

## Innovationen im Rolling-Cross-Section-Design

### Neue Ansätze zur Analyse des Wahlverhaltens

#### Kurzfassung

Die Wahrnehmungen und Einstellungen von Wählern werden in einer sich immer schneller wandelnden Zeit stärker durch mediale Ereignisse geprägt als jemals zuvor. Unser Erhebungsansatz versucht, diesen Herausforderungen der modernen Wahlforschung zu begegnen. Basierend auf dem innovativen Rolling-Cross-Section-Design (RCS) befragten wir mit dem Ipsos Demokratiepulss ab dem 23.4.2013 in einer Kombination aus Festnetz- und Mobilfunkstichprobe täglich 50 Wahlberechtigte zur politischen Stimmung in Deutschland (Wahlabsicht, Kanzlerpräferenz etc.). Durch die zusätzliche Erhebung über Mobilfunk gelingt dabei ein besseres Abbild der Grundgesamtheit, da auch die *Mobile-Only* und die in der Wissenschaft bisher wenig beachteten *Mostly-Mobile* entsprechend erreicht werden.

Höhere Repräsentativität wird ferner durch die Anwendung des RCS-Designs erzeugt, wodurch selbst schwer erreichbare Personen ausreichend stark in der Stichprobe vertreten sind. Designbedingt lassen sich die Effekte von Medienereignissen auf Tagesbasis dynamisch untersuchen. So hatte die Forderung Sigmar Gabriels zur Einführung eines Tempolimits auf Autobahnen im Mai 2013 keine deutlichen Veränderungen in der Wahlabsicht bewirkt. Lediglich die Piratenpartei konnte hiervon signifikant profitieren. Einziger Nutznießer des gescheiterten Drohnen-Projekts Euro Hawk waren die Grünen.

Die Güte des RCS-Designs kann mit einer kontinuierlichen Abarbeitung in der Feldphase maximiert und durch eine innere bzw. äußere RCS-Effizienz exakt quantifiziert werden. Die Umsetzung konstanter Fallzahlen pro Tag mindert die innere RCS-Effizienz und benötigt eine Aggregation auf Wochenbasis. Durch das rollierende Design sinkt die Varianz der Inklusionswahrscheinlichkeiten und die zu analysierende Stichprobe hat eine deutlich höhere Qualität.

## Inhalt

1. Einleitung	214
2. Theoretische Grundlagen	215
a) Allgemeine Stichprobentheorie	215
b) Kombination von Festnetz- und Mobilfunkstichproben	216
c) Das Rolling-Cross-Section Design	217
d) Medien und Wahlverhalten	220
3. Methodische Grundlagen	221
a) Vorstellung der Dual-Frame-Gewichtung	221
b) Definition von Effizienzmaßen in einem RCS-Design	222
4. Methodische Analysen	226
a) Mobile-Only und Mostly-Mobile im Vergleich	226
b) RCS-spezifische Analysen	229
5. Inhaltliche Analysen	233
a) Interpretation verschiedener politischer Zeitreihen	233
b) Test der Ereignisse auf signifikante Effekte	235
6. Fazit	237

## 1. Einleitung

Nachdem führende Meinungsforschungsinstitute bei der niedersächsischen Landtagswahl 2013 fälschlicherweise ein knappes Scheitern der FDP an der Fünf-Prozent-Hürde prognostizierten, titelte Die Welt „So verschliefen die Demoskopon den Trendwechsel.“ Bedenken über die Qualität von Meinungsumfragen sind so alt wie das Metier selbst, aber die Zeiten sind schwieriger geworden für die politische Meinungsforschung. Das Forschungsfeld ist in den letzten Jahren durch zwei Entwicklungen geprägt worden, die eng miteinander verknüpft sind. Zum einen werden politische Einstellungen immer stärker durch tagesaktuelle Ereignisse geprägt. Nirgendwo wird dies so deutlich wie bei Landtags- oder Bundestagswahlen. Wähler mit tief verwurzelten Bindungen an eine politische Partei werden immer seltener (Schoen/Weins 2005). Gleichzeitig wächst der Anteil von Wählern, die ihre Entscheidung erst kurz vor der Wahl treffen und ihre Stimme nach taktischen Erwägungen vergeben (Schmitt-Beck 2009). Zum anderen hat der Einfluss von politischen Kampagnen und der Medienberichterstattung in dem Maße zugenommen, wie politische Einstellungen an Stabilität verlieren. Dabei wurde deutlich, dass massenmediale Ereignisse vor allem indirekt wirken, indem sie Wähler informieren,

mobilisieren und die wahrgenommene Wichtigkeit politischer Probleme steuern (Brettschneider 2005). Mit der zunehmenden Dynamik politischer Einstellungen und Wahlentscheidungen stoßen traditionelle Meinungsumfragen allerdings an ihre Grenzen, da sie lediglich eine Momentaufnahme der politischen Stimmung zu einem bestimmten Zeitpunkt bieten. Dadurch fehlt die Basis, um kurzfristige Veränderungen der öffentlichen Meinung konkreten Politikentscheidungen und Ereignissen zuzuschreiben und langfristige Trends zu erkennen.

Allerdings haben sich nicht nur die Rahmenbedingungen politischer Meinungsbildung gewandelt, sondern herkömmliche Meinungsumfragen geraten auch verstärkt durch das veränderte Kommunikationsverhalten der Wähler unter Druck. Die zunehmende Verbreitung von Mobilfunktelefonen stellt klassische Telefonbefragungen vor erhebliche Herausforderungen. Die Anzahl von Festnetzanschlüssen ist rückläufig und neueste Studien schätzen den Anteil von Personen, die nur noch über ihr Mobiltelefon erreichbar sind (*Mobile-Only*), auf rund 12 Prozent der deutschen Bevölkerung (ADM 2012). Und auch unter denjenigen, die noch über einen Festnetzanschluss verfügen, gibt es inzwischen zahlreiche Personen, die hauptsächlich über ihr Mobiltelefon erreichbar sind (*Mostly-Mobile*) (Häder/Häder/Kühne 2012). Die Schwierigkeit besteht darin, dass sich exklusive Mobilfunknutzer systematisch von Personen mit reinem Festnetzanschluss unterscheiden. Reine Mobilfunknutzer sind tendenziell jünger, vorwiegend männlich und insbesondere in den neuen Bundesländern stark vertreten.

Unsere Erhebung begegnet diesen Herausforderungen; wir kontaktieren die wahlberechtigte Bevölkerung Deutschlands im Rolling-Cross-Section-Design auf Basis einer *Dual-Frame*-Stichprobe und befragen sie zu politischen Themen. Durch die tägliche Generierung von etwa 50 Interviews messen wir kurzfristige Veränderungen und erreichen designbedingt Zielgruppen, die über herkömmliche Festnetzstichproben nicht erreicht werden können.

## 2. Theoretische Grundlagen

### a) Allgemeine Stichprobentheorie

Stichproben sollten in der Regel so gebildet werden, dass die Auswahl der Untersuchungsobjekte aus der Grundgesamtheit zufällig erfolgt. In einem maximal repräsentativen Stichprobendesign besteht der Stichprobenraum<sup>1</sup> nur aus sämtlichen

1 Ein Stichprobenraum beinhaltet alle Elemente der Grundgesamtheit, die theoretisch in die Stichprobe gelangen könnten. Hierbei unterscheidet er sich von der Auswahlgesamtheit, die auch Objekte beinhalten kann, die nicht zur Grundgesamtheit gehören (Oversampling).

Elementen der Grundgesamtheit. Die Güte des Stichprobenraums kann dabei quantitativ durch die Mächtigkeit oder Kardinalität beschrieben werden. Der Raum besitzt demnach eine große Mächtigkeit, wenn fast alle Elemente der Grundgesamtheit auch Bestandteil des Stichprobenraumes sind (Hoops/Schnapp 2013).

Außerdem ist von großer Bedeutung, dass die Objekte des Stichprobenraums eine möglichst identische Inklusionswahrscheinlichkeit in die Stichprobe besitzen. Bereits durch die variierende Haushaltsgröße ist dies in der Realität meistens nicht der Fall. Trotzdem ist es das Ziel einer Erhebung mit höchstem Anspruch an Repräsentativität, eben diese Varianz der Auswahlwahrscheinlichkeiten auf ein Minimum zu beschränken. Nur dann ermöglichen die Analyseergebnisse ein möglichst unverzerrtes Abbild der Grundgesamtheit.

In der aktuellen Forschung gibt es verschiedene Designs und Techniken, welche die Mächtigkeit des Stichprobenraums erhöhen beziehungsweise die Varianz der Inklusionswahrscheinlichkeiten in die Stichprobe minimieren. Wir erreichen eine höhere Mächtigkeit des Stichprobenraums durch die Berücksichtigung von Mobilstichproben. Zudem werden die Auswahlwahrscheinlichkeiten der *Hard-to-Reach*-Personen in die Stichprobe durch das rollierende Design erhöht, was zu einer Reduktion der Varianz der Inklusionswahrscheinlichkeiten und damit geringeren Verzerrungen führt.

## **b) Kombination von Festnetz- und Mobilfunkstichproben**

Stichproben, die auf das Festnetz (kurz: FN) beschränkt sind, erzeugen meist kein repräsentatives Abbild der Grundgesamtheit, da es Personen gibt, die zur Grundgesamtheit gehören, aber über keinen Festnetzanschluss verfügen und somit nicht die Möglichkeit haben, in die Stichprobe zu gelangen. Ist keine repräsentative Projektion möglich, können Analysen nur mit einer sehr geringen Güte geleistet werden. Da ein erheblicher Teil der Wahlberechtigten entweder gänzlich oder überwiegend auf dem Handy erreichbar ist, werden in der Forschung heutzutage Zufallsbefragungen über Festnetz mit einer Stichprobe über Mobilfunk (kurz: MO) ergänzt, damit auch die sogenannten *Mostly-Mobile* beziehungsweise *Mobile-Only* ausreichend beziehungsweise überhaupt in die Stichprobe gelangen (Hoffmann 2007). Dieser Ansatz wird in der Forschung *Dual-Frame* genannt. Die Anwendung dieses Ansatzes führt zu einer höheren Repräsentativität.

