
Ökologische Inferenz und hybride Modelle

Schätzung der Wählerwanderung in Mehrparteiensystemen

André Klima



München 2016

Ökologische Inferenz und hybride Modelle
Schätzung der Wählerwanderung in
Mehrparteiensystemen

Dissertation
an der Fakultät für Mathematik, Informatik und Statistik
der Ludwig-Maximilians-Universität
München

vorgelegt von
André Klima
aus Erding

München, den 01.08.2016

Erster Berichterstatter: Prof. Dr. Helmut Küchenhoff
Zweiter Berichterstatter: Prof. Dr. Martin Elff
Tag der Disputation: 20.12.2016

Zusammenfassung

Diese Arbeit widmet sich Methoden zur Schätzung von Wählerwanderungen in Mehrparteiensystemen. Die Schätzung kann dabei anhand von Individual- und/oder Aggregatdaten erfolgen. In dieser Arbeit wird sich auf Verfahren konzentriert, bei denen die Schätzung (auch) mit Aggregatdaten erfolgt. Bei der ökologischen Inferenz werden nur Aggregatdaten verwendet, während hybride Modelle Individual- und Aggregatdaten nutzen. Aus beiden Modellklassen werden in dieser Arbeit Modelle für die Schätzung von Wählerwanderung vorgestellt. Ziel der Arbeit ist die Vorstellung und Evaluierung der Modelle. Es zeigt sich, dass die Anwendung der ökologischen Inferenz nicht ohne Probleme ist. Es ist möglich mehrere alternative Wählerwanderungen zu schätzen, die sich zum Teil deutlich unterscheiden. Ein grundlegendes Problem ist auch die Frage nach dem besten Modell. Mit einer umfangreichen Simulationsstudie wird gezeigt, dass die Modelle unterschiedlich gut geeignet sind. Die insgesamt beste Schätzung unter den verglichenen Modellen kann mit dem Multinomial-Dirichlet Modell erreicht werden. Bei den hybriden Modellen wird neben den hybriden Multinomial-Normal Modell auch das neue hybride Multinomial-Dirichlet Modell betrachtet. Bei der Evaluierung der Modelle zeigt sich, dass die Schätzung mit den hybriden Modellen besser ist. Dies gilt im Vergleich zur ökologischen Inferenz und fast immer zur Hochrechnungen. Weiterhin wird gezeigt, dass bereits mit wenig Individualdaten und in kleineren Regionen auch mit spezifischen Vorwissen die Schätzung verbessert werden kann. Eine Überprüfung der Kreditäritätsintervalle beim Multinomial-Dirichlet Modell zeigt jedoch eine nur unzureichende Überdeckung. Mit einem auf parametrischen Bootstrap basierenden Korrekturverfahren kann eine deutliche Verbesserung erreicht, das Problem aber nicht gelöst werden. Bei der Betrachtung der praktischen Anwendung zeigen sich weitere Problemfelder. Bei den Aggregatdaten nimmt die Datenaufbereitung eine wichtige Rolle ein. Besonders relevante Punkte werden mit einer Simulationsstudie identifiziert. Anhand der Münchner Nachwahlbefragung wird auf wichtige Aspekte bei der Erhebung und mögliche Probleme mit den Individualdaten eingegangen. Dies umfasst eine Diskussion des beobachteten Antwortverhaltens. Bei der Schätzung der Wählerwanderung zwischen den Bundestagswahlen 2009 und 2013 in München zeigt sich die Eignung des hybriden Multinomial-Dirichlet Modells in der praktischen Anwendung. Die Ergebnisse belegen, dass die FDP vor allem an die CSU und die Grünen an die SPD Wähler verloren haben. Die erstmals angetretene AfD konnte von fast allen Parteien Wähler gewinnen.

Abstract

This thesis addresses methods suitable for estimating voter transitions in multi-party systems. The estimation can be either based on individual-level data and/or aggregate data; in this thesis, only methods using (also) aggregate data for the estimation will be considered. The ecological inference only employs aggregate data, while hybrid models use both individual-level and aggregate data. Methods that can be used for estimating voter transitions from both classes of models are introduced. The aim of this thesis is to present and evaluate these models. It becomes evident that it is not possible to use the ecological inference without difficulties. Multiple alternative voter transitions can be estimated, and some of them are quite different from each other. The question “which model is the best” is also an essential problem. It is possible to show that the different models exhibit different suitability using an extensive simulation study. The overall best estimation quality of all compared models can be achieved with the multinomial-Dirichlet model. Of the hybrid models, the new hybrid multinomial-Dirichlet model is considered, along with the hybrid multinomial-normal model. The evaluation of these two models shows that the estimation quality of hybrid models is superior to the one that can be achieved with ecological inference, but also almost always better than the estimation quality achieved with individual-level data. Moreover, the estimation can be improved using a relatively low amount of individual-level data, or even with specific prior knowledge in smaller regions. However, an examination of the credibility intervals from the multinomial-Dirichlet model reveals that the coverage rate is too low. While a clear improvement of the coverage rate can be achieved using a correction procedure based on parametric bootstrap, the problem is not solved completely. A consideration of the practical application of these models reveals further problem areas. The pre-processing of the data holds an important role. The most important aspects of data pre-processing are identified with a simulation study. Important issues related to the gathering of individual-level data are discussed using the Munich Exit-Poll study. This discussion also encompasses the answer pattern and the non-response rates. The estimation of the voter transition for Munich between the federal elections in 2009 and 2013 demonstrates the suitability of the hybrid multinomial-Dirichlet model in a practical application. The results show that the liberal party (FDP) lost mainly to the conservative party (CSU), and the Green Party (Die Grünen) lost primarily to the social democrats (SPD). The first time running new right party AfD attracted voters from almost all other parties.

Danksagung

Besonders bedanken möchte ich mich bei Helmut Küchenhoff, der mich über die letzten Jahre bei der Erstellung dieser Arbeit sehr gut betreut hat und mir mit fachlichen Rat immer zur Seite stand, sowie auch mit der richtigen Kombination aus Geduld und Druck maßgeblich dazu beigetragen hat, dass diese Arbeit entstanden ist. Bei Prof. Martin Elff möchte ich mich dafür bedanken, dass er sich bereit erklärt als zweiter Berichterstatter für diese Arbeit zu fungieren.

Weiterhin möchte ich mich bei Paul W. Thurner bedanken, der maßgeblich mit hinter diesem interessanten Forschungsprojekt stand, von dem ich viel über die politikwissenschaftliche Seite des Projektes lernen konnte und der mit seinen wertvollen Beiträgen mit zum Gelingen beigetragen hat.

Auch möchte ich mich beim Statistischen Amt der Landeshauptstadt München bedanken, die mit ihren interessanten Fragestellungen und den daraus entstandenen Forschungsprojekt diese Arbeit sehr bereichert haben und überdies uns wichtige Informationen bereitwillig zur Verfügung stellten.

Dank gehört auch allen, die im Rahmen der interdisziplinären Forschungs- und Lehrkooperation von Paul W. Thurner und Helmut Küchenhoff, vor deren Hintergrund diese Arbeit entstanden ist, mit eigenen Projekten unseren Wissensstand über die untersuchten Fragestellungen erhöht haben und mit ihren Einsatz auch maßgeblich zum Gelingen der Münchner Nachwahlbefragung beigetragen haben. Besonders erwähnen möchte ich Thomas Schlesinger und Mirjam Selzer, ohne das dies die Leistung der anderen mindern soll.

Bedanken möchte ich mich auch bei allen Kollegen im Statistischen Beratungslabor, von denen ich in den letzten Jahren viel Lernen durfte und mit denen die Zusammenarbeit immer eine Freude war, sowie allen Mitarbeitern am Institut für Statistik, die mich mit Hilfestellungen, Problemlösungen und Ratschlägen unterstützt haben und stets für eine besonders angenehme Arbeitsatmosphäre gesorgt haben.

And last but not least: Thank you, family.

Inhaltsverzeichnis

I. Einleitung und Struktur der Arbeit.....	1
II. Ökologische Inferenz.....	5
II.1. Einführung	5
II.1.1. Geschichte der ökologischen Inferenz.....	5
II.1.2. Wählerwanderungsanalysen.....	8
II.2. Darstellung des Schätzproblems bei der ökologischen Inferenz.....	10
II.3. Methodenbeschreibung.....	16
II.3.1. Ökologische Regression nach Goodman.....	16
II.3.1.1. Allgemeine Beschreibung.....	16
II.3.1.2. Erweiterungen des Modells.....	17
II.3.1.3. Probleme bei der praktischen Anwendung.....	18
II.3.2. Thomsens Probit/Logit Modell.....	22
II.3.2.1. Modell für den 2x2 Fall.....	22
II.3.2.2. Erweiterung für den RxC Fall.....	24
II.3.2.3. Parks Erweiterungen.....	26
II.3.3. Iterative / Adaptive Modellen.....	28
II.3.3.1. Allgemeine Beschreibung.....	28
II.3.3.2. Modell von Andreadis und Chadjipadelis.....	30
II.3.4. Hierarchische Modelle.....	33
II.3.4.1. Allgemeine Beschreibung.....	33
II.3.4.2. Multinomial–Dirichlet Modell.....	36
II.3.4.3. Multinomial-(log-)Normal Modell.....	38
II.3.4.4. Vorwissen bei der Schätzung von Wählerwanderungen.....	40
II.4. Modellvergleich und Evaluierung ökologische Inferenz.....	43
II.4.1. Datengrundlage.....	43
II.4.2. Allgemeine Modelleigenschaften.....	46
II.4.2.1. Maß für den Modellvergleich.....	46
II.4.2.2. Modell-interne Evaluierung.....	47
II.4.2.3. Festlegungen und Einstellungen für die Simulationsstudie.....	56
II.4.3. Simulationsstudie Modellvergleich.....	60

II.4.3.1. Szenarien für die Simulationsstudie.....	60
II.4.3.2. Ergebnisse der Simulationsstudie.....	64
II.5. Ökologische Inferenz: Schlussbetrachtung.....	69
III. Hybride Modelle.....	71
III.1. Einführung.....	71
III.1.1. Kombination von Individual- und Aggregatdaten.....	71
III.1.2. Darstellung des Schätzproblems bei hybriden Modellen.....	75
III.2. Methodenbeschreibung.....	79
III.2.1. Hybrides Multinomial-(log-)Normal Modell.....	79
III.2.2. Hybrides Multinomial-Dirichlet Modell.....	81
III.3. Simulationsstudie Evaluierung hybrider Modelle.....	85
III.3.1. Szenarien für die Simulationsstudie.....	85
III.3.2. Festlegungen für die Simulationsstudie.....	89
III.3.3. Ergebnisse der Simulationsstudie.....	90
III.4. Erweiterte Betrachtung des (hybriden) Multinomial-Dirichlet Modells.....	94
III.4.1. Simulationsstudie Priori-Wissen und Umfang der Individualdaten.....	94
III.4.1.1. Szenarien für die Simulationsstudie.....	95
III.4.1.2. Ergebnisse der Simulationsstudie.....	100
III.4.2. Kreditibilitätsintervalle beim (hybriden) Multinomial-Dirichlet Modell.....	103
III.4.2.1. Güte der Kreditibilitätsintervalle.....	104
III.4.2.2. Korrektur der Kreditibilitätsintervalle.....	109
III.5. Hybride Modelle: Schlussbetrachtung.....	117
IV. Praktische Fragestellungen und Anwendungsbeispiel.....	120
IV.1. Datenaufbereitung bei der ökologischen Inferenz.....	120
IV.1.1. Kritische Aspekte bei der Datenaufbereitung.....	121
IV.1.2. Simulationsstudie zur Auswirkung der Datenaufbereitung.....	125
IV.2. Münchner Nachwahlbefragung.....	130
IV.2.1. Planung der Nachwahlbefragung.....	131
IV.2.2. Ergebnisse der Nachwahlbefragung.....	134
IV.3. Schätzung der Wählerwanderung in München: Bundestagswahlen 2009/2013...	139
IV.3.1. Datengrundlage und Datenaufbereitung.....	140
IV.3.2. Modellparameter und -evaluierung.....	142
IV.3.3. Modellergebnisse.....	144
V. Schlussbemerkungen.....	148

VI. Quellen.....	152
VI.1. Literaturverzeichnis.....	152
VI.2. Datenverzeichnis.....	156
VI.3. Software / Pakete.....	157
VII. Abbildungsverzeichnis.....	158
VIII. Tabellenverzeichnis.....	162
IX. Berücksichtigte Manuskripte und Forschungsergebnisse.....	166
X. Anhang.....	169
X.1. Beschreibung des Modells von Grofman und Merrill.....	169
X.2. Multinomial-Normal Modell: Konvergenz der Ketten.....	170
X.3. Multinomial-Dirichlet Modell: Korrekturverfahren Kreditabilitätsintervalle.....	171
X.4. Fragebögen bei der Münchner Nachwahlbefragung.....	174

I. Einleitung und Struktur der Arbeit

Diese Arbeit widmet sich Methoden zur Schätzung von Wählerwanderungen in Mehrparteiensystemen. Dafür wird das Wahlverhalten von Wählern bei mindestens zwei Wahlen betrachtet und analysiert. Entsprechende Analysen liefern Informationen über die Veränderungen in der politischen Landschaft und den Erfolg von Parteien bzw. Kandidaten. Die Schätzung von Wählerwanderungen kann dabei anhand von Individual- und/oder Aggregatdaten erfolgen. Diese Arbeit konzentriert sich auf Verfahren, bei denen die Schätzung der Wählerwanderung anhand von Aggregatdaten erfolgt. Verfahren, bei denen die Schätzung primär oder ausschließlich mit Individualdaten durchgeführt wird, werden nicht behandelt.

Die zur Schätzung geeigneten Verfahren können in zwei allgemeine Modellklassen unterteilt werden. Bei der ökologischen Inferenz erfolgt die Schätzung des interessierenden Zusammenhangs alleinig anhand der Aggregatdaten. Die Besonderheit bei dieser Modellklasse ist, dass mit aggregierten Daten auf das Verhalten von Individuen geschlossen werden soll. Eine neuere Entwicklung stellen die sogenannten hybriden Modelle dar, bei denen Individual- und Aggregatdaten in einem gemeinsamen Modell zur Schätzung verwendet werden. Bei der Schätzung von Wählerwanderungen handelt es sich bei den Aggregatdaten meistens um die amtlichen Wahlergebnisse der betrachteten Wahlen auf Wahlkreis- oder Wahlbezirksebene, während die Individualdaten gesondert, z. B. im Rahmen einer Nachwahlbefragung, erhoben werden. Beide Modellklassen, die ökologische Inferenz und die hybriden Modelle, sind dabei nicht auf die Schätzung von Wählerwanderungen festgelegt. Sie können und werden auch in anderen Anwendungsgebieten genutzt.

Der inhaltliche Hintergrund der durchgeführten Betrachtungen stammt aus der Politikwissenschaft. Mit den amtlichen Wahlergebnissen können nur die absoluten Gewinne oder Verluste einer Partei oder eines Kandidaten ermittelt werden, jedoch ist es nicht möglich, ohne weitere Annahmen die Veränderungen genauer zu beschreiben. Mit Wählerwanderungsanalysen lassen sich Veränderungen in der politischen Parteienlandschaft detaillierter erklären. Sie stellen daher ein wichtiges Hilfsmittel zur Beschreibung der Dynamik in politischen Systemen dar. Weiterhin ist es möglich, so den Erfolg von Parteien bzw. Kandidaten, Wahlprogrammen oder Wahlkämpfen zu analysieren und zu bewerten. Entsprechende Betrachtungen können sich dabei historischen Fragestellungen, z. B. der Wählerwanderung hin zur NSDAP in der Weimarer Republik,

oder aktuellen Fragestellungen, z. B. von welchen Parteien gewinnt die Alternative für Deutschland die meisten Wähler, widmen.

Wählerwanderungsanalysen stellen aber auch einen bedeutenden – und populären – Teil der Wahlberichtserstattung dar. Der ARD präsentiert z. B. in Zusammenarbeit mit dem Umfrageinstitut Infratest dimap auf Individualdaten basierende Analysen bei Bundestags- und Landtagswahlen. Die Ergebnisse werden nur für das gesamte Wahlgebiet präsentiert. Es besteht jedoch auch eine Nachfrage nach Wählerwanderungsanalysen auf feinerer Ebene und ein Teil der deutschen Städte führt dementsprechend eigene Wählerwanderungsanalysen durch. Eine Erhebung von Individualdaten kann dabei nicht von allen Städten durchgeführt werden, so dass bei diesen die Schätzung der Wählerwanderungen nur anhand von Aggregatdaten erfolgt.

Die im Folgenden vorgestellten Betrachtungen entstanden eingebettet im Rahmen einer langfristigen Forschungs- und Lehrkooperation des Lehrstuhls für empirische Politikforschung, Prof. Paul W. Thurner, und des Statistischen Beratungslabors der LMU, Prof. Helmut Küchenhoff. Während dieser Zusammenarbeit kam es auch zu einer Kooperation mit dem Statistischen Amt der Landeshauptstadt München, bei welcher der Frage nachgegangen wurde, welche Methode der ökologischen Inferenz zur Schätzung von Wählerwanderungen genutzt werden sollte. Das als Ergebnis der Untersuchungen empfohlene Multinomial-Dirichlet Modell wird jetzt vom Statistischen Amt für die Schätzung der Wählerwanderung in der Wahlnacht herangezogen. Weiterhin ermöglichte die Zusammenarbeit mit dem Statistischen Amt einen Einblick in die praktischen Fragestellungen bei der Schätzung von Wählerwanderungen. Den Fragestellungen, die dabei die Datenaufbereitung betreffen, wurde ebenfalls in der Untersuchung nachgegangen. Beide Analysen sind Teil dieser Arbeit.

Ein Ziel der Forschungs- und Lehrkooperation war die Erweiterung des Methodenpools für die Schätzung von Wählerwanderungen. Der Fokus lag dabei vor allem auf hybriden Modellen, da in diesen das größte Potential für zukünftige Entwicklungen gesehen wird. Für diesen relativ neuen Modelltyp gibt es jedoch nur wenige praktische politikwissenschaftliche Anwendungen und keine davon widmete sich der Schätzung von Wählerwanderungen in Mehrparteiensystemen. Es bestand daher die Notwendigkeit, zur praktischen Überprüfung der Modelle selber Individualdaten zu erheben. Dies erfolgte im September 2013 bei der Landtags- und Bundestagswahl in München im Rahmen der Münchner Nachwahlbefragung. Aufgrund der identifizierten guten Eignung des Multinomial-Dirichlet Modells zur Schätzung von Wählerwanderungen wurde dieses zu

einem hybriden Modell erweitert, evaluiert und mit zur Schätzung herangezogen.

Die Ergebnisse der Untersuchungen werden in dieser Arbeit in drei großen Blöcken vorgestellt. Kapitel II. widmet sich vor allem der ökologischen Inferenz. Einführend wird kurz die Geschichte von ökologischen Analysen mit dem Ziel individuelles Verhalten zu inferieren vorgestellt. Anschließend wird auf das Konzept von Wählerwanderungsanalysen eingegangen. Nach diesen beiden eher allgemeineren Kapiteln erfolgt eine Darstellung des Schätzproblems bei der ökologischen Inferenz. Teil davon ist auch die Einführung der in dieser Arbeit genutzten Notation. Anschließend werden mehrere Modelle für die ökologische Inferenz vorgestellt und kurz erklärt. Bei der Auswahl der Modelle wurde berücksichtigt, inwieweit diese zur Schätzung von Wählerwanderungen herangezogen werden oder inwieweit sie dafür geeignet sind. Den Hauptteil des Kapitels stellt die Evaluierung und der Vergleich der vorgestellten Methoden dar. Einen entsprechenden Vergleich, der auch die neueren Methoden mit berücksichtigt, gab es zuvor nicht. Im ersten Schritt werden die Modelle an sich evaluiert und auf die innere Stabilität der Schätzungen, in Abhängigkeit von den Festlegungen, überprüft. Dabei werden die echten Wahldaten für München für die Bundestagswahlen 2005 und 2009 verwendet. Anschließend wird im Rahmen einer Simulationsstudie die Schätzgüte der Modelle verglichen. Durch den systematischen Vergleich der Modelle soll die Eignung der Modelle bei der Schätzung von Wählerwanderungen evaluiert werden. Den Abschluss des Kapitels bildet eine kurze Zusammenfassung der wichtigsten Ergebnisse.

Die hybriden Modelle werden in Kapitel III. eingeführt. Das Kapitel beginnt mit einer allgemeinen Einführung bevor das eigentliche Schätzproblem bei hybriden Modellen bei Wählerwanderungen in Mehrparteiensystemen dargestellt wird. Dabei wird auch auf die angenommene Struktur der Individualdaten eingegangen. Anschließend werden zwei hybride Modelle vorgestellt und kurz erläutert. Es handelt sich bei beiden Modellen um hybride Erweiterungen von bereits in Kapitel II. eingeführten Modellen. Das dabei vorgestellte hybride Multinomial-Dirichlet Modell entstand im Rahmen dieses Forschungsvorhabens. Nach der Vorstellung erfolgt eine Evaluierung und Vergleich beider Modelle. Ein Schwerpunkt bei dieser Betrachtung ist die allgemeine Eignung von hybriden Modellen zur Schätzung von Wählerwanderungen im Vergleich zur ökologischen Inferenz und auch zur Hochrechnungen aus Individualdaten. Der Modellvergleich erfolgt erneut anhand einer Simulationsstudie. Im restlichen Kapitel wird eine weitergehende Evaluierung des (hybriden) Multinomial-Dirichlet Modells präsentiert. Der erste Teil widmet sich dabei vor allem Vorwissen und dem notwendigen Umfang der Individualdaten. Der zweite Teil beschäftigt sich mit der Güte der Kreditabilitätsintervalle

des Modells. Diese Betrachtung erfolgt anhand des Multinomial-Dirichlet Modells, sie ist aber auch für die hybride Erweiterung des Modells relevant. Da dabei eine nur unzureichende Überdeckung der Kreditabilitätsintervalle festgestellt wird, werden vier Korrekturverfahren vorgestellt und auf ihre Eignung überprüft. Den Abschluss des Kapitels bildet eine kurze Zusammenfassung der wichtigsten Ergebnisse.

Den praktischen Fragen bei der Schätzung von Wählerwanderungen mit der ökologischen Inferenz und mit hybriden Modellen widmet sich Kapitel IV.. Der erste Teil des Kapitels beschäftigt sich mit einem Teil der Datengrundlage, den aggregiert vorliegenden Wahlergebnissen. Die möglicherweise auftretenden Probleme werden kurz eingeführt. Angesprochen werden auch die inhaltlichen Implikationen und Vorschläge, wie mit den Problemen bei der Datenaufbereitung umgegangen werden kann. Die aus der Datenaufbereitung resultierenden Konsequenzen für die Schätzgüte werden anschließend mit einer Simulationsstudie evaluiert. Der zweite Teil des Kapitels widmet sich den Individualdaten. Die Betrachtungen erfolgen dabei anhand der Münchner Nachwahlbefragung. Zuerst werden die Überlegungen bei der Konzeption und bei der Umsetzung der Nachwahlbefragung vorgestellt. Anschließend werden ein paar der gewonnenen Erkenntnisse bezüglich der Datengüte präsentiert. Den dritten und letzten Teil des Kapitels stellt die Schätzung der Wählerwanderung bei den Bundestagswahlen 2009 und 2013 in München mit dem hybriden Multinomial-Dirichlet Modell dar. Es werden die für die Schätzung verwendeten Daten vorgestellt, die Modelleigenschaften evaluiert und die geschätzte Wählerwanderung kurz präsentiert.

Den Abschluss der Arbeit bildet Kapitel V.. In den Schlussbemerkungen werden die wichtigsten Ergebnisse der gesamten Arbeit kurz zusammengefasst und miteinander in Verbindung gesetzt. Ebenso werden die noch offenen Probleme und der weitere Forschungsbedarf kurz angesprochen.

Große Teile dieser Arbeit wurden bereits vorab veröffentlicht oder basieren auf noch unveröffentlichten Manuskripten. Ebenso entstand diese Arbeit eingebettet in einer größeren Forschungs- und Lehrkooperation. Eine genaue Darstellung der übernommenen Inhalte und ein Versuch der Abgrenzung der erbrachten Forschungsleistungen befindet sich in Kapitel IX.. Am Anfang von Kapiteln, deren Inhalt zum Teil bereits vorab veröffentlicht wurde, wird entsprechend darauf aufmerksam gemacht. Noch unveröffentlichte Manuskripte werden nur in Kapitel IX. erwähnt. Das genutzte Vorgehen bei der Zitierung bei der Übernahme von Inhalten aus den Vorab-Publikationen und den unveröffentlichten Manuskripten wird ebenfalls in Kapitel IX. erläutert.

II. Ökologische Inferenz¹

II.1. Einführung

Die ökologische Inferenz, auch Ebenen-überschreitende Inferenz genannt, beschreibt das Konzept aus aggregierten Daten Schlüsse über Zusammenhänge auf einer niedrigeren, d. h. weniger stark aggregierten Ebene, zu treffen. Meistens, aber nicht ausschließlich, wird mit der ökologischen Inferenz versucht, Zusammenhänge auf der Ebene von Individuen mit aggregierten Daten aufzudecken, beziehungsweise zu schätzen.² Der Begriff „ökologisch“ leitet sich dabei aus der geschichtlichen Benennung der zugrundeliegenden Gebietseinheiten ab.³ Durch die charakterisierende Ebenen-überschreitende Inferenz unterscheidet sich die ökologische Inferenz von der „klassischen“ Inferenz, bei der Schlüsse auf derselben Ebene, der Ebene der Daten, getätigt werden. Diese Unterscheidung erlaubt auch recht einfach eine Differenzierung und Abgrenzung zwischen einer ökologischen Analyse, bei der Aussagen über die betrachteten Gebietseinheiten an sich gewonnen werden sollen, und der ökologischen Inferenz, wo aus zum Teil vergleichbaren Informationen Zusammenhänge für Teile der betrachteten Gebietseinheiten, z. B. deren Einwohner, geschätzt werden sollen.

II.1.1. Geschichte der ökologischen Inferenz

Die ersten bekannten Analysen bei denen versucht wurde, eine ökologische Fragestellung zu beantworten, stammen aus den Anfängen des 20. Jahrhunderts. Bereits 1919 erschien eine Arbeit von Ogburn und Goltra, in welcher diese versuchten, mittels ökologischer Korrelation und einer Regressionsanalyse die Gender Gap beim Wahlverhalten in Oregon zu ermitteln. Ebenfalls noch in der ersten Hälfte des 20. Jahrhunderts beschäftigten sich z. B. Bernstein (1932) und Gehlke und Biehl (1934) mit solchen Analysen. Aber erst in der zweiten Hälfte des 20. Jahrhunderts wurde sich erneut

-
- 1 Teile dieses Kapitels wurden bereits vorab in Klima, André / Thurner, Paul W. / Molnar, Christoph / Schlesinger, Thomas / Küchenhoff, Helmut: Estimation of Voter Transitions Based on Ecological Inference: An Empirical Assessment of Different Approaches. *Advances Statistical Analysis*, 100(2), 2016, S. 133–159, sowie in Klima, André / Küchenhoff, Helmut / Thurner, Paul W.: Analysis of voter transition using ecological data: Comparison of different approaches for Munich election data. *Proceedings of the 28th International Workshop on Statistical Modelling (Palermo, 08.07.–12.07.2013)*, 1, 2013, S. 203–207 veröffentlicht. Dies betrifft alle Teile dieses Kapitels, vor allem aber die Methodenvorstellung in Kapitel II.3. und den Methodenvergleich in Kapitel II.4..
 - 2 Vgl. Cho, Wendy K. Tam / Manski, Charles F.: Cross – Level / Ecological Inference, in: Brady, Henry E. / Collier, David / Box-Steffensmeier, Janet M.: *Oxford Handbook of Political Methodology*, Oxford, 2008, S. 547.
 - 3 Vgl. King, Gary: *A Solution to the Ecological Inference Problem. Reconstructing Individual Behavior from Aggregate Data*, Princeton, 1997, S. XV.

und verstärkt dieser Fragestellung gewidmet.⁴

Zu Beginn der zweiten Hälfte des 20. Jahrhunderts gab es mehrere Entwicklungen, die die ökologische Inferenz für mehrere Jahrzehnte prägen sollten und heute noch beeinflussen. 1950 veröffentlichte Robinson eine Arbeit⁵, in welcher er die Eignung der ökologischen Korrelation zur Schätzung von Zusammenhängen auf Individualebene bewertete. Bei dem vom ihm betrachteten Beispiel ergab sich ein Ergebnis, dass auf inhaltlicher Ebene als vernichtend betrachtet werden kann: Die ökologische Korrelation war eindeutig positiv, während die Korrelation auf Individualebene negativ war. Eine Übertragung der ökologischen Korrelation, bestimmt mit den für die Gebiete vorliegenden aggregierten Daten, auf Ebene der Individuen würde daher zu einem falschen inhaltlichen Schluss führen. Robinson prägte mit diesem Ergebnis den Begriff des „ökologischen Fehlschlusses“, der genau dies beinhaltet: Ein Zusammenhang auf Aggregatebene muss nicht notwendigerweise dem Zusammenhang auf der Individualebene entsprechen. Robinsons Erkenntnis hatte eine starke Auswirkung auf die Nutzung von ökologischen Analysen und soll mit dazu beigetragen haben, dass Inferenz basierend auf Individualdaten zum dominierenden Verfahren in den Sozialwissenschaften in den Folgejahren geworden ist.⁶

1953 folgten zwei Arbeiten, die dafür gesorgt haben, dass die ökologische Inferenz, trotz Robinsons Erkenntnis, weiterhin genutzt wurde. Beide Arbeiten nahmen ebenfalls maßgeblich Einfluss auf die Entwicklungen in den folgenden Jahrzehnten und sollen lange Zeit als konkurrierend aufgefasst worden sein. Goodman stellte die ökologische Regression⁷ und Duncan und Davis die Methode der Ränder vor⁸. Bei der ökologischen Regression werden die interessierenden Größen als Parameter einer Regression geschätzt, während bei der Methode der Ränder die maximal vorhandene Information über die interessierenden Größen aus den ökologischen Daten extrahiert wird. Goodmans ökologische Regression und dessen Erweiterungen sollten für lange Zeit das Standardmodell der ökologischen Inferenz sein.⁹

Beginnend mit der zweiten Hälfte der 90er Jahren ist wieder eine verstärkte

4 Vgl. Achen, Christopher H. / Shively, W. Phillips: Cross-level inference, Chicago, 1995, S. 8.

5 Robinson, William S.: Ecological correlations and the behavior of individuals. *American Sociological Review*, 15(3), 1950, S. 351–357.

6 Vgl. King, 1997, S. 5.

7 Goodman, Leo A.: Ecological Regressions and Behavior of Individuals. *American Sociological Review*, 18(6), 1953, S. 663–664.

8 Duncan, Otis Dudley / Davis, Beverly: An alternative to ecological correlation. *American Sociological Review*, 18(6), 1953, S. 665–666.

9 Vgl. Achen / Shively, 1995, S. 29.

Beschäftigung mit und aktive wissenschaftliche Diskussion über die ökologischen Inferenz beobachtbar. Besonders einflussreich war Gary Kings 1997 erschienenes Buch „A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data“¹⁰. King stellte in diesem Buch seine Methode für die ökologischen Inferenz vor, bei der ihm zufolge Duncan und Davis's Methode Ränder mit einem regressionsbasierenden Ansatz kombiniert werden. Obwohl Kings Methode nach der Veröffentlichung stark kritisiert wurde¹¹, stellt die Veröffentlichung des Buches den Beginn einer verstärkten Methodendiskussion dar und eine Vielzahl von neuen Lösungsansätzen wurde in den Folgejahren vorgestellt. Eine Betrachtung der Herkunft der Forscher, die sich dabei mit der ökologischen Inferenz beschäftigen, zeigt, dass vor allem aus zwei Anwendungsgebieten Beiträge stammen.

Das erste große Anwendungsgebiet der ökologischen Inferenz, aus dem in den letzten Jahren wertvolle Beiträge zur Methodenweiterentwicklung gekommen sind,¹² stellt die Epidemiologie dar. Ziel ist es z. B., das Auftreten einer Krankheit anhand von Kovariablen der Umgebung zu beschreiben. Nicht immer liegen aber Messungen auf Individualebene vor, zum Teil sind nur Informationen über das Auftreten der Krankheit und Informationen von Individuen, auf-aggregiert über Gebietseinheiten, vorhanden. In diesem Fall müssen Methoden der ökologischen Inferenz herangezogen werden, um Informationen über mögliche Risikofaktoren gewinnen zu können. Gekennzeichnet sind diese Analysen davon, dass meistens mehr als ein möglicher Risikofaktor bei der Modellierung berücksichtigt wird.

Ein zweites großes Anwendungsgebiet stellt die Politikwissenschaft dar.¹³ Auch hier stehen nicht immer Individualdaten in entsprechendem Umfang oder entsprechender Güte

10 King, 1997.

11 Siehe z. B. Cho, Wendy K. Tam: Iff the Assumption Fits...: A Comment on the King Ecological Inference Solution. *Political Analysis*, 7(1), 1998, S. 143–163, Cho, Wendy K. Tam / Gaines, Brian J.: The Limits of Ecological Inference. The Case of Split-Ticket Voting, in: *American Journal of Political Science*, Volumen 48, Nummer 1, 2004, S. 152–171, Anselin, Luc / Cho, Wendy K. Tam: Spatial Effects and Ecological Inference. *Political Analysis*, 10(3), 2002, S. 276–297 und Freedman, David A.: Ecological Inference and the Ecological Fallacy, Technical Report No. 549, Berkely, 1999.

12 Siehe z. B. Salway, Ruth / Wakefield, Jon: A hybrid model for reducing ecological bias. *Biostatistics*, 9(1), 2008, S. 1–17, Smoot, Elizabeth / Haneuse, Sebastien: On the Analysis of Hybrid Designs that Combine Group- and Individual-Level Data, *Biometrics*, 71(1), 2015, S. 227–236 und Wakefield, Jon: Ecological inference for 2x2 tables, *Journal of Royal Statistical Society A*, 167(3), 2004, S. 385–445.

13 Siehe z. B. Andreadis, Ioannis / Chadjipadelis, Theodore: A Method for the Estimation of Voter Transition Rates. *Journal of Elections, Public Opinion and Parties*, 19(2), 2009, S. 203–218, Greiner, D. James / Quinn, Kevin M.: RxC ecological inference: bounds, correlations, flexibility and transparency of assumptions. *Journal of the Royal Statistical Society*, 172(1), 2009, S. 67–81, King, Gary / Rosen, Ori / Tanner, Martin A.: Binomial-Beta Hierarchical Models for Ecological Inference. *Sociological Methods & Research*, 28(1), 1999, S. 61–90 und Rosen, Ori / Jiang, Wenxin / King, Gary / Tanner, Martin A.: Bayesian and frequentist inference for ecological inference: the R x C case. *Statistica Neerlandica*, 55(2), 2001, S. 134–156.

für die interessierenden Sachverhalte zur Verfügung. Dies umfasst z. B. geschichtliche Betrachtungen, wie die Analyse des Wahlverhalten im Dritten Reich.¹⁴ Es sind aber auch bei aktuellen Fragestellungen nicht immer Individualdaten im ausreichenden Maße vorhanden. Dies kann zum Teil auf rechtliche Einschränkungen zurückgeführt werden, aber auch zu hohe Kosten können einer Erhebung entgegenstehen.¹⁵ Während eines der primären Anwendungsgebiete der ökologischen Inferenz in der Politikwissenschaft in den USA die Analyse des Ethnien-spezifischen Wahlverhaltens darstellt, liegt der Forschungsschwerpunkt in Deutschland hauptsächlich bei der Analyse des Wahlverhaltens von sozialen Gruppen und der Wählerwanderung zwischen zwei Wahlen.

