Anatomie eines Umfrage-Desasters – oder: Wie geschätzt 40-50 Leutchen deutsche Politik steuern

Laut einer gestern veröffentlichten und viele hundert Male rezitierten Forsa-Umfrage favorisieren 76 Prozent der FDP-Wähler Rainer Brüderle für den Posten des FDP-Vorsitzenden im Vergleich mit Philipp Rösler. Nur acht Prozent sehen das nicht so. Auf welcher Basis entstehen eigentlich solche Umfragen, die die Politik unseres Landes entscheidend mitbestimmen? Im konkreten Fall liegt der Verdacht nahe, dass bestenfalls drei bis vier Dutzend Befragte zu anschließend medial und politisch bis auf die Knochen ausgeschlachteten Umfrageergebnissen führen.

[Bitte lesen Sie auch die m.M.n. spannenden Kommentare für ein besseres Verständnis und auch (berechtigte!) Kritik an meiner Argumentation!]

Zugegeben: Den glücklichsten Eindruck macht Philipp Rösler auch auf politische Laien wie mich nicht. Die gestern veröffentlichten Umfrageergebnisse des Meinungsforschungsinsituts Forsa fand ich aber aus analytischer Sicht doch interessant. Demnach halten 76 Prozent der FDP-Wähler Rainer Brüderle für den besseren FDP-Vorsitzenden als Philipp Rösler. Diese Umfrage wurde laut Google News digital rund 1000 mal in den letzten 24 Stunden aufgegriffen; Spiegel Online nannte es "ein Umfragedebakel", der Focus "ein Desaster", N24 "vernichtend". Die Liste ließe sich fortsetzen.

Ein vernichtendes Debakel ist das in der Tat. Aber für wen?

Denn zunächst passiert etwas ulkiges: Laut der Umfrage halten 76 Prozent der FDP-*Wähler* Brüderle für den besseren. So steht es in der Originalquelle und auch der Pressemitteilung. Die Nachrichtenagentur Reuters schreibt allerdings am 2.1. um 11:28h : "Nach einer neuen Umfrage des Meinungsforschungsinstituts Forsa für das Magazin "stern" halten 76 Prozent der *Parteimitglieder* den Vorsitzenden der FDP-Bundestagsfraktion, Rainer Brüderle, für den besseren Parteivorsitzenden. Um 15 Uhr liefert Reuters ein Update – und spricht erneut von "*Parteimitgliedern*". Um 15:34 Uhr korrigiert sich die Agentur, es handele sich um 76 Prozent der "FDP-*Wähler*".

Da haben aber laut Google News Schlagwortsuche 232 Nachrichtenquellen, darunter Focus, Spiegel, Handelsblatt, Welt… die Agenturmeldung schon prominent online aufgegriffen und sprechen davon, 76 Prozent der *Parteimitglieder* hielten Brüderle für den besseren. Und so steht das auch jetzt noch bei den meisten online, und auf dieses Basis wurde munter analysiert und kommentiert. Ein, wie ich finde, nicht unwesentlicher Unterschied, wenn man kurz vor einem wegweisenden Parteitreffen stundenlang Wähler und Mitglieder durcheinanderwirft.

Aber gut: 76 Prozent der *FDP-Wähler* halten Brüderle also für den besseren, auch wenn das nicht überall durchgedrungen ist.

Doch wie befragt man eigentlich FDP-Wähler? Ausweislich der Originalquelle – die Kollegen vom "Stern" und "RTL" haben dazu das Meinungsforschungsinstitut Forsa beauftragt – wurden für diese Umfrage "*1003 repräsentativ ausgesuchte Bundesbürger*" befragt. Also nicht 1003 repräsentativ ausgesuchte FDP-Wähler. Und nicht Wahlberechtigte. Sondern: Bundesbürger.