Es gibt eine Vielzahl an Publikationen, die sich mit den *Mobile-Only* beschäftigen. So ist bekannt, dass der *Mobile-Only* übermäßig oft männlich, jung (Graeske/Kunz 2009: 59 f.), unverheiratet, schlecht gebildet und Geringverdiener ist (Keeter 2006: 91 f.). Zudem bilden vor allem Migranten (Pennay 2010: 7), Studenten und

Arbeitslose sowie kaum Rentner diesen exklusiven Handynutzer (Link u. a. 2007: 828 f.). In Deutschland gibt es zudem vor allem in den neuen Bundesländern viele *Mobile-Only* (Hunsicker/Schroth 2007: 175).

Auch im Hinblick auf die politische Partizipation unterscheiden sich *Mobile-Only* massiv von Personen mit Festnetzanschluss. Laut einer Befragung zum amerikanischen Wahlkampf 2004 haben die *Mobile-Only* häufig den letztendlich unterlegenen John Kerry gewählt und hatten zudem übermäßig oft eine liberale Weltanschauung. Als wichtigste Wahlkampfthemen wurden häufiger Steuern und Arbeit genannt, dagegen nur selten Terrorismus und die Situation im Irak (Keeter 2006: 94 ff.). *Mobile-Only* in einer Erhebung in Australien sehen übermäßig oft Arbeit, Klimaveränderung und Immigration als das wichtigste Problem im Land. Die Gesundheit hingegen erhält eine niedrigere Wichtigkeit (Pennay 2010: 10). Zudem hat der *Mobile-Only* ein geringeres politisches Interesse (vgl. Hunsicker/Schroth 2007: 175) und verfolgt weniger aufmerksam Wahlen. Er ist häufiger zu beschäftigt, um zu wählen, und glaubt zudem, dass seine Wahl sowieso nichts bewirkt. Außerdem ist der exklusive Handynutzer übermäßig oft Erstwähler (Keeter u. a. 2007: 779).

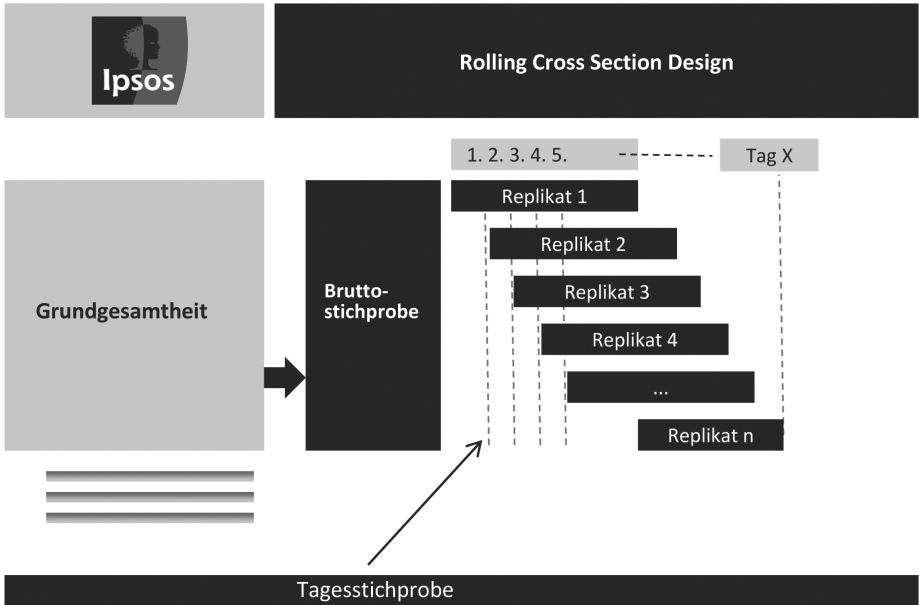
Über den *Mostly-Mobile* ist hingegen wenig bekannt. Es gibt hierfür auch keine etablierten Skalen, um diesen Dual-User entsprechend zu identifizieren.

### c) Das Rolling-Cross-Section Design

Die Grundidee des RCS-Designs ist einfach (siehe Abbildung 1): Eine normale Zufallsstichprobe wird nach strengen methodischen Anforderungen so über einen festgelegten Zeitraum verteilt, dass die Befragten jedes einzelnen Tages für sich genommen eine repräsentative Stichprobe aus der Wählerschaft darstellen (Holst/Schmitt-Beck/Faas 2006). Um dies zu erreichen, wird die Ausgangstichprobe zufällig in Teilstichproben – sogenannte Replikate – zerlegt. Diese Replikate werden wiederum an einem zufällig bestimmten Erhebungstag ins Feld gegeben und nach einem zuvor festgelegten Kontaktschema abgearbeitet. Das bedeutet, dass für jedes Replikat die gleichen Regeln gelten, sowohl was die Anzahl und Zeiten der täglichen Kontaktaufnahme mit den Befragten betrifft als auch den Zeitraum, in dem das Replikat aktiv im Feld bleibt.

In der Praxis wird nicht jeder Befragte eines Replikats am ersten Tag der Aktivierung interviewt, sondern lediglich diejenigen Personen, die relativ leicht erreichbar sind (*Easy-to-Reach*). Interviews mit schwer erreichbaren Personen (*Hard-to-Reach*) werden meist erst nach mehreren Kontaktversuchen realisiert. Infolgedessen

Abbildung 1: Umsetzung einer RCS-Erhebung in der Praxis



Quelle: Holst/Schmitt-Beck/Faas 2006.

beginnen sich im Feldverlauf die Teilstichproben zu überlappen. Das heißt, die Stichprobe eines jeden Tages besteht aus Interviewten verschiedener Replikate. Charakteristisch für das RCS-Design ist daher ein sehr strenges Kontaktprotokoll, um sicherzustellen, dass sich jede Tagesstichprobe aus einem konstanten Anteil an leicht und schwer erreichbaren Befragten der einzelnen Replikate zusammensetzt.

Die hohe Dynamik und Flexibilität des RCS-Designs machen es zu einem der vielversprechendsten neuen Ansätze der politischen Meinungsforschung. Bisher kommt der Ansatz vor allem in der akademischen Wahlforschung zum Einsatz. Das RCS-Design in seiner heutigen Form wurde erstmals 1988 in der *Canadian Election Study* (CES) eingesetzt. Seine Wurzeln reichen aber bis zu einem Pilotprojekt der *American National Election Study* (ANES) zurück, in deren Rahmen zum ersten Mal eine rollierende Querschnittsbefragung auf Wochenbasis durchgeführt wurde (Johnston/Brady 2002; Kenski 2004 a). Nach diesem Startschuss wurde das RCS-Design schnell zu einem anerkannten Verfahren, das bis heute in vielen nationalen Wahlstudien einen angestammten Platz besitzt, etwa Neuseeland (*New Zealand Election Study*, NZES), Großbritannien (*British Election Study*, BES), Italien (*Italian National Election Study*, ITANES) und den USA (*National Annenberg Election*

*Study*, NAES). Im Rahmen der Bundestagswahl 2005 wurde von der Universität Mannheim in Zusammenarbeit mit Ipsos die erste deutsche Wahlstudie auf Basis des RCS-Designs umgesetzt (Schmitt-Beck/Faas/Holst 2006). Aufgrund ihres Erfolges wurde eine RCS-Umfrage ein fester Bestandteil der *German Longitudinal Election Study* (GLES), der bisher umfassendsten nationalen Wahlstudie in Deutschland (Schmitt-Beck u. a. 2010). Sowohl bei der Bundestagswahl 2009 als auch bei der kommenden Bundestagswahl im Herbst 2013 wurde Ipsos mit der Umsetzung beauftragt.

Im Gegensatz zu traditionellen Querschnitts- oder Trendumfragen bietet ein RCS-Ansatz eine Reihe von Vorteilen: Er ermöglicht ein tagesaktuelles Tracking der Veränderung von politischen Einstellungen. So können kurzfristige Schwankungen der öffentlichen Meinung ebenso wie langfristige Trends identifiziert werden. Außerdem erlaubt er, aktuelle politische Ereignisse unmittelbar mit Veränderungen der politischen Stimmung in Beziehung zu setzen. Das gilt ebenfalls für unvorhergesehene Ereignisse, die großen Einfluss auf die öffentliche Meinung haben, wie beispielsweise politische Skandale oder wirtschaftliche Krisen. Die Feldarbeit läuft kontinuierlich, wodurch sich auf aktuelle Ereignisse und Entwicklungen schnell und flexibel reagieren lässt, beispielsweise indem neue Fragen aufgenommen werden. Die Ergebnisse können bereits während der laufenden Feldzeit analysiert werden. Dabei sind die Daten beliebig kumulierbar, das heißt sie können auf jeder vom Forscher festgelegten Zeitbasis ausgewertet werden, auch wochen-, monats- oder sogar jahresweise. RCS-Erhebungen sind sehr anschlussfähig, da sie mit zahlreichen anderen Daten verknüpft werden können, beispielsweise Inhaltsanalysen von klassischen Print- oder Onlinemedien (*Social Listening*).

In einem guten RCS-Design sind die leicht erreichbaren Elemente der Grundgesamtheit nicht übermäßig oft in der Stichprobe vertreten, sondern ebenfalls auch ein hinreichend großer Anteil von *Hard-to-Reach*-Populationen (Hoops/Schnapp 2013). Ohne Anwendung eines solchen Designs haben leicht erreichbare Bevölkerungsgruppen (z. B. Rentner) eine höhere Inklusionswahrscheinlichkeit in die Stichprobe als schwer erreichbare Sub-Populationen (wie z. B. Geschäftsleute, die viel in Meetings und auf Reisen sind). Durch die Anwendung eines RCS-Designs wird somit die Varianz der Inklusionswahrscheinlichkeiten zugunsten der Repräsentativität minimiert. Damit kann das RCS-Design die Nachteile klassischer Erhebungsverfahren bei der Messung dynamischer Ereignisse kompensieren (Johnston/Brady 2002).