Diese Arbeit widmet sich vor allem der letzteren Fragestellung und dieses Kapitel primär der Anwendbarkeit der ökologischen Inferenz zur Schätzung von Wählerwanderungen in Mehrparteiensystemen. Das grundlegende Schätzproblem ist jedoch in diesen Fällen auf einer abstrahierten Ebene beinahe identisch zu anderen Sachverhalten, auch wenn sich fundamental andere Schwierigkeiten aufgrund abweichender Datensituationen bei der Anwendung ergeben.

II.1.2. Wählerwanderungsanalysen

Die Betrachtung der Wählerwanderung stellt dabei die Betrachtung des Wahlverhaltens von Wahlberechtigten bei mindestens zwei Wahlen dar. Bei den Wahlen kann sich ein Wahlberechtigter jeweils für eine Wahlalternative, z. B. eine Partei oder auch die Nichtwahl, entscheiden. Die Wahlentscheidung kann bei den betrachteten Wahlen identisch sein, oder kann sich unterscheiden. Es ist z. B. möglich, dass ein Nichtwähler bei einer Wahl bei der darauf folgenden Wahl für eine Partei stimmt.

Im Allgemeinen werden bei entsprechenden Analysen Wahlen betrachtet, die nicht am gleichen Tag stattfinden, es handelt sich daher um eine Betrachtung des Wahlverhaltens über die Zeit. In der Mehrzahl der Fälle werden gleichwertige Wahlen, d. h. Wahlen auf der gleichen Ebene, wie z. B. zwei Bundestagswahlen, betrachtet. Es ist aber auch möglich, Wahlen auf unterschiedlicher Ebene, z. B. eine Bundestags- und eine Kommunalwahl, zu analysieren. Das betrachtete Gebiet kann dabei das gesamte Wahlgebiet oder ein Teil des Wahlgebietes sein, z. B. kann bei einer Bundestagswahl die Wählerwanderung für das

14 Siehe z. B. King, Gary / Rosen, Ori / Tanner, Martin A./ Wagner, Alexander F.: Ordinary economic voting behavior in the extraordinary election of Adolf Hitler. *Journal of Economic History* 68(04), 2008, S. 951–996.

Für eine ökologische Analyse ohne Anspruch das individuelle Wahlverhalten zu erklären, siehe z. B. Thurner, Paul W. / Klima, André / Küchenhoff, Helmut: Agricultural structure and the rise of the Nazi Party reconsidered. *Political Geography* 44, 2015, S. 50–63.

15 Vgl. King, 1997, S. 8–11.

gesamte Bundesgebiet, d. h. für alle Wahlberechtigten, oder für ein Teil des Gebietes, wie z. B. für eine Großstadt wie München, ermittelt bzw. geschätzt werden.

Mit der Wählerwanderungsanalyse wird versucht, die aus dem individuellen Wahlverhalten resultierenden Bewegungen zwischen den einzelnen Wahlalternativen im betrachteten Gebiet zu beschreiben. Bei den Analysen ist es möglich, wie auch primär in dieser Arbeit, die direkten Bewegungen zwischen den Parteien zu betrachten. Alternativ können auch die saldierten Wanderungsströme zwischen den Parteien genutzt werden. Bei letzteren werden die gegenläufigen Wanderungsströme, d. h. die Gewinne von und die Verluste an eine andere Partei, gegeneinander aufgerechnet. Eine entsprechende Darstellung wird z. B. bei der ARD-Wahlberichtserstattung genutzt, welche auf die geschätzte Wählerwanderung von Infratest dimap zurückgreift. Abbildung 1 zeigt die geschätzten Salden für die Bundestagswahlen 2009 und 2013 für die Union (CDU und CSU). So ermittelte Wanderungsbewegungen zwischen den Parteien bieten einen Anhalt für die Veränderungen in der politischen Landschaft und den Erfolg von Kandidaten bzw. Parteien.¹⁶

Wählerwanderungsanalysen werden im Allgemeinen basierend auf zwei unterschiedlichen Datenquellen geschätzt: Befragungsdaten oder die aggregierten

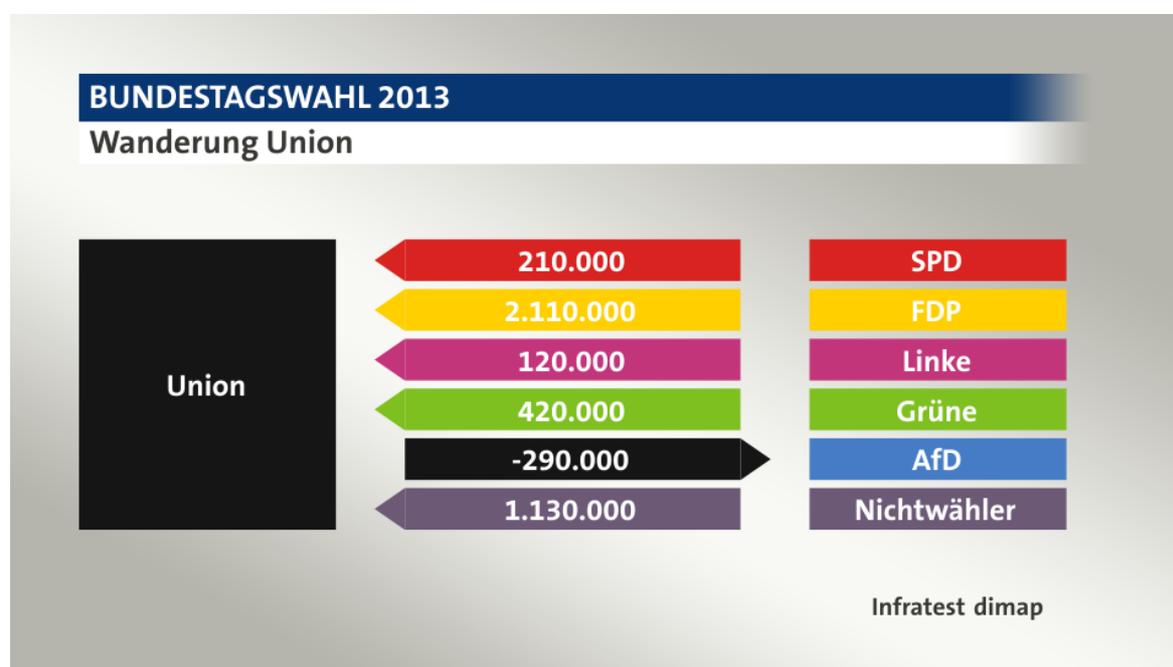


Abbildung 1: Wählerströme von bzw. zur Union (CDU, CSU) bei der Bundestagswahl 2013 im Vergleich zur Bundestagswahl 2009. Dargestellt sind die saldierten Wählerwanderungen, d. h. die absoluten Gewinne bzw. Verluste der Union. (Quelle: <https://wahl.tagesschau.de/wahlen/2013-09-22-BT-DE/analyse-wanderung.shtml>)

16 Vgl. Hilmer, Richard / Kunert, Michael: Wählerwanderung: Das Modell von Infratest dimap. In: Wahlen und Wähler. VS Verlag für Sozialwissenschaften, 2005, S. 134.

amtlichen Wahlergebnisse. Infratest dimap greift bei seinen Wählerwanderungsanalysen primär auf Befragungsdaten zurück.¹⁷ Es handelt sich dabei also um eine Hochrechnung aus Individualdaten und nicht um eine Anwendung der ökologischen Inferenz.

Falls anstelle von Individualdaten die primäre Datenquelle die vorhandenen aggregierten amtlichen Wahlergebnisse sind, muss zur Schätzung der Wählerwanderung auf Methoden der ökologischen Inferenz zurückgegriffen werden. Ausgehend von den bekannten Rändern der Wählerwanderungstabelle, den beobachteten Wahlergebnissen, wird mittels ökologischer Inferenz versucht, die inneren Zellen der Tabelle valide zu schätzen. Im Folgenden wird dieses Schätzproblem kurz formal eingeführt und einige der zur Schätzung von Wählerwanderungen genutzten Methoden kurz vorgestellt.

II.2. Darstellung des Schätzproblems bei der ökologischen Inferenz

Die Untersuchung einer Wählerwanderung mit R Parteien bei der Wahl zum Zeitpunkt t und C Parteien bei der Wahl zum Zeitpunkt t+1 kann mit einer RxC Tabelle, siehe Tabelle 1, veranschaulicht werden. Die Wähler, die bei beiden Wahlen die gleiche Partei¹⁸ gewählt haben, werden dabei als loyale Wähler betrachtet, die Wähler, die eine veränderte Wahlentscheidung aufzeigen, werden als Wechselwähler angesehen. In Tabelle 1 repräsentieren die Zellen auf der Diagonalen die loyalen Wähler, während die Wechselwähler durch die Zellen repräsentiert werden, die nicht auf der Diagonalen liegen. Im Allgemeinen wird bei der ökologischen Inferenz davon ausgegangen, dass die betrachtete Population bei beiden Wahlen identisch ist.

	CSU 2013	SPD 2013	...	NW 2013	Wahl 2009
CSU 2009					23,56%
SPD 2009					14,17%
...					...
NW 2009					26,60%
Wahl 2013	26,91%	17,02%	...	28,80%	100,00%

Tabelle 1: Darstellung der Datensituation bei der ökologischen Inferenz, die hellgrau markierten Ränder sind bekannt, die weißen inneren Zellen unbekannt. Die angegebenen Wahlergebnisse entsprechen dem Ergebnis der Bundestagswahl 2009 respektive 2013 in München, Nichtwähler werden mit NW abgekürzt, ein Teil der Tabelle wird nicht dargestellt. (Datenquelle: Stadt München¹⁹).

17 Vgl. Hilmer / Kunert, 2005, S. 140–144.

18 Der Begriff Partei wird im Folgenden auch für alle möglichen Wahlentscheidungen / Alternativen, d. h. auch für die Nichtwähler, genutzt.

19 Statistisches Amt der Landeshauptstadt München. Amtliches Endergebnis der Bundestagswahl am 27.09.2009, Erst- und Zweitstimmenergebnisse der Landeshauptstadt München, 2009. <http://www.mstatistik-muenchen.de/themen/wahlen/wahlberichterstattung/btw2009/index.html> (zuletzt aufgerufen am: 03.07.2016) und

Statistisches Amt der Landeshauptstadt München. Amtliches Endergebnis der Bundestagswahl am 22.09.2013, Erst- und Zweitstimmenergebnisse der Landeshauptstadt München, 2013. <https://www.muenchen.de/rathaus/Stadtfinfos/Statistik/Wahlen/Bundestagswahl.html> (zuletzt aufgerufen

Mit nur einer vorliegenden Tabelle handelt es sich um ein nicht schätzbares Problem, da mindestens $(R-1) \times (C-1)$ Parameter/Werte zu schätzen sind, während nur eine Beobachtung vorliegt. Das gleiche gilt, wenn mehrere Gebietseinheiten betrachtet werden, aber für jede betrachtete Tabelle ein eigener Zusammenhang angenommen wird. In diesem Fall wäre es notwendig, für jede betrachtete Tabelle $(R-1) \times (C-1)$ Parameter/Werte zu schätzen, während sie jeweils nur eine Beobachtung darstellt. Dieser Sachverhalt ist hauptsächlich dafür verantwortlich, dass dem Schätzproblem der ökologischen Inferenz eine fundamentale Unbestimmtheit zugeschrieben wird.²⁰

Tabelle 2 und 3 stellen einen weiteren Aspekt dar, der ebenfalls mit zum Problem der Unbestimmtheit bei der ökologischen Inferenz zählt. Es werden in den beiden Tabellen zwei unterschiedliche fiktive Wählerwanderungen für die gleiche Wahl dargestellt. Beide können zu den beobachteten Wahlergebnissen in Tabelle 1 geführt haben. Tabelle 2 zeigt dabei den angenommenen Fall, bei dem sich alle Wähler der dargestellten Parteien bei der Wahl im Jahr 2009 bei der Wahl im Jahr 2013 erneut für die gleiche Partei entschieden haben und die Zugewinne alleine von den nicht dargestellten kleineren Parteien stammen, d. h. die Wähler der großen Parteien werden als loyal ihrer Partei gegenüber angesehen.

Tabelle 3 zeigt einen fast umgekehrten Fall. Alle CSU-Wähler des Jahres 2009 entscheiden sich 2013 nicht erneut für die CSU und es wird weiterhin angenommen, dass die ehemaligen CSU-Wähler für die SPD oder die Nichtwähler gestimmt haben. Die Wählerschaft der SPD im Jahr 2013 besteht daher bei dieser angenommenen Wählerwanderung nur aus ehemaligen CSU-Wählern, während die Stimmen der CSU bei der Wahl 2013 von ehemaligen SPD-Wählern bzw. Nichtwählern stammen. Die beiden dargestellten Alternativen können anhand der bekannten Wahlergebnisse, dargestellt in den Rändern der Tabellen, nicht verworfen werden, da beide die von ihnen aufgestellten Restriktionen erfüllen. Zusätzlich stellen diese beiden Beispiele nur zwei der vielen

	CSU 2013	SPD 2013	...	NW 2013	Wahl 2009
CSU 2009	23,56%	0,00%	...	0,00%	23,56%
SPD 2009	0,00%	14,17%	...	0,00%	14,17%
...	3,35%	2,85%	...	2,20%	...
NW 2009	0,00%	0,00%	...	26,60%	26,60%
Wahl 2013	26,91%	17,02%	...	28,80%	100,00%

Tabelle 2: Darstellung einer möglichen Wählerwanderung in München zwischen den Wahlen 2009 und 2013, angenommen werden hohe Loyalitätsraten. Alle Wahlberechtigten mit einer Entscheidung für die CSU, die SPD und die Nichtwähler der Wahl 2009 entscheiden sich bei der Wahl 2013 identisch.

am: 03.07.2016).

20 Vgl. Cho / Manski, 2008, S. 549–551.

	CSU 2013	SPD 2013	...	NW 2013	Wahl 2009
CSU 2009	0,00%	17,02%	...	6,54%	23,56%
SPD 2009	14,17%	0,00%	...	0,00%	14,17%
...	0,00%	0,00%	...	8,40%	...
NW 2009	12,74%	0,00%	...	13,86%	26,60%
Wahl 2013	26,91%	17,02%	...	28,80%	100,00%

Tabelle 3: Darstellung einer möglichen Wählerwanderung in München zwischen den Wahlen 2009 und 2013. Angenommen werden niedrige Loyalitätsraten. Alle Wähler der CSU und der SPD der Wahl 2009 entscheiden sich bei der Wahl 2013 für eine andere Partei. Die Hälfte der Nichtwähler 2009 wählt auch 2013 nicht.

theoretisch zulässigen Wählerwanderungen bei diesen Rändern dar.²¹

Beide Tabellen stellen dabei den Anteil der Wähler dar, die eine bestimmte Parteienkombination gewählt haben. So bedeutet der Wert in der Zelle [CSU 2009, CSU 2013] in Tabelle 2 das 23,56% aller Wahlberechtigten für diese Kombination gestimmt haben, d. h. in beiden Wahlen die CSU gewählt haben. In der Summe ergeben alle Werte in der Tabelle eins und pro Zeile bzw. Spalte das beobachtete amtliche Wahlergebnis.

Im Folgenden wird für die beobachteten relativen Wahlergebnisse, die Ränder der Tabelle, eine Schreibweise der Form $P_{\text{CSU},2009}$ verwendet, im Beispiel handelt es sich um das Wahlergebnis der CSU bei der Wahl im Jahr 2009, d. h. $P_{\text{CSU},2009} = 23.56\%$ in Tabelle 1. Wenn keine Partei direkt spezifiziert werden soll, werden Formulierungen der Form $P_{p,2009}$ genutzt. Der Suffix p gibt dann entweder die Zeile oder Spalte in der Kreuztabelle an, während das Jahr erneut die Wahl kennzeichnet. Wenn anstelle von relativen Werten die absoluten Werte betrachtet werden, wird auf Formulierung der Form $N_{\text{CSU},2009}$ zurückgegriffen. Für die inneren Zellen der Tabelle wird analog dazu $P_{\text{CSU}2009,\text{CSU}2013}$ bzw. $N_{\text{CSU}2009,\text{CSU}2013}$ genutzt. Die erste Partei im Suffix bezeichnet dabei die Partei bei der Wahl zum Zeitpunkt t , die zweite Partei die Wahl zum $t+1$. Falls der Parteiname verwendet wird, ist zu besserer Lesbarkeit das Jahr ergänzt. Wenn nur die Werte eines bestimmten Gebietes i gemeint sind, wird zusätzlich jeweils der Suffix i ergänzt.

Eine alternative und ebenfalls oft genutzte Darstellung des Zusammenhangs zwischen beiden Wahlen ist die Betrachtung der Übergangsraten²² zwischen den Parteien. Dabei handelt es sich um den relativen Anteil der Wähler einer Partei r zur Wahl t , die sich bei der Wahl zum Zeitpunkt $t+1$ für die Partei c entschieden haben. Tabelle 4 zeigt die jeweiligen Übergangsraten $\beta_{r,c}$ für Tabelle 1. Die Benennung der jeweiligen Kombination folgt dabei den oben erwähnten Prinzipien für die absoluten bzw. relativen Zellwerte. Es sei darauf hingewiesen, dass es sich bei der Darstellung in Tabelle 4 nicht um eine echte

21 Vgl. Elff, Martin / Gschwend, Thomas / Johnston, Ron J.: Ignoramus, ignorabimus? On uncertainty in ecological inference. *Political Analysis*, 16(1), 2008, S. 72–73.

22 In dieser Arbeit wird der Begriff Übergangswahrscheinlichkeit synonym zu Übergangsrate verwendet.

	CSU 2013	SPD 2013	...	NW 2013	Wahl 2009
CSU 2009	$\beta_{CSU2009,CSU2013}$	$\beta_{CSU2009,SPD2013}$...	$\beta_{CSU2009,NW2013}$	23,56%
SPD 2009	$\beta_{SPD2009,CSU2013}$	$\beta_{SPD2009,SPD2013}$...	$\beta_{SPD2009,NW2013}$	14,17%
...
NW 2009	$\beta_{NW2009,CSU2013}$	$\beta_{NW2009,SPD2013}$...	$\beta_{NW2009,NW2013}$	26,60%
Wahl 2013	26,91%	17,02%	...	28,80%	100,00%

Tabelle 4: Darstellung der Datensituation bei der ökologischen Inferenz, die hellgrau markierten Ränder sind bekannt, die weißen inneren Zellen unbekannt. In den inneren Zellen sind die jeweiligen Übergangsraten ergänzt. Es handelt sich um eine vereinfachte Darstellung, da sich in den inneren Zellen nur die Übergangsraten befinden. Der eigentliche Zellwert ergibt sich aus:

$$\beta_{r,c} \times P_{r,2009}$$

Kreuztabelle handelt, da die inneren Zellen in der Summe nicht den Rändern entsprechen.

Bei Tabelle 2 würde die Übergangsrate $\beta_{CSU2009,CSU2013}$, d. h. der Anteil der Wähler der CSU im Jahr 2009 die 2013 erneut der CSU ihre Stimme gegeben haben, eins betragen. Bei Tabelle 3 würde $\beta_{CSU2009,CSU2013} = 0$ gelten, da sich keiner der 2009er Wähler der CSU 2013 erneut für diese Partei entschieden hätte. Aufgrund der Konstruktion der Übergangsraten als Anteil des Zeilenrandes lässt sich direkt der folgende Zusammenhang aufstellen:

$$\sum_{c=1}^C \beta_{r,c} = 1, \quad \forall r \in \{1, \dots, R\} \quad (1)$$

Gleichung (1) besagt, dass die Zeilensumme aller Übergangsraten für jede Zeile jeweils eins ergibt, sie also eine Partition des Wahlergebnisses zum Zeitpunkt t beschreiben. In dieser allgemeinen Darstellung wird die Position in der Tabelle genutzt, r steht dabei für die Zeile, also die Partei zum Zeitpunkt t , während c die Spalte, die Partei zum Zeitpunkt $t+1$ beschreibt. Neben diesem Zusammenhang in jeder Zeile, lässt sich auch ein Zusammenhang für jede Spalte angeben. Es gilt²³:

$$P_{c,t+1} = \sum_{r=1}^R \beta_{r,c} * P_{r,t}, \quad \forall c \in \{1, \dots, C\} \quad (2)$$

Gleichung (2) beschreibt auf formale Weise einen wahren Zusammenhang zwischen den Wahlergebnissen der beiden Wahlen mit den Übergangsraten. $P_{r,t}$ steht dabei, wie zuvor eingeführt, für das beobachtete (relative) Wahlergebnis bei der Wahl zum Zeitpunkt t für die r -te Partei, während $P_{c,t+1}$ das beobachtete (relative) Wahlergebnis der Partei c bei der Wahl $t+1$ beschreibt. Die Wählerschaft einer Partei bei der Wahl zum Zeitpunkt $t+1$ lässt sich als Summe der Wahlergebnisse jeder Partei zum Zeitpunkt t , multipliziert mit der jeweiligen Übergangsrate zur Partei c darstellen. Inhaltlich bedeutet dies, dass jeder Wähler einer Partei bei der zweiten Wahl eine Wahlentscheidung bei der ersten Wahl hatte.

23 Vgl. z. B. King, 1997, S. 93.

Gleichungen (3) und (4) zeigen diese Verbindung zur besseren Veranschaulichung für das Wahlergebnis der CSU im Jahr 2013:

$$P_{CSU,2013} = \beta_{CSU2009,CSU2013} * P_{CSU,2009} + \dots + \beta_{NW2009,CSU2013} * P_{NW,2009} \quad (3)$$

$$26,91 \% = \beta_{CSU2009,CSU2013} * 23,56 \% + \dots + \beta_{NW2009,CSU2013} * 26,60 \% \quad (4)$$

Ziel der ökologischen Inferenz ist es, entweder die inneren Zellen von Tabelle 1 oder die Übergangsraten in Tabelle 4 zu schätzen. Falls eine der beiden Größen bekannt ist, lässt sich die jeweils andere Größe aus den Rändern problemlos berechnen.

Wie bereits erwähnt, sind bei Vorliegen von nur einer Tabelle mehr Parameter/Werte zu schätzen, als Beobachtungen vorhanden sind und viele Lösungen sind theoretisch zulässig. Aber auch in diesem Fall, wie exemplarisch in Tabelle 2 und 3 veranschaulicht, besitzen die Ränder bereits Informationen über die interessierenden inneren Zellen.²⁴ Es können nur zwischen 0,00% und 72,24% der 2009er CSU-Wähler im Jahr 2013 für die SPD gestimmt haben, während zwischen 0,00% und 100,00% aller 2009er CSU-Wähler auch im Jahr 2013 erneut für die CSU gestimmt haben können. Die Idee, die aus den Rändern resultierenden Schranken für die interessierenden Größen zu berechnen, stammt von Duncan und Davis²⁵ und ist als Methode der Ränder bekannt. In vielen Anwendungsfällen ist das durch die so berechneten Schranken für die interessierenden Größen angebbare Intervall aber zu groß, um von praktischen Interesse und Nutzen zu sein.²⁶

Die zentrale Idee bei der ökologischen Inferenz ist, dass nicht nur die Informationen einer Tabelle genutzt werden. Für die Schätzung werden die Ränder von mehreren RxC Tabellen, also die Ergebnisse von vielen Wahlgebieten, verwendet. Es liegt dann nicht mehr nur eine Tabelle 1 vor, sondern N Tabellen in vergleichbarer Form, eine für jedes betrachtete Gebiet. Bei der Schätzung von Wählerwanderungen handelt es sich bei den betrachteten Gebieten meistens um die Wahlergebnisse auf Wahlkreis- bzw. Wahlbezirksebene. Die Schätzung des individuellen Wahlverhaltens wird dann von der Streuung zwischen den betrachteten Wahlgebieten abgeleitet. Wie bereits angedeutet, handelt es sich aber ohne weitere Restriktion weiterhin um ein nicht schätzbares Problem, da für jedes hinzugenommene Gebiet i (R-1) x (C-1) Parameter/Werte zu schätzen wären, während das Gebiet selber nur eine weitere Beobachtung darstellt. Aus diesem Grund nehmen alle Modelle der ökologischen Inferenz eine bestimmte Struktur zwischen den betrachteten Gebieten an, und die daraus resultierenden Restriktionen ermöglichen erst die

24 Vgl. Elff / Gschwend / Johnston, 2008, S. 72–73.

25 Vgl. Duncan / Davis, 1953.

26 Vgl. Andreadis / Chadjipadelis, 2009, S. 206.

Schätzung. Die dabei getroffenen Annahmen unterscheiden sich jedoch zum Teil recht deutlich zwischen den einzelnen Modellen.

II.3. Methodenbeschreibung

In diesem Kapitel werden mehrere Modelle für die ökologischen Inferenz kurz vorgestellt. Es handelt sich dabei um eine Auswahl und nicht um eine vollständige Darstellung aller existierenden Modelle. Als Aufnahmekriterium wurde die Nutzung dieser Modelle in der Praxis und Eignung der Modelle für die Schätzung von Wählerwanderung im RxC Fall herangezogen.

II.3.1. Ökologische Regression nach Goodman

Die ökologische Regression nach Goodman²⁷ gehört zu den ältesten Verfahren der ökologischen Inferenz und wird auch heute noch genutzt.²⁸ Ausgangspunkt des Modells ist der in Gleichung (3) dargestellte Zusammenhang zwischen erster und zweiter Wahl. Eine solche Gleichung kann für jede Partei für die betrachtete Region, aber auch für jedes betrachtete Gebiet innerhalb der Region aufgestellt werden. Im Gebiet i besitzt die Gleichung für die CSU im Jahr 2013 die folgende Form:

$$P_{CSU,2013,i} = \beta_{CSU2009,CSU2013,i} * P_{CSU,2009,i} + \dots + \beta_{NW2009,CSU2013,i} * P_{NW,2009,i} \quad (5)$$

Gleichung (5) veranschaulicht erneut, dass ohne weitere Restriktionen eine Schätzung der Übergangswahrscheinlichkeiten nicht möglich ist. Es sind alleine für die CSU im Gebiet i R Parameter zu bestimmen. Es liegt aber nur eine Beobachtung, das Wahlergebnisse für das Gebiet i , vor.

II.3.1.1. Allgemeine Beschreibung

Um eine Schätzbarkeit zu ermöglichen, wird bei der ökologischen Regression nach Goodman angenommen, dass die Übergangswahrscheinlichkeiten in der gesamten betrachteten Region identisch sind. Es gilt daher:

$$\beta_{r,c,i} = \beta_{r,c}, \quad \forall r \in \{1, \dots, R\}, c \in \{1, \dots, C\}, i \in \{1, \dots, N\} \quad (6)$$

Die Annahme von gleichen Übergangswahrscheinlichkeiten in allen Gebietseinheiten ist auch als Konstanzannahme bekannt. Dies führt zusammen mit Gleichung (5) direkt zur folgender Modellgleichung der ökologischen Regression:

$$P_{CSU,2013,i} = \beta_{CSU2009,CSU2013} * P_{CSU,2009,i} + \dots + \beta_{NW2009,CSU2013} * P_{NW,2009,i} + \epsilon_i \quad (7)$$

27 Andere Benennungen für die ökologische Regression nach Goodman sind Goodman Regression oder nur ökologische Regression.

28 Siehe z. B.: King, 1997, S. 37–38 und Sora Institut: SORA Wählerstromanalysen, online: <http://www.sora.at/themen/wahlverhalten/wahlanalysen/waehlerstromanalysen.html> (zuletzt aufgerufen am: 13.04.2016, Archiv: <http://archive.is/NP6Jf>).

Gleichung (7) beinhaltet im Gegensatz zu Gleichung (3) und (5) einen Fehlerterm ϵ_i , da im Allgemeinen aufgrund der im Modell angenommenen gleichen Parameter in jeder Gebietseinheit nicht länger eine fehlerfreie Darstellung des Wahlergebnisses der zweiten Wahl möglich ist. Dies gilt auch bei Gültigkeit der Modellspezifikation und einer Herleitung des Modells ausgehend von der Individualebene²⁹. Die Modellgleichung veranschaulicht aber auch, dass eine Schätzung bei Gültigkeit des Modells mittels linearer Regression möglich ist und die Regressionskoeffizienten in diesem Fall direkt als Übergangswahrscheinlichkeiten interpretiert werden können.

Goodman selber war sich der Bedeutung seiner Modellannahme im Klaren. Es sagte zur Übertragung von Ergebnissen aus ökologischen Analysen auf die Individualebene:

„[...] *in very special circumstances* the study of regression between ecological variables may be used to make inference concerning the behaviour of individuals“³⁰

Goodman zufolge ist eine Übertragung von Ergebnissen, eine Ebenen-überschreitende Inferenz, also nur dann möglich, wenn besondere Bedingungen erfüllt sind. Bei seinem Modell handelt es sich um die Gültigkeit der Konstanzannahme. Er geht jedoch nicht von einer allgemeinen Anwendbarkeit des Modells aus. Trotz dieser Einschränkung stellte Goodmans Grundmodell lange Zeit das Standardmodell für die ökologische Inferenz dar³¹.

II.3.1.2. Erweiterungen des Modells

Das grundlegende Problem bei der ökologischen Regression nach Goodman ist, dass die Konstanzannahme im Allgemeinen als nicht besonders plausibel betrachtet werden kann. Dies gilt erst recht für die Schätzung von Wählerwanderungen. Es ist sehr unwahrscheinlich, dass in allen betrachteten Gebieten die Übergangsraten zwischen zwei Parteien absolut identisch sind. Jedoch stellt genau dies die Grundannahme im Modell und für die Schätzung dar. Aus diesem Grund wurde in der Vergangenheit die einfache Goodman Regression erweitert, um eine größere Flexibilität zu erreichen.

Die Erweiterungen versuchen vor allem durch eine abweichende Modellierung die Konstanzannahme, siehe (6), für die Übergangswahrscheinlichkeiten zu lockern:

$$P_{CSU,2013,i} = \beta'_{CSU2009,CSU2013,i} * P_{CSU,2009,i} + \dots + \beta'_{NW2009,CSU2013,i} * P_{NW,2009,i} + \epsilon_i \quad (8)$$

mit z. B.:

$$\beta'_{r,c,i} = \gamma_{r,c,0} + \gamma_{r,c,1} * x_i \quad (9)$$

29 Vgl. Achen / Shively, 1995, S. 46–47.

30 Goodman, 1953, S. 663, zitiert nach Cho / Manski, 2008, S. 554.

31 Vgl. King, 1997, S. 37–38.

Anstelle von global konstanten Übergangswahrscheinlichkeiten $\beta_{r,c}$ wird bei dieser Art von Erweiterung von gebietsspezifischen Übergangswahrscheinlichkeiten $\beta'_{r,c,i}$ ausgegangen. Es wird weiterhin angenommen, dass die Variabilität zwischen den Gebieten allein durch eine, wie in Gleichung (9), oder mehrere Kovariablen erklärt wird. Gegeben einer bestimmten Kovariablenausprägung wird jedoch erneut von konstanten Übergangswahrscheinlichkeiten ausgegangen. Diese Art der Erweiterung wurde bereits 1959 von Goodman vorgeschlagen.³² Es ist möglich, getrennt gerechnete Modelle für als homogen angenommene Teile der betrachteten Region als Spezialfall dieser Erweiterung aufzufassen.

Falls angenommen wird, dass die Übergangswahrscheinlichkeiten vom Parteienergebnis der ersten Wahl abhängen und dies im Modell entsprechend berücksichtigt werden soll, besteht beim erweiterten Modell ein Identifikationsproblem. Es ist nicht mehr möglich, alle Parameter eindeutig zu bestimmen. Um trotzdem eine Identifizierbarkeit zu erreichen, müssen zusätzliche Annahmen getroffen werden.³³ Dieser Fall ist insofern relevant, da bei der ökologischen Inferenz in politikwissenschaftlichen Anwendungsfällen eine entsprechende Abhängigkeit durchaus als relevantes Problem angesehen werden kann. Es ist nicht vollkommen unplausibel anzunehmen, dass bestimmte Übergangswahrscheinlichkeiten, wie z. B. die Loyalitätsrate einer Partei, von Wahlergebnis der Vorwahl abhängig sind. Allgemein wird bei Vorliegen einer solchen Struktur von Aggregationsbias gesprochen.³⁴

Andere Erweiterungen der einfachen Regression nach Goodman beschäftigen sich mit der Struktur der Fehler und erlauben z. B. zufällige Fehler bei den einzelnen Übergangswahrscheinlichkeiten³⁵ oder nutzen Bayesianische Schätzverfahren, um bei der Schätzung der Übergangswahrscheinlichkeiten die Restriktionen für diese berücksichtigen zu können³⁶.

II.3.1.3. Probleme bei der praktischen Anwendung

Bei der ökologischen Regression nach Goodman zeigen sich bei der praktischen Anwendung mehrere Problemfelder. Die zugrundeliegenden Modellannahmen sind allein mit aggregierten Daten nicht überprüfbar, jedoch hängt die Interpretation der Parameter

32 Vgl. Achen / Shively, 1995, S. 39–44.

33 Vgl. Achen / Shively, 1995, S. 118–123.

34 Vgl. Ambühl, Mathias: Methoden zur Rekonstruktion von Wählerströmen aus Aggregatdaten. Bundesamt für Statistik, Neuenburg, 2003, S. 18.

35 Vgl. z. B. Achen / Shively, 1995, S. 53–55.

36 Vgl. Lau, Olivia / Moore, Ryan .T. / Kellermann, Micheal: eiPack: RxC ecological inference and higher-dimension data management. R News, 7(2), 2007, S. 45.

stark von diesen ab³⁷. Weiterhin ist es nicht möglich, die Güte der Modellschätzungen, unabhängig davon, ob es sich um eine einfache oder erweiterte Regression nach Goodman handelt, mit den üblicherweise nur vorliegenden Informationen zu überprüfen. Als Entscheidungskriterium kann nur die – angenommene – Plausibilität der Schätzung oder externes Wissen herangezogen werden.³⁸

Neben diesen inhaltlichen Problemen besteht zusätzlich noch die methodische Problematik, dass bei der Regression nach Goodman unklar ist, ob eine gewichtete oder ungewichtete Regression zu bevorzugen ist. Öfters wird die Bevölkerung N_i der betrachteten Gebietseinheiten zur Gewichtung bei der Regression verwendet. Dies geschieht in der Hoffnung, die Heterogenität der interessierenden Größen richtig zu gewichten.³⁹ Die Nutzung von N_i zur Gewichtung wird aber auch kritisch und als nicht gerechtfertigt angesehen.⁴⁰

Für den Vergleich der Schätzung mit der Regression nach Goodman mit anderen Modellen und aus inhaltlicher Sicht bei der Interpretation der Ergebnisse stellen aber Schätzwerte außerhalb der logischen Grenzen für Übergangswahrscheinlichkeiten das größte Problem dar. Es ist nicht unüblich, dass bei der Regression nach Goodman für $\beta_{r,c}$ Werte geschätzt werden, die kleiner null oder größer eins sind⁴¹. Übergangswahrscheinlichkeiten können aber nur im Intervall $[0,1]$ liegen. Ambühl führt als eine mögliche Ursache für solche unplausiblen Schätzwerte das Vorhandensein von Aggregationsbias an⁴².

Zur Korrektur von Schätzwerten außerhalb der logischen Grenzen schlägt Ambühl die Nutzung eines als „iterative proportional fitting“ (IPF) - Algorithmus bekannten Verfahrens vor.⁴³ Bei diesem Algorithmus handelt es sich eine Methode, die entwickelt wurde, um die inneren Zellen einer Kreuztabelle an bekannte äußere Ränder anzupassen. Erstmals vorgeschlagen wurde es 1940 von Deming und Stephan⁴⁴.

Vor der Anwendung des IPF-Algorithmus zur Adjustierung der geschätzten Wählerwanderung muss jedoch in einem ersten Schritt eine Anpassung der inneren Zellen

37 Vgl. z. B. Freedman, 1999, S. 3–5.

38 Es sei darauf hingewiesen, dass diese Nicht-Überprüfbarkeit für alle Modelle der ökologischen Inferenz gilt. Bei der Regression nach Goodman geht es vor allem um die Frage, welches Modell zu bevorzugen sei, wenn einfache und – ein oder mehrere - erweiterte Modelle geschätzt werden.

