

**Meinungsforschung
und
Meinungsmache
am
Beispiel
des
Bundestagswahlkampfes
2005**

**Was taugt die
„Sonntagsfrage“?**

**Vortrag Universität Koblenz-Landau,
9./10. September 2005**

**Von
Prof. Dr. Joachim Hofmann-Göttig**

Meinungsforschung und Meinungsmache am Beispiel des Bundestagswahlkampfes 2005

Was taugt die „Sonntagsfrage“?

Vortrag Universität Koblenz-Landau, 9./10. September 2005
Von Prof. Dr. Joachim Hofmann-Göttig¹

Es ist schon wahr, was die Wochenzeitung *DIE ZEIT* kürzlich diagnostizierte: „Noch nie wurden vor einer Bundestagswahl so viele Umfragen veröffentlicht wie in diesem Sommer.“²

Es vergeht nahezu kein Tag, an dem nicht der Wählerschaft neuestes Demoskopisches aufgetischt wird, Verdauliches und Unverdauliches. An einem Tag – vier Wochen vor der Bundestagswahl³ - verkündeten die beiden führenden politischen Wochenmagazine widersprüchliche Botschaften.

Widersprüchliche Botschaften in Medien

Der *FOCUS* machte in für dieses Magazin nicht untypischer Art dem vermuteten eigenen politischen Klientel Mut. Er präsentierte die aktuell publizierten partei-politischen Stimmungszahlen mit der Überschrift: „Schwarz-Gelb kurz vor dem Ziel“.⁴



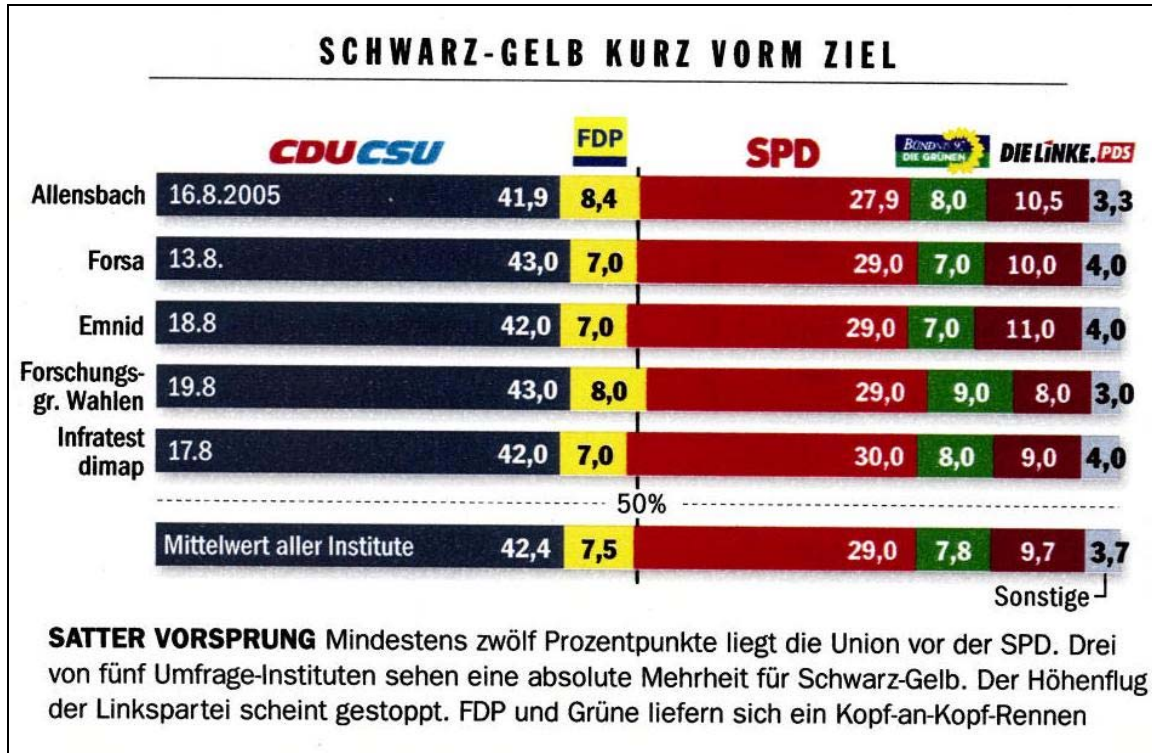
¹ Lehrbeauftragter am Institut für Soziologie der Universität Koblenz-Landau. Der Verfasser beschäftigt sich seit über drei Jahrzehnten mit der Wahlforschung; vgl. Monografien *Joachim Hofmann-Göttig*, Die jungen Wähler, Frankfurt 1984; ders., Emanzipation mit dem Stimmzettel – 70 Jahre Frauenwahlrecht in Deutschland, Bonn 1986; ders., Die neue Rechte: Die Männerparteien, Bonn 1989 (weitere Veröffentlichungen in Bibliografie in: www.mbfj.rlp.de).

Für die Lösung mancher Software-Probleme bei der Verknüpfung von WORD, EXCEL, Power Point dem World Wide Web ist *Barbara Loos* zu danken.

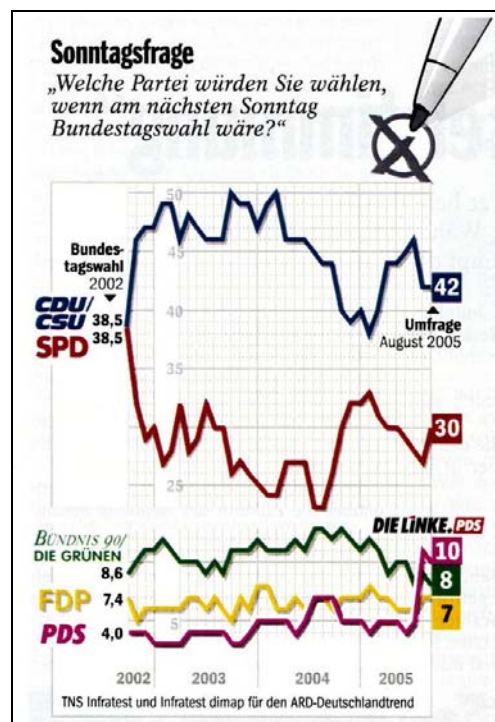
² *Matthias Krupa*, Wahlen nach Zahlen, *DIE ZEIT* Nr. 33/2005; zitiert nach Internet <http://zeus.zeit.de/text/2005/33/Demoskopen>; zugegriffen am 19.8.05.

³ Die akademische Wahlforschung hat gerade noch rechtzeitig vor der neuen Bundestagswahl ihre traditionelle Analyse (das „Blaue Bändchen“, so genannt wegen der traditionellen Einbindfarbe) der letzten Bundestagswahl vorlegen können (vgl. *Jürgen W. Falter/Oscar W. Gabriel/Bernhard Wessels* (Hrsg.), Wahlen und Wähler, Analysen aus Anlass der Bundestagswahl 2002, Wiesbaden Juli 2005.

⁴ *FOCUS*, Nr. 34 vom 22. August 2005, S. 24.



DER SPIEGEL hingegen scheint eine andere partei-politische Orientierung seiner Leserschaft als das Konkurrenzblatt aus München bedienen zu wollen. Auch dort wird eine aktuelle „Sonntagsfrage“ („Wen würden Sie wählen, wenn am nächsten Sonntag Bundestagswahl wäre?“) von „Infratest“ präsentiert, die Schwarz-Gelb vorn sieht.



Aber die Botschaft ist das Gegenteil von „Schwarz-Gelb kurz vor dem Ziel“: Im Inhaltsverzeichnis wird die Umfrage angetextet mit der Aussage „Knapper Vorsprung für Schwarz-Gelb“ und auch der eigentliche Artikel steht unter der Überschrift „Die große Unsicherheit“.



Und *DER SPIEGEL* verstärkt die Botschaft, dass trotz mehrheitlicher demoskopischer Befunde noch nicht alles gelaufen ist durch die Präsentation einer Zusatzinformation, wonach sich laut *Infratest*-Erhebung im August 2005 erst 48 Prozent der Wählerschaft entschieden hätten, „welche Partei sie bei der nächsten Bundestagswahl wählen werden“. Exakt so viel, eben auch 48 Prozent, sagten im Vor-Wahl-Monat „Nein, bin noch nicht entschlossen“. Und erstaunlich genug: Mit dem Fortgang des „Wahlkampfes“ attestiert *Infratest* nicht etwa eine wachsende Sicherheit der Entscheidungsfindung, sondern eine leicht steigende Tendenz der Unentschlossenheit.⁵

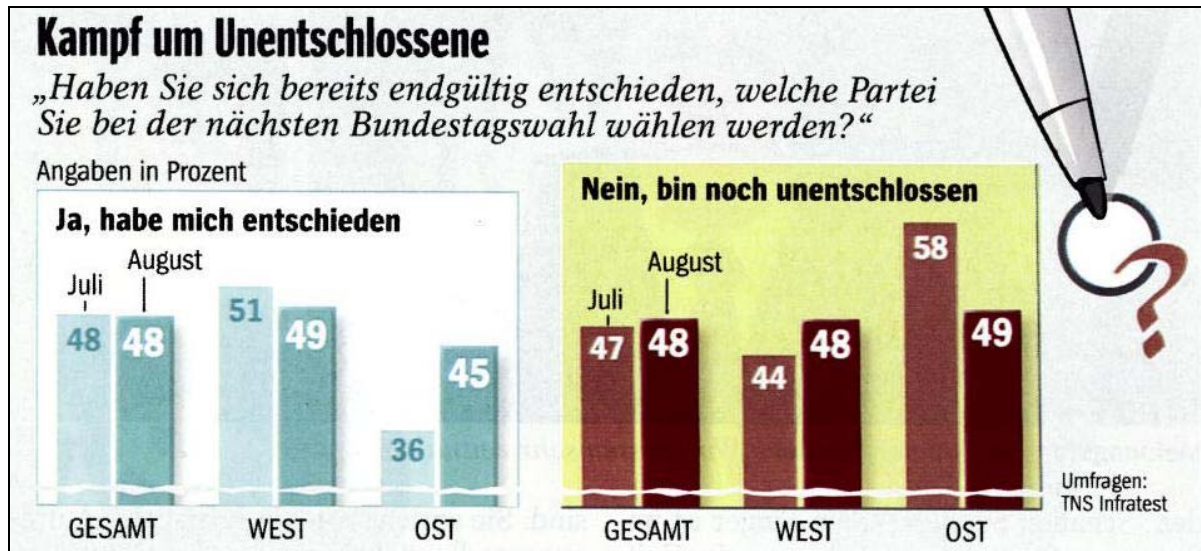
Deutschland

MEINUNGSFORSCHUNG

Die große Unsicherheit

Wird Angela Merkel Kanzlerin einer schwarz-gelben Regierung? Oder kommt die Große Koalition?
Einen Monat vor der Wahl hat sich fast die Hälfte der Bürger noch nicht endgültig entschieden. Kleine Verschiebungen der Stimmanteile können erhebliche Folgen haben.

⁵ *DER SPIEGEL*, Nr. 34 vom 22. August 2005, S. 22 ff.



Manipulation der Medien

Widersprüchliche Interpretationen von Umfragen in den Medien gehören zum Handwerk der Medienmacher. In der Schlussphase des letzten Bundestagswahlkampfes 2002 wurde es besonders auffällig. Am Wahltag machte *Bild am Sonntag* mit einer „gezinkten“ Umfrage auf, um die Botschaft zu plazieren, es sei noch alles offen: Union und SPD wurden bei 37 Prozent verortet, FDP und BÜNDNIS 90/DIE GRÜNEN bei 8 Prozent, ein „Kanzlerkrimi“ wurde inszeniert.⁶ Das Blatt berief sich auf eine angebliche „Geheimumfrage“ von *Infratest*.

Das Institut dementierte prompt.

Das entscheidende Dementi folgte wenige Stunden später, durch das amtliche Endergebnis selbst. Zwar lagen die beiden großen Volksparteien mit 38,5 Prozent tatsächlich gleichauf, die PDS verpasste wie erwartet die Sperrklausel und FDP und „Grüne“ lagen fast gleichauf, aber eben nur fast. Unterschritt die FDP die angesagten 8 Prozent um 0,6 Punkte, so übersprangen die „Grünen“ diese Marke mit 0,6 Punkte.⁷

Das sind 0,6 Punkte zwischen den beiden kleinen Parteien, entscheidend zwischen „Himmel und Hölle“.

⁶ *Bild am Sonntag*, 22. September 2002, Titelblatt.

⁷ *Infratest dimap*, Wahlrend, Bundestags-, Landtags- und Europawahlen in Deutschland. Alle Ergebnisse seit 1946, Mai 2004, S. 155.

