreichen anderen Komponenten (vgl. www.gles.eu) auch eine RCS-Umfrage einzuschließen, die durch eine als Nachwahlbefragung realisierte zweite Panel-Welle angereichert wurde. Die Variablen der Nachwahl-Welle können als Messungen des kumulativen Einflusses des gesamten Wahlkampfes interpretiert und als solche mit unterschiedlich weit in der Wahlkampfperiode zurückliegenden Vorwahl-Messungen derselben Variablen verglichen werden. Überdies lassen sich verschiedene für die Wahlkampfdynamik wesentliche Daten erst nach der Wahl sinnvoll erheben, nicht zuletzt natürlich die tatsächliche Wahlentscheidung selbst.

3. Die RCS-Studie zur Bundestagswahl 2013

In idealtypischer Betrachtung zeichnet sich das RCS-Design durch eindrucksvolle Eleganz aus. Seine praktische Umsetzung ist jedoch ungewöhnlich anspruchsvoll und stellt Forscher als auch Erhebungsinstitute vor erhebliche Herausforderungen. Im Folgenden wird beschrieben, wie die im Rahmen des GLES-Projektes durchgeführte RCS-Studie realisiert wurde.

3.1 Grundkonzept

Die RCS-Studie zur Bundestagswahl 2013 wurde als CATI-Erhebung in der Zeit vom 08. Juli bis zum 21. September 2013 durchgeführt; sie umfasste also einen Zeitraum von 76 Tagen und endete am letzten Tag vor der Bundestagswahl. Im Rahmen der Studie wurden 7.882 zufällig ausgewählte Personen aus der Grundgesamtheit der deutschsprachigen, in Privathaushalten mit mindestens einem Festnetzanschluss lebenden Bevölkerung der Bundesrepublik Deutschland, die zur Bundestagswahl 2013 wahlberechtigt war, befragt. Die unmittelbar nach der Bundestagswahl realisierte 2. Panelwelle erbrachte 5.353 Wiederholungsinterviews. Die Feldarbeit oblag dem Institut Ipsos GmbH (Hamburg). Die Stichprobenziehung

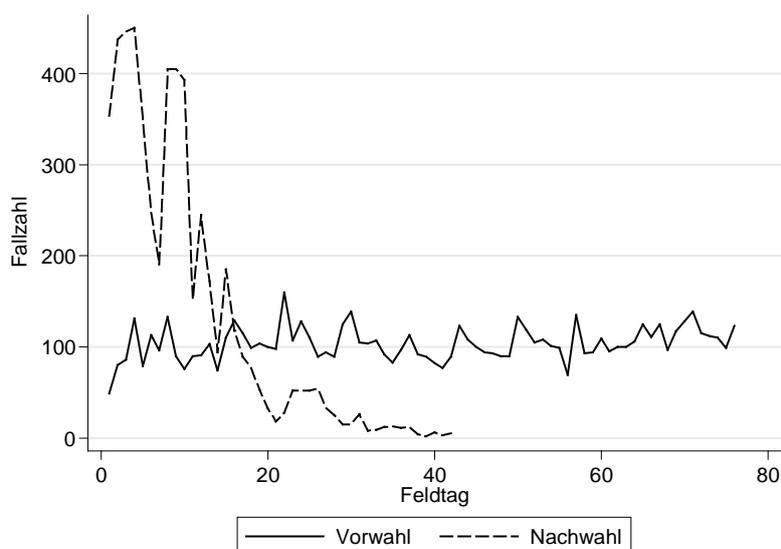
erfolgte als ADM-Telefonstichprobe gemäß dem Gabler/Häder-Modell. Angestrebt war ein Tagesmittel von 100 Interviews. Um ausreichend Reservenummern vorzuhalten, wurde eine sehr umfangreiche Bruttostichprobe gezogen, die 239.400 Nummern umfasste. Diese wurden zufällig in 76 Replikate zu je 3.150 Telefonnummern unterteilt. Insgesamt kamen 211.575 Nummern zum Einsatz. Nach Abzug der stichprobenneutralen Ausfälle (Nummer existiert nicht; kein Privathaushalt; keine Zielperson, d.h. bei der Bundestagswahl 2013 wahlberechtigte Person, im Haushalt; Sprachprobleme; Anschluss einem Fax oder Modem zugeordnet) umfasste die Nettostichprobe 49.896 Nummern. Somit betrug die Ausschöpfung aus der Nettostichprobe 15,8 Prozent. Gegenüber 2009 (20,0 Prozent) hat die Ausschöpfung somit leicht abgenommen.

Um auf schwankende Ausschöpfungsquoten an bestimmten Wochentagen reagieren zu können, ohne die RCS-Logik aufgeben zu müssen, wurden die Replikate ihrerseits wiederum zufällig in „Scheiben“ zu je 100 Telefonnummern zerlegt. Auch diese Teilstichproben waren aufgrund ihrer zufälligen Ziehung beliebig kumulierbar; es handelte sich gleichsam um „Replikate zweiter Ordnung“. Dadurch wurde die Möglichkeit geschaffen, die täglichen Teilstichproben in einer dem RCS-Design angemessenen Weise ad hoc durch Hinzuspielen oder Weglassen einzelner Scheiben zu vergrößern oder zu verkleinern, um so die erreichten Tagesfallzahlen feiner steuern zu können. Die Festlegung der Zahl der an jedem Erhebungstag neu einzuspielenden Scheiben erfolgte jeweils zeitnah in Reaktion auf die aktuelle Feldentwicklung und die bis dahin gesammelten Erfahrungen mit dem Verlauf der Erhebung. Dies ist eine bei RCS-Studien übliche Vorgehensweise; solche Studien erfordern grundsätzlich ein sehr dichtes (d.h. tägliches) Feldmonitoring und je nach Feldverlauf gegebenenfalls auch kontinuierliches Nachsteuern.

3.2 Feldverlauf

RCS-Studien können in der Erhebungsphase also keinesfalls „sich selbst überlassen“ werden. Die Feldentwicklung bedarf permanenter Beobachtung und gegebenenfalls zeitnaher Adjustierung der Zahl der täglich eingespielten Scheiben. Ansonsten bestünde ein erhebliches Risiko, das Ziel eines gleichmäßigen Feldverlaufs auf dem angestrebten Niveau täglicher Fallzahlen zu verfehlen.

Abbildung 1: Tagesfallzahlen RCS-Studie 2013 (im Vergleich zur Nachwahlwelle)



Zum Vergleich sind auch die täglichen Fallzahlen der Nachwahl-Welle dargestellt, die – beginnend am Montag nach der Wahl – wie eine typische Querschnittsbefragung durchgeführt wurde: Ihr Ziel war die Abarbeitung der Wiederholungsbefragungen in möglichst kurzer Zeit. Der Kurvenverlauf zeigt an, dass dieses Ziel erreicht wurde: Insgesamt konnten 80 Prozent der Nachwahl-Interviews in den ersten zwei Wochen nach der Bundestagswahl durchgeführt werden. Danach fiel die Zahl der täglichen Interviews stark ab und näherte sich im Zeitverlauf zusehends der Nulllinie an. Nach insgesamt 42 Erhebungstagen wurde das Feld geschlossen.

Tabelle 1: Anzahl der pro Feldtag aktivierten Scheiben

Replikant	Datum	Anzahl Scheiben	Replikant	Datum	Anzahl Scheiben
1	08.07.2013	27	39	15.08.2013	23
2	09.07.2013	38	40	16.08.2013	23
3	10.07.2013	38	41	17.08.2013	27
4	11.07.2013	38	42	18.08.2013	27
5	12.07.2013	32	43	19.08.2013	25
6	13.07.2013	35	44	20.08.2013	25
7	14.07.2013	32	45	21.08.2013	25
8	15.07.2013	28	46	22.08.2013	25
9	16.07.2013	28	47	23.08.2013	25
10	17.07.2013	28	48	24.08.2013	27
11	18.07.2013	28	49	25.08.2013	27
12	19.07.2013	30	50	26.08.2013	25
13	20.07.2013	33	51	27.08.2013	25
14	21.07.2013	33	52	28.08.2013	25
15	22.07.2013	30	53	29.08.2013	25
16	23.07.2013	31	54	30.08.2013	25
17	24.07.2013	31	55	31.08.2013	27
18	25.07.2013	31	56	01.09.2013	27
19	26.07.2013	31	57	02.09.2013	25
20	27.07.2013	33	58	03.09.2013	25
21	28.07.2013	33	59	04.09.2013	25
22	29.07.2013	30	60	05.09.2013	25
23	30.07.2013	28	61	06.09.2013	25
24	31.07.2013	28	62	07.09.2013	27
25	01.08.2013	27	63	08.09.2013	27
26	02.08.2013	26	64	09.09.2013	26
27	03.08.2013	29	65	10.09.2013	28
28	04.08.2013	29	66	11.09.2013	28
29	05.08.2013	27	67	12.09.2013	28
30	06.08.2013	27	68	13.09.2013	28
31	07.08.2013	25	69	14.09.2013	30
32	08.08.2013	25	70	15.09.2013	30
33	09.08.2013	25	71	16.09.2013	28
34	10.08.2013	25	72	17.09.2013	28
35	11.08.2013	25	73	18.09.2013	26
36	12.08.2013	23	74	19.09.2013	28
37	13.08.2013	23	75	20.09.2013	28
38	14.08.2013	23	76	21.09.2013	30

Tabelle 1 dokumentiert die Anzahl der pro Tag aktivierten Scheiben während der Vorwahl-Studie. Nachdem am ersten Feldtag zunächst 2700 Nummern aktiviert wurden, waren es an den drei darauf folgenden Tagen jeweils 3800 Nummern, um den Bereich der angestrebten Durchschnittsfallzahl möglichst rasch zu erreichen. An Tag 4 wurden mehr als 100 Interviews erzielt, woraufhin die Scheibenzahl schrittweise wieder reduziert wurde. Mit Hilfe detaillierter täglicher Feldberichte wurde das Feldgeschehen von Beginn an eng überwacht. Sofern auf Basis der aktuellen Feldentwicklung eine Anpassung des Nummerneinsatzes

als geboten erschien, erfolgte im Rahmen von Telefonkonferenzen eine Abstimmung zwischen Studienleitung und Erhebungsinstitut über Höhe und Timing der Anpassung, unter Berücksichtigung der Implikationen für die Strukturgleichheit der Stichproben. Diese Praxis der behutsamen Justierungen führte, wie die Abbildung zeigt, zu einem recht gleichmäßigen Feldverlauf mit im Mittel rund um die 100 realisierten Interviews pro Tag.