39 Vgl. Achen / Shively, 1995, S. 57–61.

40 Vgl. King, 1997, S. 61.

41 Vgl. z. B. Achen / Shively, 1995, S. 15 und S. 74–76, sowie King, 1997, S. 15–16.

42 Vgl. Ambühl, 2003, S. 18.

43 Vgl. Ambühl, 2003, S. 53–55. Das Verfahren wird auch von Infratest dimap bei ihrem Wählerwanderungsmodell genutzt, vgl. dazu Hilmer / Kunert, 2005, S. 144.

44 Vgl. Deming, W. Edwards / Stephan, Frederick F.: On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *Ann. Math. Stat.* 11(4), 1940, S. 439–442.

erfolgen, mit dem Ziel, dass nach der Anpassung alle Werte im Intervall $[0,1]$ liegen. Diese Anpassung kann auf mehrere Arten erfolgen, z. B.:

$$\beta_{r,c}^a := \max(0, \min(1, \beta_{r,c})), \quad \forall r \in \{1, \dots, R\}, c \in \{1, \dots, C\} \quad (10)$$

oder

$$\beta_{r,c}^a := \frac{\beta_{r,c} - \min(0, \beta_{r,1}, \dots, \beta_{r,C})}{\max(1, \beta_{r,1}, \dots, \beta_{r,C}) - \min(0, \beta_{r,1}, \dots, \beta_{r,C})}, \quad \forall r \in \{1, \dots, R\}, c \in \{1, \dots, C\} \quad (11)$$

$\beta_{r,c}^a$ steht dabei für die Übergangswahrscheinlichkeiten nach der Anpassung, während $\beta_{r,c}$ die mit der ökologischen Regression nach Goodman erhaltenen geschätzten Übergangswahrscheinlichkeiten beschreibt. Bei Gleichung (10) werden Werte außerhalb der logischen Grenzen auf den entsprechenden möglichen Randwert gesetzt, d. h. Werte kleiner null werden auf null und Werte größere eins werden auf eins gesetzt. Werte innerhalb der logischen Grenzen werden bei diesem Vorgehen nicht verändert. Dieser Ansatz wurde von Ambühl vorgeschlagen.⁴⁵ Gleichung (11) passt jeweils alle Werte einer Zeile an und bewahrt dabei den relativen Abstand zwischen den Schätzwerten.

Während der erste Anpassungsschritt garantiert, dass die angepassten Schätzwerte logisch als Übergangswahrscheinlichkeiten interpretiert werden können, kann ein zweiter Anpassungsschritt notwendig sein, um die in Gleichung (1) dargestellte Restriktion zu erfüllen:

$$\beta_{r,c}^{aa} := \frac{\beta_{r,c}^a}{\sum_{k=1}^C \beta_{r,k}^a}, \quad \forall r \in \{1, \dots, R\}, c \in \{1, \dots, C\} \quad (12)$$

Gleichung (12) führt dazu, dass die durch die Zeilenränder implizierten Restriktionen erfüllt sind. Dies ist jedoch anschließend nicht notwendigerweise für die Spaltenränder der Fall.

Zur abschließenden Anpassung wird der bereits zuvor erwähnte IPF-Algorithmus verwendet. Bei diesem werden abwechselnd die Spalten und Zeilen mit einem geeigneten Faktor solange angepasst, bis die Abweichung der aus den inneren Zellen berechneten Ränder von den beobachteten Rändern, den Wahlergebnissen, einen vorgegebenen Grenzwert unterschreitet. Während die Verwendung des IPF-Algorithmus garantiert, dass die abschließend so erhaltene Wählerwanderungstabelle zum einen nur valide Übergangswahrscheinlichkeiten beinhaltet und zum anderen die Ränder auch dem

⁴⁵ Vgl. Ambühl, 2003, S. 53.

beobachteten Wahlergebnis entsprechen, sind die statistischen Eigenschaften der Schätzwerte nach der hier dargestellten Anpassung unklar. Weiterhin unterscheiden sich die geschätzten Wählerwanderungstabellen nach der in Gleichung (10) dargestellten Anpassung im Allgemeinen von den Wählerwanderungstabellen nach der in Gleichung (11) dargestellten Anpassung. Es handelt sich daher nicht um eine triviale Entscheidung, welches Verfahren zur Anpassung genutzt wird.

II.3.2. Thomsens Probit/Logit Modell

Søren R. Thomsen schlug 1987 in seinem Buch „Danish Elections 1920-79: A Logit Approach to Ecological Analysis and Inference“⁴⁶ ein Modell zur Schätzung von Wählerwanderungen in Mehrparteiensystemen vor. Das von ihm vorgestellte Modell wurde vor allem für die Schätzung von Wählerwanderungen entwickelt und kann nicht als allgemeines Modell für die ökologische Inferenz angesehen werden⁴⁷.

II.3.2.1. Modell für den 2x2 Fall

Thomsens Ausgangspunkt für die Entwicklung seines Modells ist die kritische Bewertung vor allem eines Modellaspekts, dass die regressionsbasierenden Verfahren der ökologischen Inferenz gemeinsam haben. Diese Modelle betrachten weitestgehend das Ergebnis der ersten Wahl als deterministisch, während die zweite Wahl als ein Ergebnis der ersten Wahl angesehen wird. Thomsen bevorzugt einen „symmetrischen“ Ansatz, bei dem beide Wahlergebnisse als von den gleichen latenten Variablen⁴⁸ erklärt betrachtet werden.⁴⁹ Während diese latenten Variablen aber bei der Motivation und Herleitung des Modells eine wichtige Rolle spielen, ist ihre Kenntnis und Spezifikation für den eigentlichen Schätzprozess bei der ökologischen Inferenz nicht notwendig.⁵⁰

Das Modell wird von Thomsen für eine relativ homogene Region und für den 2x2 Fall hergeleitet. Das bedeutet, dass bei beiden Wahlen jeweils nur zwei Wahlalternativen betrachtet werden können.⁵¹ Die Wahlergebnisse für beide Parteien P1 und P2 werden im Folgenden als E1 und E2 bezeichnet, ein Suffix bezeichnet die entsprechende Wahl, 1 steht dabei für die erste und 2 für die zweite Wahl. Thomsen leitet zuerst die Verteilung der Wahrscheinlichkeit in beiden Wahlen für eine bestimmte Parteienkombination zu stimmen, hier die gleiche Partei P1, her:

$$P(P1_1, P1_2) = \Phi_2(\Phi^{-1}(E1_1), \Phi^{-1}(E1_2), \rho_t) \quad (13)$$

Es handelt sich um eine bivariate Normalverteilung, von der zwei der drei Parameter

46 Thomsen, Søren R.: Danish elections 1920–79: A Logit Approach to Ecological Analysis and Inference, Aarhus, 1987.

47 Die Motivation für das Modell ist die Schätzung der Wählermobilität, d. h. von Wählerwanderungen. Thomsen betrachtet aber z. B. eine Nutzung seines Modells zur Schätzung von gruppenspezifischen Wahlverhaltens durchaus für möglich, vgl. Thomsen, 1987, S. 76.

48 Mit latenten Variablen meint Thomsen Eigenschaften der Wähler, wie z. B. deren politische Einstellung, die ihre Wahlentscheidung maßgeblich beeinflussen, vgl. z. B. Thomsen, 1987, S. 49.

49 Vgl. Thomsen, 1987, S. 46–47.

50 Vgl. Thomsen, 1987, S. 61–64.

51 Es muss sich dabei bei Thomsens Modell um zwei echte Wahlalternativen handeln, eine Betrachtung von 2x2 Tabellen in der Form: [Partei1, $\overline{\text{Partei1}}$][Partei2, $\overline{\text{Partei2}}$] ist nicht möglich, vgl. Thomsen, 1987, S. 52 und S. 65–71.

bekannt sind. Die beiden bekannten Parameter sind die Probit transformierten Wahlergebnisse der betrachteten Partei(en) für beide Wahlen. Der unbekannte Parameter ist die Korrelation ρ_t . Bei Kenntnis der Korrelation ρ_t wäre $P(PI_1, PI_2)$ bestimmbar und somit eine der inneren Zellen der Wählerwanderungstabelle bekannt. Aufgrund der bekannten Ränder würden sich anschließend alle anderen Werte der Wählerwanderungstabelle direkt berechnen lassen.⁵²

Thomsens Schlüsselannahme ist die Schätzbarkeit der tetrachorische Korrelation ρ_t über die ökologische Probit-Korrelation ρ_e .

$$\hat{\rho}_t = \rho_e = \text{corr}(\Phi^{-1}(EI_1^e), \Phi^{-1}(EI_2^e)) \quad (14)$$

EI_1^e bzw. EI_2^e stehen in Gleichung (14) für die beobachteten Wahlergebnisse für die Partei P1 in den betrachteten Gebietseinheiten. Mit der von Thomsen angenommenen Gleichsetzung von $\rho_t = \rho_e$ reichen die bekannten aggregierten Daten aus, um die individuelle Wählerwanderung mit Hilfe von Verteilung (13) zu schätzen. Anstelle der direkten Nutzung der bivariaten Normalverteilung in (13) nutzt Thomsen jedoch aufgrund von Rechenzeitproblemen eine Approximation der tetrachorischen Korrelation zur Berechnung der interessierenden Größe $P(PI_1, PI_2)$.⁵³

Thomsens Ausgangspunkt für die Herleitung seines Modells stellen zwei binäre Modelle, eines für jede Wahl, mit latenten Variablen auf Individualebene dar. Thomsen geht dabei davon aus, dass beide Wahlentscheidungen, gegeben die latenten Variablen, als unabhängig angesehen werden können. Aus mathematischen Gründen nutzt er dabei Probit und keine Logit-Modelle. Um die Verteilung (13) herleiten zu können und die Gleichsetzung in (14) zu erreichen, sind nach Thomsen zusätzliche Annahmen über die Verteilung der latenten Variablen innerhalb der betrachteten Gebietseinheiten und zwischen den Gebietseinheiten, sowie der Varianzstruktur dieser Verteilungen notwendig. Thomsen tätigt dabei die Annahme, dass das Verhältnis der Varianz zwischen den Individuen innerhalb der Gebiete und die Varianz zwischen den Gebieten für alle latenten Variablen gleich ist. Diese Annahme wird von ihm „Isomorphismus“ genannt und stellt seine Strukturannahme zwischen den Gebieten dar, die notwendig für die Schätzbarkeit ist. Für die Gleichsetzung in (14) nimmt Thomsen weiterhin an, dass die Variabilität zwischen den Individuen in den Gebieten deutlich größer ist, als die Variabilität zwischen den betrachteten Gebieten.⁵⁴

52 Vgl. Thomsen, 1987, S. 56–58.

53 Vgl. Thomsen, 1987, S. 61–64.

54 Vgl. Thomsen, 1987, S. 49–55 und S. 62–63.

Die von Thomsen getätigten Annahmen werden von Ambühl kritisch gesehen. Vor allem sieht Ambühl in einem von Thomsen getätigten Grenzübergang, bei der das Verhältnis der Variabilität zwischen den Individuen in den Gebieten und die Variabilität zwischen den betrachteten Gebieten als gegen unendlich laufend angesehen wird⁵⁵ und welcher die Gleichsetzung in (14) erlaubt, in Anbetracht der Implikationen bei anderen von Thomsen genutzten Verteilungen ein Problem.⁵⁶ Dieser Schwachpunkt bei der Herleitung muss jedoch nicht implizieren, dass das Modell nicht geeignet ist, da alleine die approximative Gültigkeit von (13) und (14) für die Anwendbarkeit ausreichen.

Won-ho Park betrachte z. B. in seiner 2008 eingereichten Dissertation ebenfalls Thomsens Modell und widmet sich dabei auch der Gleichsetzung in (13). Das von ihm aufgestellte Kriterium ist dabei allgemeiner als das von Thomsen. So sei es möglich, $\rho_t = \rho_e$ anzunehmen, wenn das Varianzverhältnis von aggregierter Varianz, d. h. zwischen den Gebietseinheiten, und totaler Varianz, d. h. für alle Individuen in der Region, für die von den latenten Variablen abhängigen und von den latenten Variablen nicht abhängigen Varianzkomponenten identisch sei. Eine Möglichkeit diese Bedingung zu erfüllen stellt Thomsens Annahme von einer homogenen Region dar, eine andere erlaubt Variabilität innerhalb der Region, wobei jedoch die Variabilität zwischen den Gebieten größer sein muss als in den Gebieten.⁵⁷

II.3.2.2. Erweiterung für den RxC Fall

Thomsens Modell ist nicht direkt im allgemeinen RxC Fall anwendbar, da die theoretische Herleitung des Modells nur für den 2x2 Fall gilt und das 2x2 Modell nur bei der Betrachtung von jeweils zwei echten Wahlalternativen Gültigkeit hat. Um trotzdem eine Schätzung von Wählerwanderungen im RxC Fall zu ermöglichen, schlägt Thomsen einen Algorithmus vor, der die Schätzung durch wiederholtes Anwenden seines 2x2 Modells erlauben soll.

Ausgangspunkt des vorgeschlagenen Algorithmus sind die beobachteten Wahlergebnisse in den einzelnen Gebieten innerhalb der betrachteten Region, siehe z. B. Tabelle 1. Thomsens Modell wäre nur direkt anwendbar, wenn Teiltabellen in der Form von [CSU2009, SPD2009]x[CSU2013, SPD2013] betrachtet werden könnten. Für diese Teiltabellen sind jedoch die Ränder unbekannt. Im ersten Schritt nutzt daher Thomsen

55 Vgl. Thomsen, 1987, S. 63.

56 Vgl. Ambühl, 2003, S. 43–44.

57 Park, Won-ho: Ecological Inference and Aggregate Analysis of Elections, The University of Michigan, Dissertation, 2008, S. 34–38.

Betrachtungen der Form $[\text{CSU2009}, \overline{\text{CSU2009}}] \times [\text{CSU2013}, \overline{\text{CSU2013}}]$, also kollabierte Tabellen, um Schätzer für die inneren Zellen jeder Gebietseinheit zu erhalten. Die so berechneten Schätzwerte werden für den weiteren Algorithmus dann als Daten angesehen.⁵⁸

Mit den so geschätzten Werten ist es möglich, sich Ränder von 2×2 Tabellen der Form $[\text{CSU2009}, \text{SPD2009}] \times [\text{CSU2013}, \text{SPD2013}]$ zu berechnen und das Schätzverfahren anzuwenden. Um nicht alle möglichen 2×2 Tabellen betrachten zu müssen, legt Thomsen für jede Wahl eine Referenzpartei fest. Dieser Schritt reduziert die Zahl der zu betrachteten 2×2 Tabellen auf $(R-1) \times (C-1)$ Tabellen. Thomsen empfiehlt die Verwendung einer neutralen Partei, wie z. B. der Nichtwähler, als Referenzpartei. Mit den berechneten Werten für die einzelnen Gebiete wird anschließend die ökologische Logit⁵⁹-Korrelation ρ_e für jede Tabelle der Form $[\text{Partei1}, \text{Referenzpartei1}] \times [\text{Partei2}, \text{Referenzpartei2}]$ bestimmt. Für den Algorithmus sind zusätzlich noch erste Schätzwerte für die inneren Zellen für die Gesamtregion notwendig, diese werden über eine Betrachtung von Tabellen der Form $[\text{Partei1}, \text{Referenzpartei1}] \times [\text{Partei2}, \text{Referenzpartei2}]$ ⁶⁰ ermittelt. Der anschließende Teil des Algorithmus wird dann solange iteriert, bis das vorgegebene Konvergenzkriterium erreicht wird.⁶¹

Im ersten Schritt werden die inneren Zellen für die Parteien angepasst, die nicht als Referenzparteien festgelegt worden sind. Dazu wird die zuvor berechnete ökologische Korrelation zwischen den jeweils betrachteten Parteien und den Referenzparteien genutzt. Da die so erhaltenen bisherigen Schätzwerte für die inneren Zellen nicht notwendigerweise den Randrestriktionen entsprechen, erfolgt in einem weiteren Schritt eine Anpassung der Zellen für die Referenzparteien. Dabei wird das Verhältnis zwischen beobachtetem Rand und aktuellem Rand aus den Schätzungen zur Adjustierung der Werte für die Referenzparteien verwendet. Entsprechend wird im letzten Schritt⁶² die Zelle $[\text{Referenzpartei1}, \text{Referenzpartei2}]$ angepasst, wobei jedoch die gesamte Tabelle als Referenz genutzt wird.⁶³

Für die Schätzung hat Thomsen die Softwarelösung Ecol⁶⁴ entwickelt. Neben den

58 Vgl. Thomsen, 1987, S. 73.

59 Thomsen nutzt im Algorithmus die Logit-Transformation anstelle der Probit-Transformation und berechnet eine gewichtete Korrelation, vgl. Thomsen, 1987, S. 60.

60 Der Schätzalgorithmus nutzt nur Zellen, bei denen entweder eine Wählerwanderung von oder zur jeweils festgelegten Referenzpartei ermittelt wird, vgl. Thomsen, 1987, S. 75.

61 Vgl. Thomsen, 1987, S. 73–74 und S. 76.

62 Dieser letzte Schritt wird nicht explizit von Thomsen angegeben, ist jedoch notwendig, um bei der eigenen Implementierung des Algorithmus Schätzwerte zu erlangen, die identisch mit denen von Thomsens Softwarelösung sind.

63 Vgl. Thomsen, 1987, S. 74–76.

64 Thomsen, Søren R. / Frandsen, Annie G. / Kristmar, Thomas / Lauritsen, Per / Sørensen, Michael B.: Ecol: Version 3, Aarhus University, 1995.

gerade erläuterten Algorithmus verfügt das Programm über ein zweites Schätzverfahren, bei dem die Schätzung nicht nur für die gesamte Region, sondern für jedes betrachtete Gebiet erfolgt. Dabei wird angenommen, dass die ökologische Logit-Korrelation auch der tetrachorischen Korrelation der Individuen in jeder Gebietseinheit entspricht. Der Vorteil des alternativen Schätzverfahrens ist, dass so auch gebietsspezifische Schätzwerte ermittelt werden können. Zusätzlich besteht die Option, für Stichprobeneffekte zu korrigieren.

II.3.2.3. Parks Erweiterungen

Neben der Fragestellung, wann die Gleichsetzung von $\rho_t = \rho_e$ möglich ist, stellt Park auch Erweiterungen von Thomsens Probit/Logit Modell vor. Die erste Erweiterung betrifft den Schätzalgorithmus im RxC Fall und kann eher als eine Vereinfachung der Methodik als echte Erweiterung verstanden werden.

Anstelle der iterativen Schätzung, die Thomsen vorgeschlagen hat, schlägt Park die Nutzung von IPF⁶⁵ vor. Die erste Schätzung der inneren Zellen erfolgt wie bei Thomsen durch Betrachtungen von Tabellen der Form [CSU2009, $\overline{\text{CSU2009}}$]x[CSU2013, $\overline{\text{CSU2013}}$]. Von der geschätzten Wählerwanderung wird nur die direkt geschätzte Zelle, im Beispiel [CSU2009, CSU2013] genutzt. Anschließend wird der IPF-Algorithmus verwendet, um die so geschätzten inneren Zellen an die bekannten Ränder anzupassen. Parks Vorschlag hat zum einen den Vorteil, dass das Schätzverfahren deutlich vereinfacht und somit auch klarer ist, zum anderen ist es bei diesem Vorgehen nicht mehr notwendig, eine Referenzpartei festzulegen.⁶⁶

Die zweite Erweiterung von Thomsens Modell durch Park versucht die Berücksichtigung von (binären) Kovariablen zu ermöglichen. Die Grundidee des Modells ist die Annahme von getrennten Verteilungen der Form (13) für jede Ausprägung der Kovariablen. Die gesamte Verteilung ist dann die Mischung dieser einzelnen Verteilungen. Dies führt dazu, dass anstelle eines Zusammenhangs zwischen den latenten Variablen und dem Wahlergebnis, für jede Ausprägung der Kovariablen ein eigener Zusammenhang angenommen wird. Für Parks Ansatz ist es notwendig, explizit die Parameter dieser Probit/Logit Modelle der Wahlentscheidung zu schätzen. Da die Ausprägung der dafür notwendigen Variablen auf Individualebene nicht bekannt ist, werden diese im Modell durch die Probit-transformierten Parteienanteilen auf Gebietsebene ersetzt, welche Park zufolge als Gebietsmittel der latenten Variablen aufgefasst werden können.⁶⁷

65 Siehe Kapitel II.3.1.3. für eine kurze Vorstellung des IPF-Algorithmus.

66 Vgl. Park, 2008, S. 55–58.

67 Vgl. Park, 2008, S. 81–90.

Park widmet sich in seiner Arbeit auch einer weiteren Einschränkung von Thomsens Modell. Es ist nicht möglich, die Schätzunsicherheit anzugeben, da Thomsens Modell nur Punktschätzer für die interessierenden Größen liefert. Park stellt zwar einen Ansatz für die Bestimmung von Konfidenzintervalle für die Schätzung mit Thomsens Modell vor, jedoch sollen diese in der praktischen Anwendung eine zu geringe Überdeckung besitzen. Die so ermittelten Konfidenzintervalle dürften daher im Allgemeinen nicht nutzbar sein.⁶⁸

68 Vgl. Park, 2008, S. 63–65.

II.3.3. Iterative / Adaptive Modellen

Das Modell von Thomsen gehört im RxC Fall bereits zu einer größeren Klasse von Modellen für die ökologische Inferenz. Allgemein wird in der ökologischen Inferenz zwischen Modellen für den 2x2 Fall und den allgemeineren RxC Fall unterschieden. Nicht jedes Modell, das für den 2x2 Fall entwickelt wurde, lässt sich ohne weiteres auch auf den RxC Fall erweitern. Daher gibt es Bemühungen, die Schätzung von RxC Tabellen mittels wiederholter Anwendung von 2x2 Verfahren zu ermöglichen.

II.3.3.1. Allgemeine Beschreibung

Modelle, die eine RxC Tabelle durch wiederholte Anwendung eines 2x2 Verfahrens versuchen zu schätzen, werden als iterative bzw. adaptive Modelle bezeichnet. Die Bezeichnung Modell für diese Schätzverfahren kann jedoch etwas irreführend sein, da es sich öfters um Algorithmen handelt, die eine Anwendungslogik für Methoden der ökologischen Inferenz für den 2x2 Fall darstellen und dabei nicht notwendigerweise von einer bestimmten Methode abhängig sind. Zumindest muss zwischen der Methode, der Anwendungslogik – dem Algorithmus, und dem zur Schätzung herangezogenen Verfahren unterschieden werden. Thomsens Probit/Logit Modell für den RxC Fall stellt insofern ein Sonderfall dar, da der iterative Algorithmus explizit auf sein Modell für den 2x2 Fall abgestimmt ist.

Grundsätzlich stellt sich bei der mehrfachen Anwendung von 2x2 Verfahren die Frage, ob diese parallel⁶⁹ genutzt werden sollen, d. h. mehrere Subtabellen werden unabhängig voneinander geschätzt, oder ob diese mehrstufig, d. h. bei späteren Betrachtungen werden die vorherigen Ergebnisse mit verwendet bzw. sie basieren auf diesen, angewendet werden sollen. Beide Prinzipien haben klar erkennbare Nachteile. Bei paralleler Anwendung stellt sich die Frage, was bei Schätzwerten geschehen soll, bei denen die geschätzten Ränder nicht den beobachteten Rändern entsprechen oder was geschehen soll, wenn Restriktion, wie z. B. Gleichung (1), nicht eingehalten werden. Bei einer mehrstufigen Anwendung hängen offensichtlich alle folgenden Schätzungen von den zuvor erhaltenen Ergebnissen ab.

Ein Beispiel für ein mehrstufiges Verfahren stellt Kings Vorschlag für die Anwendung seines 1997 vorgestellten 2x2 Verfahrens für 2x3 Tabellen dar. Das Verfahren soll kurz anhand von Tabelle 5, bei der es sich um eine kollabierte Version von Tabelle 1 handelt, dargestellt werden. Im ersten Schritt werden zwei Spalten, z. B. SPD 2013 und Andere

⁶⁹ Der Begriff „parallel“ wird genutzt, da theoretisch bei diesem Modelltyp eine parallele Berechnung möglich wäre.

	CSU 2013	SPD 2013	Andere 2013	Wahl 2009
CSU 2009				23,56%
Andere 2009				76,44%
Wahl 2013	26,91%	17,02%	56,07%	100,00%

Tabelle 5: Darstellung der Datensituation bei der ökologischen Inferenz im 2x3 Fall, die hellgrau markierten Ränder sind bekannt, die weißen inneren Zellen unbekannt. Es handelt sich um eine kollabierte Version von Tabelle 1, bei der die Nichtwähler und alle Parteien außer der CSU (2009) bzw. der CSU und SPD (2013) zur Kategorie „Andere“ zusammengefasst worden sind. (Datenquelle: Stadt München)

2013 zur Kategorie $\overline{\text{CSU 2013}}$ zusammengefasst. Die Ränder der resultierenden Tabelle lassen sich direkt aus den bekannten Rändern der 2x3 Tabelle berechnen. Die Berechnung der Wählerwanderung für die Tabelle $[\text{CSU 2009}, \overline{\text{Andere 2009}}] \times [\text{CSU 2013}, \overline{\text{CSU 2013}}]$ ist mit dem 2x2 Verfahren möglich. Diese erste Schätzung ermöglicht es, Schätzwerte für die zuvor unbekannt Ränder der Tabelle $[\text{CSU 2009}, \overline{\text{Andere 2009}}] \times [\text{SPD 2013}, \overline{\text{Andere 2013}}]$ anzugeben. Eine Anwendung der 2x2 Methodik bei dieser Tabelle wäre daher jetzt möglich und alle inneren Werte von Tabelle 5 können so bestimmt werden. Bei seinen Algorithmus berücksichtigt King weiterhin noch die Unsicherheit der ersten Schätzung.⁷⁰

Neben dem mehrstufigen Verfahren schlägt King auch eine paralleles Verfahren für RxC Tabellen vor. Dabei werden jeweils kollabierte 2x2 Tabellen der Form $[\text{CSU 2009}, \overline{\text{CSU 2009}}] \times [\text{CSU 2013}, \overline{\text{CSU 2013}}]$ betrachtet. Wenn die Modellschätzungen als valide angenommen werden, müssen auf diese Art und Weise mindestens $(R-1) \times (C-1)$ innere Zellen bestimmt werden, um eine vollständige Wählerwanderung berechnen zu können. Bei Tabelle 5 würde z. B. die Kenntnis von zwei inneren Zellen ausreichen. Alle anderen Werte würden sich direkt aus den dann bekannten inneren Zellen und den Rändern ergeben. Eine Möglichkeit wäre es z. B. die Tabellen $[\text{CSU 2009}, \overline{\text{CSU 2009}}] \times [\text{CSU 2013}, \overline{\text{CSU 2013}}]$ und $[\text{CSU 2009}, \overline{\text{CSU 2013}}] \times [\text{SPD 2013}, \overline{\text{SPD 2013}}]$ zu betrachten. King betrachtet ein entsprechendes Vorgehen als nicht für alle RxC Tabellen geeignet, da es möglich ist, zu nicht plausiblen Ergebnissen zu gelangen.⁷¹

Karen E. Ferree diskutiert die bei der mehrfachen Anwendung von 2x2 Tabellen auftretenden Problem. Dazu gehört z. B., dass ein Kollabieren von Alternativen die Verteilungsannahme der 2x2 Modell verletzen oder auch zu Aggregationsbias führen kann. Zur Umgehung von einigen dieser Probleme schlägt sie z. B. die Nutzung von Kovariablen oder eine bewusste Auswahl der Reihenfolge der betrachteten kollabierten 2x2 Tabellen vor.⁷²

70 Vgl. King, 1997, S. 150–153.

71 Vgl. King, 1997, S. 264–267.

72 Vgl. Ferree, Karen E.: Iterative Approaches to R x C Ecological Inference Problems. Where They Can Go Wrong and One Quick Fix. *Political Analysis*, 12(2), 2004, S. 145–147 und S. 153–158.

II.3.3.2. Modell von Andreadis und Chadjipadelis

Ein Modell aus dieser Klasse, das speziell für die Schätzung von Wählerwanderungen entwickelt wurde, stammt von Ioannis Andreadis und Theodore Chadjipadelis.⁷³ Sie greifen dabei auf die von King vorgeschlagenen Prinzipien für die mehrfache Anwendung von 2x2 Methoden zurück. Das von ihnen vorgeschlagene Modell gehört dabei zur Kategorie der mehrstufigen Verfahren, da die Ergebnisse der ersten Schätzungen bei den folgenden Schätzungen berücksichtigt werden. Die Besonderheit ihres Modells stellt dabei die Wahl der jeweils betrachteten kollabierten Tabelle dar. Von den Autoren selber wird das Modell als rekursives Verfahren bezeichnet.

Eine Motivation hinter dem Modell von Andreadis und Chadjipadelis ist eine typische Eigenschaft in Mehrparteiensystemen. Eine kleine Zahl von Parteien kann jeweils relativ viele Stimmen auf sich vereinigen, während viele kleinere Parteien vergleichsweise wenig Stimmen erlangen können. Dies impliziert, dass in den Rändern der RxC Tabelle unterschiedlich viel Information über die betrachteten Parteien vorhanden ist. Während es durchaus möglich ist, dass für größere Parteien Restriktionen aus den Rändern resultieren, sind bei kleineren Parteien keine bzw. fast keine Restriktionen zu erwarten. Aus diesem Grund seien die Schätzungen für kleinere Parteien im Allgemeinen fehlerbehafteter. Wenn jedoch bereits die Wählerwanderung eines relevanten Anteils der Wahlberechtigten bekannt sein würde, wäre es möglich, dass eine kleine Partei, bezogen auf die Wähler mit noch unbekannter Wählerwanderung, als „große“ Partei angesehen werden kann.⁷⁴

Das vorgeschlagene Modell betrachtet in jeder Iteration kollabierte Tabellen der Form [CSU 2009, CSU 2009]x[CSU 2013, CSU 2013], d. h. für jede Wahl jeweils eine Partei gegen den Rest. Dabei wird immer das aktuelle „Best Pair“ betrachtet.

$$Best\ Pair = \underset{(r,c)}{argmax}(C_{r,c}^*) \quad (15)$$

mit

$$C_{r,c}^* = (P_{CSU,2009}, P_{SPD,2009}, \dots, P_{NW,2009}) * corr(E2009, E2013) * (P_{CSU,2013}, P_{SPD,2013}, \dots, P_{NW,2013})^T \quad (16)$$

Die Einträge der Form $P_{CSU,2009}$ stehen dabei für die relativen Wahlergebnisse der Parteien bei den beiden betrachteten Wahlen in der Region, der jeweilige Vektor beschreibt daher die relative Stärke der Parteien bei der Wahl. $corr(E2009, E2013)$ steht die RxC

73 Andreadis / Chadjipadelis, 2009.

74 Vgl. Andreadis / Chadjipadelis, 2009, S. 206–207.

Korrelationsmatrize der Wahlergebnisse der Parteien auf Gebietsebene. Das Best Pair beschreibt daher die Parteienkombination, bei der die größte ökologische Korrelation der Wahlergebnisse zwischen den beiden Wahlen, gewichtet durch die Stärke der Partei(en), gemessen durch die Wahlergebnisse, beobachtet wird. Die Gewichtung mit dem Wahlergebnis wird von Andreadis und Chadjipadelis durchgeführt, um tendenziell die Wählerwanderungen von größeren Parteien zuerst zu schätzen.⁷⁵

Der rekursive Algorithmus betrachtet in der ersten Iteration das mit (15) und (16) bestimmte Best Pair. Für diese Parteienkombination wird mit der über alle anderen Zeilen und Spalten kollabierten Tabelle und den gewählten 2x2 Verfahren die Wählerwanderung geschätzt. Die so geschätzten Wähler für die betrachtete Parteienkombination werden für die folgenden Iterationen des Algorithmus als bereits erklärt betrachtet. Dies bedeutet, dass der Schätzwert für die innere Zelle der Wählerwanderungstabelle gespeichert wird und zur Vorbereitung der nächsten Iteration die gebietsspezifischen Schätzungen von den Rändern der Gebietseinheiten, den beobachteten aggregierten Wahlergebnissen, abgezogen werden. Anschließend wird mit (15) und (16) und den neuen Rändern das nächste Best Pair bestimmt. Der Algorithmus iteriert solange, bis das Abbruchkriterium, ein vorgegebener Anteil an erklärten Wahlberechtigten, erreicht wird.⁷⁶

Eine Implikation des Algorithmus ist, dass in den ersten Iterationen, aufgrund der Definition des Best Pairs, tendenziell Wählerwanderungen von großen Parteien, vor allem deren Loyalitätsraten, betrachtet werden. Für diese Parteien liegen bereits Anfangs relativ viel Information in den Rändern vor. In jeder folgenden Iteration wird der Anteil der noch zu erklärenden Wahlberechtigten niedriger, und es ist möglich anzunehmen, dass die Ränder in den späteren Iterationen, die den in den vorhergehenden Iterationen noch nicht erklärten Wahlberechtigten entsprechen, mehr Informationen über die kleineren Parteien beinhalten. Weiterhin ist es möglich und auch beabsichtigt, dass der Algorithmus mehrmals eine bestimmte Parteienkombination zum Best Pair bestimmt und so die Schätzung der Wählerwanderung verbessert.⁷⁷

Das von Andreadis und Chadjipadelis genutzte 2x2 Verfahren⁷⁸ stammt von Bernard Grofman und Samuel Merrill⁷⁹ und ist in der Lage gebietsspezifische Schätzwerte für die

75 Vgl. Andreadis / Chadjipadelis, 2009, S. 210–211.

76 Vgl. Andreadis / Chadjipadelis, 2009, S. 212–213.

77 Vgl. Andreadis / Chadjipadelis, 2009, S. 208 und S. 210–211.

78 Eine kurze Erläuterung des Modellansatzes von Grofman und Merrill befindet sich im Anhang, siehe Anhang X.1..

79 Grofman, Bernard / Merrill, Samuel: Ecological Regression and Ecological Inference. In: King, Gary / Rosen, Ori / Tanner, Martin A. (Editoren): Ecological Inference: New Methodological Strategies, Cambridge, 2004, S. 123–143.

Wählerwanderung zu bestimmen. Diese Eigenschaft ist für den von Andreadis und Chadjipadelis vorgeschlagenen Algorithmus insofern notwendig, da bei jeder Iteration eine Anpassung der noch nicht erklärten Wähler erfolgt, und für diese Anpassung valide gebietsspezifische Schätzwerte zur Verfügung stehen müssen. Jedoch ist der Algorithmus nicht auf das genutzten 2x2 Verfahren festgelegt. Die Autoren schreiben dazu:

„For this task in our method any of the known methods for the solution of the 2×2 ecological inference problem could be used.“⁸⁰

Die Autoren zogen zwei 2x2 Verfahren für ihre Anwendung in die engere Auswahl⁸¹. Da beide Verfahren jedoch ihre Stärke aus der Information in den Rändern ziehen sollen, ist es Andreadis und Chadjipadelis zufolge sinnvoller, die verfügbaren Ressourcen in einer Erhöhung der nutzbaren Gebietseinheiten, als in komplizierte, aufwändige und intensive Verfahren zur Modellierung der Daten zu investieren.⁸²

Thomas Kellermann schlug 2011 eine Veränderung des von Andreadis und Chadjipadelis vorgeschlagenen Algorithmus vor. Während diese in ihrer Argumentation von einer tendenziell hohen Loyalität, im Vergleich zu den Wählerwanderungen zwischen den Parteien, ausgehen,⁸³ garantiert ihr Algorithmus nicht, dass die Loyalitätsraten in den ersten Iterationen auch alle geschätzt werden. Dies kann Kellermann zufolge unter Umständen zu Problemen und unrealistischen Schätzwerten führen. Kellermann greift diesen Aspekt daher auf und schlägt vor, dass in den ersten Iterationen erzwungen wird, die Loyalitätsraten einmal zu schätzen. Die Auswahl der Reihenfolge erfolgt dabei weiterhin über die Best Pair Methode. Andere Zellen sollen erst anschließend betrachtet werden.⁸⁴ Kellermanns Erweiterung ist insofern für diese Arbeit relevant, dass sein Modell von einigen Städten in Deutschland zur Schätzung der Wählerwanderung genutzt wird.

