

Die Vermessung der Dynamik

*Eine rollierende Panelstudie
im Vorfeld der baden-
württembergischen
Landtagswahl 2011*

Measuring Dynamics

*A Rolling Panel Study in the
run-up to the Baden-
Wuerttemberg state election
2011*

Thorsten Faas und Johannes N. Blumenberg

Zusammenfassung

Eine rollierende Panelstudie vereint die Vorzüge des Rolling Cross-Section (RCS) Designs mit denen eines Panels. Wie das RCS-Design erlaubt es eine rollierende Panelstudie, tagesgenaue Veränderungen in den Einstellungen und Verhaltensmustern der Bevölkerung zu entdecken. Zugleich bleibt die paneltypische Möglichkeit, Prädiktoren von Veränderungen auf der individuellen Ebene zu analysieren, erhalten.

Im vorliegenden Beitrag werden zunächst die jeweiligen Vor- und Nachteile der beiden klassischen Forschungsdesigns – RCS und Panel – skizziert, um daraus die Möglichkeit einer Kombination beider Ansätze abzuleiten. Anschließend wird die „Wahlstudie Baden-Württemberg 2011“ als konkreter Versuch, eine solche Studie zu realisieren, beschrieben: Dies schließt sowohl eine Diskussion der Qualität der Daten als auch von spezifischen Analysepotenzialen ein.

Abstract

The rolling panel study combines the advantages of rolling cross-section designs with those of classic panel designs. Using a rolling panel, it is possible to scrutinize aggregate changes of attitudes and behavioural intentions on a day-by-day basis, as is the case for rolling cross-section studies. But the design also allows us to analyse individual changes and their causes – typical features of panel studies.

The present article starts with a discussion of the advantages and disadvantages of the two classic designs – panel and RCS. Based on that, we will deduce a solution to merge both approaches into a single integrated design. Following that, the “Election Study Baden-Wuerttemberg 2011” will be presented as an attempt to actually realize such a merged design. The discussion will exemplify possible analyses based on such data, but also an evaluation of data quality.

1 Einleitung^{1,2}

Der wahlpolitische Prozess in Deutschland im Allgemeinen sowie die Beziehungen zwischen Wählern und Parteien im Speziellen sind seit einigen Jahren deutlichen Veränderungen unterworfen. Abnehmende Wahlfreude, zunehmende Wankelmütigkeit und eine Verlagerung der individuellen Wahlentscheidung bis spät in den Wahlkampf hinein als Grundlage dieser Phänomene sind die Kennzeichen des heutigen Elektorats. Selbstverständlichkeiten im Wahlprozess, die es früher einmal gegeben hat, sind selten geworden. Wähler „verhalten“ sich offenkundig nicht mehr nur, sie handeln aktiv und entscheiden sich – für oder gegen eine Partei, aber auch zunehmend für oder gegen die Teilnahme an einer Wahl. Wahlkämpfe und Wahlabende sind infolgedessen eher durch Unübersichtlichkeit als durch Selbstverständlichkeiten geprägt. An die Stelle von Stabilität tritt zunehmende Dynamik.

Diese fundamentalen Verschiebungen bleiben nicht ohne Folgen für die akademische Wahlforschung. Die Beschreibung und Erklärung von Stabilität haben gänzlich andere Erfordernisse als die Beschreibung und Erklärung von Veränderung und Dynamik. Dies gilt sowohl in theoretischer als auch in methodischer Hinsicht. Aus theoretischer Sicht stehen sich auf der einen Seite mikro- und makrosoziologische Modelle, auf der anderen Seite informationsorientierte Modelle aus der Politischen Kommunikationsforschung und der Politischen Psychologie gegenüber. Ein Fokus auf (langfristige) Stabilität dominiert erstere, während der Umgang mit Informationen, die per se kurzfristiger Natur sind, letztere prägt. Forschungsgeschichtlich ist dabei eine deutliche Verschiebung hin zu informationsorientierten Modellen erkennbar – wenig verwunderlich angesichts der skizzierten Entwicklungen.

Allerdings weist – im Gegensatz zur Erfassung von (unterstellter) Stabilität, für deren Zweck eine einmalige Messung von (vermeintlich) stabilen Merkmalen ausreicht – die Erfassung von Dynamik besondere Erfordernisse auf, was nicht zuletzt auch ein Blick in den Werkzeugkasten politischer Psychologen zeigt. Dynamische Prozesse der Informationsverarbeitung überhaupt abbilden zu können, ist keine Selbstverständlichkeit, weshalb Psychologen auch bevorzugt auf das kontrollierte Setting von Laborexperimenten zurückgreifen.

1 Wir danken den beiden anonymen Gutachtern für ihre hilfreichen Kommentare und Anregungen sowie Michelle Kohler und Judith Schenk für ihre Unterstützung bei der Finalisierung des Manuskripts.

2 Die Struktur des vorliegenden Beitrags orientiert sich an den Methodenberichten zu den beiden Rolling Cross-Section Studien, die im Vorfeld der Bundestagswahlen 2005 und 2009 unter Leitung von Rüdiger Schmitt-Beck realisiert wurden (siehe hierzu Schmitt-Beck/Faas/Holst 2006; Schmitt-Beck/Faas/Wolsing 2010).

Möchte man in der Welt von Bevölkerungsumfragen bleiben, so zeigt sich, dass bislang stets eines von zwei verfügbaren Instrumenten eingesetzt worden ist, um Dynamik – gerade mit Blick auf Wahlkämpfe – abzubilden: das Paneldesign und das Design rollierender Querschnitte. Beide Verfahren bringen große Vorteile mit sich, haben aber auch Nachteile: Erlauben Paneldaten die Beobachtung *individueller* Veränderung von Einstellungen und Verhaltensabsichten zwischen zwei oder mehreren Zeitpunkten, fehlt in der Regel die Sensibilität für *tagesaktuelle* Veränderungen. Genau umgekehrt verhält es sich beim Rolling Cross-Section Design (RCS; vgl. Johnston 2001; Johnston/Brady 2002). Hier wird es durch die tagesgenaue Aussteuerung der Befragung möglich, die Veränderung der Aggregatverteilung über den Befragungszeitraum hinweg feinkörnig (bis hin zu tagesgenauer Beobachtung) zu untersuchen. Offen muss dabei allerdings bleiben, bei welchen Personen Veränderungen auftreten, da jeden Tag andere Personen befragt werden. Vor- und Nachteile der beiden Designs verhalten sich demnach komplementär zueinander – eine Kombination beider Verfahren erscheint daher höchst wünschenswert. Allerdings steht der Nachweis, dass eine solche Kombination tatsächlich möglich ist, bislang noch aus.

An dieser Stelle setzt der vorliegende Beitrag an: Im Mittelpunkt steht die „Wahlstudie Baden-Württemberg 2011“ – eine Studie, die rund um die Landtagswahl in Baden-Württemberg vom 27. März 2011 realisiert werden konnte. In diesem Rahmen wurde die Idee einer Panelstudie so mit den grundlegenden Prinzipien einer Rolling Cross-Section Studie verbunden, dass am Ende eine Studie resultierte, die die Vorzüge beider Ansätze vereinen soll.

Der vorliegende Beitrag gliedert sich wie folgt: Zunächst werden die jeweiligen Vor- und Nachteile der beiden klassischen Forschungsdesigns (Panel vs. RCS) in aller Kürze skizziert, um daraus die Möglichkeit einer Kombination beider Ansätze abzuleiten. Anschließend wird im dritten Abschnitt die „Wahlstudie Baden-Württemberg 2011“ als konkreter Versuch, eine solche Studie zu realisieren, beschrieben. Die folgenden Kapitel beschäftigen sich zunächst mit der resultierenden Datenqualität dieser Studie (im Sinne einer notwendigen Bedingung), um abschließend an einem ausgewählten Beispiel das besondere Analysepotenzial der Daten aus dem integrierten Design aufzuzeigen.

2 Forschungsdesigns zur Erfassung von Wahlkampfdynamik

2.1 Möglichkeit 1: Panelstudien

Dass Paul Lazarsfeld heute – aus substanzwissenschaftlicher Sicht – mit dem mikrosoziologischen Ansatz und damit eher stabilem Wählerverhalten verbunden wird, ändert nichts an der Tatsache, dass er – aus methodischer Sicht – einen Meilenstein zur Erforschung von (möglicher) Dynamik von Wahlkämpfen gelegt hat: Lazarsfeld hat mit seinen Mitwirkenden in den 1940er-Jahren das Paneldesign in der Wahlforschung verankert (Lazarsfeld/Berelson/Gaudet 1968). In der Erwartung, Wahlkampfdynamik zu finden (was vor allem im Untertitel seiner einschlägigen Studie – „How the Voter Makes Up His Mind in a Presidential Campaign“ – zum Ausdruck kommt), hat er in Erie County eine Panelbefragung (also das mehrfache Befragen eines identischen Personenkreises) realisiert (siehe Tabelle 1). So wollten Lazarsfeld und sein Team den Entscheidungsprozess der Bürger von Erie County sichtbar und erklärbar machen. Dass sich dieser Prozess, wie die Ergebnisse der Studie zeigten, bei der großen Mehrheit der damaligen Probanden in der empirischen Realität als eher statisch-stabil erwies, ändert an dem Potenzial des Lazarsfeld'schen Designs überhaupt nichts – auch wenn es praktisch zur Folge hatte, dass sein Design mangels substanzieller Notwendigkeit lange Zeit wenig Nachahmer fand.