Nun plätschtert die FDP in der Sonntagsfrage laut den ebenfalls von Forsa ermittelten Wahltrends derzeit bei rund vier Prozent. Das heißt: Wenn ich 1003 Bundesbürger repräsentativ aussuche, lande ich aktuell bei rund 40 Menschen, die sich derzeit als FDP-Wähler zu erkennen geben. [Das ist natürlich nur eine grobe Schätzung*, wir wissen weder, wie FDP-Wähler definiert sind, noch, wie mit Nichtwählern umgegangen wird]

Bleiben wir aber bei der Schätzung von vier Prozent von 1003, denn es heißt ja übereinstimmend bei den Instituten, dass die FDP in der Sonntagsfrage auf vier Prozent käme. Und diese so ermittelten 40 FDP-Wähler frage ich dann, ob sie Herrn Brüderle eventuell für den besseren FDP-Vorsitzenden halte? Ein interessanter Ge-

danke. Es können aber auch – das mal als wilde Theorie – 50 gewesen sein, wie eine kleine Zahlenknobelei nahe legt **. Es könnten aber, ganz optimistisch geschätzt, auch 140 gewesen sein. Zum Beispiel dann, wenn ich als FDP-Wähler nicht definiere, wer jetzt FDP wählt, sondern, wer die FDP 2009 gewählt hat, wenn ich die Nichtwähler komplett auslasse und somit eine große Frust-Gruppe hinzu nehme (danke für die Hinweise dazu). Es könnten auch methodisch definierte "Stammwähler" sein. Aber derzeit ist die Gruppe der FDP-Wähler ja nun mal vier Prozent, und dabei werden die schon "extrapoliert", denn in der politischen Stimmung schneidet die FDP noch *schlechter* ab als die vier Prozent, nämlich, je nach Institut, neigen derzeit nur zwei bis drei Prozent der FDP zu.

Entscheidend wäre aus methodischer Sicht nicht nur, wie viele FDP-Wähler befragt wurden. Sondern auch, wie die Frage gestellt wurde: Lautete sie: Wen fänden Sie besser als FDP-Vorsitzenden: Brüderle oder Rösler? Oder lautete sie: Fänden Sie Herrn Brüderle besser als FDP-Vorsitzenden (als Rösler)? Diese letztgenannte Frage hätte für mich methodisch den Charakter einer Suggestivfrage, und ich fürchte, genau so ist sie gestellt worden, denn im Originaltext heißt es: "Nur acht Prozent sehen das nicht so" [" Brüderle wäre ein besserer Vorsitzender als Rösler", das klingt nach ja/nein-Frage]. Ob die Leute auch gefragt wurden, ob sie Rösler für den besseren Parteichef als Brüderle halten, um Verzerrungen herauszunehmen, wissen wir leider nicht. Jedenfalls: Wenn Sie 100 Leute im Restaurant fragen, ob sie lieber Fisch oder Fleisch hätten, erhalten Sie auch andere Ergebnisse als bei der Frage, ob sie statt des Fischs lieber das Fleisch hätten.

Gewiss: Für das Urteil, dass Rösler nicht den besten Stand parteiintern hat, muss man kein Politikexperte sein. Dazu bedarf es auch keiner Umfrage. Aber: Mir dreht sich der Magen um, wenn ich darüber nachdenke, dass die Befragung von vielleicht 40 bis 50 "zufällig ausgewählten" oder in einer Kohorte ermittelten FDP-Wählern derart meinungsbildend wirkt. (Disclaimer: Wenn meine These stimmt, dass hier nur von einer anfänglichen Stichprobe von 1003 Leuten weitergerechnet wurde). Methodisch mag das gerade noch akzeptabel sein. Wäre es auch, wenn man sagt: FDP Wähler ist der, der 2009 FDP gewählt hat und somit die Grundgesamtheit etwas, aber vermutlich nicht dramatisch höher wäre bei einer 1000er-Stichprobe als 30 bis 50. Wenn es stringent wäre, was heraus kam.

Ist es das? Die Frage Rösler vs. Brüderle wurde auch schon in früheren Umfragen gestellt. Und jetzt wird es lustig. Denn als Datenbasis für die gestern veröffentlichte Zahl, nach der 76 Prozent der FDP-Wähler Brüderle für den besseren halten, diente eine Forsa-Umfrage, die laut Originalquelle am 20. und 21. Dezember durchgeführt wurde unter eben jenen 1003 Befragten.

Am 2. Weihnachtstag schrieb der Spiegel noch online , gerade einmal 48 Prozent der FDP-Wähler hielten Brüderle für den besseren FDP-Chef. Herausgefunden hat das – Forsa. Und befragt hat man dazu Menschen zwischen 17. und 21. Dezember, dafür aber auch 2500. Andere Quellen sprechen von 48 Prozent der "starken" FDP-Wähler, die Brüderle favorisieren.

