

„Meinungsforschung: Wissenschaft oder Lotterie?“

Markus Bohn, Kornwestheim

Immer wenn im Bund oder in einem Bundesland Wahlen anstehen, dann dürfen professionelle Wahrsager ihre Kunst einer breiten Öffentlichkeit werbewirksam vorführen und sie bekommen dafür auch noch einen hübschen Batzen Geld. So zwischen 50.000 und 100.000 DM. Diese gutbezahlten Reklameveranstaltungen sind dabei für das Orakel fast ohne Risiko: Denn wenn es etwas Falsches prophezeit, dann hat eben der Wähler in letzter Minute seine zuvor bekundete Meinung geändert. Irgend ein Anlaß, der für diesen Meinungsumschwung verantwortlich gemacht werden kann, findet sich in der heißen Phase eines Wahlkampfes allemal. Aber wenn die Prognose einigermaßen hinlänglich ist, und das ist tatsächlich oft der Fall, dann läßt sich damit eindrucksvoll vorführen, wie zuverlässig so ein Meinungsforschungsinstitut arbeitet. Diese vermeintliche Zuverlässigkeit wird dann oft auch für andere Umfragen, z.B. für kommerzielle Marktanalysen in Anspruch genommen, und viele Kunden der Prozenzorakel glauben sogar daran.

Vor knapp drei Wochen hat unser badenwürttembergischer Regierungschef das Ergebnis einer Umfrage bekanntgegeben, die ein Frankfurter Institut im Auftrag des Staatsministeriums angestellt hatte. Auf's zehntel Prozent genau werden den vier Landtagsparteien hier ihre Stimmanteile für den 20. März vorhergesagt. Lothar Späth freut sich über die für ihn günstige Prognose besonders, weil dieses Frankfurter Institut vor vier Jahren den CDU-Anteil in der Tat auf 0,1 Prozent genau getroffen hat. Dieter Spöri von der SPD hingegen hält die Zahlen für „getürkt“. Und der neue Regierungssprecher wiederum kontert diesen Vorwurf mit dem Hinweis, daß „Herr Spöri überhaupt nicht weiß, was es mit diesen Umfragen auf sich hat.“

Nun, offenbar ist keiner der Herren mit den statistischen Gesetzen vertraut, die solchen Weissagungen zugrundeliegen. Denn daß die Wahlpropheten das letzte Mal beim CDU-Anteil ins Schwarze getroffen haben, ist Zufall. Und daß solche Zahlen speziell bei Wahlprognosen getürkt werden, das muß man nicht erst unterstellen, sondern das ist gängige Praxis. Das machen die Meinungsforscher übrigens, ohne daß man ihnen dazu eigens einen Auftrag erteilt. Nur nennt man

das in der Branche nicht „Türken“, sondern man spricht wissenschaftlich vom „Gewichten“. Wir kommen darauf noch zurück.

Zunächst einmal gilt es aber zu begründen, weshalb das purer Zufall ist, wenn eine Wahlprognose wenigstens bei einer Partei aufs zehntel Prozent genau stimmt. Sehr eindrucksvoll ist dies in einer umfangreichen Arbeit des Wuppertaler Mathematikprofessors Fritz Ulmer beschrieben, die demnächst veröffentlicht wird. Professor Ulmer hat damit zwar keine prinzipiell neue Erkenntnis zutage gefördert, aber es ist sein Verdienst, daß er mit Nachdruck an grundlegende Gesetze der Statistik erinnert und anhand von umfangreichen Computerberechnungen beispielhaft zeigt, was Meinungsforschung kann und was sie nicht kann.

Im Mittelpunkt seiner Kritik steht dabei ein Begriff, den man eigentlich als Etikettenschwindel bezeichnen muß, und das ist der sogenannte „repräsentative Querschnitt“. Die meisten werden mit diesem Begriff „repräsentativ“ die Vorstellung verbinden, hier handle es sich sozusagen um ein verkleinertes Abbild der Gesellschaft, also um eine gezielte Auswahl von Bundesbürgern, die nach Alter, Geschlecht, Bildung, Einkommen usw. die Gesamtheit repräsentieren und deren Meinung deshalb auch zuverlässig die Meinung aller widerspiegelt, sei es jetzt auf ein einzelnes Bundesland bezogen oder auf die gesamte Republik. Aber genau das stimmt eben nicht.