Neben diesen Vorteilen besitzen RCS-Designs allerdings den Nachteil, dass Veränderungen von politischen Einstellungen nur auf der aggregierten und nicht auf der individuellen Ebene analysiert werden können. Im Gegensatz zu Paneldesigns

können also lediglich Aussagen über Wandlungsprozesse in der gesamten Wählerschaft getroffen werden. Dieser Nachteil kann jedoch durch die Einführung einer zusätzlichen Panelkomponente behoben werden, wie dies beispielsweise bei der RCS-Studie 2009 der GLES realisiert wurde.

#### **d) Medien und Wahlverhalten**

Dem Einfluss der Massenmedien auf das Wahlverhalten wurde in der Wahlforschung lange Zeit kaum Beachtung geschenkt (siehe Brettschneider 2005). Während Belege für eine direkte Medienwirkung im Sinne einer Veränderung politischer Einstellungen tatsächlich selten sind, konnte die neuere Forschung zeigen, dass die Berichterstattung der Medien vor allem indirekt Wirkung entfaltet, indem sie die Relevanz von Themen- und Kandidatenorientierung beeinflussen kann.

Im Rahmen des sozialpsychologischen Modells des Wählerverhaltens (Campbell u. a. 1960; Schoen/Weins 2005) wird die Wahlentscheidung der Bürgerinnen und Bürger einerseits durch die affektive Bindung an eine politische Partei und andererseits durch kurzfristige Einstellungen gegenüber politischen Streitfragen und dem personellen Angebot der politischen Parteien, etwa den Spitzenkandidaten, bestimmt. Wie die Spitzenkandidaten bewertet werden und wie die Problemlösungskompetenz der Parteien beurteilt wird, hängt stark von der Bindung an eine Partei ab, die die Wahrnehmung wie ein Filter in Richtung der präferierten Partei verzerrt. Die Parteibindung wirkt hierbei als langfristig stabilisierender Faktor, der meist schon früh im Laufe der politischen Sozialisation erworben wird und kaum durch Medien veränderbar erscheint. Allerdings können Medien bestimmen, welche Themen und Kandidateneigenschaften als wichtig betrachtet werden. Diese Annahme basiert auf der Agenda-Setting-Funktion der Massenmedien (McCombs/Shaw 1972; Weaver/McCombs/Shaw 2004). Je nachdem, wie häufig ein bestimmtes politisches Problem in den Medien betont wird, umso mehr beeinflusst es auch die öffentliche Agenda. Auf der individuellen Ebene verändert sich dadurch die subjektive Wichtigkeit dieses Thema in der Bevölkerung durch den sogenannten Priming-Effekt (Miller/Krosnick 1996; für Deutschland siehe Schoen 2004). Dies bedeutet, dass Wähler bei ihrer Meinungsbildung auf Kriterien zurückgreifen, die momentan geistig präsent sind. Falls über ein Thema berichtet wird, bei dem eine bestimmte Partei einen Kompetenzvorteil hat, dann kann sich damit letztendlich auch die Wahlabsicht ändern. Wenn andererseits beispielsweise vermehrt über Koalitionsaussagen der Parteien berichtet wird, kann dies bei den Wählern taktische Überlegungen in den Vordergrund rücken. Falls diejenigen Aspekte, über die berichtet wird, eindeutig positiv oder negativ besetzt sind, hat dies auch – neben der



Relevanz – Auswirkungen auf die Bewertung von Themen bzw. Kandidaten, etwa durch eine skandalisierende Berichterstattung. Der RCS-Ansatz bietet hierbei ganz neue Möglichkeiten, solche Medien-Effekte innerhalb kurzer Intervalle zu identifizieren.

### 3. Methodische Grundlagen

#### a) Vorstellung der Dual-Frame-Gewichtung

Da pro Haushalt nur eine Person befragt wird, liefert das Stichprobenverfahren eine Haushaltsstichprobe, das heißt jeder Haushalt der Grundgesamtheit, der über einen Festnetz- oder Mobilfunkanschluss verfügt, hat die Chance, in die Stichprobe zu gelangen. Bedingt durch eine heterogene Anzahl an Mobiltelefonen und Festnetznummern pro Haushalt haben die Elemente der Grundgesamtheit zum Teil stark variierende Inklusionswahrscheinlichkeiten in die Stichprobe. Aus Effizienzgründen empfiehlt der Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V. (kurz: ADM) ein Mischungsverhältnis von 60-70% Festnetz-Interviews.

Um jene designbedingten Effekte für die Analysen zu bereinigen, erhalten die Fälle in der späteren Gewichtung ein Pre-Designgewicht, welches durch die Anzahl der „Zielpersonen im Haushalt“ ( $Z_{HH}$ ), der „Zielpersonen, die das angerufene Handy nutzen“ ( $Z_{Mobil}$ ), der „Festnetznummern, über die der Haushalt erreichbar ist“ ( $k^F$ ) und der „Handynummern, über die der Befragte persönlich erreichbar ist“ ( $k^C$ ) berechnet wird. Damit lässt sich für die  $i$ -te Person die Wahrscheinlichkeit bestimmen, in die Stichprobe zu gelangen:

$$(1) \quad \pi_i \approx k_i^F \cdot \frac{m^F}{M^F} \cdot \frac{1}{Z_{HH}} + k_i^C \cdot \frac{m^C}{M^C} \cdot \frac{1}{Z_{Mobil}}$$

Neben der realisierten Stichprobengröße im Festnetz ( $m^F$ ) und im Mobilnetz ( $m^C$ ) hängt dieser Wert auch von der Anzahl der Festnetz- ( $M^F$ ) bzw. Mobilfunknummern ( $M^C$ ) im Auswahlrahmen ab (Gabler/Ayhan 2007: 42 ff.). Nach einer Proportionalisierung beider Samples, die notwendig ist, damit die *Landline-Only*, das heißt diejenigen mit ausschließlich Festnetzanschluss, im Vergleich zu den *Mobile-Only* nicht übermäßig stark in der Stichprobe vertreten sind, wird über die typischen Merkmale, wie Geschlecht, Alter, Bundesland und Ortsgröße, eine Redressement-Gewichtung durchgeführt. Damit wird die Stichprobenstruktur an die amtliche Statistik angeglichen.

**b) Definition von Effizienzmaßen in einem RCS-Design**

In einer perfekten RCS-Erhebung müssten die Interviews eines jeden Tages eine identische Mischung aus mehreren Replikaten darstellen (siehe Tabelle 1). Somit sollten in der Theorie an jedem Tag gleich viele Interviews (in unserem Beispiel 56%) aus dem frischen Replikat generiert werden. Die absoluten Zahlen können dabei durchaus variieren. Jede Tagesstichprobe müsste aus genau 13,7% Interviews des Replikalters 1, aus 7,9% Interviews des Replikalters 2 etc. bestehen.

*Tabelle 1: Fiktive Interviewverteilung als Schema eines optimalen RCS-Designs*

Rep.-alter	Gesamt %	Tag 1		Tag 2		...		Tag n	
		N	%	N	%	N	%	N	%
0	56,0	560	56,0	1120	56,0	...		1120	56,0
1	13,7	137	13,7	274	13,7	...		274	13,7
2	7,9	79	7,9	158	7,9	...		158	7,9
3	5,6	56	5,6	112	5,6	...		112	5,6
...	...	...	...	...	...	...		...	...
14+	0,2	2	0,2	4	0,2	...		4	0,2

Diese theoretische Überlegung ist jedoch in der Praxis ohne feste Quoten, die wiederum dem Bestreben einer echten Zufallsstichprobe widersprechen, nicht umsetzbar. Trotzdem sollte es das Ziel sein zu erreichen, dass jede Tagesstichprobe eine möglichst ähnliche Mischung aus mehreren Replikaten darstellt.

Um die Strukturgleichheit zu quantifizieren, haben wir folgende RCS-Effizienz (innere Güte) definiert: Sei  $x$  der relative Anteil von Interviews mit dem  $i$ -ten Replikalter in der  $j$ -ten Tagesstichprobe, dann kann durch die Summe der betragsmäßigen Abweichungen zu den anderen Tagesstichproben ein RCS-Effizienzmaß für die Stichprobe  $j \geq m+1$  definiert werden:

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

Die Gesamt-RCS-Effizienz des Designs lässt sich durch den Mittelwert der  $n-m$  verschiedenen Tageseffizienzen quantifizieren:

$$(3) \quad I_{ges} = 1 - \frac{\sum_{k>j \geq m+1} \sum_{i=0}^m |x_{ij} - x_{ik}|}{(n-m) \cdot (n-m-1)}$$

Durch die wochentagsspezifische Ausschöpfung sind einzelne Tage leider nur bedingt miteinander vergleichbar, wodurch das Erreichen einer perfekten RCS-Struktur lediglich theoretischer Natur sein kann. Deshalb erweitern wir diese RCS-Effizienz (innere Güte) zu einer RCS-II-Effizienz (äußere Güte), die für jeden Tag angibt, wie homogen die Vor- und Folgewoche sind, denn durch Aggregation zu einer Woche werden die wochentagsspezifischen Effekte aufgelöst (insofern keine Feiertage vorhanden sind) und entsprechende Stichproben gewinnen an Homogenität bezüglich des Replikalters.

