

Orakel der Neuzeit: Was leisten Wahlbörsen, Wählerstromanalysen und Wahltagshochrechnungen?

„Eine Wahl ist eine Veranstaltung zur Überprüfung demoskopischer Vorhersagen.“
(Robert Lembke)

Wahlbörsen, Wählerstromanalysen und Hochrechnungen werden zwar in der Öffentlichkeit viel beachtet, in der politikwissenschaftlichen Forschung aber nur selten diskutiert, weil die Methodologie jeweils komplex und ungewohnt ist. Alle Verfahren weisen eine Reihe spezifischer Probleme auf, die die Wahlsoziologie nur in Kooperation mit den Experimental Economics, aus denen die Wahlbörsen kommen, und mit der Statistik, aus denen die Wählerstromanalyse kommt, lösen kann.

Wahlbörsen haben nach ihren anfänglichen Prognose-Erfolgen eine Größe und eine Öffentlichkeitswirkung erlangt, die die Vorhersagegenauigkeit vermutlich wieder stark beeinträchtigen. In unserem Beitrag fassen wir die jüngste Geschichte der Political Stock Markets zusammen und geben einige Design-Empfehlungen.

Wählerstromanalysen, die auch die Basis für Wahltagshochrechnungen liefern, sind mit einer Reihe von methodischen Herausforderungen konfrontiert. Wir belegen mit österreichischen Wahldaten die andernorts geäußerte Vermutung, dass einer adäquaten Gruppierung von Gemeinden zur Vermeidung eines aggregate bias die höchste Aufmerksamkeit geschenkt werden sollte.

Einleitung: Wozu Wahlprognosen?

Eine Prognose ist die Vorhersage von in der Zukunft liegenden Entwicklungen auf Basis von in der Gegenwart vorliegenden Informationen. Wenige Prognosen werden in der Öffentlichkeit so viel beachtet wie die Vorhersagen von Wahlergebnissen, schon allein wegen der viel diskutierten Frage, inwieweit deren Veröffentlichung das Verhalten der WählerInnen beeinflusst. Unter anderem deswegen sind die Anforderungen an die Genauigkeit oft ungleich höher als bei anderen sozialwissenschaftlichen Experten.

Wir beschreiben im ersten Abschnitt dieses Beitrags den in den Experimental Economics

entwickelten Prognose-Ansatz der sogenannten „Wahlbörsen“ oder *Political Stock Markets*. Erst seit zwei Jahrzehnten im Einsatz, trat nach ersten spektakulären Erfolgen eine gewisse Ernüchterung über die Vorhersagekraft der Methode ein – was möglicherweise gerade an den unerwünschten Nebeneffekten der Popularität dieses Instruments liegt, wie etwa Manipulationsversuchen durch Parteizentralen. Die im Anschluss geschilderte Methode der Wählerstromanalyse ist nach wie vor mit einer Vielzahl methodischer Herausforderungen konfrontiert, wobei unserer Ansicht nach die mit Abstand größte ist, Gebiete bzw. Gruppen von Gemeinden mit homogenen Trends zu definieren, in denen kein *aggregate bias* die Schätzergebnisse verzerrt. Obwohl bei der Wähler-

stromanalyse noch viele Verfahrensfragen heiß diskutiert werden, hat sich ihr Einsatz bei der Wahltagshochrechnung – einer speziellen Art der Prognose, die nur am Wahltag selbst aufgrund vorliegender Ergebnisse getätigt werden kann – bislang jedenfalls bewährt. Im dritten Abschnitt beschreiben wir, wie wir das Verfahren bisher bei Wahlenden eingesetzt und welche Erfahrungen wir unter diesen besonderen, für WissenschaftlerInnen ungewöhnlich spannenden Umständen gemacht haben.

1. Befragen oder handeln lassen?

In den 80er Jahren entwickelten Ökonomen an nordamerikanischen Universitäten¹ erstmals Wahlbörsen, zunächst als experimentelle Programme, um StudentInnen mit Börsemechanismen vertraut zu machen, später als öffentliche Experimente mit dem expliziten Anspruch, Wahlergebnisse vorherzusagen. Der in jedem Fall unbestrittene Erfolg der Wahlbörsen in Sachen Öffentlichkeitswirkung liegt an zwei angenehmen Eigenschaften:

- Wahlbörsen funktionieren bereits mit einem kleinen Kreis an TeilnehmerInnen – im Prinzip genügen bereits mehr als ein paar Dutzend –, die nicht repräsentativ für die wahlberechtigte Population sein müssen, was die Kosten relativ gering hält.
- Wahlbörsen liefern jeden Tag Ergebnisse, die breitenwirksam über Tageszeitungen oder das Internet publiziert werden können.

Exkurs: Funktionsweise der Wahlbörsen

Die den Autoren bekannten Wahlbörsen sind allesamt „Double Auction Märkte“: Die Kaufangebote (Bids) für eine bestimmte Partei werden hierbei in eine Warteschlange (Queue) gestellt, ebenso die Verkaufsangebote (Asks). Alle Verkaufsangebote, deren Preis unter den höchsten Kaufangeboten liegt, werden abgearbeitet (und umgekehrt), solange bis die beiden Queues sich nicht mehr überschneiden. In einem funktionierenden Markt sind die Queues in der Nähe des Marktpreises sehr dicht belegt,

d.h. es müssen z.B. relativ viele Kaufangebote gemacht werden, um den Preis deutlich zu erhöhen. Unrealistische Angebote führen meist gar nicht zu einer Transaktion.

Die HändlerInnen treten in den Markt ein, indem sie eine bestimmte Anzahl Pakete mit je einer Aktie aller kandidierenden Parteien² zu einem festgelegten Preis erwerben. Die „Bank“ kauft auch stets Bündel mit allen Parteien zum Marktpreis zurück.³ Alle bisher durchgeführten Wahlbörsen begrenzten die Marktmacht einzelner TeilnehmerInnen durch eine fixe Anzahl an Kapital bzw. Obergrenzen. Bei den Wahlbörsen, bei denen mit „echtem Geld“ gespielt wird, ist diese Begrenzung nicht zuletzt deswegen notwendig, weil diese eigentlich das staatliche Glücksspielmonopol in Frage stellen und nur durch niedrige Beträge (bzw. die Betonung wissenschaftlichen Interesses) die Kulanz der Behörden erwirken konnten.

Der Handel schließt stets, bevor die Wahlergebnisse bekannt werden, häufig um 0.00 Uhr des Wahltags. Der Prozentsatz in gültigen Stimmen nach Bekanntgabe des offiziellen Wahlergebnisses determiniert dann den Wert, zu dem die Bank die Parteien-Aktien von den HändlerInnen wieder zurückkauft.

Der Grundgedanke hinter den Wahlbörsen ist, dass Börsenaggregate Informationen widerspiegeln: Die HändlerInnen (oder neudeutsch „Trader“) nehmen aus ihrer Umgebung (Familie, Arbeit) und über Medien Stimmungen wahr, die sie zum Abstoßen oder Zukaufen von Parteienaktien⁴ animieren. Der aktuelle Börsenkurs ist ein Gleichgewicht, das als Instrument der Informationsgewinnung interpretiert werden kann: Je nach der Effizienz des Marktes reflektieren seine Preise entweder das Wissen dessen, der am besten informiert ist oder sogar ein noch höheres Niveau – laut ökonomischer Theorie spiegeln die Marktpreise ein Informationslevel wider, der das durchschnittliche Niveau der TeilnehmerInnen übersteigt (Smith 1982).

Grundsätzlich wird zwischen öffentlichen – also jedem/r zugänglichen – sowie privaten Informationen unterschieden. Für Wahlbörsen sind primär öffentliche Informationen relevant, denn kein/e HändlerIn kennt das Wahlergebnis im voraus, womit niemand über verlässliche

private Informationen verfügt. Es gibt zwar „politische Insider“, die etwa über unveröffentlichte Umfragen oder Stimmungsberichte aus Parteizentralen verfügen – nachdem sie das Wahlergebnis aber ebenso wenig im voraus kennen, handelt es sich nicht um „Insider“ im finanzwirtschaftlichen Sinn (Huber 2001, 226).

Da anzunehmen ist, dass auch öffentlich verfügbare Informationen nicht allen MarktteilnehmerInnen bekannt sind, dürften Informationsniveau und Erwartungen von HändlerIn zu HändlerIn stark variieren. Außerdem interpretieren nicht alle TraderInnen dieselbe Information auf die gleiche Weise. Selbst wenn MarktteilnehmerInnen eine Information (z.B. eine neue Umfrage) erhalten, kann es sein, dass sie diese nicht in ihre Erwartungen einbeziehen bzw. anders bewerten als andere Akteure: „Händler und Forscher gleichermaßen sind also nicht nur mit einer asymmetrischen Informationsverteilung konfrontiert, sondern sogar das Ausmaß dieser Asymmetrie ist unbekannt“ (Huber 2001, 225).

Niemand kennt das Wahlergebnis im Vorhinein und niemand kann mit Sicherheit bestimmen, welche Informationen relevant sind. Weder TeilnehmerInnen noch VeranstalterInnen können einschätzen, wie zutreffend die Erwartungen der einzelnen HändlerInnen sind. Dies impliziert ein hohes Maß an Unsicherheit für alle TeilnehmerInnen.