Wer in diesen repräsentativen Querschnitt aufgenommen wird und somit stellvertretend für alle Gleichgesinnten die Fragebögen der Meinungsforscher ausfüllen darf, das bestimmen nicht irgendwelche Soziologen aufgrund bestimmter Merkmale, sondern das bestimmt im besten Fall allein der Kollege Zufall. Professor Ulmer verwendet hier bewußt und wirklich völlig zurecht den Begriff „Lotterie“. Der sogenannte repräsentative Querschnitt wird per Lotterie zusammengestellt und das ist einer der Gründe, wenn auch nicht der einzige, weshalb das Ergebnis zwar zufällig richtig sein kann, aber beileibe nicht richtig sein muß.

Nun, warum lösen die Meinungsforscher 1.000 oder 2.000 Bundesbürger aus, an die sie bei einer

Umfrage herantreten? Warum versucht man nicht, tatsächlich ein verkleinertes Abbild der Gesellschaft auszuwählen? Die Antwort ist einfach: erstens ist das unmöglich und zweitens würde es auch gar nichts nützen. Unmöglich ist solch ein Abbild schon rein von den Zahlen her. Selbst wenn man sich z.B. allein auf die Merkmale Geschlecht, Alter, Konfession, Schulbildung, Familienstand, Beruf, Wohngegend und Einkommen beschränken würde, so gibt es hier schon weit mehr als 1.000 Kombinationsmöglichkeiten. Es kann also gar nicht jede dieser Möglichkeiten bei einer Umfrage vertreten sein.

O-Ton: „Das ist vollkommen hoffnungslos, mit 1.000 oder 2.000 Befragten ein verkleinertes Abbild bezüglich 40, 50, 60 oder 100 Merkmalen oder noch mehr zu erstellen. Das ist eine Illusion, das hat die Statistik den Meinungsforschern nie versprochen.“

Hinzu kommt aber, und das ist mindestens genauso wichtig, daß man weder aus Einzelmerkmalen noch aus einer bestimmten Kombination von Merkmalen auf die Meinung des Betreffenden schließen kann. Ein bestimmter verheirateter katholischer Bäcker mit Mittlerer Reife ist halt nicht für alle katholischen Bäcker mit Mittlerer Reife repräsentativ.

Und deshalb bleibt tatsächlich gar nichts anderes übrig als auszulosen, wen man befragt, nach der Devise: Je zufälliger, desto besser. Es gibt hier verschiedene Lotteriemethoden, und die sind durchaus von unterschiedlicher Güte. Aber das soll uns hier nicht weiter interessieren. Am Prinzip ändert sich dadurch nicht viel.

Nehmen wir einmal an, es handelt sich tatsächlich um eine perfekte Zufallsauswahl, bei der niemand bevorzugt oder benachteiligt wird. Das ist das Beste, was überhaupt passieren kann. Wie groß ist dann die Unsicherheit einer per Umfrage ermittelten Prozentzahl?

Um dies zu klären, hat Professor Ulmer Meinungsumfragen zur Bundestagswahl auf dem Computer simuliert. Er hat dazu die tatsächlichen Wahlergebnisse eingegeben, also so getan, als hätte er alle abgegebenen Stimmzettel in einer riesigen Lostrommel. Dann ließ er den Computer viele Male jeweils 1.000 oder 2.000 Stimmzettel ziehen nach einer reinen Zufallsauswahl. Jede dieser Ziehungen entspricht dabei einer Umfrage, die unter absolut idealen Bedingungen durchgeführt wurde, wie man sie in Wirklichkeit nie vorfindet. Keiner, der die Auskunft verweigert; keiner, der sich noch nicht entschieden hat; keiner, der es sich dann am Wahltag doch noch anders überlegt. Sondern lauter ideale Stichproben bzw. in Anführungszeichen „repräsentative Querschnitte“ der tatsächlich abgegebenen Stimmen.