3.3 Bearbeitung der Replikate

Nicht nur auf der soeben beschriebenen Ebene einzelner Feldtage, sondern auch eine Ebene darunter, nämlich bei der Bearbeitung der einzelnen Replikate am Tag und in den Tagen nach ihrer Freigabe, verlangt das RCS-Design ein strikt standardisiertes Vorgehen. Durch intensive Feldarbeit soll jede Teilstichprobe so weit wie möglich ausgeschöpft werden, da gerade schwerer erreichbare Personen, mit denen Interviews erst nach mehreren vergeblichen Kontaktversuchen zustande kommen, für die Repräsentativität der pro Tag realisierten Interviews von zentraler Bedeutung sind. Essentiell ist daher, dass die Bearbeitungsroutine für die einzelnen Replikate einheitlich ist und genau definierten, konstanten Regeln folgt.

Für die Bundestagswahlstudie 2013 wurde beschlossen, dass jedes Replikat 14 Tage lang aktiv bleiben sollte, um eine möglichst gute Ausschöpfung zu erreichen. In den ersten beiden Tagen sollte jede nicht sofort stichprobenneutral ausgeschiedene Telefonnummer, bei der noch kein Interview realisiert (oder ein Termin vereinbart) werden konnte, vier Mal angerufen werden, an den folgenden zwölf Tagen jeweils zwei Mal.² Bezüglich der Uhrzeiten der Kontaktversuche war dabei ein komplexes Rotationsschema zu beachten, um die Kontaktversuche gleichmäßig über den Tag zu verteilen. Telefoniert wurde wochentags von 10 bis 21 Uhr, samstags und sonntags von 11 bis 20 Uhr.

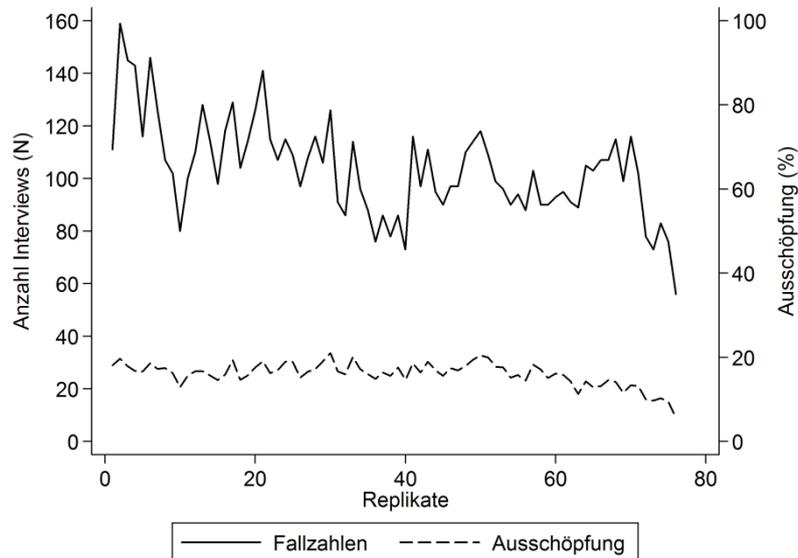
Besetzte Nummern – unter denen offenkundig eine Kontaktperson aktuell erreichbar ist – wurden bereits nach 30 Minuten erneut angewählt, um die gegebene Kontaktmöglichkeit zu nutzen. Zudem wurden Terminvereinbarungen prioritär behandelt und durchbrachen den skizzierten Kontakt-Algorithmus. Dies gilt für die Kontaktfrequenz pro Tag ebenso wie für den Zeitraum, für den eine Nummer aktiv bleibt, denn Terminvereinbarungen wurden auch über den definierten zweiwöchigen Zeitraum hinaus wahrgenommen. Schließlich wurden – als zusätzliche Maßnahme mit dem Ziel optimaler Ausschöpfung – gezielte Konversionsversuche „weicher“ Verweigerer (Personen, die z.B. mit der Begründung, keine Zeit oder kein Interesse zu haben, ein Interview ablehnten, jedoch keine weiteren Kontaktversuche untersagten) unternommen: Sie sollten zwei Tage nach ihrer ursprünglichen Ablehnung noch einmal kontaktiert werden.

Abbildung 2 zeigt das Resultat der Befragung. Die absolute Zahl der Interviews, die aus jedem der 76 täglichen Replikate hervorgegangen sind, schwankte zwischen 56 aus dem am letzten Samstag freigegebenen Replikat und 159 aus dem 2. Replikat. Beide Extreme sind wenig überraschend: Dass gerade das 2. Replikat zu besonders vielen Interviews führte, ist vor allem darauf zurückzuführen, dass es die hohe Zahl von 38 Scheiben umfasste. Generell gilt natürlich: Je höher der Bruttoansatz, desto mehr Interviews, und das kommt durchweg in Abbildung 2 zum Ausdruck. Dass demgegenüber das letzte Replikat – wie auch generell die Replikate aus den letzten Tagen vor der Wahl – vergleichsweise geringe Interviewzahlen lieferte, ist ebenfalls designbedingt, denn diese Replikate konnten nicht für dieselbe Zeitdauer offen bleiben wie diejenigen, die so rechtzeitig ins Feld gegangen waren, dass sie noch für die volle Periode von 14 Tagen abgearbeitet werden konnten. Da die Vorwählerhebung naturgemäß mit dem Vorabend der Wahl ihr Ende finden musste, „trunkierte“ der Wahltag die Bearbeitung dieser Replikate. Für die Analyse-

² Das bedeutet eine deutliche Reduktion der Kontaktfrequenz gegenüber früheren deutschen RCS-Studien. Diese wurde notwendig wegen geänderter rechtlicher Rahmenbedingungen für die Umfrageindustrie.

möglichkeiten der Daten ist dieses für RCS-Studien designbedingt normale Phänomen unbedenklich, denn diese orientieren sich ja nicht an den Replikaten, sondern an den Erhebungstagen.

Abbildung 2: Realisierte Fallzahlen und Ausschöpfung aus Replikaten



Das standardisierte Bild der Nettoausschöpfungsquoten der Teilstichproben, das Abbildung 2 ebenfalls zeigt, ist demgegenüber – wie es sein soll – sehr viel gleichmäßiger: Die Effekte unterschiedlicher Scheibenzahlen sind hier verschwunden; dagegen bleibt – da designbedingt – der Rückgang der Ausschöpfung zum Ende der Feldzeit hin auch in dieser Darstellungsweise sichtbar. Die Gleichmäßigkeit bei Ausschöpfung der Replikate zeigt an, dass der Kontakalgorithmus in der Feldpraxis kontinuierlich abgearbeitet wurde. Berichtstabellen über die Einhaltung des Kontaktschemas waren Bestandteil der täglichen Feldberichte und in der Tat erfolgten 98,9% der Kontaktversuche im vorgesehenen Zeitfenster (bei 15 minütiger Fehlertoleranz). In der Feldpraxis wurde das angestrebte ideale Kontaktschema also beinahe in Perfektion umgesetzt. Die gleichmäßige Abarbeitung des Kontaktschemas könnte auch ein Grund dafür sein, wieso 2013 im Vergleich zu 2009 weniger Kontaktversuche ausreichten, um ein Interview zu erzielen. Nach bis zu 14 Versuchen lagen 2013 bereits 90 Prozent der Interviews vor; 2009 waren es da erst 80 Prozent (Abbildung 3).

Analog zu Abbildung 3 stellt Abbildung 4 für die RCS-Studien 2005, 2009 und 2013 gegenüber, wie viele Tage nach Replikatsstart die vorhandenen Interviews durchgeführt wurden. Dabei wurden für 2013 nur jene Nummern berücksichtigt, die bis zum 63. Erhebungstag (2005: bis 21. Erhebungstag; 2009 bis 47. Erhebungstag) aktiviert wurden. Nummern aus Replikaten, die danach ins Feld gingen, konnten wegen des herannahenden Wahltermins nicht mehr volle 14 (2009, 2013) bzw. 21 Tage (2005) lang ausgeschöpft werden. Der Graph, der die Anteile realisierter Interviews in Abhängigkeit der vergangenen Tage nach Replikatsfreigabe in der RCS-Studie 2013 angibt, verläuft in der Abbildung zunächst oberhalb der entsprechenden Graphen für 2005 und 2009. Rund 48 Prozent der Interviews wurden 2013 sofort am Tag der Freigabe ihres Replikats realisiert (2005: 34 Prozent, 2009: 37 Prozent). Die 50-Prozent-Marke wurde 2013 bereits am zweiten Tag erreicht (2005, 2009: dritter Tag). Eine Woche nach der Aktivierung des zugehörigen Replikats waren 2013 82 Prozent, 2009 75 Prozent der Interviews realisiert; 2005 konnten bis dahin 78 Prozent durchgeführt werden.