Wahlbörsen dürften zwar im Sinne der finanzwissenschaftlichen Theorie zu Vorhersagen führen, die an Genauigkeit die durchschnittlichen Vorhersagen der TraderInnen übersteigen und somit im ökonomischen Sinne Effizienzkriterien erfüllen, durch den inhärenten Mangel an echten Insidern können die Aktienkurse der Parteien jedoch weitab vom tatsächlichen Ergebnis liegen.

Wie erfolgreich sind Wahlbörsen als Prognose-Instrument?

Es gibt viele „Erfolge“, die sich Wahlbörsen jedenfalls auf die Fahnen heften können, wie etwa die Page Impressions auf den Betreiberseiten zu erhöhen, den Mechanismus eines Aktienmarktes auch außerhalb von Volkswirt-

schaftslehre-Seminaren anschaulich zu vermitteln, oder auch das allgemeine Politik-Interesse zu steigern. In diesem Beitrag wollen wir zur Evaluierung des Erfolges von Wahlbörsen jedoch nur ein Kriterium heranziehen: Sind Wahlbörsen bei der Vorhersage von Wahlergebnissen besser als (publizierte) Meinungsumfragen? Konkret: Liegt der Schlusskurs der Wahlbörsen-Aktien näher beim Wahlergebnis als die knapp vor der Wahl veröffentlichten Umfragen?

Als Benchmark für die Wahlbörse-Prognosen können nur veröffentlichte Umfragen herangezogen werden. Für Massenmedien erstellte Umfragen weisen meist kleine Stichproben auf und werden teilweise von Firmen erstellt, deren methodische Kompetenz nur bedingt als erwiesen gelten kann. In Österreich und auch in anderen Ländern werden gerade die am sorgfältigsten durchgeführten Umfragen von politischen Parteien beauftragt, die dieselben nur in Ausnahmefällen publizieren. Mit anderen Worten: Die Konkurrenz in punkto Prognosegenauigkeit ist für die Wahlbörsen insofern nicht überwältigend, als gerade die veröffentlichten Umfragen die weniger präzisen sind.

Grundsätzlich ist festzuhalten, dass Umfragen nur bedingt geeignet sind, den Ausgang einer Wahl vorherzusagen. Dies hat eine Reihe von völlig unterschiedlichen Gründen.

Zunächst haben bei Umfragen nicht alle Wahlberechtigten die gleiche Chance, ihre Präferenz in einem Interview mitzuteilen. Bei Telefonumfragen, der weitaus häufigsten Grundlage von veröffentlichten Prognosen, können etwa WählerInnen ohne Telefon oder mit Geheimnummer gar nicht erreicht werden.⁵ Unter denjenigen, die in einem Telefonverzeichnis stehen, weisen nicht alle das gleiche Verhalten bezüglich der Annahme von Anrufen auf – wer z.B. jeden Anruf zuerst vom Anrufbeantworter beantworten lässt, hat eine geringere Wahrscheinlichkeit, zu einem Interview überredet zu werden. Mobile Befragte werden deutlich seltener erreicht, auch bei den üblichen Standards von fünf Kontaktversuchen. In Zeiten zunehmender Single-Haushalte wird es auch immer schwieriger, über eine andere Person im Haushalt einen Interviewtermin zu vereinbaren. Wenn es schließlich doch zu einem Gespräch mit der

Zielperson kommt, verweigert stets ein Teil der Angesprochenen die Beantwortung politischer Fragen.

Nachdem alle diese Phänomene in bestimmten soziodemographischen Gruppen häufiger auftreten als in anderen, weist praktisch jede aktualisierte Stichprobe gegenüber der Population beträchtliche Verzerrungen auf. Die soziodemographischen Unebenheiten der Stichproben können (und müssen) zwar mit Gewichtsungsverfahren ausgeglichen werden, aber diese sind nur so zuverlässig wie die vorliegenden Angaben über die Population und führen außerdem dazu, dass in manchen gesellschaftlichen Gruppen das Gewicht einer Person mitunter mehr als verdoppelt werden muss.

Weiters ist die Umfrageforschung mit dem Problem konfrontiert, dass die Befragten in der Interview-Situation Angaben über zukünftiges Verhalten machen. Auch wenn die Befragten nach bestem Wissen und Gewissen antworten, können sie natürlich nicht vorhersehen, ob sie ihre Meinung bis zum Wahltag noch ändern bzw. ob sie tatsächlich wählen gehen werden – Nichtwählen gilt verbreitet als sozial unerwünschtes Verhalten, das aber in der Interview-situation durch Deklaration für eine Partei elegant geleugnet werden kann.⁶ Die Deklarationsbereitschaft ist schließlich nicht bei AnhängerInnen aller Parteien gleich hoch – in Österreich wurde etwa jahrelang der tatsächliche Support für die Freiheitliche Partei unterschätzt.⁷

Selbst wenn es nun aber ein perfektes WählerInnen-Telefonbuch gäbe, eine perfekte Ausschöpfung erzielbar wäre (d.h. mit jeder gezogenen Person könnte ein Interview durchgeführt werden) und jede/r Befragte genau das angäbe, was er/sie am Wahlsonntag wirklich tun wird, müssen wir die der schließenden Statistik inhärenten Konfidenzintervalle vulgo „Schwankungsbreiten“ berücksichtigen. Bei einem n von 1000 (was für eine veröffentlichte Umfragen gar nicht so wenig ist, in vielen Fällen liegt die Befragtenanzahl nicht über 500) beträgt die Schwankungsbreite einer Partei, die in der Stichprobe auf 40% kommt, +/-3%, d.h. der wahre Wert liegt mit einer 95%-igen Wahrscheinlichkeit zwischen 37% und 43%.⁸

Die Umfragen, die unmittelbar vor Wahlen veröffentlicht werden und hier als Vergleichsmaßstab für die Prognosegenauigkeit der Wahlbörsen herangezogen werden, beruhen auf Interviews, die oft bis zu einer Woche oder noch länger vor dem Wahltag durchgeführt werden. Bei Wahlbörsen kann – ganz abgesehen davon, dass sie von den geschilderten methodologischen Fallstricken der Umfrageforschung nicht betroffen sind – der Schlusskurs üblicherweise bis um Mitternacht vor dem Wahltag noch korrigiert werden, die HändlerInnen verarbeiten also quasi bis zur letzten Sekunde Informationen. Gute Bedingungen, so scheint es, um in punkto Prognoseperformance besser abzuschneiden.

Noch wurde das Versprechen allerdings nicht eingelöst, der Befund ist – zumindest in Europa – ernüchternd: Wahlbörsen prognostizieren im allgemeinen nicht besser als Umfragen (Huber 2001, 183ff.). Ein einheitlicher Trend ist schwer festzustellen – in manchen Fällen war die Vorhersagegenauigkeit der Börsen etwas besser (z.B. bei der österreichischen Bundespräsi-

Tabelle 1: Vergleich des Mean Average Error (durchschnittlicher Prognosefehler pro Partei) von Wahlbörsen und Umfragen bei österreichischen Wahlen

Wahlgang	Wahlbörsen	Umfragen
Nationalratswahl 1994	1,2	1,3
Nationalratswahl 1995	2,7	2,2
Landtagswahl Steiermark 1995	3,0	2,3
Gemeinderatswahl Wien 1996	1,2	1,3
Wahlen zum EU-Parlament 1996	2,3	2,0
Bundespräsidentenwahl 1998	1,7	2,2
Wahlen zum EU-Parlament 1999	2,3	1,7
Nationalratswahl 1999	0,9	0,6
Gemeinderatswahl Wien 2001	3,1	2,1

Quellen: Huber 2001, <http://ebweb.tuwien.ac.at/apsm>, eigene Berechnungen

dentenvahl 1998 oder bei der Wahl in Nordrhein-Westfalen 2000), hin und wieder fast exakt gleich (z.B. Dänemark 1990), und bisweilen auch deutlich schlechter (wie bei den österreichischen Nationalratswahlen 1995 und 1999). Den Autoren ist kein einziger Fall in Europa bekannt, wo bei einem „Versagen“ der Umfragen die Wahlbörsen einen Wahlausgang relativ genau vorhergesagt hätten. Mit anderen Worten: In Österreich brachten Wahlbörsen bis dato „keine signifikant besseren Prognoseergebnisse als Umfragen“ (Huber 2001, 98).

Können wir die Wahlbörsen daher ad acta legen? Noch nicht, denn es gab auch einige spektakuläre Erfolge zu vermelden, zum Beispiel bei den US-Präsidentenwahlen 1988 und 1992 (Forsythe et al. 1992). Beinahe punktgenaue Vorhersagen von fast unter Ausschluss der Öffentlichkeit durchgeführten Trading Experiments standen bei diesen beiden Wahlen in scharfem Kontrast zu Umfragen, die teilweise spektakulär weit neben dem tatsächlichen Wahlausgang lagen.