Nun, wie groß ist hier der mögliche Fehler, wie verlässlich sind die so ermittelten Prozentzahlen? Diese Frage liegt nahe, aber sie ist falsch gestellt. Professor Ulmer:

O-Ton: „Fehler ist eigentlich nicht das richtige Wort, weil der Eindruck entsteht, als hätte man ungenau gemessen - was mit mehr Sorgfalt zu vermeiden wäre. Hier handelt es sich aber um ein anderes Phänomen. Die Befragten werden ja ausgelost. Und was heißt bei einer Lotterie schon Fehler? Die Abweichung kann mal gravierend sein, mal überhaupt nicht ins Gewicht fallen. Man mußte also eher wie beim Lotto von Chance sprechen und dann die Frage stellen: Wie groß ist die Chance, daß eine Prognose von einem bestimmten Typ richtig ist?“

Also: Nehmen wir einmal an: 1.000 Befragte, vier Parteien, und das Ergebnis soll auf 0,1 Prozent genau stimmen:

Wie groß ist dann die Wahrscheinlichkeit, daß man diese Genauigkeit bei einer einzigen Ziehung aus unserer fiktiven Trommel mit den tatsächlich abgegebenen Stimmzetteln erreicht? Professor Ulmer hat ermittelt: Diese Wahrscheinlichkeit ist praktisch Null, um genau zu sein: 0,006 Prozent. Aber auch wenn man das Ergebnis nur auf ganze Prozent genau wissen will bei den beiden großen Parteien und aufs halbe Prozent genau bei der FDP und bei den Grünen, dann beträgt die Chance, dieses Ergebnis zu treffen, nur 1,2 Prozent. Das heißt, selbst unter diesen idealen Umfragebedingungen ohne jeden Interviewfehler, ohne Falschangaben, ist eine aufs Prozent genau angegebene Prognose mit fast 99-prozentiger Wahrscheinlichkeit falsch. Nur unrealistisch große Spielräume könnten den Meinungsforschern aus dieser mißlichen Lage helfen:

O-Ton: „Wenn Sie eine Wahlprognose mit plus oder minus drei Prozent abgeben für diese vier Parteien - plus-minus drei Prozent für die großen und plus-minus zwei Prozent für die kleinen oder so was -, dann ist kein Mensch mehr an Ihrer Prognose interessiert. Plus-minus drei Prozent für die CDU würde bedeuten, ja irgendwo zwischen 42 und 48 Prozent. Würden Sie dafür 50.000 Mark ausgeben? Und selbst wenn die Leute jetzt bei Parteien ein solche Aussage machen: CDU auf plus-minus drei Prozent, SPD auf plus-minus drei Prozent und Grüne und FDP auf plus-minus zwei Prozent, dann ist diese Aussage immer noch mit einer Wahrscheinlichkeit von vielleicht 20 Prozent falsch.“

Nun werden Sie zu Recht einwenden, das kann doch alles gar nicht wahr sein, schließlich fallen die Wahlprognosen ja gar nicht so schlecht aus, wie es der Mathematikprofessor Ulmer aufgrund der

Theorie und aufgrund seiner Computersimulation vorhersagt. Richtig: Aber deshalb stimmt die Theorie, stimmen die Simulationen trotzdem. Nur: Wahlprognosen werden allesamt „getürkt“, sprich „gewichtet“. Die Meinungsforschungsinstitute haben dazu verschiedene „Hokuspokusysteme“ entwickelt, die auf den ersten Blick durchaus einen wissenschaftlichen Eindruck erwecken, in Wahrheit aber Taschenspielertricks sind. Daß Wahlprognosen oft recht genau ausfallen, wenigstens hierzulande, hat einen ganz einfachen Grund:

O-Ton: „In der Bundesrepublik sind die Verhältnisse relativ stabil. Da geht es mal bei einer Partei einige wenige Prozente rauf und runter. Und Sie wissen auch aus historischen Gründen, in welchen Größenordnungen sich die CDU und die SPD bewegen; und diese Information, die da eingeht, ist mehr als aus einer Meinungsumfrage herauskommen kann. Das ist in de facto wirklich mehr eine Fortschreibung an das letzte Wahlergebnis. Ich meine, was die messen und was die hinterher publizieren oder auf den Bildschirm bringen, das hat doch sehr wenig miteinander zu tun. Die Sache ist weitgehend spekulativ und sie funktioniert von Zeit zu Zeit, einfach weil die Verhältnisse stabil sind. Wenn Sie große Veränderungen hätten, wie manchmal in England oder in Frankreich, von Wahl zu Wahl, dann würden den Leuten die Felle sehr bald davonschwimmen.“