„Was würden Sie wählen, wenn am Sonntag Wahlen wären?“



- BamS veröffentlicht am Wahlsonntag eine angebliche Umfrage von „Infra-test“: Kopf an Kopf!
- „Infra-test“-Dementi vom 22. Sept. 2002: „Die Meldungen sind schlicht falsch!“

9./10. September 2005

Prof. Dr. Joachim Hofmann-Göttig:
Meinungsforschung

8

Demoskopen als Politikberater

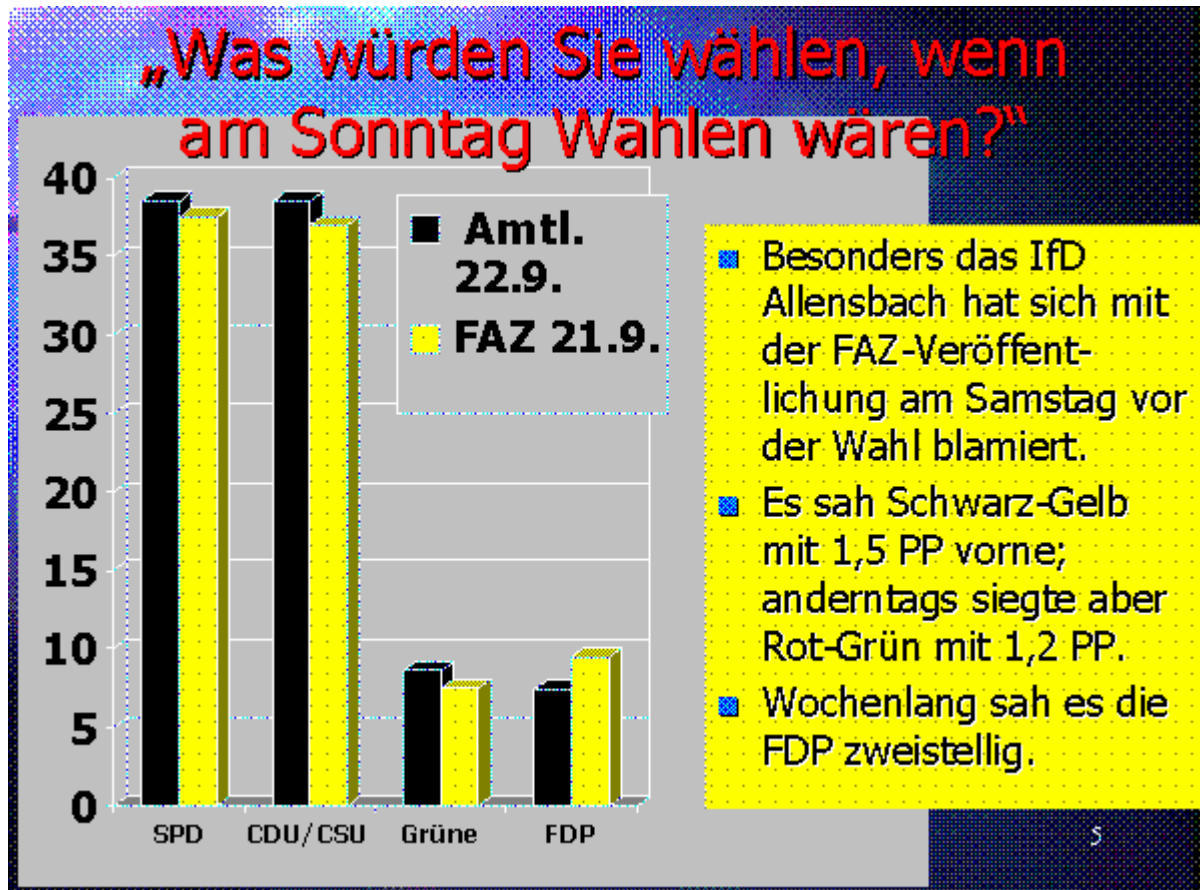
Demoskopie sind gefragte Experten/innen auf dem Stimmungsmarkt der Parteien, Wähler/innen und Medien. Wer Jahrzehnte lang Mehrheitsmeinungen erspürt, ist als Strategieberater/in von Parteien umschwärmt. Das galt viele Jahre für Elisabeth Noelle-Neumann und Renate Köcher vom Allensbacher Institut für Demoskopie, die transparent ihre Herzen auf dem rechten Fleck trugen. Ihre Freude über das demoskopische Hoch der Union ging so weit, dass sie schon Monate vor der letzten Bundestagswahl Schwarz-Gelb als die sicheren Sieger ausriefen, woran sie noch heute schwer zu „knabbern“ haben.⁸ Wochenlang sah das Institut die FDP im Ergebnis zweistellig. Und selbst einen Tag vor der Wahl veröffentlichte Allensbach noch einen Vorsprung von Schwarz-Gelb vor Rot-Grün von immerhin 1,5 Prozentpunkten.⁹ Das Dementi folgte anderntags durch das Wahlergebnis, da nämlich war Rot-Grün mit 1,2 Prozentpunkten vorn.¹⁰ Man wird sich ja mal täuschen dürfen, vor allem, wenn man sich der hohen Treffsicherheit der Prognosen bei Bundestagswahlen im Hause Allensbach sonst gern rühmt.¹¹

⁸ DER SPIEGEL, a.a.O., S. 26.

⁹ Institut für Demoskopie, Allensbach, in: FAZ vom 21. September 2002.

¹⁰ Forschungsgruppe Wahlen, Wahlergebnisse in Deutschland 1946 – 2004, Januar 2005, S. 100.

¹¹ Elisabeth Noelle-Neumann/Thomas Petersen, Alle, nicht jeder – Einführung in die Methoden der Demoskopie, München, 2. Aufl. 1998, S. 22 ff.; Elisabeth Noelle-Neumann, Der Beitrag der Wahlforschung zur Methoden-



© Joachim Hofmann-Göttig

Es ist für eine/n Wissenschaftler/in nicht schlimm, gelegentlich Fehler zu machen. Für Demoskopien, die sich als Politikberater/innen engagieren, gehört das zum nahezu unvermeidlichen Berufsrisiko. Unakzeptabel ist es aber, wenn man nicht sich selbst, sondern andere für Fehler verantwortlich macht. *Noelle-Neumann* hielt ihre damalige Minderheitenmeinung über die starke FDP auch Tage nach der Wahl aufrecht und bezeichnete die tatsächliche FDP-Schwäche als „die größte Überraschung der Bundestagswahl“.¹²

Ihre Kollegin *Renate Köcher* ging noch einen Schritt weiter. Nicht *Allensbach* hätte die FDP zu hoch eingeschätzt, sondern die anderen Institute zu niedrig und hätten damit angeblich das Klima beeinflusst, ließ sie die staunende Öffentlichkeit wissen.¹³ Die berüchtigte *Allensbacher* „Schweigespirale“ ließ grünen.

Mit dieser – hart formuliert – selbstgerechten Rechtfertigungsvariante stimmte sie freilich die Demoskopien-Zunft wenig fröhlich. „Abenteuerlich“ nannte dies der Chef der Forschungsgruppe Wahlen, *Matthias Jung*.¹⁴

entwicklung in der empirischen Sozialforschung, in: *Klein/Jagodziniski/Mochmann/Ohr* (Hrsg.), 50 Jahre empirische Wahlforschung in Deutschland, Wiesbaden 2000, S.61.

¹² *Elisabeth Noelle-Neumann* in: *FAZ* vom 25. September 2002.

¹³ *Renate Köcher* in: *DIE ZEIT* vom 26. September 2002.

¹⁴ *Matthias Jung*, Pressemitteilung der *Forschungsgruppe Wahlen* vom 26. September 2002.

Und auch die Union war im Nachhinein von den Gefälligkeitsanalysen nicht angetan, das habe den „Schaden für die Union“, so ihr Wahlkampfberater *Michael Spreng*, „nur erhöht“.¹⁵

Nach diesen Erfahrungen sind die Demoskopien, besonders die aus *Allensbach*, im diesjährigen Bundestagswahlkampf 2005 vorsichtiger geworden, Stimmungsbilder vor den Wahlen bereits für „Prognosen“ zu halten, obwohl „Wahlprognosen“, insbesondere die berühmten „18.00 Uhr-Prognosen“ der Fernsehanstalten nicht auf Stimmungszahlen vor den Wahlen, sondern auf „exit polls“¹⁶, also Befragungen der Urnenwähler, beruhen.

Aus der Deckung der Wissenschaft heraus wagte sich im laufenden Bundestagswahlkampf 2005 bisher nur einer: *Manfred Güllner*, der bekennende *Schröder-Fan* und Chef des Meinungsforschungsinstituts *Forsa*, ließ sich bereits Wochen vor der Bundestagswahl mit der These öffentlich zitieren, Schröders SPD habe eigentlich „keine Chance“ mehr. Selbst wenn alle Unentschlossenen nunmehr zur SPD fänden, wären mehr als 33 Prozent nicht drin.¹⁷

Polemisch formuliert ist zu sagen: Da staunt der Fachmann und der Laie wundert sich. Es ist nicht überliefert, wie der Angesprochene auf diese öffentlichen Analysen reagierte, ob er *Güllners* Empfehlung folgte und sich „in einen Trance-Zustand versetzte“.

Die Botschaft beider Erfahrungen mit Demoskopien in den Bundestagswahlkämpfen 2002 und 2005 ist eindeutig: Pluralität ist unverzichtbar. Solange interesselgeleitete Analysen durch demoskopische Vielfalt validiert werden können und werden, sind die Risiken im Meinungsforschungsgeschäft zu beherrschen. Eine Konzentration auf wenige Anbieter/innen der Dienstleistung Meinungsforschung oder auch auf wenige Auftraggeber wäre demokratie-theoretisch gefährlich.

Der Wert der „Sonntagsfrage“

Fest steht: Meinungsforschung in Wahlkämpfen reduziert sich keinesfalls auf die „Sonntagsfrage“. Aber das öffentliche Interesse an Demoskopie ist sehr stark auf diese eine Frage hin konzentriert. Die Ergebnisse der „Sonntagsfrage“ können – so haben wir gesehen - in Medien kräftig unterschiedlich interpretiert, sie können manipuliert, sie können selbst von Sozialwissenschaftlern fehlerhaft, sogar manipulativ eingesetzt werden. Und dennoch sind sie in einer modernen Medienwelt unverzichtbar geworden.

¹⁵ *Michael Spreng*, *DIE ZEIT* vom 10. Oktober 2002.

¹⁶ *Dieter Roth*, *Empirische Wahlforschung*, Opladen 1998, S. 82 f.

¹⁷ *Manfred Güllner*, in: *DIE ZEIT* Nr. 33 vom 11. August 2005.

Sie beleben das Interesse am Wahlkampf. Sie dienen vielen „Nicht-Stamm-Wählern“ als Orientierungshilfe, vor allem im Hinblick auf die Fünf-Prozent-Klausel. Das könnte im laufenden Wahlkampf sehr anschaulich anhand der Berichterstattung über die neue „Linkspartei“ gezeigt werden, an deren Einzug in den nächsten Bundestag angesichts der Umfragedaten öffentlich niemand zweifelt.

Je geringer der Anteil von Stammwählern, je kurzfristiger sich Wahlentscheidungen bilden, je größer die Variabilität der Entscheidung, desto größer ist die Bedeutung der Meinungsforschung, insbesondere der Sonntagsfrage für den Wahlentscheid. Das ist Fakt, auch wenn die herrschende Lehre der Wahlforschung dies – so wird in einem kürzlich erschienenen Standardwerk der Wahlforschung berichtet – noch nicht akzeptieren will.¹⁸

Wie verlässlich ist die Sonntagsfrage?

Wenn für Parteistrategen, für die Medien, aber eben auch für die Wählerschaft selbst die Ergebnisse der Sonntagsfrage immer wichtiger werden, dann bedarf es einer Einschätzung ihrer Verlässlichkeit. Das ist entscheidend, um adäquat mit diesem Instrument umgehen zu können.

Dabei steht außer Zweifel, dass die „Sonntagsfrage“ immer verlässlicher wird, je näher der Sonntag rückt, an dem tatsächlich gewählt wird, denn dann

- ist die Sonntagsfrage nicht mehr hypothetisch,
- haben sich immer mehr Interviewte mit der Frage tatsächlich beschäftigt,
- verringert sich ein Stimmungsumschwung durch aktuelle Ereignisse.

Sehr zeitnahe Sonntagsfragen vor Wahlen lassen eine zufrieden stellende Verlässlichkeit erhoffen. Genügend Probleme bleiben dennoch:

- Weiterhin hat es die Meinungsforschung mit wenig stabilen, festen Absichten zu tun,
- Gibt es Unentschlossene und Nichtwähler/innen, die sich möglicherweise nicht proportional über die Parteien verteilen,
- Ist die Aussagebereitschaft über immerhin geschützte Daten (Wahlgeheimnis) indifferent,
- Ist die Bekennerfreude zu gesellschaftlich diskriminierten Wahlentscheidungen (z.B. rechtsextremen Parteien) nicht ausgeprägt,
- Und über all dies haben wir es mit den üblichen, wahrscheinlichkeitstheoretischen Fehlertoleranzen bei Schlüssen von Stichproben auf die Grundgesamtheit zu tun. Allein dieser Faktor wird in der Öffentlichkeit gern übersehen.

¹⁸ Jürgen Falter/Harald Schoen (Hrsg.), Handbuch Wahlforschung, Wiesbaden, Juli 2005; S. 615 f. Allerdings differenzierter: Frank Brettschneider, Demoskopie im Wahlkampf – Leitstern oder Irrlicht, in: Klein u.a. (Hrsg.), a.a.O., S. 477 ff.

Wie verlässlich können also zeitnah publizierte Sonntagsfragen das tatsächliche Wahlergebnis prognostizieren?

Die Sonntagsfrage zur Bundestagswahl 2005

Werfen wir also einen Blick auf die aktuell veröffentlichten Vorwahlauffragen.¹⁹ Ein Vergleich der kürzlich 2005 publizierten Stimmungsbilder zwischen *Allensbach*, *Emnid*, *Forsa*, *Forschungsgruppe Wahlen*, *GMS* und *Infratest* sieht durchgängig eine Regierungsmehrheit für Schwarz-Gelb, zumindest der Mandate (*Emnid*, *Infratest dimap*), überwiegend auch der Prozentanteile (*Allensbach*, *Forsa*, *Forschungsgruppe Wahlen*, *GMS*).²⁰

Institut	<u>Allensbach</u>	<u>Emnid</u>	<u>Forsa</u>	<u>Forsch'gr. Wahlen</u>	<u>GMS</u>	<u>Infratest-dimap</u>
Veröffentlichung	02.09.2005	07.09.2005	03.09.2005	02.09.2005	05.09.2005	01.09.2005
CDU/CSU	41,7 %	42 %	43 %	43 %	42 %	43 %
SPD	29,6 %	32 %	31 %	32 %	32 %	32 %
GRÜNE	8,1 %	7 %	7 %	7 %	8 %	7 %
FDP	8,0 %	7 %	7 %	7 %	7 %	6 %
Die Linke. PDS	9,7 %	9 %	9 %	8 %	8 %	9 %
Sonstige	2,9 %	3 %	3 %	3 %	3 %	3 %

Ist angesichts dieser durchgängigen, eindeutigen demoskopischen Lage so kurz vor einer Wahl der tatsächliche Wahlerfolg von Schwarz-Gelb nur noch Formsache?

Wie verlässlich ist die Sonntagsfrage?