Tabelle 1 Das Lazarsfeld'sche Paneldesign, Basis von „The People's Choice“ (in Anlehnung an Lazarsfeld/Berelson/Gaudet 1968: 4)

Zeit	Mai	Juni	Juli	Aug.	Sep.	Okt.	Nov.
Panel		600	600	600	600	600	600
	3000						
Kontrollgruppe			A 600	B 600		C 600	

Die Genialität des Designs ergibt sich dabei nicht nur aus der mehrfachen Befragung eines identischen Personenkreises, sondern zeigt sich insbesondere in der zusätzlichen Berücksichtigung von Kontrollgruppen. Von den ursprünglich im Mai rekrutierten 3000 Personen sollte nämlich nur eine kleine Gruppe von 600 fortan in monatlichen Abständen wiederholt interviewt werden. Diese Gruppe eröffnete den Forschern die Möglichkeit, den Entscheidungsprozess dieses Personenkreises auf individueller Ebene nachzuverfolgen.

Zugleich aber birgt die mehrfache Befragung von Personen die Gefahr, dass das Befragen selbst Einfluss auf die Probanden und ihre Orientierungen nimmt (Sobol 1959)³. Zu verhindern ist eine solche Konditionierung im Rahmen einer Panelstudie natürlich nicht; aber durch entsprechende Kontrollgruppen lässt sie sich zumindest sichtbar machen. Lazarsfeld und sein Team führten daher parallel zu den Panelbefragungen der Monate Juli, August und Oktober zusätzliche Befragungen von Personen durch, die zwar ebenfalls im Mai rekrutiert worden waren, aber eben nicht an den weiteren Panelwellen teilgenommen hatten. Treten bei gleichzeitiger Befragung der Kern- und einer Kontrollgruppe Unterschiede in den Ergebnissen zwischen beiden Gruppen auf, so deutet dies auf einen Effekt der Befragung selbst hin.

Eine solche Panelkonditionierung ist nicht die einzige Möglichkeit zur Erklärung von Unterschieden zwischen Kontroll- und Kerngruppe. Alternativ kommt in Frage, dass sich die beiden Gruppen nach einiger Zeit in ihrer *Zusammensetzung* systematisch unterscheiden. Das sollte zwar seitens der Forscher tunlichst vermieden werden, ist aber mitunter nicht zu verhindern – etwa weil Personen aus der Kerngruppe des Panels infolge von Krankheit, Umzug, Tod oder schlicht der fehlenden Bereitschaft heraus, weiterhin interviewt zu werden, ausscheiden. In diesem Fall spricht man von „Panelmortalität“ (Stadtmüller 2009: 111).

Kurzum: Eine Panelstudie hat den großen Vorteil, Dynamik auf der Ebene einzelner Personen messen zu können, was allerdings zugleich die Gefahr von Konditionierungseffekten bei genau diesen Personen mit sich bringt. Auch ein gewisses Maß an Panelmortalität (und insbesondere die damit verbundene Gefahr einer systematischen Verzerrung) scheinen unvermeidlich.

Darüber hinaus weisen Panelstudien auch Schwächen auf, wenn es um die *Erklärung* von Wandel geht – insbesondere wenn dieser Wandel durch Rückbezug auf tatsächliche Ereignisse im Wahlkampf erklärt werden soll (vgl. Schmitt-Beck/Faas/Wolsing 2010). Das Lazarsfeld'sche Design sah, wie Tabelle 1 zeigt, Befragungen in monatlichem Abstand vor. Ergeben sich nun Unterschiede zwischen zwei Erhebungszeitpunkten, so sind diese (wohl) auf Ereignisse, die in diesem Monat zwischenzeitlich stattgefunden haben, zurückzuführen. Eine präzisere Zuschreibung, etwa zu Einzelereignissen, ist in der Regel nicht möglich, was im Sinne einer sauberen Erklärung nicht gänzlich befriedigend ist.

3 Die Probanden können durch die Befragung beispielsweise etwas über den Gegenstand lernen, sich gezielt auf weitere Befragungsteile vorbereiten oder erst durch die Befragung selbst Einstellungen zu bestimmten Themen entwickeln (Lenz 2009; Toepel/Das/van Soest 2009; Das/Toepel/van Soest 2011).

Der Taktfrequenz sind allerdings sowohl aus logistischen als auch aus ethischen Gründen (gegenüber den Befragten) Grenzen gesetzt; sie lässt sich nicht beliebig reduzieren. Nun könnte man einwenden, dass wichtige Einzelereignisse antizipierbar sind und die Panelwellen entsprechend terminiert werden können, um bestimmte, ereignisbezogene Ursache-Wirkungs-Vermutungen einem stringenteren (Panel-)Test zu unterziehen. Tatsächlich finden sich solche Studien, etwa in Form von Kurzfristpanels im Kontext von Fernsehduellen (vgl. Faas/Maier 2011). Kennzeichen der so analysierbaren Ereignisse ist allerdings genau, dass a) sie lange im Voraus bekannt sind und b) theoretische Erwartungen hinsichtlich möglicher Effekte existieren.

Treten jedoch unerwartete Ereignisse auf, sind die Auswirkungen nicht theoretisch klar umrissen oder finden innerhalb kürzerer Zeit mehrere Ereignisse statt, stößt ein Paneldesign zwangsläufig an seine Grenzen (vgl. Schmitt-Beck/Faas/Wolsing 2010: 17). Dass Wahlkämpfe in der heutigen Zeit tatsächlich durch unvorhersehbare Ereignisse geprägt sind, hat nicht zuletzt die Wahl in Baden-Württemberg exemplarisch gezeigt: Die Atomkatastrophe von Fukushima ist ein prototypisches Beispiel dafür, wie unvorhergesehene Ereignisse einen Wahlkampf prägen, aber in einem Panel nicht abgebildet werden können. Um (mögliche) Effekte dennoch prüfen zu können, scheinen Designalternativen zwingend notwendig.

2.2 Möglichkeit 2: Rollierende Querschnitte

Solche Alternativen, die sowohl dem Problem der Taktung als auch der (Nicht-)Antizipierbarkeit von Ereignissen begegnen, existieren durchaus, nämlich in Form des Rolling Cross-Section Designs (Johnston 2001; Johnston/Brady 2002; Romer et al. 2006; Schmitt-Beck/Faas/Holst 2006; Schmitt-Beck/Faas/Wolsing 2010). Die Grundidee solcher Erhebungen besteht darin, „eine Querschnittsbefragung kontrolliert so über einen vorab definierten Zeitraum zu spreizen, dass auch die Befragten jedes einzelnen Tages jeweils eine Zufallsstichprobe aus der Grundgesamtheit konstituieren“ (Schmitt-Beck/Faas/Wolsing 2010: 3). Praktisch bedeutet das, dass die Stichprobenziehung zweistufig erfolgt: Im ersten Schritt wird eine herkömmliche Zufallsstichprobe, zum Beispiel von Telefonnummern gezogen, die dann allerdings im zweiten Schritt in weitere Einladungspakete („Replikate“) unterteilt wird. Pro Tag des anvisierten Beobachtungszeitraums wird ein Einladungspaket freigegeben und an diesem – wie auch den folgenden Tagen – nach einem strikten Kontaktprotokoll abgearbeitet.

Die tagesgenaue Steuerung hat das Ziel, den Beobachtungszeitraum in standardisierter Form feinkörnig zu untersuchen. Der Umstand, dass die einzel-

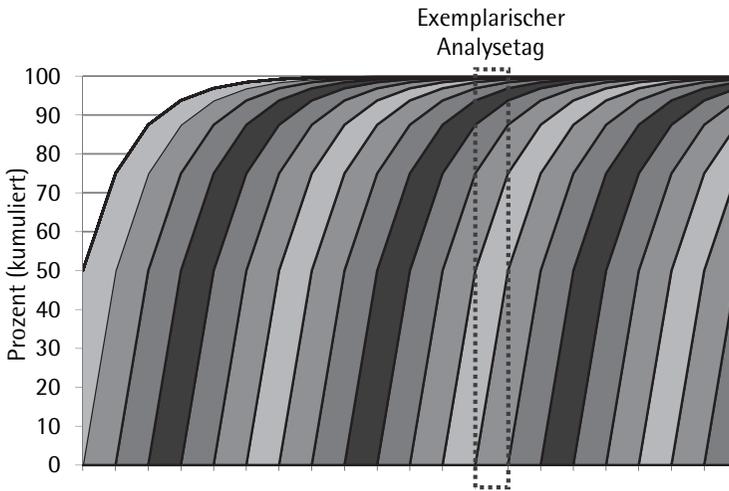
nen Einladungspakete nicht nur am Tag ihrer Freigabe, sondern auch in den darauf folgenden Tagen in streng kontrollierter Weise abgearbeitet werden, dient der Erhöhung der Repräsentativität und Varianz der resultierenden Daten. Es ist davon auszugehen, dass unmittelbar am Tag der Freigabe zunächst vor allem leicht erreichbare Personen erfolgreich kontaktiert und interviewt werden können. Erst durch die erweiterte Kontaktoutine über diesen ersten Tag hinaus werden auch schwer erreichbare Personen letztlich mit einbezogen. Praktisch schiebt man systematisch eine Welle schwer erreichbarer Personen vor sich her, was für jeden einzelnen Befragungstag bedeutet, dass man sowohl leicht als auch schwer erreichbare Personen tatsächlich interviewt – letztere aus den an den Vortagen freigegebenen Paketen.

Durch die zufällige Verteilung der Ausgangsstichprobe auf die einzelnen Replikate und die gleichförmige Abarbeitung entstehen nach kurzer Anlaufzeit strukturgleiche *Tagesstichproben*. Diese sind nicht mehr (wie zunächst die Replikate) über den Tag ihrer Freigabe, sondern über den Tag des tatsächlichen Interviews definiert. Überzufällige Unterschiede zwischen zwei Tagesstichproben sind – bei gegebener Strukturgleichheit – nur durch Ereignisse zu erklären, die zwischen den beiden Tagen stattgefunden haben. Das RCS-Design weist an dieser Stelle deutliche Parallelen zu einer experimentellen Logik (mit all ihren Vorteilen) auf, wobei das experimentelle Treatment aus den Ereignissen eines Tages besteht.

