



## Dortmunder Methodik der Wählerwanderungsanalyse

### Fragestellung

Wahlergebnisse werfen - neben der reinen Betrachtung von Gewinnen und Verlusten gegenüber der vorangegangenen Wahl - auch die Frage nach dem „Woher und Wohin“ der Wählerschaft auf. Stamm- und Wechselwählerschaft sollen beschrieben und quantifiziert werden. Bei der Wechselwählerschaft interessiert es, deren veränderte Parteienpräferenz sichtbar zu machen.

Das Ziel einer Wählerwanderungsanalyse ist damit zu ermitteln, wie viele Wähler\*innen von Partei A (frühere Wahl) zu Partei B (jetzige Wahl) „gewandert“ oder bei Partei A verblieben sind.

Aus den Wahlergebnissen liegen, sofern keine gebietsspezifischen Wahltagsbefragungen stattgefunden haben, aber nur Flächendaten (Aggregatdaten auf der Ebene der Wahl- bzw. Stimmbezirke) vor. Aus diesen ist nicht direkt ersichtlich, wie die Wähler\*innen tatsächlich „gewandert“ sind. In einem unwahrscheinlichen Extremfall könnten zum Beispiel alle Wähler\*innen der Partei A der früheren Wahl auch bei der aktuellen Wahl wieder für Partei A gestimmt haben. Oder umgekehrt: Kein\*e ehemalige\*r Wähler\*in der Partei A hat erneut für diese Partei votiert.

### Verwendetes Wählerwanderungsmodell

Zur Analyse der Wählerwanderung gibt es grundsätzlich zwei verschiedene Ansätze: Die Berechnung der Wanderung aus Umfragedaten (Individualdaten) sowie die Schätzung aus kleinräumigen Wahlergebnissen (Aggregatdaten). Da in Dortmund bei der Bundestagswahl 2021 keine repräsentative Wahltagsbefragung durchgeführt wurde, wird auf den zweiten Ansatz zurückgegriffen. Zur Schätzung des Wanderungsgeschehens wird dabei ein ökologisches Inferenzmodell herangezogen. Dessen Grundannahme ist es, von Daten einer möglichst kleinräumigen Fläche (hier die Wahl- bzw. Stimmbezirke) auf individuelle (personenbezogene) Parameter (hier die Parteienpräferenz einer aktuellen gegenüber einer früheren Wahl) zu schließen.

Das dabei angewandte Multinomial-Dirichlet Modell von Rosen et al. (2001) ist ein hierarchisches Bayesianisches Modell zur Schätzung der Wählerwanderung und besitzt nach Klima (2016, S. 68) die beste Schätzgüte unter einer Vielzahl vergleichbarer ökologischer Inferenzmodelle. Zur technischen Durchführung der Berechnung wird die von der deutschen Städtestatistik in Auftrag gegebene Software-Implementierung (eiPack) von Lau, Moore und Kellermann (2020) im Programmpaket R (R Core Team, 2020) eingesetzt.

### Datenaufbereitung: Zuordnung der Briefwahl zur räumlichen Basis, den Urnenwahllokalen

In der Regel wird mehreren Wahl- bzw. Stimmbezirken (im Folgenden „Urnenwahllokale“ genannt) ein Briefwahlbezirk zugeordnet. Zur Berechnung der Wählerwanderung mit einem ökologischen

Inferenzmodell ist ein „vollständiges“ Wahlergebnis, das die Briefwahl auf der Ebene der Urnenwahllokale beinhaltet, notwendig.

Die aktuellen Briefwahlergebnisse liegen jedoch nur als Summe ihrer zugehörigen Urnenwahllokale vor.

Sofern die einem Briefwahlbezirk zugeordneten Urnenwahllokale im Hinblick auf Wahlbeteiligung und Parteienpräferenz als homogen bezeichnet werden können, ist es möglich, die Briefwahl proportional auf die zugehörigen Urnenwahllokale zu „verteilen“ (gemäß den Wahlberechtigten mit Sperrvermerk). Diese Vorgehensweise entspricht der bisherigen Version der verwendeten Software-Implementierung (eiPack).