Aus dieser Pionierzeit der Wahlbörsen rührt wohl auch ihr Ruf als Prognose-Wundermittel. Jedoch bereits im Jahr 1996 wurden die Propheten dieser neuen Methode wieder auf den Boden der Realität zurückgeholt: Den Abstand zwischen Bill Clinton und Bob Dole prognostizierten Umfragen wie Wahlbörsen um rund neun Prozent vom tatsächlichen Ergebnis entfernt – diese Ernüchterung kam umso überraschender, als es sich dabei um die von der TeilnehmerInnenzahl, dem eingesetzten Kapital und der Laufzeit bis dato größte Wahlbörse der Geschichte handelte.

Aber gerade das war möglicherweise das Problem. Während in Umfragen bei gleich bleibender Stichprobenqualität die Anzahl der Befragten im Prinzip nie groß genug sein kann, dürften Wahlbörsen dagegen mit einer limitierten Zahl von HändlerInnen optimal funktionieren. Dieser zunächst überraschende Befund findet eine Erklärung darin, dass eine massive Zahl von platzierten Orders den Spielraum des/r Einzelnen einschränkt und so der Markt für die „Marginal Traders“, die von der Kumulation zahlreicher kleiner Transaktionsgewinne leben und damit die Kurse immer wieder in Richtung

ihrer Erwartungen drängen, unattraktiv wird (Huber 2001, 195ff.).

Auch ein zweites Problem tritt gerade erst durch den Erfolg der Wahlbörsen auf: Durch Medien, die ja oft als Betreiber auftreten,⁹ gelangen die Ergebnisse – im Gegensatz zu den Pionierbörsen der Experimental Economics – in Print- und Online-Ausgaben in Echtzeit an die Öffentlichkeit. Bereits 1995 hat sich erstmals eine Parteizentrale auf eine Wahlbörse gestürzt, um Kurse massiv nach oben zu treiben (Ortner 1996, 36). Diese Untugend (die übrigens sogar demobilisierende Effekte auf die eigene Wählerschaft haben könnte, was aber der mittleren Funktionärsschicht in der Hitze des Wahlkampfgefechts wohl selten bewusst zu sein scheint) setzte sich seitdem bei den meisten in der Öffentlichkeit bekannten Wahlbörsen fort und verzerrte die Ergebnisse zum Teil massiv.

Wenn die Wahlbörsen auch in Europa eine ernstzunehmende Konkurrenz zu den Umfragen bilden wollen, hieße die beste Strategie wahrscheinlich „back to the roots“, sprich nur teilweise publik gemachte Märkte mit einer begrenzten Zahl an HändlerInnen durchzuführen. Bei diesen muss sichergestellt sein, dass sie den Börsemechanismus verstehen¹⁰ und auf dem Markt vordringlich das Ziel verfolgen, ihren Gewinn zu maximieren, sei es aus monetären Interessen oder aus spielerischem Ehrgeiz. (Überraschenderweise dürften Wahlbörsen mit „Spielgeld“ beinahe gleich gut funktionieren wie Wahlbörsen, in die tatsächlich eingezahlt wird, Huber 2001, 74.)

In der Literatur wird immer wieder auf den vermeintlichen Vorteil der Wahlbörsen hingewiesen, dass es relativ gleichgültig ist, wie der HändlerInnenkreis soziodemographisch und von den politischen Präferenzen her zusammengesetzt sei. Dies widerspricht unseres Erachtens aber der theoretischen Grundannahme, dass Märkte in der Lage sind, in der Gesellschaft verstreute Informationen zu aggregieren. Überspitzt gesagt stellt sich die Frage, wie zum Beispiel ein Markt von Wiener StudentInnen einen Stimmungsumschwung in der westösterreichischen Landbevölkerung mitbekommen soll.¹¹ Eine gewisse regionale und soziodemographische Streuung ist daher wahrschein-

lich förderlich, wobei es nicht um Repräsentativität geht, sondern wohl eher darum, die wichtigsten Milieus innerhalb der Wahlberechtigten abzudecken.

Ein weiterer immer wieder betonter Vorteil von Wahlbörsen, nämlich dass die Parteipräferenzen der HändlerInnen nicht der Verteilung in der Bevölkerung entsprechen müssen, wird bisweilen auch so ausgelegt, dass die politischen Orientierungen der Trader unerheblich sind. Begleitende Untersuchungen zu Aktienmärkten haben jedoch ergeben, dass die Partei- bzw. Kandidatenpräferenzen durchaus eine Rolle für das Handelsverhalten spielen – zumindest außerhalb der strikt gewinnorientierten Gruppe der Marginal Traders (Forsythe et al. 1992; Huber 2001, 77). Dies würde bedeuten, dass die Parteipräferenzen in der Traderpopulation grundsätzlich gleich verteilt sein sollten – also gerade nicht wie in der Bevölkerung –, damit jede Aktie sozusagen den gleichen Rückenwind durch Parteipräferenzen erhält.

Was hier jedoch ins Spiel kommt, ist das Wahlbörsen inhärente Phänomen überhöhter Kurse von kleineren Parteien (Huber 2001, 90ff.); die Erklärung dafür liegt darin, dass geringer bewertete Aktien bei gleichem Kapital mehr Spielraum für Transaktionen und potenzielle Gewinne bieten. Dies spräche für eine stärkere Repräsentation der AnhängerInnen größerer Parteien, um den „Run“ auch von rational handelnden Marginal Traders auf die kleineren Werte zu kompensieren. In jedem Fall erscheint es jedoch kontraproduktiv, wenn sich, wie in Österreich meistens der Fall, die TraderInnen mehrheitlich aus den AnhängerInnen kleinerer Parteien (Grüne und LIF) zusammensetzen.

Welche weitere Schlussfolgerungen ergeben sich aus den bisherigen Erfahrungen mit Wahlbörsen? Nachdem die Prognosegenauigkeit mit steigender Parteienanzahl, also der Anzahl unterschiedlicher Aktien, abnehmen dürfte (Berg et al. 1996), könnten Wahlbörsen auch für leicht geänderte Fragestellungen eingesetzt werden. In Mehrparteiensystemen wie dem österreichischen bzw. dem deutschen bietet sich etwa die Frage an, ob das Wahlergebnis eine linke (rot-grüne) oder eine rechte (schwarz-blaue bzw. schwarz-gelbe) Mehrheit bringen wird. Für sol-

che bipolaren Fragestellungen eignen sich sogenannte „Winner-Takes-It-All“ (WTA) Märkte, die in den USA bereits erprobt wurden.

Schließlich ist noch zu erwähnen, dass nicht nur die Schlusskurse bei Wahlbörsen interessant sind; während des Wahlkampfes ist dieses Instrument möglicherweise in der Lage, gleich schnell oder sogar schneller als Umfragen gewisse Trendwenden zu erahnen, z.B. dass eine Partei eine andere in der Wählergunst überholt (Huber 2001, 165). Ob sie hier tatsächlich einen Vorsprung gegenüber der Demoskopie haben, wäre allerdings noch nachzuweisen.

Was noch nicht systematisch analysiert wurde, ist die Sensibilität der Wahlbörsen für außerordentliche Ereignisse im Verlauf der Kampagne. Es wäre gleichermaßen methodisch herausfordernd wie inhaltlich lohnend, anhand vergangener Wahlkämpfe zu untersuchen, ob markante Ereignisse wie die TV-Konfrontationen der KandidatInnen oder die Inhalte der Medienberichterstattung signifikante Effekte auf die Parteien-Kurse aufweisen.

Zusammenfassend kann die Wahlbörse aus der Sicht der empirischen Wahlforschung als ein hochinteressanter Methodenimport bezeichnet werden, der aber noch schuldig geblieben ist, auf reproduzierbare Weise einen Mehrwert gegenüber traditionellen Methoden wie Umfragen zu liefern. Sorgfältige Designs bei zukünftigen Wahlen werden weisen, ob die nach wie vor bestehenden Hoffnungen, dass dieser Mehrwert doch vorhanden ist, berechtigt sind.

2. Welche Ströme liegen dahinter?

Bei Wahlergebnissen ergibt sich die gleichermaßen spannende wie schwierig zu beantwortende Frage, auf welche Wählerbewegungen die relativen Stimmengewinne und -verluste der Parteien tatsächlich zurückzuführen sind. In diesem Abschnitt beschreiben wir „ökologische“ Modelle, die mit Hilfe multipler Regressionsverfahren auf die Wanderungsströme schließen, die hinter den Veränderungen zwischen zwei Wahlergebnissen liegen. Als Datenmaterial dienen nicht Angaben über individuelles Wahlverhalten, wie sie aus Umfragen

erhältlich sind, sondern die amtlichen Ergebnisse von Stimmbezirken (z.B. Wahlsprengel, Gemeinden oder Bezirke). Diese Daten nennt man „Aggregatdaten“, weil dabei die Einzelentscheidungen der WählerInnen eines Stimmbezirkes bereits zusammengezählt – aggregiert – wurden. Der Begriff „ökologisch“ hat sich eingebürgert, weil hier ein Phänomen durch sein Umfeld (seine Umwelt) erklärt wird.

Die Idee hinter dem Verfahren ist, vereinfacht ausgedrückt: Wenn eine Partei bei einer Wahl genau in jenen Gemeinden stark ist, wo eine andere Partei bei der vorhergehenden Wahl stark war, interpretieren wir das als Wählerbewegung, sprich dass viele WählerInnen zwischen diesen Parteien gewechselt haben.