Abbildung 1 verdeutlicht die Logik noch einmal grafisch. Ihr liegt die Annahme zugrunde, dass 50 Prozent eines Replikats am Tag seiner Freigabe erreicht und interviewt werden, von den verbleibenden Personen nochmals 50 Prozent am Folgetag usw. Wie der exemplarische Erhebungstag zeigt, resultieren schon nach wenigen Tagen gut durchmischte Stichproben aus leicht und schwer erreichbaren Personen, die sich im Vergleich der verschiedenen Stichproben strukturell in nichts unterscheiden als dem Erhebungstag und seinem spezifischen Kontext.

Die Vorteile dieses Designs gegenüber dem Lazarsfeld'schen Paneldesign sind offensichtlich: Zu grobe Taktfrequenzen gibt es hier nicht – im Gegenteil: Tagesgenau lassen sich Bewegungen der öffentlichen Meinung nachzeichnen. Dabei beschränken sich die Möglichkeiten keineswegs darauf, nur zwei Tagesstichproben zu vergleichen. Die einzelnen Tagesstichproben lassen sich beliebig kumulieren – nach formalen Kriterien (z. B. Wochen) oder auch rund um Ereignisse. Antizipierbare ebenso wie völlig überraschende Ereignisse können in dieser Datenstruktur sichtbar ihre Spuren hinterlassen, sofern das Fragenprogramm entsprechend offen gestaltet ist.

Abbildung 1 Logik einer Rolling Cross-Section Studie



Es ist mit dem RCS-Design jedoch – und dies im Gegensatz zu einem klassischen Paneldesign – auch ein Nachteil verbunden: Diese Spuren werden nur im Aggregat oder zumindest in Teilgruppen hinterlassen. Kausale Inferenzen auf individueller Ebene sind unmöglich (vgl. Lenz 2009). Stattdessen greift eine quasi-experimentelle Logik, die bekanntlich auch auf Gruppenvergleichen (zwischen der Experimental- und der Kontrollgruppe) aufbaut: Unterschiede zwischen (Tages-)Gruppen lassen sich sehr präzise identifizieren. Der Frage allerdings, wer innerhalb der Gruppe auf ein Ereignis reagiert hat und warum, kann man sich nur über die Analyse von Subgruppen nähern. Auf die individuelle Ebene einzelner Befragter und den dort wirksamen Mechanismus kommt man mit einem RCS-Design nicht.

2.3 Rollierende Panelwellen als Möglichkeit 3?

Die bisherigen Ausführungen haben gezeigt: Panel- wie Rolling Cross-Section Studien haben ihre spezifischen Stärken und Schwächen. Entwicklungen auf der individuellen Ebene und die damit verbundenen Möglichkeiten kausaler Inferenz lassen sich nur mit Hilfe mehrfacher Befragungen identischer Personen nachzeichnen. Allerdings dürfen Befragungspersonen dabei nicht überstrapaziert werden, was zwangsläufig eine gröbere Beobachtungsfrequenz bei Panelstudien zur Folge hat. Zudem eignen sich Panelbefragungen besonders gut für antizipierbare Ereignisse, die aber – wie ein Blick gerade auf Wahlkämpfe der jüngeren Vergangenheit

bestätigt – nur eine Teilmenge (vermutlich) relevanter Ereignisse umfassen. Hier wiederum liegt der große Vorteil von RCS-Studien, mittels derer sich auch längere Zeiträume sehr präzise beobachten lassen. Kurzum: Stärken und Schwächen verhalten sich komplementär zueinander.

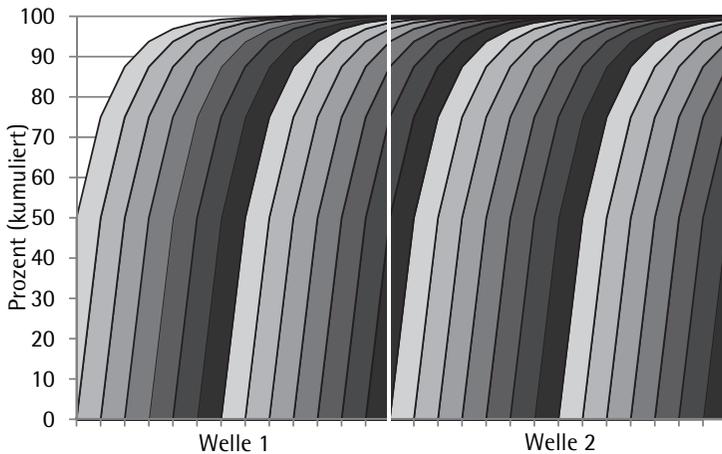
Ein Blick in die Forschungspraxis zeigt, dass bislang zumeist entweder die eine oder die andere Variante zur Beobachtung von Wahlkampfdynamik zum Einsatz gekommen ist. Mitunter – wie im Falle der German Longitudinal Election Study (GLES) – werden im Rahmen einer modularen Struktur einer Wahlstudie auch beide Varianten gleichzeitig realisiert, ohne allerdings unmittelbar miteinander in Verbindung zu stehen.⁴

Dabei ist eine Integration beider Prinzipien in einem integrierten Studiendesign durchaus vorstell- und machbar. So werden schon in der bisherigen Praxis RCS-Vorwahlumfragen mit einer zweiten Panelwelle kombiniert, die nach der Wahl erhoben wird, um damit Schlüsse auf individueller Ebene zu ermöglichen (vgl. Johnston/Brady 2002). Die bislang vorliegenden Studien erlauben es allerdings nicht, intraindividuelle Entwicklungen innerhalb der Vorwahlperiode zu beobachten; zudem ist im skizzierten erweiterten RCS-Design die Anzahl der realisierten Wellen mit $N=2$ minimal.

Doch auch eine weiterreichende Integration beider Ansätze erscheint möglich. Ausgehend vom Lazarsfeld'schen Paneldesign können auch einzelne Panelwellen in rollierender Art und Weise gestaltet werden. Abbildung 2 deutet dies in schematischer Weise an. Zu sehen ist eine Reihe von Panelwellen, die in sich wiederum in kontrollierter Art und Weise – einer RCS-Logik folgend – ins Feld gegeben und abgearbeitet werden. Die Panelwellen fungieren also sowohl als Panel als auch als rollierende Querschnitte, ohne dass eine Eigenschaft die andere massiv beeinflusst. Aus der Perspektive einer Panellogik heraus ändert sich also lediglich die Feldsteuerung des Panels; aus der RCS-Perspektive werden an den einzelnen Tagen nicht fortlaufend *neue* Replikate, sondern ab einem bestimmten Tag wieder alte Replikate (aus Vorwellen) erneut ins Feld gegeben. Auf diese Weise ändert sich zwar das konkrete Vorgehen bei der Durchführung der Designs, der Kern der Designs wird davon jedoch nicht tangiert.

4 Im Rahmen der GLES gibt es darüber hinaus auch noch ein spezifisches Modul zu Fernsehduellen, in dessen Rahmen ein vierwelliges Paneldesign, rund um dieses wichtigste Einzelereignis in heutigen Wahlkämpfen, realisiert wird.

Abbildung 2 Das Rollierende Panel



Bevor am Beispiel der „Wahlstudie Baden-Württemberg 2011“ nun detaillierter auf die konkrete Ausgestaltung einer solchen Studie eingegangen wird, ist abschließend noch ein Hinweis nötig. Bislang wurde die Frage des Erhebungsmodus gänzlich ausgeblendet. Dabei finden sich für die beiden Grundprinzipien durchaus typische Modi: Lazarsfeld hat sein Paneldesign mit persönlichen Interviews realisiert; RCS-Studien werden zumeist telefonisch durchgeführt.

In jüngerer Vergangenheit sind für beide Designs auch vermehrt Onlineumfragen eingesetzt worden, wie etwa im Kontext der amerikanischen, britischen oder auch der deutschen Wahlstudie. Generell kommen Onlineumfragen in sehr unterschiedlichen Gewändern (mit sehr unterschiedlichen Qualitätsansprüchen) daher (vgl. Schonlau/Fricke/Elliott 2002). Wenn für Wahlstudien auf Onlineumfragen zurückgegriffen wird, so bauen diese meist auf der Idee so genannter Access Panels auf. Internetnutzer können sich hier – je nach Panel – selbstständig als Mitglieder solcher Access Panels registrieren und stehen damit fortan für zukünftige Onlinebefragungen zur Verfügung. Soll dann eine konkrete Onlinebefragung realisiert werden, können Personen – gegebenenfalls quotiert – aus diesem Pool von Befragungswilligen dazu eingeladen werden.

Natürlich sind mit Onlineumfragen gewisse Einschränkungen hinsichtlich ihrer Repräsentativität verbunden, nicht zuletzt, weil Personen ohne Internetzugang zwangsläufig außen vor bleiben. Zudem ist unklar, welchen Qualitätsansprüchen die Befragten und ihr Antwortverhalten letztlich genügen. Entsprechend ist die Verwendung solcher Studien nicht unumstritten (vgl. Chang/Krosnick 2009; Faas/Schoen 2006; Sanders et al. 2007).

Dennoch bieten Online-Befragungen gerade für Panelstudien (im Lazarsfeld'schen Sinne) sehr günstige Voraussetzungen, da die Teilnahmebereitschaft der Probanden für das Gelingen maßgeblich ist. Diese ist bei Access Panels vergleichsweise hoch und kann gegebenenfalls durch entsprechende Incentives noch weiter erhöht werden. Insgesamt scheint daher diese Variante einer Onlinebefragung für das uns hier interessierende integrierte Design durchaus geeignet, wohl wissend, dass damit Einbußen hinsichtlich der Repräsentativität verbunden sind.

Mit einem ähnlichen Zielkonflikt sah sich Lazarsfeld bei seiner lokalen Panelstudie in Erie County konfrontiert, den sein Team und er mit dem Satz „In any case, we were studying the *development* of votes and not their distribution“ (Lazarsfeld/Berelson/Gaudet 1968: 3) entschieden. Dieser Sichtweise haben auch wir uns mit Blick auf die „Wahlstudie Baden-Württemberg 2011“ angeschlossen, die es im Folgenden vorzustellen gilt.