¹⁹ Als Quelle für die folgende Analyse dient ausschließlich (mit einer Ausnahme, vgl. Fußnote 19) die sehr verdienstvolle und sehr aktuelle Zusammenstellung von *Wilko Zicht* in: www.wahlrecht.de/Umfragen; im Internet am 26. August 2005, am 31. August für die Bundestagswahl 2005. Die Daten dieser Quelle konnten nicht überprüft werden; sie liegen laut Impressum in der Verantwortung des Autors, der auf Nachfrage die Datenbank als „Hobbytätigkeit“ bezeichnete, ohne im Übrigen über seinen beruflichen Hintergrund weitere Informationen zu geben. Das ist bedauerlich, weil Internetrecherchen natürlich auch Quellenprüfungen verlangen. Bei der Datenübertragung für die vorliegenden Matrizen wurden daher zumindest Plausibilitätsprüfungen vorgenommen und die Zeilen-Prozentuierung auf 100 überprüft. Dabei konnte z.B. festgestellt werden, dass die „Sonstigen“ beim amtlichen Endergebnis bei der Landtagswahl in Mecklenburg-Vorpommern nicht wie ausgewiesen auf 6,8 Prozent, sondern nur auf 4,3 Prozent kamen. Für Rheinland-Pfalz fehlen die Vergleichswerte des amtlichen Endergebnisses der LTW 2001 (sie wurden aus *Infratest dimap*, a.a.O., S. 64 entnommen). Aber insgesamt machte die Datenbank einen zuverlässigen Eindruck.

²⁰ Datenbasis: www.wahlrecht.de/Umfragen, im Internet am 7. September 2005 (die Kontrolle ergab, dass die aktuelle Datenbank täglich gepflegt wird).

Während die Meinungsforschungsinstitute die Qualität ihrer förmlichen Prognosen regelmäßig evaluieren, wurde der Frage nach der Verlässlichkeit von Vorwahlumfragen bislang nach hiesiger Kenntnis nur vereinzelt und nicht systematisch nachgegangen.²¹

Test der Prognosefähigkeit der Sonntagsfrage

Schauen wir uns einmal die jeweils in den letzten zehn Tagen vor dem Wahltag publizierten Stimmungsmessungen für die letzten Landtagswahlen in Deutschland, die jüngste Europawahl und die letzte Bundestagswahl an.

Vergleichen wir die jeweiligen Messungen mit dem amtlichen Endergebnis, berechnen die jeweiligen Abweichungen, die Summe der Abweichungen, die durchschnittlichen Abweichungen bei der Summe der Umfragen pro Wahl und deren Abweichungssumme und werfen für die 19²² zu untersuchenden Wahlen der Jahre 2001 - 2005 einen Blick auf die Höchstwerte der Abweichungen. Dann wissen wir genauer, was exakt Stimmungsbilder immerhin sehr kurzfristig – nur wenige Tage - vor einer Wahl wert sind.²³

Die Einbeziehung nur jener Umfragen, die in den letzten zehn Tagen vor der jeweiligen Wahl publiziert wurden, war eine strikt einzuhaltende Vorgabe, auch wenn es gelegentlich verführerisch wirkte, ein oder zwei Tage weiter zurück zu gehen, um damit zahlreiche weitere Umfragen in die Analyse einbeziehen zu können. Aber der Begriff „zeitnahe Umfrage“ zur Wahl musste definiert und strikt eingehalten werden, um objektiv und nicht tendenziös an die Untersuchungsfrage herangehen zu wollen.²⁴

Zu diesen 19 Wahlen liegen 47 Umfragen innerhalb des genannten 10-Tage-Zeitraumes vor. Diese werden hinsichtlich ihrer Prognosefähigkeit getestet (evaluiert).

²¹ Vgl. Grabmal für die unbekannte Fehlprognose, in: www.wahlprognosen-info.de; im Internet am 25. August 2005. Wegen der ausgeprägten unwissenschaftlichen Polemik kann man diese Quelle für wissenschaftliche Untersuchungen kaum heranziehen, lediglich als Datenquelle für die Umfragen zur Bundestagswahl 2002 wurde sie verwendet.

²² Die Zahl 19 ergibt sich daraus, dass alle Länder in die Analyse einbezogen werden sollten. In dem Untersuchungszeitraum 2001 – 2005 fanden aber in Hamburg zwei Bürgerschaftswahlen statt, ferner war die Europa- und die Bundestagswahl 2002 einzubeziehen.

²³ Bewusst wird darauf verzichtet, mit abstrakten Streumassen wie Varianzen oder Standardabweichungen zu arbeiten, weil die Orientierung am arithmetischen Mittelwert gegenstandsadäquater erscheint. Hilfreich für die Entscheidung über das gewählte Vorgehen waren Gespräche mit dem Chef des Münchner Meinungsforschungsinstituts „polis“ Walter Ruhland.

²⁴ In den Untersuchungs-Tabellen wird soweit möglich auch der jeweilige Befragungszeitraum dargestellt. Er bleibt aber für die Frage der Einbeziehung in die vorliegende Untersuchung unbeachtlich. Entscheidend hierfür ist ausschließlich das Veröffentlichungsdatum.

Wahlkalender: Deutschland 2001 -2005

Europa-, Bundestags- und Landtagswahlen

Nr.	Wahl in ...	Am ...	Zahl der getesteten Umfragen	Summe der Fehler- toleranz	Summe des mittleren Fehlers
1	Nordrhein-Westfalen	22.05.2005	4	10,3	6,0
2	Schleswig-Holstein	20.02.2005	4	11,7	8,5
3	Brandenburg	19.09.2004	2	13,6	17,4
4	Sachsen	19.09.2004	2	13,7	17,1
5	Saarland	05.09.2004	1	10,6	9,8
6	Thüringen	13.06.2004	3	12,3	19,9
7	Europawahl	13.06.2004	3	12,5	18,7
8	Hamburg I	29.02.2004	5	12,2	5,7
9	Bayern	21.09.2003	1	10,9	4,2
10	Bremen	25.05.2003	1	12,4	15,4
11	Niedersachsen	02.02.2003	2	10,4	4,3
12	Hessen	02.02.2003	2	10,7	3,4
13	Mecklenburg-Vorpommern	22.09.2002	2	11,7	7,5
14	Bundestagswahl	22.09.2002	5	11,8	6,2
15	Sachsen-Anhalt	21.04.2002	2	13,3	16,7
16	Berlin	21.10.2001	3	13,1	14,7
17	Hamburg II	23.09.2001	2	13,3	11,0
18	Baden-Württemberg	25.03.2001	2	11,6	8,9
19	Rheinland-Pfalz	25.03.2001	1	10,8	4,1
Summe			47	11,9	10,5

© Joachim Hofmann-Göttig

Dazu werden die Prozentpunkt-Differenzen der jeweiligen Umfrage vom amtlichen Endergebnis der wenige Tage später stattgefundenen Wahl berechnet für jede einzelne Partei und für die Gesamtheit der Parteien bei einer Wahl. Liegen zu einer Wahl mehr als eine Umfrage vor, was bei 15 der 19 untersuchten Wahlen der Fall ist, dann werden ergänzend die mittleren Differenzen errechnet, die in den Matrizen als „mittlere Differenz“ der einschlägigen Vorwahlumfragen ausgewiesen sind.²⁵

In einem weiteren Untersuchungsschritt kann auf dieses Merkmal ein Benchmarking der beteiligten Institute vorgenommen werden. Letzteres bezieht sich ausdrücklich nur auf die Prognosetauglichkeit von zeitnahen Vorwahlumfragen, ausdrücklich nicht auf die *Exit Polls* gestützten „18.00 Uhr-Prognosen“.

²⁵ Vgl. im Einzelnen die Matrizen der 19 Wahlen im Anhang.

Wann kann von einer „prognosetauglichen Sonntagsfrage“ gesprochen werden?

Umfragen messen stets nur Stimmungsbilder zum Testzeitpunkt. Ist der Zeitpunkt allerdings unmittelbar vor einer Wahl, geht die Öffentlichkeit davon aus, dass diese Umfrage dem tatsächlichen Ergebnis relativ nahe kommt.

Allerdings ist der Schluss von einer Stichprobe - in der Regel in etwa 1.000 Befragte - auf die Grundgesamtheit der etwa 61,5 Millionen deutschen Wahlberechtigten auch bei bester Untersuchungsanlage stets mit einem Fehlerrisiko behaftet. Statistisch gesehen bedeutet die Berichterstattung eines Merkmales, dass die Forscher mit einer Wahrscheinlichkeit von 95 Prozent (= übliches Signifikanzniveau von 5 Prozent) den tatsächlichen Wert innerhalb eines Vertrauensintervalls von +/- x Prozentpunkten erfassen.

Lediglich das Institut *Infratest dimap* hat es sich zur begrüßenswerten Gewohnheit werden lassen, auf diesen Tatbestand durch regelmäßigen Abdruck der „Fehlertoleranztabelle“ aufmerksam zu machen.²⁶