80 Andreadis / Chadjipadelis, 2009. S. 212.

81 Anmerkung: Es handelt sich um Kings 1997 vorgestelltes Modell für den 2x2 Fall und Grofman und Merrills Modell aus dem Jahr 2004.

82 Vgl. Andreadis / Chadjipadelis, 2009. S. 212.

83 Vgl. Andreadis / Chadjipadelis, 2009. S. 207.

84 Vgl. Kellermann, Thomas: Vom Wahlergebnis zur Wählerwanderung: Welche Wähler wechselten wie ihre Entscheidung, Stadtforschung und Statistik, 2011(1), 2011, S. 38.

II.3.4. Hierarchische Modelle

Ein möglicher Ansatz die notwendigen Annahmen bei der ökologischen Inferenz flexibler zu formulieren, ist ein mehrstufiger, hierarchischer Modellaufbau. Bei einem solchen Modell kann eine Strukturannahme für jedes einzelne Gebiet aufgestellt werden. Um eine Schätzbarkeit zu gewährleisten, muss jedoch, wie bei den zuvor vorgestellten Modellen, zusätzlich auch noch eine Annahme getroffen, die die einzelnen Gebietseinheiten verbindet. Auch wenn Modelle, die hierarchisch aufgefasst werden können, bereits zuvor vorgeschlagen worden sind⁸⁵, machte vor allem Kings Methode für die ökologische Inferenz⁸⁶ diesen Ansatz populär. Zur Zeit handelt es sich um den Modellansatz der ökologischen Inferenz, der die meisten Weiterentwicklungen in den letzten Jahren gesehen hat.

II.3.4.1. Allgemeine Beschreibung

Kings 1997 vorgeschlagenes Modell kann zu einem als Random Effekt Modell aufgefasst werden.⁸⁷ Es ist jedoch auch möglich, dass Modell als hierarchisches Modell zu interpretieren. Bei diesem für den 2x2 Fall entwickelten Modell wird auf der Ebene der betrachteten Gebietseinheiten die deterministische Information in den Rändern genutzt. Für jedes Gebiet werden nur Übergangswahrscheinlichkeiten als valide angenommen, die diese Restriktionen erfüllen. Die Restriktion lässt sich dabei als Gerade im Raum der Übergangswahrscheinlichkeiten auffassen. Als gemeinsame Verteilung der Übergangswahrscheinlichkeiten der Gebiete wird eine bivariate Normalverteilung angenommen, die auf das Einheitsquadrat $[0,1] \times [0,1]$ trunziert wird. Diese Verteilungsannahme verbindet die einzelnen Gebietseinheiten und erlaubt so die Schätzung mit dem Modell.⁸⁸

1999 wurde von King, Rosen und Tanner das Binomial-Beta Modell für die ökologische Inferenz im 2x2 Fall vorgestellt. Es handelt sich dabei explizit um ein hierarchisches Bayesianisches Modell, dass zur Schätzung auf Markov-Ketten-Monte-Carlo (MCMC) Methoden zugreift.⁸⁹ Grundlegend basieren hierarchische Modellen auf folgenden Prinzip:

85 Vgl. z. B. Brown, Philip J. / Payne, Clive D.: Aggregate Data, Ecological Regression, and Voting Transitions. *Journal of the American Statistical Association*, Volumen 81, Nummer 394, 1986, S. 452, aber auch Erweiterungen der Goodman Regression, die eine Abhängigkeit der Übergangswahrscheinlichkeiten von Kovariablen erlauben, können entsprechend aufgefasst werden.

86 King, 1997.

87 Siehe z. B. Gschwend, Thomas: Ökologische Inferenz. In: Behnke, Joachim / Gschwend, Thomas / Schindler, Delia / Schnapp, Kai-Uwe (Editoren): *Methoden der Politikwissenschaft: neuere qualitative und quantitative Analyseverfahren.*, Baden-Baden, 2006, S. 230.

88 Vgl. King, 1997, S. 93–99.

89 King / Rosen / Tanner, 1999, S. 61–90.

$$Y_i | \beta \sim p_1(y | \beta_i) \quad (17)$$

$$\beta_i \sim p_2(\beta | \theta) \quad (18)$$

Im Allgemeinen soll eine abhängige Variable Y durch einen Parameter(vektor) θ erklärt werden. Es ist jedoch nicht immer ohne weiteres möglich, eine entsprechende Verteilung direkt und geschlossen zu spezifizieren. Von hierarchische Modellen wird gesprochen, wenn die Verteilung von $Y|\theta$ nicht geschlossen, sondern mehrstufig aufgebaut wird. Gleichung (17) und (18) zeigen eine zweistufige Konstruktion. Auf der ersten Ebene (17) hängt die Verteilung von Y_i von einem Parameter(vektor) β_i ab. Der Parameter(vektor) β_i wiederum folgt einer Verteilung, die von einem Parameter(vektor) θ , der unabhängig von i ist, abhängt (18). Mit beiden Verteilungen lässt sich wiederum die interessierende Verteilung von $Y|\theta$ bestimmen.⁹⁰

$$p(y | \theta) = \int_{-\infty}^{\infty} p_1(y | \beta) p_2(\beta | \theta) d\beta = \int_{-\infty}^{\infty} p(y, \beta | \theta) d\beta \quad (19)$$

Die hierarchische Konstruktion des Zusammenhangs hat den Vorteil, dass Zusammenhänge modelliert werden können, für die keine geschlossene Verteilung spezifiziert werden kann. Es ist nur notwendig, die einzelnen Ebenen des Modells spezifizieren zu können. Es ist daher möglich, mittels hierarchischen Modellen kompliziertere Zusammenhänge abzubilden. Die Berechnung des Integrals in (19) stellt jedoch eine zusätzliche Schwierigkeit dar, da im Allgemeinen eine Berechnung des Integrals und eine direkte Spezifikation der interessierenden Verteilung nicht möglich ist. Zur Schätzung wird daher oft auf Bayesianische Schätzverfahren, vor allem MCMC-Methoden zurückgegriffen.⁹¹

Beim hierarchischen Binomial-Beta Modell wird für die einzelnen Gebietseinheiten jeweils eine Binomialverteilung angenommen. Dabei werden jedoch nicht die inneren Zellen, sondern eine Randsumme, bei Wählerwanderungen entspricht dies dem Wahlergebnis der zweiten Wahl, betrachtet. Der Parameter der Binomialverteilungen ist jeweils eine gewichtete Summe der Übergangswahrscheinlichkeiten und dem Rand, der hier dem relativen Wahlergebnis bei der ersten Wahl entspricht. Als gebietsübergreifende Verteilung der Übergangswahrscheinlichkeiten werden bei diesem Modell zwei unabhängige Beta-Verteilungen angenommen. Die Hyperprioris der Parameter der Beta-

90 Vgl. King / Rosen / Tanner, 1999, S. 69 und Robert, Christian: The Bayesian choice: from decision-theoretic foundations to computational implementation. Springer Science & Business Media, 2007, S. 460–462.

91 Vgl. King / Rosen / Tanner, 1999, S. 69–70 und Robert, 2007, S. 301–303.

Verteilungen sind Exponentialverteilungen.⁹² Wakefield analysierte 2004 dieses Modell und kritisierte die Wahl der Verteilung auf der ersten Ebene. Anstelle der Binomialverteilung für die Randsumme schlägt Wakefield eine Faltung von zwei Binomialverteilungen vor, die die inneren Zellen repräsentieren. Weiterhin untersuchte er auch die Auswirkung der Wahl der gemeinsamen Verteilung der Übergangswahrscheinlichkeiten und der Hyperprioris.⁹³

Gemeinsam haben hierarchische Bayesianische Modelle, die mittels MCMC-Methoden geschätzt werden, dass bestimmte Größen vorab festgelegt werden müssen. Die am wenigsten kritische Größe stellt dabei die Anzahl der genutzten Ziehungen, d. h. die für die Schätzung der Kennwerte der Posteriori-Verteilung verwendeten Werte der Markov-Kette, dar. Es muss nur ein Wert festgelegt werden, der eine adäquate Bestimmung der interessierenden Größe erlaubt. Kritischer sind die Festlegung des Burn-Ins und des Thinnings. Der Burn-In legt fest, wie viele Werte am Anfang der Kette verworfen werden, da die Kette als noch nicht gegen die Posteriori-Verteilung konvergiert angenommen wird. Thinning definiert die Anzahl der Beobachtungen, die zwischen zwei Ziehungen aus der Kette verworfen werden.⁹⁴ Eine sinnvolle Festlegung der beiden Werte ist notwendig, da z. B. ein zu niedriger Burn-In zu einer verzerrten Schätzung führen kann. Am Anfang ist die Kette noch von den Startwerten abhängig und nicht gegen die Posteriori-Verteilung konvergiert.⁹⁵ Das Thinning dient dazu, relativ unabhängige Ziehungen aus der Posteriori zu erhalten. Benachbarte Beobachtungen einer Markov-Kette sind im Allgemeinen korreliert und können nicht als unabhängig angesehen werden.⁹⁶ Wenn jedoch genügend Werte einer Kette zwischen zwei Ziehungen verworfen werden, können die Ziehungen als unabhängig betrachtet werden. Alternativ wird Thinning auch dazu verwendet, um längere Ketten berechnen zu können. Indem nicht jeder Wert gespeichert wird, kann der Speicherbedarf in kontrollierbaren Rahmen gehalten werden.

Eine weitere festzulegende „Größe“ sind die Priori-Verteilungen, und dabei vor allem die Wahl der Parameter der Hyperprioris bei hierarchischen Bayesianischen Modellen. Dies stellt eine Stärke der Bayesianischen Modellierung dar, da spezifisches Vorwissen und die Unsicherheit über dieses Vorwissen explizit im Modell berücksichtigt werden kann. Im

92 Vgl. King / Rosen / Tanner, 1999, S. 71–72. Auf eine formale Darstellung wird verzichtet, da das Binomial-Beta Modell und das nachfolgend vorgestellte Multinomial-Dirichlet Modell eine vergleichbare Modellstruktur besitzen.

93 Vgl. Wakefield, 2004, S. 390–391, S. 400–404 und S.406–417.

94 Vgl. Robert, 2007, S. 314.

95 Vgl. Robert, 2007, S. 302.

96 Vgl. Robert, 2007, S. 302.

Allgemeinen werden jedoch meistens nicht-informative Priori-Verteilungen gewählt. Nichtsdestoweniger stellt die Wahl der Parameter der Hyperprioris eine Entscheidung dar, die bewusst getroffen werden sollte, da sie Einfluss auf die Modellschätzung haben kann. In Kapitel II.3.4.4. wird kurz auf die Wahl geeigneter Parameter bei der Schätzung von Wählerwanderungen eingegangen.

II.3.4.2. Multinomial–Dirichlet Modell

Das 2001 vorgestellte Multinomial-Dirichlet Modell von Rosen, Jiang, King und Tanner⁹⁷ stellt eine direkte Erweiterung des Binomial-Beta Modells für den RxC Fall dar. Bei einer Anwendung des Multinomial-Dirichlet Modells im 2x2 Fall entspricht dieses daher auch dem Binomial-Beta Modell. Es handelt sich bei dem Modell ebenfalls um ein hierarchisches Bayesianisches Modell,⁹⁸ bei dem die Schätzung mittels MCMC-Methoden erfolgt. Der Vorteil des Multinomial-Dirichlet Modells gegenüber dem Binomial-Beta Modell ist, dass ersteres direkt für Wählerwanderungen im RxC Fall genutzt werden kann, während für eine Anwendung des Binomial-Beta Modell ein wie in Kapitel II.3.3. beschriebener Algorithmus genutzt werden müsste.

Auf der ersten Ebene des Modells, den einzelnen Gebietseinheiten, wird für das Ergebnis der zweiten Wahl eine Multinomialverteilung angenommen.⁹⁹

$$(N_{CSU,2013,i}, N_{SPD,2013,i}, \dots, N_{NW,2013,i}) \sim MNL(\theta_{CSU,i}, \theta_{SPD,i}, \dots, \theta_{NW,i}, N_i) \quad (20)$$

$N_{Partei,2013,i}$ entspricht der Anzahl der Wähler für die jeweilige Partei bei der zweiten Wahl im Gebiet i , N_i der Anzahl der Wähler im Gebiet i . Er wird wie beim Binomial-Beta Modell angenommen, dass die gebietsspezifischen Parameter $\theta_{Partei,i}$ als eine gewichtete Summe der Übergangswahrscheinlichkeiten dargestellt werden können. Es gilt z. B. für $\theta_{CSU,i}$:¹⁰⁰

$$\begin{aligned} \theta_{CSU,i} &= \beta_{CSU2009,CSU2013,i} * P_{CSU,2009,i} + \dots + \beta_{NW2009,CSU2013,i} * P_{NW,2009,i} \\ &= \sum_{r=1}^R \beta_{r,CSU2013,i} * P_{r,2009,i} \end{aligned} \quad (21)$$

Der in Gleichung (21) dargestellte Zusammenhang zwischen den Parametern $\theta_{Partei,i}$ und den Übergangswahrscheinlichkeiten und dem Wahlergebnis der ersten Wahl folgt dem in Gleichung (3) dargestellten deterministischen Zusammenhang und der Grundstruktur der ökologischen Regression nach Goodman (7). Im Gegensatz zur Regression nach Goodman

97 Vgl. Rosen / Jiang / King / Tanner, 2001.

98 Es sei darauf hingewiesen, dass von den Autoren auch eine Schätzmethode basierende auf einer nichtlinearen kleinsten Quadrate Methode vorgeschlagen wird, vgl. Rosen / Jiang / King / Tanner, 2001, S. 142–147.

99 Vgl. Rosen / Jiang / King / Tanner, 2001, S. 137.

100 Vgl. Rosen / Jiang / King / Tanner, 2001, S. 137.

wird aber von gebietsspezifischen Übergangswahrscheinlichkeiten ausgegangen und es wird der Erwartungswert des zweiten Wahlergebnisses betrachtet.

Um eine Schätzbarkeit des Modells zu ermöglichen, ist auf der zweiten Ebene des Modells eine Annahme über die Verteilung der Übergangswahrscheinlichkeiten zwischen den Gebieten notwendig. Folgend der Überlegungen beim Binomial-Beta Modell wird für jede Zeile eine gemeinsame Dirichlet-Verteilung angenommen. Zusätzlich wird angenommen, dass a priori die einzelnen Zeilen unabhängig voneinander sind. Es gilt daher z. B. für die erste Zeile im Gebiet i :¹⁰¹

$$\begin{aligned} & (\beta_{CSU2009, CSU2013, i}, \beta_{CSU2009, SPD2013, i}, \dots, \beta_{CSU2009, NW2013, i}) \\ & \sim \text{Dirichlet}(\alpha_{CSU2009, CSU2013}, \alpha_{CSU2009, SPD2013}, \dots, \alpha_{CSU2009, NW2013}) \end{aligned} \quad (22)$$

Die Nutzung der Dirichlet-Verteilung stellt sicher, dass die Übergangswahrscheinlichkeiten in jeder Zeile in der Summe 1 ergeben, die in Gleichung (1) aufgestellte Restriktion ist daher automatisch erfüllt. Im Gegensatz zu den gebietsspezifischen Übergangswahrscheinlichkeiten handelt es sich bei den Parametern der Dirichlet-Verteilung um globale Parameter, d. h., sie gelten für jedes Gebiet. Aufgrund möglicherweise vorhandener Unterschiede in der Zahl der Wahlberechtigten in den Gebietseinheiten, sind die Erwartungswerte der Dirichlet-Verteilung jedoch im Allgemeinen nicht mit den globalen Übergangswahrscheinlichkeiten identisch. Als Hyperpriori für die Parameter der Dirichlet-Verteilung wird von Rosen, Jiang, King und Tanner jeweils eine Exponentialverteilungen angenommen. Z. B. gilt für den Parameter $\alpha_{CSU2009, CSU2013}$:¹⁰²

$$\alpha_{CSU2009, CSU2013} \sim \exp(\lambda_{CSU2009, CSU2013}) \quad (23)$$

Während es möglich ist, wie in Gleichung (23), für jeden Parameter der Dirichlet-Verteilungen eine eigene Hyperpriori zu definieren, wird im Allgemeinen für alle Parameter eine identische Verteilung gewählt. Es gilt daher:

$$\alpha_{r,c} \sim \exp(\lambda), \quad \forall r \in \{1, \dots, R\}, c \in \{1, \dots, C\} \quad (24)$$

Als Wert für λ wird von den Autoren 0,5 vorgeschlagen und im Beispiel genutzt.¹⁰³

Eine Softwareimplementierung des Multinomial-Dirichlet Modells für R ¹⁰⁴ stellt das R-

101 Vgl. Rosen / Jiang / King / Tanner, 2001, S. 137. Im Artikel wird ein Modell mit Kovariablen vorgestellt, die hier präsentierte Form geht von einem Modell ohne Kovariablen aus und ist entsprechend angepasst worden, vgl. daher auch: King / Rosen / Tanner, 1999, S. 72.

102 Vgl. Rosen / Jiang / King / Tanner, 2001, S. 138 und King / Rosen / Tanner, 1999, S. 72.

103 Vgl. Rosen / Jiang / King / Tanner, 2001, S. 139 und King / Rosen / Tanner, 1999, S. 72.

104 R Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2013, URL: <http://www.R-project.org/>.

Paket `eiPack`¹⁰⁵ dar. Abweichend vom 2001 vorgeschlagenen Modell erlaubt diese Implementierung eine etwas allgemeinere Spezifikation der Hyperprioris. Anstelle einer für alle Parameter genutzten Exponentialverteilung wird eine Gamma-Verteilung für die Hyperprioris genutzt. Die Exponentialverteilung kann als Spezialfall der Gamma-Verteilung angesehen werden.¹⁰⁶

Die Softwareimplementierung von Thomas Schlesinger im R-Paket `eiwild`¹⁰⁷ erlaubt eine noch flexiblere Definition der Hyperprioris. Anstelle einer identischen Verteilung für alle Parameter kann für jeden Parameter eine eigene Hyperpriori spezifiziert und somit zellspezifisches Vorwissen angegeben werden.

II.3.4.3. Multinomial-(log-)Normal Modell

Ein weiteres hierarchisches Bayesianisches Modell, das für den RxC Fall vorgeschlagen wurde, ist das Multinomial-Normal Modell von Greiner und Quinn. Das im Jahr 2009 vorgestellte Modell weicht bei den Verteilungsannahmen auf beiden Ebenen deutlich vom Multinomial-Dirichlet Modell ab. Die Schätzung des Modells erfolgt jedoch wie beim Multinomial-Dirichlet Modell auch mittels MCMC.¹⁰⁸

Beim Multinomial-Normal Modell werden die inneren Zellen der Wählerwanderungstabelle direkt als interessierende Größe¹⁰⁹ betrachtet und modelliert. Anstelle einer Verteilung für den Rand, für das beobachtete Wahlergebnis, wird für jede Zeile der Wählerwanderungstabelle eine Multinomialverteilung angenommen. Die einzelnen Multinomialverteilungen für jede Zeile werden dabei a priori als unabhängig betrachtet. Es gilt daher z. B. für die erste Zeile:¹¹⁰

$$\begin{aligned} & (N_{CSU\ 2009,CSU\ 2013,i}, \dots, N_{CSU\ 2009,NW\ 2013,i}) \\ & \sim MNL(\beta_{CSU\ 2009,CSU\ 2013,i}, \dots, \beta_{CSU\ 2009,NW\ 2013,i}, N_{CSU,2009,i}) \end{aligned} \quad (25)$$

$N_{CSU\ 2009,CSU\ 2013,i}$ steht dabei für die Wähler, die im Gebiet i in beiden Wahlen für die CSU gestimmt haben, $N_{CSU,2009,i}$ entspricht der Zahl der Wähler für die CSU im Gebiet i im Jahr 2009. Im Gegensatz zum Multinomial-Dirichlet Modell, bei dem die linke Seite von

105 Lau / Moore / Kellermann, 2007 und Lau, Olivia / Moore, Ryan T. / Kellermann, Michael: `eiPack`: Ecological Inference and Higher-Dimension Data Management, R package version 0.1-7, 2012. <http://CRAN.R-project.org/package=eiPack>.

106 Vgl. Lau / Moore / Kellermann, 2007, S. 46.

107 Schlesinger, Thomas: Kombination von Aggregat- und Individualdaten bei der Analyse von RxC-Tafeln. Neue Implementierungen in R, Masterarbeit, 2013 und Schlesinger, Thomas: `eiwild`: Ecological Inference with individual and aggregate data. R package version 0.6.7, 2014. <http://CRAN.R-project.org/package=eiwild>.

108 Vgl. Greiner / Quinn, 2009.

109 Andere interessierende Größen, wie z. B. die Übergangswahrscheinlichkeiten, werden dann aus den geschätzten inneren Zellen der Wählerwanderungstabelle berechnet, vgl. Greiner / Quinn, 2009, S. 68.

110 Vgl. Greiner / Quinn, 2009, S. 70.

Gleichung (20) bekannt ist, werden beim Multinomial-Normal Modell nicht beobachtete Größen betrachtet. Die einzig bekannte Größe in Gleichung (25) ist die Zahl der Wähler für die jeweils betrachtete Partei.

Die Modellierung in der Form von Gleichung (25) garantiert, dass die Zeilenrestriktionen erfüllt sind. Zur Wahrung der Spaltenrestriktion ist eine weitere Annahme notwendig. Die aus den einzelnen Multinomialverteilungen resultierende Spaltensumme muss den beobachteten Rand entsprechen. Es muss daher auch gelten, dass:¹¹¹

$$\sum_{r=1}^R N_{r,c,i} = N_{c,2013,i}, \quad \forall c \in \{1, \dots, C\} \quad (26)$$

$N_{c,2013,i}$ steht dabei für den beobachteten Rand, d. h. die Zahl der Wähler der Partei in der Spalte c bei der zweiten Wahl im Gebiet i . Im Modell werden diese Restriktionen durch Indikatorfunktionen abgebildet.¹¹²

Auf der zweiten Ebene des Modells wird ebenfalls ein Zusammenhang zwischen den Gebieten hergestellt, um die Schätzung zu ermöglichen. Dafür erfolgt in einem ersten Schritt eine multivariate Logit-Transformation der Übergangswahrscheinlichkeiten:¹¹³

$$\left(\omega_{CSU2009, CSU2013, i}, \dots, \omega_{CSU2009, Sonst2013, i} \right) = \left(\log \left(\frac{\beta_{CSU2009, CSU2013, i}}{\beta_{CSU2009, NW2013, i}} \right), \dots, \log \left(\frac{\beta_{CSU2009, Sonst2013, i}}{\beta_{CSU2009, NW2013, i}} \right) \right) \quad (27)$$

Für die Transformation muss zuerst eine Referenzpartei festgelegt werden, Greiner und Quinn nutzten dafür die „Enthalter“, was bei einer Wählerwanderung den Nichtwählern (NW) entspricht. Die Transformation sorgt auch dafür, dass die Dimension reduziert wird und dass anstelle von $R \times C$ Übergangswahrscheinlichkeiten $R \times (C-1)$ transformierte Parameter betrachtet werden.¹¹⁴

Für alle Logit-transformierten Übergangswahrscheinlichkeiten wird eine gemeinsame, d. h. eine $R \times (C-1)$ -variate, Normalverteilung angenommen. Diese Modellierung erlaubt im Gegensatz zu den zeilenweise unabhängig angenommenen Dirichlet-Verteilungen beim Multinomial-Dirichlet Modell a priori eine flexiblere Berücksichtigung von möglichen Korrelationen zwischen den einzelnen Übergangswahrscheinlichkeiten, innerhalb einer Zeile, aber auch zwischen den einzelnen Zeilen. Diese größere Flexibilität führt aber zu einer deutlich größeren Zahl von zu schätzenden Parametern auf dieser Ebene. Während

111 Vgl. Greiner / Quinn, 2009, S. 72.

112 Vgl. Greiner / Quinn, 2009, S. 72.

113 Vgl. Greiner / Quinn, 2009, S. 70.

114 Vgl. Greiner / Quinn, 2009, S. 70.

beim Multinomial-Dirichlet Modell $R \times C$ α -Parameter zu bestimmen sind, müssen beim Multinomial-Normal Modell insgesamt $R \times (C-1)$ Erwartungswerte und $0,5 \times R \times (C-1) \times (R \times (C-1) + 1)$ Varianzkomponenten bestimmt werden.

Als Hyperpriori für die Erwartungswerte der Normalverteilung nutzen Greiner und Quinn eine Normalverteilung, für die Varianz wird eine inverse Wishart-Verteilung angenommen. In der aktuellen Implementierung des Modells im R-Paket `RxCcolInf`¹¹⁵ ist es nur möglich, bei den beiden Hyperprioris eine Diagonalstruktur für die Varianz(komponenten) zu spezifizieren, d. h. es ist nicht möglich eine Korrelationsstruktur zwischen den Mittelwerten oder den Varianzkomponenten a priori anzunehmen. Es ist jedoch möglich, für die Erwartungswerte spezifische Priori-Parameter in der Hyperpriori anzugeben.

II.3.4.4. Vorwissen bei der Schätzung von Wählerwanderungen

Die Wahl der Priori-Verteilungen und der Parameter dieser Verteilungen stellt im Allgemeinen bei Bayesianischen Modellen eine relevante und wichtige Entscheidung dar. Oft wird bei der praktischen Anwendung auf konjugierte oder nicht-informative Verteilungen zurückgegriffen.¹¹⁶

Bei der Schätzung von Wählerwanderungen mit hierarchischen Bayesianischen Modellen stellt sich die Frage, welche Verteilungen möglichst wenig Vorwissen, also Annahmen über die Verteilung der zu schätzenden Größen, beinhalten. Nichtwissen oder Vorwissen wird durch die Parameter- und Verteilungswahl bei den Verteilungen auf der Ebene der Hyperprioris der Modelle festgelegt. Es greift aber zu kurz, nur die Verteilungen auf dieser Ebene zu betrachten;¹¹⁷ bei hierarchischen Modellen ist es auch notwendig, die Auswirkungen der Wahl auf allen Ebenen des Modells zu bewerten. Über die angenommene Struktur auf den anderen Ebenen beeinflusst die Wahl der Hyperprioris auch das Priori-Wissen auf diesen Ebenen.

Bei den beiden in Kapitel II.3.4.2. bzw. II.3.4.3. vorgestellten Modellen werden für die Ebenen der Modelle unterschiedliche Verteilungen angenommen. Jedoch haben beide Modelle gemeinsam, dass es nicht möglich ist, echte nicht-informative Priori-Parameter zu wählen. Dies liegt an der besonderen Struktur der Daten. Die Problematik bei der Wahl der Priori-Parameter soll kurz am Beispiel des Multinomial-Dirichlet Modells dargestellt

115 Greiner, D. James / Baines, Paul / Quinn, Kevin M.: `RxCcolInf`. R x C Ecological Inference With Optional Incorporation of Survey Information. R package version 0.1-3., 2013, <http://CRAN.R-project.org/package=RxCcolInf>.

116 Vgl. Robert, 2007, S. 105–106.

117 Vgl. Wakefield, 2004, S. 406–407.

werden. Tabelle 1 zeigt die Datensituation auf der ersten Ebene der Modelle.

Auf der Ebene von Tabelle 1 würde ein möglichst nicht-informatives Priori-Wissen keine der Zellen der Tabelle bevorzugen. Da beim Multinomial-Dirichlet Modell die Übergangswahrscheinlichkeiten betrachtet werden, müssten diese daher a priori als gleichwertig angesehen werden. Dies impliziert aber, dass der Priori-Erwartungswert der Übergangswahrscheinlichkeiten nicht 0,5 sein kann, sondern genau $1/C$ entspricht, d. h. von der Anzahl der Spalten der Tabelle abhängt.

Auf der Ebene der Übergangswahrscheinlichkeiten würde wiederum ein möglichst nicht-informative Priori-Wissen keinen möglichen Wert bevorzugen, d. h. alle möglichen Übergangswahrscheinlichkeiten im Intervall $[0,1]$ würden a priori als gleichwertig betrachtet. Dies würde eine Verteilung erfordern, die den Erwartungswert 0,5 hat und den gesamten möglichen Bereich $[0,1]$ abdeckt. Abgesehen vom Spezialfall $C = 2$ sind jedoch beide Bedingungen für nicht-informatives Vorwissen auf Ebene der Tabelle bzw. der Übergangswahrscheinlichkeiten nicht miteinander vereinbar.¹¹⁸

Abbildung 2 zeigt den Unterschied zwischen einer nicht-informativen, d. h. alle Spalten werden als gleichwertig angesehen, und einer informativen Wahl der Parameter für die Hyperprioris. Bei der ersten Zeile werden für alle Spalten die gleichen Parameter für die Gamma-Verteilung ($\text{shape} = 4$, $\text{rate} = 2$) angenommen. Diese gewählten Werte entsprechen den Standardwerten in den R-Paketen `eiPack`¹¹⁹ und `eiwild`¹²⁰. Für alle Spalten ist daher das Vorwissen identisch, ein Großteil der Masse der Verteilungen ist kleiner als 0,5 und der Erwartungswert der Verteilungen ist $1/7$. In der zweiten Zeile wird zellspezifisches Vorwissen für die erste Spalte angenommen. Dieses wird durch die gewählten Parameter ($\text{shape} = 8$, $\text{rate} = 0,5$) ausgedrückt. Für alle anderen Spalten werden die gleichen Parameter wie in der ersten Zeile verwendet. Die Priori-Wahrscheinlichkeit für größere Werte ist aufgrund der Änderung für die Übergangswahrscheinlichkeit in der ersten Spalte deutlich größer als bei der für alle Spalten identischen Wahl.¹²¹ Die höhere Priori-

118 Vgl. Klima, André / Schlesinger, Thomas / Küchenhoff, Helmut / Thurner Paul W.: Combining aggregate data and exit-polls for the estimation of voter transitions, unveröffentlichtes Manuskript, 2016, S. 9–11.

119 Lau / Moore / Kellermann, 2012.

120 Schlesinger, 2014.

121 Falls spezifisches Vorwissen spezifiziert werden soll, muss berücksichtigt werden, dass unterschiedliche Parameter zu einer scheinbar identischen Verteilung auf der in Abbildung 2 gezeigten Ebene für die Übergangswahrscheinlichkeiten führen können. Abbildung 2 zeigt die Dichten für alle Übergangswahrscheinlichkeiten und berücksichtigt daher die Streuung der Verteilungen zwischen den Gebieten (2. Ebene des Modells) und die Streuung der Verteilungen innerhalb der Gebiete (1. Ebene des Modells). Eine große Variabilität bei der Verteilung zwischen den Gebieten mit niedriger Variabilität in den Gebieten kann in dieser Darstellung eine ähnliche Dichte zeigen, wie eine niedrige Variabilität zwischen den Gebieten mit hoher Variabilität in den Gebieten. Jedoch unterscheidet sich das angenommene Vorwissen in beiden Fällen.

Priorwissen beim Multinomial-Dirichlet Modell:
Darstellung der Auswirkung der Wahl der Parameter der Hyperprioris

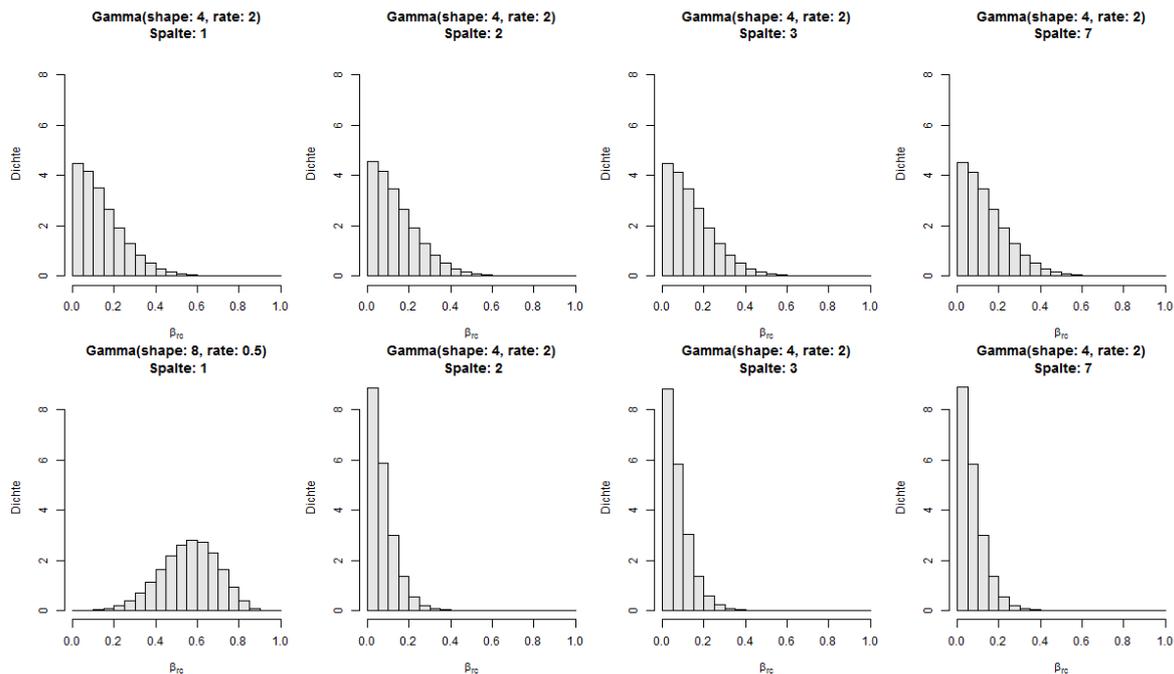


Abbildung 2: Darstellung der Auswirkung der Wahl der Parameter bei den Hyperprioris beim Multinomial-Dirichlet Modell; die erste Zeile zeigt eine für allen Spalten identische Wahl, die zweite Zeile zeigt für die erste Spalte eine abweichende Wahl, so dass für diesen Parameter eher höhere Übergangswahrscheinlichkeiten a priori erwartet werden. Die Spalten vier bis sechs werden nicht angezeigt, sind jedoch identisch mit Spalte drei. Die gewählten Parameter der Gamma-Verteilung werden jeweils in der Überschrift der Histogramme angegeben. Bestimmt mit der Funktion `prioriPlot` aus dem R-Paket `eiwild`.

Wahrscheinlichkeit für größere Werte für die erste Spalte impliziert aber auch, dass für alle anderen Spalten im Vergleich zu ersten Zeile eher niedrigere Übergangswahrscheinlichkeiten erwartet werden.

Bei der Schätzung von Wählerwanderungen kann im Allgemeinen davon ausgegangen werden, dass die Übergangswahrscheinlichkeiten für die großen und etablierten Parteien nicht der ersten Zeile von Abbildung 2 entsprechen und die zweite Zeile mit zellspezifischen Vorwissen näher an der Wahrheit liegt. Die Loyalitätsrate der Wähler dieser Parteien ist im Allgemeinen deutlich größer als die Übergangsraten zu anderen Parteien. Das eigentlich nicht-informative Vorwissen würde daher a priori etwas implizieren, was an sich als sehr unwahrscheinlich betrachtet werden muss. Inwiefern die Berücksichtigung von informativen zellspezifischen Vorwissen die Schätzgüte des Modells beeinflusst, wird in Kapitel III.4.1. mit untersucht.

II.4. Modellvergleich und Evaluierung ökologische Inferenz

Im Kapitel II.3. werden einige Modelle für die ökologischen Inferenz vorgestellt, die entweder explizit für die Schätzung von Wählerwanderungen in Mehrparteiensystemen entwickelt worden sind oder aufgrund ihrer Modellstruktur als dafür besonders geeignet angesehen werden können. In diesem Kapitel werden diese detailliert auf ihre Eignung untersucht. Dabei werden die zuvor vorgestellten Erweiterungen aber nur zum Teil mit berücksichtigt. Diese Einschränkung ist notwendig, um die Zahl der betrachteten Modelle auf einem handhabbaren Niveau zu halten.