Abbildung 3: Anzahl der Kontaktversuche bis zum Interview

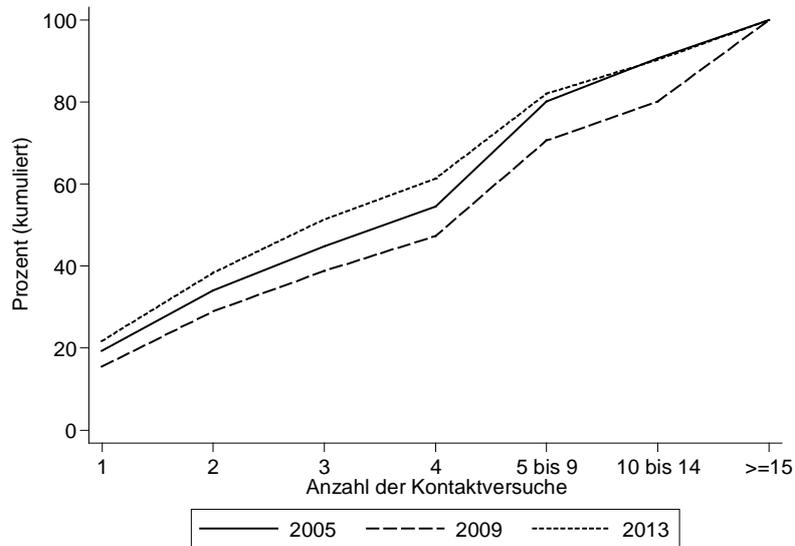
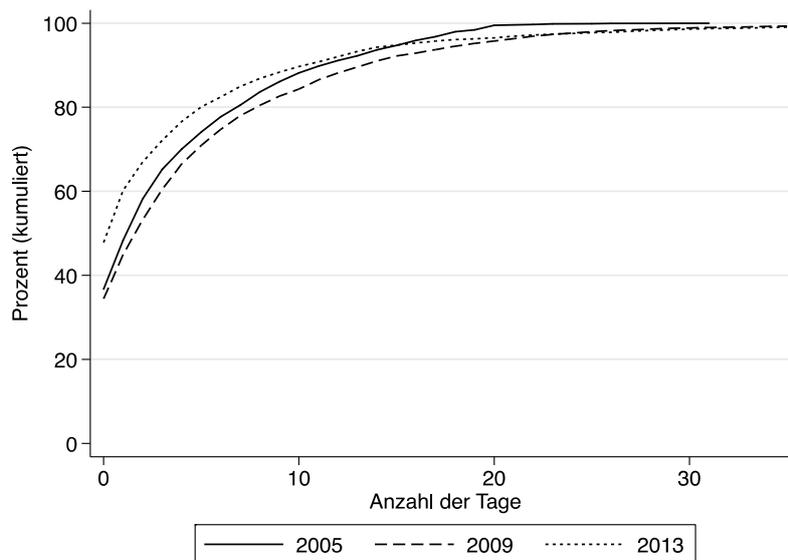


Abbildung 4: Tage zwischen realisiertem Interview und Replikatfreigabe



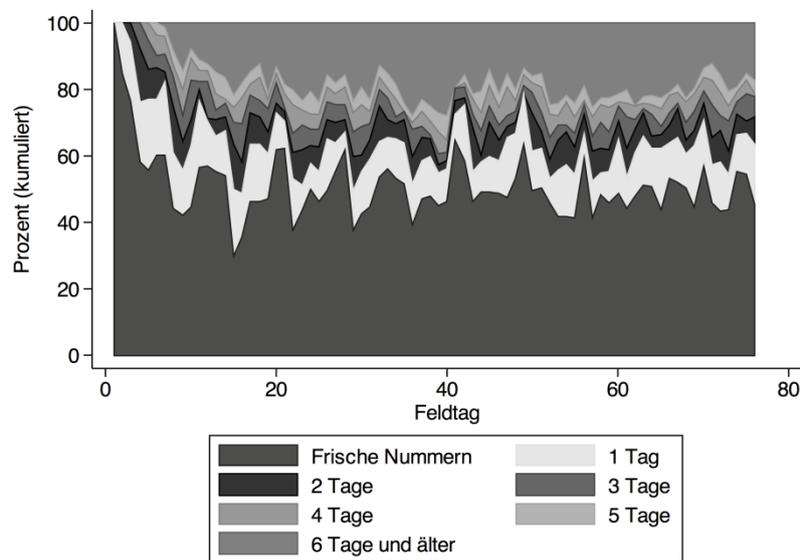
Die Versuche, „weiche Verweigerer“ durch nochmaliges Kontaktieren zur Interviewteilnahme zu bewegen, erwiesen sich 2013 hingegen als weniger erfolgreich: So konnten insgesamt nur 61 Interviews auf diesem Weg gewonnen werden. Dies entspricht einer Konversionsrate von einem Prozent (2009: 4%, 2005: 14,1%). Angesichts dieser äußerst geringen Konversionsrate der RCS-Studie 2013 gegenüber früheren Wahlen, die auf geänderte Feldbedingungen zurückzuführen sind, ist die Sinnhaftigkeit solcher Konversionsversuche bei künftigen Studien in Frage gestellt.

4. Zur Qualität der RCS-Daten

4.1 Strukturgleichheit der Tagesstichproben

Es ist bei graphischen Analysen von RCS-Daten üblich, die Befragten der ersten Erhebungstage auszuschließen, da designbedingt erst eine gewisse Zeit vergehen muss, ehe sich die gewünschte Mischung aus leicht und schwer erreichbaren Personen einstellt. Diesen Sachverhalt illustriert Abbildung 5, in der der tageweise Anteil von Replikaten an den Tagesstichproben nach Alter der Replikate dargestellt wird: Am ersten Feldtag stehen zwangsläufig nur die an diesem Tag aktivierten Nummern zur Verfügung (vgl. Abschnitt 2 zum Design), wodurch schwer erreichbare Personen (Vollzeitberufstätige) in der ersten Tagesstichprobe in der Regel unterrepräsentiert sind und leicht zugängliche Segmente (Rentner, Arbeitslose) übermäßig vorhanden sind. Die zweite Tagesstichprobe setzt sich zu 85 Prozent aus an diesem Tag aktivierten Nummern und zu 15 Prozent aus Nummern des Vortags zusammen. Hinsichtlich des Replikatsalters sind beide Stichproben somit zu 85% strukturgleich. Bis einschließlich des dritten Feldtages dominieren die besonders häufig kontaktierten Nummern des jeweils aktuellen Feldtags sowie des Vortags die Struktur der täglich resultierenden Stichproben – und damit jene Zielpersonen, die mit relativ wenigen Kontaktversuchen erreicht werden konnten. Erst ab dem vierten Feldtag wird die angestrebte Mischung aus leicht und schwer erreichbaren Personen annähernd erreicht. Da die Erreichbarkeit bei telefonischen Befragungen mit Merkmalen verknüpft ist, die auch inhaltlich von Bedeutung sind, beginnen die folgenden Analysen mit dem vierten Feldtag unserer Erhebung.

Abbildung 5: Anteil von Replikaten an den täglichen Interviews nach Alter der Replikate



Das tägliche Aktivieren neuer Telefonnummern und das strikte Einhalten des vorgegebenen Kontaktschemas dienen dem Ziel, die Interviews in den verschiedenen Tagesstichproben mit möglichst homogenen Verteilungen des Replikatsalters zu erheben. Unter Replikatsalter definieren wir die Differenz zwischen Aktivierungs- und Befragungszeitpunkt in Tagen. Wird eine Rufnummer in das Kontaktschema aufgenommen und am gleichen Erhebungstag zum Interview, erhält der Fall das Replikatsalter Null. Ein erfolgreiches Interview am Folgetag würde das Alter 1 erzeugen.

Durch Umsetzung des RCS-Designs wird erreicht, dass das Verhältnis leicht erreichbarer Segmente und schwer zugänglicher Personen in den erhobenen Tagesstichproben nur zufallsbedingt variiert, wodurch sich Veränderungen der politischen Meinungsbildung nicht auf die Erreichbarkeit zurückführen lassen.

Theoretisch müsste die Verteilung des Replikalters für eine perfekte RCS-Struktur an jedem Tag identisch sein. Es sollte also täglich der gleiche Anteil an Interviews mit dem Replikalter i erzielt werden. Tabelle 2 zeigt anhand der ab Tag 15 gemessenen Gesamtverteilung des Replikalters in der RCS-Vorwahlstudie die für jeden Tag anzustrebenden Fälle nach dem Zeitpunkt der Aktivierung.

Tabelle 2: Soll-Verteilung der Interviews nach Replikalter und Tag

Replikatsalter i in Tagen	Gesamt %	Tag 15		...		Tag 20		Tag 21		...		Tag 76	
		N	%	N	%	N	%	N	%	N	%	N	%
0	48%	48	48%	...		48	48%	48	48%	...		48	48%
1	12%	12	12%	...		12	12%	12	12%	...		12	12%
2	7%	7	7%	...		7	7%	7	7%	...		7	7%
3	5%	5	5%	...		5	5%	5	5%	...		5	5%
4	5%	5	5%	...		5	5%	5	5%	...		5	5%
...
14+	7%	7	7%	...		7	7%	7	7%	...		7	7%
Summe	100%	100	100%	...		100	100%	100	100%	...		100	100%

Insgesamt etwa 48% dieser Interviews wurden noch am Aktivierungstag erhoben und weitere 12% am Folgetag. Durchschnittlich 7% der Fälle konnten erst nach 14 Tagen oder später erfolgreich befragt werden. Damit sind diese Tagesstichproben hinsichtlich des Replikalters nur zu 48% strukturgleich mit dem ersten Tag, der eine äußerst geringe Tageseffizienz aufweisen würde. Abweichungen in den Antworten der ersten und einer späteren Stichprobe könnten durch die unterschiedliche Erreichbarkeit bedingt sein.