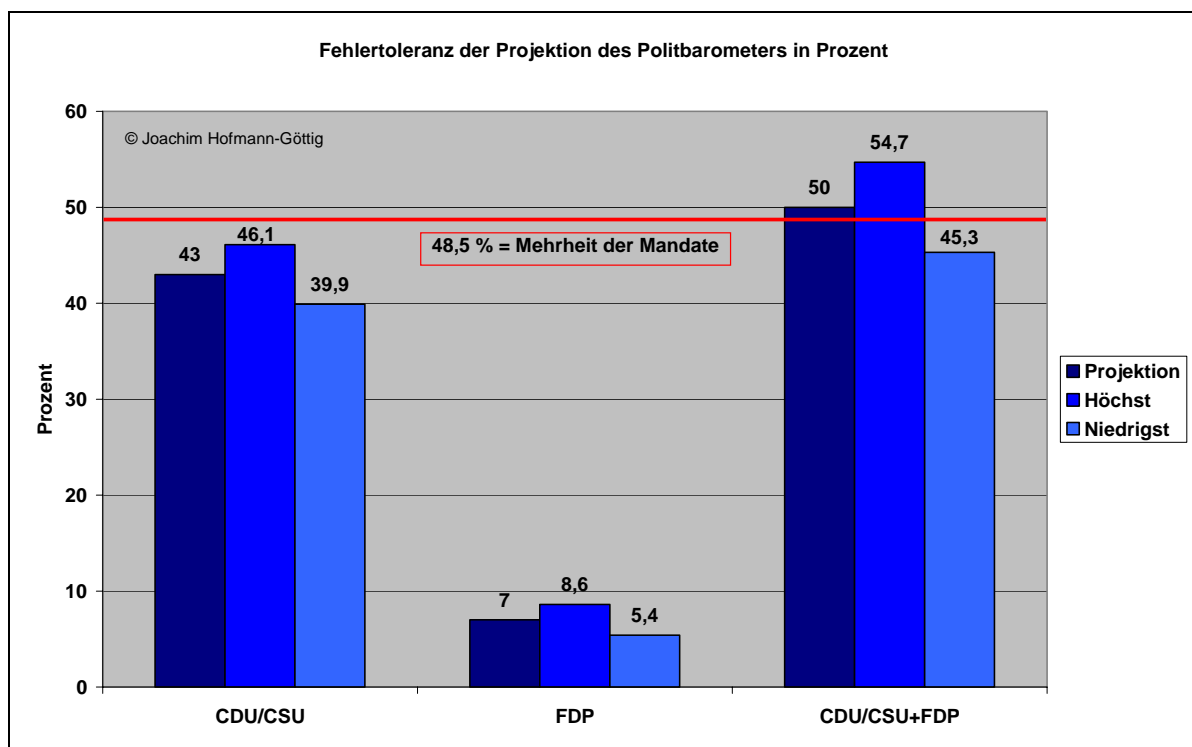
Fehlertoleranztabelle																				
Lesebeispiel:		In einer Stichprobe von 2500 Personen aus einer Grundgesamtheit von N= 61.500.000 Wahlberechtigten sei ein Anteil von 10% für ein Merkmal ermittelt worden. Dann liegt der wahre Wert der Grundgesamtheit mit 95%iger Wahrscheinlichkeit im Intervall 10% ± 1,2%. Die Tabelle basiert auf der Formel für das Konfidenzintervall bei Zufallsstichproben ohne Zurücklegen: $s(p) = t \cdot \sqrt{p(1-p)/(n-1)\sqrt{1-n/N}}$ In Feldern mit '---' beträgt das Konfidenzintervall mehr als die Hälfte des Anteils und obige Formel ist nicht mehr anwendbar.																		
Größe der Stichprobe n=	Anteilswerte in der Stichprobe																			
	1%	2%	3%	4%	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%	60%	70%	80%	90%	95%	
100	---	---	---	---	---	7,1%	8,0%	8,6%	9,1%	9,5%	9,8%	9,9%	10,0%	9,8%	9,1%	8,0%	---	---	---	
200	---	---	---	---	---	4,2%	5,0%	5,6%	6,1%	6,4%	6,7%	6,8%	7,0%	7,0%	6,8%	6,4%	5,6%	4,2%	---	
300	---	---	---	---	2,5%	3,4%	4,1%	4,6%	4,9%	5,2%	5,4%	5,6%	5,7%	5,7%	5,6%	5,2%	4,6%	3,4%	2,5%	
400	---	---	---	1,9%	2,1%	3,0%	3,5%	3,9%	4,3%	4,5%	4,7%	4,8%	4,9%	4,9%	4,8%	4,5%	3,9%	3,0%	2,1%	
500	---	---	---	1,7%	1,9%	2,6%	3,1%	3,5%	3,8%	4,0%	4,2%	4,3%	4,4%	4,4%	4,3%	4,0%	3,5%	2,6%	1,9%	
600	---	---	1,4%	1,6%	1,7%	2,4%	2,9%	3,2%	3,5%	3,7%	3,8%	3,9%	4,0%	4,0%	3,9%	3,7%	3,2%	2,4%	1,7%	
700	---	---	1,3%	1,5%	1,6%	2,2%	2,7%	3,0%	3,2%	3,4%	3,5%	3,6%	3,7%	3,7%	3,6%	3,4%	3,0%	2,2%	1,6%	
800	---	1,0%	1,2%	1,4%	1,5%	2,1%	2,5%	2,8%	3,0%	3,2%	3,3%	3,4%	3,5%	3,5%	3,4%	3,2%	2,8%	2,1%	1,5%	
900	---	0,9%	1,1%	1,3%	1,4%	2,0%	2,3%	2,6%	2,8%	3,0%	3,1%	3,2%	3,3%	3,3%	3,2%	3,0%	2,6%	2,0%	1,4%	
1000	---	0,9%	1,1%	1,2%	1,4%	1,9%	2,2%	2,5%	2,7%	2,8%	3,0%	3,0%	3,1%	3,1%	3,0%	2,8%	2,5%	1,9%	1,4%	
1100	---	0,8%	1,0%	1,2%	1,3%	1,8%	2,1%	2,4%	2,6%	2,7%	2,8%	2,9%	2,9%	3,0%	2,9%	2,7%	2,4%	1,8%	1,3%	
1200	---	0,8%	1,0%	1,1%	1,2%	1,7%	2,0%	2,3%	2,5%	2,6%	2,7%	2,8%	2,8%	2,8%	2,8%	2,6%	2,3%	1,7%	1,2%	
1300	---	0,8%	0,9%	1,1%	1,2%	1,6%	1,9%	2,2%	2,4%	2,5%	2,6%	2,7%	2,7%	2,7%	2,7%	2,5%	2,2%	1,6%	1,2%	
1400	---	0,7%	0,9%	1,0%	1,1%	1,6%	1,9%	2,1%	2,3%	2,4%	2,5%	2,6%	2,6%	2,6%	2,6%	2,4%	2,1%	1,6%	1,1%	
1500	---	0,7%	0,9%	1,0%	1,1%	1,5%	1,8%	2,0%	2,2%	2,3%	2,4%	2,5%	2,5%	2,5%	2,5%	2,3%	2,0%	1,5%	1,1%	
1600	0,5%	0,7%	0,8%	0,9%	1,0%	1,1%	1,5%	1,8%	2,0%	2,1%	2,2%	2,3%	2,4%	2,4%	2,5%	2,4%	2,2%	2,0%	1,5%	1,1%
1700	0,5%	0,7%	0,8%	0,9%	1,0%	1,4%	1,7%	1,9%	2,1%	2,2%	2,3%	2,3%	2,4%	2,4%	2,3%	2,2%	1,9%	1,4%	1,0%	
1800	0,5%	0,8%	0,8%	0,9%	1,0%	1,4%	1,7%	1,8%	2,0%	2,1%	2,2%	2,3%	2,3%	2,3%	2,3%	2,1%	1,8%	1,4%	1,0%	
1900	0,4%	0,8%	0,8%	0,9%	1,0%	1,4%	1,6%	1,8%	1,9%	2,1%	2,1%	2,2%	2,2%	2,3%	2,2%	2,1%	1,8%	1,4%	1,0%	
2000	0,4%	0,8%	0,7%	0,9%	1,0%	1,3%	1,6%	1,8%	1,9%	2,0%	2,1%	2,1%	2,2%	2,2%	2,1%	2,0%	1,8%	1,3%	1,0%	
2500	0,4%	0,5%	0,7%	0,8%	0,9%	1,2%	1,4%	1,6%	1,7%	1,8%	1,9%	1,9%	2,0%	2,0%	1,9%	1,8%	1,6%	1,2%	0,9%	
3000	0,4%	0,5%	0,6%	0,7%	0,8%	1,1%	1,3%	1,4%	1,6%	1,6%	1,7%	1,8%	1,8%	1,8%	1,8%	1,6%	1,4%	1,1%	0,8%	
4000	0,3%	0,4%	0,5%	0,6%	0,7%	0,9%	1,1%	1,2%	1,3%	1,4%	1,5%	1,5%	1,6%	1,6%	1,6%	1,4%	1,2%	0,9%	0,7%	
5000	0,3%	0,4%	0,5%	0,5%	0,6%	0,8%	1,0%	1,1%	1,2%	1,3%	1,3%	1,4%	1,4%	1,4%	1,4%	1,3%	1,1%	0,8%	0,6%	
6000	0,3%	0,4%	0,4%	0,5%	0,6%	0,8%	0,9%	1,0%	1,1%	1,2%	1,2%	1,3%	1,3%	1,3%	1,2%	1,2%	1,0%	0,8%	0,6%	
8000	0,2%	0,3%	0,4%	0,4%	0,5%	0,7%	0,8%	0,9%	0,9%	1,0%	1,0%	1,1%	1,1%	1,1%	1,1%	1,0%	0,9%	0,7%	0,5%	
10000	0,2%	0,3%	0,3%	0,4%	0,4%	0,6%	0,7%	0,8%	0,8%	0,9%	0,9%	1,0%	1,0%	1,0%	1,0%	0,9%	0,8%	0,6%	0,4%	
11000	0,2%	0,3%	0,3%	0,4%	0,4%	0,6%	0,7%	0,7%	0,8%	0,9%	0,9%	0,9%	0,9%	0,9%	0,9%	0,9%	0,7%	0,6%	0,4%	
20000	0,1%	0,2%	0,2%	0,3%	0,3%	0,4%	0,5%	0,6%	0,6%	0,6%	0,7%	0,7%	0,7%	0,7%	0,7%	0,6%	0,6%	0,4%	0,3%	
25000	0,1%	0,2%	0,2%	0,2%	0,3%	0,4%	0,4%	0,5%	0,5%	0,6%	0,6%	0,6%	0,6%	0,6%	0,6%	0,6%	0,5%	0,4%	0,3%	
30000	0,1%	0,2%	0,2%	0,2%	0,2%	0,3%	0,4%	0,5%	0,5%	0,5%	0,5%	0,6%	0,6%	0,6%	0,6%	0,5%	0,5%	0,3%	0,2%	

Infratest dimap

²⁶ *Infratest dimap*, Deutschland-Trend, August 2005, Anhang, letzte Seite. Auf diese Fehlertoleranztabelle stützt sich die folgende Analyse. Sie ist auch Grundlage für die Matrizen. Auch „*Forsa*“ benennt Fehlertoleranzen (vgl. „*Stern*“ Nr. 36 v. 01.09.2005, S. 20, S. 38 f., S. 40).

Wenn also in einer Befragung von $N = 1.000$, z.B. 50 Prozent angeben, sie wollten CDU/CSU wählen, dann liegt der Wert in der Grundgesamtheit der 61,5 Millionen wahlberechtigten Deutschen mit einer Wahrscheinlichkeit von 95 Prozent $\pm 3,1$ Prozentpunkte, also innerhalb des Konfidenzintervalls von 46,9 bis 53,1 Prozent. Noch schlimmer ist diese statistische Unsicherheit für die kleinen Parteien. Wenn etwa die FDP bei 6 Prozent gemessen wird, dann liegt die Fehlertoleranz bei $\pm 1,5$ Prozentpunkten.

Diese Fehlertoleranzen können an einem Realbeispiel, der Projektion der *Forschungsgruppe Wahlen* demonstriert werden.²⁷



Projizierte 50 Prozent für Schwarz-Gelb können eine sehr auskömmliche absolute Mehrheit mit 54,7 Prozent bedeuten, aber eben auch mit 45,3 Prozent Nichterreicherung der Mehrheit der Mandate (bei 3 Prozent für Sonstige: 48,5 Prozent).

Mit diesen Fehlertoleranzen kann sich die Zunft der politischen Meinungsforscher natürlich nicht zufrieden geben. Zu ihrem Handwerkszeug gehören viele Instrumente, die das Fehlerrisiko unter diese Margen drücken sollen:

Die wichtigsten Maßnahmen sind die Trendanalyse und die Kumulierung von mehreren Testzeitpunkten. Je häufiger im Feld geforscht wird und je größer die Stichprobe über den üblichen $N=1.000$ ist, desto geringer sind die zufälligen Fehler.

²⁷ Datenbasis: Forschungsgruppe Wahlen, ZDF-Politbarometer, September I/05..

So haben die *Forschungsgruppe Wahlen* mit dem „*Politbarometer*“ für das ZDF und *Infratest dimap* mit dem „*Deutschland-Trend*“ für die ARD unmittelbar vor der Bundestagswahl ihren gewöhnlich monatlichen Rhythmus der Erhebungen verkürzt, *Forsa* ist mindestens einmal pro Woche im Feld.²⁸

Dazu kommt, dass Rohdaten der Stichproben erfahrungsgestützt bearbeitet werden, denn hinter den Computern stehen erfahrene Forscher/innen, die ihr Handwerk über Jahrzehnte solide gelernt haben.

Deshalb ist eine untersuchungsmethodische Entscheidung grundsätzlich eher demoskopiefreundlich, wenn wir von unmittelbar vor Wahlen veröffentlichten Umfragen erwarten wollen, dass sie sich wenigstens im Rahmen der Fehlertoleranztafel (für Stichproben von $N=1.000$) bewegen.

Demoskopie nur Kaffee-Satz-Leserei?

Für die einen ist Demoskopie „Kaffee-Satz-Leserei“, für die anderen sind Vorwahlumfragen fast der Ersatz des tatsächlichen Wahlaktes.

Wir wollen die Frage objektiv beantworten:

Wie nah sind die zeitnah veröffentlichten Umfragen am amtlichen Endergebnis?

Sind die Abweichungen innerhalb der statistischen Fehlertoleranz?

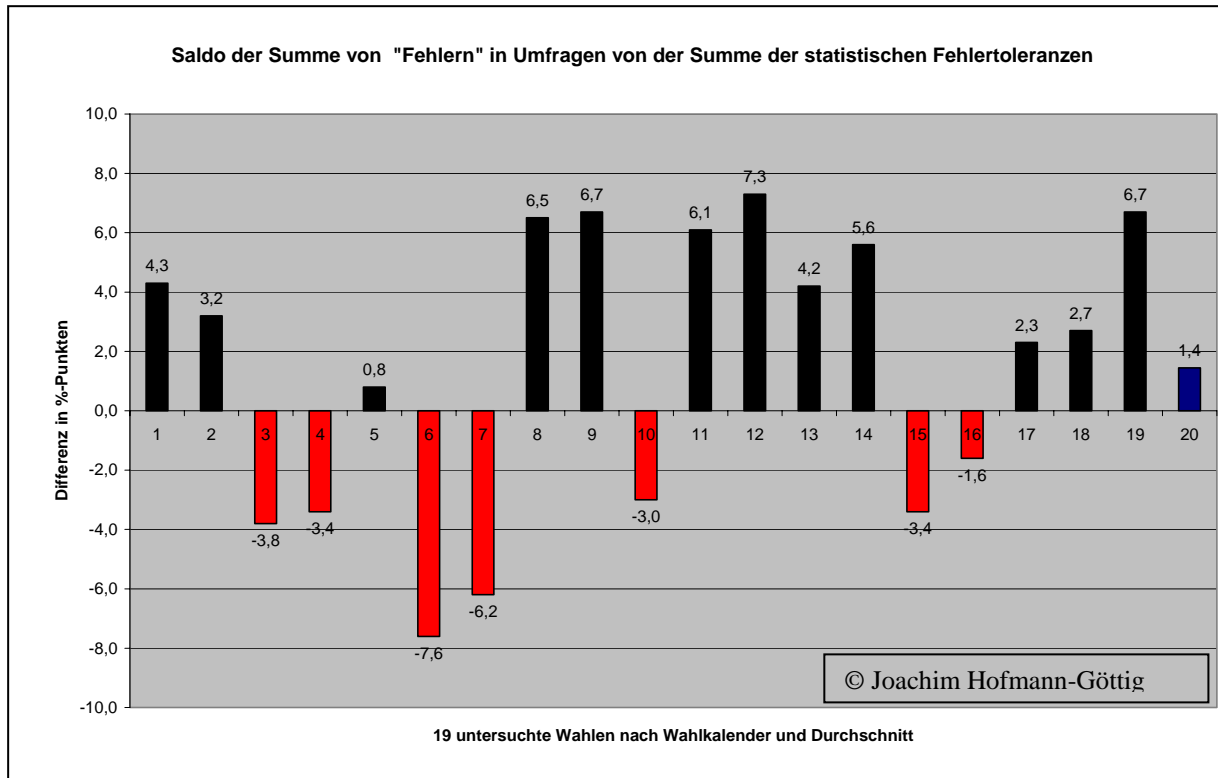
Um diese Fragen beantworten zu können wurde für jedes gemeldete Parteiresultat der 19 untersuchten Wahlen die statistische Fehlertoleranz berichtet und summiert.

Der Mittelwert des summierten tolerablen Fehlers liegt bei 11,9 Prozentpunkten. Der Mittelwert des tatsächlichen summierten Fehlers der 47 untersuchten Umfragen zu den 19 Wahlen liegt bei 10,5 Prozentpunkten.

Das heißt, generell liegen die Abweichungen der Umfragen knapp, aber deutlich innerhalb des statistisch tolerablen Bereichs. Dieser Befund lässt für eine ganz generelle Demoskopieschelte im Hinblick auf ihre Prognosefähigkeit keinen Raum.

Gehen wir in der Analyse von der Gesamtheit der Wahlen eine Ebene tiefer in die 19 Einzelwahlen.

²⁸ *Manfred Güllner*, Methodische Entwicklungen in der Empirischen Wahlforschung, in: *Klein u.a.* (Hrsg.), a.a.O., S.564 ff.



Nr.	Wahl in ...	Am ...
1	Nordrhein-Westfalen	22.05.2005
2	Schleswig-Holstein	20.02.2005
3	Brandenburg	19.09.2004
4	Sachsen	19.09.2004
5	Saarland	05.09.2004
6	Thüringen	13.06.2004
7	Europawahl	13.06.2004
8	Hamburg I	29.02.2004
9	Bayern	21.09.2003
10	Bremen	25.05.2003
11	Niedersachsen	02.02.2003
12	Hessen	02.02.2003
13	Mecklenburg-Vorpommern	22.09.2002
14	Bundestagswahl	22.09.2002
15	Sachsen-Anhalt	21.04.2002
16	Berlin	21.10.2001
17	Hamburg II	23.09.2001
18	Baden-Württemberg	25.03.2001
19	Rheinland-Pfalz	25.03.2001
Summe		

Wir sehen, dass sich bei 12 Wahlen, also der großen Mehrheit, die Summe des mittleren Fehlers im Toleranzbereich bewegt. Das heißt, bei knapp 2/3 der analysierten Vorwahlumfragen hielten sich die Abweichungen der erhobenen Stimmungsbilder unmittelbar vor Wahlen im vertretbaren Bereich. Das Ergebnis der Evaluierung ist zunächst einmal grundsätzlich positiv.