3 „Wahlstudie Baden-Württemberg 2011“

3.1 Grundkonzept

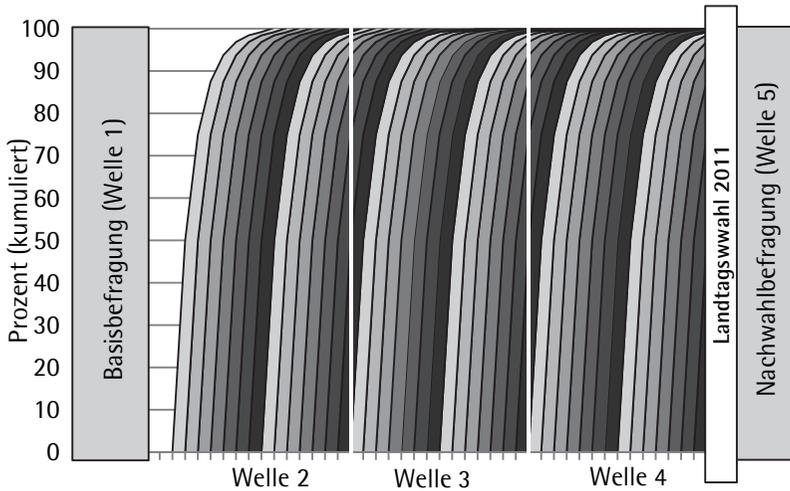
Aufbauend auf den RCS-Studien, die im Vorfeld der Bundestagswahlen 2005 und 2009 realisiert werden konnten (Schmitt-Beck/Faas/Holst 2006; Schmitt-Beck/Faas/Wolsing 2010), haben wir im Umfeld der baden-württembergischen Landtagswahl 2011 eine Studie realisiert, die in der skizzierten Art und Weise Panelwellen mit dem Prinzip rollierender Querschnitte verbindet: die „Wahlstudie Baden-Württemberg 2011“ (Abbildung 3).⁵ Insgesamt waren darin fünf Befragungswellen vorgesehen: eine Basismessung noch Ende 2010 (Panelwelle 1), drei Wahlkampfwellen vor der Wahl (Panelwellen 2 bis 4), die jeweils in einer RCS-Logik realisiert werden sollten sowie schließlich eine fünfte Welle nach der Landtagswahl (Panelwelle 5).

Die Basismessung war vom 18. November 2010 bis zum 2. Februar 2011 im Feld. Ihr Ziel war es, die Einstellungen der Befragten zunächst möglichst frei von Wahlkampfeffekten zu messen. Gleichzeitig ist erfahrungsgemäß die Mortalität bei einer Panelbefragung zwischen der ersten und zweiten Welle am größten. Da Mortalität jedoch einen erheblichen Effekt auf das rollierende Design hätte, sollte diese mit Blick auf die Wahlkampfwellen selbst auf ein Minimum reduziert werden – vor allem durch die vorgeschaltete Basismessung. Insgesamt konnten im Rahmen der

5 Die Studie ist im Rahmen des Juniorprofessorenprogramms des Landes Baden-Württemberg gefördert worden.

Basismessung 3.147 vollständige Interviews realisiert werden. Rekrutiert wurden die Befragten dabei aus den Access Panels von YouGov (n=2.600) und Respondi (n=547) unter der Maßgabe, die Bevölkerung Baden-Württembergs in ihrer Struktur möglichst gut abzubilden.⁶

Abbildung 3 Studiendesign der Wahlstudie Baden-Württemberg 2011



Diese Personen bilden den Ausgangspool von Personen für die dynamische Betrachtung des unmittelbaren Vorwahlzeitraums. Beobachtet wurden die letzten sechs Wochen vor der Landtagswahl, die am 27. März 2011 stattfand. Drei Panelwellen, jeweils gespreizt über einen Zeitraum von zwei Wochen, waren dafür vorgesehen. Dazu wurden die 3.147 Personen, die den Fragebogen der Basismessung vollständig ausgefüllt hatten, zufällig in 14 gleich große Replikate unterteilt, von denen täglich eines – beginnend am 13. Februar 2011 – zur Befragung freigegeben wurde.⁷

6 Sowohl die lange Feldzeit der ersten Welle (von November bis in die ersten Februartage hinein) als auch die Rekrutierung aus zwei verschiedenen Access Panels sind Folge eines eher schleppenden Verlaufs der Rekrutierung von Probanden für die erste Welle. Access Panels – obwohl sie mehrere Tausend Personen umfassen – stoßen derzeit offenkundig für Studien mit einem regionalen Bezug (wie die „Wahlstudie Baden-Württemberg“) noch an Grenzen. Diesem Problem sind wir durch die lange Feldzeit und den zusätzlichen Rückgriff auf ein zweites Access Panel begegnet.

7 Tatsächlich wurden, wie auch Tabelle 2 zeigt, nur 3.078 Personen zur zweiten Welle eingeladen, da sich zwischen den beiden Wellen insgesamt 59 Personen aus den Access Panels abgemeldet hatten.

An den einzelnen Tagen erhielten die Mitglieder der jeweiligen Tagespakete morgens eine Einladung zur Teilnahme an der Studie per E-Mail zugesandt („Einladung zu einer wissenschaftlichen Studie“). Drei Tage später wurden diejenigen, die bis dahin noch nicht reagiert hatten, an die Studie erinnert, nochmal drei Tage später folgte eine zweite Erinnerung. Weitere drei Tage später wurde das Paket geschlossen, danach war eine Teilnahme der restlichen Personen dieses Tagespakets nicht mehr möglich. Jede Zielperson hatte also beginnend mit ihrem jeweiligen Einladungstag genau neun Tage Zeit, um an der Befragung teilzunehmen.⁸ Durch die Erinnerungen und die lange Öffnung der Befragung wurde sichergestellt, dass auch schwer erreichbare Personen ausreichende Möglichkeiten erhielten, an der Umfrage teilzunehmen, was sich positiv auf die letztlich resultierende Varianz der Tagesstichproben auswirken sollte.

Der Design-Unterschied zu einer reinen RCS-Studie, in deren Rahmen für jeden Tag der Feldzeit ein neues Replikat ins Feld gegeben worden wäre, und damit die Verbindung des Panel- mit dem RCS-Prinzip stellt sich nach genau zwei Wochen ein. Nach 14 Tagen nämlich – also beginnend am 27. Februar 2011 – erhielten die Befragten die Einladung zur zweiten, nach weiteren 14 Tagen – ab dem 13. März 2011 – zur dritten und letzten Wahlkampfwellen.⁹ Die zufälligen Zuweisungen zu einzelnen Tagespaketen für die erste Wahlkampfwellen bestimmten dabei auch die Zugehörigkeiten zu den Paketen der folgenden Vorwählerhebungen.^{10, 11}

- 8 Vor dem Hintergrund der Tatsache, dass Befragte bis zu neun Tage Zeit hatten, um an einer Befragungswelle teilzunehmen, sind an dieser Stelle zwei Hinweise angebracht: Erstens kann es infolgedessen passieren, dass Personen zwar zur Feldzeit der ersten Wahlkampfwellen eingeladen werden, aber erst zur Feldzeit der zweiten Welle teilnehmen. An dieser Stelle vermischen sich die beiden Prinzipien und der genaue Umgang mit den Daten hängt von der konkreten Fragestellung ab. Zweitens setzt die Landtagswahl vom 27. März den Vorwählerbefragungen ein natürliches Ende, so dass Personen, die erst spät zur letzten Wahlkampfwellen eingeladen wurden, nicht mehr die vollen neun Tage Zeit hatten, um unsere Fragen zu beantworten. Für das RCS-Design ist das unproblematisch, für das Paneldesign allerdings fehlen hier systematisch Befragte dieser Welle im Vergleich zu den Vorwellen.
- 9 Um es aus einer anderen Perspektive noch einmal auf den Punkt zu bringen: Am 15. Feldtag würde man in einem klassischen RCS-Design ein neues Replikat mit gänzlich neuen Befragten ins Feld geben; im Design rollierender Panelwellen wird am 15. Feldtag kein neues, sondern erneut das Replikat des ersten Feldtags ins Feld gegeben.
- 10 Der Fragebogen war über den gesamten Vorwahlzeitraum weitgehend konstant. Gleichwohl haben wir einige Fragen von Welle zu Welle ausgetauscht, um angesichts der relativ kurzen Abstände zwischen den Wellen für eine gewisse Abwechslung für die Befragten zu sorgen. Der konkrete Fragebogen, den die Befragten jeweils erhielten, wurde dabei durch den Zeitpunkt des erstmaligen Aufrufens der Studie bestimmt. Wurde also eine Person am letzten Samstag der zweiten Welle eingeladen, nahm aber erst am darauf folgenden Montag an der Studie teil, so erhielt sie den Fragebogen der dritten Welle. Im Durchschnitt dauerten die Befragungen 20 Minuten.
- 11 Parallel zur vierten Panelwellen wurde, ebenfalls einer RCS-Logik folgend, eine neue Befragungswelle mit knapp 1.000 Befragten gestartet, um – wie bei Lazarsfeld/Berelson/Gaudet (1968) – eine mögliche Panelkonditionierung der Kerngruppe des Panels potenziell sichtbar machen zu können. Darauf wollen wir an dieser Stelle aber nicht näher eingehen.