II.4.1. Datengrundlage

Um einen ersten Eindruck von der Eignung der Modelle bei der Schätzung von Wählerwanderungen zu erlangen, wird mit den ausgewählten Modellen, siehe Tabelle 6, die Wählerwanderung zwischen den Bundestagswahlen 2005¹²² und 2009¹²³ in München geschätzt. Dabei wird auf die amtlichen Wahlergebnisse für beide Wahlen auf Wahlbezirksebene¹²⁴ zurückgegriffen.

Da bei der ökologischen Inferenz jedoch angenommen wird, dass bei beiden Wahlen die gleichen Gebiete betrachtet werden, sind die amtlichen Wahlergebnisse nicht direkt verwendbar. Ein veränderter Zuschnitt der Wahlbezirke, wie in München zwischen den Bundestagswahlen 2005 und 2009 geschehen, erfordert es, über die Zeit konstante Wahlbezirke zu bilden. Die Zusammenlegung erfolgte unter Zuhilfenahme der

Kurzname	Modell	Kapitel
Goodman	Ökologischen Regression nach Goodman	II.3.1.
Thomsen	Thomsens Probit/Logit Modell	II.3.2.
And & Cha	Rekursiver Algorithmus von Andreadis und Chadjipadelis	II.3.3.2.
Kellermann	Kellermanns Version von And & Cha	II.3.3.2.
M-D M	Multinomial-Dirichlet Modell	II.3.4.2.
RxCeol	Multinomial-Normal Modell	II.3.4.3.

Tabelle 6: Evaluierung der Methoden für die ökologische Inferenz, Übersicht über die weiter betrachteten Modell in Kapitel II.4.. Die erste Spalte zeigt den in Grafiken und zum Teil auch im Text genutzten Kurznamen des Modells, die zweite Spalte zeigt den Modellnamen und die dritte Spalte verweist auf das Kapitel, in dem das Modell vorgestellt wurde.

122 Statistisches Amt der Landeshauptstadt München. Amtliches Endergebnis der Bundestagswahl am 18.09.2005, Erst- und Zweitstimmenergebnisse der Landeshauptstadt München, 2005. <http://www.mstatistik-muenchen.de/themen/wahlen/wahlberichterstattung/btw2005/index.html> (zuletzt aufgerufen am: 08.06.2016).

123 Statistisches Amt der Landeshauptstadt München, 2009.

124 2005: 654 Urnenwahlbezirke; 2009: 664 Urnenwahlbezirke.

Wahlbezirkseinteilung¹²⁵. Wenn sich der Umriss eines Wahlbezirkes nur in relativ geringen Umfang zwischen 2005 und 2009 geändert hat, wird dieser als über die Zeit konstant betrachtet. Bei größeren Änderungen im Zuschnitt wurden solange Wahlbezirke zu einem Gebiet zusammengefasst, bis das Gesamtgebiet als über die Zeit konstanter Wahlbezirk betrachtet werden kann.

Dieser Datenaufbereitungsschritt reduziert die Zahl der Wahlbezirke auf 314. Die meisten dieser über die Zeit konstanten Wahlbezirke verfügen über eine ähnliche Größe, d. h. eine vergleichbare Zahl an Wahlberechtigten. Der Median liegt bei 1376,5 Wahlberechtigten, jedoch entstanden auch einige sehr große Wahlbezirke mit bis zu 56926 Wahlberechtigten. Die Briefwähler, die in gesonderten Briefwahlbezirken erfasst werden, sind ihren Urnenwahlbezirken zugeordnet worden. Dabei wird angenommen, dass es zwischen den Urnenwahlbezirken, die einen Briefwahlbezirk bilden, keine strukturellen Unterschiede bei den Briefwählern gibt. Veränderungen bei der Zahl der Wahlberechtigten zwischen den beiden Wahlen werden über Anpassungen bei den Nichtwählern aufgefangen. Eine etwas ausführlicher Betrachtung dieser Annahmen erfolgt in Kapitel IV.1..

Abbildung 3 zeigt die Wahlergebnisse bei beiden Wahlen für die gebildeten 314 über die Zeit konstanten Wahlbezirke. Betrachtet werden die Zweitstimmen für die Christlich-Soziale Union in Bayern e.V. (CSU), die Sozialdemokratische Partei Deutschlands (SPD), die Freie Demokratische Partei (FDP), Bündnis 90/Die Grünen (Grüne), die Partei Die Linke (DieLinke). Alle anderen Parteien werden zu Sonstigen (Sonst) zusammengefasst. Die Nichtwähler (NW) beschreiben hier die Wahlberechtigten, die nicht oder ungültig gewählt haben, sowie mögliche Veränderungen bei der Zahl der Wahlberechtigten im Laufe der Zeit. Bei der Bundestagswahl 2009 zeigt sich bei der CSU und der SPD in fast allen Wahlbezirken eine Verschlechterung des Wahlergebnisses. Die kleineren Parteien, FDP, Grüne und die Linke, sowie die Sonstigen, verbessern sich fast überall bei der zweiten Wahl. Die Zahl der Nichtwähler sinkt tendenziell,¹²⁶ wobei jedoch keine klare Struktur zwischen den Wahlbezirken zu erkennen ist.

125 Die Bildung der über die Zeit konstanten Wahlbezirke erfolgte im Rahmen eines statistischen Praktikums, vgl. Alkaya, Levent / Bracher, Johannes / Seibold, Heidi / Shao, Shuai: Analyse von Wahlen in München, Praktikumsbericht, 2012, S. 1–2.

126 Die Reduktion der Nichtwähler ist auf die durchgeführte Datenaufbereitung zurückzuführen. Veränderungen bei der Zahl der Wahlberechtigten (2005: 840166, 2009: 895101) in den Wahlbezirken werden über die Nichtwähler aufgefangen. Für viele Wahlbezirke gilt, dass es bei der zweiten Wahl mehr Wahlberechtigte als bei der ersten Wahl gab. Wenn nur die Wahlberechtigten der jeweiligen Wahl betrachtet werden, ist die Wahlbeteiligung 2009 niedriger (2005: 76,9%, 2009: 73,4%). Statistisches Amt der Landeshauptstadt München, 2005 und Statistisches Amt der Landeshauptstadt München, 2009.

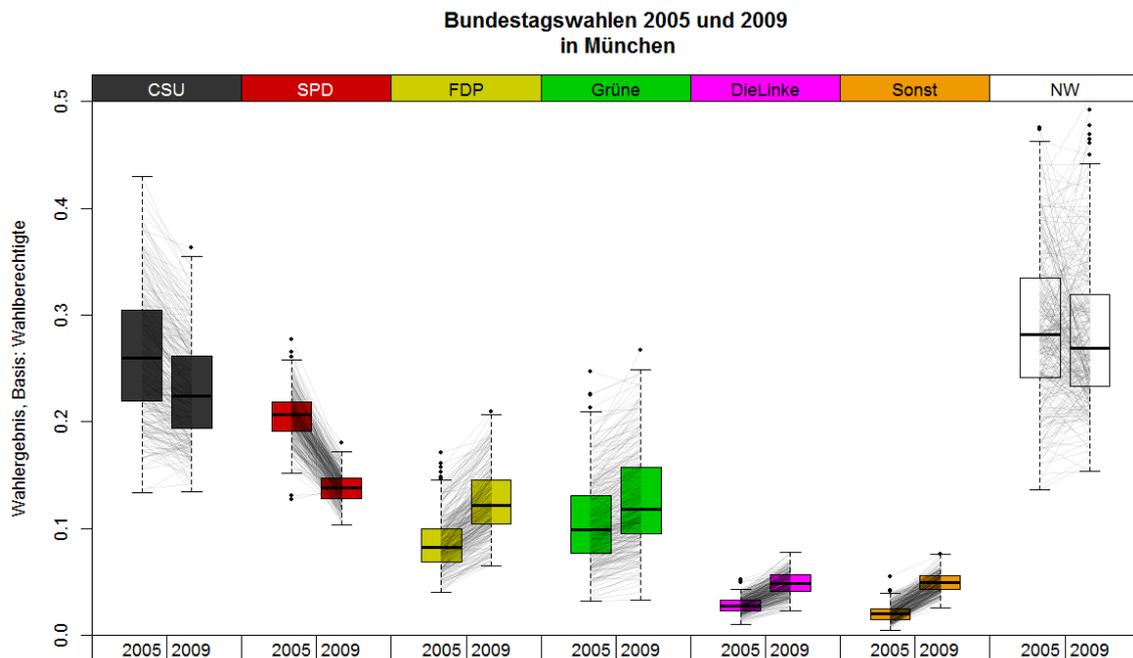


Abbildung 3: Wahlergebnisse (Zweitstimme) bei den Bundestagswahlen 2005 und 2009 in München für die gebildeten 314 über die Zeit konstanten Wahlbezirke. Briefwähler sind auf die Urnenwahlbezirke verteilt. Sonstige (Sonst) umfasst alle nicht explizit betrachteten Parteien, Nichtwähler (NW) beinhalten die Nichtwähler, ungültige Stimmen und Veränderungen bei der Zahl der Wahlberechtigten. Jede Linie repräsentiert bei jeder Partei jeweils einen Wahlbezirk. (Datenquelle: Stadt München)

II.4.2. Allgemeine Modelleigenschaften

Mit den so aufbereiteten Daten ist es möglich, die Wählerwanderung zwischen den Bundestagswahlen 2005 und 2009 zu schätzen. Für die Anwendung der meisten Modelle müssen jedoch noch Modellparameter festgelegt werden oder es gibt weitere Festlegungen, die möglicherweise die Schätzung beeinflussen. Um die Auswirkungen dieser Festlegungen zu überprüfen, werden die Parameterwahl und andere Festlegungen im ersten Schritt für jedes Modell gesondert evaluiert.

II.4.2.1. Maß für den Modellvergleich

Die potentielle Auswirkung dieser Entscheidungen wird dabei anhand der geschätzten Wählerwanderungen bewertet. Bei einer relativ unwichtigen Entscheidung sollte es keine großen Unterschiede zwischen den geschätzten Wählerwanderungen geben, relevante Entscheidungen würden wiederum maßgeblich Einfluss auf die Schätzung haben und die geschätzten Wählerwanderungen würden sich deutlich voneinander unterscheiden. Um den Unterschied zwischen den Schätzungen zu quantifizieren, wird in den folgenden Betrachtungen der absolute Abstand AD zwischen den geschätzten Wählerwanderungen genutzt:

$$AD(T_1, T_2) = \sum_{r=1}^R \sum_{c=1}^C |T_{1,r,c} - T_{2,r,c}| \quad (28)$$

T_1 und T_2 stehen in Gleichung (28) für die beiden betrachteten Wählerwanderungstabellen, $T_{1,r,c}$ und $T_{2,r,c}$ für die inneren Zellen der Tabellen. Dabei handelt es sich um Tabellen in der Form von Tabelle 1, bei der die inneren Zellen relative Häufigkeiten repräsentieren. Der AD stellt die Summe der zellweise berechneten absoluten Unterschiede zwischen beiden Tabellen dar. Er kann minimal null, bei identischen Tabellen, und maximal zwei annehmen. Die Hälfte des ADs gibt an, wie viel Prozent der Wahlberechtigten in einer Tabelle anders verteilt werden müssten, um zwei identische Tabellen zu erhalten, d. h. bei einem AD von 0,2 sind 10% der Wahlberechtigten in einer Tabelle im Vergleich zur anderen Tabelle anders verteilt. Das Maß beschreibt vor allem die Ähnlichkeit von zwei Tabellen, es ist jedoch nur bedingt geeignet, um die zellspezifische Ähnlichkeit der Schätzungen zu quantifizieren. Offensichtlich impliziert ein AD von 0,2 bei einer 2x2 Tabelle tendenziell eine größere zellspezifische Abweichung als in einer 7x7 Tabelle.¹²⁷

127 Ein ähnliches Maß wurde bereits von Thomsen vorgeschlagen, vgl. Thomsen, 1987, S. 85.

II.4.2.2. Modell-interne Evaluierung

Bei der ökologischen Regression nach Goodman, siehe Kapitel II.3.1., gibt es mehrere Festlegungen, die vor der Schätzung getroffen werden müssen. In Gleichung (5) wird das Modell mit relativen Häufigkeiten $P_{r,2009,i}$ dargestellt. Anstelle der relativen Häufigkeiten, dem Standardmodell, ist es jedoch auch möglich, die Zahl der Wahlberechtigten $N_{r,2009,i}$ zu nutzen. Zusätzlich stellt sich auch die Frage, ob eine gewichtete oder ungewichtete Regression gerechnet werden soll. Auch nach der Schätzung sind gegebenenfalls weitere Anpassungen notwendig. Wie erwähnt, ist nicht garantiert, dass alle geschätzten Übergangswahrscheinlichkeiten im Intervall $[0,1]$ liegen und somit inhaltlich interpretiert werden können. Dies ist bei ca. 20¹²⁸ der 49 zu schätzenden Koeffizienten hier der Fall. Zwei mögliche Algorithmen, mit denen die Schätzwerte korrigieren werden können, wurden zuvor bereits vorgestellt.¹²⁹

Aus diesen Festlegungen resultieren insgesamt acht Möglichkeiten, die ökologische Regression nach Goodman durchzuführen. Tabelle 7 stellt den AD zwischen den erhaltenen geschätzten Wählerwanderung für diese acht Modelle dar. Der größte AD zwischen zwei Modellen ist 0,56. Es handelt sich um den Unterschied zwischen der gewichteten Regression mit relativen Häufigkeiten und der Korrektur nach Gleichung (10) und der ungewichteten Regression mit der Zahl der Wahlberechtigten und der Korrektur nach Gleichung (11).

Modell		Korrektur nach (10)				Korrektur nach (11)				
		P, g	P, u	N, g	N, u	P, g	P, u	N, g	N, u	
nach (10)	Korrektur	P, g	0,00	0,07	0,24	0,41	0,20	0,18	0,42	0,56
		P, u	0,07	0,00	0,19	0,36	0,22	0,17	0,40	0,54
		N, g	0,24	0,19	0,00	0,21	0,33	0,25	0,32	0,45
		N, u	0,41	0,36	0,21	0,00	0,42	0,38	0,27	0,35
nach (11)	Korrektur	P, g	0,20	0,22	0,33	0,42	0,00	0,10	0,29	0,43
		P, u	0,15	0,17	0,25	0,38	0,10	0,00	0,27	0,41
		N, g	0,42	0,40	0,32	0,27	0,29	0,27	0,00	0,16
		N, u	0,56	0,54	0,45	0,35	0,43	0,41	0,16	0,00

Tabelle 7: Evaluierung der Methoden für die ökologische Inferenz, ökologische Regression nach Goodman, Schätzung der Wählerwanderung bei den Bundestagswahlen 2005 und 2009 in München, Unterschied zwischen den Schätzungen bei möglichen Festlegungen, dargestellt ist der AD zwischen den geschätzten Wählerwanderungstabellen. P steht für Nutzung der relativen Häufigkeiten, N für die Zahl der Wahlberechtigten, g steht für eine gewichtete Regression, u für eine ungewichtete Regression, für Details zu den Korrekturverfahren für Schätzwerte außerhalb des Intervalls, siehe Kapitel II.3.1.3..

128 Die genaue Zahl unterscheidet sich bei den betrachteten Modellen, die Größenordnung ist jedoch ähnlich.

129 Vgl. Kapitel II.3.1.3..

Die meisten ADs liegen jedoch im Bereich von 0,20 – 0,30. Die niedrigsten Unterschiede gibt es zwischen der gewichteten und ungewichteten Regression bei Nutzung der relativen Häufigkeiten innerhalb eines Korrekturverfahrens. Das verwendete Korrekturverfahren hat einen relevanten Einfluss auf die Schätzung.

Bei der Nutzung von Thomsens Probit/Logit Modell, siehe Kapitel II.3.2., ist die Festlegung einer Referenzpartei für den RxC Algorithmus notwendig. Es muss für jede der beiden betrachteten Wahlen jeweils eine Referenzpartei festgelegt werden, so dass insgesamt R x C Möglichkeiten bestehen. Im Folgenden wird jedoch für beide Wahlen immer die gleiche Referenzpartei verwendet. Thomsen empfiehlt die Nutzung einer großen neutralen Partei, bei der Schätzung von Wählerwanderungen die Nichtwähler.¹³⁰ Alternativ kann Parks Erweiterung, siehe Kapitel II.3.2.3., genutzt werden, bei der auf eine Festlegung von Referenzparteien verzichtet werden kann. Daraus resultieren acht Möglichkeiten, das Modell zu spezifizieren.

Der AD zwischen den geschätzten Wählerwanderungen ist in Tabelle 8 dargestellt. Wie bei der ökologischen Regression nach Goodman sind auch hier deutliche Unterschiede zwischen den einzelnen Schätzungen festzustellen. Den größten AD von 0,52 gibt es zwischen der Schätzung mit der FDP bei beiden Wahlen als Referenzpartei und den Sonstigen bei beiden Wahlen als Referenzpartei. Die meisten Werte sind auch hier größer als 0,2. Den geringste AD gibt es zwischen dem Modell mit den Nichtwählern bei beiden Wahlen als Referenzpartei und Parks Vorschlag, bei dem auf eine Festlegung von Referenzparteien verzichtet werden kann.

Ref-Partei	CSU	SPD	FDP	Grüne	DieLinke	Sonst	NW	Park
CSU	0,00	0,22	0,29	0,25	0,37	0,46	0,34	0,42
SPD	0,22	0,00	0,20	0,19	0,45	0,52	0,39	0,47
FDP	0,29	0,20	0,00	0,23	0,44	0,50	0,37	0,45
Grüne	0,25	0,19	0,23	0,00	0,40	0,44	0,32	0,39
DieLinke	0,37	0,45	0,44	0,40	0,00	0,32	0,21	0,24
Sonst	0,46	0,52	0,50	0,44	0,32	0,00	0,20	0,17
NW	0,34	0,39	0,37	0,32	0,21	0,20	0,00	0,10
Park	0,42	0,47	0,45	0,39	0,24	0,17	0,10	0,00

Tabelle 8: Evaluierung der Methoden für die ökologische Inferenz, Thomsens Probit/Logit Modell, Schätzung der Wählerwanderung bei den Bundestagswahlen 2005 und 2009 in München, Unterschied zwischen den Schätzungen für mögliche Wahlen der Referenzparteien, dargestellt ist der AD zwischen den geschätzten Wählerwanderungstabellen. Bei beiden Wahlen wird jeweils die gleiche Referenzpartei genutzt, Park steht für Parks Vorschlag ohne Referenzpartei, siehe Kapitel II.3.2.3..

¹³⁰ Vgl. Thomsen, 1987, S. 73–74 und S. 76.

Modell	Standard	Standard, a	Complete	Complete, a	Standard, d
Standard	0,00	0,03	0,27	0,30	0,44
Standard, a	0,03	0,00	0,27	0,30	0,47
Complete	0,27	0,27	0,00	0,05	0,59
Complete, a	0,30	0,30	0,05	0,00	0,63
Standard, d	0,44	0,47	0,59	0,63	0,00

Tabelle 9: Evaluierung der Methoden für die ökologische Inferenz, Thomsens Probit/Logit Modell, Schätzung der Wählerwanderung bei den Bundestagswahlen 2005 und 2009 in München, Unterschied zwischen den Schätzungen bei den weiteren Festlegungen in der Software, dargestellt ist der AD zwischen den geschätzten Wählerwanderungstabellen. Bei allen Modellen werden die Nichtwähler als Referenzpartei für beide Wahlen genutzt, Standard ist der in Kapitel II.3.2. beschriebenen Algorithmus, Complete die in der Software vorhanden alternative Implementierung, a steht für Modelle, bei denen für Stichprobeneffekte korrigiert wurde, d (=0,5) falls die Korrelationsmatrizen abgeschwächt werden.

Die Wahl der Referenzpartei ist jedoch nicht die einzig notwendige Festlegung in Thomsens Softwareimplementierung Ecol. Zwei Parameter sind dabei vor allem von Interesse. Der Parameter deflat erlaubt eine Abschwächung der geschätzten Korrelationen und wird multiplikativ auf die geschätzten Korrelationsmatrizen angewandt, während der Parameter adjust mögliche Stichprobeneffekte korrigieren soll, die bei kleinen N_i auftauchen können. Zusätzlich steht noch die alternative Modellimplementierung, vergleiche Kapitel II.3.2.2., zur Verfügung, die etwas von Standardmodell abweicht. Beim sogenannten complete-Algorithmus wird für jedes Gebiet eine Verteilung analog zu (13) betrachtet und es wird angenommen, dass die Gleichsetzung der ökologische Korrelation (14) für jedes einzelne Gebiet gilt, was dazu führt, dass auch Schätzungen auf Gebietsebene möglich sind.¹³¹

Erneut zeigen sich deutliche Unterschiede zwischen den einzelnen Schätzungen, einzig die Korrektur auf Stichprobeneffekte mit adjust zeigt einen vernachlässigbaren Effekt. Die Schätzungen zwischen den beiden Algorithmen unterscheidet sich recht deutlich, der AD ist aber von der Größenordnung noch vergleichbar mit den ADs in Tabelle 8. Bei kleineren Werten beim Parameter deflat sind keine Änderungen bei der geschätzten Wählerwanderung feststellbar, jedoch zeigt Tabelle 9, dass größere Werte die Schätzung maßgeblich beeinflussen können.

Der rekursive Algorithmus von Andreadis und Chadjipadelis, siehe Kapitel II.3.3.2., verlangt weitestgehend keine vorab-Festlegungen. Einzig das Abbruchkriterium für den Algorithmus muss angegeben werden. Diese Entscheidung ist jedoch weitestgehend unkritisch, wenn der Wert des akzeptierten, nicht durch das Modell erklärten Restes nicht

¹³¹ Vgl. Thomsen / Frandsen / Kristmar / Lauritsen / Sørensen, 1995.

zu groß gewählt wird, da ein früheres Beenden des Algorithmus die vorhergehenden Iterationen nicht beeinflusst. Jedoch können zwei Aspekte beim Algorithmus geändert werden und diese Änderung können Einfluss auf die Schätzung haben. Zum einen haben Andreadis und Chadjipadelis sich nicht auf ein 2x2 Modell definitiv festgelegt. Sie nutzen in ihrer Anwendung eines der Modelle von Grofman und Merrill, erwähnen aber z. B. Kings 2x2 Verfahren von 1997 als Alternative. Das heißt, dass 2x2 Modell kann als eine mögliche Festlegung angesehen werden. Zum anderen zeigt die Abänderung des Modells durch Kellermann auf, dass auch die exakte Gestaltung des Algorithmus als Festlegung angesehen werden kann. Kellermann erwartet durch seine Veränderung der Reihenfolge eine Verbesserung der Schätzung von Wählerwanderungen und motiviert die Abänderung entsprechend.

Tabelle 10 zeigt den AD zwischen den geschätzten Wählerwanderungen bei drei betrachteten Modellvarianten. Es handelt sich um das von Andreadis und Chadjipadelis vorgeschlagene Modell in der vorgeschlagenen Form, das Modell mit den ursprünglichen Algorithmus bei Nutzung des Multinomial-Dirichlet Modells¹³² und ein Modell mit Kellermanns geänderten Algorithmus bei Nutzung von Grofman und Merrills 2x2 Modell. Die Änderung des Algorithmus hat einen nur vergleichsweise geringen Einfluss auf die Schätzung, jedoch zeigt sich bei der Nutzung eines alternativen 2x2 Modells ein relevanter Unterschied zwischen den geschätzten Wählerwanderungen. Bei einer genauen Betrachtung des Schätzprozesses zeigt sich, dass der große Unterschied auf eine geänderte Schätzreihenfolge in den ersten Iterationen zurückgeführt werden kann. Weiterhin ist auffällig, dass bei allen drei Modellvarianten sehr hohe Übergangswahrscheinlichkeiten von deutlich über 95% geschätzt werden, die in dieser Höhe als nicht realistisch angesehen werden müssen.

Modell	And & Cha, Standard	And & Cha, M-D M	Kellermann, Standard
And & Cha, Standard	0,00	0,32	0,10
And & Cha, M-D M	0,32	0,00	0,28
Kellermann, Standard	0,10	0,28	0,00

Tabelle 10: Evaluierung der Methoden für die ökologische Inferenz, Andreadis und Chadjipadelis rekursives Modell, Schätzung der Wählerwanderung bei den Bundestagswahlen 2005 und 2009 in München, Unterschied zwischen den Schätzungen bei möglichen Festlegungen, dargestellt ist der AD zwischen den geschätzten Wählerwanderungstabellen. And & Cha zeigt die Nutzung des von Andreadis und Chadjipadelis vorgeschlagenen Algorithmus an, Kellermann die Nutzung dessen Algorithmus an. Bei Standard wird das Modell von Grofman und Merrill genutzt, M-D M zeigt die Nutzung des Multinomial-Dirichlet Modells an.

¹³² Da es sich bei den betrachteten Tabellen jeweils um 2x2 Tabellen handelt, entspricht dieses hier dem Binomial-Beta Modell.

Modell	1000/1/1000	10000/10/1000	10000/100/2000	50000/1000/1000	50000/3000/1000
1000/1/1000	0,00	0,57	0,67	0,69	0,69
10000/10/1000	0,57	0,00	0,14	0,15	0,15
10000/100/2000	0,67	0,14	0,00	0,03	0,02
50000/1000/1000	0,69	0,15	0,03	0,00	0,01
50000/3000/1000	0,69	0,15	0,02	0,01	0,00

Tabelle 11: Evaluierung der Methoden für die ökologische Inferenz, Multinomial-Dirichlet Modell, Schätzung der Wählerwanderung bei den Bundestagswahlen 2005 und 2009 in München, Unterschied zwischen den Schätzungen bei unterschiedlicher Parameterwahl, dargestellt ist der AD zwischen den geschätzten Wählerwanderungstabellen. Der erste Wert gibt den Burn-In an, der zweite Werte das Thinning und der letzte Wert die Zahl der betrachteten Ziehungen. 1000/1/1000 entspricht den Standardwerten des R-Pakets eiPack.

Beim hierarchischen Multinomial-Dirichlet Modell, siehe Kapitel II.3.4.2., sind vor allem die drei vorgestellten Parameter Burn-In, Thinning und die Zahl der betrachteten Ziehungen aus den Markov Ketten von Interesse.¹³³ Es kann festgehalten werden, dass die Standardwerte¹³⁴ deutlich zu niedrig sind und mit diesen Werten noch nicht von einer Konvergenz der Ketten ausgegangen werden kann.

Tabelle 11 zeigt den AD zwischen den Schätzungen bei unterschiedlicher Parameterwahl. Vor allem die Schätzung mit den Standardwerten unterscheidet sich recht deutlich von den Schätzungen mit größeren Burn-In und Thinning. Ab einem Burn-In von 10000 und einem Thinning von 200 gibt es nur noch sehr geringe Unterschiede zwischen den Schätzungen. Zwischen den beiden Modellen mit den längsten Ketten gibt es keine echten Unterschiede, beide Schätzungen können als identisch angesehen werden. Als geeignet für diese Daten kann daher ein Burn-In von 50000, ein Thinning von 1000 und 1000 betrachtete Ziehungen angesehen werden, auch wenn die Autokorrelation zwischen zwei Ziehungen bei dieser Parameterwahl noch nicht bei allen Koeffizienten vernachlässigbar klein ist.

Bei einer wiederholten Schätzung mit dem Multinomial-Dirichlet Modell konvergieren die Ketten in etwa gegen die gleichen Verteilungen, der AD ist fast immer kleiner 0,02. Abbildung 4 zeigt die Markov Ketten für drei ausgewählte Übergänge bei einer wiederholten Modellschätzung mit den gleichen Parametern¹³⁵. Der zuvor angesprochene Burn-In Wert von 50000 wird durch eine graue Linie in den linken Grafiken markiert. Alle drei Ketten erreichen bereits vor diesen Wert ein stabiles Niveau. Es können zwar numerische Abweichungen bei den geschätzten Übergangswahrscheinlichkeiten

¹³³ Die Auswirkung abweichender Priori-Parameter bei den Hyperprioris wird in diesem Kapitel nicht untersucht.

¹³⁴ Burn-In: 1000, Thinning: 1, Ziehungen: 1000.

¹³⁵ Burn-In: 0, Thinning: 1000, Ziehungen: 2000.

festgestellt werden, diese sind aber auf einer inhaltlichen Ebene weitestgehend vernachlässigbar.

Bei einer detaillierten Betrachtung einzelner Ketten fällt jedoch auf, dass es bei einigen

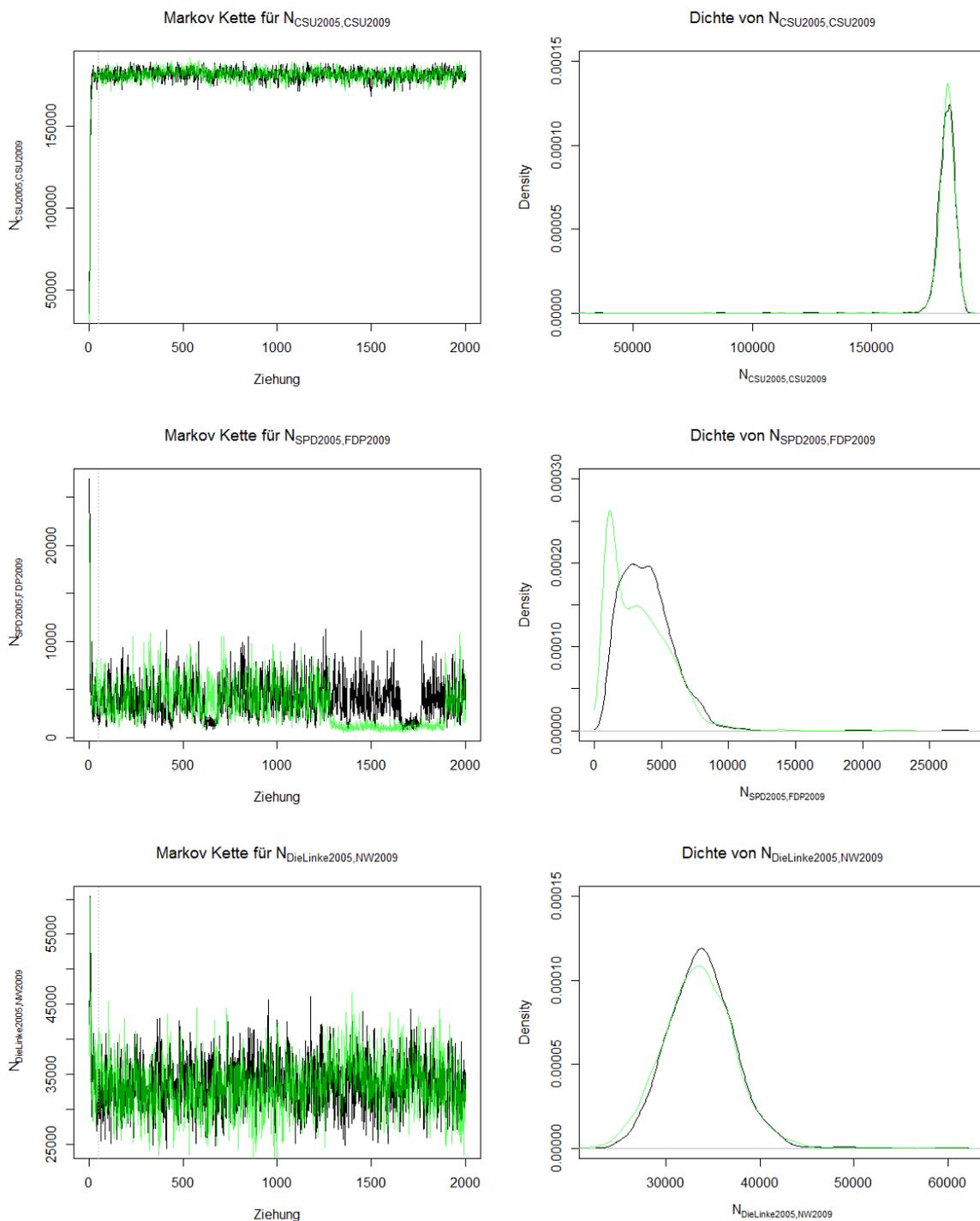


Abbildung 4: Evaluierung der Methoden für die ökologische Inferenz, Multinomial-Dirichlet Modell, Schätzung der Wählerwanderung bei den Bundestagswahlen 2005 und 2009 in München, Vergleich von zwei Modellschätzungen, Modell 1: schwarz, Modell 2: grün. Die Parameter bei beiden Modellen sind identisch, Burn-In: 0, Thinning: 1000, Ziehungen: 2000. Die linken Grafiken zeigen die Markov Ketten für drei ausgewählten Übergänge, die rechten Grafiken die zugehörigen Dichten, die graue Linie in den linken Grafiken markiert den Wert 50000.

von ihnen zu einer kurzfristigen Abweichung vom eigentlich erreichten stabilen Niveau kommen kann. Diese Ketten scheinen temporär einer anderen Verteilung zu folgen, bis sie wieder zurück zum vorher erreichten Niveau springen. Dieses Phänomen taucht vor allem bei kleineren Übergängen auf und ist in Abbildung 4 deutlich bei den Ketten für $N_{SPD2005,FDP2009}$ zu erkennen.

Bei der Nutzung des hierarchischen Multinomial-Normal Modells, siehe Kapitel II.3.4.3., sind die gleichen drei Parameter – Burn-In, Thinning und Zahl der betrachteten Ziehungen – wie beim Multinomial-Dirichlet Modell zu wählen. Zusätzlich kommt bei diesem Modell noch die Wahl einer Referenzpartei bei der zweiten Wahl hinzu. Greiner und Quinn betrachten die Enthalter, bei Wählerwanderungen die Nichtwähler, als natürliche Wahl für die Referenzpartei. Dieser Empfehlung wird bei der Überprüfung der Konvergenz der Ketten gefolgt.

Auch beim Multinomial-Normal Modell sind die Standardwerte für Burn-In und Thinning unzureichend.¹³⁶ Die Korrelation zwischen benachbarten Ziehungen ist bei diesem Modell jedoch deutlich größer als beim Multinomial-Dirichlet Modell, so dass bei einigen Übergängen auch bei einem Thinning von 10000 noch von einer moderaten Korrelation zwischen benachbarten Ziehungen ausgegangen werden muss. Bei vergleichbaren Werten wie beim Multinomial-Dirichlet Modell kann noch nicht von einer Konvergenz der Ketten ausgegangen werden, da ADs von über 0,05 beobachtet werden können. Der AD zwischen Modellen mit sehr lange Ketten¹³⁷ ist jedoch relativ niedrig und liegt unter 0,02.

Jedoch ist fraglich, ob es bei diesen Modellen bereits zu einer echten Konvergenz der Markov Ketten gekommen ist. Bei einigen der Ketten ist zu beobachten, dass diese sehr lange sehr nahe null verlaufen, um dann deutlich anzusteigen. Abbildung 5 zeigt die Ketten für drei ausgewählten Übergänge bei einer Gesamtkettenlänge von 10000000¹³⁸. Der Übergang von der CSU zur CSU ist bei beiden Modell identisch. Jedoch zeigt sich beim Übergang von der CSU zur FDP bei den Ketten ein abweichender Verlauf zwischen den beiden Modellen. Die Kette von Modell 1, schwarze Linie, nimmt sehr lange den Wert null an, um am Ende, nach 700 Ziehungen – bei einem Thinning von 10000 –, auf einen Wert

136 Die Spezifikation der Parameter beim Multinomial-Normal Modell unterscheidet sich vom Multinomial-Dirichlet Modell. Dargestellt werden im Folgenden nicht die Werte, die beim Funktionsaufruf übergeben werden müssen, sondern Werte, die vergleichbar zum Multinomial-Dirichlet Modell in ihrer Interpretation sind. Die Standardwerte für die Parameter sind, Burn-In: 10000, Thinning: 1000, Ziehungen: 990.

137 Burn-In: 2000000, Thinning: 10000, Ziehungen: 1000.

138 Burn-In: 0, Thinning: 10000, Ziehungen: 1000.

von ca. 5000 anzusteigen. Die Kette von Modell 2, grüne Linie, befindet sich die gesamte Zeit ungefähr um einen Wert von 2500.