Bei sieben Wahlen müssen allerdings „*Fehlprognosen*“ diagnostiziert werden: Bei den Landtagswahlen in Brandenburg (- 3,8 Prozentpunkte), Sachsen (- 3,4 PP), Berlin (- 1,6 PP), Sachsen-Anhalt (- 3,4 PP) und Bremen (- 3,0 PP) sind die Summen der Fehler deutlich über der Toleranz. Bei der Europawahl (- 6,2 Prozentpunkte) und der Landtagswahl in Thüringen (-7,6 Prozentpunkte) müssen sogar sehr krasse Abweichungen festgestellt werden.

Gehen wir also wiederum eine Ebene tiefer und betrachten die mittleren Abweichungen der Gesamtheit der Umfragen pro Wahl zu den einzelnen Parteien.

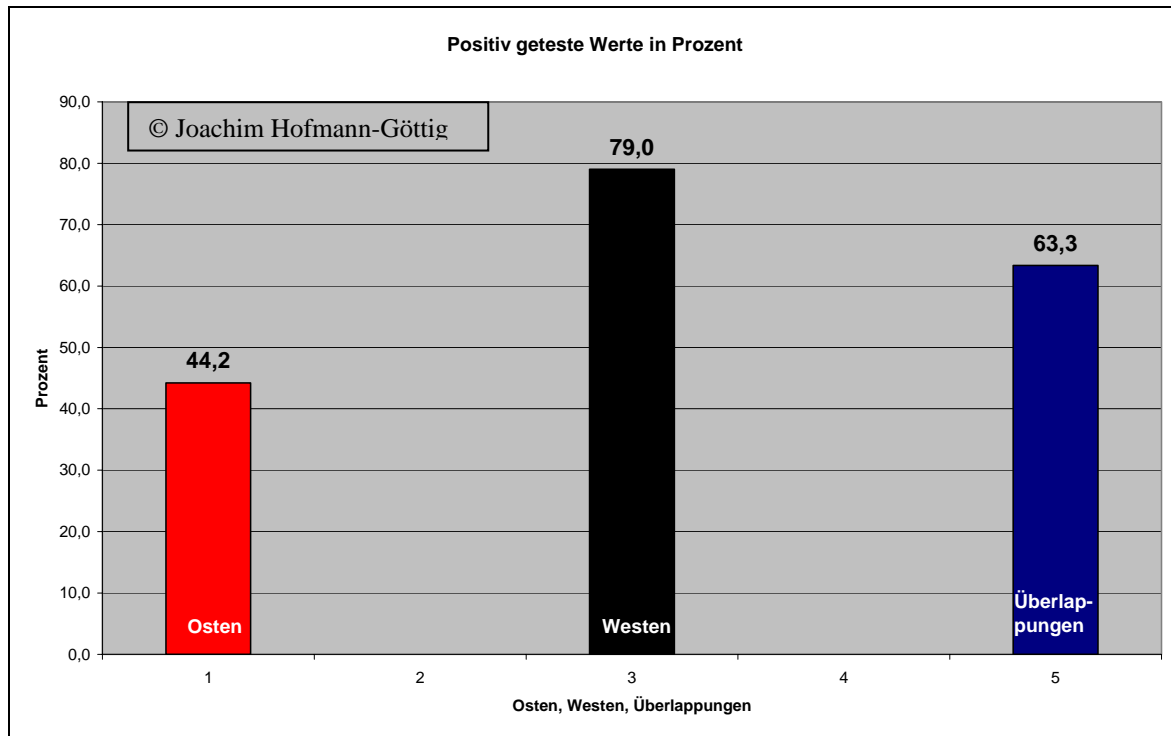
Auch hier dominiert die positive Seite. In 81 Fällen haben die Umfrageinstitute durchschnittlich die Parteianteile mit tolerablen Fehlern gemessen (d.h. 70,4 Prozent der Umfragen lagen im tolerablen Bereich). In 34 Fällen (=29,6 Prozent) gelang dies allerdings nicht.

Es könnte vermutet werden, dass die Umfrageinstitute besonders im Osten Schwierigkeiten haben, prognosefähige Stimmungen zu messen.

Diese Hypothese kann verifiziert werden. 15 Fälle der 34 Abweichungen waren im Osten zu verzeichnen, nur 12 im Westen (7 Überlagerungen = Bundestags- und Europawahlen, Berlin-Wahl). Es sind in den Matrizen die relativen Bilanzen pro Wahl (Anteil der positiv getesteten Werte in Prozent) ausgewiesen. Diese Daten können nach Gebietseinheiten (Osten, Westen, Überlappungen) gruppiert und arithmetisch gemittelt werden.

Während im Westen durchschnittlich 79,0 Prozent der Daten positiv getestet wurden, ist im Osten nur eine Minderheit (44,2 Prozent) positiv getestet. Die Überlappungen (Europawahl, Bundestagswahl, Wahl zum Abgeordnetenhaus in Berlin) liegen mit 63,3 Prozent in der Mitte.

Im Osten sind die Probleme mit der Demoskopie also erkennbar ausgeprägter. Die Probleme beschränken sich aber keineswegs auf den Osten Deutschlands.



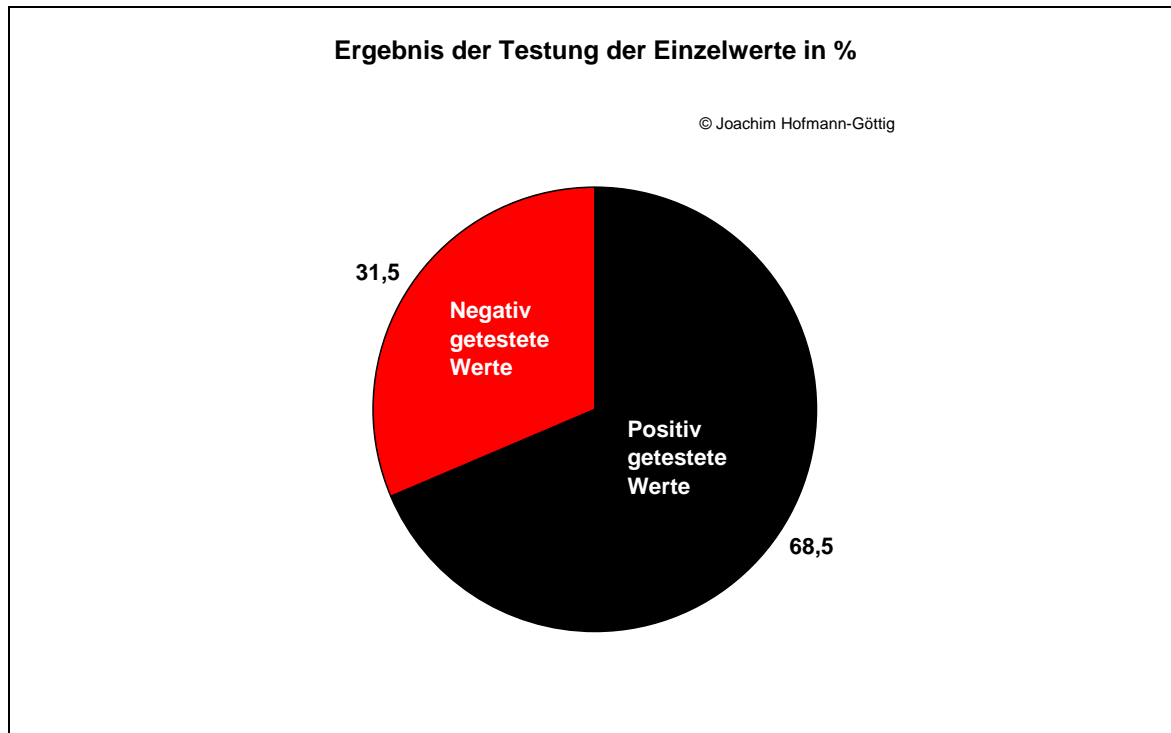
Es könnte ferner die Hypothese aufgestellt werden, dass sich die Probleme in Stadtstaaten im Vergleich zu Flächenstaaten wegen der höheren Wählermobilität massieren.

Die Hypothese kann falsifiziert werden: In Hamburg, Bremen und Berlin werden die Parteien durchschnittlich in 19 Fällen tolerabel, nur in acht Fällen falsch prognostiziert. Auch diese werden anhand der Einzelwerte geprüft. Die Einzelwerte zeigen, dass es keinen Zusammenhang zwischen Genauigkeit der Vorhersagen durch Umfragen und Dichte des Wohngebietes gibt. Teilen wir die Umfragen nach Stadtstaaten, Flächenländern und überlappende Wahlen (Bundestags- und Europawahlen), so sehen wir in allen drei Kategorien etwa 67 Prozent tolerable Werte ohne Unterschied.

Schließlich gehen wir noch einen Schritt tiefer in das Material und analysieren die Abweichungen der einzelnen Umfragen.

Insgesamt berichten die 47 Umfragen 286 Messungen für die Parteien. Jeder einzelne Wert wurde mit dem beschriebenen Verfahren im Rahmen der Fehler-toleranz getestet. Den Wahlmatrizen können die absoluten Testergebnisse entnommen werden. Dort sind auch die relativen Bilanzen ausgewiesen, die wir bereits oben zweimal herangezogen haben.

Von den insgesamt 286 Messwerten sind 196 im tolerablen Bereich zutreffend (das sind 68,5 Prozent). Auch dieser Wert ist zunächst positiv zu bewerten.



Man könnte ferner die Hypothese aufstellen, dass „Fehlprognosen“ vor allem durch Schwierigkeiten der Vorhersage für die extremen Parteien hervorgerufen werden.

Diese Hypothese kann durch einfache Durchsicht der Matrizen ohne aufwendige Rechnungen verworfen werden.

Benchmarking für die Institute

Die Institute stehen im Wettbewerb miteinander, auf dem Markt und in der Öffentlichkeit. Für die Öffentlichkeit interessant ist die Frage, wie verlässlich die Institute im Vergleich zueinander mit zeitnahen Umfragen das spätere Wahlergebnis vorhersagen.

Wir legen dazu eine Liste der im Feld der 47 analysierten Umfragen beteiligter Institute auf und zählen für jede Umfrage des Instituts in den Matrizen die positiv und die negativ getesteten Werte. Der relative Anteil der positiv getesteten Werte pro Institut ergibt die Rangfolge des hier verwendeten Benchmarking. In einem zweiten Schritt werfen wir einen Blick auf die höchsten Abweichungen, die pro Institut zwischen Umfrage-Werten für die Parteien und amtlichem Resultat zu verzeichnen sind.

Bei dem Benchmarking muss allerdings berücksichtigt werden, dass die Institute sehr unterschiedlich häufig im Feld tätig waren. Deshalb wird hier zunächst die Benchmarking-Tabelle geordnet nach der Zahl der getesteten Umfragen vorangestellt.