Nach der Wahl wurden die Teilnehmer aller vorherigen Wellen zu einer fünften Welle eingeladen, welche unmittelbar nach der Wahl gestartet wurde und bis zum 18. April lief. Im Mittelpunkt dieser Nachwahlwelle standen die retrospektive Perspektive der Wahlentscheidungsfindung und vor allem die Wahlentscheidung selbst als tatsächliches Verhalten.¹² Tabelle 2 gibt Aufschluss über die genauen Feldzeiten, Einladungen und resultierenden Fallzahlen. Wie sich hier zeigt, hat sich die Vorschaltung einer Basismessung bewährt: Die Rekrutierung der Befragten für die Basismessung war in der Tat beschwerlich (und die Rücklaufquote entsprechend niedrig); mit dem Eintreten in die unmittelbare Vorwahlphase mit ihren rollierenden Wellen aber stellt sich eine erfreulich hohe und vor allem stabile Rücklaufquote von Welle zu Welle ein.¹³

Tabelle 2 Die Panelwellen der Wahlstudie

	W1 Basismessung	W2 Wahlkampf	W3 Wahlkampf	W4 Wahlkampf	W5 Nachwahl
Feldzeit	18.11. bis 02.02.	13.02. bis 26.02.	27.02. bis 12.03.	13.03. bis 26.03.	28.03. bis 18.04.
Eingeladene Personen	11989	3078	3052	3048	2916
Vollständige Interviews (Prozent der Einladungen)	3147 (26,3 %)	2230 (72,4 %)	2059 (67,5 %)	1895 (62,2 %)	2130 (73,1 %)

3.2 Feldverlauf und Datenqualität

Sowohl die Panel- als auch die RCS-Elemente stellen an die Erhebungsphase ganz eigene Herausforderungen. Während im Panel die erreichten Fallzahlen über die Wellen hinweg stabil gehalten werden sollten, ist die strukturelle Konstanz der einzelnen Tagesstichproben aus RCS-Sicht zentral. Dabei ist weiterhin zu bedenken, dass durch die Basismessung das Teilnehmerfeld für die rollierenden Wellen von Beginn an festgelegt war und keinerlei Nachjustierungen während der unmittelba-

- 12 Aufgrund der besonderen Situation nach der Wahl mit der Übernahme der Regierungsgeschäfte durch den ersten grünen Ministerpräsidenten in Deutschland wurde vom 17. Mai bis zum 25. Mai 2011 noch eine sechste Panelwelle durchgeführt – auf diese wollen wir aber hier nicht näher eingehen.
- 13 Dabei ist für die Rücklaufquote der vierten Welle unmittelbar vor der Wahl zu bedenken, dass diese Welle bedingt durch die Wahl vom 27. März 2011 trunziert wurde, was die Rücklaufquote im Vergleich zu den vorhergehenden Wellen zwangsläufig schmälert. Die leicht rückläufige Zahl *eingeladener* Personen, die Tabelle 2 ebenfalls ausweist, ist dabei auf Abmeldungen aus dem Access Panel zurückzuführen.

ren Vorwahlphase mehr ermöglichte. Dem Feldverlauf kommt daher für das Gelingen der Studie zentrale Bedeutung zu.

Für die Panelkomponente stellt sich also besonders die Frage nach dem Ausmaß und der Systematik einer möglichen *Panelmortalität*. Je mehr Wellen eine Studie insgesamt hat und je größer die Distanz zwischen den einzelnen Wellen ist, desto höher ist die Wahrscheinlichkeit, dass Probanden aus dem Panel ausscheiden und sich so der durchgehend befragte Personenkreis (in potenziell systematischer Weise) verkleinert.

Welche Muster ergeben sich aus dieser Logik heraus am Beispiel der „Wahlstudie Baden-Württemberg 2011“? Tabelle 3 zeigt die Teilnahmemuster für die 3.147 ursprünglich befragten Personen der Basismessung. Dabei wird deutlich, dass es – erwartungsgemäß – Panelmortalität gegeben hat und dass diese besonders deutlich beim Übergang von Welle 1 zu Welle 2 auftrat: In knapp einem Fünftel der Fälle finden wir ein Muster, demzufolge eine Person an der ersten Welle teilgenommen hat, danach aber für weitere Wellen nicht mehr zu gewinnen war.

Erfreulicherweise aber weist Tabelle 3 ein anderes Muster (mit fast 50 Prozent) als dominant aus: Nahmen nämlich Probanden an mehr als einer Panelwelle teil, so verblieben sie in der Regel auch für *alle* folgenden Wellen im Panel.¹⁴ Aus der Sicht unseres integrierten Designs ist das Muster, wonach eine Person zwar zusätzlich zur ersten Welle noch an Welle 2 teilgenommen hat, danach aber nicht mehr für weitere Wellen zu gewinnen war, problematisch. Dieses Muster tritt bei rund fünf Prozent der betrachteten Personen auf. Inwieweit systematische Faktoren mit dieser Form der Panelmortalität verbunden sind, werden wir im nächsten Schritt prüfen. Darüber hinaus zeigten sich noch weitere Muster der Teilnahme, die sich aber nicht mehr in bestimmten Clustern konzentrieren. Alles in allem ist das Niveau der Panelmortalität damit erfreulich gering und konzentriert sich vor allem auf den Übergang von Welle 1 zu Welle 2, was den Erwartungen entspricht, aber das Design rollierender Panelwellen nicht tangiert, welches erst mit Welle 2 einsetzt.

14 Zudem ist an dieser Stelle anzumerken, dass am Tag der Landtagswahl alle Pakete für die letzte Vorwahlbefragungswelle geschlossen wurden. Ein Teil der Personen hatte also insgesamt weniger Zeit, um den Fragebogen der Welle 4 abzuschließen, was die Fallzahlen der vierten Welle negativ beeinflusst. Ein Befragungsmuster, bei dem die Teilnahme an der vierten Welle fehlt, deutet in diese Richtung und ist immerhin auch in 6,5 Prozent aller Fälle zu verzeichnen.

Tabelle 3 Muster der Panelteilnahme

Muster	Anteil der Befragten				
	W1	W2	W3	W4	W5
X	X	X	X	X	47,3 %
X					18,1 %
X	X	X		X	6,5 %
X	X				4,7 %
Andere Muster					23,4 %

Unter Teilnahme werden dabei nur vollständig beendete Interviews gefasst.

Tabelle 4 Prädiktoren der Teilnahme an der/den Folgewelle(n)

Modell	1	2	3	4	5	6
	W1:W2	W2:W3	W3:W4	W4:W5	W1:W5	W2:W5
Alter	0,015 ***	0,004	0,014 **	0,008	0,016 ***	0,012 ***
Frau	0,158	-0,078	-0,035	-0,113	0,005	-0,120
Politisches Interesse	0,110 *	0,108	0,080	0,108	0,115 *	0,056
Niedrige Bildung	-0,024	-0,001	-0,010	-0,111	-0,076	0,060
Hohe Bildung	0,259 **	0,092	0,225	-0,040	0,154	0,148
LTW06: Nichtwahl	0,053	-0,124	-0,452 *	-0,058	-0,161	-0,322 *
Stuttgart 21: Gegner	0,079	0,098	0,182	0,312	0,149	0,120
Links-Rechts-Selbsteinstufung	0,005	-0,032	-0,006	-0,036	-0,022	-0,026
N	2933	2052	1768	1541	2872	2038
Pseudo R ²	0,015 ***	0,007	0,019 ***	0,010	0,020 ***	0,015 ***

Logitkoeffizienten; Signifikanzniveaus: * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$

Bei der Betrachtung von Panelmortalität ist deren Ausmaß eine wichtige Komponente; daneben wäre aber insbesondere dann ein Problem für die Studie damit verbunden, wenn sie *selektiv* auftreten würde, bestimmte Personen also systematisch häufiger aus dem Panel ausschieden als andere (Kaspar 2009: 85). Um das Ausmaß selektiver Mortalität empirisch zu ermitteln, haben wir für die jeweiligen Übergänge von Welle zu Welle, für die Teilnahme an allen Wellen sowie für die Teilnahme an den Panelwellen 2 bis 5 jeweils logistische Regressionsmodelle geschätzt. Die abhängige Variable bildet dabei die vollständige Teilnahme an der Folge- beziehungsweise den folgenden Panelwellen (Tabelle 4). Die einbezogenen Erklärungsfaktoren folgen einer eher explorativen Logik aus dem Kontext der Studie; einbezogen haben wir Alter, Geschlecht, Bildung, politisches Interesse einer Person ebenso wie ihre Wahlteilnahme bei der Landtagswahl 2006, ihre Links-Rechts-Selbstein-

stufung sowie ihre Position zu Stuttgart 21. Ausgegraute Prädiktoren (und Modelle insgesamt) in Tabelle 4 erwiesen sich als statistisch nicht signifikant.

Was zeigen nun die Ergebnisse? Erfreulich ist zunächst die durchweg äußerst geringe Erklärungskraft der Modelle; an einigen Stellen *innerhalb des unmittelbaren Vorwahlzeitraums* ist die Erklärungskraft sogar nicht signifikant von Null verschieden. Die höchste Erklärungskraft ergibt sich bei der Betrachtung der Panelmortalität insgesamt, also im Vergleich von Welle 5 mit Welle 1. Bei dieser Gesamtbetrachtung machen höheres Alter und höheres politisches Interesse einen Verbleib im Panel wahrscheinlicher; dies gilt gerade auch für den Übergang von Welle 1 zu Welle 2. Diese Korrelate der Panelmortalität decken sich mit Befunden aus anderen Panelstudien (vgl. etwa Sobol 1959; Kaspar 2009). Erfreulich ist allerdings, dass diejenigen Faktoren, die sich bei Betrachtung *einzelner* Übergänge als signifikante Prädiktoren erweisen, durchaus von diesem globalen Muster abweichen und sich von Welle zu Welle unterscheiden – sofern sie überhaupt auftreten. Verstärkungseffekte im Verlauf der Studie scheint es demnach nicht zu geben.

Insgesamt bleibt aus einer Panellogik heraus festzuhalten, dass bei der „Wahlstudie Baden-Württemberg 2011“ zunächst nach der Basismessung ein großes Teilnehmerfeld bereitstand. Beim Übergang in die dynamische Vorwahlphase ist zwar eine gewisse Panelmortalität zu verzeichnen, die aber – was Niveau und Systematik betrifft – auf erfreulich niedrigem und damit akzeptablem Niveau geblieben ist.