Entscheidender für die Strukturgleichheit als die absolute Verteilung der Interviews bezüglich des Replikalters ist deren relatives Verhältnis. So hätten für eine vollständige Strukturgleichheit an Tag 20 auch durchaus 200 Fälle erhoben werden können, wenn gleichzeitig 96 (48%) davon aus frischen, 24 (12%) aus am Vortag freigegebenen Nummern,... und 14 (7%) mit Rufnummern des Alters 14 oder höher generiert worden wären. Tabelle 3 zeigt nun einen Auszug der tatsächlichen Verteilung des Replikalters.

Tabelle 3: Ist-Verteilung der Interviews nach Replikalter und Tag

Replikatsalter i in Tagen	Gesamt %	Tag 15		...		Tag 20		Tag 21		...		Tag 76	
		N	%	N	%	N	%	N	%	N	%	N	%
0	48%	33	30%	...		62	62%	61	62%	...		56	49%
1	12%	22	20%	...		11	11%	8	8%	...		22	19%
2	7%	15	14%	...		3	3%	3	3%	...		10	9%
3	5%	7	6%	...		6	6%	2	2%	...		8	7%
4	5%	5	5%	...		3	3%	4	4%	...		1	1%
...
14+	7%	1	1%	...		0	0%	3	3%	...		10	9%
Summe	100%	110	100%	...		100	100%	98	100%	...		123	100%

Am Wochenende der Tage 20 und 21 wurden insgesamt 100 bzw. 98 Fälle erhoben. Dies entspricht verglichen mit der Gesamtstichprobe zu vielen Interviews mit frischen Nummern (jeweils 62% anstatt 48%) und zu wenig Fällen mit höherem Replikatsalter (8-11% anstatt 12%, 3% anstatt 7%,..., 0-3% anstatt 7%). Der Tag 15 zeigt eine ganz andere Struktur. So wurden relativ viele (110) Interviews generiert, jedoch zu wenige aus frischen Rufnummern (30% anstatt 48%) sowie außerordentlich viele mit am Wochenende davor aktivierten Nummern ($x_{1,15} = 0.2$ anstatt $\bar{x}_1 = 0.12$ bzw. $x_{2,15} = 0.14$ anstatt $\bar{x}_2 = 0.07$).

Die Höhe dieser Strukturgleichheit lässt sich für einen beliebigen Tag j durch Differenzberechnungen zum arithmetischen Mittel zudem exakt beziffern:

$$(1) \quad I_j^* = 1 - \sum_{i=0}^m |x_{ij} - \bar{x}_i| \cdot 2^{-1} \in [0,1]$$

Viel aussagekräftiger als der Abstand zur Gesamtstruktur sind jedoch die Differenzen zu den Tagesstichproben. Diese können mit der RCS-Effizienz für eine Untersuchung von n Feldtagen exakt quantifiziert werden (Hoops/Glantz/Michael 2013):

$$(2) \quad I_j = 1 - \sum_{k>m} \sum_{i=0}^m |x_{ij} - x_{ik}| \cdot (n-m-1)^{-1} \cdot 2^{-1} \in [0,1]$$

Wenden wir diese Maßzahl für $j=20$ an, werden zunächst sukzessiv die Abstände der Verteilung des Replikatsalters zu den Tagen 15 bis 76 berechnet. Damit ergibt sich eine RCS-Tageseffizienz von 79 Prozent:

$$(3) \quad I_{20} = 1 - |0.62 - 0.30| + |0.11 - 0.20| + \dots + |0 - 0.09| \cdot 122^{-1} = 0,79$$

Das heißt Tag 20 ist bezüglich des Replikalters zu 79% strukturgleich mit den anderen Tagesstichproben und eignet sich am besten für einen paarweisen Vergleich mit einer anderen zufällig ausgewählten Stichprobe der Studie. Die Gesamt-RCS-Effizienz kann nun analog dazu berechnet werden:

$$(4) \quad I_{ges} = 1 - \frac{\sum_{k>j \geq m+1} \sum_{i=0}^m |x_{ij} - x_{ik}|}{(n-m) \cdot (n-m-1)} \in [0,1]$$

Die Vorwahlstudie zeigt mit dieser Methode eine Gesamt-RCS-Effizienz in Höhe von 82%. Der Tag 15 hat dabei die niedrigste Strukturgleichheit von 71%, der Tag 26 hat mit 86% die höchste Effizienz.

In der Verteilung des Replikatsalters zeigt sich von Mittwoch bis Freitag die höchste Strukturgleichheit mit einer mittleren Effizienz von 82,4 bis 83,3% (siehe Tabelle 4). Verantwortlich für den Großteil an Heterogenität ist tatsächlich der Übergang des Wochenendes auf den Wochenbeginn. So haben die Sonntage mit durchschnittlich 79,6% die geringste wochentagsspezifische RCS-Effizienz. Ursache hierfür liegt in der Vielzahl von Interviews mit frischen Nummern (58% anstatt 48%) und der geringen Anzahl an Interviewrealisierungen mit Rufnummern höheren Alters. Vermutlich erzeugt durch den niedrigen Anteil an Terminvereinbarungen, die auf die Sonntage fallen, sowieso die höheren Replikatsgrößen. So haben am Sonntag 70% der Befragten ein Replikatsalter von höchstens einem Tag, am Montag jedoch nur 55%.

Tabelle 4: Ist-Verteilung der Interviews mit n>14 nach Replikatsalter und Wochentag

Replikatsalter i in Tagen	Gesamt %	Montag %	Dienstag %	Mittwoch %	Donnerstag %	Freitag %	Samstag %	Sonntag %
0	48%	42%	45%	47%	49%	49%	51%	58%
1	12%	13%	12%	11%	12%	13%	13%	12%
2	7%	10%	7%	7%	7%	6%	6%	4%
3	5%	7%	7%	5%	4%	6%	5%	3%
4	5%	5%	6%	5%	4%	5%	3%	3%
...
14+	7%	7%	7%	7%	8%	7%	7%	5%
N	6591	1128	1048	962	948	867	892	746
Effizienz	81,7%	80,6%	81,2%	83,3%	82,4%	82,5%	81,7%	79,6%

Zum Teil konträre Gründe erklären die niedrigere Effizienz an Montagen. So finden am Wochenstart immens viele Terminvereinbarungen statt (siehe Tabelle 5), die zu einer großen Anzahl an Interviews mit einem hohen Replikatsalter führen. Zudem werden übermäßig viele Interviews mit Nummern generiert, die am Freitag bis Sonntag davor aktiviert wurden.

Tabelle 5: Wochentagsspezifische Tagesausschöpfungen

Wochentag	Durchschnittliche Anzahl von Interviews	Anteil Interviews aus Terminvereinbarung	Anteil Interviews aus direkter & fester Terminvereinbarung
Montag	119,1	51,5%	11,5%
Dienstag	110,7	52,5%	10,9%
Mittwoch	102,2	51,5%	11,0%
Donnerstag	106,3	47,6%	10,4%
Freitag	94,3	47,3%	10,9%
Samstag	100,7	40,2%	6,2%
Sonntag	91,7	30,5%	3,8%
Gesamt	103,7	46,5%	9,5%

In der Studie 2009 schwankten die RCS-Tageseffizienzen zwischen 42% und 64%. Die Gesamt-RCS-Effizienz lag bei 58,6%. Damit besitzen die Daten der aktuellen Studie eine wesentlich höhere Strukturgleichheit.

Dennoch können einige Punkte weiterhin optimiert werden. So wäre es denkbar, am Montag mehr frische Nummern zu aktivieren und am Sonntag entsprechend weniger. Letzteres hätte jedoch negative Auswirkungen auf die Tagesfallzahl, die durch Reduktion der Erhebungstage kompensiert werden könnte. Höhere Tagesfallzahlen würden sich auch positiv auf die Strukturgleichheit auswirken, denn mit steigendem Stichprobenumfang verlieren Zufallsschwankungen an Einfluss.

Unsere Analysen zeigen zudem Schwankungen im Anteil der generierten Nummern pro Replikat. Zwar sind diese äußerst minimal, die Extremwerte liegen zwischen 82% und 85,5%. Dennoch würden konstante Werte die Strukturgleichheit weiter fördern. Außerdem wäre es überlegungswert, Termine nur systematisch nach festen Regeln zu vergeben und gleichmäßig auf die Wochentage zu verteilen. Dieser Schritt könnte jedoch die Ausschöpfung negativ beeinflussen.

4.2 Soziodemographische Strukturgleichheit

Das Kernanliegen einer RCS-Studie besteht wie skizziert darin, eine Stichprobe zu realisieren, für deren Mitglieder nicht nur die Inklusion selbst, sondern auch der Tag, an dem sie interviewt werden, per Zufall bestimmt wird. Wenn dies gelingt, dürfen die Verteilungen demographischer Variablen im Feldverlauf nicht variieren. Die Tagesstichproben sollen ja strukturgleich sein. Systematische Trends dürfen nicht erkennbar werden, lediglich zufallsbedingte Fluktuationen sind zulässig. Inwieweit dies in der praktischen Umsetzung der RCS-Studie zur Bundestagswahl 2013 gelungen ist, wird im Folgenden anhand von sechs soziodemographischen Merkmalen aufgezeigt. Hierfür empfehlen sich zunächst Trendplots der interessierenden Merkmalsanteile bzw. Mittelwerte. Soll eine solche Analyse auf der Ebene einzelner Stichprobentage durchgeführt werden, entsteht das Problem, dass die Tagesstichproben jeweils relativ klein und daher zwangsläufig mit einem recht großen Zufallsfehler behaftet sind. Dem kann durch Verfahren der Datenglättung entgegengewirkt werden. Dazu greifen wir hier auf das LOWESS-Glättungsverfahren (Bandbreite 0,5) zurück (vgl. Cleveland 1994).