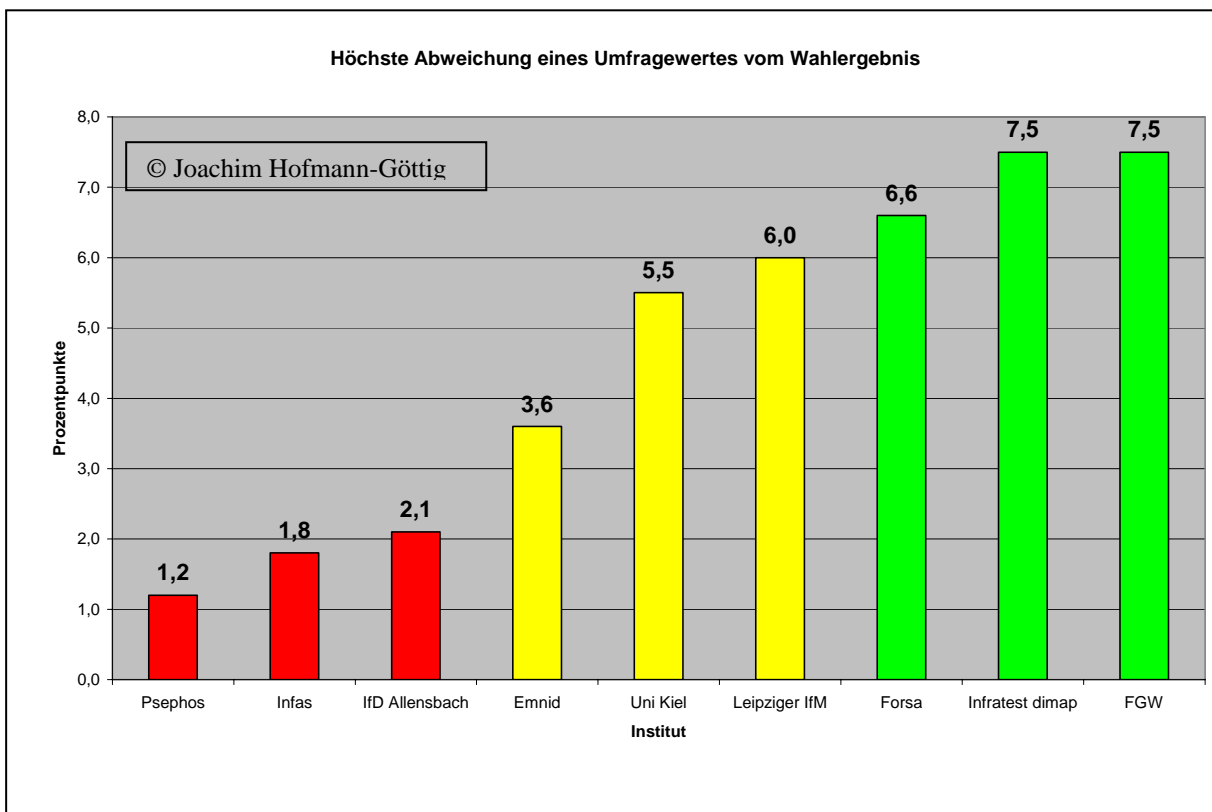
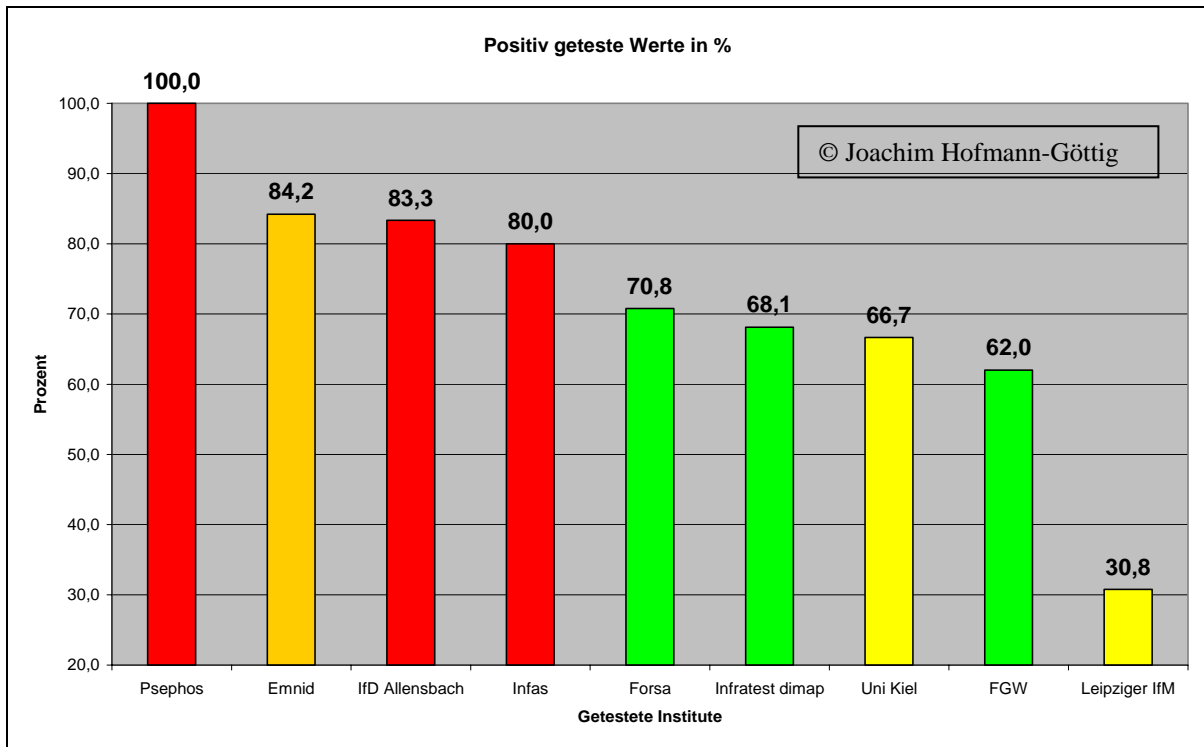
Benchmarking der Institute						
Name Institut	Zahl der Umfragen	Positive Werte	Neg. Werte	Sum.	Rel. pos. Anteil %	Höchster neg. Wert
Psephos	1	7	0	7	100,0	1,2
IfD Allensbach	1	5	1	6	83,3	2,1
Infas	1	4	1	5	80,0	1,8
Uni Kiel	1	4	2	6	66,7	5,5
Leipziger IfM	2	4	9	13	30,8	6,0
Emnid	3	16	3	19	84,2	3,6
FGW	8	31	19	50	62,0	7,5
Forsa	15	63	26	89	70,8	6,6
Infratest dimap	15	62	29	91	68,1	7,5
Summe	47	196	90	286	68,5	
© Joachim Hofmann-Göttig						

Es wird deutlich: Insgesamt werden hier neun Institute getestet. *Forsa* und *Infratest dimap* sind mit jeweils 15 Umfragen beteiligt. Vier Institute (*Psephos*, *Allensbach*, *Infas*, *Uni Kiel*) sind jeweils nur mit einer Umfrage vertreten. Das *Leipziger Institut für Marktforschung* ist zwar mit zwei Umfragen im Test, in beiden Fällen aber mit ostdeutschen Landtagswahlen (Sachsen und Thüringen), für die, wie gezeigt, die politische Meinungsforschung noch schwieriger ist.

Insoweit muss das Benchmarking mit methodischen Einschränkungen interpretiert werden. Ordnen wir die Benchmarking-Tabelle nach der Reihenfolge der Institute nach dem relativen Anteil positiv getesteter Werte.

Psephos, *Emnid*, *Allensbach* und *Infas* stehen in dieser Rangreihe oben und die *Leipziger* (nicht weiter überraschend) unten. Interessanter ist die Rangreihe unter jenen Instituten, die mit zweimal 15 (*Forsa* und *Infratest*) und einmal 8 (*Forschungsgruppe Wahlen*) im Test beteiligt waren. Diesen Test gewinnt (knapp) *Forsa* vor *Infratest dimap* (u.a. ARD-Deutschland-Trend), die beide überraschend deutlich die *Forschungsgruppe Wahlen* (ZDF-Politbarometer) hinter sich lassen.

In einem weiteren Schritt werden die höchsten Abweichungen pro Institut zur Sortierung herangezogen.



Nicht weiter überraschend ist, dass an der Spitze dieses Benchmarking Institute liegen, die nur selten im Feld waren, gleichwohl weist ein Institut wie *Psephos*, das u.a. auch die Quartalsumfragen für die pfälzische Tageszeitung „*Rheinpfalz*“ verantwortet, einen maximalen Fehler von nur 1,2 Prozentpunkten auf. Die höchsten Fehler machen diejenigen, die am meisten forschen, auch das ist nicht überraschend. Schon eher erstaunlich ist, dass die maximalen Fehler bei *Infra-test dimap* und bei der *Forschungsgruppe Wahlen* (mit jeweils 7,5 Prozentpunkte) durchaus bemerkenswert hoch ausfielen und höher als bei *Forsa* (6,6 Prozentpunkte) waren.

Fazit:

Welchen Prognosewert haben die aktuellen Umfragewerte für die Bundestagswahl 2005?

Ist die Wahl schon entschieden?

Die deutschen Meinungsforschungsinstitute sind sich einig: „Wenn am nächsten Sonntag Bundestagswahlen wären“, so lautet ihre beliebte „Sonntagsfrage“, dann könnte sich *Angela Merkel* auf eine solide Bundestagsmehrheit einer schwarz-gelben Koalition stützen. Im Übrigen hätten wir im neuen Deutschen Bundestag eine sozialdemokratische Oppositionspartei um die 30 Prozent und neben *BÜNDNIS 90/DIE GRÜNEN* die neue *Linke/PDS* sicher im Bundestag zu erwarten. Für die neue Linke, so die Ergebnisse der Sonntagsfrage, seien die Zeiten zweistelliger Prozentergebnisse vorbei.

Ist die wirkliche „Sonntagsfrage“ der 61,5 Millionen Wahlberechtigter, die tatsächliche Bundestagswahl, also nur noch Formsache? Könnte man sich die teuren Wahlen am besten sparen zugunsten des Durchschnittswertes der demoskopischen Daten der letzten zehn Tage vor einer Wahl?

Die Öffentlichkeit nimmt politische Umfragedaten sehr umfassend auf. Die Einnahmen verwechseln sie fast mit der Wirklichkeit. Für die anderen sind sie so verlässlich wie das Lesen aus dem Kaffee-Satz. Die meisten ahnen, dass die Wahrheit wohl in der Mitte liegt. Aber wo genau? Das sollte erstmals empirisch wissenschaftlich gesichert ohne Vorurteile erhoben werden. Dazu wurden die in einer Internet-Datenbank²⁹ gesammelten 47 Umfragen zu den letzten Landtagswahlen in allen 16 Bundesländern, der letzten Europa- und der letzten Bundestagswahl (insgesamt 19 Wahlen), die jeweils höchstens zehn Tage vor der jeweiligen Wahl veröffentlicht wurden, mit den tatsächlichen Wahlergebnissen verglichen. Es galt heraus zu finden, wie genau diese aktuellen Umfragen das tatsächliche Wahlergebnis prognostizieren.

Ein Umfragewert für eine Partei wurde für tolerabel eingeschätzt, wenn er sich innerhalb der statistischen Fehlertoleranztafel³⁰ bewegte (d.h. für die großen Parteien +/- 3 Prozentpunkte, für die kleinen mindestens 1,5 Prozentpunkte).

Werden die 19 untersuchten Wahlen mit ihren 47 Umfragen als Einheit gesehen, kann generell festgestellt werden, dass die Demoskopie die Wahlen mit einem tolerablen Fehler vorher sagen konnten. Der durchschnittlich summierte Fehler

²⁹ Datenbasis: www.wahlrecht.de/ Aktuell/ Sonntagsfrage, zugegriffen am 31.8.2005.

³⁰ Basis: *Infratest dimap*, Deutschland-Trend, August 2005, letzte Seite.

lag bei 10,5 Prozentpunkten (bis zu 11,9 läge dieser Wert im tolerablen Bereich, der reale Wert liegt also mit 1,4 Prozentpunkten im positiven Bereich).

Werden die 19 Wahlen einzeln betrachtet, ist erkennbar, dass dies dort auch für 12 Wahlen (also einer knappen 2/3-Mehrheit) gilt.

Bei sieben Wahlen müssen allerdings „*Fehlprognosen*“ diagnostiziert werden: Bei den Landtagswahlen in Brandenburg (- 3,8 Prozentpunkte), Sachsen (- 3,4 PP), Berlin (- 1,6 PP), Sachsen-Anhalt (- 3,4 PP) und Bremen (- 3,0 PP) sind die Summen der Fehler deutlich über der Toleranz. Bei der Europawahl (- 6,2 Prozentpunkte) und der Landtagswahl in Thüringen (-7,6 Prozentpunkte) müssen sogar sehr krasse Abweichungen festgestellt werden.

Werden die Analysen für die Gesamtheit der Umfragen auf die einzelnen Parteiateile fortgeführt, so liegen 70,4 Prozent der von den Umfragen durchschnittlich für eine Wahl erhobenen Parteiateile im tolerablen Bereich. Dabei ist augenfällig, dass sich die Umfrageinstitute im Osten Deutschlands schwerer tun als im Westen.

Wird der Blick auf die einzelnen Umfragen und deren Messergebnisse für die Parteien gelenkt (286 Messwerte), können auch hier 2/3 tolerable Messungen (68,5 Prozent) festgestellt werden.

Die Prognosetauglichkeit der Sonntagsfragen der neun verschiedenen Institute in den 47 getesteten Umfragen schwankte allerdings beachtlich (zwischen 100 Prozent = *Psephos* und 30,8 Prozent = *Leipziger Institut für Marktforschung*), was aber auch methodisch erklärt werden muss (geringe Umfrageanzahl, regionale Einseitigkeiten). Von den Instituten, die mit vielen Umfragen im Untersuchungszeitraum 2001 – 2005 im Feld waren, war *Forsa* „*Gruppensieger*“, vor *Infratest dimap* und deutlich vor der *Forschungsgruppe Wahlen*.

Bei den Instituten gab es beachtliche „*Höchst-Abweichungen*“ zwischen Umfrage und Wahlergebnis. Das ging bis zu 6,6 Prozentpunkten bei *Forsa* und 7,5 Prozentpunkten bei *Infratest* und der *Forschungsgruppe Wahlen*. Diese hohen Abweichungen waren allerdings regelmäßig Landtagswahlen im Osten Deutschlands geschuldet.

Aber auch bei einer bundesweiten Wahl, der Europawahl vom 13.6.2004, kam es zu beachtlichen Abweichungen: Hier wurde die SPD von allen drei Forschungsinstituten (*Forsa*, *FGW*, *Infratest*) mit durchschnittlich 6,5 Prozentpunkten überschätzt.

Auch bei anderen Wahlen kam es vor, dass alle Forschungsinstitute intolerable Fehler in dieselbe Richtung vermelden mussten: So wurde bei der Landtagswahl in Schleswig-Holstein vom 20.2.2005 die CDU von allen vier Instituten (*Forsa*, *Uni Kiel*, *FGW*, *Infratest*) um durchschnittlich 3,4 Prozentpunkte unterschätzt.

Erstaunlicherweise gab es sogar in stabilen westlichen Ländern einheitliche Fehlprognosen. Bei der Landtagswahl in Baden-Württemberg am 25.3.2001 wurde die CDU von beiden Instituten (*Forsa, Infratest*) um 3,8 Prozentpunkte unterschätzt.

Demgegenüber waren die Umfragen bei der letzten Bundestagswahl am 22.9.2002 sehr nah am amtlichen Endergebnis. Lediglich *Allensbach* musste feststellen, dass es die FDP mit 2,1 Prozentpunkten überschätzt hat.

Unter dem Strich kann festgestellt werden, dass die kurz vor Wahlen veröffentlichten Stimmungszahlen im Regelfall zufrieden stellend das spätere Wahlergebnis prognostizieren, allerdings im Rahmen der üblichen Fehlertoleranz für Hochrechnungen von Stichproben auf die Gesamtwählerschaft. Die Genauigkeit ist für Bundestagswahlen höher als für Europa- und für Landtagswahlen, im Westen höher als im Osten. Selbst die einheitliche Messung aller Institute schützt nicht zuverlässig vor Fehlprognosen.

Die gegenwärtigen Umfragen lassen einen Sieg von Schwarz-Gelb zwar sehr wahrscheinlich erscheinen, keineswegs aber als „sicher“. Unsicher ist vor allem der Zuspruch der Linkspartei, besonders im Osten. Da der hohe Zuspruch für die Union insbesondere auch enttäuschter SPD-Klientel zuzurechnen ist, würde ein zweistelliges Abschneiden der Linkspartei auch zu Lasten der Union gehen.

So genau die Demoskopie heute insgesamt misst, die Wahrheit kennen wir erst, wenn „*am nächsten Sonntag tatsächlich Bundestagswahlen sind*“.

Und das ist gut so.