Das zur Berechnung der entsprechenden Parameter eingesetzte Verfahren heißt multiple Regression: „Regression“, weil wir die Parteienergebnisse der neuen Wahl auf die Parteienergebnisse der alten Wahl regredieren (zurückführen), und „multiple“, weil wir das Ergebnis einer Partei bei der neuen Wahl gleichzeitig mit allen Parteienergebnissen der alten Wahl in Beziehung setzen. Dabei kommen entweder Ordinary-Least-Square-(OLS-) oder Maximum-Likelihood-Schätzmethoden zur Anwendung (Hofinger/Ogris 1996; King 1997). Allerdings verletzen die speziellen Eigenschaften von aggregierten Wahlraten – nämlich spatiale Autokorrelation, Multikollinearität und Heteroskedastizität – so ziemlich alle Voraussetzungen, die für die Anwendung dieser Schätzverfahren gelten (Ogris 1993, 113).

Spatiale (räumliche) *Autokorrelation* bezeichnet das Phänomen, dass es in benachbarten Gemeinden zu ähnlichen Unter- oder Überschätzungen der Ergebnisse kommen kann. Die Ergebnisse der einzelnen Gemeinden sind oft nicht voneinander unabhängig (z.B. indem Zentren auf die Umlandgemeinden „ausstrahlen“). Dies ist der Fall, wenn regionale Subtrends auftreten.

Multikollinearität tritt auf, wenn erklärende Variablen stark miteinander korrelieren. Im Fall der Wählerstromanalyse ist dies offensichtlich, nachdem viele Parteien in ähnlichen Regionen stark sind; so erreichen zum Beispiel in Österreich in urbanen Gebieten FPÖ-, Grün-, LIF-

und NichtwählerInnen höhere Stimmenanteile als in ländlichen Regionen.

Unter *Heteroskedastizität* versteht man je nach Wertebereich der unabhängigen Variablen divergierende Varianzen in den Residuen, also in den Abweichungen zwischen den geschätzten und den tatsächlichen Werten. Dazu kann es zum Beispiel kommen, wenn das Wahlverhalten in großen Gemeinden mobiler ist als in kleineren Gemeinden. So findet ein großer Teil des Wählerwechsels bei Wahlen mit der „Gruppe“ der NichtwählerInnen statt (Hofinger/Jenny/Ogris 2000). Vor allem bei Wahlen und in Regionen mit niedriger Wahlbeteiligung kann die Mobilisierung von NichtwählerInnen und die Demobilisierung von WählerInnen zu NichtwählerInnen die häufigste Form des Wechselwählens sein (Hofinger/Ogris/Breitenfelder 2000). Dies ist ein Phänomen mit besonderer Bedeutung in größeren Gemeinden und daher auch oft eine Quelle für Heteroskedastizität in den Residuen.

Regionale Subtrends und Multikollinearität können dazu führen, dass das Ausmaß der Veränderungen falsch eingeschätzt wird. Heteroskedastizität und spatiale Autokorrelation verzerren die Schätzung von Schwankungsbreiten. Das kann unter anderem den unangenehmen Effekt haben, dass ein kleines Konfidenzintervall fälschlicherweise eine hohe Sicherheit der Ergebnisse suggeriert.

Das in der praktischen Anwendung der linearen Regression (OLS) heikelste Problem besteht jedoch darin, dass die resultierenden Beta-Koeffizienten oft nicht als Übergangswahrscheinlichkeiten interpretiert werden können. Wahrscheinlichkeiten machen nur in der Bandbreite zwischen 0 und 1 Sinn. Bei negativen Koeffizienten haben wir das Problem, dass keine negative Anzahl von WählerInnen der Partei x zur Partei y gegangen sein kann; bei Koeffizienten größer 1 (die meistens in der Diagonale auftreten, also bei den Behalteraten der Parteien) können wir nicht folgern, dass mehr als 100% der WählerInnen der Partei x sie ein zweites Mal gewählt haben.

Um dieses Problem zu vermeiden, gibt es zwei Strategien: entweder die Koeffizienten zu korrigieren oder ein anderes Regressionsverfahren

wie einen Maximum-Likelihood-Schätzer zu verwenden, das nur Wahrscheinlichkeiten zwischen 0 und 1 zulässt (Neuwirth 1984).

Die Autoren verwenden bewusst ein Verfahren, das „böse“ Koeffizienten kleiner 0 oder größer 1 zulässt: Wenn uninterpretierbare Koeffizienten vorkommen, sollen sie unserer Ansicht nach nicht von vorneherein durch das Verfahren unterdrückt werden, sondern Grundlage für eine Neuspezifizierung des Modells bzw. für eine Neu-Gruppierung der Aggregat-Einheiten bilden. Unterteilt man die Gemeinden dann mehrmals nach regionalen, politischen oder geographischen Kriterien und rechnet die Analyse für jede dieser Einteilungen, ist im Einzelfall abschätzbar, in welchem Ausmaß die Bedingungen für das Schätzverfahren verletzt werden.

Wir schlagen zur Einschätzung der Modellgüte den von uns so genannten Non-Interpretable-Coefficients-Index (kurz: NIC-Index) vor, der die durchschnittliche Höhe der nicht interpretierbaren Teile der Koeffizienten misst.

$$NIC = \frac{\sum_i^n (p_i - 1) - \sum_i^n q_i}{n}$$

p ... Regressionskoeffizienten größer 1

q ... Regressionskoeffizienten kleiner 0

n ... Anzahl der geschätzten Koeffizienten (z.B. 326 im Falle von 6 x 6 Parteien inkl. NichtwählerInnen, getrennt für 9 Bundesländer)

Je geringer der NIC-Index, desto besser ist es uns durch die Einteilung der Gemeinden gelungen, nicht-interpretierbare Koeffizienten zu vermeiden. Wir interpretieren dies so, dass wir dadurch Fehlschätzungen durch Multikollinearität und regionale Subtrends minimieren.

Erst wenn auf diesem Weg die Zahl der nicht-interpretierbaren Koeffizienten – und damit hoffentlich die Schätzfehler – minimiert wurden, begrenzen wir die übrigen nicht-interpretierbaren Übergangswahrscheinlichkeiten mathematisch auf Werte zwischen 0 und 1.

Die *a priori*-Unterdrückung von Koeffizienten kleiner 0 oder größer 1 durch das Verfahren erspart einem zwar die Notwendigkeit einer

nachfolgenden Korrektur, verdeckt aber wertvolle Informationen. Denn die Qualität einer Wählerstromanalyse aus Aggregatdaten hängt unserer Erfahrung nach – neben der Anzahl der Untersuchungsobjekte (Gemeinden, Sprengel etc.) – vor allem von der Sorgfalt ab, mit der diese zu homogenen Gruppen zusammengefasst werden, damit regionale oder auch andere Subtrends getrennt analysiert werden können.

Cho (1998) belegt den zweiten Punkt mit Daten, wo Zusammenhänge aus Individualdaten bekannt sind. Wenn diese Individualdaten auf Bezirksebene aggregiert und mittels ökologischer Regression Koeffizienten berechnet werden, liegen die Ergebnisse oft weit neben den „wahren“ Werten, die in diesem Fall aus den Übergängen in den Individualdaten bestehen. Schlimmer noch: Die Signifikanzstatistiken führen bisweilen in die Irre und lassen weit vom wahren Ergebnis liegende Koeffizienten als gesicherte Ergebnisse erscheinen. Cho weist nach, dass für die Berechnung von realitätskonformen Koeffizienten weniger die Art des Schätzverfahrens entscheidend ist (sie vergleicht Ordinary Least Square mit Kings Verfahren), sondern vielmehr die Vermeidung eines *aggregate bias* – was nichts anderes bedeutet, als dass die Gemeinden sinnvoll gruppiert werden müssen.

Das nach wie vor ungelöste Problem dabei ist: Es gibt keine wissenschaftliche Methode, um festzustellen, welche Gruppeneinteilungen der Aggregatergebnisse tatsächlich zu Ergebnissen führen, die näher an der Realität liegen. King (1997) schlägt visuelle Interpretation der Daten und die Anwendung von durch Vorwissen begründete Theorien vor, was Cho zu Recht als unbefriedigend kritisiert. Sie liefert auch den Nachweis, dass selbst theoretisch gut begründete Gruppierungen der Aggregateinheiten zu ökologischen Fehlschlüssen führen können.

Wir wollen an einem Beispiel mit Ergebnissen der österreichischen Nationalratswahl 1999 demonstrieren, welch bedeutenden Effekt die Einteilung in homogene Subgruppen auf die Interpretierbarkeit der Koeffizienten aufweist.