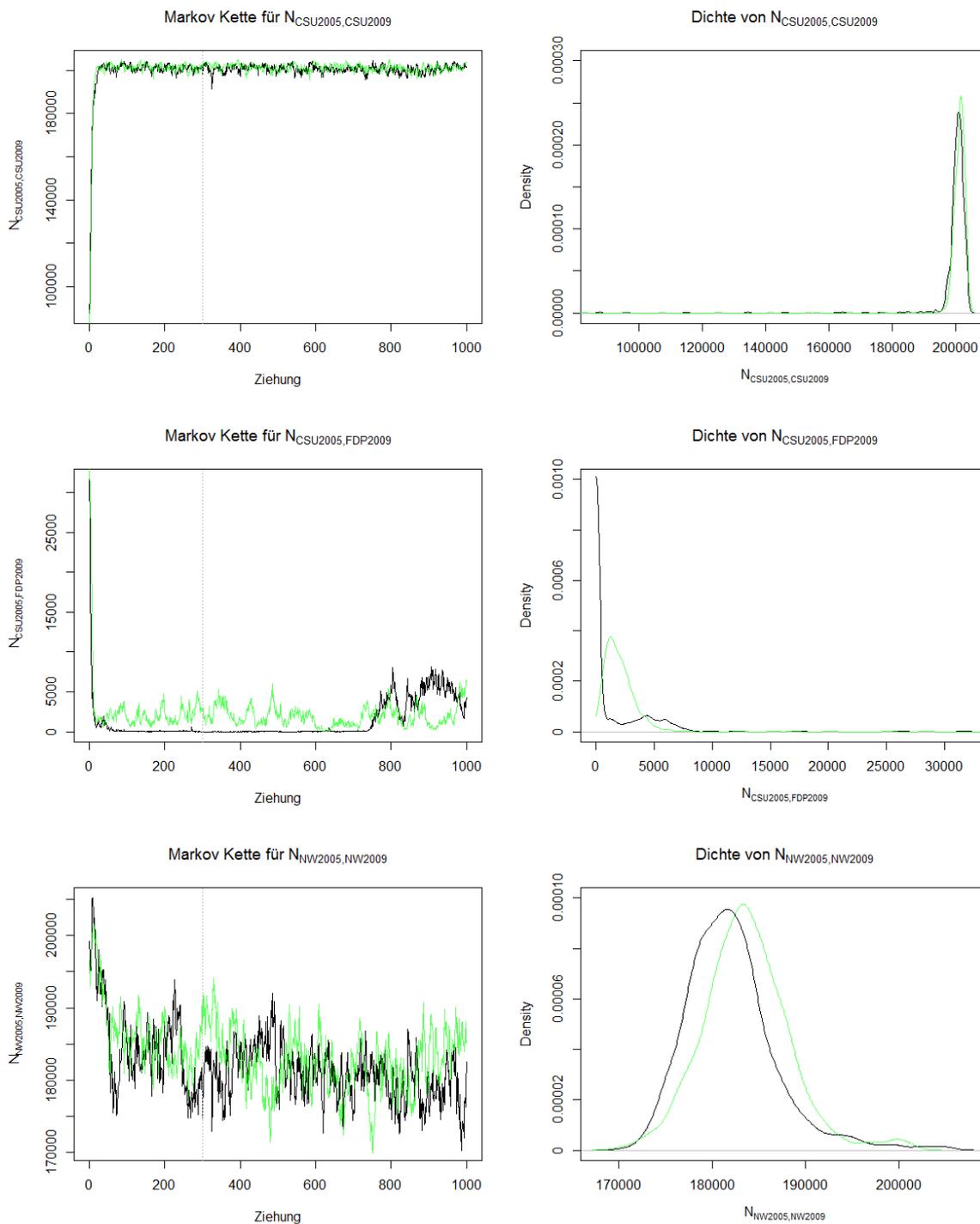


Abbildung 5: Evaluierung der Methoden für die ökologische Inferenz, Multinomial-Normal Modell, Schätzung der Wählerwanderung bei den Bundestagswahlen 2005 und 2009 in München, Vergleich von zwei Modellschätzungen, Modell 1: schwarz, Modell 2: grün. Die Parameter bei beiden Modellen sind identisch, Burn-In: 0, Thinning: 10000, Ziehungen: 1000. Die linken Grafiken zeigen die Markov Ketten für drei ausgewählten Übergänge, die rechten Grafiken die zugehörigen Dichten, die graue Linie in den linken Grafiken markiert den Wert 300000.

Beide Ketten sind daher offensichtlich (noch) nicht gegen die gleiche Verteilung konvergiert. Bei den Übergang von den Nichtwählern zu den Nichtwählern scheint vor allem bei Modell 1 noch ein Trend zu existieren, der sich in tendenziell abnehmenden Werten manifestiert. Neben den Übergängen, die als exakt null oder sehr nahe der null geschätzt werden, gilt zusätzlich, dass die mit dem Modell geschätzten Loyalitätsraten zum Teil unrealistisch hoch sind und Werte nahe eins annehmen.

Neben der Wahl der Modellparameter stellt die Wahl der Referenzpartei die zweite Festlegung beim Multinomial-Normal Modell dar. Die Referenzpartei wird jedoch nicht explizit als eigener Parameter in der Softwareimplementierung¹³⁹ übergeben. Beim Modell wird automatisch die letzte Spalte beim Funktionsaufruf als gewünschte Referenzpartei angenommen. Die Auswirkung dieser Festlegung wird überprüft, indem die Wählerwanderung mit jeweils jeder möglichen Partei als Referenzpartei erneut geschätzt wird. Für die Schätzung wird ein Burn-In von 3000000 und ein Thinning von 1000 gewählt. Insgesamt sollen 2000 Ziehungen betrachtet werden, so dass die Gesamtkettenlänge 5000000 beträgt¹⁴⁰. Ein Burn-In von 3000000 erscheint ausreichend groß, siehe Abbildung 5, um zumindest die initiale Abhängigkeit der Ketten von den Startwerten zu berücksichtigen.

Der AD zwischen den geschätzten Wählerwanderungen zwischen den Modellen mit abweichender Referenzpartei ist in Tabelle 12 angegeben. Im Gegensatz zur Wahl der Parameter hat die Wahl der Referenzpartei einen größeren Einfluss auf die Schätzung der Wählerwanderungen. Es können ADs von bis zu 0,14 zwischen Modellen mit ansonsten identischer Parametrisierung identifiziert werden. Die ADs zwischen dem Modell mit der Referenzpartei Nichtwähler, bei dem eine leicht abweichende Parametrisierung genutzt wird,¹⁴¹ und den Modellen mit anderer Referenzpartei sind noch größer und liegen zwischen 0,14 und 0,18. Eine mögliche Ursache für die Abhängigkeit der Schätzungen von der Wahl der Referenzpartei ist das durch die Modellstruktur implizit angenommene Priori-Wissen über die einzelnen Größen.

139 Vgl. Greiner / Baines / Quinn, 2013.

140 Für das Modell mit Referenzpartei Nichtwähler wird abweichend davon auf die Ergebnisse des zuvor betrachteten Modells mit Gesamtkettenlänge 10000000 zurückgegriffen. Aufgrund des gewählten Thinning Wertes von 10000 stehen 1000 Ziehungen aus den Ketten zur Verfügung. Von diesen werden die ersten 300, dies entspricht einem Burn-In von 3000000, für diese Analyse verworfen.

141 Um den Einfluss der abweichenden Parametrisierung weitestgehend zu minimieren, wird auch der AD zwischen stärker vergleichbaren Modellen betrachtet. Beim Modell für die Nichtwähler werden nur die ersten 200 Ziehungen nach dem Burn-In von 3000000 betrachtet, dies entspricht einer Gesamtkettenlänge von 5000000. Um das größere Thinning beim Modell mit den Nichtwählern zu berücksichtigen, wird bei den anderen Modellen nur jede 10 Ziehung verwendet, so dass auch bei diesen Modellen 200 Ziehungen mit einem Thinning von 10000 nach dem Burn-In von 3000000 betrachtet werden. Die ADs sind in diesem Fall minimal, d. h. um $\sim 0,01$, kleiner.

Ref-Partei	CSU	SPD	FDP	Grüne	DieLinke	Sonst	NW
CSU	0,00	0,10	0,08	0,08	0,14	0,10	0,18
SPD	0,10	0,00	0,06	0,06	0,11	0,08	0,16
FDP	0,08	0,06	0,00	0,04	0,09	0,06	0,14
Grüne	0,08	0,06	0,04	0,00	0,09	0,07	0,14
DieLinke	0,01	0,11	0,09	0,09	0,00	0,11	0,18
Sonst	0,10	0,08	0,06	0,07	0,11	0,00	0,14
NW	0,18	0,16	0,14	0,14	0,18	0,14	0,00

Tabelle 12: Evaluierung der Methoden für die ökologische Inferenz, Multinomial-Normal Modell, Schätzung der Wählerwanderung bei den Bundestagswahlen 2005 und 2009 in München, Unterschied zwischen den Schätzungen bei unterschiedlicher Wahl für die Referenzpartei, dargestellt ist der AD zwischen den geschätzten Wählerwanderungstabellen. Die Parameter für die Modelle sind Burn-In: 3000000, Thinning: 1000, Ziehungen: 2000, abweichend für das Modell mit Referenzpartei Nichtwähler: Burn-In: 0, Thinning: 10000, Ziehungen: 1000. Um einem Burn-In von 3000000 zu simulieren, werden bei den Modellen mit Referenzpartei Nichtwähler die ersten 300 Ziehungen verworfen.

II.4.2.3. Festlegungen und Einstellungen für die Simulationsstudie

Die Betrachtungen im letzten Kapitel dokumentieren, dass es bei unterschiedlichen Festlegungen auch innerhalb eines Modells zu deutlichen Unterschieden bei den geschätzten Wählerwanderungen kommen kann. Umso weniger Festlegungen getroffen werden müssen, desto stabiler ist das Modell und die Schätzung der Wählerwanderung mit dem Modell. Dies zeigt sich vor allem beim Multinomial-Dirichlet Modell, bei dem bei der Parameterwahl nur darauf geachtet werden muss, dass eine Konvergenz der Ketten erreicht wird.

Anhand der Ausführungen zu jedem Modell ist es jedoch möglich, sich für die folgenden Betrachtungen auf ein bestimmtes Modell festzulegen. Dies ist erforderlich, da in der folgenden Simulationsstudie nicht alle vorgestellten Alternativen berücksichtigt werden können. Bei den ausgewählten Modellen handelt es sich aber nicht notwendigerweise um das beste Modell im Bezug auf die Schätzgüte, da diese mit den vorliegenden Daten nicht bestimmt werden kann. Neben den oben vorgestellten Ergebnissen, sind auch externe Entscheidungsgründe bei der Auswahl mit eingeflossen.

Bei der ökologischen Regression nach Goodman wird im Folgenden das Modell mit den relativen Häufigkeiten und ungewichteter Regression genutzt. Die relativen Häufigkeiten werden gewählt, da dies dem klassischen Modell entspricht. Bei der Entscheidung, ob eine gewichtete oder ungewichtete Regression genutzt werden soll, wird den Überlegungen von King gefolgt¹⁴². Auch gibt es nur relativ geringe Unterschiede zwischen gewichteter und

142 Vgl. King, 1997, S. 61.

ungewichteter Regression bei der Nutzung der relativen Häufigkeiten. Für die Anpassung von Übergangswahrscheinlichkeiten außerhalb von $[0,1]$ wird die in Gleichung (11) dargestellte Korrektur gewählt. Im Gegensatz zur Korrektur nach (10) werden mit diesem Verfahren weniger Werte auf null gesetzt.

Bei Thomsens Probit/Logit Modell werden für beide Wahlen die Nichtwähler als Referenzpartei ausgewählt. Die Wahl der Nichtwähler wurde von Thomsen selber vorgeschlagen. Die durchgeführten Analysen ließen auch keine Gründe erkennen, die eindeutig für eine andere Wahl bei den Referenzparteien sprechen würden. Für die Simulation wird auf Thomsens Originalalgorithmus zurückgegriffen. Die Entscheidung für den Originalalgorithmus für die Simulation ist der besseren Dokumentation dieser Methode, der schnelleren Rechenzeit und größeren Stabilität¹⁴³ geschuldet. Die Korrektur auf Stichprobeneffekte wird bei der Schätzung genutzt.

Beim rekursiven Modell von Andreadis und Chadjipadelis wird der Originalalgorithmus und Kellermanns Änderung des Algorithmus, jeweils mit dem Modell von Grofman und Merrill, im Folgenden genutzt. Bei der Verwendung des Multinomial-Dirichlet Modells und des Modells von Grofman und Merrill werden jeweils Übergangswahrscheinlichkeiten geschätzt, die als unrealistisch betrachtet werden müssen. Es ist jedoch nicht möglich, eine der Schätzungen als eindeutig der anderen Schätzung überlegen zu klassifizieren. Aus diesem Grund wird die Originalversion genutzt. Kellermann Änderung des Algorithmus wird aufgrund von externen Überlegungen¹⁴⁴ mit berücksichtigt.

Beim Multinomial-Dirichlet Modell ist es notwendig, die Parameter für die Simulation festzulegen, da eine detaillierte Evaluierung eines jeden Modells nicht praktikabel ist. Für den Burn-In wird ein Wert von 50000 gewählt, Thinning wird auf 1000 festgesetzt. Insgesamt sollen 1000 Ziehungen aus den Markov Ketten gezogen werden. Diese Parameterwahl stellt bei der Schätzung der Wählerwanderung bei den Bundestagswahlen 2005 und 2009 in München einen guten Kompromiss zwischen Rechenzeit und Stabilität der Ketten dar. Auf eine Anpassung der Priori-Parameter der Hyperprioris wird verzichtet, es werden die Standardwerte¹⁴⁵ für die Simulation genutzt.

Für das Multinomial-Normal Modell wird im Folgenden ein Burn-In von 3000000 und ein Thinning von 1000 gewählt. Der Burn-In erscheint ausreichend groß, um zumindest die initiale Abhängigkeit vom Startwert zu berücksichtigen. Bei der Wahl der Thinnings

143 Vgl. Hilfetext zur Methode „complete“ in: Thomsen / Frandsen / Kristmar / Lauritsen / Sørensen, 1995.

144 Diese Analyse wurde ursprünglich im Rahmen einer Forschungskoooperation für das Statistische Amt der Landeshauptstadt München durchgeführt. Bei der Auswahl der Modelle flossen daher auch die Wünsche des Statistischen Amtes mit ein.

145 Gamma-Verteilung mit shape = 4 und rate = 2.

Modell	Goodman	Thomsen	And & Cha	Kellermann	M-D M	RxCeCol
Goodman	0,00	0,33	0,51	0,52	0,54	0,68
Thomsen	0,33	0,00	0,55	0,55	0,56	0,69
And & Cha	0,51	0,55	0,00	0,10	0,26	0,27
Kellermann	0,52	0,55	0,10	0,00	0,25	0,25
M-D M	0,54	0,56	0,26	0,25	0,00	0,24
RxCeCol	0,68	0,69	0,27	0,25	0,24	0,00

Tabelle 13: Evaluierung der Methoden für die ökologische Inferenz, Schätzung der Wählerwanderung bei den Bundestagswahlen 2005 und 2009 in München, Unterschied zwischen den Schätzungen mit den Modellen, dargestellt ist der AD zwischen den geschätzten Wählerwanderungstabellen, Tabelle 6 zeigt die Übersicht über die ausgewählten Modelle und die hier verwendeten Abkürzungen. Für die genaue Parametrisierung und die Festlegungen bei den einzelnen Modellen, siehe Kapitel II.4.2.3., abweichend davon wurde beim Multinomial-Normal Modell das Modell mit einer Gesamtkettenlänge von 10000000 (Thinning: 10000, Ziehungen: 1000) und Referenzpartei Nichtwähler genutzt Um einem Burn-In von 3000000 zu simulieren, werden bei den Modellen mit Referenzpartei Nichtwähler die ersten 300 Ziehungen verworfen.

handelt es sich um einen Kompromiss zwischen Rechenzeit und weiterhin vorhandener Korrelation zwischen benachbarten Ziehungen. Insgesamt werden 1000 Ziehungen betrachtet. Als Referenzpartei wird bei der Simulation auf die Empfehlung Nichtwähler zurückgegriffen. Wie auch beim Multinomial-Dirichlet Modell wird jeweils nur ein Modell berechnet. Bei der weiterführenden Evaluierung des Modells, die sich vor allem der Stabilität der Schätzung und der Abhängigkeit von der Referenzpartei widmet, werden längere Ketten¹⁴⁶ verwendet.

Ein Vergleich der geschätzten Wählerwanderungen mit den verschiedenen Modellen ist in Tabelle 13¹⁴⁷ dargestellt. Auffällig ist, dass die Schätzung mit der ökologischen Regression nach Goodman und Thomsens Probit/Logit Modell den größten Abstand von den anderen Modellen haben, während sie sich selbst noch am ähnlichsten sind. Der AD zwischen den Schätzungen mit den anderen Modellen ist deutlich niedriger, liegt aber immer noch bei ~0,25. Wie bereits diskutiert, hat die Nutzung des geänderte Algorithmus von Kellermann hier nur einen relativ geringen Einfluss auf die Schätzung. Die Zahlen dokumentieren, dass die Wahl der Methode einen maßgeblichen Einfluss auf die geschätzte Wählerwanderung hat. Jedoch beschreiben die dargestellten Werte nur das relative Verhältnis der Modellschätzungen zueinander.

Zwei der geschätzten Wählerwanderungen sind in Abbildung 6 dargestellt. Die obere Tabelle zeigt die Schätzung mit den Multinomial-Dirichlet Modell, die untere Tabelle die Schätzung mit Thomsens Probit/Logit Modell. Der große Unterschied zwischen beiden

146 Burn-In: 0, Thinning: 1000, Ziehungen: 10000.

147 Die Abweichungen zur Tabelle 3 in Klima / Thurner / Molnar / Schlesinger / Küchenhoff, 2016, S. 142 ergeben sich aus der Nutzung des Originalalgorithmus bei Thomsen in dieser Darstellung.

Schätzungen, der AD beträgt 0,56, zeigt sich auch in dieser Darstellung. Während z. B. das Multinomial-Dirichlet Modell ungefähr 182000 loyale CSU-Wähler schätzt, werden mit Thomsens Modell 131000 loyale CSU-Wähler geschätzt.

Da beide Schätzungen aber die durch die amtlichen Wahlergebnisse aufgestellten Restriktionen erfüllen, ist es ohne weiteres Wissen nicht möglich, eine der beiden Wählerwanderungen zu bevorzugen. Die Ausführungen in den letzten beiden Kapiteln zeigen zwar, dass es deutliche Unterschiede zwischen den Modellen und auch Modell-intern gibt, jedoch ist es mit den bisherigen Ergebnissen nicht möglich, Aussagen über die Güte der Modellschätzungen zu tätigen.

2005	CSU	181867	3182	4184	2700	1733	2156	44131
	SPD	2724	103690	4202	2258	14053	26445	32408
	FDP	1802	1440	69727	1461	1246	1506	1446
	Grüne	1864	1985	3092	82043	1658	1113	1571
	DieLinke	948	1079	875	898	17650	2018	1308
	Sonst	923	1138	834	717	4061	7869	2646
	NW	19336	13466	31682	24300	3498	3661	164770
		CSU	SPD	FDP	Grüne 2009	DieLinke	Sonst	NW

2005	CSU	130987	22982	27238	4547	2405	9836	41907
	SPD	28714	51572	10235	11484	13745	13089	56909
	FDP	27809	4416	37866	5200	313	539	2466
	Grüne	2523	11026	16531	56335	2455	904	3531
	DieLinke	715	3112	724	2057	6340	2333	9492
	Sonst	1914	1513	476	496	1343	2266	10177
	NW	16638	31240	21556	34238	17292	15740	123974
		CSU	SPD	FDP	Grüne 2009	DieLinke	Sonst	NW

Abbildung 6: Evaluierung der Methoden für die ökologische Inferenz, Schätzung der Wählerwanderung bei den Bundestagswahlen 2005 und 2009 in München, dargestellt sind die geschätzten absoluten Zahlen für jede Kombination. Unterschiede vom Wahlergebnis sind Algorithmus bedingt. oben: Multinomial-Dirichlet Modell (Burn-In: 50000, Thinning: 1000, Ziehungen: 1000), unten: Thomsens Probit/Logit Modell mit Originalalgorithmus, Nichtwählern als Referenzpartei und Adjustierung für Stichprobeneffekte.

II.4.3. Simulationsstudie Modellvergleich

Um die Güte der Modelle bestimmen zu können, ist die Kenntnis der wahren Wählerwanderung notwendig. Diese ist jedoch im Allgemeinen unbekannt. Die beobachteten Wahlergebnisse alleine reichen für eine Bewertung nicht aus, da, wie bereits ausgeführt, viele Tabellen zu denselben Rändern führen können. Auch wenn Individualdaten vorliegen, ist fraglich, ob diese eine geeignete Referenz darstellen können.¹⁴⁸ Aus diesem Grund soll in diesem Kapitel die Eignung der Modelle bei der Schätzungen von Wählerwanderungen im Rahmen einer Simulationsstudie evaluiert werden. Berücksichtigt werden in der Simulationsstudie die in Tabelle 6 aufgelisteten Modelle, die Festlegungen für jedes Modell sind in Kapitel II.4.2.3. dargelegt.

II.4.3.1. Szenarien für die Simulationsstudie

Ausgangspunkt für die Simulation sind erneut die Wahlergebnisse der Bundestagswahlen 2005 und 2009 in München. Betrachtet werden wie zuvor die Zweitstimmenergebnisse der Parteien. Es werden ebenfalls die gleichen 314 über die Zeit konstanten Wahlbezirke¹⁴⁹ betrachtet. Um sicherzustellen, dass das Wahlergebnis der ersten Wahl einer echten Wahl nahe kommt, wird für die Simulation das Ergebnis der Bundestagswahl 2005 als Ausgangspunkt genutzt. Zur Vereinfachung wird im Folgenden in der Simulation davon ausgegangen, dass bei beiden Wahlen die Bevölkerung identisch bleibt und dass es keine Briefwähler gibt bzw. die genaue Zuordnung der Briefwähler zu ihren Urnenwahlbezirk möglich ist.

Für die Simulation der Wählerwanderung wird auf die von Infratest dimap geschätzte Wählerwanderung für die gesamte Bundesrepublik zurückgegriffen.¹⁵⁰ Mit dieser gesamtdeutschen Wählerwanderung ist es jedoch nicht möglich, ausgehend vom Wahlergebnis bei der Bundestagswahl 2005, ein Wahlergebnis bei der Bundestagswahl 2009 zu simulieren, dass mit dem beobachteten Wahlergebnis in der Stadt München vergleichbar ist. Aus diesem Grund wurden die Übergangsraten solange angepasst, bis zumindest das simulierte Wahlergebnis auf Stadtebene mit dem echten Wahlergebnis vergleichbar ist. Die resultierende Wählerwanderung für die Simulation, gezeigt werden die Übergangswahrscheinlichkeiten, ist in Abbildung 7 zu sehen. Wie in Abbildung 3

148 Eine kurze Diskussion der Probleme mit Individualdaten bei der Schätzung von Wählerwanderungen befindet sich in Kapitel III.1.1., vergleiche auch Kapitel IV.2..

149 Vgl. Kapitel II.4.1..

150 Während in der ARD Berichterstattung nur die Salden angegeben werden, beinhalten andere Publikationen die geschätzten Übergänge, vgl. z. B. Neu, Viola: Bundestagswahl in Deutschland am 27. September 2009, Konrad-Adenauer-Stiftung, 2009, S. 94–103.

Angenommen Wählerwanderung für die Simulation
Modellvergleich ökologische Inferenz

2005	CSU	0.65	0.03	0.10	0.05	0.01	0.04	0.12
	SPD	0.06	0.45	0.05	0.10	0.05	0.04	0.25
	FDP	0.02	0.05	0.75	0.05	0.01	0.02	0.10
	Grüne	0.04	0.04	0.05	0.75	0.00	0.05	0.07
	DieLinke	0.02	0.10	0.05	0.05	0.40	0.20	0.18
	Sonst	0.02	0.02	0.20	0.10	0.20	0.30	0.16
	NW	0.15	0.09	0.05	0.03	0.07	0.05	0.56
		CSU	SPD	FDP	Grüne 2009	DieLinke	Sonst	NW

Abbildung 7: Evaluierung der Methoden für die ökologische Inferenz, angenommene Wählerwanderung für München bei der Simulation, dargestellt sind die Übergangswahrscheinlichkeiten. Die Werte basieren auf der Wählerwanderung für die Bundesrepublik Deutschland von Infratest dimap für diese beiden Wahlen, siehe Kapitel II.4.3.1. für weitere Details.

gezeigt, haben die beiden großen Parteien, die CSU und die SPD, relativ starke Verluste bei der Wahl 2009 erlitten. Dies drückt sich auch in den Übergangswahrscheinlichkeiten aus. Es wird angenommen, dass 65% der CSU-Wähler im Jahr 2005 bei der Wahl im Jahr 2009 erneut für die CSU gestimmt haben. Die angenommene Loyalitätsrate bei der SPD ist noch niedriger und beträgt 45%.

Mit dieser angenommenen Wählerwanderung ist es möglich, dass individuelle Wahlverhalten alle Wahlberechtigten zu simulieren. In jeden Wahlbezirk wird die Wahlentscheidung bei der zweiten Wahl mit den für den Wahlbezirk angenommenen Übergangswahrscheinlichkeiten simuliert. Diese Berechnung erfolgt zeilenweise, d. h. für jede Partei einzeln, mit Hilfe einer oder mehrerer Multinomialverteilung. Aus den so berechneten inneren Zellen lässt sich das simulierte Wahlergebnis bei der Bundestagswahl 2009 für den Wahlbezirk ermitteln.

Für die Schätzung der Wählerwanderung mittels ökologischer Inferenz werden nur die Aggregatdaten, das Wahlergebnis bei der Bundestagswahl 2005 und das simulierte Wahlergebnis für die Bundestagswahl 2009, genutzt. Die inneren Zellen der Wählerwanderungstabellen werden nach der Schätzung für die Bestimmung der Güte der Schätzung verwendet. Als Maß für die Schätzgüte wird erneut der AD, siehe

Gleichung (28), verwendet. Es wird in der Simulation nur die Güte der Schätzung für die gesamte Region, für die Wählerwanderung in der Stadt München, bewertet.¹⁵¹ Bei der Simulation gibt der Wert jedoch an, wie viel Prozent der Wahlberechtigten falsch zugeordnet sind, es handelt sich also um ein absolutes Maß der Schätzgüte.

Insgesamt werden vier unterschiedliche Szenarien, die sich in der angenommenen Struktur der Verteilung der Übergangswahrscheinlichkeiten zwischen den Gebieten unterscheiden, betrachtet. Szenario Sim 1 folgt dabei den Annahmen der ökologischen Regression nach Goodman und geht von konstanten Übergangswahrscheinlichkeiten aus, d. h., für jeden Wahlbezirk wird die in Abbildung 7 dargestellte Wählerwanderung angenommen.

$$\beta_{r,c,i} = \beta_{r,c}, \quad \forall r \in \{1, \dots, 7\}, c \in \{1, \dots, 7\}, i \in \{1, \dots, 314\} \quad (29)$$

Gleichung (29) veranschaulicht die Konstanzannahme bei diesen Szenarien, in jedem Wahlbezirk i entsprechen die Übergangswahrscheinlichkeiten $\beta_{r,c,i}$ den globalen Übergangswahrscheinlichkeiten $\beta_{r,c}$. Dieses Szenario wird als nicht besonders realistisch eingeschätzt, sollte aber zumindest bei der ökologischen Regression nach Goodman einen guten Schätzer für den minimal erreichbaren Fehler liefern.

Sim 2 folgt relativ nahe der Verteilungsannahme des Multinomial-Dirichlet Modells. Anstelle von konstanten Übergangswahrscheinlichkeiten wie in Sim 1 wird bei diesem Szenario davon ausgegangen, dass die Übergangswahrscheinlichkeiten zeilenweise einer gemeinsamen Dirichlet-Verteilung folgen. Die Erwartungswerte dieser sieben Dirichlet-Verteilungen entsprechen den in Abbildung 7 gezeigte Übergangswahrscheinlichkeiten. Für die Parameter der Dirichlet-Verteilungen wird angenommen:

$$\alpha_{r,c} = 10 * \beta_{r,c}, \quad \forall r \in \{1, \dots, 7\}, c \in \{1, \dots, 7\} \quad (30)$$

In Gleichung (30) entsprechen die Parameter der Dirichlet-Verteilungen den $\alpha_{r,c}$, die $\beta_{r,c}$ sind die in Abbildung 7 gezeigte Übergangswahrscheinlichkeiten. Der Faktor 10 wird verwendet, um eine als adäquat empfundene Variabilität der Übergangswahrscheinlichkeiten zwischen den Gebieten zu erreichen.

In Sim 3 wird nicht länger von einer homogenen Bevölkerung in den Wahlbezirken ausgegangen. Es wird angenommen, dass es in jeden Wahlbezirk drei unterschiedliche Bevölkerungsgruppen gibt, die sich im Mittel in ihrem Wahlverhalten unterscheiden. In der Simulation drückt sich dies vor allem in unterschiedlichen Loyalitätsraten aus. Die Anteile

151 Auf eine Betrachtung der zum Teil vorliegenden Schätzungen auf Wahlbezirksebene wird verzichtet, da mit dieser Simulation vor allem die allgemeine Eignung der Modelle evaluiert werden soll.

der drei Bevölkerungsgruppen werden mit einer Dirichlet-Verteilung simuliert.¹⁵² Für jede der drei Bevölkerungsgruppen wird eine eigene Wählerwanderung angenommen. Die Übergangswahrscheinlichkeiten der einzelnen Populationen in den Wahlbezirken folgen dabei jeweils einer gemeinsamen Dirichlet-Verteilung.

$$\alpha_{r,c,k} = 10 * \beta_{r,c,k}, \quad \forall r \in \{1, \dots, 7\}, c \in \{1, \dots, 7\}, k \in \{1, 2, 3\} \quad (31)$$

Gleichung (31) zeigt die Parameter der 7 x 3 Dirichlet-Verteilungen. Im Vergleich zu Gleichung (30) ist der Index k hinzugekommen, der die Bevölkerungsgruppe angibt. Diese Modellierung impliziert, dass die Verteilung der Übergangswahrscheinlichkeiten zwischen den Wahlbezirken nicht länger einer gemeinsamen Dirichlet-Verteilung entspricht. Es handelt sich vielmehr in jeden Wahlbezirk um eine Mischung von drei Dirichlet-Verteilungen, wobei sich das Mischungsverhältnis zwischen den Wahlbezirken unterscheidet.

Das letzte Szenario, Sim 4, geht erneut von einer Bevölkerung in den Wahlbezirken aus. Jedoch wird für dieses Szenario das Vorhandensein von Aggregationsbias¹⁵³ angenommen. Für die Simulation gilt, dass die Wahrscheinlichkeit erneut bei der zweiten Wahl für die CSU bzw. die SPD zu stimmen, vom Wahlergebnis des jeweiligen „politischen Lagers“ bei der ersten Wahl abhängt. Es wird angenommen, dass die Loyalitätsrate von CSU-Wählern größer in Wahlbezirken ist, in denen 2005 die CSU und FDP stärker abgeschnitten hat. Für die SPD wird hingegen von niedrigeren Loyalitätsraten ausgegangen, wenn 2005 die SPD und DieLinke stärker waren. Diese Annahme bedeutet, dass die ersten beiden Zeilen in Abbildung 7 in jedem Gebiet anderes sind.

$$\begin{aligned} \alpha_{r,c,i} &= 100 * \beta_{r,c,i}, & \forall r \in \{1, 2\}, c \in \{1, \dots, 7\}, i \in \{1, \dots, 314\} \\ \alpha_{r,c} &= 100 * \beta_{r,c}, & \forall r \in \{3, \dots, 7\}, c \in \{1, \dots, 7\} \end{aligned} \quad (32)$$

Gleichung (32) zeigt die Parameter der Dirichlet-Verteilungen für diese Simulation. Im Vergleich zu Gleichung (30) ist für die ersten beiden Zeilen der Wahlbezirksindikator i hinzugekommen. Auch wird eine größere Variabilität der Übergangswahrscheinlichkeiten zwischen den Wahlbezirken angenommen. Die Aufnahme des Gebietsindikator i bei den ersten beiden Zeilen impliziert, dass nicht länger von gemeinsamen Dirichlet-Verteilungen ausgegangen werden kann. Die angenommene Wählerwanderung in einem Wahlbezirk i entstammt zwei gebietsspezifischen und fünf gemeinsamen Dirichlet-Verteilungen.

152 Für 25% der Bevölkerung wird von niedrigeren Loyalitätsraten ausgegangen, während 35% der Bevölkerung eine höhere Loyalität ihrer 2005 gewählten Partei gegenüber haben. Das Wahlverhalten der verbleibenden 40% entspricht im Mittel Abbildung 7. Die Verteilung zur Simulation der Bevölkerungsanteile in den Wahlbezirken lautet: Dirichlet(2,5; 4,0; 3,5).

153 Vgl. Kapitel II.3.1.2..

Szenario	Name	Verteilung innerhalb der Gebiete	Verteilung zwischen den Gebieten	Mehrere Bevölkerungen	Aggregationsbias
Sim 1	Goodman	Multinomialverteilung	keine (konstant)	Nein	Nein
Sim 2	M-D M	Multinomialverteilung	Dirichlet	Nein	Nein
Sim 3	Bevölkerungen	Multinomialverteilung	Dirichlet (Mischung)	Ja (3)	Nein
Sim 4	Aggregationsbias	Multinomialverteilung	Dirichlet (gebietspezifisch)	Nein	Ja

Tabelle 14: Evaluierung der Methoden für die ökologische Inferenz, Übersicht der betrachteten vier Datenszenarien in der Simulation zum Methodenvergleich der Verfahren für die ökologischen Inferenz, dargestellt sind die Schlüsseleigenschaften der Szenarien. Details zu den einzelnen Szenarien finden sich in Kapitel II.4.3.1.

Von den vier Szenarien folgen zwei, Sim 1 und Sim 2, relativ exakt der Verteilungsannahme, z. B. der Annahme konstanter Übergangswahrscheinlichkeiten der ökologischen Regression nach Goodman, eines der betrachteten Modelle. Zwei Szenarien, Sim 3 und Sim 4, entsprechen keiner der Verteilungsannahmen und es wird nicht von einer gemeinsamen Verteilung der Übergangswahrscheinlichkeiten zwischen den Gebieten ausgegangen. Bei Sim 3 handelt es sich auf Gebietsebene um eine Mischung von drei Verteilungen, während bei Sim 4 die Verteilungen von zwei Zeilen der Tabelle gebietspezifisch sind. Tabelle 14 fasst die vier Szenarien kurz zusammen und zeigt eine Übersicht der jeweiligen Schlüsseleigenschaften.

II.4.3.2. Ergebnisse der Simulationsstudie

Für jedes der im vorherigen Kapitel vorgestellten Szenarien wurden 100 Datensätze generiert. Bei vier Szenarien bedeutet dies, dass insgesamt 400 Datensätze zu evaluieren sind. Für jeden der 400 Datensätze wurde mit jeden der sechs in Tabelle 6 aufgelisteten Modelle die Wählerwanderung geschätzt, so dass insgesamt 2400 geschätzte Wählerwanderungen für die Analyse zur Verfügung stehen. Da die simulierten Individualdaten für die Evaluierung verwendet werden können, ist es möglich, die Güte der Schätzungen zu bewerten und zu vergleichen. Abbildung 8 präsentiert den AD der geschätzten Wählerwanderungen von den simulierten Individualdaten für die jeweils 100 Datensätze pro Szenario und Modell. Es gilt, dass der halbe AD angibt, wie viel Prozent der Wahlberechtigten bei der Schätzung falsch zugeordnet werden.