Elisabeth Noelle-Neumann wird im „Rheinischen Merkur“ mit folgender Aussage zitiert:

„Zwischen dem, was wir an Rohergebnissen erhalten, und dem, was wir als Prognose veröffentlichen, liegt manchmal eine Differenz von zehn oder elf Prozent. Denken Sie sich, wir würden die um zehn oder elf Prozent abweichenden Ergebnisse veröffentlichen!“

Das heißt im Klartext: Das, was per Umfrage tatsächlich ermittelt wird, die sogenannten „Rohergebnisse“, liegen oft um runde 10 Prozent daneben, wenn das veröffentlichte Ergebnis ungefähr hinkommt. Und das entspricht in der Größenordnung recht gut der Fehlerbreite, die von Statistikexperten diesem Instrument Meinungsumfrage zugebilligt wird. In fast allen Bereichen, wo solche Umfragen heute erhoben werden, in der Marktforschung z.B., oder auch wenn die Hörerzahlen der Rundfunkanstalten ermittelt werden, muß man mit Ungenauigkeiten in dieser Größenordnung rechnen. Denn hier kommt zu dem Lotteriefehler, der durch die Auslosung des sogenannten „repräsentativen Querschnitts“ bedingt ist, noch der Interviewfehler hinzu. Der entsteht dadurch, daß manche Befragten unwissentlich oder manchmal auch absichtlich falsche Antworten geben.

Nun, den Meinungsforschern selbst kann dies egal sein, solange sie trotzdem Kunden haben, die ihren Prozentzahlen Glauben schenken. Anders als bei Wahlen „schlägt bei solchen Marktanalysen keine Stunde der Wahrheit“, lassen sich die Prognosen nicht an der Realität messen. Und mehrere unabhängige Parallelumfragen leistet sich wohl kaum ein Auftraggeber.

So bleibt die Streubreite solcher Umfragen den Auftraggebern meist verborgen. Größte Vorsicht ist geboten, wenn Meinungsforscher behaupten, es sei etwas „statistisch signifikant“, z.B. ein Trend, der aus zwei Umfragen in einem gewissen zeitlichen Abstand abgeleitet wird.

O-Ton: „Da werden einfach so Milchmädchenrechnungen durchgeführt, die der Komplexität des Problems - das ist eigentlich der entscheidende Punkt - in keiner Weise angepaßt sind. Das sind so ganz einfache Rechnungen über Signifikanz, die unter Annahmen gemacht werden, bei denen man sich - wenn man sich das wirklich überlegt - an den Kopf greifen muß und sagt, das hat einfach keinen Sinn.“

In unserer Wirtschaftsordnung bestimmt bekanntlich die Nachfrage das Angebot. Den Meinungsforschern ist deshalb kaum ein Vorwurf zu machen. Die Mahnung des Mathematikers richtet sich deshalb auch mehr an die Kunden:

O-Ton: „Man will eine Prozentzahl und man kriegt diese Prozentzahl, aber sie hat keinerlei wissenschaftliche Basis. Das ist eine reine Spekulation. Was man machen muß ist: man muß die Erwartungen reduzieren. Dieses Instrumentarium vermag nicht das zu leisten, was die Leute von ihm erhoffen. Die Leute erhoffen das von ihm, weil Jahr für Jahr diese Prozentzahlen oft mit einer Nachkommastelle auf dem Bildschirm oder in der Zeitung abgedruckt werden. Und diese Zahlen erwecken eine Genauigkeitsvorstellung, die grundsätzlich unmöglich ist mit diesen Meßmethoden, das ist der entscheidende Punkt. Das ist grundsätzlich unmöglich, mit diesen Meßmethoden eine solche Genauigkeit zu erzielen.“