Man kann es fast nicht glauben, aber ein und dasselbe Meinungsforschungsinstitut ermittelt (vorausgesetzt, die Zahlen stimmen) vom 17. bis 21. Dezember, dass 48 Prozent der FDP-Wähler Brüderle für den besseren Parteichef halten, um am 2. Januar gibt das Institut bekannt, dass man am 20. und 21. Dezember ermittelt habe, dass 76 Prozent der FDP-Wähler Brüderle für den besseren halten.

Welches Schweinderl' hätten's denn jetzt gerne – die 76 Prozent, die ein Debakel für Rösler und ein Triumph für Brüderle sind und vom 20. bis 21. Dezember erhoben wurden? Oder die zwischen dem 17. und 21. Dezember ermittelten 48 Prozent, die ein Debakel für Brüderle wären, weil er's offenbar auch nicht besser kann laut der befragten (starken?) FDP-Wähler? Denn interessanterweise konnte Brüderle damit ein ganzes Jahr lang nicht von der allerorts diganostizierten schlechten Arbeit Röslers profitieren: Vor ziemlich genau einem Jahr hielten auch 48 Prozent Brüderle für den besseren. Damals war das übrigens, ich zitiere n-tv, "kein klares Votum".

Und als wäre das alles nicht schon irre genug, macht das "Handelsblatt" heute (Freitag) seine Zeitung mit der Geschichte auf: "Brüderle ante portas", und als Beleg (samt Grafik auf der Seite 1) werden diese beiden Zah-

len verglichen: “Während 76 Prozent der der FDP Wähler Brüderle für den besseren Parteichef halten…. (…) Im Dezember 2011 hatten nur 48 Prozent der FDP-Anhänger für einen Parteichef Brüderle vortiert.”

Statistisch wäre jedenfalls die Streuung von schlappen 28 Prozentpunkten zwischen 48 und 76 Prozent keine so große Überraschung. Weder in einer Umfrage, noch, wenn sie auf Basis der Unterscheidung starker/schwacher FDP-Wähler zustande käme. Die Kollegen der “ZEIT” haben in einem ebenso kurzen wie treffenden Artikel vor knapp einem Jahr auf die statistischen Probleme hingewiesen, die sich stellen, wenn eine Partei – wie die FDP - in Befragungen in die Nähe der “Zufälligkeitsgrenze” rückt in Meinungsumfragen. Das sehr schöne Stück ist hier nachzulesen, unter drei Prozent spricht der Chef eines Meinungsforschungsinstituts von “Zahlenartefakten”.

Hinzu kommt demokratisch, dass die Veröffentlichung von Meinungsumfragen einen signifikanten Einfluss auf das tatsächliche Wahlverhalten und die Meinungsbildung hat. Vereinfacht: Wenn veröffentlicht wird, dass eine bestimmte Partei den Einzug in ein Parlament nicht schaffen wird oder ein Politiker besonders schlechte Umfragewerte erhält, dann führt dies dazu, dass die Menschen diese Partei tendenziell nicht wählen werden und diesen Politiker ebenfalls schlecht beurteilen. Denn sie sind mit ihrem Wahlverhalten und ihrer Meinung gerne “mit der Mehrheit” unterwegs und ungerne mit einem Loser.

Zu diesem Problem gibt es die hübsche Anekdote, dass das Allensbacher-Institut die Republikaner 1992 und 1996 zweimal wider den vorliegenden Umfragen auf unter fünf Prozent taxiert hat, um Wähler von einem Kreuz bei den Partei abzuhalten. Und wie fahren wohl die Delegierten zum Dreikönigstreffen der FDP, wenn kurz zuvor die Runde macht, mehr als drei Viertel der Leute da (der *Mitglieder*) halten Brüderle für den besseren? Das hat für mich – und ich habe zur FDP, Rösler und Brüderle überhaupt keine persönliche oder politische Meinung – den Charakter, einem taumelnden Menschen auf dem Weg zur Bühne mal eben mit Anlauf ins Kreuz zu springen.