Anhang

© Joachim Hofmann-Göttig		
Nr.	Wahl in ...	Am ...
1	Nordrhein-Westfalen	22.05.2005
2	Schleswig-Holstein	20.02.2005
3	Brandenburg	19.09.2004
4	Sachsen	19.09.2004
5	Saarland	05.09.2004
6	Thüringen	13.06.2004
7	Europawahl	13.06.2004
8	Hamburg I	29.02.2004
9	Bayern	21.09.2003
10	Bremen	25.05.2003
11	Niedersachsen	02.02.2003
12	Hessen	02.02.2003
13	Mecklenburg-Vorpommern	22.09.2002
14	Bundestagswahl	22.09.2002
15	Sachsen-Anhalt	21.04.2002
16	Berlin	21.10.2001
17	Hamburg II	23.09.2001
18	Baden-Württemberg	25.03.2001
19	Rheinland-Pfalz	25.03.2001

1. Landtagswahl in Nordrhein-Westfalen am 22.5.2005

Institut	Quelle	Zeitraum	Datum	CDU	SPD	Grüne	FDP	Sonstige	Summe
Amtl. Ergebnis			22.05.2005	44,8	37,1	6,2	6,2	5,7	100,0
Forsa	RTL/n-tv	18.05.2005	19.05.2005	43,0	36,0	7,0	7,0	7,0	100,0
Infratest dimap	WDR	12.-13.05.2005	15.05.2005	43,0	37,0	7,5	7,5	5,0	100,0
Infas	Infas	10.-12.05.2005	13.05.2005	43,0	36,0	7,0	8,0	6,0	100,0
FGWahlen	ZDF	9.-12.05.2005	13.05.2005	44,0	35,0	9,0	7,0	5,0	100,0
Diff. Forsa				-1,8	-1,1	0,8	0,8	1,3	5,8
Diff. Infratest				-1,8	-0,1	1,3	1,3	-0,7	5,2
Diff. Infas				-1,8	-1,1	0,8	1,8	0,3	5,8
Diff. FGW				-0,8	-2,1	2,8	0,8	-0,7	7,2
Mittlere Diff.				1,6	1,1	1,4	1,2	0,8	6,0
Statistische Fehlertoleranz in Prozentpunkten				3,1	3,0	1,4	1,4	1,4	10,3
Absolute Zahl der positiv getesteten Werte									18
Absolute Zahl der negativ getesteten Werte									2
Relative Bilanz (Anteil der positiv getesteten Werte in Prozent)									90,0

Umfragedaten aus: www.wahlrecht.de/Umfragen, im Internet am 31.8.2005; eigene Berechnungen; © Joachim Hofmann-Göttig

2. Landtagswahl in Schleswig-Holstein am 20.2.2005

Institut	Quelle	Zeitraum	Datum	CDU	SPD	Grüne	FDP	SSW	Sonstige	Summe
Amtl. Ergebnis			20.02.2005	40,2	38,7	6,2	6,6	3,6	4,7	100,0
Forsa	n-tv		16.02.2005	37,0	40,0	6,0	7,0	4,0	6,0	100,0
Uni Kiel	Uni Kiel	17.1.-11.2.05	15.02.2005	37,3	39,4	10,6	7,0	3,3	2,4	100,0
FGWahlen	ZDF	8.-10.2.05	11.02.2005	37,0	40,0	7,0	7,0	4,0	5,0	100,0
Infratest dimap	NDR	7.-9.2.05	10.02.2005	36,0	41,0	7,5	7,0	3,0	5,5	100,0
Diff. Forsa				-3,2	1,3	-0,2	0,4	0,4	1,3	6,8
Diff. Uni Kiel				-2,9	0,7	4,4	0,4	-0,3	-2,3	11,0
Diff. FGW				-3,2	1,3	0,8	0,4	0,4	0,3	6,4
Diff. Infratest				-4,2	2,3	1,3	0,4	-0,6	0,8	9,6
Mittlere Diff.				3,4	1,4	1,7	0,4	0,4	1,2	8,5
Statistische Fehlertoleranz in Prozentpunkten				3,0	3,0	1,5	1,6	1,2	1,4	11,7
Absolute Zahl der positiv getesteten Werte										19
Absolute Zahl der negativ getesteten Werte										5
Relative Bilanz (Anteil der positiv getesteten Werte in Prozent)										79,2

Umfragedaten aus: www.wahlrecht.de/Umfragen, im Internet am 31.8.2005; eigene Berechnungen; © Joachim Hofmann-Göttig

3. Landtagswahl in Brandenburg am 19.9.2004

Institut	Quelle	Zeitraum	Datum	CDU	SPD	Grüne	FDP	PDS	DVU	Sonstige	Summe
Amtl. Ergebnis			19.09.2004	19,4	31,9	3,6	3,3	28,0	6,1	7,7	100,0
FGWahlen	ZDF	6.-9.9.04	10.09.2004	23,0	29,0	6,0	5,0	27,0	6,0	4,0	100,0
Infratest dimap	Märk. Allgemeine	6.-8.9.04	09.09.2004	24,0	27,0	5,0	4,0	31,0	5,0	4,0	100,0
Diff. FGW				3,6	-2,9	2,4	1,7	-1,0	-0,1	-3,7	15,4
Diff. Infratest				4,6	-4,9	1,4	0,7	3,0	-1,1	-3,7	19,4
Mittlere Diff.				4,1	3,9	1,9	1,2	2,0	0,6	3,7	17,4
Statistische Fehlertoleranz in Prozentpunkten				2,5	2,8	1,2	1,1	2,8	1,5	1,7	13,6
Absolute Zahl der positiv getesteten Werte											4
Absolute Zahl der negativ getesteten Werte											10
Relative Bilanz (Anteil der positiv getesteten Werte in Prozent)											28,6

Umfragedaten aus: www.wahlrecht.de/Umfragen, im Internet am 31.8.2005; eigene Berechnungen; © Joachim Hofmann-Göttig

4. Landtagswahl in Sachsen am 19.9.2004

Institut	Quelle	Zeitraum	Datum	CDU	SPD	Grüne	FDP	PDS	NPD	Sonstige	Summe
Amtl. Ergebnis			19.09.2004	41,1	9,8	5,1	5,9	23,6	9,2	5,3	100,0
Leipziger Institut f. Marktforschung	Leipziger Volkszeitung	7.-9.9.04	10.09.2004	44,0	14,0	7,0	6,0	19,0	7,0	3,0	100,0
FGWahlen	ZDF	6.-9.9.04	10.09.2004	47,0	11,0	6,0	4,0	19,0	9,0	4,0	100,0
Diff. Leipziger				2,9	4,2	1,9	0,1	-4,6	-2,2	-2,3	18,2
Diff. FGW				5,9	1,2	0,9	-1,9	-4,6	-0,2	-1,3	16,0
Mittlere Diff.				4,4	2,7	1,4	1,0	4,6	1,2	1,8	17,1
Statistische Fehlertoleranz in Prozentpunkten				3,0	1,9	1,4	1,5	2,7	1,8	1,4	13,7
Absolute Zahl der positiv getesteten Werte											6
Absolute Zahl der negativ getesteten Werte											8
Relative Bilanz (Anteil der positiv getesteten Werte in Prozent)											42,9

Umfragedaten aus: www.wahlrecht.de/Umfragen, im Internet am 31.8.2005; eigene Berechnungen; © Joachim Hofmann-Göttig

5. Landtagswahl im Saarland am 5.9.2004

Institut	Quelle	Zeitraum	Datum	CDU	SPD	Grüne	FDP	Sonstige	Summe
Amtl. Ergebnis			05.09.2004	47,5	30,8	5,6	5,2	10,9	100,0
Infratest dimap	ARD	23.-25.8.04	26.08.2004	51,0	30,0	7,0	5,0	7,0	100,0
Diff. Infratest				3,5	-0,8	1,4	-0,2	-3,9	9,8
Mittlere Diff.				3,5	0,8	1,4	0,2	3,9	9,8
Statistische Fehlertoleranz in Prozentpunkten				3,1	2,8	1,4	1,4	1,9	10,6
Absolute Zahl der positiv getesteten Werte									3
Absolute Zahl der negativ getesteten Werte									2
Relative Bilanz (Anteil der positiv getesteten Werte in Prozent)									60,0

Umfragedaten aus: www.wahlrecht.de/Umfragen, im Internet am 31.8.2005; eigene Berechnungen; © Joachim Hofmann-Göttig

6. Landtagswahl in Thüringen am 13.6.2004

Institut	Quelle	Zeitraum	Datum	CDU	SPD	Grüne	FDP	PDS	Sonstige	Summe
Amtl. Ergebnis			13.06.2004	43,0	14,5	4,5	3,6	26,1	8,3	100,0
FGWahlen	ZDF	1.-3.6.04	04.06.2004	44,0	22,0	5,0	5,0	21,0	3,0	100,0
Infratest dimap	WDR	29.5.-3.6.-04	04.06.2004	45,0	20,0	4,5	4,0	22,0	4,5	100,0
IfM Leipzig	Freies Wort		03.06.2004	49,0	20,0	4,0	3,0	21,0	3,0	100,0
Diff. FGW				1,0	7,5	0,5	1,4	-5,1	-5,3	20,8
Diff. Infratest				2,0	5,5	0,0	0,4	-4,1	-3,8	15,8
Diff. Leipzig				6,0	5,5	-0,5	-0,6	-5,1	-5,3	23,0
Mittlere Diff.				3,0	6,2	0,3	0,8	4,8	4,8	19,9
Statistische Fehlertoleranz in Prozentpunkten				3,1	2,2	1,3	1,2	2,7	1,8	12,3
Absolute Zahl der positiv getesteten Werte										7
Absolute Zahl der negativ getesteten Werte										11
Relative Bilanz (Anteil der positiv getesteten Werte in Prozent)										38,9

Umfragedaten aus: www.wahlrecht.de/Umfragen, im Internet am 31.8.2005; eigene Berechnungen; © Joachim Hofmann-Göttig

7. Europawahl am 13.6.2004

Institut	Quelle	Zeitraum	Datum	CDU	SPD	Grüne	FDP	PDS	Sonstige	Summe
Amtl. Ergebnis			13.06.2004	44,5	21,5	11,9	6,1	6,1	9,8	99,9
Forsa	Stern	1.-4.6.04	08.06.2004	45,0	28,0	13,0	6,0	4,0	4,0	100,0
FGWahlen	ZDF	1.-3.6.04	04.06.2004	49,0	27,0	10,0	5,0	4,0	5,0	100,0
Infratest dimap	WDR	29.5.-3.6.04	04.06.2004	47,0	29,0	12,0	5,0	4,0	3,0	100,0
Diff. Forsa				0,5	6,5	1,1	-0,1	-2,1	-5,8	16,1
Diff. FGW				4,5	5,5	-1,9	-1,1	-2,1	-4,8	19,9
Diff. Infratest				2,5	7,5	0,1	-1,1	-2,1	-6,8	20,1
Mittlere Diff.				2,5	6,5	1,0	0,8	2,1	5,8	18,7
Statistische Fehlertoleranz in Prozentpunkten				3,1	2,5	2,0	1,5	1,5	1,9	12,5
Absolute Zahl der positiv getesteten Werte										8
Absolute Zahl der negativ getesteten Werte										10
Relative Bilanz (Anteil der positiv getesteten Werte in Prozent)										44,4

Umfragedaten aus: www.wahlrecht.de/Umfragen, im Internet am 31.8.2005; eigene Berechnungen; © Joachim Hofmann-Göttig

8. Hamburger Bürgerschaftswahl am 29.2.2004

Institut	Quelle	Zeitraum	Datum	CDU	SPD	GAL	FDP	Offen -sive	Pro DM/ Schill	Sonstige	Summe
Amtl. Ergebnis			29.02.2004	47,2	30,5	12,3	2,8	0,4	3,1	3,6	99,9
Forsa	Stern	20.-26.2.04	27.02.2004	44,0	30,0	14,0	4,0	1,0	4,0	3,0	100,0
Emnid	Die Welt	24.02.2004	25.02.2004	46,0	30,0	13,0	3,0	1,0	3,0	4,0	100,0
Forsa	Hamburger Morgenpost	18.-23.2.04	25.02.2004	45,0	29,0	14,0	4,0	1,0	3,0	4,0	100,0
Psephos	Abendblatt	19.-22.2.04	24.02.2004	46,0	30,0	12,0	3,5	1,0	3,5	4,0	100,0
FGWahlen	ZDF	16.-19.2.04	20.02.2004	47,0	29,0	13,0	4,0	0,0	3,0	4,0	100,0
Diff. Forsa I				-3,2	-0,5	1,7	1,2	0,6	0,9	-0,6	8,7
Diff. Emnid				-1,2	-0,5	0,7	0,2	0,6	-0,1	0,4	3,7
Diff. Forsa II				-2,2	-1,5	1,7	1,2	0,6	-0,1	0,4	7,7
Diff. Psephos				-1,2	-0,5	-0,3	0,7	0,6	0,4	0,4	4,1
Diff. FGW				-0,2	-1,5	0,7	1,2	-0,4	-0,1	0,4	4,5
Mittlere Diff.				1,6	0,9	1,0	0,9	0,6	0,3	0,4	5,7
Statistische Fehlertoleranz in Prozentpunkten				3,1	2,8	2,0	1,1	0,9	1,1	1,2	12,2
Absolute Zahl der positiv getesteten Werte											31
Absolute Zahl der negativ getesteten Werte											4
Relative Bilanz (Anteil der positiv getesteten Werte in Prozent)											88,6

Umfragedaten aus: www.wahlrecht.de/Umfragen, im Internet am 31.8.2005; eigene Berechnungen; © Joachim Hofmann-Göttig