Perfekte RCS-Daten sind in der Realität unmöglich. Aber wir können sehr gute Daten durch eine konstante Behandlung der Teilstichproben erhalten. Der prozentuale Anteil der Anrufe eines Tages auf frische Nummern sollte beispielsweise stets identisch sein. Sei  $a$  der relative Anteil der Anwahlversuche bei Nummern mit dem Replikalter  $i$  zum Ende des  $j$ -ten Tages, so kann die RCS-Effizienz durch folgende A-Effizienz determiniert werden:

$$(4) \quad A_{ges} = 1 - \frac{\sum_{k>j \geq m+1} \sum_{i=0}^m |a_{ij} - a_{ik}|}{(n-m) \cdot (n-m-1)}$$

Außerdem wird eine identische Mischung der Replikate in den interviewten Tagesstichproben wahrscheinlicher bei einem möglichst gleichen Verhältnis der Teilstichproben zu Beginn des Tages. Hieraus ableitbar ist eine weitere Determinante der RCS-Effizienz. Sei  $n$  der relative Anteil von aktiven Fällen mit dem  $i$ -ten Replikalter zum Ende der  $j$ -ten Tagesstichprobe, so kann äquivalent zur RCS-Effizienz eine Teilstichprobeneffizienz (N-Effizienz) formuliert werden:

$$(5) \quad N_{ges} = 1 - \frac{\sum_{k>j \geq m+1} \sum_{i=0}^m |n_{ij} - n_{ik}|}{(n-m) \cdot (n-m-1)}$$

Beide Effizienzmaße haben einen Einfluss auf die RCS-Effizienz. Bedingt durch Auffälligkeiten in der N-Effizienz könnte theoretisch die A-Effizienz durch gezielten Einsatz im Telefonstudio beeinflusst werden, um die RCS-Effizienz demen-

sprechend zu erhöhen. Um die Grundprinzipien des RCS-Designs nicht zu verletzen, verzichten wir jedoch auf diesen Schritt. Wir streben an, beide Effizienzmaße lediglich durch eine möglichst konstante Umsetzung eines vorab definierten Kontaktschemas zu erhöhen.

Um die Repräsentativität der Teilstichproben und eine möglichst gleichmäßige Anzahl an Tagesinterviews zu gewährleisten, muss jede Teilstichprobe nach einem genau festgelegten Kontaktschema mit einer hohen Kontaktdichte abgearbeitet werden. Gemäß diesen Ideen erfolgte die Kontaktierung nach einem rollierenden Muster (siehe Tabelle 2).

*Tabelle 2: Kontaktschema in der Erhebung*

Tag	Montag bis Freitag			Samstag und Sonntag		
1	10-14 Uhr	14-18 Uhr	18-21 Uhr	11-14 Uhr	14-17 Uhr	17-20 Uhr
2	10-14 Uhr	14-18 Uhr	18-21 Uhr	11-14 Uhr	14-17 Uhr	17-20 Uhr
3	10-16 Uhr	16-21 Uhr		11-15 Uhr	15-20 Uhr	
...	...	...		...	...	
14	10-16 Uhr	16-21 Uhr		11-15 Uhr	15-20 Uhr	

Dabei wird werktags von 10-21 Uhr und an Wochenenden bzw. an Feiertagen von 11-20 Uhr angerufen. Jede Telefonnummer bleibt maximal 14 Tage aktiv. Die Kontaktierung erfolgt am 1. und 2. Einspieltag dreimal über den Tag (an Werktagen von 10-14 Uhr, 14-18 Uhr, 18-21 Uhr) verteilt. Vom 3. bis 14. Einspieltag wird zweimal am Tag angerufen. An Werktagen beginnen die Anwahlversuche um 10 Uhr bzw. 16 Uhr und enden um 16 Uhr bzw. 21 Uhr. Diese Anrufe werden über die Tage rollierend jeweils von Tag zu Tag um eine Stunde zeitversetzt durchgeführt, wobei der letzte Startpunkt um 14 Uhr bzw. 19 Uhr ist. Äquivalent dazu findet die Abarbeitung am Wochenende statt.

Der Algorithmus basiert auf der Annahme, dass tatsächlich eine gültige Nummer vorhanden ist, bei der der Telefonhörer nicht abgenommen wird. Alle anderen möglichen Ereignisse (ungültige Nummer, Fax oder Modem, besetzt oder Terminvereinbarung etc.) beenden diesen Algorithmus. Eine Ausnahme stellen Nummern dar, die vom System als „besetzt“ erkannt werden. Diese werden 30 Minuten später noch einmal angerufen. Wird bis zum 14. Feldtag kein eindeutiges Ergebnis erzielt, wird diese Nummer als stichprobenneutraler Ausfall gewertet. Terminvereinbarungen werden auch über die 14-Tage-Frist hinaus verabredet und entsprechend wahrgenommen. Bei jedem Kontaktversuch wird von der Telefonnummernverwaltung das entsprechende Ergebnis (zum Beispiel ein Interviewabbruch oder Besetztzeichen)

in Form eines Tipcodes aufgezeichnet und die betreffende Telefonnummer in unterschiedlichen Queues<sup>2</sup> abgelegt.

Der Ipsos Demokratiepuls wird kontinuierlich als tägliche RCS-Befragung der wahlberechtigten deutschen Bevölkerung durchgeführt. Nach mehreren Pretest-Wellen, die ab Oktober 2012 durchgeführt wurden, startete die Umfrage am 23. April 2013 mit insgesamt 1.200 frischen Nummern, davon 600 aus dem Festnetz- und 600 aus dem Mobilfunk-Frame (siehe Tabelle 3).

Tabelle 3: Replikatsgrößen, prognostizierte & tatsächliche Interviewanzahl nach Tag

Tag	Datum	Replikatsgrößen FN/MO	Interviews SOLL/IST	Tag	Datum	Replikatsgrößen FN/MO	Interviews SOLL/IST
1	Di, 23.04.	600/600	33/23	15	Di, 07.05.	800/600	50/50
2	Mi, 24.04.	600/600	49/40	16	Mi, 08.05.	800/600	50/49
3	Do, 25.04.	600/600	49/34	17	Do, 09.05.	1.200/900	50/46
4	Fr, 26.04.	800/800	49/42	18	Fr, 10.05.	700/500	50/56
5	Sa, 27.04.	1.000/750	50/50	19	Sa, 11.05.	1.000/750	50/46
6	So, 28.04.	1.000/750	50/38	20	So, 12.05.	1.200/900	50/45
7	Mo, 29.04.	700/550	50/63	21	Mo, 13.05.	500/350	50/50
8	Di, 30.04.	700/550	50/46	22	Di, 14.05.	800/600	50/65
9	Mi, 01.05.	1.100/800	50/44	23	Mi, 15.05.	800/600	50/56
10	Do, 02.05.	700/550	50/58	24	Do, 16.05.	800/600	50/43
11	Fr, 03.05.	800/600	50/50	25	Fr, 17.05.	900/650	50/57
12	Sa, 04.05.	1.000/750	50/49	26	Sa, 18.05.	1.000/750	50/59
13	So, 05.05.	1.200/900	50/48	27	So, 19.05.	1.200/900	50/55
14	Mo, 06.05.	500/350	50/57	28	Mo, 20.05.	1.200/900	50/50

Die Replikatsgrößen werden so gewählt, dass unter einer gleichmäßigen Abarbeitung nach Algorithmus etwa 50 Interviews pro Tag erzeugt werden. Dies wurde bereits am fünften Erhebungstag erreicht (siehe IST-Wert bei Interviews). Um ein optimales Verhältnis von 60% Festnetz- und 40% Mobilfunk-Interviews zu erhalten, haben wir ab Tag 5 mehr Festnetz- als Mobilfunk-Nummern aktiviert. Damit

2 Ausschlaggebend für die Zuordnung einer Nummer zu einem Queue ist in der Regel der letzte Anwahlversuch. So gelangen beispielsweise Rufnummern, zu denen konkrete Termine vereinbart werden konnten, per Algorithmus automatisch in „qapp“, während Ausfälle in „qausf“ gespeichert und somit nicht mehr angerufen werden.

haben wir nach der Anlaufphase bislang etwa 51,2 Interviews pro Tag erhoben. Die Extremwerte wurden an den Tagen 6 bzw. 22 mit jeweils 38 bzw. 65 Interviews erreicht. Um einer höheren Tagesausschöpfung an Montagen bzw. einer niedrigen Interviewwahrscheinlichkeit am Wochenende bzw. an Feiertagen entgegenzuwirken, werden die Größen zudem wochentagsspezifisch gewählt. Wegen signifikanten Unterschieden zwischen beiden Erhebungsframes in der Ausschöpfung und im Anteil qualitätsneutraler Ausfälle führen wir eine separate Schätzung für Festnetz und Mobilfunk durch.

*Tabelle 4: Replikatsgrößen für Festnetz und Mobilfunk nach Wochentag*

Wochentag	Festnetz	Replikatsgrößen	Mobile
Montag	500		350
Dienstag	800		600
Mittwoch	800		600
Donnerstag	800		600
Freitag	900		650
Samstag	1.000		750
Sonntag	1.200		900

Die Replikatsgrößen werden derart geschätzt, dass die Abweichungen zum Ziel von 50 Interviews minimiert werden, unter der Nebenbedingung eines Mindestmaßes an RCS-Effizienz. Für eine maximale Strukturgleichheit müssten wir am Wochenende weniger Nummern aktivieren. Wegen des geringen Stichprobenumfangs in den einzelnen Tagesstichproben verzichten wir auf diesen Schritt. Feiertage in der Woche behandeln wir in der Regel wie Sonntage und erhöhen die Replikatsgröße um 50 Prozent. Für eine konstante Stichprobengröße wird die Größe des Folgetages auf einen Feiertag in der Woche um 25% reduziert, weil durch den Feiertag bedingt übermäßig viele Nummern unausgeschöpft blieben. Um die nun leicht geringere Anzahl an aktiven Nummern auszugleichen, wird die Replikatsgröße am Folgetag des Folgetages um 10% erhöht. Damit erreichen wir eine sehr gute Prognosegüte bei einer mittleren Abweichung von 5,1 Interviews zum Ziel von fünfzig Interviews.