Wir nehmen das Ergebnis der SPÖ bei der Nationalratswahl 1999 und berechnen die Herkunft der Stimmen von 1995. Datenbasis sind

die 2358 Gemeinden außerhalb Wiens und die 23 Wiener Gemeindebezirke. Die Koeffizienten wurden mittels linearer multipler Regression ermittelt, abhängige Variable war der Anteil der SPÖ-Stimmen an den Wahlberechtigten 1999, unabhängige Variablen waren die Anteile der Parteien an den Wahlberechtigten 1995. Ungültige Stimmen und NichtwählerInnen von 1995 wurden zu einer Partei der NichtwählerInnen zusammengefasst, ebenso wurden KPÖ, DU, Nein und SWG zu einer Liste „Sonstige“ summiert. Bei der Berechnung mittels Weighted Ordinary Least Square (WOLS) bildet die Anzahl der Wahlberechtigten von 1999 die Gewichtsvariable.

Wenn wir alle Gemeinden in einen Topf werfen, ist das Ergebnis so gut wie uninterpretierbar (Tabelle 2): 5 von 7 Koeffizienten (also 71%) sind negativ – im Sinne von Wählerströmen wären nur zwei Koeffizienten interpretierbar, nämlich mit der ÖVP (5% der ÖVP-WählerInnen von 1995 haben 1999 die SPÖ gewählt) und mit der SPÖ, also die Behalterate (88% der SPÖ-WählerInnen von 1995 sind 1999 bei ihrer Partei geblieben). Der NIC-Index ergibt 0,09, d.h.

die Koeffizienten schießen im Durchschnitt um 0,09 über den interpretierbaren Bereich hinaus.

Eine regionale Einteilung der Gemeinden nach Bundesländern verschafft nur geringe Linderung des Problems (Tabelle 3): Noch immer ist die Hälfte aller Koeffizienten negativ, wobei diese Werte bis -0,87 reichen. Der NIC-Index ist mit 0,10 sogar etwas höher.

Die Einteilung der Gemeinden nach der relativen Stärke der SPÖ hat dagegen einen bedeutenderen Effekt (Tabelle 4): Nur mehr 8 von 35, also 23% der Koeffizienten sind negativ, das Minimum erreicht bloß -0,38, und der NIC-Index ist auf 0,04 gefallen.

Die Einteilung nach der Größe der abhängigen Variable (der SPÖ-Stärke) hat also dazu geführt, dass das Problem der nicht als Übergangswahrscheinlichkeiten interpretierbaren Koeffizienten zumindest mehr als halbiert werden konnte. Dies sollte das Vertrauen in die mit dieser Gemeinde-Einteilung gewonnenen Schätzergebnisse wesentlich erhöhen. Zu einem höheren Vertrauen in die Ergebnisse der Schätzungen, die durch Gruppierung der Gemeinden gewonnen wurden, trägt auch der Umstand bei,

Tabelle 2: Multiple Regression des SPÖ-Ergebnisses bei der Nationalratswahl 1999 auf die Ergebnisse der Parteien bei der Nationalratswahl 1995

	N	SPÖ	ÖVP	FPÖ	Grüne	LIF	Sonst.	Nichtw.
alle Gemeinden	2381	0,88	0,05	-0,02	-0,36	-0,03	-0,18	-0,07

Werte sind unstandardisierte Regressions-Koeffizienten

Tabelle 3: Multiple Regression des SPÖ-Ergebnisses bei der Nationalratswahl 1999 auf die Ergebnisse der Parteien bei der Nationalratswahl 1995, nach Bundesland

	n	SPÖ	ÖVP	FPÖ	Grüne	LIF	Sonst.	Nichtw.
Burgenland	171	0,92	0,02	-0,02	-0,04	-0,34	-0,59	0,09
Kärnten	132	0,87	0,15	-0,02	-0,66	0,02	0,12	-0,10
NÖ	573	0,87	0,02	-0,01	-0,07	-0,01	0,12	-0,05
OÖ	445	0,86	0,05	0,02	-0,04	-0,30	0,48	-0,13
Salzburg	119	0,82	0,04	-0,01	-0,22	-0,22	0,04	0,06
Steiermark	543	0,87	0,02	0,03	0,06	-0,44	-0,01	-0,12
Tirol	279	0,80	0,02	0,06	-0,19	-0,10	-0,20	-0,06
Vorarlberg	96	0,78	0,01	-0,01	-0,09	-0,22	0,61	-0,02
Wien	23	0,73	-0,01	0,30	-0,79	0,39	-0,87	-0,04

Werte sind unstandardisierte Regressions-Koeffizienten

Tabelle 4: Multiple Regression des SPÖ-Ergebnisses bei der Nationalratswahl 1999 auf die Ergebnisse der Parteien bei der Nationalratswahl 1995, nach Anteil der SPÖ an den Wahlberechtigten 1999

	n	SPÖ	ÖVP	FPÖ	Grüne	LIF	Sonst.	Nichtw.
SPÖ < 15,218%	476	0,65	0,04	0,04	-0,20	0,08	0,22	-0,02
SPÖ < 19,778%	476	0,42	0,12	0,12	0,02	0,10	0,09	0,09
SPÖ < 24,783%	477	0,57	0,11	0,10	-0,04	0,05	0,37	0,07
SPÖ < 31,830%	476	0,64	0,14	0,12	-0,26	0,12	0,33	0,06
SPÖ > 31,830%	476	0,85	0,17	-0,01	-0,38	0,02	-0,38	-0,09

Werte sind unstandardisierte Regressions-Koeffizienten

dass die Varianz der Residuen dadurch kleiner geworden ist: Betrug die Standardabweichung des Schätzfehlers in der über alle Gemeinden gerechneten Gleichung noch 0,023, sinkt sie in den getrennt nach Bundesländern bzw. SPÖ-Stärke gerechneten Varianten deutlich (auf 0,017 bzw. 0,016).

Bleibt noch die Frage, was mit den verbliebenen negativen Koeffizienten zu tun ist, um zu Aussagen über Wählerbewegungen zu kommen. Ein möglicher Weg ist die sogenannte Ridge-Regression (Füle 1994), bei der die Behalteraten, also die Korrelation einer Partei bei der neuen Wahl mit ihrem Ergebnis bei der alten Wahl, so lange gesenkt und auf die negativen Koeffizienten verteilt werden, bis diese Werte über Null erreichen. Uns erscheint es allerdings sinnvoller, die Behalterate nur dann zu senken, wenn sie größer als 1 geschätzt wird. Unabhängig davon sollten die negativen Koeffizienten als Werte knapp bei Null, also sehr geringe Übergänge, interpretiert werden.

Zur Ehrenrettung der Koeffizienten > 1 oder < 0 sei noch gesagt, dass sie – auch wenn sie Stoff für die Alpträume des Wählerstrom-Analytikers abgeben – wahlsoziologisch durchaus sinnvoll interpretierbar sind. Betrachten wir dies etwa für den hypothetischen Fall, in dem der Koeffizient zwischen SPÖ 1995 und SPÖ 1994 größer als 1, also z.B. $p_{ij} = 1,12$. Im Sinne des statistischen Modells von Übergangswahrscheinlichkeiten würde man das übersetzen mit: „Von 100 SPÖ-Stimmen hat die SPÖ 112 bekommen“ – in dieser Form eine sinnlose Aussage. Im Sinne eines wahlsoziologischen Modells der Mobilisierung aber könnte man dies so formulieren: „Wo die SPÖ 100 Stimmen hat-

te, hat sie zusätzliche 12 Stimmen für sich mobilisiert.“ Die Stärke der SPÖ bei der aktuellen Wahl wird so auf die Stärke der SPÖ bei der letzten Wahl zurückgeführt. Wahlsoziologisch ist der Regressionskoeffizienten über 1 interpretierbar, als Übergangswahrscheinlichkeit jedoch nicht.

Solche Wählerbewegungen sind etwa bei einem „Swing back“ durch Remobilisierung früherer Verluste feststellbar. Wenn eine Partei in ihren Hochburgen nach starken Verlusten bei der letzten Wahl bei der aktuellen Wahl durch starke Mobilisierung ehemalige WählerInnen wieder für sich gewinnt (wie etwa die SPÖ 1995 durch einen an ihren Kernschichten orientierten Mobilisierungswahlkampf), machen solche Koeffizienten auch inhaltlich Sinn.

Auch Aussagen wie: „Pro 100 Stimmen der Partei X bei der letzten Wahl hat die Partei Y jetzt um 10 Stimmen weniger“ sind wahlsoziologisch interpretierbar. Am Beispiel in Tabelle 2: „Wo bei der letzten Wahl besonders viele Grün-WählerInnen waren, hat die SPÖ Stimmen abgegeben.“

Wenn es also durch Zusammentreffen von SPÖ und Grün-WählerInnen zu Unzufriedenheit innerhalb der Wählerschaft der SPÖ kommt, weil von dieser etwa ein massiveres Auftreten gegen Ausländerfeindlichkeit gewünscht wird, kann dies auch zu Stimmenverlusten führen. Es kommt dann zu einer Demobilisierung von SPÖ-WählerInnen, die mit der Stärke der Grünen zusammenhängt. Wenn nun dieser Demobilisierungseffekt das Wechselwählen von Grün zu Rot überlagert, dann ist ein negativer Koeffizient die Folge (wie z.B. in Tabellen 2 bis 4).

Wählerstromanalyse oder Umfragen?