Evaluation der Methoden der ökologischen Inferenz: Ergebnisse der Simulation

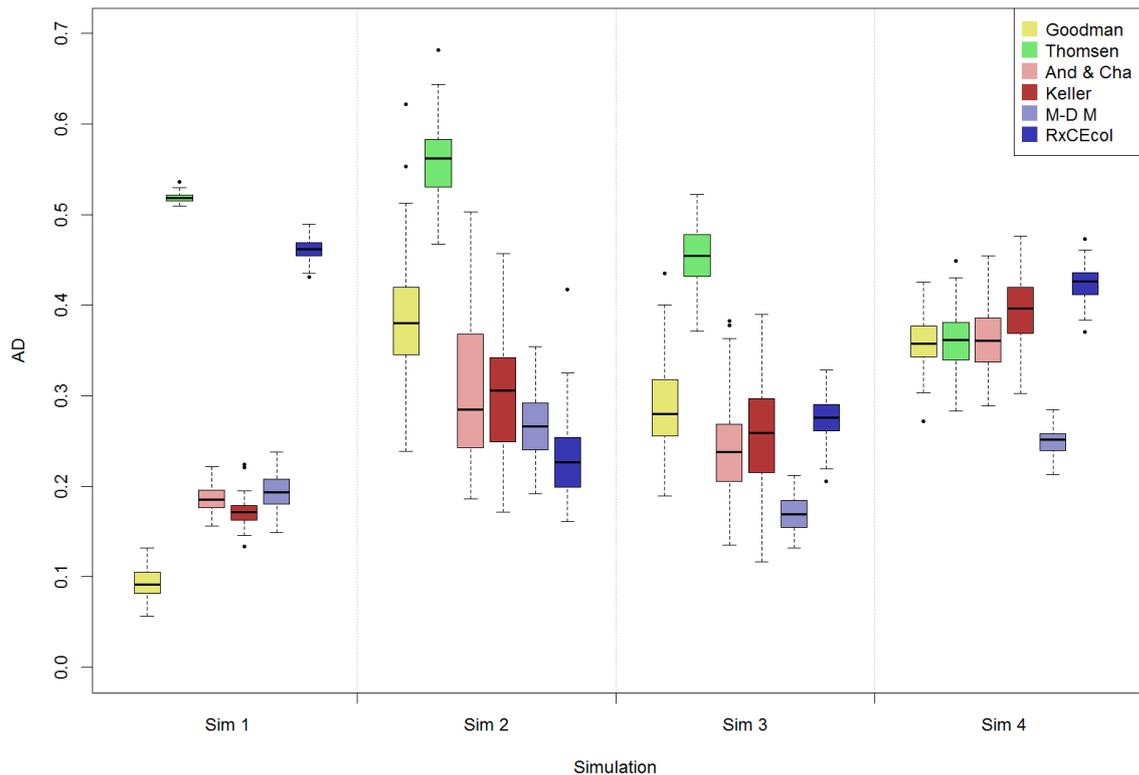


Abbildung 8: Evaluierung der Methoden für die ökologische Inferenz, Ergebnisse der Simulationsstudie, dargestellt ist der AD der Schätzung mit den Modellen zu den simulierten Individualdaten. Für jedes Szenario wurden 100 Datensätze generiert und evaluiert. Ein Überblick der Szenarien befindet sich in Tabelle 14, die berücksichtigten Modelle sind in Tabelle 6 dargestellt.

Bei Szenario Sim 1, welches den Annahmen der ökologischen Regression nach Goodman folgt, schneidet auch die ökologische Regression nach Goodman am besten ab. Die Schätzgüte mit diesem Modell ist deutlich am höchsten und die ADs liegen um 0,1. Die beiden rekursiven Algorithmen, die Modelle von Andreadis und Chadjipadelis und Kellermann, sowie das Multinomial-Dirichlet Modell liegen ungefähr auf einem vergleichbaren Niveau mit ADs von in etwa 0,2. Deutlich schlechter ist die Schätzung mit Thomsens Probit/Logit Modell und dem Multinomial-Normal Modell. Die Schätzgüte mit der ökologischen Regression nach Goodman kann aufgrund der Nähe von Modell und Szenario als eine Art „Optimum“ für die Schätzung insgesamt angesehen werden, d. h. ein AD von 0,1 wäre ein sehr gutes Schätzergebnis.

Bei Sim 2 wird die Güte der Schätzung tendenziell schlechter, nur beim Multinomial-Normal Modell zeigt sich eine deutlich bessere Schätzgüte als bei Sim 1. Das Multinomial-Normal Modell stellt auch mit einen AD von im Mittel unter 0,25 das beste Modell in diesem Szenario dar. Gefolgt wird es vom Multinomial-Dirichlet Modell, nach dessen

Verteilungsannahmen das Szenario konstruiert wurde. Der AD der Schätzungen liegt bei diesem Modell im Mittel über 0,25. Noch etwas schlechter schneiden die beiden rekursiven Algorithmen mit ADs von im Mittel größer 0,3 ab. Deutlich schlechter ist die ökologische Regression nach Goodman und Thomsens Probit/Logit Modell.

Beim Szenario mit den drei angenommenen Bevölkerungsgruppen, Sim 3, wird die beste Schätzgüte mit den Multinomial-Dirichlet Modell erreicht. Die ADs der Schätzungen mit diesem Modell sind in den meisten Fällen kleiner 0,2. Die beiden rekursiven Algorithmen erreichen im Mittel die zweit- und drittbeste Schätzgüte, wobei die Güte der Schätzungen mit dem Originalalgorithmus tendenziell etwas besser ist. Die ökologische Regression nach Goodman und das Multinomial-Normal Modell weisen eine vergleichbare Schätzgüte auf, jedoch ist die Variabilität bei Goodman größer. Erneut die schlechteste Güte zeigt Thomsens Probit/Logit Modell.

Die insgesamt – über alle Modelle gesehen – schlechteste Schätzgüte ist bei Sim 4 zu finden. Das beste Modell ist hier erneut das Multinomial-Dirichlet Modell, dass in etwa die gleiche Güte wie bei Sim 2 erreicht. Alle anderen Modelle weisen eine deutlich schlechtere Güte auf. Auf vergleichbaren Niveau sind Thomsens Probit/Logit Modell, die ökologische Regression nach Goodman und der rekursive Algorithmus von Andreadis und Chadjipadelis. Die Schätzgüte liegt bei diesen drei Modelle im Mittel bei 0,35. Der rekursive Algorithmus von Kellermann zeigt eine etwas schlechtere Schätzgüte. Das schlechteste Modell in diesem Szenario ist das Multinomial-Normal Modell mit einem AD von im Mittel über 0,4.

Aufgrund der vermuteten Konvergenzproblematik und der Abhängigkeit von der Referenzpartei erfolgte auch eine erneute Evaluierung des Multinomial-Normal Modells im Rahmen der Simulation. Für diese Betrachtungen wird der erste Datensatz von Sim 2 genutzt, also von dem Szenario, bei dem das Multinomial-Normal Modell am besten abgeschnitten hat. Für die Evaluierung wird zum einen die Schätzung mit der Referenzpartei Nichtwähler 10 mal wiederholt, zum anderen wird das Modell auch mit jeder anderen möglichen Referenzpartei geschätzt. Die Gesamtkettenlänge der Modelle beträgt 10000000, wobei die ersten 3000000 Ziehungen als Burn-In verworfen werden und das Thinning 1000 beträgt.

Tabelle 15 zeigt den AD der Schätzungen zwischen den 10 gerechneten Modellen mit Referenzpartei Nichtwähler. Während bei der Modellevaluierung bei einer vergleichbaren Kettenlänge, vgl. Kapitel II.4.2.2., nur noch geringe Unterschiede zwischen den Schätzungen festgestellt werden konnten, finden sich hier deutlich größere ADs.

	1	2	3	4	5	6	7	8	9	10	Sim
1	0,00	0,11	0,15	0,12	0,11	0,13	0,11	0,12	0,12	0,10	0,21
2	0,11	0,00	0,10	0,11	0,10	0,09	0,08	0,10	0,12	0,09	0,21
3	0,15	0,10	0,00	0,14	0,11	0,11	0,13	0,09	0,10	0,12	0,24
4	0,12	0,11	0,14	0,00	0,14	0,08	0,10	0,12	0,13	0,13	0,20
5	0,11	0,10	0,11	0,14	0,00	0,12	0,11	0,12	0,09	0,14	0,21
6	0,13	0,09	0,11	0,08	0,12	0,00	0,11	0,10	0,11	0,11	0,24
7	0,11	0,08	0,13	0,10	0,11	0,11	0,00	0,12	0,13	0,13	0,21
8	0,12	0,10	0,09	0,12	0,12	0,10	0,12	0,00	0,09	0,11	0,22
9	0,12	0,12	0,10	0,13	0,09	0,11	0,13	0,09	0,00	0,13	0,21
10	0,10	0,09	0,12	0,13	0,14	0,11	0,13	0,11	0,13	0,00	0,21

Tabelle 15: Evaluierung der Methoden für die ökologische Inferenz, Ergebnisse der Simulationsstudie, weiterführende Evaluierung des Multinomial-Normal Modells mit dem ersten Datensatz von Sim 2, Konvergenz der Ketten, dargestellt ist der AD zwischen den Schätzungen, die letzte Spalte gibt die Schätzgüte an. Burn-In: 3000000, Thinning: 1000, Ziehungen: 7000.

Im Mittel beträgt der Abstand zwischen den Schätzungen der Modelle $\sim 0,11$. Es kann daher (noch) nicht von einer Konvergenz der Ketten ausgegangen werden. Dies bedeutet entweder, dass eine Gesamtkettenlänge von 10000000 noch zu niedrig ist und das Modell nur sehr langsam konvergiert, oder dass noch allgemeinere Konvergenzprobleme bestehen. In Abbildung 26 im Anhang findet sich ein Vergleich von jeweils drei Ketten zwei dieser Modelle. Die Schätzgüte, siehe letzte Spalte von Tabelle 15, ist jedoch bei allen Modellen in etwa vergleichbar und relativ gut. Die Kombination der Ketten aller 10 Modelle führt zu einer marginal besseren Schätzgüte.

Bei der Wahl einer anderen Partei als Referenzpartei, siehe Tabelle 16, zeigt sich ein ähnliches Bild wie bereits in Tabelle 12: Die Wahl der Referenzpartei beeinflusst

Ref-Partei	CSU	SPD	FDP	Grüne	DieLinke	Sonst	NW	Sim
CSU	0,00	0,26	0,28	0,19	0,18	0,25	0,26	0,29
SPD	0,26	0,00	0,21	0,18	0,21	0,23	0,22	0,24
FDP	0,28	0,21	0,00	0,25	0,27	0,23	0,25	0,23
Grüne	0,19	0,18	0,25	0,00	0,16	0,22	0,21	0,25
DieLinke	0,18	0,21	0,27	0,16	0,00	0,23	0,19	0,27
Sonst	0,25	0,23	0,23	0,22	0,23	0,00	0,18	0,27
NW	0,26	0,22	0,25	0,21	0,19	0,18	0,00	0,21

Tabelle 16: Evaluierung der Methoden für die ökologische Inferenz, Ergebnisse der Simulationsstudie, weiterführende Evaluierung des Multinomial-Normal Modells mit dem ersten Datensatz von Sim 2, Abhängigkeit von der Referenzpartei, dargestellt ist der AD zwischen den Schätzungen, die letzte Spalte gibt die Schätzgüte an. Burn-In: 3000000, Thinning: 1000, Ziehungen: 7000.

maßgeblich die Schätzung. Jedoch ist der durchschnittliche Abstand zwischen den Modellen mit $\sim 0,22$ hier ungefähr doppelt so groß. Inwieweit dies auf die identifizierten Konvergenzprobleme zurückgeführt werden kann, ist offen. Die Güte der Schätzungen unterscheidet sich auch recht deutlich zwischen den einzelnen Referenzparteien und liegt zwischen 0,21 und 0,29. Die beste Güte erreicht das Modell mit den Nichtwählern als Referenzpartei.

Die Ergebnisse der Simulation zeigen auf, dass es deutliche Unterschiede zwischen den Modellen bei der Schätzgüte gibt. Über alle vier Szenarien besitzt das Multinomial-Dirichlet Modell die beste Schätzgüte. In Sim 1 und Sim 2 wird es jeweils nur von einem Modell geschlagen, welche dann bei anderen Szenarien eine deutlich schlechtere Performance aufweisen. Die ökologische Regression nach Goodman hat nur bei Sim 1 eine gute Schätzgüte, während das Multinomial-Normal Modell vor allem bei Sim 1 und Sim 4 eine deutlich schlechtere Güte aufweist. Beim Multinomial-Normal Modell kommen zusätzlich noch die Abhängigkeit von der Referenzpartei und die offene Frage bezüglich der Konvergenz der Ketten hinzu. Deutlich absetzen kann sich das Multinomial-Dirichlet Modell bei Sim 3 und Sim 4, den beiden Szenarien, die als am realistischen bewertet werden können. Der AD zwischen den geschätzten Wählerwanderungen mit dem Multinomial-Dirichlet Modell und der bekannten Wählerwanderung liegt bei der Simulation zwischen 0,15 und 0,3, d. h. bis zu 15% der Wahlberechtigten werden bei der Schätzung einem falschen Übergang zugeordnet.

II.5. Ökologische Inferenz: Schlussbetrachtung

Die ökologische Inferenz hat weiterhin in der Forschung und in der Praxis ihre Daseinsberechtigung. In der Forschung ist es nicht immer möglich, auf Individualdaten zurückzugreifen. Entweder liegen diese nicht vor und können auch nicht mehr erhoben werden, ein typisches Problem bei der Analyse von vergangenen Ereignissen, oder es sprechen andere Gründe, z. B. rechtliche, politische oder finanzielle, gegen eine Erhebung. Außerhalb der Forschung sind es vor allem rechtliche und finanzielle Gründe, die gegen eine Erhebung von Individualdaten sprechen. Bei der Analyse von Wählerwanderungen steht aber mit den amtlichen Wahlergebnissen eine günstige Alternative mit guter Datenqualität zur Verfügung.

Die durchgeführten Betrachtungen zeigen aber, dass die Anwendung der ökologischen Inferenz nicht ohne Probleme ist. Ein grundlegendes Problem ist dabei auch die Frage nach dem besten Modell. Wie in Kapitel II.3. gezeigt, existiert mehr als ein Modell für die ökologische Inferenz. In Kapitel II.4.2.2. kann gezeigt werden, dass es möglich ist, selbst mit einem Modell mehrere alternative Wählerwanderungen zu schätzen, die sich dabei zum Teil deutlich voneinander unterscheiden. Ein ähnliches Bild zeigt sich beim Vergleich der Modelle untereinander. Allgemein ist jedoch die Güte der Schätzung nicht quantifizierbar, da die dafür notwendigen Informationen nicht vorhanden sind. Welches Modell und welche Schätzung sind daher zu bevorzugen?

Mit den vorgestellten Erkenntnissen ist eine Bewertung der Modelle möglich. Die ökologische Regression nach Goodman zeigt in der Simulation eine schlechte Schätzgüte, auch ist offen, was bei Schätzwerten außerhalb der logischen Grenzen geschehen soll. Bei Thomsens Probit/Logit Modell stellt neben der schlechten Schätzgüte in der Simulation auch die deutliche Abhängigkeit der Schätzung von der Referenzpartei ein Problem dar. Auch wenn Thomsen die Nichtwähler vorgeschlagen hat, gibt es keine echten Argumente, die für die eine oder andere Partei sprechen, d. h. bei diesem Modell gibt es miteinander konkurrierende und zum Teil sich deutlich unterscheidende Schätzungen. Aufgrund der identifizierten Probleme muss daher die Eignung dieser beiden Modelle zur Schätzung von Wählerwanderungen in Mehrparteiensystemen äußerst kritisch gesehen werden.

Das Multinomial-Normal Modell erreicht in einigen Szenarien der Simulation eine gute, einmal sogar die beste Schätzgüte. Bei anderen Szenarien ist die Schätzgüte deutlich schlechter. Jedoch implizieren die identifizierten Schwierigkeiten bei der Konvergenz der Ketten und die Abhängigkeit der Schätzung – wenn auch im schwächeren Maße als bei

Thomsens Probit/Logit Modell – von der Wahl der Referenzpartei, dass bei einer Nutzung des Modells sehr vorsichtig vorgegangen werden muss. Die Konvergenz des Modells sollte sichergestellt sein und auch die Auswirkung der Wahl der anderen Parteien als Referenzpartei untersucht werden. Falls die Wahl der Referenzpartei maßgeblich Einfluss auf die Schätzung nimmt, ist auch bei diesem Modell nicht klar, welche Schätzung zu präferieren ist.

Mit dem rekursiven Algorithmen und dem Multinomial-Dirichlet Modell kann jeweils eine gute Schätzung in der Simulation erreicht werden. Das Multinomial-Dirichlet Modell besitzt dabei in der Simulation über alle Szenarien die höchste Schätzung der verglichenen Modelle. Weiterhin zeigen sich bei den Voruntersuchungen bei den rekursiven Algorithmen Probleme, da unrealistisch hohe Loyalitätsraten mit den Modellen geschätzt werden. Aufgrund der insgesamt besseren Schätzung und der geringeren Schwierigkeiten bei der Anwendung ist daher das Multinomial-Dirichlet Modell zur Schätzung von Wählerwanderungen in Mehrparteiensystemen zu empfehlen.

III. Hybride Modelle

III.1. Einführung

Eine neuere Entwicklung bei der Schätzung von Wählerwanderungen, oder im Allgemeinen bei der Nutzung von Aggregatdaten, stellen die sogenannten hybriden Modelle dar. Es handelt sich bei diesem Modelltyp um eine Erweiterung bzw. Weiterentwicklung der ökologischen Inferenz, bei dem neben den Aggregatdaten auch Individualdaten bei der Schätzung verwendet werden. Durch die Hinzunahme von Individualdaten kann jedoch streng genommen nicht mehr in allen Fällen von einer ökologischer Inferenz bei diesen Modellen gesprochen werden.

Während bei der ökologischen Inferenz die Schätzung alleine auf aggregierten Daten basiert, werden bei hybriden Modellen Individual- und Aggregatdaten in einem gemeinsamen Modell kombiniert. D. h. auch, dass – zumindest bei einem Teil der Modelle – auch Informationen auf der interessierenden Ebene vorliegen und die Inferenz nicht notwendigerweise als Ebenen-überschreitend angesehen werden kann. Bei den vorgeschlagenen Modellen, die für die Schätzung von Wählerwanderungen geeignet sind und die im Folgenden vorgestellt werden, handelt es sich vor allem um Erweiterungen von hierarchisches Bayesianisches Modellen.

III.1.1. Kombination von Individual- und Aggregatdaten

Das grundlegende Prinzip der ökologischen Inferenz ist die Nutzung von Aggregatdaten zur Schätzung von Zusammenhängen auf einer niedrigeren Ebene. Bei der Schätzung von Wählerwanderungen wird auf die amtlichen Wahlergebnisse zurückgegriffen, um die Wählerwanderung von Individuen zu schätzen. Bei den zu schätzenden Größen handelt es sich um die inneren Zellen einer Kreuztabelle, bei der die Ränder die Wahlergebnisse der beiden betrachteten Wahlen repräsentieren (vgl. Tabelle 1). Aufgrund der Nicht-Überprüfbarkeit der Modellannahmen und der Abhängigkeit der Schätzungen vom verwendeten Modell, wie auch den genauen Festlegungen innerhalb eines Modells, ist eine Angabe der Schätzgüte im Allgemeinen nicht möglich. Zusätzlich kommt noch die fundamentale Unsicherheit über die Methode ökologische Inferenz, die seit Robinsons Artikel über den „ökologischen Fehlschluss“ besteht, hinzu. Im vorherigen Kapitel konnte aber bereits gezeigt werden, dass es möglich ist, mittels ökologischer Inferenz Wählerwanderungen zu schätzen. Als bestes Modell konnte dabei das Multinomial-Dirichlet Modell identifiziert werden.

Die Schätzung von Wählerwanderungen erfolgt alternativ auch anhand von Individualdaten. Diese können auf unterschiedliche Art und Weise, z. B. mit einer Telefonumfrage, einer Nachwahlbefragung oder im Rahmen einer Panelstudie, erhoben werden.¹⁵⁴ Jede dieser Erhebungsmethoden ist mit eigenen Schwierigkeiten behaftet.¹⁵⁵ Es ist z. B. offensichtlich, dass bei einer Telefonumfrage vor der Wahl nur die Wahlabsicht abgefragt werden kann, jedoch nicht das aktuelle Wahlverhalten bei der Wahl. Bei Nachwahlbefragungen, die direkt nach der Wahl durchgeführt werden, ist es möglich, die Wahlentscheidung zu erheben, wenn der Befragte korrekt Auskunft gibt. Jedoch ist auf diesem Wege nur ein Teil der Wahlberechtigten erreichbar, da die Nichtwähler der zweiten Wahl und die Briefwähler in den meisten Fällen nicht erfasst werden können. Panelstudien verfolgen eine bestimmte Personengruppe über die Zeit und die Validität der so erhobenen Daten hängt vor allem von der Zusammensetzung des Panels und dem wann und wie die Wahlentscheidung erhoben wird ab. Falls z. B. Personen mit höherer Loyalität mit geringere Wahrscheinlichkeit aus dem Panel ausscheiden, werden im Panel nach einiger Zeit Wahlberechtigte mit höherer Loyalität im Vergleich zum Beginn überrepräsentiert sein.

Neben diesen fundamentalen Schwierigkeiten bei der Erhebung der Daten gibt es bei der Schätzung der Wählerwanderung mit Individualdaten zusätzliche Schwierigkeiten, die mit der besonderen Ausgangslage zusammenhängen. Im Allgemeinen werden zwei Wahlen betrachtet, die mehrere Jahre auseinanderliegen. Abgesehen von Panelstudien, bei denen die gleiche Personengruppe bei beiden Wahlen befragt werden kann, muss bei Telefonumfragen und Nachwahlbefragungen die Wahlentscheidung bei der Vorwahl mittels Erinnerungsfrage erhoben werden. Erinnerungsfragen sind insofern nicht unproblematisch, da bekannt ist, dass mehrere Effekte zu einer falschen Angabe führen können.¹⁵⁶ Es ist z. B. möglich, dass ein Wahlberechtigter sich nicht mehr korrekt an seine letzte Wahlentscheidung erinnern kann und ersatzweise die aktuelle Wahlentscheidung auch für die letzte Wahl angibt.

In der Summe führen diese Schwierigkeiten dazu, dass auch die Schätzung der Wählerwanderung mittels Individualdaten als nicht problemfrei angesehen werden muss.¹⁵⁷

154 Vgl. Klima / Schlesinger / Küchenhoff / Thurner, 2016, S. 2.

155 Vgl. z. B. Greiner, D. James / Quinn, Kevin M.: Long Live the Exit Poll. *Daedalus*, 141(4), 2012, S. 15–18.

156 Vgl. Klima / Schlesinger / Küchenhoff / Thurner, 2016, S. 2.

157 In Deutschland wird bei Bundestagswahlen und bei zumindest einen Teil der Landtagswahlen eine Schätzung der Wählerwanderung von Infratest dimap durchgeführt. Die Forschungsgruppe Wahlen präsentiert keine Wählerwanderungsanalysen.

Zusätzlich gilt, wie bei der ökologischen Inferenz, dass die Schätzgüte, die bei der Schätzung aus Individualdaten fundamental von der im Allgemeinen nicht verifizierbaren Güte der erhobenen Daten abhängt, nicht überprüft werden kann. Während bei der ökologischen Inferenz vor allem die Frage offen ist, inwiefern der Schluss von den aggregierten Daten auf die Individualebene zulässig ist, stellt sich hier die Frage, inwiefern der Schluss von der erhobenen Stichprobe auf die gesamte Wählerschaft valide ist.

Aufgrund der jeweiligen Probleme der beiden möglichen Datenquellen scheint es naheliegend zu sein, beide Informationsquellen, d. h. Individual- und Aggregatdaten, zur Schätzung heranzuziehen. Die Hoffnung dabei ist, dass beide Informationsquellen zusammen eine bessere Schätzung erlauben und dass die jeweils mit den einzelnen Datenarten verbundenen Schwierigkeiten bei der Schätzung so weniger ins Gewicht fallen. Weiterhin spricht für eine Kombination, dass bei einer Schätzung mit Individual- und Aggregatdaten, aufgrund der vorhandenen Informationen in Letzteren, für eine adäquate Schätzung von Wählerwanderungen weniger Individualdaten als bei einer Hochrechnung alleine aus Individualdaten notwendig sein könnten und daher entsprechende Untersuchungen kostengünstiger durchzuführen wären. Die Kombination der beiden Datenquellen kann dabei auf vielfältige Art und Weise erfolgen.

Die Wählerstromanalysen von Infratest dimap basieren z. B. vor allem aus den Daten der am Wahltag durchgeführten Nachwahlbefragung. Es handelt sich daher primär um eine Schätzung aus Individualdaten. Zusätzlich werden aber Informationen aus anderen Datenquellen ergänzt, um die im Rahmen einer Nachwahlbefragung nicht zu erhebenden Größen bestimmen zu können. Dazu gehören z. B. Informationen aus Telefonumfragen, um die Nichtwähler bei der zweiten Wahl schätzen zu können. Weiterhin nutzt Infratest dimap jedoch auch die beiden Wahlergebnisse als Ränder der Tabelle, um die aus den Individualdaten geschätzten inneren Zellen zu adjustieren.¹⁵⁸ Es handelt sich daher nicht um eine rein auf Individualdaten basierende Schätzung.

Der Begriff hybrides Modell ist aber vor allem mit hierarchischen Bayesianischen Modellen verbunden, die auch bei der ökologischen Inferenz genutzt werden. Dabei kann es sich um ein spezifisches hybrides Modell handeln, oder um eine Erweiterung eines Modells für die ökologische Inferenz, bei welchen es möglich ist, auch Individualdaten mit bei der Schätzung zu berücksichtigen. Eine entsprechende Erweiterung wurde 2004 von Jon Wakefield für sein 2x2 Modell vorgeschlagen. Bei diesem Modell werden bei Gebieten mit Individualdaten beide Datenquellen genutzt, während die Gebiete ohne

158 Vgl. Hilmer / Kunert, 2005, S. 140–144.

Individualdaten wie bei der ökologischen Inferenz in das Modell eingehen.¹⁵⁹

Bei der Betrachtung von hybriden Modellen ist es jedoch notwendig, zwei verschiedene Modelltypen zu unterscheiden. Die beiden Modelltypen unterscheiden sie vor allem in der Art der vorliegenden Informationen auf der Individualebene. Es gibt Modelle, wie das 2004 von Wakefield vorgeschlagene, bei dem die Individualdaten den gleichen Zusammenhang beschreiben, der auch mittels ökologischer Inferenz geschätzt werden soll. Bei der Schätzung von Wählerwanderungen würde dies dem individuellen Wahlverhalten bei beiden Wahlen entsprechen. Da in diesem Fall die Individualdaten den zu schätzenden Zusammenhang beschreiben, wird mit diesen hybriden Modellen, streng genommen, keine ökologische Inferenz durchgeführt.

Vor allem in der Epidemiologie ist aber auch eine andere Art von hybriden Modellen bekannt. Die Individualdaten beschreiben bei diesen Modellen nicht den interessierenden Zusammenhang, sondern beschreiben die Struktur einer Kovariablen oder das gemeinsame Vorkommen von Kovariablen. Bei epidemiologischen Untersuchungen kann z. B. das Risiko an einer bestimmten Krankheit zu erkranken mittels mehrere Kovariablen, den vermuteten Risikofaktoren, geschätzt werden. Wenn nur ökologische Daten vorliegen, ist jeweils nur der Anteil der Bevölkerung bekannt, der eine bestimmte Ausprägung einer kategorialen Kovariablen hat. Bei metrischen Kovariablen ist wiederum nur ein Mittelwert für das Gebiet, aber nicht notwendigerweise die Streuung auf Individualebene bekannt. Weiterhin geht aus den Aggregatdaten meistens nicht hervor, welche Zusammenhänge zwischen den Kovariablen auf der Individualebene bestehen. Ein entsprechendes Modell wurde z. B. von Salway und Wakefield vorgeschlagen¹⁶⁰. Da bei diesen Modellen der interessierende Zusammenhang nicht von den Individualdaten beschrieben wird, erfolgt mit diesen hybriden Modellen weiterhin eine ökologische Inferenz.

Für die Schätzung von Wählerwanderungen ist vor allem der erste Typ der hybriden Modelle von Interesse, d. h., es wird auf Individualebene das Wahlverhalten bei beiden Wahlen erhoben. Aus diesem Grund konzentrieren sich die folgenden Darstellungen auch nur auf diesen Modelltyp. Im nächsten Kapitel erfolgt eine formale Einführung dieser Art von hybriden Modellen. Dafür wird auf die Darstellung von Wakefield zurückgegriffen¹⁶¹, diese jedoch auf den RxC Fall und den Anwendungsfall Schätzung von Wählerwanderung übertragen. Zwei hybriden Modelle, die zur dieser Modellklasse gehören, werden anschließend kurz eingeführt.

159 Vgl. Wakefield, 2004, S. 418–421.

160 Vgl. Salway / Wakefield, 2008, S. 2–3 und S. 7–9.

161 Vgl. Wakefield, 2004, S. 418.

III.1.2. Darstellung des Schätzproblems bei hybriden Modellen

Bei der Schätzung von Wählerwanderungen mit hybriden Modellen ist die Ausgangssituation zum Teil direkt vergleichbar mit der bei der Anwendung der ökologischen Inferenz in diesen Kontext. Erneut werden R Parteien bei der Wahl zum Zeitpunkt t und C Parteien bei der Wahl zum Zeitpunkt $t+1$ betrachtet. Für beide Wahlen liegen in N Gebietseinheiten die Wahlergebnisse vor. Es handelt sich prinzipiell um die gleichen Aggregatdaten, die auch bei der ökologischen Inferenz für die Schätzung verwendet werden.

	CSU 2013	SPD 2013	...	NW 2013	Wahl 2009
CSU 2009			...		345
SPD 2009			...		199
...
NW 2009			...		380
Wahl 2013	321	201	...	377	1073

Tabelle 17: Darstellung der Datensituation bei der Anwendung der hybriden Modelle bei der Schätzung von Wählerwanderungen, gezeigt werden die Aggregatdaten im Gebiet i , die Ränder entsprechen den Wählern bei der jeweiligen Wahl, die hellgrau markierten Ränder sind bekannt, die weißen inneren Zellen unbekannt. Es handelt sich um fiktive Daten.

Die vorliegenden Aggregatdaten lassen sich, siehe Tabelle 17, erneut in Tabellenform darstellen. Für die Darstellung werden hier die absoluten Zahlen verwendet, während in Tabelle 1 relative Werte genutzt werden. Die Ränder der Tabelle sind erneut bekannt, während die inneren Zellen, die die interessierenden Größen repräsentieren, erneut unbekannt sind. Entsprechende Tabellen können für alle N Gebietseinheiten aufgestellt werden.

Zusätzlich zu den Aggregatdaten liegen auch noch Befragungsdaten vor, welche die zweite Datenquelle für die hybriden Modelle darstellen. Es sei für diese Darstellung angenommen, dass diese ebenfalls auf der Ebene der einzelnen Gebiete vorliegen. Weiterhin wird davon ausgegangen, dass alle Befragten die Wahlentscheidung bei beiden Wahlen angeben konnten und alle Wahlberechtigten bei der Erhebung auch theoretisch erreicht werden konnten. Tabelle 18 zeigt die Befragungsdaten für das Gebiet i :

	CSU 2013	SPD 2013	...	NW 2013	Wahl 2009
CSU 2009	138	8	...	26	211
SPD 2009	12	87	...	14	147
...
NW 2009	20	16	...	156	235
Wahl 2013	195	151	...	241	681

Tabelle 18: Darstellung der Datensituation bei der Anwendung der hybriden Modelle bei der Schätzung von Wählerwanderungen, gezeigt werden die Individualdaten im Gebiet i , die Zahlen entsprechen den befragten Wählern mit der jeweiligen Kombination/Partei, die hellgrau markierten inneren Zellen sind bekannt. Es handelt sich um fiktive Daten.

Bei den Individualdaten sind die inneren Zellen und die Ränder der Tabelle bekannt, d. h. keine Zelle der Tabelle ist unbekannt. In den meisten Fällen wird jedoch nur ein Teil der Wahlberechtigten befragt werden, so dass die Ränder bei den Individualdaten nicht den Rändern bei den Aggregatdaten entsprechen. Im Allgemeinen werden auch nicht in allen N Gebiete Individualdaten erhoben, entsprechende Informationen auf Individualebene liegen daher für $M \leq N$ Gebiete vor. Bei Nachwahlbefragungen ist M üblicherweise deutlich kleiner als N .

Eine grundlegende Frage bei hybriden Modellen ist, wie diese beiden Informationsquellen sinnvoll in einem Modell miteinander verbunden werden können. Dabei muss bei den Daten berücksichtigt werden, dass die für die Individualdaten befragten Wahlberechtigten auch Teil der Aggregatdaten sind. Dies bedeutet, dass die für Tabelle 18 befragten Wahlberechtigten eine Teilmenge der in Tabelle 17 betrachteten Wahlberechtigten darstellen. Es erscheint daher nicht sinnvoll, die vorliegenden Daten direkt ohne weitere Aufbereitung zu nutzen.

Wakefield schlug 2004 eine Erweiterung seines hierarchischen Bayesianischen 2x2 Modells für die ökologische Inferenz vor. Diese Modellerweiterung erlaubt die Berücksichtigung von Individualdaten bei der Schätzung, es handelt sich daher um ein hybrides Modell. Wakefield zieht bei diesem Modell die erhobenen Individualdaten von den Aggregatdaten ab.¹⁶² Dieser Datenaufbereitungsschritt schließt aus, dass es zu einer doppelten Berücksichtigung von Wahlberechtigten bei der Schätzung kommt. Nach einer entsprechenden Korrektur der Ränder bei den Aggregatdaten enthalten diese nur noch die Wahlentscheidung der Wahlberechtigten, für die keine Informationen in den Individualdaten vorliegen. Allgemein gilt daher folgendes für ein Gebiet i , für das Individualdaten und Aggregatdaten vorliegen,¹⁶³ nach der Datenaufbereitung:

Individualdaten					Aggregatdaten				
	CSU 2013	...	NW 2013	Summe		CSU 2013	...	NW 2013	Summe
CSU 2009	$M_{CSU,CSU,i}$...	$M_{CSU,NW,i}$	$M_{CSU,2009,i}$	CSU 2009				$N_{CSU,2009,i}$ - $M_{CSU,2009,i}$
...
NW 2009	$M_{NW,CSU,i}$...	$M_{NW,NW,i}$	$M_{NW,2009,i}$	NW 2009				$N_{NW,2009,i}$ - $M_{NW,2009,i}$
Summe	$M_{CSU,2013,i}$...	$M_{NW,2013,i}$	M_i	Summe	$N_{CSU,2013,i}$ - $M_{CSU,2013,i}$...	$N_{NW,2013,i}$ - $M_{NW,2013,i}$	$N_i - M_i$

Tabelle 19: Datenaufbereitung beim hybriden Modell nach Wakefield in einem Gebiet i mit Individualdaten (links) und Aggregatdaten (rechts), die hellgrau markierten Zellen sind bekannt oder können mit den vorhandenen Informationen berechnet werden. Darstellung angepasst auf den Anwendungsfall Wählerwanderung in Mehrparteiensystemen (RxC Tabellen).

162 Vgl. Wakefield, 2004, S. 418.

163 Vgl. Wakefield, 2004, S. 418.

Die Darstellung in Tabelle 19 zeigt auch direkt auf, was geschieht, wenn für ein Gebiet keine Individualdaten vorhanden sind. Die Tabelle für die Individualdaten enthält in allen Zellen null und es ist keine Anpassung der Ränder bei den Aggregatdaten notwendig. Umgekehrt gilt, dass bei einer Befragung aller Wahlberechtigten durch diese Anpassung die Ränder bei den Aggregatdaten auf null gesetzt werden.