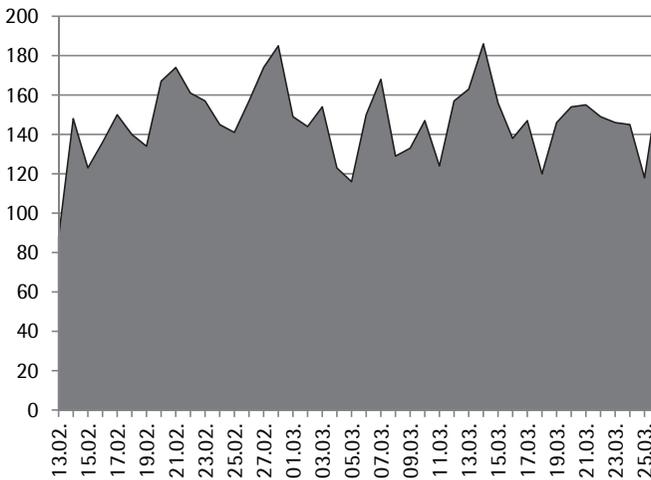
Im nächsten Schritt wollen wir nun unsere Aufmerksamkeit auf Feldverlauf und Durchführung der *rollierenden* Panelwellen lenken, also die RCS-Logik in den Vordergrund rücken. Ziel an dieser Stelle müssen strukturgleiche Tagesstichproben sein; um diese zu erreichen, folgte die genaue Taktung der Einladungen und Erinnerungen, wie oben skizziert, einem strengen, automatischen Algorithmus. Konnte damit tatsächlich ein gleichmäßiges, strukturgleiches Bild auch bei tageweiser Betrachtung erreicht werden?

Werfen wir dazu zunächst einen Blick auf die pro Feldtag resultierenden Fallzahlen. Den allerersten Befragungstag und den Tag der Wahl ausgeblendet¹⁵, wurden im Durchschnitt 149 Interviews realisiert. Die Standardabweichung betrug

15 Am letzten Befragungstag, dem Wahltag, wurden keine neuen Personen mehr zur Befragung eingeladen und die noch geöffneten Einladungspakete wurden kurz nach Öffnung der Wahllokale vormittags geschlossen. Die ersten Befragungstage werden bei der Untersuchung von rollierenden Querschnitten üblicherweise ausgeschlossen, da das RCS-Design einige Tage Anlauf benötigt, bis sich die gewünschten Fallzahlen und die gewünschte Mischung aus leicht und schwer erreichbaren Personen tatsächlich einstellt. Üblicherweise wird hierfür eine Zeit von fünf Tagen einkalkuliert. Im Falle der Wahlstudie waren die Fallzahlen jedoch bereits ab dem zweiten Tag ausreichend hoch, um sie in die Analyse einzubeziehen, auch wenn die optimale Durchmischung zu diesem Zeitpunkt noch nicht gegeben war (siehe auch Abbildung 5).

16,8 Interviews. Die geringste Anzahl realisierter Interviews wurde am Samstag, den 5. März, mit 116, die größte am Montag, den 14. März, mit 186 erreicht, wie Abbildung 4 zeigt.¹⁶ Insgesamt stellt sich, was die resultierenden Fallzahlen betrifft, ein einigermaßen gleichmäßiger Feldverlauf ein, der in seiner Form und „Zackigkeit“ durchaus den Verläufen reiner RCS-Studien entspricht.

Abbildung 4 Entwicklung der Teilnehmerzahlen

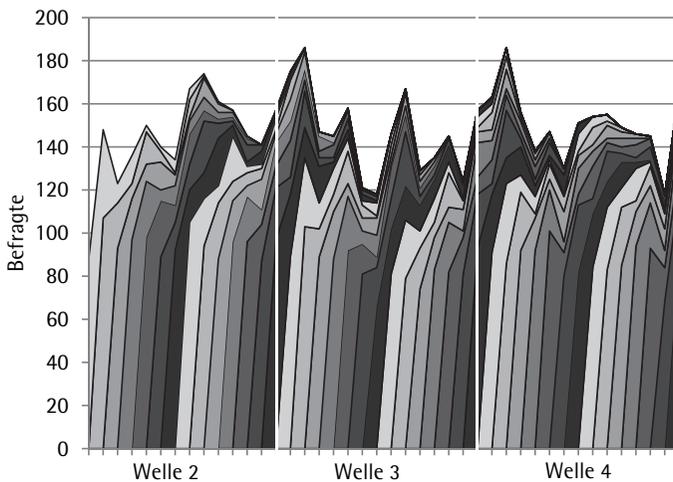


Allerdings ist die Zahl der Interviews pro Tag nur ein Indikator für einen gleichmäßigen Feldverlauf; darüber hinaus kommt es auf eine gute Mischung von leicht und schwer erreichbaren Personen in den einzelnen Tagesstichproben an, die in ihrer Struktur identisch sein sollen. Wie Abbildung 5 zeigt, greift tatsächlich für das rollierende Panel die für eine Rolling Cross-Section typische Logik, die Abbildung 1 idealtypisch aufzeigt. An jedem einzelnen Tag werden sowohl leicht als auch schwer erreichbare Personen zur Umfrage eingeladen. Während viele der leicht erreichbaren Personen direkt am Tag der Einladung an der Umfrage teilnehmen, braucht es bei den eher schwer erreichbaren Personen mitunter mehrere Kontaktversuche (Erinnerungen), bis sie an der Umfrage teilnehmen. Jedenfalls nehmen nicht alle

16 Dass der beste Tag auf einen Montag und der schlechteste auf einen Samstag fällt, ist durchaus typisch: Während des gesamten Befragungszeitraums wurden an Wochenenden geringere Tagesfallzahlen erreicht als an Werktagen. Dies ist unter anderem darauf zurückzuführen, dass ein Teil der Personen am Wochenende weniger häufig E-Mails abrufen. Auf die Frage „Wenn Sie jetzt einmal speziell an Wochenenden denken, wie häufig rufen Sie dann – im Vergleich zu Werktagen – Ihre E-Mails ab?“ antworteten 12 Prozent der Personen mit „häufiger“, 59 Prozent mit „genauso häufig“ und 28 Prozent mit „weniger häufig“.

Personen eines Pakets am gleichen Tag an der Umfrage teil, sondern verteilen sich auf den Einladungstag sowie die Folgetage. Dies führt zu Beginn der rollierenden Phase zu einer kurzen Anlaufphase der Studie. Schon nach wenigen Tagen stellt sich für die Tagesstichproben eine gute Mischung von leicht und schwer erreichbaren Personen ein (Abbildung 5), die dem Augenschein nach sehr konstant über den Feldverlauf hinweg bleibt. Dies gilt gerade auch für die kritischen Übergänge zwischen den einzelnen Panelwellen, die in Abbildung 5 durch vertikale weiße Linien markiert sind.

Abbildung 5 Struktur der Tagesstichproben

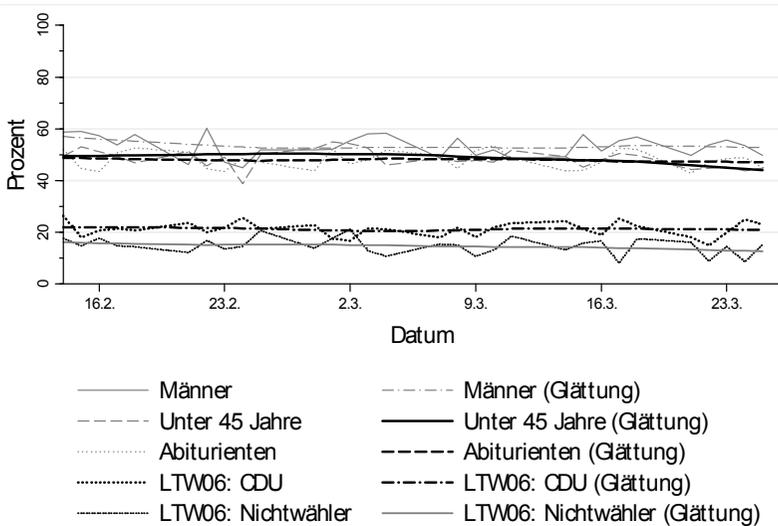


Strukturgleiche Tagesstichproben sind eine notwendige Voraussetzung für weitergehende inhaltliche Analysen. Nur wenn tatsächlich sichergestellt ist, dass sich diese Struktur nicht ändert, können beobachtbare Veränderungen von politischen Einstellungen und Verhaltensabsichten tatsächlich als Konsequenz des Wahlkampfes und seiner Ereignisse interpretiert werden. Abbildung 5 legt nahe, dass diese notwendige Bedingung tatsächlich erfüllt ist; aber wir können dies noch stringenter überprüfen.

Wenn die einzelnen Tagesstichproben strukturgleich sind, sollten stabile Merkmale von Befragten (zum Beispiel ihr Alter oder ihr Geschlecht) in den einzelnen Tagesstichproben – abgesehen von zufälligen Schwankungen – identisch verteilt sein. Zur Überprüfung der Datenqualität unter diesen Gesichtspunkten können wir daher entsprechende Trendlinien für Verteilungen ausgewählter stabiler Merkmale betrachten. Die Linien sollten flach sein. Da für die einzelnen Tage nur rela-

tiv geringe Fallzahlen (von rund 150 Interviews) zur Verfügung stehen, wollen wir nicht bloß tageweise Mittelwerte betrachten, sondern auch geglättete Trendlinien. Wir haben dazu das LOWESS-Verfahren (mit einer Bandbreite von 0,5) verwendet; dabei wird jeder Punkt durch ein polynomiales Regressionsmodell an eng benachbarte Punkte angepasst.