9. Landtagswahlen in Bayern am 21.9.2003

Institut	Quelle	Zeitraum	Datum	CSU	SPD	Grüne	FDP	Freie Wähler	Sonstige	Summe
Amtl. Ergebnis			21.09.2003	60,7	19,6	7,7	2,6	4,0	5,4	100,0
Infratest dimap	Bayr. Rundf.	6.-10.9.03	11.09.2003	59,0	20,0	8,0	4,0	4,0	5,0	100,0
Diff. Infratest				-1,7	0,4	0,3	1,4	0,0	-0,4	4,2
Mittlere Diff.				1,7	0,4	0,3	1,4	0,0	0,4	4,2
Statistische Fehlertoleranz in Prozentpunkten				3,0	2,5	1,7	1,1	1,2	1,4	10,9
Absolute Zahl der positiv getesteten Werte										5
Absolute Zahl der negativ getesteten Werte										1
Relative Bilanz (Anteil der positiv getesteten Werte in Prozent)										83,3
Umfragedaten aus: www.wahlrecht.de/Umfragen , im Internet am 31.8.2005; eigene Berechnungen; © Joachim Hofmann-Göttig										

10. Bremer Bürgerschaftswahl am 25.5.2003

Institut	Quelle	Zeitraum	Datum	CDU	SPD	Grüne	FDP	Schill	DVU	Sonstige	Summe
Amtl. Ergebnis			25.05.2003	29,9	42,3	12,8	4,2	4,3	2,3	4,2	100,0
Infratest dima	ARD		16.05.2003	36,0	38,0	12,0	5,0	2,0	2,0	5,0	100,0
Diff. Infratest				6,1	-4,3	-0,8	0,8	-2,3	-0,3	0,8	15,4
Mittlere Diff.				6,1	4,3	0,8	0,8	2,3	0,3	0,8	15,4
Statistische Fehlertoleranz in Prozentpunkten				2,8	3,0	2,1	1,2	1,2	0,9	1,2	12,4
Absolute Zahl der positiv getesteten Werte											4
Absolute Zahl der negativ getesteten Werte											3
Relative Bilanz (Anteil der positiv getesteten Werte in Prozent)											57,1
Umfragedaten aus: www.wahlrecht.de/Umfragen , im Internet am 31.8.2005; eigene Berechnungen; © Joachim Hofmann-Göttig											

11. Landtagswahl in Niedersachsen am 2.2.2003

Institut	Quelle	Zeitraum	Datum	CDU	SPD	Grüne	FDP	Sonstige	Summe
Amtl. Ergebnis			02.02.2003	48,3	33,4	7,6	8,1	2,5	99,9
Forsa I	RTL	27.-30.1.03	31.01.2003	48,0	33,0	8,0	7,0	4,0	100,0
Forsa II	Stern	20.-24.1.03	29.01.2003	48,0	35,0	8,0	6,0	3,0	100,0
Diff. Forsa I				-0,3	-0,4	0,4	-1,1	1,5	3,7
Diff. Forsa II				-0,3	1,6	0,4	-2,1	0,5	4,9
Mittlere Diff.				0,3	1,0	0,4	1,6	1,0	4,3
Statistische Fehlertoleranz in Prozentpunkten				3,1	2,9	1,7	1,7	1,0	10,4
Absolute Zahl der positiv getesteten Werte									8
Absolute Zahl der negativ getesteten Werte									2
Relative Bilanz (Anteil der positiv getesteten Werte in Prozent)									80,0

Umfragedaten aus: www.wahlrecht.de/Umfragen, im Internet am 31.8.2005; eigene Berechnungen; © Joachim Hofmann-Göttig

12. Landtagswahl in Hessen am 2.2.2003

Institut	Quelle	Zeitraum	Datum	CDU	SPD	Grüne	FDP	Sonstige	Summe
Amtl. Ergebnis			02.02.2003	48,8	29,1	10,1	7,9	4,1	100,0
Forsa I	RTL		31.01.2003	49,0	29,0	11,0	8,0	3,0	100,0
Forsa II	Stern	20.-24.1.03	29.01.2003	51,0	29,0	10,0	7,0	3,0	100,0
Diff. Forsa I				0,2	-0,1	0,9	0,1	-1,1	2,4
Diff. Forsa II				2,2	-0,1	-0,1	-0,9	-1,1	4,4
Mittlere Diff.				1,2	0,1	0,5	0,5	1,1	3,4
Statistische Fehlertoleranz in Prozentpunkten				3,1	2,8	1,9	1,7	1,2	10,7
Absolute Zahl der positiv getesteten Werte									10
Absolute Zahl der negativ getesteten Werte									0
Relative Bilanz (Anteil der positiv getesteten Werte in Prozent)									100,0

Umfragedaten aus: www.wahlrecht.de/Umfragen, im Internet am 31.8.2005; eigene Berechnungen; © Joachim Hofmann-Göttig

13. Landtagswahl in Mecklenburg-Vorpommern am 22.9.2002

Institut	Quelle	Zeitraum	Datum	CDU	SPD	Grüne	FDP	PDS	Sonstige	Summe
Amtl. Ergebnis			22.09.2002	31,4	40,6	2,6	4,7	16,4	4,3	100,0
Emnid	Antenne MV		16.09.2002	33,0	37,0	3,0	3,0	19,0	5,0	100,0
Infratest dimap	NDR		12.09.2002	33,0	40,0	2,0	4,0	17,0	4,0	100,0
Diff. Emnid				1,6	-3,6	0,4	-1,7	2,6	0,7	10,6
Diff. Infratest				1,6	-0,6	-0,6	-0,7	0,6	-0,3	4,4
Mittlere Diff.				1,6	2,1	0,5	1,2	1,6	0,5	7,5
Statistische Fehlertoleranz in Prozentpunkten				2,8	3,0	1,0	1,4	2,3	1,2	11,7
Absolute Zahl der positiv getesteten Werte										9
Absolute Zahl der negativ getesteten Werte										3
Relative Bilanz (Anteil der positiv getesteten Werte in Prozent)										75,0

Umfragedaten aus: www.wahlrecht.de/Umfragen, im Internet am 31.8.2005; eigene Berechnungen; © Joachim Hofmann-Göttig

14. Bundestagswahl am 22.9.2002

Institut	Quelle	Zeitraum	Datum	CDU/ CSU	SPD	Grüne	FDP	PDS	Sonstige	Summe
Amtl. Ergebnis			22.09.2002	38,5	38,5	8,6	7,4	4,0	3,0	100,0
Allensbach	FAZ		21.09.2002	37,0	37,5	7,5	9,5	4,5	4,0	100,0
Forsa			20.09.2002	37,5	39,0	7,0	7,5	4,3	4,7	100,0
Emnid			16.09.2002	37,0	39,0	7,0	8,0	5,0	4,0	100,0
Infratest dimap	ARD		13.09.2002	36,0	38,5	8,0	8,5	4,7	4,3	100,0
FGWahlen	ZDF		13.09.2002	37,0	40,0	7,0	7,5	4,5	4,0	100,0
Diff. Allensbach				-1,5	-1,0	-1,1	2,1	0,5	1,0	7,2
Diff. Forsa				-1,0	0,5	-1,6	0,1	0,3	1,7	5,2
Diff. Emnid				-1,5	0,5	-1,6	0,6	1,0	1,0	6,2
Diff. Infratest				-2,5	0,0	-0,6	1,1	0,7	1,3	6,2
Diff. FGW				-1,5	1,5	-1,6	0,1	0,5	1,0	6,2
Mittlere Diff.				1,0	0,7	0,7	0,8	0,6	1,2	6,2
Statistische Fehlertoleranz in Prozentpunkten				3,0	3,0	1,8	1,7	1,2	1,1	11,8
Absolute Zahl der positiv getesteten Werte										27
Absolute Zahl der negativ getesteten Werte										3
Relative Bilanz (Anteil der positiv getesteten Werte in Prozent)										90,0

Umfragedaten aus: www.wahlrecht.de/Umfragen, im Internet am 31.8.2005; eigene Berechnungen; © Joachim Hofmann-Göttig

15. Landtagswahl in Sachsen-Anhalt am 21.4.2002

Institut	Quelle	Zeitraum	Datum	CDU	SPD	Grüne	FDP	PDS	Schill	Sonstige	Summe
Amtl. Ergebnis			21.04.2002	37,3	20,0	2,0	13,3	20,4	4,5	2,5	100,0
Forsa	Magdeburger Volksstimme		19.04.2002	34,0	24,0	2,0	10,0	22,0	6,0	2,0	100,0
Infratest dimap	MDR		11.04.2002	32,0	25,0	2,0	9,0	23,0	5,0	4,0	100,0
Diff. Forsa				-3,3	4,0	0,0	-3,3	1,6	1,5	-0,5	14,2
Diff. Infratest				-5,3	5,0	0,0	-4,3	2,6	0,5	1,5	19,2
Mittlere Diff.				4,3	4,5	0,0	3,8	2,1	1,0	1,0	16,7
Statistische Fehlertoleranz in Prozentpunkten				3,0	2,5	0,9	2,1	2,5	1,3	1,0	13,3
Absolute Zahl der positiv getesteten Werte											5
Absolute Zahl der negativ getesteten Werte											9
Relative Bilanz (Anteil der positiv getesteten Werte in Prozent)											35,7

Umfragedaten aus: www.wahlrecht.de/Umfragen, im Internet am 31.8.2005; eigene Berechnungen; © Joachim Hofmann-Göttig

16. Wahlen zum Berliner Abgeordnetenhaus am 21.10.2001

Institut	Quelle	Zeitraum	Datum	CDU	SPD	Grüne	FDP	PDS	Sonstige	Summe
Amtl. Ergebnis			21.10.2001	23,8	29,7	9,1	9,9	22,6	4,9	100,0
Forsa II	Die Woche		18.10.2001	25,0	35,0	9,0	7,0	18,0	5,0	99,0
Infratest dimap	SFB		12.10.2001	25,0	33,0	10,0	9,0	19,0	4,0	100,0
Forsa II	Die Woche		11.10.2001	26,0	36,0	10,0	7,0	16,0	5,0	100,0
Diff. Forsa I				1,2	5,3	-0,1	-2,9	-4,6	0,1	14,2
Diff. Infratest				1,2	3,3	0,9	-0,9	-3,6	-0,9	10,8
Diff. Forsa II				2,2	6,3	0,9	-2,9	-6,6	0,1	19,0
Mittlere Diff.				1,5	5,0	0,6	2,2	4,9	0,4	14,7
Statistische Fehlertoleranz in Prozentpunkten				2,6	2,8	1,8	1,9	2,6	1,4	13,1
Absolute Zahl der positiv getesteten Werte										10
Absolute Zahl der negativ getesteten Werte										8
Relative Bilanz (Anteil der positiv getesteten Werte in Prozent)										55,6

Umfragedaten aus: www.wahlrecht.de/Umfragen, im Internet am 31.8.2005; eigene Berechnungen; © Joachim Hofmann-Göttig

17. Hamburger Bürgerschaftswahl am 23.9.2001

Institut	Quelle	Zeitraum	Datum	CDU	SPD	GAL	FDP	Schill	Regen- bogen	Sonstige	Summe
Amtl. Ergebnis			29.02.2004	26,2	36,5	8,6	5,1	19,4	1,7	2,5	100,0
Forsa	Die Welt Hamburger		20.09.2001	28,0	36,0	11,0	5,0	14,0	3,0	3,0	100,0
Infratest dimap	Abendblatt		13.09.2001	28,0	35,0	9,0	5,0	16,0	2,0	5,0	100,0
Diff. Forsa				1,8	-0,5	2,4	-0,1	-5,4	1,3	0,5	12,0
Diff. Infratest				1,8	-1,5	0,4	-0,1	-3,4	0,3	2,5	10,0
Mittlere Diff.				1,8	1,0	1,4	0,1	4,4	0,8	1,5	11,0
Statistische Fehlertoleranz in Prozentpunkten				2,7	3,0	1,8	1,4	2,5	0,9	1,0	13,3
Absolute Zahl der positiv getesteten Werte											9
Absolute Zahl der negativ getesteten Werte											5
Relative Bilanz (Anteil der positiv getesteten Werte in Prozent)											64,3

Umfragedaten aus: www.wahlrecht.de/Umfragen, im Internet am 31.8.2005; eigene Berechnungen; © Joachim Hofmann-Göttig

18. Landtagswahl in Baden-Württemberg am 25.3.2001

Institut	Quelle	Zeitraum	Datum	CDU	SPD	Grüne	FDP	REP	Sonstige	Summe
Amtl. Ergebnis			25.03.2001	44,8	33,3	7,7	8,1	4,4	1,7	100,0
Forsa	Stuttgarter Zeitung		20.03.2001	41,0	32,0	8,0	9,0	6,0	4,0	100,0
Infratest dimap	SWR		15.03.2001	41,0	35,0	8,5	8,5	5,0	2,0	100,0
Diff. Forsa				-3,8	-1,3	0,3	0,9	1,6	2,3	10,2
Diff. Infratest				-3,8	1,7	0,8	0,4	0,6	0,3	7,6
Mittlere Diff.				3,8	1,5	0,6	0,7	1,1	1,3	8,9
Statistische Fehlertoleranz in Prozentpunkten				3,1	2,9	1,7	1,7	1,3	0,9	11,6
Absolute Zahl der positiv getesteten Werte										8
Absolute Zahl der negativ getesteten Werte										4
Relative Bilanz (Anteil der positiv getesteten Werte in Prozent)										66,7

Umfragedaten aus: www.wahlrecht.de/Umfragen, im Internet am 31.8.2005; eigene Berechnungen; © Joachim Hofmann-Göttig

19. Landtagswahl in Rheinland-Pfalz am 25.3.2001

Institut	Quelle	Zeitraum	Datum	CDU	SPD	Grüne	FDP	Sonstige	Summe
Amtl. Ergebnis			25.03.2001	35,3	44,7	5,2	7,8	6,9	99,9
Infratest dimap	SWR-Ländersache		15.03.2001	35,0	43,0	6,0	9,0	7,0	100,0
Diff. Infratest				-0,3	-1,7	0,8	1,2	0,1	4,1
Mittlere Diff.				0,3	1,7	0,8	1,2	0,1	4,1
Statistische Fehlertoleranz in Prozentpunkten				3,0	3,1	1,4	1,7	1,6	10,8
Absolute Zahl der positiv getesteten Werte									5
Absolute Zahl der negativ getesteten Werte									0
Relative Bilanz (Anteil der positiv getesteten Werte in Prozent)									100,0

Umfragedaten aus: www.wahlrecht.de/Umfragen, im Internet am 31.8.2005; eigene Berechnungen; © Joachim Hofmann-Göttig