Bei sorgfältiger Anwendung bringt die ökologische Wählerstromanalyse plausible Ergebnisse (Hofinger/Ogris 1996; 2000). Für die Nationalratswahl 1999 zeigt das Verfahren die starken direkten Verluste der Regierungsparteien SPÖ und ÖVP an die Oppositionsparteien, vor allem an die FPÖ. Mangelnde Mobilisierung früherer WählerInnen war eine der Hauptursachen für die Niederlage der SPÖ. Die ÖVP hat ihr relativ gutes Ergebnis vor allem durch eine hohe Behalterate und geringe Verluste durch Wahlenthaltung erzielt. Diese zentralen Aussagen werden auch durch Umfragen bestätigt.

Das Verfahren der ökologischen Wählerstromanalyse hat gegenüber Umfragen mehrere Vor- und Nachteile (vgl. Hoschka/Schunck 1975 und 1982; Küchler 1983; Neuwirth 1984; Seller 1992; Ogris 1993; Hofinger/Ogris 1996; Hofinger/Ogris 1999).

Die Vorteile der ökologischen Regression sind:

- Es gibt keine Probleme mit der Datenerhebung, wie sie bei Umfragen vorkommen, also keine Stichprobenfehler, keine Antwortverweigerung, keine falsche Angaben etc.
- Durch Analyse der amtlichen Ergebnisse bekommt man auch Ergebnisse für regionale Subeinheiten (z. B. den Wiener Gemeindebezirk Favoriten), die von Umfragen üblicherweise nicht erfasst werden.
- Das Verfahren ist kostengünstiger: Die Kosten betragen bei Zugriff auf das amtliche Wahlergebnis nur einen Bruchteil der Aufwendungen einer Umfrage.
- Das Verfahren ist auch auf längst vergangene Wahlen anwendbar (Falter 1991).
- Die Wählerstromanalyse kann im Gegensatz zu Exit Polls die NichtwählerInnen berücksichtigen.

Die ökologische Wählerstromanalyse hat aber auch mehrere Nachteile:

- Sie operiert nur mit Aggregatdaten – daraus abgeleitete Aussagen über Einzelpersonen bergen daher das Risiko des „ökologischen Fehlschlusses“.

- Das Verfahren ermöglicht keine Aufschlüsselung des Wahlverhaltens nach Alter, Geschlecht, Beruf etc., sondern ausschließlich nach Typen von Gemeinden (z. B. Industriegemeinden, Städte, FPÖ-Hochburgen, Fremdenverkehrsgemeinden etc.) oder Regionen.
- Das Verfahren kann kaum auf die demographische Veränderung des Wahlkörpers Rücksicht nehmen: Die den Schätzungen zugrundeliegenden Gemeindeergebnisse sind ja nicht nur Folge von Wählerwanderungen, sondern auch von Zugängen durch Zuzug und ErstwählerInnen sowie von Abgängen durch Wegzug und Verstorbene; in jeder ökologischen Wählerstromanalyse sind diese Veränderungen implizit enthalten, auch wenn sie als Veränderung nur im bei beiden Wahlen stimmberechtigten Elektorat erscheinen. Eine Lösung hierfür wäre der relativ aufwändige Ansatz, für jede Gemeinde demographische Zu- und Abgänge zu schätzen.
- Die ökologische Wählerstromanalyse liefert für kleinere Parteien relativ unsichere Ergebnisse.
- Die Daten verletzen mehrfach die geltenden statistischen Rahmenbedingungen für die angewandten Verfahren und können zu Fehlschätzungen und falschen Schwankungsbreiten führen.

Es ist zu hoffen, dass die hier beschriebenen methodischen Probleme mit der Zeit durch Weiterentwicklungen des Verfahrens besser in den Griff zu bekommen sind. Da verbesserte Methoden jederzeit auf alte Wahlkosten anwendbar sind, könnten Berechnungen auch neues Licht auf längst vergangene Wahlen werfen.

Um Schwankungsbreiten effektiv zu schätzen, brauchen wir mehr Verständnis über die Auswirkungen der Daten-Probleme. Dazu sollten zusätzliche Schätzgrößen entwickelt werden, die Aussagen darüber machen, bei welcher Einteilung der Gemeinden die geringsten Ausmaße an Heteroskedastizität, Autokorrelation und Multikollinearität vorliegen.

Die Hauptaufgabe bei der Weiterentwicklung des Verfahrens liegt unserer Ansicht nach jedoch weniger bei der Optimierung der Schätzverfahren als bei der wissenschaftlich gesicher-

ten Identifizierung von sinnvollen Einteilungen der Aggregateinheiten, um uninterpretierbare Koeffizienten zu vermeiden und regionale oder andere Subtrends möglichst exakt und effektiv zu lokalisieren.

3. Hochrechnung – Showdown am Wahlabend

Dass die Wählerstromanalyse als Methode der Hochrechnung verwendet wird, ist in der Literatur nur wenig dokumentiert – und ihre Anwendung nur für Österreich bekannt.

In Österreich gibt es an Wahltagen keinen allgemein gültigen Wahlschluss. Bereits am späten Vormittag wird in den Wahllokalen einiger kleiner Gemeinden ausgezählt – um 15.00 Uhr haben meist in mehr als der Hälfte aller Gemeinden die Wahllokale schon geschlossen. Die Wahlkommissionen beginnen unverzüglich nach dem Wahlschluss die Ergebnisse zu ermitteln – und so steht den HochrechnerInnen frühzeitig ein Datensatz zur Verfügung, der die vorläufigen Ergebnisse von mehreren hundert Gemeinden enthält.

Die Autoren arbeiten selbst seit Mitte der 80er Jahre mit dem Verfahren der Wählerstromanalyse. Zu Beginn der 90er Jahre kam die Idee auf, dass man dieses Verfahren auch für die Hochrechnung verwenden könnte. Im Frühjahr 1994 haben die Autoren dieses Verfahren erstmals im Auftrag des ORF für die Hochrechnung von Landtagswahlen eingesetzt. Seither wurden damit auch bei Nationalratswahlen im In- und Ausland, bei Bundespräsidentenwahlen, bei Wahlen zum Europäischen Parlament, bei der EU-Volksabstimmung sowie bei zahlreichen weiteren Landtagswahlen Erfahrungen gesammelt.

Ein dokumentierter Vergleich verschiedener Prognoseverfahren hat die Vorteile der Wählerstromanalyse bei diesen Datenstrukturen gezeigt (Neuwirth 1994). Es wird dabei mit der Annahme gearbeitet, dass die Übergänge von einer Partei zu den anderen in den nicht ausgezählten Gemeinden gleich sind wie in den bereits ausgezählten. Auch nach unserer Erfahrung ist dieses Verfahren weit präziser als Ansätze wie eine

einfache Regression oder eine simple Differenzschätzung, vor allem wenn genügend ausgezählte Gemeinden (ca. 150–200) vorhanden sind: Der durchschnittliche Schätzfehler liegt ab 200 ausgezählten Gemeinden selten über 1% pro Partei, und nur vereinzelt liegen Vorhersagen für einzelne Parteien um mehr als 1,5% neben dem Endergebnis. Nach Wahlschluss, wenn die erste Hochrechnung „on air“ geht, liegt der durchschnittliche Vorhersagefehler pro Partei meist unter 0,5% und der maximale Fehler selten über 1%.

Der Zeitdruck, unter dem hierbei Entscheidungen gefällt werden müssen, erinnert eher an Hollywood-Streifen wie „Mission Impossible“ als an den Arbeitsalltag eines Forschers. Manchmal bleibt nur eine Minute, um die hereinkommenden Daten zu verarbeiten und die Hochrechnungswerte an die Grafik weiter zu geben. Unter diesen Umständen haben wir über die Jahre versucht, so viel wie möglich zu automatisieren – am Wahlabend selbst zu entscheiden, welche Variante die plausiblere ist, birgt die Gefahr, sich durch vor der Wahl veröffentlichte Umfragen oder auch eigene Parteipräferenzen in die falsche Richtung locken zu lassen.

Die SORA-Software¹² rechnet am Wahlabend bis zu 18 verschiedene Varianten von Wählerstromanalysen. Diese Varianten werden im Vorhinein festgelegt und unterscheiden sich durch die Einteilung der Gemeinden in verschiedene Gruppen, differenziert zum Beispiel nach Anzahl der Wahlberechtigten, regionalen Kriterien oder nach der Stärke einzelner Parteien. (Theoretisch wäre es möglich, erst am Wahlabend Gemeinde-Unterteilungen „on the fly“ zu errechnen, was wir aber im Sinne einer Risiko-Minimierung unterlassen.)

Auf diese Weise erhalten wir bis zu 18 verschiedene Ergebnisse, die sich zwar selten in ihrer substantiellen Aussage unterscheiden, aber naturgemäß variieren. Konkret betrug zum Beispiel bei der EU-Wahl 1996 zu Zeitpunkten, an denen noch weniger als ein Fünftel der Wahlberechtigten erfasst waren, die Standardabweichung der Vorhersagen für die drei größeren Parteien 0,9. Um 17.30 Uhr, als 58% der Wahlberechtigten erfasst waren, die Ergebnisse aus

den meisten Städten aber noch ausstanden, sank die Standardabweichung der Vorhersagen auf deutlich geringere Werte, wie etwa 0,4 (SPÖ), 0,3 (ÖVP), 0,2 (FPÖ) und 0,1 (LIF).