Abbildung 6: Verteilungsstabilität demographischer Merkmale

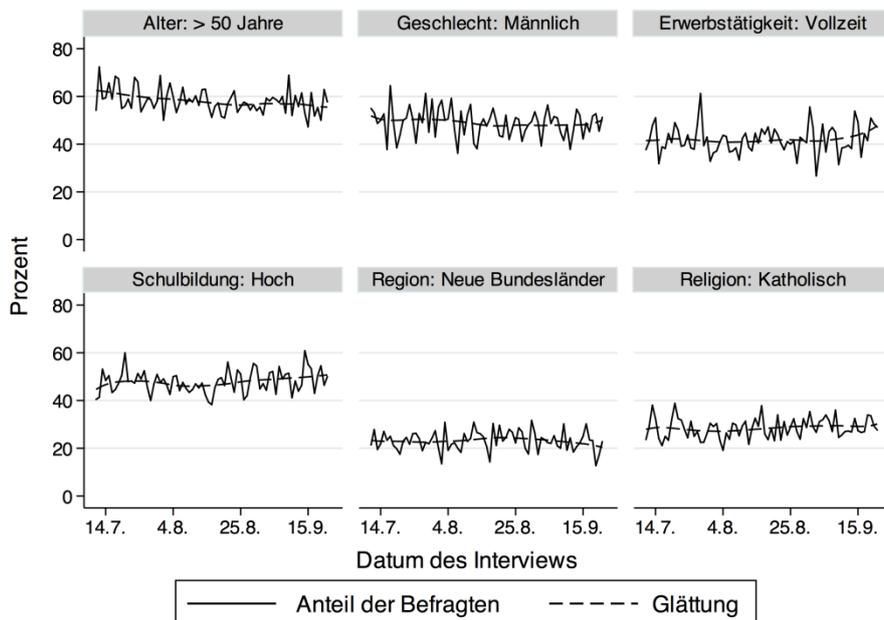


Abbildung 6 zeigt, dass die Verteilungen des Lebensalters (Anteil der Befragten unter 60 Jahre), des Geschlechts, des der Schulbildung (Hochschul- bzw. Fachhochschulreife), des Erwerbsstatus (Vollzeit erwerbstätig) der regionalen Herkunft der Befragten (Wohnort in den neuen Bundesländern) sowie der Konfessionszugehörigkeit (römisch-katholisch) im wesentlichen über den gesamten Erhebungszeitraum hinweg gleich geblieben sind. Allenfalls geringfügige Abweichungen lassen sich erkennen.

Tabelle 6: Demographische Merkmale: Tagespaarvergleiche (Anteile signifikanter Abweichungen im Tagespaarvergleich ($p < .05$); Angaben in Prozent)

	alle Wochentage	ohne Wochenenden
Geschlecht	6,8	9,5
Alter	8,3	8,4
Schulbildung: Hoch	1,2	0,4
Erwerbsstatus: Vollzeit	5,1	2,0
Region:	2,6	3,6
Neue Bundesländer		
Religion: Katholisch	1,6	0,2
Anzahl Paare	2628	1326

In Ergänzung zur graphischen Analyse haben wir diese Ergebnisse auch inferenzstatistisch abgesichert. Für *alle möglichen Paare* von Tagesstichproben (mit Ausnahme der ersten drei Feldtage) haben wir den Anteil der oben genannten Merkmale mittels logistischer bzw. für das mittlere Alter mittels linearer Regressions in den beiden jeweils betrachteten Tagesstichproben verglichen. In die Modelle haben wir jeweils neben der Konstanten nur eine einzige Indikatorvariable einfließen lassen, die die Zugehörigkeit zu *einer* der beiden jeweils verglichenen Tagesstichproben anzeigt. In Tabelle 6 sind die Anteile signifikanter Abweichungen zwischen jeweils zwei Feldtagen an sämtlichen Zweier-Kombinationen von Tagesstichproben dargestellt. Eine signifikante Abweichung liegt dann vor, wenn der Koeffizient der Indikatorvariablen, die zwischen den beiden Feldtagen trennt, außerhalb des 95-Prozent-Konfidenzintervalls liegt. Rein zufallsbedingt würde man in bis zu fünf Prozent aller Kombinationen von Feldtagen Abweichungen erwarten.

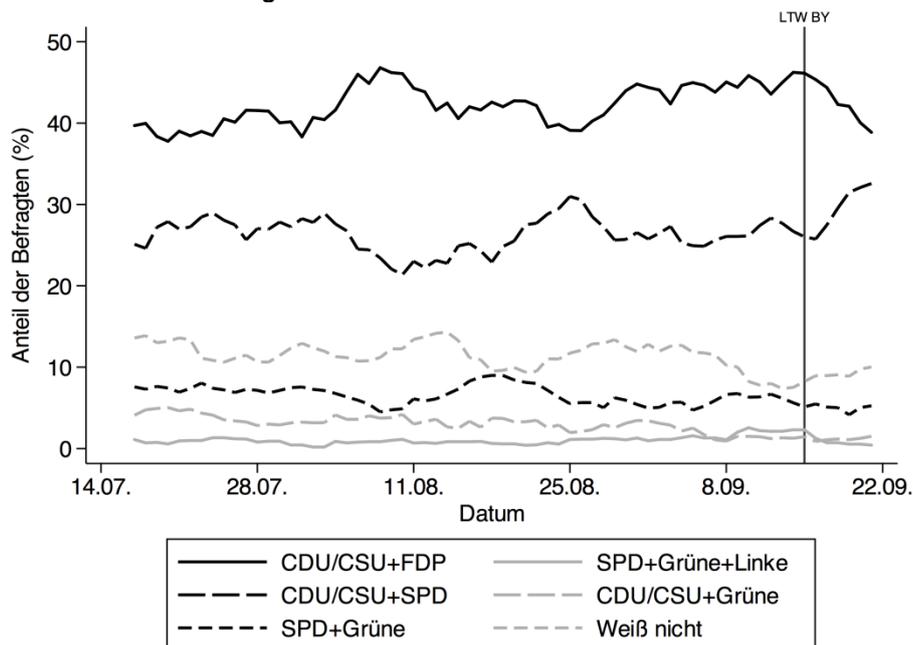
Die Tabelle zeigt lediglich für drei Merkmale geringfügige Überschreitungen dieser Marge: das Geschlecht, das Alter und der Erwerbsstatus. Schließt man die Wochenenden aus, an denen besondere Bedingungen herrschen, weil bestimmte Bevölkerungsgruppen, insbesondere Berufstätige, leichter zu erreichen sind als unter der Woche, während aber gleichzeitig die Befragungsbereitschaft insgesamt sinkt, so verbessert sich der Wert für den Erwerbsstatus deutlich. Allerdings erhöhen sich geringfügig die Abweichungen beim Alter und Geschlecht der Befragten. Diese Ergebnisse sind insgesamt zufriedenstellend; das Gros der Abweichungen beruht demnach auf zufälligen Schwankungen der sozialstrukturellen Merkmale.

5. Analysepotenzial

Der nachfolgende Abschnitt diskutiert unter deskriptiver Perspektive einige exemplarische Befunde zum Bundestagswahlkampf 2013, um das Analysepotenzial der RCS-Daten zu veranschaulichen. Wie stabil bzw. volatil waren die politischen Orientierungen der Bürger während des Wahlkampfs? Frühere RCS-Studien zu Bundestagswahlen haben ergeben, dass speziell die Erwartungen über die nächste Regierungskoalition im Laufe des Wahlkampfs einer erheblichen Dynamik unterworfen sind (Faas/Schmitt-Beck 2007, Krewel/Schmitt-Beck/Wolsing 2011). Kann dies auch für den Bundestagswahlkampf 2013 gelten? Abbildung 7 stellt dazu die Entwicklung der Koalitionserwartungen im Zeitverlauf dar:³ Ähnlich wie bei den zwei vorangegangenen Wahlen rechneten die Bürger zunächst überwiegend mit einem schwarz-gelben Regierungsbündnis; alternativ dazu wurde auch eine Große Koalition für möglich gehalten. In der Vorwahlwoche näherten sich die Anteile für die beiden Koalitionsoptionen – mutmaßlich infolge der Landtagswahl in Bayern am Sonntag vor der Bundestagswahl – jedoch stark an. Zugleich nahm auch der Anteil der „Weiß nicht“-Antworten, der im Laufe des Wahlkampfs erst zurückgegangen war, wieder leicht zu. Der Wahlausgang wurde also in der letzten Woche vor der Wahl zunehmend wieder als offen angesehen.

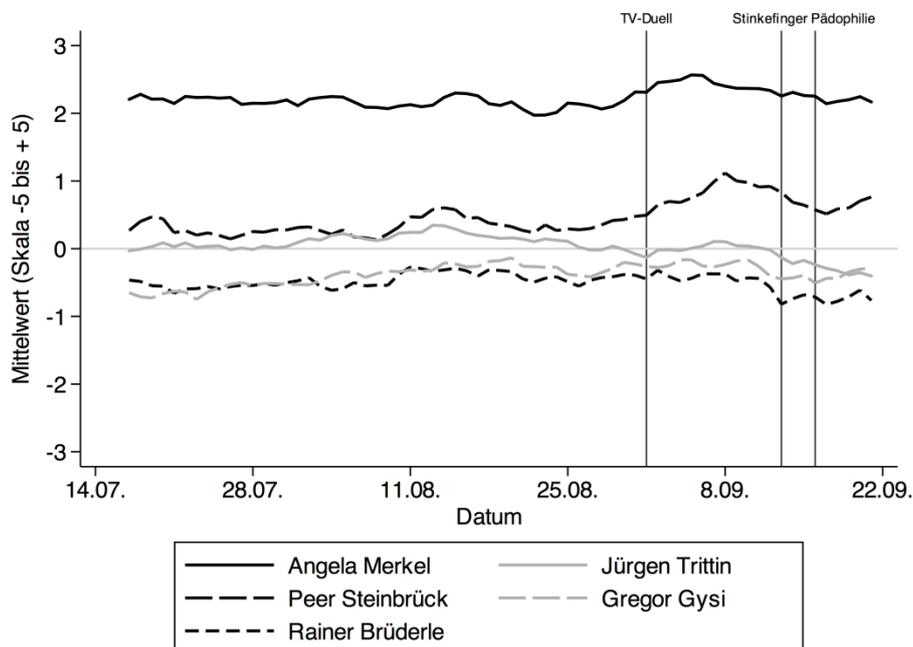
3 Dargestellt sind hier und im folgenden jeweils die nach Schulbildung und Auswahlwahrscheinlichkeit gewichteten Durchschnittswerte der jeweils zurückliegenden sieben Tage.