Spannend wäre, zu wissen, woher der dramatische Anstieg der “Brüderle-Fans” von 48 auf 76 Prozent kommt, wenn man den Befragungszeitraum von 17. bis 21.12. mit dem 20. und 21.12. vergleicht. Oder erklärt der sich nur aus der Unterscheidung zwischen “starken” und “schwachen” FDP-Wählern und stammt aus des statistisch erheblich verlässlicheren 2500er-Stichprobe? Und überhaupt: Wie viele FDP-Wähler wurden denn nun befragt?

Zur Größe der Stichprobe habe ich Donnerstag bei Forsa angerufen – leider war aber niemand da, der meine Frage beantworten konnte. Ich wurde gebeten, eine kurze Mail zu schicken. Das habe ich getan mit einer sehr einfachen Frage: Wie groß war die Grundgesamtheit (N) der FDP-Wähler, auf die Sie Ihre Unfrageergebnisse stützen? ~~Ich habe die Frage heute früh erneuert – allerdings keine Antwort erhalten. Sollte die Zahl nachgereicht werden, trage ich sie hier natürlich nach.~~

Nachtrag: Forsa teilt mit, die Grundgesamtheit der FDP-Wähler betrage 6,3 Millionen. Das stimmt natürlich. Ich habe daraufhin meine Frage nochmals präzisiert. Mich interessiert natürlich die Zahl der FDP-Wähler der Stichprobe, auf deren Basis die 76 Prozent errechnet wurden. Stand Dienstag, 14 Uhr habe ich darauf noch immer keine Antwort erhalten.

* Wenn ein Meinungsforschungsinstitut veröffentlicht, dass vier Prozent der 1003 Befragten am Sonntag die FDP wählen werden, heißt das nicht, dass tatsächlich 40 Leute die “Sonntagsfrage” entsprechend beantwortet haben. Das geht schon mal los bei der Frage, wie man überhaupt Wähler definiert. Stammwähler? Der Partei, der er gerade zuneigt? Die er gewählt hat? Die er wählen wird? Oder: Die er vermutlich wählen wird, weil man als Meinungsforscher (und das tun sie ja wirklich) besser weiß, was jemand wählen wird, selbst wenn er das selbst noch nicht weiß? Denn egal, wie ich die Stichprobe konzipiere – als Zufallsstichprobe (Straße/Telefon), bei denen 1000 Leute zufällig ausgewählt werden – oder, indem ich 1000 Leute definiere, die laut meinen Daten ein Abbild der Bevölkerung bilden: Ich habe immer Verzerrungen drin. Konkret: Die CDU/CSU hat viele Wähler in der älteren Generation, die leichter erreichbar sind, wenn ich eine Zufallsstichprobe per Telefon mache. Junge Leute wählen oft “grün”, sind aber nur schwer erreichbar. “Linke” und “Grüne” sind

skeptischer, was Meinungsumfragen angeht aus Datenschutzgründen. NPD-Wähler outen sich ungern am Telefon oder auf der Straße in Befragungen. Ostdeutsche sind skeptischer. Und Leute, die sehr gerne an Meinungsumfragen teilnehmen am Telefon oder gegen Geld, sind ein anderer Schlag Mensch als der "Durchschnittsbürger". So etwas wird dann nach einer geheimen Formel "geglättet".

** Eine kleine statistische Spekulation: Die Umfrage spricht davon, dass 76 Prozent Brüderle für den besseren FDP-Chef halten, acht Prozent sehen das nicht so, was mit den restlichen 16 Prozent ist [100-[76+8]], steht in der Originalquelle nicht. Was zu einer kleinen Zahlenspielerei einlädt: 76, 16 und 8 Prozent – vielleicht handelt es sich ja nicht um rechnerisch rund 40 befragte FDP-Mitglieder, sondern tatsächlich um genau 50. Denn man kann glatt und ohne Runden 76, 16 und 8 Prozent von 50 Befragten bilden: 38 Befragte fänden demnach Brüderle besser, 4 eher Rösler, 8 haben keine Meinung [korrigiert, danke]. Plausibel wäre das bei einer 1000er-Befragung. Oder um 100.

Disclaimer: Ich arbeite bei den G+J Wirtschaftsmedien, die wiederum – wie der "Stern" als Co-Auftraggeber der Umfrage – zum G+J Verlag Hamburg gehören. Dieser Artikel analysiert ausschließlich die Methodik der Forsa-Umfrage und soll keinerlei Kritik an den sehr geschätzten Kollegen des "Stern" sein!

About these ads