Wie sehr die Vorhersage für eine Partei variiert, sagt pikanterweise nichts über die Genauigkeit der Prognose aus. Im oben genannten Beispiel war die Korrelation zwischen der Standardabweichung und dem Fehler der Prognose sogar negativ (-0,11) – so war die SPÖ zwar die Partei mit der höchsten Varianz der Vorhersagen, aber gleichzeitig dem geringsten mittleren Schätzfehler (0,16%).

Die beste Hochrechnung liefert unserer Erfahrung nach im Allgemeinen ein gewichtetes Mittel aus allen gerechneten Varianten. Als Gewicht, mit der eine Variante in das Mittel eingeht, verwenden wir den Kehrwert des weiter oben beschriebenen NIC-Index. Das bedeutet, dass jene Schätzungen, die wenig uninterpretierbare Koeffizienten aufweisen, ein größeres Gewicht erhalten und umgekehrt.

Eine Garantie für eine exakte Hochrechnung kann es dennoch nicht geben. Wenn die WählerInnen in den großen Gemeinden und insbesondere in Wien, dessen Ergebnisse bei Wahlschluss stets ausstehen, nicht im Trend der kleineren Gemeinden liegen, kann das beste Verfahren nicht erraten, wie dieser urbane Gegentrend aussieht. Aber dieser Fall von stark gegenläufigen Trends ist selten – und wenn diese doch zu einer weniger präzisen Hochrechnung führen sollten, liegt Trost für alle Beteiligten darin, dass der Wahlabend länger spannend bleibt und zumindest das Fernsehpublikum auf seine Rechnung kommt.

4. Ausblick

Die Methodenentwicklung für Wählerstromanalysen, Hochrechnungen und Wahlbörsen ist noch lange nicht an ihrem Endpunkt angekommen. Eine Reihe von neuen Forschungsfragen ist zu stellen, an einigen alten Fragen ist weiter zu arbeiten.

Im Bereich der Signifikanzstatistik für die Wählerstromanalysen könnte „Bootstrapping“ (eine iterative Methode, die Parameter aus

immer wieder neu gezogenen Teilstichproben der Daten schätzt, vgl. u.a. Beran/Ducharme 1991) ein Weg sein, um Schwankungsbreiten besser zu schätzen. Dadurch könnten möglicherweise jene Probleme der Signifikanzschätzung, die aus spatialer Autokorrelation und Multikollinearität herrühren, verringert werden.

Den möglichen Bias in den Schätzkoeffizienten könnte man eventuell dadurch verringern, dass die Ergebnisse der Wählerstromanalyse bereits auf der untersten Aggregatebene (z.B. Gemeinden) an die Randverteilungen angepasst werden. Derzeit produziert die Wählerstromanalyse ja Schätzungen, die für einzelne Gemeinden ein vom endgültigen Ergebnis abweichendes Resultat vorhersagt – diese Fehler werden dann meist erst auf der Ebene des Gesamtergebnisses korrigiert. Wenn bereits auf der niedrigsten Ebene die Parameterschätzung einem Randsummen-Fitting unterzogen wird, könnte das Verfahren zusätzliche Informationen verwerten.

Beide Wege zusammen wären wahrscheinlich geeignet, um die Unsicherheiten im hier diskutierten Hochrechnungsverfahren zu vermindern.

Die Erforschung der Wahlbörsen und ihre Validität als Prognoseinstrument steht überhaupt erst am Anfang. Eine der Fragen richtet sich an die AnalytikerInnen sozialer Netzwerke – inwieweit nehmen mehr oder weniger geschlossene Kreise am „Traden“ teil? Je mehr der Handel durch soziale, regionale oder sonstige Netzwerke beeinflusst wird, desto mehr stellt sich die Frage, ob dem System wichtige Stimmungen der Bevölkerung entgehen.

Die Frage, inwieweit Parteipräferenzen die Erwartungshaltungen hinsichtlich des Wahlausganges beeinflussen, ließe sich relativ leicht durch Umfragen klären. Wenn Erwartungshaltungen und Parteipräferenzen stark miteinander korrelieren, wird es wichtig, repräsentative Trader-Kollektive zusammenzustellen. Das würde aber einen großen organisatorischen und finanziellen Aufwand mit sich bringen, und der viel gepriesene Kostenvorteil der Wahlbörsen wäre wieder dahin. Jedenfalls zu empfehlen ist, wieder mit experimentellen Designs wie zu den „Urzeiten“ der Political Stock Markets zu arbeiten, um die Vorhersagekraft

dieser Methode abseits des großen Medienrummels in Ruhe ausloten zu können.

ANMERKUNGEN

- 1 Federführend ist die University of Iowa, vgl. <http://www.uiowa.edu>
- 2 Die kleineren Parteien werden meist zu einer Aktie „Sonstige“ zusammengefasst; in Österreich waren sechs Aktien (SPÖ, ÖVP, FPÖ, Grüne, LIF, Sonstige) der häufigste Fall.
- 3 Dies hat unter anderem den nicht unwichtigen Effekt, dass sich der Gesamtwert der Parteienaktien in einem funktionierenden Politik-Markt stets sehr nah dem Ausgabewert bewegt. Falls dem kurzfristig nicht so ist, können die HändlerInnen durch bloßes Ver- und Zurückkaufen an die und von der Bank Gewinne realisieren, bis die Kurse in Summe wieder den Ausgabewert erreichen. Dadurch kann z.B. bei einem Bündel-Ausgabewert von 1 Euro der in Cent gemessene Aktienkurs eleganterweise als der Prozentsatz interpretiert werden, den der Markt für diese Partei vorhersagt.
- 4 Eigentlich sind es Futures (Huber 2001, 18), aber in diesem Kontext ist immer von Stocks = Aktien die Rede.
- 5 In letzter Zeit setzen daher viele Institute Random Digital Dialing (RDD) ein, ein Verfahren, bei dem nicht gelistete Nummern durch zufällige Variation von im Verzeichnis enthaltenen Anschlüssen erreicht werden sollen.
- 6 Nichtwählen gehört überhaupt zu den demoskopisch am schwersten erfassbaren Phänomenen (Falter/Schumann 1994).
- 7 Dass die FPÖ in Rohdaten häufig unterschätzt wird, wird jedoch nach Ansicht der Autoren oft zu monokausal auf die mangelnde Deklarationsbereitschaft ihrer AnhängerInnen zurückgeführt. Die FPÖ wies stets die größte Unterstützung bei einer der am schwersten zu erreichenden soziodemographischen Gruppen, nämlich jungen Männern mit geringer formaler Bildung auf (vgl. Plasser/Ullrich 2000). Nach Erfahrungen der Autoren ist durch eine soziodemographische Gewichtung, die die Bildung auch als Gewichtungskriterium inkludiert, und eine anschließende Recall-Gewichtung der tatsächliche Anteil der FPÖ nicht schwieriger vorherzusagen als der der anderen Parteien – zumindest seit dem Eintritt der FPÖ in die Bundesregierung. Was unserer Ansicht nach nur mit Vorsicht angewandt werden sollte, sind Modelle, die die Nicht-Deklarierten bestimmten Parteien zuordnen, in der Annahme, dass diese Befragten das Wahlverhalten vor den InterviewerInnen verheimlichen wollen. Die Mehrzahl der Nicht-Deklarierten am Ende eines längeren Interviews über politische Fragen geht wahrscheinlich tatsächlich nicht wählen. Wenn der Anteil der Nicht-Deklarierten sehr deutlich über dem zu erwartenden Anteil an Nicht-WählerInnen liegt, kann dies vor al-

lem an sehr kurzen Interviews oder dem Einschalten der „Sonntagsfrage“ in eine Mehr-Themen-Umfrage liegen – der abrupte Schwenk zur Wahlfrage verleitet in diesen Fällen viele Befragte zur Antwortverweigerung.

- 8 Eine leicht programmierbare Computersimulation verdeutlicht anschaulich das Phänomen der Schwankungsbreiten: Aus einer großen Population, in der exakt 40% für eine Partei x stimmen, wurden zehn Mal zufällig 1000 Individuen gezogen. Der Anteil der Partei x erreichte in den zehn Stichproben folgende Prozent-Werte: 37,0 – 40,8 – 37,3 – 40,1 – 40,4 – 41,0 – 42,0 – 41,9 – 41,1 – 36,5. Selbst diese perfekten Ziehungen produzierten also einen durchschnittlichen Fehler von 1,7%.
- 9 Z.B. liefen bei der Wiener Gemeinderatswahl 2001 auf DiePresse.at, DerStandard.at und Kurier.at jeweils eigene Wahlbörsen.
- 10 Gerade die Unerfahrenheit der EuropäerInnen mit Aktienmärkten (in Österreich besitzen bloß 8% der Bevölkerung Aktien, vgl. profil 05/2002 vom 28.1.) führt dazu, dass sie sich bei Wahlbörsen offensichtlich häufig aus blankem Unwissen gegen ihre eigenen Interessen verhalten, während ja in den USA weit mehr als die Hälfte der Haushalte einen Teil ihres Vermögens in Aktien anlegt.
- 11 Außer über Umfragen, die aber zum Einen selten so genau differenzieren und zum Anderen ja nicht alleiniger Informationsinput von Wahlbörsen sein können, wenn dieselben zusätzliche Prognosegenauigkeit erreichen wollen.
- 12 SORA steht für Institute for Social Research and Analysis, das von den Autoren in Wien gegründete Forschungsinstitut.