Abbildung 6 Verteilungsstabilität ausgewählter soziodemographischer Merkmale (Bandbreite 0,5)



Die Ergebnisse dieser Analysen zeigt Abbildung 6. Auch wenn die Trendlinien nicht in allen Fällen gänzlich flach sind, so bleiben alles in allem die Verteilungen des Lebensalters, des Geschlechts, des höchsten Schulabschlusses und der Parteiwahl bei der vorhergehenden Landtagswahl im Feldverlauf sehr stabil. Systematische Trends lassen sich kaum ausmachen.

Diese visuellen Analysen lassen sich noch weiter statistisch unterfüttern, indem – in Anlehnung an Schmitt-Beck/Faas/Wolsing (2010) – aus allen verfügbaren Tagesstichproben jeweils zwei herausgegriffen und die jeweils interessierenden Anteile (mit Hilfe logistischer Regressionsmodelle) miteinander verglichen werden. Vergleicht man auf diese Weise alle möglichen Paare von Tagesstichproben (konkret: die sich daraus jeweils ergebenden Anteile der skizzierten Variablen), so sollten diese sich nur zufällig unterscheiden. Bei einem angelegten statistischen Signifikanzniveau von fünf Prozent würde man rein zufallsbedingt in bis zu fünf Prozent aller Kombinationen von Feldtagen Abweichungen erwarten, aber nicht mehr.

Tabelle 5 listet die tatsächlich beobachtbaren Abweichungsanteile basierend auf allen möglichen Zweierkombinationen der Tagesstichproben auf. Die Ergebnisse zeigen, dass sich nur in einem Fall – nämlich dem Anteil der Nichtwähler bei der Landtagswahl 2006 – deutlich überzufällige Abweichungen ergeben; das deckt sich mit den Analysen der Panelmortalität oben. Nichtwähler sind uns im Verlauf der Vorwahlbeobachtung in etwas überdurchschnittlichem Maße „abhanden“ gekommen. Auch für das Geschlecht ergibt sich ein zu hoher Wert an Abweichungen, der allerdings mit 5,1 Prozent nur minimal erhöht ist.

Tabelle 5 Anteile signifikanter Abweichungen im Paarvergleich von Tagesstichproben

Prozentualer Anteil Abweichungen im Tagespaarvergleich ($p > 0,05$)	
Geschlecht	5,1
Alter	2,6
Abiturienten	2,4
LTW06: CDU	1,8
LTW06: Nichtwähler	6,7
Anzahl Paare	820

Der Großteil der Abweichungen ist demnach auf rein zufällige Streuungen zurückzuführen. Die für eine RCS-Studie notwendigen Voraussetzungen konnten alles in allem erfüllt werden, auch in der Variante der rollierenden Panelstudie.

4 Analysepotenziale

Wie bereits ausgeführt, erweitert das rollierende Panel die Analysemöglichkeiten der verknüpften Designs um die Vorzüge des jeweils anderen. So können nicht nur Trends und Dynamiken tagesgenau nachgezeichnet werden (RCS), es ist auch feststellbar, auf welchen individuellen Faktoren diese beruhen (Panel). Um dies zu verdeutlichen, möchten wir abschließend an einem Beispiel das Analysepotenzial der Daten aufzeigen. Wir greifen dazu auf ein Ereignis zurück, das den Wahlkampf und die Wahl in Baden-Württemberg nachhaltig beeinflusst hat – nämlich das Atomunglück von Fukushima, das sich zwei Wochen vor dem Wahltag ereignet hat.

Grundsätzlich ist die Atomkraft seit vielen Jahren eines der wichtigsten Positionsissues im politischen Wettbewerb in Deutschland. In der Folge weist das Thema bei Parteien und ihren Anhängern ein hohes Maß an Stabilität auf (Thur-

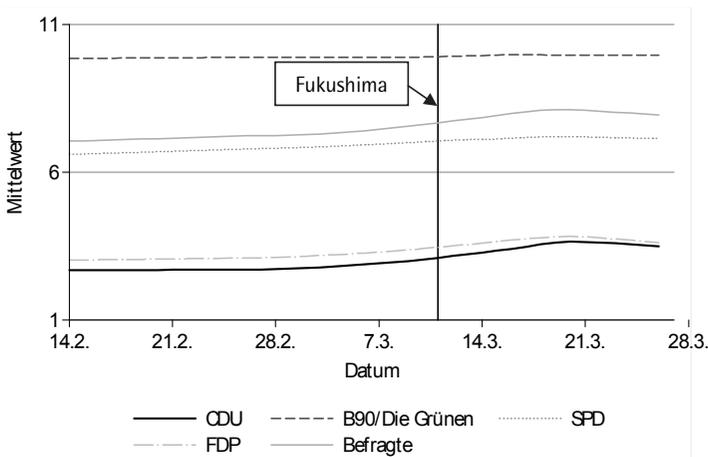
ner 2010: 338). Während der Feldzeit der „Wahlstudie Baden-Württemberg 2011“ kam es jedoch zu eben jenem desaströsen Ereignis in Japan, das an dieser Stabilität gerüttelt hat. In Abbildung 7 sind die mit dem LOWESS-Verfahren geglätteten Trendlinien zur Frage

„Sollte die Kernenergie weiter ausgebaut oder sollten alle Kernkraftwerke sofort abgeschaltet werden? Wie stehen Ihrer Meinung nach die Parteien dazu, wenn Sie eine Skala von 1 bis 11 verwenden?“

dargestellt. Die Skala reichte dabei von 1 „Weiterer Ausbau der Kernenergie“ bis 11 „Sofortige Abschaltung aller Kernkraftwerke“. Bis zum Zeitpunkt des Unglücks, in der Abbildung dargestellt als senkrechte durchgezogene Linie, gab es, wie vermutet, nur sehr wenige Veränderungen im Zeitverlauf. B90/Die Grünen wurden klar in der Nähe des sofortigen Ausstiegs eingeordnet, die Position der SPD war gemäßigter und CDU sowie FDP befanden sich in den Augen der Wähler relativ nahe beim Skalenspunkt „Weiterer Ausbau der Kernenergie“. Die Position der Befragten lag im Mittel zwischen der Position der SPD und jener der Grünen – allerdings bei erheblicher interindividueller Varianz.

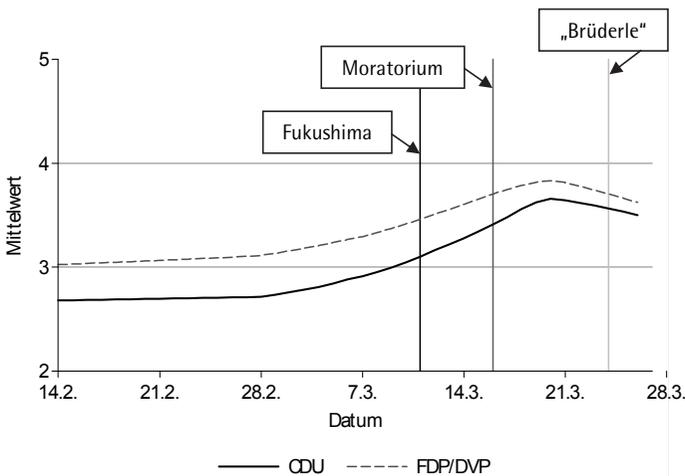
Mit dem Unglück veränderten sich diese Positionen jedoch. Insbesondere die Befragten näherten sich im Mittel noch weiter der Position der Grünen an. Doch auch die wahrgenommenen Positionen der CDU und der FDP veränderten sich merklich.

Abbildung 7 Wahrgenommene Parteipositionen und eigene Einstellung zur Atomkraft (LOWESS-Kurven; Bandbreite 0,5)



Verfeinert man die Skala (Abbildung 8), werden diese Veränderungen noch deutlicher sichtbar. Daneben fällt auf, dass die Veränderung bei Union und FDP ab etwa dem 19. März wieder zum Erliegen kommt, tendenziell sogar wieder zurückgeht. Die weiter rechts eingezeichneten Linien geben Aufschluss über mögliche Gründe. Am 15. März verkündete Angela Merkel, dass es ein Moratorium, also die Aussetzung der Laufzeitverlängerung deutscher Atomkraftwerke, geben würde. Kurz nach diesem Zeitpunkt erreichen die (geglätteten) Mittelwerte für Union und FDP ihre Höchstwerte. Bereits eine Woche später jedoch ist eine Stabilisierung der Position zu erkennen. Als am 23. März (Linie 3) zudem bekannt wurde, dass Rainer Brüderle vor Industrievertretern verlauten ließ, es handle sich bei dem Moratorium lediglich um eine Wahlkampf-Taktik, dürfte dies die Trendumkehr weiter verstärkt haben.

Abbildung 8 Parteipositionen zur Atomkraft – Details (LOWESS-Kurven; Bandbreite 0,5)



Offen bleibt bis zu diesem Punkt allerdings die Frage, wessen individuelle Positionen und Wahrnehmungen für die Veränderungen im Aggregat verantwortlich sind. Mittels der Paneldaten kann nun explizit überprüft werden, wer diese Personen waren. Hierzu werden zunächst die Differenzen zwischen den Antworten der Befragungswellen zwei und drei¹⁷ – also vor der Atomkatastrophe – auf der einen

17 Die Atomkatastrophe ereignete sich am vorletzten Befragungstag der Befragungswelle 3. Für diejenigen Personen, die an diesem Tag oder danach in Welle 3 befragt wurden, wurde die

Seite, Befragungswelle vier auf der anderen Seite berechnet. Insgesamt gaben 47,9 Prozent der Befragten nach Fukushima einen höheren (und damit atomkritischeren) Wert an; 36,8 Prozent nannten den gleichen Wert wie zuvor. Im Mittel veränderte sich die Position um +0,8 Punkte auf der elfstufigen Skala.