#### **4. Methodische Analysen**

##### **a) Mobile-Only und Mostly-Mobile im Vergleich**

Durch das zusätzliche Mobilfunk-Sample wird auch die Gruppe der *Mobile-Only* zum Wahlverhalten befragt. Darüber hinaus kann auch stärker die Klasse der *Mostly-Mobile* abgebildet werden, die zwar über Festnetz verfügen, aber „ausschließlich

über Mobilfunk<sup>3</sup> erreichbar sind. Aus allen Befragten ergibt sich ein Anteil von 14,8% *Mobile-Only* und 6% *Mostly-Mobile*.<sup>3</sup>

Beide Typen sind übermäßig oft männlich und deutlich jünger als die Gesamtstichprobe (siehe Tabelle 5). Bei den *Mostly-Mobile* ist der Anteil der Männer nochmals höher als bei den *Mobile-Only*. Zudem sind sie signifikant häufiger zwischen 30 und 49 Jahre alt.

Tabelle 5: Alter und Geschlecht nach Mobile-Typ

	Mobile-Only	Mostly-Mobile	Gesamt
Männer	61,6%	75,2%	51,5%
bis 29 Jahre	34,4%	27,8%	16,8%
30-49 Jahre	39,1%	46,6%	34,0%
ab 50 Jahre	26,5%	25,5%	49,2%
N	312	126	2.102

Der *Mobile-Only* ist mehr als doppelt so oft in der Gruppe der unter 30-jährigen vertreten. Zudem befindet er sich übermäßig oft in Ausbildung oder Umschulung. Auch ist der *Mobile-Only* häufiger arbeitslos und eher schlecht gebildet. Der *Mostly-Mobile* hingegen ist deutlich häufiger voll berufstätig und kaum arbeitslos. Sein höchster Schulabschluss hat eine ähnliche Verteilung wie in der Gesamtstichprobe mit leichter Tendenz zum Abitur.

Tabelle 6: Berufstätigkeit und Schulabschluss nach Mobile-Typ

	Mobile-Only	Mostly-Mobile	Gesamt <sup>4</sup>
voll berufstätig	49,6%	74,9%	43,2%
teilweise berufstätig	11,2%	5,1%	12,5%
in Ausbildung/Umschulung	16,4%	12,3%	9,6%
Rentner/Pensionär	10,0%	6,8%	27,7%
arbeitslos	12,9%	0,9%	7,1%
kein Schulabschluss	3,3%	1,0%	2,6%
Haupt-/Realschulabschluss	61,1%	54,9%	55,6%
Fachhochschulreife	10,3%	12,2%	12,8%
Abitur	25,4%	31,9%	29,1%

3 Bezogen auf die Erhebung im Zeitraum vom 23.4.2013 bis 2.6.2013.

4 Abweichungen zur amtlichen Statistik existieren, weil beide Merkmale in der Gewichtung nicht berücksichtigt wurden. Eine Gewichtung auf diesen Merkmalen wurde bewusst nicht vorgenommen, um die Unterschiede in den einzelnen Untersuchungsgruppen deutlicher darzustellen.

Selbst in der Haushaltsgröße und beim Haushaltseinkommen gibt es signifikante Unterschiede. Der Haushalt des *Mobile-Only* verdient eher schlecht, wohingegen der Privathaushalt des *Mostly-Mobile* finanziell deutlich besser gestellt ist (siehe Tabelle 7).

Tabelle 7: Haushaltsgröße und -einkommen nach Mobile-Typ

	Mobile-Only	Mostly-Mobile	Gesamt
unter 2.000 Euro	70,4%	43,3%	50,7%
2.000 – 5.000 Euro	26,9%	45,5%	41,5%
über 5.000 Euro	2,7%	11,3%	7,8%
Anzahl Haushaltsmitglieder	2,10	2,78	2,45

Verglichen mit der Gesamtstichprobe haben beide Typen ein weniger ausgeprägtes politisches Interesse, was die höheren Mittelwerte in Tabelle 8 bei diesem Merkmal zeigen. Dies ist auch konsistent mit der Absicht der Wahlbeteiligung. Der *Mobile-Only* ist sich zudem seiner Wahlentscheidung relativ sicher, der *Mostly-Mobile* hingegen noch unentschlossener.

Tabelle 8: Politische Partizipation und Wahlverhalten nach Mobile-Typ

	Mobile-Only	Mostly-Mobile	Gesamt
Sicherheit der Wahlabsicht	1,74	1,99	1,85
Politisches Interesse	3,07	2,82	2,75
Wahlbeteiligungsabsicht	1,94	1,84	1,57

*Anmerkungen:* Die Sicherheit der Wahlabsicht wird gemessen auf einer Skala von 1 = „ganz sicher“ bis 5 = „sehr unsicher“. Politikinteresse und Wahlbeteiligungsabsicht sind analog codiert. Sämtliche in dieser Tabelle dargestellten Zahlen entsprechen arithmetischen Mittelwerten.

Für die *Mobile-Only* ist die Arbeitslosigkeit das größte politische Problem Deutschlands (siehe Tabelle 9), vermutlich weil sie selbst übermäßig oft arbeitslos sind (vgl. Tabelle 6). Ebenfalls wird die Frage der Integration bzw. Migration im Vergleich zur Gesamtstichprobe häufig als größtes Problem genannt. Altersarmut, die Eurokrise und eine schlechte Bildungspolitik werden hingegen seltener erwähnt (siehe Tabelle 9).



Tabelle 9: Größtes politisches Problem in Deutschland nach Mobile-Typ

	Mobile-Only	Mostly-Mobile	Gesamt
Arbeitslosigkeit	21,9%	15,4%	14,9%
Bildungspolitik	3,2%	8,1%	5,3%
Energiewende/Umweltpolitik	3,5%	4,7%	4,2%
Eurokrise	15,5%	20,8%	18,4%
Integration/Migration	10,2%	3,3%	5,6%
Niedriglöhne	5,7%	6,2%	6,3%
Renten/Altersarmut	1,6%	3,5%	3,4%
Soziale Ungerechtigkeit	7,7%	8,2%	9,4%
Steuern	3,8%	4,4%	3,4%

Damit finden wir für die *Mobile-Only* in Deutschland in den Punkten Arbeitslosigkeit, Steuern und Integration/Migration ein ähnliches Antwortverhalten wie Keeter (2006) für die USA und Pennay (2010) für Australien. Nur im Punkt Umweltpolitik kommen wir zu einem anderen Ergebnis.

Die *Mostly-Mobile* bezeichnen vor allem die Eurokrise als größtes Problem. Zudem wird auch häufiger die Bildungs-, Umwelt- und Steuerpolitik genannt, vermutlich bedingt durch das höhere Einkommen und die bessere Bildung. Integration sowie Migration werden verglichen mit der Gesamtstichprobe dagegen deutlich seltener als größtes Problem genannt.

Zusammenfassend kann festgehalten werden, dass unsere soziodemografischen Analyseergebnisse weitestgehend konsistent sind mit anderen Arbeiten (Link u. a. 2007). Zudem weichen beide Typen – *Mobile-Only* und *Mostly-Mobile* – in der politischen Partizipation und in der Meinungsbildung von der Grundgesamtheit ab. Dies macht die Erhebung von Mobilfunkstichproben in der Wahlforschung für repräsentative Aussagen unabdingbar.

## b) RCS-spezifische Analysen

Für eine hohe RCS-Effizienz sollte eine gleichmäßige zeitliche Streuung der Anwahlversuche stattfinden. Tabelle 10 als ein Auszug zeigt, dass am ersten Erhebungstag insgesamt 1.952 Anrufe getätigt wurden, davon 412 (= 21%) zwischen 10-11 Uhr.

Tabelle 10: Zeitliche Streuung der Anwahlversuche im Demokratiepuls

Uhrzeit	Gesamt	Tag 01		Tag 02		Tag 03		Tag 04		Tag 05	
	%	%	N	%	N	%	N	%	N	%	N
10-11 Uhr	6%	21%	412	16%	427	12%	398	13%	543	0%	0
11-12 Uhr	9%	26%	498	18%	486	18%	574	18%	775	0%	0
12-13 Uhr	10%	15%	289	14%	374	23%	735	20%	867	0%	7
13-14 Uhr	8%	2%	36	7%	201	1%	22	3%	142	42%	2.098
14-15 Uhr	13%	7%	130	6%	156	12%	374	4%	162	10%	519
15-16 Uhr	9%	7%	141	6%	160	8%	260	6%	272	1%	33
16-17 Uhr	12%	5%	95	4%	106	0%	12	8%	330	7%	356
17-18 Uhr	7%	1%	23	8%	216	0%	3	8%	323	6%	293
18-19 Uhr	16%	2%	36	7%	185	14%	449	11%	478	34%	1.710
19-20 Uhr	7%	5%	104	14%	373	11%	356	5%	231	1%	32
20-21 Uhr	2%	10%	188	1%	15	1%	20	3%	109	0%	0
Gesamt	126.040	1.952		2.699		3.203		4.232		5.048	

Es ist ersichtlich, dass am fünften Tag keine Anwahlversuche vor 13 Uhr erfolgt sind. Dies liegt an der speziellen Behandlung der Wochenenden, an denen wir zu Werktagen abweichende Anrufzeiten umsetzen. Zu Beginn der Erhebung sah das Kontaktschema am Wochenende noch einen Erstanruf ab 13 Uhr vor. Für eine höhere RCS-Effizienz wurde die Anrufzeit am Wochenende im Feldverlauf auf 11 bis 20 Uhr ausgedehnt (siehe Tabelle 2). Die insgesamt 126.040 durchgeführten Wahlversuche streuen sehr gut. So entfallen zwischen zwei und 16% der Wahlversuche auf die einzelnen Zeitintervalle.