LITERATURVERZEICHNIS

- Beran, Rudolf/Gilles R. Ducharme* (1991). *Asymptotic Theory for Bootstrap Methods in Statistics*, Montreal.
- Berg, Joyce/Robert Forsythe/Thomas Rietz* (1996). *What Makes Markets Predict Well? Evidence from the Iowa Election Markets*, in: *Albert Wulff/Werner Güth/Peter Hammerstein/Benny Moldovanu/Eric Van Damme* (Hg.): *Understanding Strategic Interaction. Essays in Honor of Reinhard Selten*, New York, 444–463.
- Cho, Wendy K. Tam* (1998). *Iff [!] the Assumption Fits...: A Comment on the King Ecological Inference Solution*, in: *Political Analysis*, 7(1), 143–163.
- Falter, Jürgen W.* (1991). *Hitlers Wähler*. München.
- Falter, Jürgen W./Siegfried Schumann* (1994). *Der Nichtwähler – das unbekannte Wesen*, in: *Max Kaase/Hans-Dieter Klingemann* (Hg.): *Wahlen und Wähler. Analysen aus Anlaß der Bundestagswahl 1990*, Opladen, 161–213.
- Forsythe, Robert/Forrest Nelson/George Neumann/Jack Wright* (1992). *Anatomy of an Experimental Political*

- Stock Market, in: *American Economic Review*, 82(5), 1142–1161.
- Fülle*, Erika (1994). Estimating Voter Transitions by Ecological Regression, in: *Electoral Studies* 13(4), 313–330.
- Hofinger*, Christoph/Günther *Ogris* (1996). Wählerwanderungen. Ein Vergleich fünf verschiedener Wählerstromanalysen anlässlich der Nationalratswahl 1995, in: *Fritz Plasser/Peter A. Ulram/Günther Ogris* (Hg.): *Wahlkampf und Wählerentscheidung. Analysen zur Nationalratswahl 1995* (Schriftenreihe des Zentrums für angewandte Politikforschung; 10), Wien, 315–341.
- Hofinger*, Christoph/Günther *Ogris* (1999). Wie treu sind blaue Wählerherzen? Eine statistische Analyse des Wahlverhaltens der FPÖ-Wähler bei den österreichischen Nationalratswahlen in den neunziger Jahren, in: *Fritz Plasser/Oscar W. Gabriel/Jürgen Falter/Peter A. Ulram* (Hg.): *Wahlen und politische Einstellungen in Deutschland und Österreich* (= Empirische und methodologische Beiträge zur Sozialwissenschaft; 17), Frankfurt/Bern, 129–144.
- Hofinger*, Christoph/Günther *Ogris* (2000). Die Analyse der Wählerströme bei der Nationalratswahl 1999, in: *SWS-Rundschau*, 40(2), 125–139.
- Hofinger*, Christoph/Günther *Ogris*/Ursula *Breitenfelder* (2000). Ein Jahr der Kontraste. Die Wählerströme bei den Nationalratswahlen am 3. Oktober, bei den Wahlen zum Europäischen Parlament und bei den Landtagswahlen in Kärnten, Salzburg und Tirol am 7. März, in: *Andreas Khol/Günther Ofner/Alfred Stirmemann* (Hg.): *Österreichisches Jahrbuch für Politik 1999*, Wien/München, 119–145.
- Hofinger*, Christoph/Marcelo *Jenny*/Günther *Ogris* (2000). Steter Tropfen höhlt den Stein. Die Wählerströme 1999 im Kontext der 80er und 90er Jahre, in: *Fritz Plasser/Peter A. Ulram/Franz Sommer* (Hg.): *Das österreichische Wahlverhalten* (Schriftenreihe des Zentrums für angewandte Politikforschung; 21), Wien, 117–140.
- Hoschka*, Peter/Hermann *Schunck* (1975). Schätzung von Wählerwanderungen: Puzzlespiel oder gesicherte Ergebnisse?, in: *Politische Vierteljahresschrift*, 16(4), 491–539.
- Hoschka*, Peter/Hermann *Schunck* (1982). Das Puzzlespiel der Wählerwanderungen: Noch immer ungelöst, in: *Zeitschrift für Parlamentsfragen*, 13(1), 113–115.
- Huber*, Jürgen (2001). Wahlbörsen. Preisbildung auf Politischen Märkten zur Vorhersage von Wahlergebnissen, Dissertation, Universität Innsbruck.
- King*, Gary (1997). A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data, Princeton.
- Küchler*, Manfred (1983). Die Schätzung von Wählerwanderungen. Neue Lösungsversuche, in: *Max Kaase/Hans-Dieter Klingemann* (Hg.): *Wahlen und politisches System. Analysen aus Anlass der Bundestagswahl 1980*, Opladen, 632–650.
- Neuwirth*, Erich (1984). Schätzung von Wahlübergangswahrscheinlichkeiten, in: *Manfred J. Holzer* (Hg.): *Wahlanalyse. Hypothese, Methode und Ergebnisse*, München, 197–211.
- Neuwirth*, Erich (1994). Prognoserechnung am Beispiel der Wahlhochrechnung, in: *Peter Mertens* (Hg.): *Prognoserechnung*, Würzburg/Wien.
- Ogris*, Günther (1993). Die Wählerstromanalyse ist etwas besser als ihre Kritik. Replik auf Daniel Sellaer ‚Die Wählerstromanalyse. Anspruch und Wirklichkeit‘ aus der SWS-Rundschau 3/1992, 417–428, in: *SWS-Rundschau*, 33(1), 109–114.
- Ortner*, Gerhard (1996). Experimentelle Aktienmärkte als Prognoseinstrument – Qualitätskriterien der Informationsverarbeitung in Börsen am Beispiel Political Stock Markets, Dissertation, Universität Wien.
- Plasser*, Fritz/Peter A. *Ulram* (2000). Rechtspopulistische Resonanzen: Die Wählerschaft der FPÖ, in: *Fritz Plasser/Peter A. Ulram/Franz Sommer* (Hg.): *Das österreichische Wahlverhalten* (Schriftenreihe des Zentrums für angewandte Politikforschung; 21), Wien, 225–241.
- Sellaer*, Daniel (1992). Die Wählerstromanalyse. Anspruch und Wirklichkeit, in: *SWS-Rundschau*, 32(3), 417–428.
- Smith*, Vernon (1982). Markets as Economizers of Information: Experimental Examination of the ‚Hayek Hypothesis‘, in: *Economic Inquiry* 20(2), 165–179.

AUTOREN

Christoph HOFINGER, geb. 1967, seit 1996 zusammen mit Günther Ogris geschäftsführender Gesellschafter von SORA (Institute for Social Research and Analysis); Forschungsschwerpunkte: Individual- und Aggregatdatenanalysen des österreichischen Parteiensystems sowie Arbeitsmarktforschung und Arbeitszufriedenheitsmessung. Neuere Publikationen: *Hofinger*, Christoph/Marcelo *Jenny*/Günther *Ogris* (2000). *Steter Tropfen höhlt den Stein. Die Wählerströme 1999 im Kontext der 80er und 90er Jahre*, in: *Fritz Plasser/Peter A. Ulram/Franz Sommer* (Hg.): *Das österreichische Wahlverhalten* (Schriftenreihe des Zentrums für angewandte Politikforschung; 21), Wien, 117–140; *Hofinger*, Christoph/Günther *Ogris* (1999). *Wie treu sind blaue Wählerherzen? Eine statistische Analyse des Wahlverhaltens der FPÖ-Wähler bei den österreichischen Nationalratswahlen in den neunziger Jahren*, in: *Fritz Plasser/Oscar W. Gabriel/Jürgen Falter/Peter A. Ulram* (Hg.): *Wahlen und politische Einstellungen in Deutschland und Österreich* (=Empirische und methodologische Beiträge zur Sozialwissenschaft; 17), 129–144.

Adresse: Institute for Social Research and Analysis, SORA Ogris & Hofinger OEG, Gumpendorfer Str. 5/7, A-1060 Wien
E-mail: ch@sora.at Web: www.sora.at

Günther OGRIS, geb. 1960 in Wien, Studium der sozialwissenschaftlichen Datenanalyse an der Universität Essex, GB. 1983 bis 1996 freier Mitarbeiter des Instituts für empirische Sozialforschung (IFES). 1996 Gründung des Institute for Social Research and Analysis (SORA) gemeinsam mit Christoph Hofinger. 2000 bis 2002 Wissenschaftlicher Leiter des Institutes für Jugendforschung. Seit 1986 Wählerstromanalysen und

seit 1994 Hochrechnungen – gemeinsam mit Christoph Hofinger – für den ORF.

Adresse: Institute for Social Research and Analysis,
SORA Ogris & Hofinger OEG, Gumpendorfer Str. 5/
7, A-1060 Wien
E-mail: go@sora.at Web: www.sora.at