Die große Herausforderung der Briefwahlorganisation, samt coronakonformer Auszählung, hat 2021 jedoch eine andere Vorgehensweise bei der Einteilung der Briefwahlbezirke gefordert. In der Konsequenz sind häufiger als bisher Urnenwahllokale, mit zu erwartendem heterogenem Abstimmungsverhalten, zu gemeinsamen Briefwahlbezirken zusammengefasst worden. Die für das Modell erforderliche Homogenität ist damit nicht gegeben.

### **Notwendige Weiterentwicklung des Verfahrens**

Um die bei der Bundestagswahl 2021 zu erwartende erhebliche Zahl der Briefwahlstimmen dennoch in das Verfahren zu integrieren, wurde in der Dortmunder Statistik ein neues Verfahren zur Aufteilung der Briefwahlstimmen auf die Urnenwahllokale entwickelt. Dieses basiert im Wesentlichen auf dem „Satz von Bayes“ und der Annahme, dass sich Wähler\*innen, die sich für die Stimmabgabe per Brief entschieden haben, konditional auf ihre Parteientscheidung ähnlich verhalten, wie die Wähler\*innen im zugehörigen Urnenwahllokal. Das heißt:

Für eine Person X, einen Stimmbezirk j und eine Partei i gilt nach dem Satz von Bayes

$$P(X \text{ ist in Stimmbezirk } j | X \text{ wählt Partei } i) = \frac{P(X \text{ wählt Partei } i | X \text{ ist in Stimmbezirk } j)P(X \text{ ist in Stimmbezirk } j)}{P(X \text{ wählt Partei } i)}$$

Durch das Ersetzen der Wahrscheinlichkeit

- $P(X \text{ wählt Partei } i | X \text{ ist in Stimmbezirk } j)$  durch die relative Häufigkeit der Urnenstimmen in Stimmbezirk j, die Partei i gewählt haben und
- $P(X \text{ ist in Stimmbezirk } j)$  durch die relative Häufigkeit der Wahlberechtigten mit Sperrvermerk im betrachteten Briefwahlbezirk, die in Stimmbezirk j liegen,

ergibt sich eine Aufteilung der Briefwahlstimmen für eine Partei i auf die Stimmbezirke eines Briefwahlbezirks proportional zu

$$\frac{\text{Anzahl Urnenstimmen für Partei } i \text{ in Stimmbezirk } j}{\text{Anzahl Urnenstimmen im Stimmbezirk } j} \cdot (\text{Anzahl Wahlberechtigte mit Sperrvermerk in Stimmbezirk } j).$$

Neben der Anzahl der Wahlberechtigten mit Sperrvermerk werden also auch die Urnenwahlergebnisse in den Stimmbezirken bei der Aufteilung berücksichtigt.

**Weitere methodische Hinweise:**

- *Zur Berechnung einer Wählerwanderung mit aggregierten Daten müssen für die aktuelle Wahl und die Vorwahl räumlich identische Gebiete vorliegen. Durch eine zwischenzeitlich erfolgte Anpassung des Zuschnitts der Wahlbezirke sind die Wahlergebnisse der Bundestagswahl 2017 entsprechend umgerechnet worden.*
- *Das Multinomial-Dirichlet Modell unterstellt, dass die Population bei beiden Wahlen identisch ist. Durch die Veränderung der Wahlberechtigtenzahl sind die 2017er Ergebnisse entsprechend gewichtet worden – das heißt, das relative Wahlergebnis 2017 wird gewahrt, das absolute wird der Anzahl der Wahlberechtigten der Bundestagswahl 2021 angepasst.*

**Literatur**

Klima, A. (2016). Ökologische Inferenz und hybride Modelle: Schätzung der Wählerwanderung in Mehrparteiensystemen, München.

Lau, O., Moore, R. T. und Kellermann, M. (2020). eiPack: Ecological Inference and Higher-Dimension Data Management. R package version 0.2-1.

<https://CRAN.R-project.org/package=eiPack>

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>

Rosen, O., Jiang, W., King, G. und Tanner, M. A. (2001). Bayesian and frequentist inference for ecological inference: the R x C case. In: Statistica Neerlandica, 55, 2, S. 134-156.