Abbildung 7: Koalitionserwartungen



Die Schwankungen der Koalitionserwartungen im Zeitverlauf dürften in hohem Maße darauf zurückzuführen sein, für wie wahrscheinlich die Wähler einen Einzug der FDP in den Bundestag hielten. Bei der bayerischen Landtagswahl hatte die FDP die Fünf-Prozent-Hürde nicht erreicht, wodurch aus Sicht der Wähler offenbar auch der Einzug in den Bundestag in Frage gestellt war. Außerdem dürften wie bei früheren Wahlen veröffentlichte Meinungsumfragen eine bedeutsame Rolle für die Dynamik der Erwartungsbildung gespielt haben (Faas/Schmitt-Beck 2007, Faas/Mackenrodt/Schmitt-Beck 2008, Hoffmann/Klein 2013).

Jenseits der hoch volatilen Regierungsprospektion haben sich in der Vorgängerstudie zur Bundestagswahl 2009 auch interessante parteispezifische Dynamiken in Hinblick auf Kandidaten- und Sachfragenkompetenz-Bewertungen erkennen lassen (Krewel/Schmitt-Beck/Wolsing 2011). Vor allem der SPD-Kanzlerkandidat Steinmeier konnte im Bundestagswahlkampf 2009 an Sympathie hinzugewinnen und gegenüber Kanzlerin Merkel aufholen; gleichzeitig wurde seine Partei in der Tendenz zunehmend als kompetent beurteilt. Wie entwickelten sich Kandidatenbewertung und Sachfragenkompetenz im 2013er Wahlkampf?

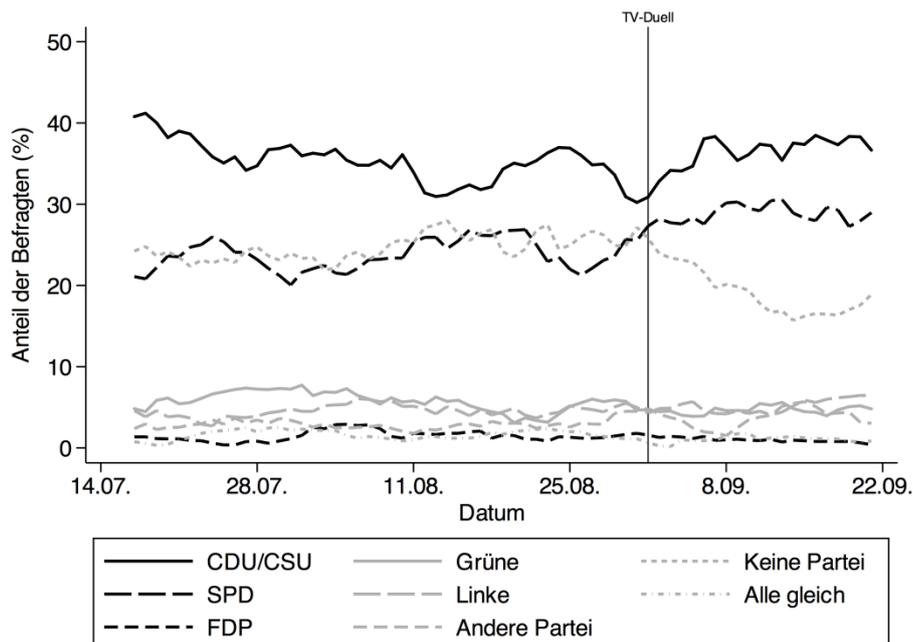
Abbildung 8: Bewertungen der Spitzenpolitiker

In Abbildung 8 sind die Bewertungen des politischen Spitzenpersonals aufgetragen. Auffällig sind die bereits 2009 beobachteten sehr hohen Popularitätswerte für Bundeskanzlerin Merkel (CDU/CSU), die in der Zeit um das TV-Duell am 1. September zwischenzeitlich moderat noch weiter anstiegen. Die Beurteilung des SPD-Kanzlerkandidaten Steinbrück fällt demgegenüber deutlich weniger positiv aus und weist vergleichsweise stärkere Schwankungen im Zeitverlauf auf: Steinbrück konnte vor allem in der Woche nach dem TV-Duell an Sympathie hinzugewinnen und gegenüber Merkel aufholen. Jedoch hielt dieser positive Effekt nicht lange an; stattdessen gingen die Beliebtheitswerte eine Woche später zunächst leicht, dann, vermutlich als Folge der "Stinkefinger"-Debatte (SZ vom 12.09.2013), noch etwas deutlicher zurück. In den letzten Vorwahltagen stiegen seine Bewertungen zwar wieder an, erreichten jedoch nicht erneut wieder das zwischenzeitlich erzielte Höchstniveau. Jürgen Trittin, einer der Spitzenkandidaten der Grünen, der mit Steinbrück zu Beginn des Wahlkampfes noch in etwa gleichauf gelegen hatte, büßte insbesondere kurz vor der Wahl erkennbar an Popularität ein. Hierbei könnte die Debatte um fragwürdige Äußerungen des Kandidaten zur Pädophilie in den frühen 1980er Jahren eine Rolle gespielt haben, die wenige Tage vor der Wahl durch Presseberichte ausgelöst wurde (Zeit vom 17.09.2013). Unmittelbar vor der Wahl wurde der Spitzenkandidat der Grünen letztlich in der Gesamttendenz sichtlich negativ beurteilt. Für die Spitzenkandidaten der Linken und der FDP, Gregor Gysi und Rainer Brüderle, zeigen sich nur geringfügige Veränderungen: Beide Kandidaten wurden während des gesamten Beobachtungszeitraums der zentralen Tendenz nach negativ bewertet. Zwar zeichnete sich zeitweise ein schwach positiver Trend ab, circa eine Woche vor der Wahl gingen die Sympathiewerte aber auch wieder etwas zurück. Insgesamt deuten die aufgezeigten zeitlichen Koinzidenzen darauf hin, dass die Entwicklungen der Kandidatenbewertungen in engem Zusammenhang mit medialen Ereignissen und Debatten standen.

Abbildung 9 zeigt die Entwicklung der Sachfragenkompetenz im Bundestagswahlkampf 2013. Wie bei der Bundestagswahl 2009 wurden vorwiegend die großen Parteien CDU/CSU und SPD als kompetent zur Lösung des wichtigsten politischen Problems erachtet; besonders gering waren im Vergleich zur vorherigen Wahl die Kompetenzwerte für die FDP. Die CDU/CSU verfügte gegenüber der SPD während des gesamten Beobachtungszeitraums über einen deutlichen Kompetenzvorsprung, jedoch konnte die SPD – trotz einiger Schwankungen – den Abstand zur CDU/CSU im Zeitverlauf insgesamt verringern: Lagen die

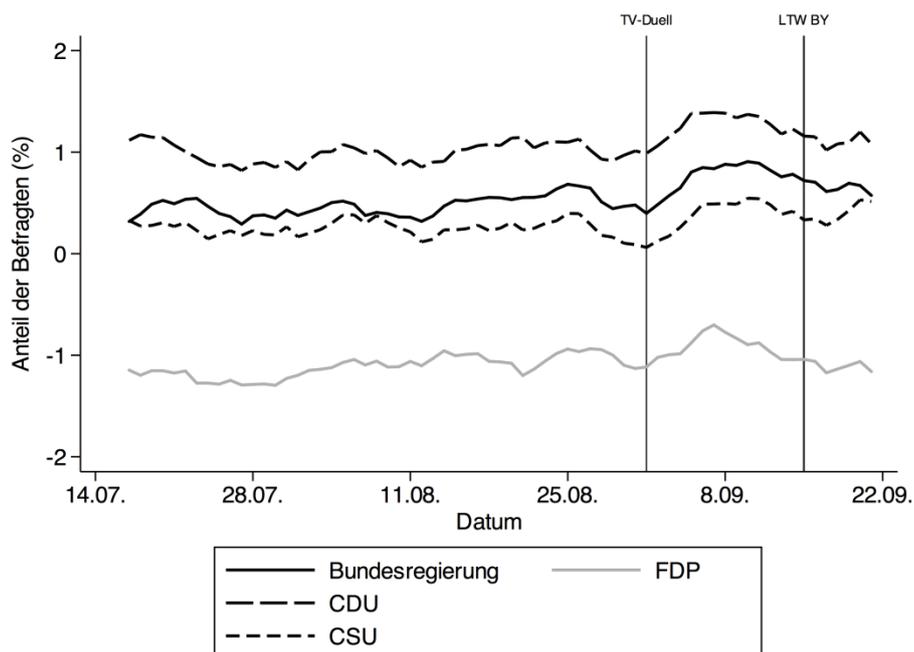
beiden großen Parteien Mitte Juli noch rund 20 Prozentpunkte auseinander, schrumpfte der Vorsprung der CDU/CSU bis zum Tag des TV-Duells auf circa 5 Prozentpunkte zusammen. Danach stiegen die Kompetenzwerte für die CDU/CSU jedoch an, so dass sich der Abstand zur SPD bis zum Wahltermin wieder vergrößerte. Gleichzeitig nahm der Anteil derjenigen, die keine Partei für kompetent hielten, in den letzten drei Wochen vor der Wahl deutlich ab. Die kleinen Parteien konnten aber auch in dieser Phase keine Kompetenzgewinne für sich verzeichnen. Für einige wichtige Kurzfristfaktoren des Wählerverhaltens ergibt sich somit insgesamt ein differenziertes Bild von Stabilität und Wandel während des Wahlkampfes: Wie bereits in früheren Wahlkämpfen offenbaren sich kandidaten- und parteispezifische Dynamiken.