Für eine hohe RCS-Effizienz ist darüber hinaus eine erfolgreiche Abarbeitung des Kontaktschemas notwendig. Tabelle 11 gibt einen Überblick über den Erfolg im Feldverlauf. Beim ersten Replik (600 Festnetz- und 600 Mobilfunknummern) hätten in der ersten Anrufphase 1.200 verschiedene Nummern angewählt werden müssen, jede Nummer genau ein Mal (SOLL-Spalte). Dies wurde auch entsprechend umgesetzt (IST-Spalte). Nach dem ersten Zeitintervall waren noch 345 Fälle für die zweite Phase aktiv, die übrigen Nummern wurden entweder zum Interview, zum Ausfall oder zum Termin. In der dritten Anrufphase hätten 281 verschiedene Kontakte angerufen werden müssen, es wurden jedoch nur 276 unterschiedliche Personen angewählt.

Tabelle 11: Überprüfung des Anruf-Algorithmus nach definiertem Kontaktschema

Anruf	% erfüllt	Rep 01		Rep 02		Rep 03		Rep 04		Rep 05	
		SOLL	IST	SOLL	IST	SOLL	IST	SOLL	IST	SOLL	IST
1	100%	1.200	1.200	1.199	1.199	1.199	1.199	1.601	1.599	1.751	1.751
2	94%	345	345	380	380	355	355	481	480	501	76
3	96%	281	276	312	305	295	293	394	391	491	490
4	100%	231	231	261	261	240	239	344	344	341	340
5	98%	204	204	240	229	216	213	313	287	311	304
6	86%	198	198	219	217	206	193	304	301	285	284
7	100%	194	194	205	205	186	185	276	276	261	260
8	100%	185	185	197	197	175	174	254	254	238	234
9	93%	182	182	197	196	161	160	238	238	225	222
10	96%	178	178	191	190	150	149	224	116	213	210
...											
30	91%	101	97	96	90	93	83	109	97	134	126
Gesamt % erfüllt	103.035 93%	5.757 93%	5.328	6.005 86%	5.190	5.475 93%	5.070	7.676 91%	6.992	7.788 88%	6.841

Insgesamt 93% (= 103.035) der Gesamtanrufe laut Algorithmus (147.904) wurden auch entsprechend umgesetzt. Eine Einzelfallanalyse ergab, dass ein Großteil der Verfehlungen durch Kapazitätsengpässe im Telefonstudio entstand mit größtenteils minimalen Abweichungen von maximal 60 Minuten.

Die Umsetzung des Kontaktschemas hat natürlich unmittelbaren Einfluss auf die Interviewrealisierung. So hat Tag 18 im ausgewählten Beispiel die größte RCS-Effizienz von 80%. Bezüglich des Replikalters weicht dieser Tag nur minimal vom Durchschnitt ab (vgl. Tabelle 12). So wurden an jenem Freitag insgesamt 57,1% der Interviews mit frischen Nummern generiert (im Vergleich zu durchschnittlichen 56%).

An Tag 21 wurden mehr Interviews (32,7%) mit dem Replikalter 1 durchgeführt als mit frischen Replikaten (26,5%). Die Verteilung der Interviews an diesem Montag weicht damit natürlich stark von den anderen Tagen ab, was sich auch in der niedrigen RCS-Effizienz von 59% widerspiegelt. Grund hierfür ist die geringe Replikatsgröße an diesem Tag (vgl. Tabelle 3) und das vorherige Wochenende. Dort wurden sehr viele Nummern aktiviert, die erst an einem Werktag erreicht werden konnten. Außerdem gab es aus der Woche davor vermutlich einige Terminvereinbarungen, die zum Interview wurden. Somit sind an diesem Tag die sehr leicht

Tabelle 12: Auszug der Interviewverteilung nach Replikatalter

Rep.- alter	Ge- samt %	Tag 16		Tag 17		Tag 18		Tag 19		Tag 20		Tag 21		Tag 22	
		%	N	%	N	%	N	%	N	%	N	%	N	%	N
0	56,0	63,3	31	63,0	29	57,1	32	47,8	22	75,6	22	26,5	13	75,6	34
1	13,7	10,2	5	6,5	3	16,1	9	15,2	7	4,4	7	32,7	16	4,4	4
2	7,9	8,2	4	8,7	4	8,9	5	17,4	8	0,0	8	10,2	5	0,0	4
3	5,6	8,2	4	0,0	0	3,6	2	2,2	1	6,7	1	4,1	2	6,7	4
4	3,7	6,1	3	2,2	1	0,0	0	0,0	0	0,0	0	10,2	5	0,0	3
5	2,9	0,0	0	4,3	2	1,8	1	0,0	0	6,7	0	2,0	1	6,7	6
6	1,5	0,0	0	0,0	0	1,8	1	0,0	0	2,2	0	2,0	1	2,2	2
7	2,5	4,1	2	13,0	6	3,6	2	0,0	0	2,2	0	2,0	1	2,2	0
8	1,1	0,0	0	2,2	1	0,0	0	4,3	2	0,0	2	0,0	0	0,0	0
9	1,3	0,0	0	0,0	0	1,8	1	2,2	1	0,0	1	4,1	2	0,0	3
10	1,1	0,0	0	0,0	0	3,6	2	6,5	3	2,2	3	0,0	0	2,2	1
11	0,7	0,0	0	0,0	0	0,0	0	4,3	2	0,0	2	2,0	1	0,0	0
12	0,4	0,0	0	0,0	0	0,0	0	0,0	0	0,0	0	2,0	1	0,0	2
13	0,8	0,0	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0	1
14+	0,2	0,0	0	0,0	0	1,8	1	0,0	0	0,0	0	0,0	1	0,0	1

erreichbaren Personen etwas unterrepräsentiert und die *Hard-to-Reach*-Personen übermäßig oft vertreten.

Bis dato erreichen wir für unsere Erhebung eine Gesamt-RCS-Effizienz von 74% (siehe Tabelle 13). Die zudem relevantere RCS-II-Effizienz liegt mit 91% deutlich höher.

Tabelle 13: Auszug der RCS-Effizienzen

	Gesamt	Tag 16	Tag 17	Tag 18	Tag 19	Tag 20	Tag 21	Tag 22
RCS-Effizienz	0,74	0,78	0,73	0,78	0,70	0,69	0,59	0,74
RCS-II-Effizienz	0,91	0,92	0,92	0,89	0,89	0,88	0,92	0,92

Sollte an einem dieser sieben Tage ein Ereignis geschehen, dessen Effekt es zu analysieren gilt, so wäre Tag 20 wegen der kleinsten RCS-II-Effizienz in Höhe von 88% dafür am wenigsten geeignet. Im Vergleich zur RCS-Effizienz haben die Werte der äußeren RCS-Effizienz eine geringere Varianz. Diese Stabilität ergibt sich aus

der Definition und nimmt mit Zunahme des Lags<sup>5</sup> leicht ab. Die RCS-Güte der erhobenen Daten kann zusammenfassend als sehr hoch eingestuft werden.

Dieses Kapitel hat die Praxistauglichkeit dieser Kombination aus *Dual-Frame*-Stichprobe und RCS-Design eindeutig nachgewiesen. Die Vorgaben nach Kontaktschema können entsprechend eingehalten werden, so dass eine ausreichende Strukturgleichheit auf Tagesebene erzeugt wird. Zudem gehen nur durch Berücksichtigung einer Mobilfunkstichprobe die zur Grundgesamtheit gehörenden *Mobile-Only* bzw. *Mostly-Mobile* überhaupt bzw. adäquat in die Stichprobe ein. Dies ist insofern von Relevanz, da sich beide Typen in der Demografie und in der politischen Partizipation signifikant von der Gesamtbevölkerung unterscheiden.

## 5. Inhaltliche Analysen

### a) Interpretation verschiedener politischer Zeitreihen

Aufgrund der relativ kleinen Tagesstichproben unterliegen Analysen auf Tagesbasis üblicherweise größeren Zufallsschwankungen. Für die folgenden Analysen wurden die Daten deshalb mithilfe robuster lokal gewichteter Regressionen (*locally weighted scatterplot smoothing*, LOWESS)<sup>6</sup> geglättet, um übergreifende Trends besser zu veranschaulichen (Cleveland 1979; Kenski 2004 b). Im Ergebnis erhalten wir eine Zeitreihe, die um datenbedingte Fluktuationen bereinigt ist. Zunächst betrachten wir die Entwicklung der Wahlabsicht bzw. Kanzlerpräferenz innerhalb der Untersuchungsperiode. Das Ziel hierbei ist es nicht, eine politische Prognose abzugeben, sondern tagesaktuelle Schwankungen in der politischen Stimmung in Deutschland aufzuzeigen. Ein Vergleich mit alternativen Umfrageergebnissen ist insofern schwierig, da diese im Gegensatz zu den hier vorgestellten Werten in der Regel über die Gewichtung hinaus von den Rohdaten abweichen, um Effekte wie soziale Erwünschtheit oder taktisches Wählen zu berücksichtigen. Nur bei Kenntnis der verwendeten Projektionen lassen sich aussagekräftige Vergleiche durchführen.