Um nun herauszufinden, mit welchen individuellen Faktoren diese Veränderungen zusammenhängen, schätzen wir explorativ eine OLS-Regression, bei der die Differenz zwischen den Positionen zwischen den beiden Zeitpunkten als abhängige Variable dient. Als unabhängige Variablen werden die soziodemographischen Variablen Alter, Geschlecht und Bildung aufgenommen. Hinzu kommt politisches Interesse als Prädiktor für Informiertheit, sowie die eigene ideologische Position, gemessen anhand der Links-Rechts-Selbsteinstufung (Tabelle 6).

Tabelle 6 Einflussfaktoren auf Veränderung der Haltung zu Atomkraft nach Fukushima (OLS-Regression)

	Größe der Veränderung	Standardfehler
Alter	0,010*	0,005
Frau	-0,120	0,140
Politisches Interesse	-0,207	0,203
Interesse am Wahlkampf	-0,008	0,150
Niedrige Bildung	-0,090	0,088
Hohe Bildung	-0,120	0,072
Links-Rechts-Selbsteinstufung	0,181***	0,032
<i>N</i>	992	
Adjusted R ²	0,036	

*Unstandardisierte Effektkoeffizienten; Signifikanzniveaus: * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$*

Wie die Ergebnisse zeigen, sind das Alter und die Links-Rechts-Selbsteinstufung signifikante Prädiktoren individueller Positionsveränderungen. Ältere Personen ebenso wie Personen, die sich politisch weiter rechts einstufen, sind atomkritischer geworden.

Die Mischung von Rolling Cross-Section und Panel ermöglicht es also, mehr Einblicke in derartige Prozesse zu bekommen, als es bei der Verwendung eines der Verfahren alleine möglich gewesen wäre. Auch wenn wir mit der Atomkraft und der Atomkatastrophe ein besonders prominentes Beispiel herausgegriffen haben, so ändert das nichts am Potenzial der Studie auch für andere Fragestellungen. Es ließe sich ebenso im Zusammenhang mit dem Ansehensverlust von Stefan Mappus, den

Einstellung in Welle 2 als Vergleichswert herangezogen.

Themenkarrieren (beispielsweise der Rolle von Stuttgart 21 während des gleichen Wahlkampfes) oder der Rolle von Emotionen während des Wahlkampfes anwenden – all dies haben wir an anderer Stelle bereits zeigen können (vgl. Blumenberg/Faas 2012a, 2012b).

5 Fazit

Ziel des Beitrags war es, die Theorie und Praxis eines rollierenden Panels vorzustellen, dessen Durchführung anhand des Beispiels „Wahlstudie Baden-Württemberg 2011“ zu beschreiben und das damit verbundene Analysepotenzial aufzuzeigen. Das rollierende Panel soll das Beste aus zwei Welten (Panel- und RCS-Studien) verbinden. Wie beim RCS-Design soll es die tagesgenaue Beobachtung von Änderungen in den Einstellungen und Verhaltensmustern der Bevölkerung abbilden, ohne dabei jedoch die Möglichkeit zu verlieren, auch „hinter die Kulissen“ zu schauen und die individuellen Gründe und Prädiktoren dieser Veränderungen zu analysieren.

Doch nützt das beste theoretische Instrument wenig, wenn es sich in der Praxis nicht bewährt. Zentral sind dabei im konkreten Fall Feldsteuerung und Feldverlauf. Weder darf es zu hoher Panelmortalität und Panelkonditionierung (aus der Perspektive der Panelkomponente) kommen, noch dürfen sich die einzelnen Tagesstichproben (aus der RCS-Perspektive) systematisch unterscheiden. Beides konnte im Fall der „Wahlstudie Baden-Württemberg 2011“ ohne merkliche Abstriche erreicht werden. Dabei hat sich vor allem die Erhebung einer Basismessung vor der ersten rollierenden Welle als günstig erwiesen: Panelmortalität hat sich so vor allem zwischen Basismessung und rollierender Phase eingestellt – berührte aber damit die eigentliche Vorwahlphase mit ihrer dynamischen Logik nicht.

Auch auf der inhaltlichen Ebene hat das neue Instrument die erwarteten Ergebnisse geliefert: Gerade vor dem Hintergrund unserer konkreten Ergebnisse aus dem Kontext der Landtagswahl in Baden-Württemberg – exemplarisch konnten wir dies am Beispiel der Veränderungen der Einstellungen zur Atomkraft in Deutschland in Folge der Atomkatastrophe in Japan zeigen – wurde deutlich, dass es einen wahren Mehrwert gegenüber den klassischen Instrumenten liefert, ohne zugleich nennenswerte Nachteile mitzubringen.

Insgesamt wäre es zu begrüßen, wenn zukünftig auch in anderen Studien über eine Integration von Panel- und RCS-Komponenten nachgedacht würde, anstatt sie als separate Komponenten zu realisieren. Die Vorteile der Integration und die Machbarkeit des Unterfangens konnten anhand der „Wahlstudie Baden-

Württemberg 2011“ demonstriert werden: Das Ganze ist mehr als die Summe seiner Teile.

Literatur

- Blumenberg, J. N. und T. Faas, 2012a: Abstimmung gut, alles gut? Der Bürger im Staat. 3-2012: 182-187.
- Blumenberg, J. N. und T. Faas, 2012b: Stuttgart 21: Einstellungen und Emotionen. in: U. Wagschal (Hg.): Machtwechsel oder Politikwechsel – Eine Bilanz der Landtagswahl 2011 in Baden-Württemberg. [im Erscheinen]
- Chang, L. und J. A. Krosnick, 2009: National Surveys Via Rdd Telephone Interviewing Versus the Internet. Comparing Sample Representativeness and Response Quality. *Public Opinion Quarterly* 73 (4): 641-678.
- Das, M., V. Toepoel und A. van Soest, 2011: Nonparametric Tests of Panel Conditioning and Attrition Bias in Panel Surveys. *Sociological Methods & Research* 40 (1): 32-56.
- Faas, T. und J. Maier, 2011: Medienwahlkampf: Sind TV-Duelle nur Show und damit nutzlos? S. 99-114 in: E. Bytzek und S. Roßteutscher (Hg.): Der unbekannte Wähler?: Mythen und Fakten über das Wahlverhalten der Deutschen. Frankfurt a. M.: Campus.
- 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. *Journal of Official Statistics* 22 (2): 177-190.
- Johnston, R., 2001: Capturing Campaigns in National Election Studies. S. 149-172 in: E. Katz und Y. Warshel (Hg.): Election studies. What's their use? Boulder, Colo: Westview Press.
- Johnston, R. und H. E. Brady, 2002: The rolling cross-section design. *Electoral Studies* 21: 283-295.
- Kaspar, H., 2009: Panelpflege, Panelmortalität und Konvertierung im Panel. S. 85-109 in: H. Schoen, H. Rattinger und O. W. Gabriel (Hg.): Vom Interview zur Analyse. Methodische Aspekte der Einstellungs- und Wahlforschung. Baden-Baden: Nomos.
- Lazarsfeld, P. F., B. Berelson und H. Gaudet, 1968: *The People's Choice. How the Voter Makes Up his Mind in a Presidential Campaign.* New York: Columbia Univ. Press.
- Lenz, G., 2009: Learning and Opinion Change, Not priming: Reconsidering the Evidence for the Priming Hypothesis. *American Journal of Political Science* 53 (4): 821-837.
- Romer, D., K. Kenski, K. Winne, C. Adasiewicz und K. H. Jamieson, 2006: Capturing campaign dynamics, 2000 and 2004. *The national Annenberg election survey.* Philadelphia: University of Pennsylvania.
- Sanders, D., H. D. Clarke, M. C. Stewart und P. Whiteley, 2007: Does Mode Matter For Modeling Political Choice? Evidence From The 2005 British Election Study. *Political Analysis* 15 (3): 257-285.
- Schmitt-Beck, R., T. Faas und C. Holst, 2006: Der Rolling Cross-Section Survey – ein Instrument zur Analyse dynamischer Prozesse der Einstellungsentwicklung: Bericht zur ersten deutschen RCS-Studie anlässlich der Bundestagswahl 2005. *ZUMA-Nachrichten* (58): 13-49.
- Schmitt-Beck, R., T. Faas und A. Wolsing, 2010: Kampagnendynamik bei der Bundestagswahl 2009. Arbeitspapiere – Mannheimer Zentrum für Europäische Sozialforschung (134).
- Schonlau, M., R. D. Fricker Jr. und M. N. Elliott, 2002. *Conducting Research Surveys via E-mail and the Web.* Santa Monica, CA: RAND.
- Sobol, M. G., 1959: Panel Mortality and Panel Bias. *Journal of the American Statistical Association* 54 (285): 52-68.
- Stadtmüller, S., 2009: Weniger Befragte, schlechtere Ergebnisse? Die Wirkungen der Panelmortalität. S. 111-130 in: H. Schoen, H. Rattinger und O. W. Gabriel (Hg.): Vom Interview

- zur Analyse. Methodische Aspekte der Einstellungs- und Wahlforschung. Baden-Baden: Nomos.
- Turner, P. W., 2010: „Issue-Unentschiedene“ und „Issue-Inkonsistente“ als Targetpopulationen? Das Beispiel Kernenergie (1987-2005). S. 333-351 in: T. Faas, K. Arzheimer, S. Roßteutscher (Hg.): Information Wahrnehmung Emotion: Politische Psychologie in der Wahl- und Einstellungsforschung. Wiesbaden: VS Verlag.
- Toepoel, V., M. Das und A. van Soest, 2009: Relating Question Type to Panel Conditioning: Comparing Trained and Fresh Respondents. *Survey Research Methods* 3 (2): 73-80.

Anschrift der Autoren

Thorsten Faas / Johannes N. Blumenberg
Institut für Politikwissenschaft
Bereich „Methoden der empirischen
Politikforschung“
Johannes Gutenberg-Universität Mainz
Hegelstraße 59
55099 Mainz
E-Mail: thorsten.faas@uni-mainz.de
blumenberg@politik.uni-mainz.de