Abbildung 9: Bewertungen der Sachfragenkompetenz



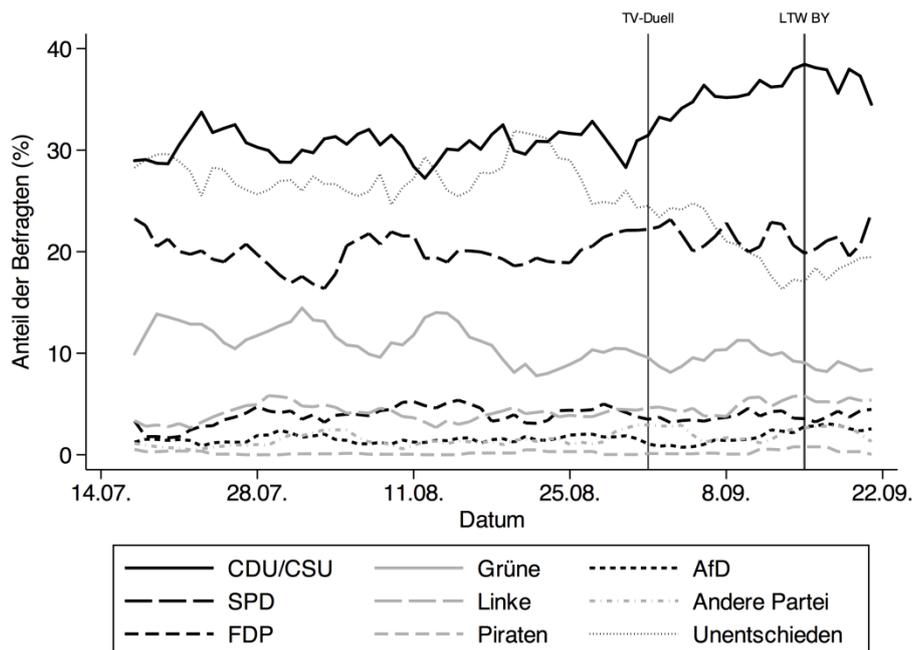
Als ein stabilisierend wirkender Gegenspieler zu den bisher betrachteten vergleichsweise volatilen politischen Orientierungen kann die retrospektive Evaluation der Regierungsleistung angesehen werden. So argumentierte Key (1966), dass Wahlkämpfen nur ein begrenztes Wirkungspotential zukommt, weil die Wähler bei ihrer Wahlentscheidung nicht primär ihre gegenwärtigen Impressionen vom Wahlkampfgeschehen einbeziehen, sondern vielmehr die Regierungsleistung der gesamten zurückliegenden Amtsperiode in Rechnung stellen. Diese Performanzbewertungen sollten nicht oder nur geringfügig vom aktuellen Wahlkampfgeschehen beeinflussbar sein. Abbildung 10 gibt Aufschluss, wie sich die Performanzbewertungen für die Bundesregierung und die einzelnen Regierungsparteien im Wahlkampf 2013 entwickelt haben. Während die Bundesregierung insgesamt überwiegend positiv bewertet wurde, ergibt sich für die einzelnen Regierungsparteien ein differenziertes Meinungsbild: Besonders positiv fiel die Leistungsbilanz der CDU aus, außerordentlich negativ jedoch jene der FDP. Die Werte für die CSU lagen dazwischen, aber noch im positiven Bereich der Skala. Sowohl für die Bundesregierung als auch für die Einzelparteien blieben die Performanzbewertungen bis zum TV-Duell im Wesentlichen stabil. In der Woche nach dem TV-Duell verbesserten sie sich jedoch zumindest vorübergehend. Obwohl die Entwicklungen für die einzelnen Parteien im Zeitverlauf überwiegend gleichförmig verliefen, kam es außerdem, wohl als Folge ihres Wahlsiegs bei der Landtagswahl in Bayern, speziell für die CSU zu einer Verbesserung der Performanzbewertungen. Damit ergibt sich in der Gesamtschau und mit Blick auf die Stabilitätshypothese ein gemischtes Bild: Obwohl die Leistungsbewertungen eine vergleichsweise hohe Stabilität erkennen lassen, scheinen sie zumindest zu einem gewissen Grad auch von Wahlkampfereignissen beeinflussbar zu sein.

Abbildung 10: Performanz der Bundesregierung und der Regierungsparteien



In Anbetracht der aufgezeigten Entwicklungen der kognitiven und affektiven Orientierungen während des Wahlkampfes stellt sich die Frage, inwiefern sich daraus Konsequenzen für die Parteipräferenzen der Bürger ergaben. Abbildung 11 stellt die Entwicklung der Wahlabsichten für die Zweitstimmen dar (Prozentuierung unter Einschluss der unentschlossenen Wähler). Hier zeigt sich, dass sich die Unterstützung für die CDU/CSU schon kurz vor dem TV-Duell deutlich zu verbessern begann. Angesichts der sehr positiven Kandidaten-, Kompetenz- und Leistungsbewertungen der beiden dominanten Regierungsparteien erscheint dies wenig verwunderlich. Die SPD-Wahlabsicht blieb bei einigen Schwankungen insgesamt auf etwa gleichbleibendem Niveau. Offenbar reichten die positiven Tendenzen bei der Kandidaten- und Kompetenzbewertung für die SPD nicht aus, um neue Wähler hinzugewinnen und den Stimmenanteil zu erhöhen. Für die Grünen zeigen sich insbesondere zu Beginn des Wahlkampfes starke Schwankungen. Spätestens ab Mitte August wird für diese Partei ein Abwärtstrend erkennbar. Für die anderen kleinen Parteien zeigte sich im Wesentlichen Stabilität. Ein „Mitleidseffekt“ für die FDP, wie ihn einige Beobachter nach dem Ausscheiden aus dem bayerischen Landtag erwartet hatten, stellte sich nicht in erkennbarer Weise ein. Allein für die Alternative für Deutschland (AfD) zeichnete sich kurz vor der Wahl noch ein Aufwärtstrend ab. Das überraschend gute Abschneiden dieser Euro-kritischen, erst ein halbes Jahr vor der Bundestagswahl gegründeten Partei, die mit 4,7 Prozent der Zweitstimmen den Einzug in den Bundestag nur knapp verfehlte, geht also primär auf Entscheidungen zurück, die Wähler erst in letzter Minute getroffen haben. Abbildung 11 zeigt darüber hinaus, dass der Anteil der Unentschlossenen in den letzten vier Wochen vor dem Wahltermin, also in der „heißen Phase“ des Wahlkampfes, deutlich zurückging. Dies deutet darauf hin, dass einem substantiellen Teil der Wählerschaft erst der offizielle Wahlkampf zu einer Entscheidung verhalf.

Abbildung 11: Entwicklung der Zweitstimmen-Wahlabsicht



Unsere deskriptiven Befunde zeigen, dass sich der Bundestagswahlkampf 2013 durch ein komplexes Gefüge von Stabilität und Dynamik auszeichnete. Der Ausgang der Wahl war für die Bürger schwer vorhersehbar: Bis unmittelbar vor der Wahl erwartete eine Mehrheit der Bürger den Fortbestand der schwarz-gelben Regierungskoalition, die jedoch letztlich nicht zustande kam. Die Einstellungen der Bürger zu den Spitzenkandidaten scheinen sich situativ im Lichte des medialen Diskurses verändert zu haben, und auch die Kompetenzbewertungen der Parteien waren dynamischem Wandel unterworfen. Diese Befunde, aber auch die aufgezeigten Entwicklungen der Wahlabsichten sprechen dafür, dass sich der Bundestagswahlkampf 2013 auf die Meinungsbildungs- und Entscheidungsprozesse der Wähler ausgewirkt hat. Allerdings war der Spielraum für Wahlkampfwirkungen auf den Wahlausgang gleichzeitig auch begrenzt: Die positiven Performanzwerte für CDU und CSU ließen einen vollständigen Regierungswechsel von Beginn an als äußerst unwahrscheinlich erscheinen. Aufgrund ihrer ausgesprochen negativen Leistungsbilanz wurde die FDP zur Hypothek für das schwarz-gelbe Regierungsbündnis und scheiterte letztlich an der Fünf-Prozent-Hürde.

6. Fazit

RCS-Umfragen sind in hohem Maße dazu geeignet, die Dynamik von Wahlkämpfen einzufangen und abzubilden. Die tagesgenaue Beobachtung erlaubt es, Effekte (und auch ihre Dauerhaftigkeit) von vorhersehbaren und unvorhersehbaren Ereignissen auf die Entwicklung der öffentlichen Meinung zu untersuchen. Die hohe zeitliche Auflösung ermöglicht es darüber hinaus, Daten aus anderen Quellen sehr genau den Umfragedaten hinzuzuspielen. Die zeitliche Reihenfolge von Ursache und Wirkung kann so präzise ermittelt werden, so dass eine kausale Interpretation der Effekte begünstigt wird. Anhaltspunkte über die zugrunde liegenden intraindividuellen Prozesse können zusätzlich mit Hilfe der begleitenden Nachwahl-Panelwelle gewonnen werden.