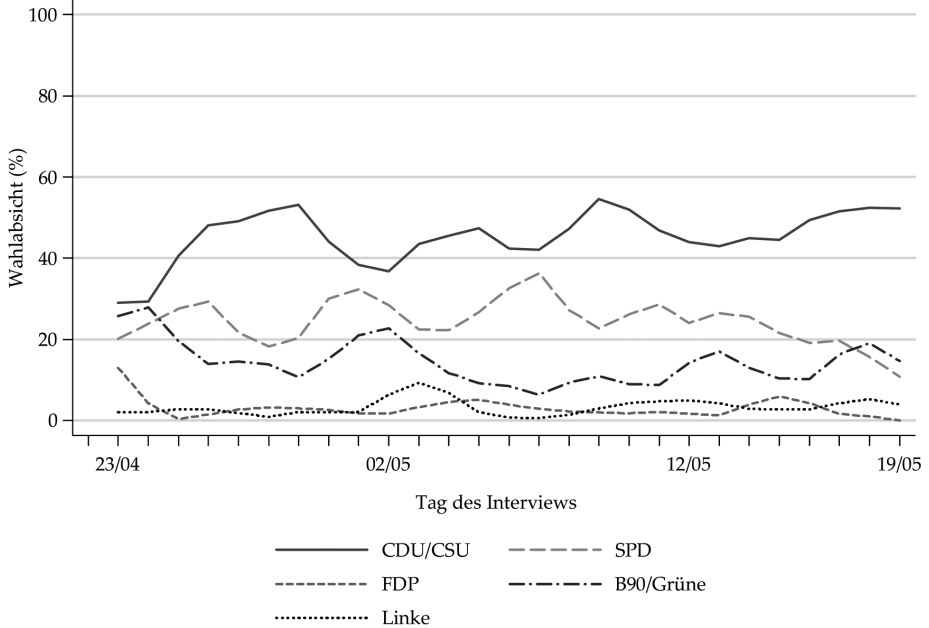
Die Abbildung 2 zeigt die Entwicklung der Wahlabsicht in unserer Untersuchungsperiode. Zwar liegt die Union in der Wählergunst deutlich vorn, die RCS-

5 Als Lag ist in der Statistik ein Operator bekannt, der eine Zeitreihe um Zeiteinheiten verschiebt. Die RCS-II-Effizienzen der Tage 16 und 17 beispielsweise berechnen sich aus Werten zweier Reihen, die mit Ausnahme der Tage 9, 16, 17 und 24 deckungsgleich sind. Diese Deckungsgleichheit nimmt mit zunehmendem Lag, also größerer Verschiebung, ab, was zu einer leicht abnehmenden Stabilität der RCS-II-Effizienz führt.

6 LOWESS ist ein nichtparametrisches Glättungsverfahren, bei dem für jeden Datenpunkt ein Schätzwert berechnet wird, in dem die jeweils benachbarten Datenpunkte mit unterschiedlichen Gewichten eingehen. Das Gewicht jedes benachbarten Punktes nimmt dabei mit zunehmender Distanz ab. Für die folgenden Analysen verwenden wir einen Glättungsparameter von 0.2.

Daten zeigen aber klare Schwankungen in der Wahlabsicht, die eine rein statische Betrachtung nicht offenbart.

Abbildung 2: Entwicklung der Wahlabsicht (geglättet nach LOWESS)

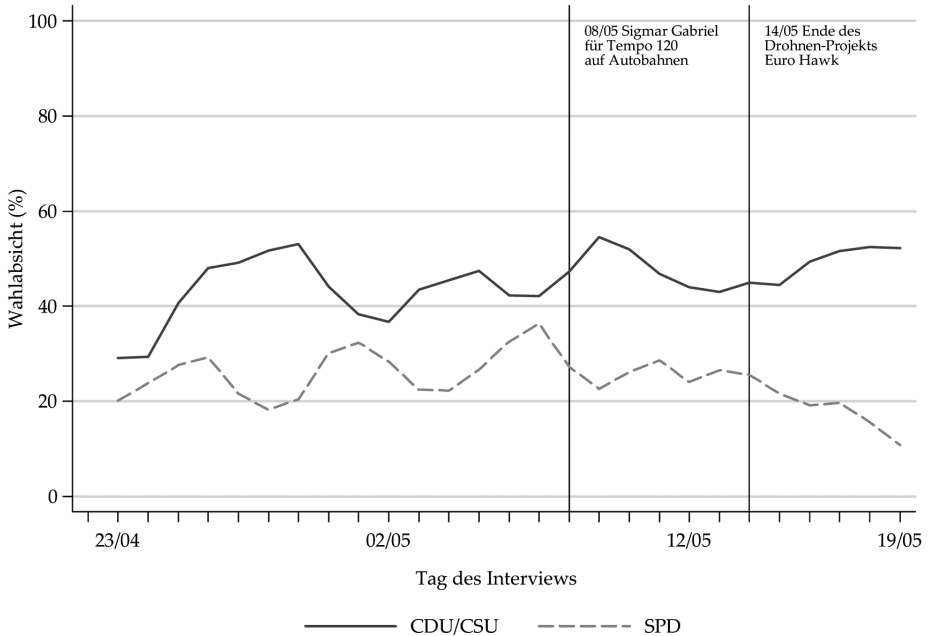


Auch die SPD und die Grünen weisen zum Teil starke Fluktuationen auf. Wie sind diese Schwankungen zu erklären?

Abbildung 3 zeigt die Verknüpfung der Parteipräferenzen bei zwei signifikanten Ereignissen im Medioumfeld, über die im Untersuchungszeitraum intensiv berichtet wurde und somit Anknüpfungspunkte an unsere theoretischen Überlegungen zu Agenda-Setting und Priming-Effekten bieten. Da wäre zum einen die Forderung des SPD-Vorsitzenden Sigmar Gabriel nach einem Tempolimit von 120 auf deutschen Autobahnen. Mit seinen Äußerungen hat er nicht nur ein umfangreiches Medienecho hervorgerufen, sondern auch durch diese in der Bevölkerung negativ besetzte Forderung einen erheblichen Einbruch in der Wahlabsicht für die SPD hervorgerufen, wie unsere Darstellung zeigt. Das andere Ereignis, das im Fokus der Berichterstattung stand, war die Bekanntgabe, dass das Drohnen-Projekt „Euro Hawk“ in einem finanziellen Debakel endet. Wie unsere Darstellung verdeutlicht, hat die beginnende Berichterstattung hierüber nicht zu einem Einbruch bei der

Union geführt, was vermutlich darauf zuzuführen ist, dass das ganze Ausmaß des Scheiterns erst Anfang Juli bekannt wurde. Nichtsdestotrotz konnte die SPD anscheinend nicht von dem sich abzeichnenden Euro-Hawk-Fiasko profitieren.

Abbildung 3: Medienereignisse und Wahlabsicht (geglättet nach LOWESS)



**b) Test der Ereignisse auf signifikante Effekte**

Aber sind die grafisch beobachteten Effekte auch statistisch signifikant? Aufgrund der geringen RCS-Effizienz und Größe einer Tagesstichprobe testen wir die Strukturbrüche auf Wochenbasis. Tabelle 14 zeigt die Verteilung der Wahlabsicht im Zeitraum vor bzw. nach dem jeweiligen Ereignis. Das Befürworten von Sigmar Gabriel für Tempo 120 auf Autobahnen an Tag 16 führt zu einer Reduktion der SPD-Wahlabsicht um 3%. Auch die Grünen als möglicher Koalitionspartner erfahren durch die Aussagen des SPD-Vorsitzenden einen Verlust von 3,4 Prozentpunkten. Vor allem die Piratenpartei, aber auch Die Linke sind auf den ersten Blick Nutznießer dieses medialen Ereignisses.

Tabelle 14: Verteilung der Wahlabsicht eine Woche vor bzw. nach dem Ereignis

	Tag 09-15	Tag 17-23	Tag 15-21	Tag 23-29
CDU/CSU	45,4%	46,4%	42,4%	45,7%
SPD	29,9%	26,9%	30,6%	23,0%
B90/Grüne	13,7%	10,3%	10,1%	15,8%
FDP	2,8%	2,3%	2,1%	3,0%
Die Linke	3,0%	4,6%	4,2%	3,5%
Piratenpartei	0,9%	3,5%	3,6%	3,6%
AfD	1,0%	1,0%	2,9%	1,0%

Der Stopp des Euro-Hawk-Projekts hat dazu geführt, dass die SPD 7,6% verliert, während die Grünen 5,7% zulegen. Die CDU/CSU gewinnt 3,3% an Stimmen, während die Wahlabsicht für die AfD um 1,9% abnimmt.

In Tabelle 15 ist ersichtlich, dass zum 10%-Niveau lediglich die Grünen signifikant von der Beendigung des Euro-Hawk-Projekts profitieren. Dies ist zudem konsistent mit der politischen Meinung der Partei, die sich von Anfang an gegen die Anschaffung der Drohnen aussprach. Darüber hinaus gibt es keinen nachweisbaren Effekt des Ereignisses auf die Wahlabsicht der anderen Parteien.

Tabelle 15: *p*-Werte paarweiser Chi-Quadrat-Tests

	„Tempo 120“	„Euro-Hawk“
CDU/CSU	0,602	0,390
SPD	0,220	0,144
B90/Grüne	0,160	0,075
FDP	0,673	0,469
Die Linke	0,555	0,693
Piratenpartei	0,072	0,917
AfD	0,944	0,452

Das Befürworten von Sigmar Gabriel für ein Tempolimit auf deutschen Autobahnen bewirkt lediglich einen zum 10%-Niveau signifikanten Anstieg der Wahlabsicht der Piratenpartei. Auch dies ist vereinbar mit der grundsätzlichen Ablehnung eines generellen Tempolimits auf Autobahnen durch die Piratenpartei (Piratenpartei 2012). Für die anderen Parteien ist kein Effekt nachweisbar. Jedoch sei hier abschließend auf den trotz der Wochenaggregation immer noch sehr geringen Stichprobenumfang hingewiesen.



## 6. Fazit

Dieser Beitrag wendet ein in der politikwissenschaftlichen Forschung etabliertes Design – das Rolling-Cross-Section Design (RCS-Design) – an, wodurch die Effekte unvorhersehbarer, medialer Ereignisse auf Tagesebene dynamisch untersucht werden können. So identifiziert unsere Analyse einen positiven Einfluss auf die Wahlabsicht der Piratenpartei nach der Forderung nach einem generellen Tempolimit auf deutschen Autobahnen durch den SPD-Vorsitzenden Sigmar Gabriel beziehungsweise einen Zuwachs der Wahlabsicht zugunsten der Grünen nach dem Euro-Hawk-Debakel.

Zudem werden in diesem Design erstmalig Festnetz- und Mobilfunkstichprobe miteinander verknüpft. Dadurch lassen sich vor allem über den noch relativ unbekannteren *Mostly-Mobile* neue Erkenntnisse gewinnen. So gehören dieser Gruppe überdurchschnittlich viele Männer im mittleren Alter mit hoher Bildung und gutem Einkommen, aber einem geringen politischen Partizipationsniveau an.

Da *Mobile-Only* und *Mostly-Mobile* stark von der Grundgesamtheit abweichen, sind zusätzliche Erhebungen über Mobilfunk für repräsentative Analysen absolut notwendig. Auch stichprobentheoretisch wird eine höhere Mächtigkeit des Stichprobenraums erzeugt; durch das RCS-Design hingegen sinkt die Varianz der Auswahlwahrscheinlichkeiten, denn im Vergleich zu alternativen Wahlumfragen werden auch schwer erreichbare Sub-Populationen befragt. Die Stichprobe erhält damit eine deutlich höhere Qualität.

Für eine perfekte RCS-Erhebung müssten jeden Tag die Interviews bzgl. ihres Replikalters homogen zueinander sein. Der Grad dieser Strukturgleichheit kann mit der RCS-Effizienz beschrieben werden. Das Ziel konstanter Tagesfallzahlen liegt im Widerspruch zur Maximierung dieser inneren Effizienz, da die Interviewwahrscheinlichkeit vom Wochentag abhängig ist. Bei großer Heterogenität oder zu kleinen Tagesstichproben kann die innere Güte zu einer RCS-II-Effizienz erweitert werden, um einen Teil der Strukturungleichheit zu bereinigen.

Durch Berücksichtigung einer replikalters- und zugleich wochentagspezifischen Ausschöpfung lassen sich die optimalen Replikatsgrößen mit einer mittleren Abweichung von etwa 10% zum Tagesziel von fünfzig Interviews gut voraussagen. Mit dem Ipsos Demokratiepuls werden kontinuierlich Daten erhoben, um die Veränderungen der politischen Stimmung in Deutschland tagessgenau zu beschreiben und zu analysieren.

## Literatur

- Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V. (ADM)*, 2012: ADM-Forschungsprojekt, „Dual-Frame-Ansätze“, Forschungsbericht, <https://www.adm-ev.de/index.php?id=forschungsprojekte> (Stand: 6.6.2013).
- Brettschneider, Frank*, 2005: Massenmedien und Wählerverhalten, in: Jürgen Falter/Harald Schoen (Hrsg.), *Handbuch Wahlforschung*, Wiesbaden, 473-500.
- Campbell, Angus/Converse, Philip/Miller, Warren/Stokes, Donald*, 1960: *The American Voter*, New York.
- Cleveland, William*, 1979: Robust Locally Weighted Regression and Smoothing Scatterplots, in: *Journal of the American Statistical Association* 74, 829-836.
- Gabler, Siegfried/Ayhan, Öztas*, 2007: Gewichtung bei Erhebungen im Festnetz und über Mobilfunk: ein Dual Frame, in: Siegfried Gabler/Sabine Häder (Hrsg.), *Mobilfunktelefonie. Eine Herausforderung für die Umfrageforschung*, ZUMA-Nachrichten 13, 47-58.
- Graeske, Jennifer/Kunz, Tanja*, 2009: Stichprobenqualität der CELLA-Studie unter besonderer Berücksichtigung der Mobile-onlys, in: Michael Häder/Sabine Häder (Hrsg.), *Telefonbefragungen über das Mobilfunknetz. Konzept, Design und Umsetzung einer Strategie zur Datenerhebung*, Wiesbaden, 57-70.
- Häder, Sabine/Häder, Michael/Kühne, Mike*, 2012: *Telephone Surveys in Europe. Research and Practice*, Wiesbaden.
- Hoffmann, Hermann*, 2007: Kombinierte Stichproben für Telefonumfragen – Ansätze in Europa, in: Siegfried Gabler/Sabine Häder (Hrsg.), *Mobilfunktelefonie. Eine Herausforderung für die Umfrageforschung*, ZUMA-Nachrichten 13, 47-58.
- Holst, Christian/Schmitt-Beck, Rüdiger/Faas, Thorsten*, 2006: Kampagnendynamik: Der Nutzen des Rolling-Cross-Section Designs für die Kommunikationsforschung mit einem Beispiel aus der Wahlforschung, in: *planung&analyse* 5, 54-58.
- Hoops, Christian/Michael, Tobias*, 2013: Effizienzmessung und Steuerung des RCS-Designs – Eine Betrachtung im Hinblick auf die Bundestagswahl, in: *DAGStat Tagungsband*, Freiburg.
- Hoops, Christian/Schnapp, Kai-Uwe*, 2013: Sampling of Hard-to-Reach Populations, in: *5<sup>th</sup> European Survey Research Association Conference Book*, Ljubljana.
- Hunsicker, Stefan/Schroth, Yvonne*, 2007: Die Kombination von Mobilfunk- und Festnetzstichproben – Eine praktische Anwendung des Dual-Frame-Ansatzes, in: *Methoden-Daten-Analysen* 1 (2), 161-182.

- Johnston, Richard/Brady, Henry*, 2002: The Rolling Cross-Section Design, in: *Electoral Studies* 21, 283-295.
- Keeter, Scott*, 2006: The Impact of Cell Phone Noncoverage Bias on Polling in the 2004 Presidential Election, in: *Public Opinion Quarterly* 70, 88-98.
- Keeter, Scott/Kennedy, Courtney/Clark, April/Tompson, Trevor/Mokrzycki, Mike*, 2007: What's Missing From National Landline RDD Surveys? The Impact of the Growing Cell-Only Population, in: *Public Opinion Quarterly* 71, 772-792.
- Kenski, Kate*, 2004 a: The Rolling Cross-Section Design, in: Daniel Romer/Kate Kenski/Paul Waldman/Christopher Adasiewicz/Kathleen Hall Jamieson (Hrsg.), *Capturing Campaign Dynamics: The National Annenberg Election Survey*, New York, 56-65.
- Kenski, Kate*, 2004 b: Visualizing Data Across the Campaign, in: Daniel Romer/Kate Kenski/Paul Waldman/Christopher Adasiewicz/Kathleen Hall Jamieson (Hrsg.), *Capturing Campaign Dynamics: The National Annenberg Election Survey*, New York, 66-79.
- Link, Michael/Battaglia, Michael/Frankel, Martin/Osborn, Larry/Mokdad, Ali*, 2007: Reaching the U.S. Cell Phone Generation. Comparison of Cell Phone Survey Results with Ongoing Landline Telephone Survey, in: *Public Opinion Quarterly* 71, 814-839.
- McCombs, Maxwell/Shaw, Donald*, 1972: The Agenda-Setting Function of Mass Media, in: *Public Opinion Quarterly* 36, 176-187.
- Miller, Joanne/Krosnick, Jon*, 1996: News Media Impact on the Ingredients of Presidential Evaluations. A Program of Research on the Priming Hypothesis, in: Diana Mutz/Paul Sniderman (Hrsg.), *Political Persuasion and Attitude Change*, Ann Arbor, 79-99.
- Pennay, Darren*, 2010: Profiling the Mobile-Phone-Only Population. Results From a Dual-Frame Telephone Survey Using a Landline and Mobile Phone Sample Frame, in: *ACSPRI Social Science Methodology Conference Book*, Sydney.
- Piratenpartei*, 2012: Positionspapier – Ablehnung eines generellen Tempolimits, <http://www.piratenpartei-nrw.de/2012/04/24/positionspapier-ablehnung-eines-generellen-tempolimits/> (Stand: 6.6.2013).
- Schmitt-Beck, Rüdiger*, 2009: Kampagnendynamik im Bundestagswahlkampf 2005, in: Oscar Gabriel/Bernhard Weißels/Jürgen Falter (Hrsg.), *Wahlen und Wähler. Analysen aus Anlass der Bundestagswahl 2005*, Wiesbaden, 146-176.
- 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/Rattinger, Hans/Rofsteuscher, Sigrid/Wefßels, Bernhard*, 2010: Die deutsche Wahlforschung und die German Longitudinal Election Study (GLES), in: Frank Faulbaum/Christof Wolf (Hrsg.), *Gesellschaftliche Entwicklungen im Spiegel der empirischen Sozialforschung*, Wiesbaden, 141-172.
- Schoen, Harald*, 2004: Der Kanzler, zwei Sommerthemen und ein Foto-Finish: Priming-Effekte bei der Bundestagswahl 2002, in: Frank Brettschneider/Jan van Deth/Edeltraud Roller (Hrsg.), *Die Bundestagswahl 2002. Analysen der Wahlergebnisse und des Wahlkampfes*, Wiesbaden, 23-50.
- Schoen, Harald/Weins, Cornelia*, 2005: Der sozialpsychologische Ansatz zur Erklärung von Wahlverhalten, in: Jürgen Falter/Harald Schoen (Hrsg.), *Handbuch Wahlforschung*, Wiesbaden, 187-242.
- Weaver, Daniel/McCombs, Maxwell/Shaw, Donald*, 2004: Agenda-Setting Research – Issues, Attributes, and Influences, in: Lynda Lee Kaid (Hrsg.), *Handbook of Political Communication Research*, Mahwah, 257-282.

Korrespondenzanschrift:

Christian Hoops  
Ipsos Public Affairs  
Sachsenstraße 6  
20097 Hamburg  
E-Mail: [Christian.Hoops@Ipsos.com](mailto:Christian.Hoops@Ipsos.com)

Dr. Alexander Glantz  
Ipsos Public Affairs  
Schwartzkopffstraße 11  
10115 Berlin  
E-Mail: [Alexander.Glantz@Ipsos.com](mailto:Alexander.Glantz@Ipsos.com)

Tobias Michael  
Ipsos Public Affairs  
Schwartzkopffstraße 11  
10115 Berlin  
E-Mail: [Tobias.Michael@Ipsos.com](mailto:Tobias.Michael@Ipsos.com)