Die praktische Umsetzung des RCS-Designs erfordert die systematische Abarbeitung eines festen Kontaktschemas sowie ein durchgehendes und detailliertes Monitoring des täglichen Feldgeschehens. Kontinuierlich muss evaluiert werden, ob der gewählte Bruttoansatz angemessen ist, und gegebenenfalls müssen rasch und wiederholt Adjustierungen vorgenommen werden. Die Interviewbereitschaft am Wochenende, insbesondere an Sonntagen, ist geringer als an Werktagen. Damit die Anzahl der realisierten Interviews an den Wochenenden nicht zu stark absinkt, ist ein leicht höherer Nummerneinsatz erforderlich, der jedoch mit Bedacht auf die Strukturgleichheit der Tagesstichproben gewählt werden muss.

Die RCS-Studie zur Bundestagswahl 2013 hat die Anforderungen dieses Designs in hoher Qualität umgesetzt. Die Feldverläufe der RCS-Wahlkampfstudie und der Nachwahl-Panelwelle zur Bundestagswahl 2013 verliefen beinahe idealtypisch: Ausgedehnt über einen Zeitraum von 76 Tagen wurden während der Vorwahlzeit tägliche Stichproben realisiert mit einem Umfang von durchschnittlich 104 Wahlberechtigten; die Nachwahl-Panelwelle konnte zügig nach der Wahl eingeholt werden. Die Dauer und Kontaktfrequenz bis zur Realisierung des Interviews stellte sich ähnlich wie bei früheren RCS-Studien dar. Der Kontaktalgorithmus wurde nahezu in Perfektion eingehalten, wozu ein enges Monitoring entscheidend beigetragen hat. Die Analysen zur Strukturgleichheit zeigen eine sehr gleichmäßige Durchmischung der einzelnen Tagesstichproben mit leicht und schwer erreichbaren Personen. Auch in Hinblick auf die soziodemographischen Merkmale lassen sich nur geringfügige Abweichungen erkennen. Die Tagesstichproben haben sich also in hohem Maße als strukturgleich erwiesen.

Exemplarische, rein deskriptiv angelegte graphische Analysen zur Entwicklung einiger ausgewählter Orientierungen im Wahlkampf lassen ein komplexes Gesamtbild von Stabilität und Dynamik erkennen. Sowohl politische Perzeptionen als auch Einstellungen scheinen durch das Wahlkampfgeschehen beeinflusst worden zu sein. In vielen Fällen zeigten sich dabei kandidaten- bzw. parteispezifische Effekte. Insbesondere in den ereignisreichen letzten vier Wochen vor der Wahl lassen sich Prozesse der Meinungsbildung und des Meinungswandels deutlich erkennen. Der von Beobachtern insgesamt als sehr, vielleicht sogar zu ruhig charakterisierte Wahlkampf scheint die Wähler in der Endphase spürbar erreicht und einiges in Bewegung gesetzt zu haben. Vieles deutet daraufhin, dass bei der Bundestagswahl 2013 vor allem die „heiße“ Phase des Wahlkampfes, insbesondere die Zeit nach dem TV-Duell der Kanzlerkandidaten, das exakt drei Wochen vor der Bundestagswahl stattfand, eine wichtige Bedeutung für die Kristallisation von Wahlpräferenzen besaß.

Literatur

Bartels, Larry M. (2006): Three Virtues of Panel Data for the Analysis of Campaign Effects, in: *Brady, Henry E. & Johnston, Richard* (Hrsg.), *Capturing Campaign Effects*. Ann Arbor: University of Michigan Press, 134-163

Bevan, Shaun/Krewel, Mona (2013): Responsive Elections: The Effect of Public Opinion on Political Campaigns, MPESA Annual Conference, Chicago, 11.4.2013 - 14.4.2013.

Brady, Henry E./Johnston, Richard/Sides, John (2006): The Study of Political Campaigns, in: *Brady, Henry E. & Johnston, Richard* (Hrsg.), *Capturing Campaign Effects*. Ann Arbor: University of Michigan Press, 1-26

Cleveland, William (1994): *The Elements of Graphing Data*. Murray Hill/NJ: At&T Bell Labs

Faas, Thorsten/Blumenberg, Johannes N. (2012): Die Vermessung der Dynamik: Eine rollierende Panelstudie im Vorfeld der baden-württembergischen Landtagswahl 2011, in: *Methoden - Daten - Analysen* (mda) 6, 157-183.

Faas, Thorsten/Mackenrodt, Christian/Schmitt-Beck, Rüdiger (2008): 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, 299-325.

Faas, Thorsten/Schmitt-Beck, Rüdiger (2007): Wahrnehmung und Wirkungen politischer Meinungsumfragen: Eine Exploration zur Bundestagswahl 2005, in: *Brettschneider, Frank, Niedermayer, Oskar & Weßels, Bernhard* (Hrsg.), *Die Bundestagswahl 2005. Analysen des Wahlkampfes und der Wahlergebnisse*. Wiesbaden: VS-Verlag, 233-267

Hoffmann, Hanna/Klein, Markus (2013): 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: *Faas, Thorsten, Arzheimer, Kai, Roßteutscher, Sigrid & Weßels, Bernhard* (Hrsg.), *Koalitionen, Kandidaten, Kommunikation*. Wiesbaden: Springer Fachmedien, 221-246

Hoops, Christian/Glantz, Alexander/Michael, Tobias (2013): Innovationen im Rolling-Cross-Section Design, in: *Zeitschrift für Politikwissenschaft* (im Erscheinen).

Johnston, Richard (2001): Capturing Campaigns in National Election Studies, in: *Katz, Elihu & Warshel, Yael* (Hrsg.), *Election Studies: What's Their Use?* Boulder: Westview, 149-172

Johnston, Richard (2008): Modeling Campaign Dynamics on the Web in the 2008 National Annenberg Election Study, in: *Journal of Elections, Public Opinion and Parties* 18, 401-412.

Johnston, Richard/Brady, Henry E. (2002): The Rolling Cross-Section Design, in: *Electoral Studies* 21, 283-295.

Johnston, Richard/Hagen, Michael G./Hall Jamieson, Kathleen (2004): *The 2000 Presidential Election and the Foundations of Party Politics*. Cambridge: Cambridge University Press

Key, Valdimer Orlando (1966): *The responsible electorate*. Cambridge, MA: Belknap

Krewel, Mona/Schmitt-Beck, Rüdiger/Wolsing, Ansgar (2011): The Campaign and its Dynamics at the 2009 German General Election, in: *German Politics* 20, 28-50.

Lazarsfeld, Paul F./Berelson, Bernard/Gaudet, Hazel (1944): *The People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign*. New York: Columbia University Press

Romer, Daniel/Kenski, Kate/Winneg, Kenneth/Adasiewicz, Christopher, et al. (Hrsg.) (2006): *Capturing Campaign Dynamics 2000 & 2004: The National Annenberg Election Survey*. Philadelphia: University of Pennsylvania Press

Schmitt-Beck, Rüdiger (2007): *New Modes of Campaigning*, in: *Dalton, Russel J. & Klingemann, Hans-Dieter* (Hrsg.), *Oxford Handbook on Political Behavior*. Oxford: Oxford University Press, 744-764

Schmitt-Beck, Rüdiger (2009): *Kampagnendynamik im Bundestagswahlkampf 2005*, in: *Gabriel, Oscar W., Weißels, Bernhard & Falter, Jürgen W.* (Hrsg.), *Wahlen und Wähler: Analysen aus Anlass der Bundestagswahl 2005*. Wiesbaden: VS-Verlag, 146-176

Schmitt-Beck, Rüdiger/Faas, Thorsten (2006): *The Campaign and its Dynamics at the 2005 German General Election*, in: *German Politics* 15, 393-419.

Schmitt-Beck, Rüdiger/Faas, Thorsten/Holst, Christian (2006): *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, 13-49.

Schmitt-Beck, Rüdiger/Faas, Thorsten/Wolsing, Ansgar (2010): *Kampagnendynamik bei der Bundestagswahl 2009: Die Rolling-Cross-Section-Studie im Rahmen der "German Longitudinal Election Study" 2009*, in: *Arbeitspapiere - Mannheimer Zentrum für Europäische Sozialforschung*.

Schmitt-Beck, Rüdiger/Partheymüller, Julia (2012): *Why Voters Decide Late: A Simultaneous Test of Old and New Hypotheses at the 2005 and 2009 German Federal Elections*, in: *German Politics* 21, 299-316.

Schmitt-Beck, Rüdiger/Rattinger, Hans/Roßteutscher, Sigrid/Weißels, Bernhard (2010): *Die deutsche Wahlforschung und die German Longitudinal Election Study (GLES)*, in: *Faulbaum, Frank & Wolf, Christof* (Hrsg.), *Gesellschaftliche Entwicklungen im Spiegel der empirischen Sozialforschung*. Wiesbaden: VS Verlag für Sozialwissenschaften, 141-171

Schmitt-Beck, Rüdiger/Tenscher, Jens (2008): *Divided We March, Divided We Fight: Trade Unions, Social Democrats, and Voters at the 2005 German General Election*, in: *Farrell, David M. & Schmitt-Beck, Rüdiger* (Hrsg.), *Non-Party Actors in Electoral Politics: The Role of Interest Groups and Independent Citizens in Contemporary Election Campaigns*. Baden-Baden: Nomos, 151-182

Schoen, Harald (2005): *Wahlkampfforschung*, in: *Falter, Jürgen W. & Schoen, Harald* (Hrsg.), *Handbuch Wahlforschung*. Wiesbaden: VS-Verlag, 503-542