Tabelle 20 zeigt Tabelle 17 nach der Datenaufbereitung. Die Werte in den Rändern sind deutlich niedriger, da für fast zwei Drittel der Wahlberechtigten Informationen in den Individualdaten vorhanden sind. Z. B. liegt für 211 der 345 CSU-Wähler des Jahres 2009 die Information über das Wahlverhalten bei der zweiten Wahl im Jahr 2013 in den Individualdaten vor. Für 134 2009er CSU-Wähler ist das Wahlverhalten bei der zweiten Wahl weiterhin unbekannt.

	CSU 2013	SPD 2013	...	NW 2013	Wahl 2009
CSU 2009			...		134
SPD 2009			...		52
...
NW 2009			...		145
Wahl 2013	126	50	...	136	392

Tabelle 20: Darstellung der Datensituation bei der Anwendung der hybriden Modelle bei der Schätzung von Wählerwanderungen, gezeigt werden die Aggregatdaten nach der Datenaufbereitung im Gebiet i , die Ränder entsprechen den Wählern bei der jeweiligen Wahl, die nicht befragt wurden, die hellgrau markierten Ränder sind bekannt, die weißen inneren Zellen unbekannt. Es handelt sich um fiktive Daten.

Diese Art der Datenaufbereitung setzt zwingend voraus, dass die Gebietszugehörigkeit bei den für die Individualdaten befragten Wahlberechtigten bekannt ist. Ohne dieses Wissen ist eine Anpassung der Ränder bei den Aggregatdaten nicht möglich. Bei Nachwahlbefragungen vor den Wahllokalen kann in Allgemeinen davon ausgegangen werden, dass die Gebietszugehörigkeit bekannt ist, während dies z. B. bei Telefonumfragen weniger gesichert¹⁶⁴ ist.

Wenn Individual- und Aggregatdaten auf Ebene der Gebiete vorliegen, ist es auch möglich, sie bei den Modellen direkt auf dieser Ebene zu verbinden. Wakefield verknüpft z. B. bei seinem hybriden Modell die beiden Datenquellen, indem er eine gemeinsame Likelihood auf Gebietsebene aufstellt.¹⁶⁵

$$L(\beta_{CSU2009, CSU2013, i}, \dots, \beta_{NW2009, NW2013, i}) = \prod_{r=1}^R p(M_{r, CSU2013, i}, \dots, M_{r, NW2013, i} | \beta_{r, CSU2013, i}, \dots, \beta_{r, NW2013, i}) * p(Aggregatdaten | \beta_{CSU2009, CSU2013, i}, \dots, \beta_{NW2009, NW2013, i}) \quad (33)$$

164 Bei Bundestagswahlen würde dies z. B. bedeuten, dass der Befragte seinen Wahlbezirk selber benennen kann oder dieser aus gegebenen Informationen hergeleitet werden kann.

165 Vgl. Wakefield, 2004, S. 418.

Gleichung (33) zeigt die so gebildete Likelihood für das Gebiet i . Diese besteht aus zwei Komponenten. Die erste Komponente stammt von den Individualdaten, bei Gleichung (33) handelt es sich um das Produkt der Verteilungen der Individualdaten für jede Zeile. Die zweite Komponente stammt von den Aggregatdaten. Wenn beide Informationsquellen vorliegen, beinhaltet die Likelihood für ein Gebiet beide Komponenten. Wenn aber z. B. keine Individualdaten vorhanden sind, besteht die Likelihood für dieses Gebiet nur aus der Aggregatdatenkomponente und wäre identisch mit der Likelihood für dieses Gebiet bei der ökologischen Inferenz.

In Fällen, wo eine genaue Zuordnung der Individualdaten zum jeweiligen Gebiet nicht möglich ist, können die Individual- und Aggregatdaten nur in der erhobenen Form verwendet werden. Ein Teil der Wahlberechtigten würde daher doppelt im Modell bei der Schätzung berücksichtigt werden. Zusätzlich resultiert daraus auch eine Einschränkung bei der Kombination der beiden Datenquellen, da diese nicht wie dargestellt auf der Ebene der Gebiete möglich ist. Die Individualdaten könnten jedoch als zusätzliches Gebiet aufgenommen werden, welches nur Individualdaten umfasst und dessen Likelihood auch nur diese Komponente beinhaltet.

Für die Verknüpfung von Individual- und Aggregatdaten ist weiterhin eine Annahme über die zu schätzenden Parameter, bei Wählerwanderungen handelt es sich dabei oft um die Übergangswahrscheinlichkeiten, in beiden Datenarten notwendig. Wakefield geht bei seinem Modell von identischen Übergangswahrscheinlichkeiten innerhalb jedes Gebiets bei Individual- und Aggregatdaten aus.¹⁶⁶ Im Falle einer Zufallsstichprobe ist diese Annahme durchaus plausibel, jedoch kann die Annahme bei anderen Erhebungsarten problematisch werden. Eine Annahme von identischen Übergangswahrscheinlichkeiten ist jedoch nicht zwingend, es ist nur notwendig, dass es eine definierte Struktur zwischen den Übergangswahrscheinlichkeiten in beiden Datenarten gibt und so beide miteinander in Verbindung gebracht werden können.

Falls aber von komplett unterschiedlichen Strukturen bei den Übergangswahrscheinlichkeiten zwischen Individual- und Aggregatdaten ausgegangen werden muss und diese sich auch nicht sinnvoll in Verbindung zueinander bringen lassen, scheint ein gemeinsames Modell nicht sinnvoll zu sein. Die jeweils eine Datenart beinhaltet keine echten Informationen für die interessierenden Parameter der anderen Datenart, sie kann daher die Schätzung dieser Werte nicht verbessern.

¹⁶⁶ Vgl. Wakefield, 2004, S. 418.

III.2. Methodenbeschreibung

In diesem Kapitel werden zwei hybride Modelle kurz vorgestellt, die zur Schätzung von RxC Tabellen und daher für die Schätzung von Wählerwanderungen in Mehrparteiensystemen geeignet sind. Bei beiden Modellen handelt es sich um hierarchische Bayesianische Modelle. Die Modelle stellen Erweiterungen der beiden in Kapitel II.3.4. vorgestellten Modelle für die ökologische Inferenz im RxC Fall dar. Ohne Individualdaten gibt es keine Unterschiede zwischen den hier vorgestellten hybriden Modellen und den Modellen für die ökologische Inferenz.

III.2.1. Hybrides Multinomial-(log-)Normal Modell

Das hybride Multinomial-Normal Modell stellt eine Erweiterung des in Kapitel II.3.4.3. vorgestellten Multinomial-Normal Modells dar. Die Möglichkeit, dass Modell entsprechend zu erweitern, wurde von den Autoren bereits 2009 angesprochen und kurz diskutiert.¹⁶⁷ 2010 stellten Greiner und Quinn explizit ihr hybrides Multinomial-Normal Modell vor, dass auf dem 2009 vorgestellten Modell für die ökologische Inferenz basiert.¹⁶⁸

Für das hybride Modell nehmen Greiner und Quinn an, dass es sich bei den Individualdaten um eine Zufallsstichprobe handelt und dass eine exakte Zuordnung zu einem Gebiet möglich ist. Für ein Teil der Gebiete liegen daher zusätzlich zu den Aggregatdaten Individualdaten in der Form von Tabelle 18 vor. Die angenommene Zufallsstichprobe stellt dabei nur eine explizite Annahme für das vorgestellte Modell dar, es handelt sich nicht um eine notwendige allgemeine Anforderung an dieses Modell. Greiner und Quinn zufolge reicht es aus, wenn die Wahrscheinlichkeit angegeben werden kann, eine bestimmte Stichprobe gegeben den unbekannt inneren Zellen, zu beobachten.¹⁶⁹

Vor der eigentlichen Schätzung erfolgt auch bei diesem Modell ein Datenaufbereitungsschritt. Es werden, wie von Wakefield vorgeschlagen, die beobachteten Individualdaten von den Aggregatdaten abgezogen und im Modell werden bei den Aggregatdaten nur noch die Wahlberechtigten berücksichtigt, für die keine Informationen in den Individualdaten vorhanden sind.¹⁷⁰

Das eigentliche Modell entspricht dann weitestgehend dem Modell für die ökologische

167 Vgl. Greiner / Quinn, 2009, S. 77–79.

168 Vgl. Greiner, D. James / Quinn, Kevin M.: Exit Polling and Racial Bloc Voting: Combining Individual-Level and R x C Ecological Data. *The Annals of Applied Statistics*, 4(4), 2010.

169 Vgl. Greiner / Quinn, 2010, S. 1780.

170 Vgl. Greiner / Quinn, 2010, S. 1781.

Inferenz. Erneut werden die inneren Zellen betrachtet und für die Aggregatdaten wird für jede Zeile eine Multinomialverteilung angenommen. Z. B. gilt für die Zeile der CSU:

$$\begin{aligned} & (N_{CSU\ 2009,CSU\ 2013,i}^*, \dots, N_{CSU\ 2009,NW\ 2013,i}^*) \\ & \sim MNL(\beta_{CSU2009,CSU2013,i}, \dots, \beta_{CSU2009,NW2013,i}, N_{CSU,2009,i}^*) \end{aligned} \quad (34)$$

mit

$$\begin{aligned} N_{CSU\ 2009,CSU\ 2013,i}^* &= N_{CSU\ 2009,CSU\ 2013,i} - M_{CSU\ 2009,CSU\ 2013,i} \\ N_{CSU,2009,i}^* &= N_{CSU,2009,i} - M_{CSU,2009,i} \end{aligned} \quad (35)$$

Die Verteilungsannahme für die Aggregatdaten ist daher, unabhängig davon ob Individualdaten vorliegen, für jedes Gebiet identisch. Ein Unterschied findet sich jedoch im Geltungsbereich der Verteilungen. Wenn keine Individualdaten vorliegen, handelt es sich um die Verteilung aller Wahlberechtigten im Gebiet. Falls Individualdaten vorliegen, beschreibt Verteilung (34) nur noch die Wahlberechtigten, für die keine Informationen in den Individualdaten vorhanden sind.

Die im Modell explizit angenommene Verteilung für die Individualdaten in einem Gebiet ist schwerer anzugeben. Für das vorgestellte Modell wird bei den Individualdaten eine Zufallsstichprobe angenommen. Greiner und Quinn geben zwar die Wahrscheinlichkeit an, eine bestimmte Stichprobe zu beobachten, die dargestellte Gleichung des Modells enthält aber keine Komponente, die sich eindeutig auf die Verteilung der Individualdaten bezieht.¹⁷¹ In der Gleichung findet sich nur eine Komponente, die die Verteilung der Aggregatdaten (34) berücksichtigt.¹⁷² Die Diskussion möglicher Erweiterungen ihres Modells im Jahr 2009 legt jedoch nahe, dass auch für die Individualdaten eine Multinomialverteilung angenommen oder zumindest in Betracht gezogen wird.¹⁷³

$$\begin{aligned} & (M_{CSU\ 2009,CSU\ 2013,i}, \dots, M_{CSU\ 2009,NW\ 2013,i}) \\ & \sim MNL(\beta_{CSU2009,CSU2013,i}, \dots, \beta_{CSU2009,NW2013,i}, M_{CSU,2009,i}) \end{aligned} \quad (36)$$

Die Annahme einer Zufallsstichprobe bei den Individualdaten impliziert auch, dass für Individual- und Aggregatdaten die gleichen Übergangswahrscheinlichkeiten angenommen werden. Dies bedeutet, dass die $\beta_{r,c,i}$ bei Aggregatdaten (34) und Individualdaten (36) als

171 Die Darstellung des Modells auf S. 1782 in Greiner / Quinn, 2010 scheint unvollständig zu sein. Im Text beziehen sich die Autoren auf zwei (von vier) Zeilen, die der Multinomialverteilungsannahme der inneren Zellen entsprechen sollen. Die entsprechende Formel besitzt jedoch nur drei Zeilen, von denen sich eindeutig nur eine auf die Multinomialverteilungsannahme bezieht. Auch nimmt der Text Bezug auf einen Bruch, der in der eigentlichen Gleichung nicht in dieser Form vorkommt.

172 Vgl. Greiner / Quinn, 2010, S. 1782.

173 Vgl. Greiner / Quinn, 2009, S. 78.

identisch angenommen werden.

Auf der zweiten Ebene des Modells nehmen Greiner und Quinn erneut eine multivariate Normalverteilung für die Logit-transformierten Übergangswahrscheinlichkeiten an. Die durchgeführte Transformation ist dabei identisch zu (27). Z. B. gilt für die Übergangswahrscheinlichkeiten von der CSU zu den anderen Parteien:¹⁷⁴

$$\begin{aligned} & (\omega_{CSU2009, CSU2013, i}, \dots, \omega_{CSU2009, Sonst2013, i}) \\ & = \left(\log \left(\frac{\beta_{CSU2009, CSU2013, i}}{\beta_{CSU2009, NW2013, i}} \right), \dots, \log \left(\frac{\beta_{CSU2009, Sonst2013, i}}{\beta_{CSU2009, NW2013, i}} \right) \right) \end{aligned} \quad (37)$$

Erneut muss für die Transformation in (37) eine Referenzpartei festgelegt werden. Auch für das hybride Modell wird die Nutzung der „Enthalter“, bei der Wählerwanderung wären dies die Nichtwähler (NW), als naheliegend vorgeschlagen.¹⁷⁵

Auf der dritten Ebene des Modells wird auch beim hybriden Multinomial-Normal Modell, vollkommen analog zum Multinomial-Normal Modell, für die Erwartungswerte der Normalverteilung der Logit-transformierten Übergangswahrscheinlichkeiten als Hyperpriori eine Normalverteilung angenommen. Die Hyperpriori der Varianzkomponente ist erneut eine inverse Wishart-Verteilung. Die in Kapitel II.3.4.3. erwähnten Einschränkungen bei der Wahl der Priori-Parameter der Hyperprioris in der aktuellen Softwareimplementierung im R-Paket RxCEcolInf¹⁷⁶ gelten auch für das hybride Modell.

III.2.2. Hybrides Multinomial-Dirichlet Modell

Das hybride Multinomial-Dirichlet Modell stellt eine Erweiterung des in Kapitel II.3.4.2. vorgestellten Multinomial-Dirichlet Modells dar. Das Modell für die ökologische Inferenz wurde erstmals 2001 vorgestellt, die hybride Erweiterung des Modells existiert seit dem Jahr 2013. Die technische Umsetzung des Modells in R erfolgte durch Thomas Schlesinger.¹⁷⁷

Die Motivation für die Erweiterung des Multinomial-Dirichlet Modells ist in den zuvor beschriebenen Schwierigkeiten beim Multinomial-Normal Modell zu finden.¹⁷⁸ Bereits von Greiner und Quinn wird die längere Rechenzeit ihres Multinomial-Normal Modells im Vergleich zum Multinomial-Dirichlet Modell angesprochen.¹⁷⁹ Dies zeigte sich auch bei unserer Nutzung ihres Modells. Erschwerend kommt noch hinzu, dass die Schätzung mit

174 Vgl. Greiner / Quinn, 2010, S. 1781–1782.

175 Vgl. Greiner / Quinn, 2010, S. 1781.

176 Greiner / Baines / Quinn, 2013.

177 Vgl. Schlesinger, 2013. Die Betreuung der Masterarbeit erfolgte mit durch den Autor.

178 Vgl. Kapitel II.4.2.2..

179 Greiner / Quinn, 2010, S. 1780.

dem Multinomial-Normal Modell möglicherweise instabil sein kann, d. h. es kommt nicht notwendigerweise zu einer Konvergenz der Ketten gegen eine Verteilung, und dass die Schätzung von der Wahl der Referenzpartei abhängt.¹⁸⁰ Vergleichbare Probleme können beim Multinomial-Dirichlet Modell nicht identifiziert werden. Aus diesen Gründen erscheint es sinnvoll, auch das Multinomial-Dirichlet Modell so zu erweitern, dass es als hybrides Modell genutzt werden kann und als Alternative zum hybriden Multinomial-Normal Modell zur Verfügung steht.

Für das hybride Multinomial-Dirichlet Modell wird wie beim hybriden Multinomial-Normal Modell angenommen, dass die Individualdaten auf Gebietsebene vorliegen, d. h. es ist möglich, sie in der Form von Tabelle 18 darzustellen. Ebenfalls wird auch angenommen, dass es keine strukturellen Unterschiede zwischen den Übergangswahrscheinlichkeiten bei Individual- und Aggregatdaten innerhalb eines Gebietes gibt. Bei einer Zufallsstichprobe wäre diese Annahme z. B. erfüllt. Die Datenaufbereitung erfolgt nach dem in Tabelle 19 dargestellten Prinzip und entspricht somit der von Wakefield¹⁸¹ vorgeschlagenen Vorgehensweise. Es werden daher auch bei diesem Modell nur noch die Wahlberechtigten bei den Aggregatdaten betrachtet, für die keine Information in den Individualdaten vorliegen.¹⁸²

Für die Verteilung der Aggregatdaten, beim Modell der Spaltenrand, der dem Wahlergebnis bei der zweiten Wahl entspricht, wird wie beim Modell für die ökologische Inferenz eine Multinomialverteilung in jedem Gebiet angenommen.¹⁸³

$$(N_{CSU\ 2013,i}^*, N_{SPD\ 2013,i}^*, \dots, N_{NW\ 2013,i}^*) \sim MNL(\theta_{CSU,i}, \theta_{SPD,i}, \dots, \theta_{NW,i}, N_i^*) \quad (38)$$

mit

$$\begin{aligned} N_{CSU,2013,i}^* &= N_{CSU,2013,i} - M_{CSU,2013,i} \\ N_i^* &= N_i - M_i \end{aligned} \quad (39)$$

und z. B.:

$$\begin{aligned} \theta_{CSU,i} &= \beta_{CSU2009,CSU2013,i} * P_{CSU,2009,i}^* + \dots + \beta_{NW2009,CSU2013,i} * P_{NW,2009,i}^* \\ &= \sum_{r=1}^R \beta_{r,CSU2013,i} * P_{r,2009,i}^* \end{aligned} \quad (40)$$

mit

180 Vgl. Kapitel II.4.3.2..

181 Vgl. Wakefield, 2004, S. 418.

182 Vgl. Klima / Schlesinger / Küchenhoff / Thurner, 2016, S. 6–7.

183 Vgl. Klima / Schlesinger / Küchenhoff / Thurner, 2016, S. 6–7.

$$P_{r,2009,i}^* = \frac{N_{r,2009,i} - M_{r,2009,i}}{N_i - M_i}, \quad \forall r \in \{1, \dots, R\} \quad (41)$$

Die Anpassung in Gleichung (39) stellt die in Tabelle 19 veranschaulichte Datenaufbereitung dar. Falls keine Individualdaten im Gebiet vorliegen, ist (38) identisch zu (20). Wenn Individualdaten für das Gebiet vorliegen, werden nur noch die Wahlberechtigten berücksichtigt, für die keine Informationen in den Individualdaten vorhanden sind und Verteilung (38) bezieht sich nur noch auf diese. Gleichung (40), die die Beziehung zwischen den Übergangswahrscheinlichkeiten und den Parametern in Verteilung (38) beschreibt, ist fast identisch zu der entsprechenden Gleichung (21) beim Multinomial-Dirichlet Modell, jedoch erfolgt auch hier bei vorhandenen Individualdaten eine Anpassung, die in Gleichung (41) dargestellt wird.

Für die Individualdaten wird für jede Zeile, d. h. für jede Partei bei der ersten Wahl, eine Multinomialverteilung für die inneren Zellen angenommen. Dabei werden die einzelnen Multinomialverteilungen für die einzelnen Parteien als unabhängig voneinander betrachtet. Für die erste Zeile in Tabelle 18, die in den Individualdaten vorhandenen Wähler der CSU bei der Wahl im Jahr 2009, wird daher folgende Verteilung angenommen:¹⁸⁴

$$\left(M_{CSU\ 2009,CSU\ 2013,i}, \dots, M_{CSU\ 2009,NW\ 2013,i} \right) \sim MNL(\beta_{CSU\ 2009,CSU\ 2013,i}, \dots, \beta_{CSU\ 2009,NW\ 2013,i}, M_{CSU,2009,i}) \quad (42)$$

Dabei wird, wie bereits erwähnt, angenommen, dass die Übergangswahrscheinlichkeiten in Gleichung (40) und bei den Verteilungen für die Individualdaten, Verteilung (42), innerhalb eines Gebietes identisch sind.

Das Zusammenführen der beiden Informationsquellen in einem Gebiet erfolgt analog zu den in Gleichung (33) dargestellten Vorschlag von Wakefield. Die Likelihood für ein Gebiet ist das Produkt aus der Likelihood der Individualdaten, die ein Produkt der Likelihoods der Verteilungen der einzelnen Zeilen darstellt, und der Likelihood der Aggregatdaten.¹⁸⁵ Dieses Vorgehen impliziert, dass die Likelihood eines Gebiets ohne Individualdaten beim hybriden Modell identisch zur Likelihood bei der ökologischen Inferenz ist.

Auf der zweiten Ebene wird beim hybriden Multinomial-Dirichlet Modell, wie auch zuvor beim Multinomial-Dirichlet Modell, eine gemeinsame Dirichlet-Verteilung für die Übergangswahrscheinlichkeiten in einer Zeile angenommen. Für die erste Zeile in einem Gebiet i gilt daher:

¹⁸⁴ Vgl. Klima / Schlesinger / Küchenhoff / Thurner, 2016, S. 6–7.

¹⁸⁵ Vgl. Klima / Schlesinger / Küchenhoff / Thurner, 2016, S. 8 und Wakefield, 2004, S. 418.

$$\begin{aligned} & (\beta_{CSU2009, CSU2013, i}, \beta_{CSU2009, SPD2013, i}, \dots, \beta_{CSU2009, NW2013, i}) \\ & \sim \text{Dirichlet}(\alpha_{CSU2009, CSU2013}, \alpha_{CSU2009, SPD2013}, \dots, \alpha_{CSU2009, NW2013}) \end{aligned} \quad (43)$$

Verteilung (43) ist die gemeinsame Verteilung der Übergangswahrscheinlichkeiten für alle N Gebiete. Die gemeinsamen Verteilungen der einzelnen Zeilen werden dabei erneut a priori als unabhängig voneinander angenommen. Auf dieser Ebene des Modells gibt es daher keinen Unterschied zwischen dem hybriden Modell und dem Modell für die ökologische Inferenz, Verteilung (43) ist identisch mit Verteilung (22).

Auf der dritten Ebene des Modells, den Hyperprioris der Parameter der gemeinsamen Verteilungen der Übergangswahrscheinlichkeiten, wird beim hybriden Modell allgemein eine Gamma-Verteilung angenommen. Das hybride Modell folgt daher der Implementierung von Lau, Moore und Kellermann.¹⁸⁶

$$\alpha_{r,c} \sim \text{Gamma}(\lambda_{r,c,1}, \lambda_{r,c,2}), \quad \forall r \in \{1, \dots, R\}, c \in \{1, \dots, C\} \quad (44)$$

Das Modell erlaubt jedoch eine zellspezifische Wahl der Hyperparameter und somit auch die Berücksichtigung von zellspezifischen Vorwissen.¹⁸⁷ Die Softwareimplementierung im R-Paket `eiwild`¹⁸⁸ erlaubt zusätzlich noch die einfache Spezifikation eines für alle Zellen identischen Vorwissens.

186 Vgl. Klima / Schlesinger / Küchenhoff / Thurner, 2016, S. 7 und Lau / Moore / Kellermann, 2007, S. 46.

187 Vgl. Kapitel II.3.4.4.

188 Schlesinger, 2014.

III.3. Simulationsstudie Evaluierung hybrider Modelle

Die Kombination von Individual- und Aggregatdaten bei hybriden Modellen erfolgt in der Hoffnung, dass im Vergleich zur Schätzung mit nur einer Datenquelle eine bessere Schätzgüte erreicht werden kann. Für das hybride Multinomial-Normal Modell konnte bereits durch Greiner und Quinn gezeigt werden, dass die Berücksichtigung von Individualdaten bei der Schätzung beim hybriden Modell zur einer Verbesserung der Schätzgüte führt. Dies ist nicht nur im Vergleich zur ökologischen Inferenz mit dem Multinomial-Normal Modell der Fall, sondern auch fast immer im Vergleich zur Schätzung mit den Individualdaten.¹⁸⁹

In der folgenden Simulationsstudie,¹⁹⁰ die sich grob an der von Greiner und Quinn orientiert, soll die Eignung des hybriden Multinomial-Dirichlet Modells detailliert evaluiert werden. Es soll dabei primär überprüft werden, ob mit dem hybriden Modell eine bessere Schätzgüte erreicht werden kann. Als Alternativen zum hybriden Modell werden in der Studie dabei die ökologische Inferenz und die einfache Hochrechnung aus den Individualdaten berücksichtigt. Für die Individualdaten selber werden insgesamt drei Erhebungsmethoden in Betracht gezogen. Weiterhin ist bekannt, dass bei der ökologischen Inferenz auch die Verteilung der Übergangswahrscheinlichkeiten zwischen den Gebieten einen Einfluss auf die Schätzgüte hat, daher wird dieser Aspekt ebenfalls mit berücksichtigt. Somit werden in dieser Simulationsstudie die folgenden drei Aspekte adressiert:

1. Schätzverfahren
2. Erhebungsmethode der Individualdaten
3. Verteilung der Übergangswahrscheinlichkeiten zwischen den Gebieten

Um eine bessere Interpretierbarkeit der Ergebnisse zu gewährleisten, wird neben dem (hybriden) Multinomial-Dirichlet Modell auch das (hybride) Multinomial-Normal Modell bei der Simulation mit berücksichtigt. Die Simulation erlaubt daher auch den Vergleich der beiden vorgestellten hybriden Modelle untereinander.

III.3.1. Szenarien für die Simulationsstudie

Die Ausgangssituation und Annahmen für diese Simulation unterscheiden sich etwas von den Annahmen der in Kapitel II.4.3. vorgestellten Simulationsstudie. Erneut wird eine Wählerwanderung zwischen zwei Wahlen angenommen, es wird jedoch eine 7x6 Situation

189 Vgl. Greiner / Quinn, 2010, S. 1785.

190 Es handelt sich um eine erweiterte Version der Simulationsstudie Schlesinger, 2013, S. 63–81.

Angenommenen Wählerwanderung für die Simulation
Evaluierung der hybriden Modelle

	CSU	SPD	DieLinke	Grüne	Sonst	NW
2013 CSU	0.80	0.02	0.03	0.03	0.05	0.07
SPD	0.04	0.79	0.03	0.03	0.05	0.06
DieLinke	0.02	0.09	0.45	0.01	0.18	0.25
FDP	0.37	0.06	0.01	0.01	0.52	0.03
Grüne	0.02	0.17	0.03	0.71	0.04	0.03
Sonst	0.03	0.06	0.03	0.01	0.82	0.05
NW	0.03	0.04	0.03	0.03	0.04	0.83
	CSU	SPD	DieLinke	Grüne	Sonst	NW

2008

Abbildung 9: Simulation zur Evaluierung der hybriden Modelle, angenommene Wählerwanderung für die Simulation, dargestellt sind die Übergangswahrscheinlichkeiten. Die Werte basieren auf der vom Statistischen Amt der Landeshauptstadt München geschätzten Wählerwanderung für die Stadt München für diese beiden Wahlen, sind jedoch zum Teil an die Bedürfnisse der Simulation angepasst. (Datenquelle: Stadt München)

betrachtet. Die FDP wird dabei nur bei der ersten Wahl gesondert aufgeführt, während sie bei der zweiten Wahl Teil der sonstigen Parteien ist. Die für die Simulation angenommene Wählerwanderung ist in Abbildung 9 dargestellt. Bei der Wahl der Übergangswahrscheinlichkeiten wurde sich grob an den für München mit dem Multinomial-Dirichlet Modell geschätzten Übergangswahrscheinlichkeiten für die Landtagswahlen 2008 und 2013 orientiert¹⁹¹. Die verwendeten Werte weichen jedoch zum Teil etwas von der Schätzung ab, um eine größere Variabilität zwischen den Gebieten erreichen zu können.

Für die Simulation werden keine echten Wahldaten als Ausgangspunkt verwendet. Die betrachtete Region und auch das Wahlergebnis bei der ersten Wahl werden beide simuliert. Angenommen wird eine kleinere Stadt mit in etwa 240000 Wahlberechtigten. Die Stadt ist in 300 in etwa gleich große Wahlbezirke¹⁹² unterteilt. Im Vergleich zur Simulation in Kapitel II.4.3. ist die Zahl der Wahlberechtigten deutlich niedriger, die Zahl der

191 Vgl. Statistisches Amt der Landeshauptstadt München. Amtliches Endergebnis der Landtagswahl am 15.09.2013, Analyse der Wählerwanderung für die Landtagswahlen 2008 und 2013, 2013. <https://www.muenchen.de/rathaus/Stadinfos/Statistik/Wahlen/Landtagswahl.html> (zuletzt aufgerufen am: 17.07.2016).

192 Bei Landtagswahlen in Bayern wird anstelle des bei Bundestagswahlen üblichen Begriffes Wahlbezirk der Begriff Stimmbezirk genutzt. Beide beschreiben bei der jeweiligen Wahl eine vergleichbare Abstimmungsebene. Zur Vereinfachung der Benennung wird jedoch durchgehend Wahlbezirk in dieser Arbeit verwendet.

Wahlbezirke jedoch in etwa vergleichbar. Die Zahl der Wahlberechtigten N_i in den Wahlbezirken wird mittels einer Poisson-Verteilung, siehe Verteilung (45), simuliert.

$$N_i \sim Po(\lambda=800), \quad \forall i \in \{1, \dots, 300\} \quad (45)$$

Jeder Wahlbezirk hat daher im Mittel um die 800 Wahlberechtigte. Die relativ homogene Größe der Wahlbezirke stellt einen weiteren Unterschied zwischen dieser Simulation und der in Kapitel II.4.3., wo einzelne Wahlbezirke deutlich größer als andere sind, dar.

Für die Simulation des Wahlergebnisses bei der ersten Wahl, der Landtagswahl 2008, wird auf eine Dirichlet-Verteilung zurückgegriffen. Mit dieser werden für jeden Wahlbezirk die Parteianteile simuliert.

$$\begin{aligned} (P_{CSU,2008,i}, \dots, P_{NW,2008,i}) \sim \\ Dirichlet(\alpha_{CSU,2008} = 50 * 0.269, \dots, \alpha_{NW,2008} = 50 * 0.266) \\ \forall i \in \{1, \dots, 300\} \end{aligned} \quad (46)$$

Um eine ausreichende Streuung zwischen den Wahlbezirken zu erreichen, werden die angenommenen mittleren Parteienergebnisse¹⁹³ bei der Dirichlet-Verteilung mit jeweils 50 multipliziert. Daraus resultiert z. B. für die CSU, dass die simulierten Wahlergebnisse in den Wahlbezirken zwischen ~10% und ~45% liegen, wobei sich der Großteil der Ergebnisse im Bereich von 22% und 31% befindet. Die so erreichte Variabilität zwischen den Wahlbezirken zeigt eine als realistisch empfundene Streuung.

Ausgehend von diesen simulierten Wahlergebnissen für die erste Wahl wird mit der in Abbildung 9 dargestellten Wählerwanderung das individuelle Wahlverhalten bei der zweiten Wahl simuliert. Dabei werden zwei Szenarien mit unterschiedlicher Variabilität der Übergangswahrscheinlichkeiten zwischen den Wahlbezirken betrachtet. Für die Simulation wird wie bei Sim 2 in Kapitel II.4.3. für jede Zeile angenommen, dass die Übergangswahrscheinlichkeiten in den Wahlbezirken einer gemeinsamen Dirichlet-Verteilung folgen. Die Parameter dieser R Dirichlet-Verteilungen haben die folgende Form:

$$\alpha_{r,c} = v * \beta_{r,c}, \quad \forall r \in \{1, \dots, 7\}, c \in \{1, \dots, 6\}, v \in \{5, 20\} \quad (47)$$

Für das Szenario, bei dem nur eine geringe Variabilität der Übergangswahrscheinlichkeiten zwischen den Wahlbezirken angenommen wird, nimmt der Faktor v in Gleichung (47) den Wert fünf an, für das Szenario mit hoher Variabilität wird für v der Wert 20 verwendet. Die $\beta_{r,c}$ entsprechen dabei den Übergangswahrscheinlichkeiten in Abbildung 9.

Zusätzlich soll neben der Variabilität der Übergangswahrscheinlichkeiten zwischen den

¹⁹³ Die angenommenen mittleren Parteienergebnisse betragen: CSU: 26,9%; SPD 24,4%; DieLinke: 8,6%; FDP: 9,2%; Grüne: 3,9%; Sonst: 3,4%; NW: 26,6%. Die Werte orientieren sich zum Teil an der Landtagswahl 2008 in München, entsprechen jedoch nicht dem Wahlergebnis.

Wahlbezirken jeweils auch noch die Auswirkung von Aggregationsbias¹⁹⁴ mit untersucht werden. Dafür wird angenommen, dass die Loyalität der CSU-Wähler in Gebieten mit einem höheren gemeinsamen Wahlergebnis für die CSU und FDP bei der ersten Wahl größer ist. Die Anpassung erfolgt dabei analog zu dem in Kapitel II.4.3. in Gleichung (32) beschriebenen Vorgehen, wobei jedoch in dieser Simulation nur eine Zeile über eine gebietsspezifische Verteilung verfügt. Zusammen bedeutet dies, dass in der Simulation vier Szenarien für die Verteilung der Übergangswahrscheinlichkeiten betrachtet werden.

Bei der Erhebung der Individualdaten werden vier Szenarien verglichen. Beim ersten Szenario wird davon ausgegangen, dass keine Individualdaten zur Verfügung stehen und die Schätzung der Wählerwanderung nur mittels ökologischer Inferenz erfolgen kann. Das zweite Szenario simuliert eine „perfekte“ Telefonumfrage. Es wird angenommen, dass bei der Umfrage 2,5% alle Wahlberechtigten befragt werden. Es wird weiterhin davon ausgegangen, dass jeder Befragte richtig Auskunft erteilt. Diese Art der Erhebung führt dazu, dass in den einzelnen Wahlbezirken zwischen 1 - 5% der Wahlberechtigten befragt werden und für die gesamte Stadt Individualdaten von im Mittel ca. 6000 Wahlberechtigten vorliegen.

Das dritte und vierte Szenario simuliert jeweils eine Nachwahlbefragung (Exit-Poll). Diese findet an 15 der 300 Wahlbezirke statt, wobei die Auswahl der ausgesuchten Wahlbezirke per Zufallsauswahl erfolgt. Es wird weiterhin angenommen, dass nur alle wählenden Wahlberechtigten der zweiten Wahl befragt werden können. Dies impliziert, dass die Nichtwähler bei der zweiten Wahl, wie bei einem echten Exit-Poll, nicht mit erhoben werden können. Auch gibt nur ein Teil der befragten Wahlberechtigten Auskunft. Im dritten Szenario wird dabei von einem relativ „perfekten“ Exit-Poll ausgegangen. Alle Befragten geben richtig Auskunft und die Wahrscheinlichkeit am Exit-Poll teilzunehmen hängt nicht von der Wahlentscheidung ab. Es wird davon ausgegangen, dass die Response-Rate bei 70% liegt, so dass im Mittel Individualdaten von 5880 Wahlberechtigten vorliegen.

Beim vierten Szenario wird zusätzlich davon ausgegangen, dass das Antwortverhalten fehlerbehaftet ist. Ausgangspunkt für Szenario vier ist dabei der für Szenario drei simulierte Exit-Poll. Es wird jedoch angenommen, dass sich nur 50% der Wahlberechtigten die Auskunft geben an die Wahlentscheidung bei der ersten Wahl richtig erinnern können. Für die Wahlberechtigten, die sich nicht mehr an ihre Wahlentscheidung bei der ersten Wahl erinnern können, wird angenommen, dass sie ihre Wahlentscheidung bei der zweiten

194 Vgl. Kapitel II.3.1.2..

Wahl auch als Wahlentscheidung bei der ersten Wahl angeben. Zusätzlich wird davon ausgegangen, dass auch die Response-Rate von der aktuellen Wahlentscheidung abhängt. Während für fast alle Parteien weiterhin eine Response-Rate von 70% angenommen wird, gilt diese nicht mehr für die CSU- und SPD-Wähler bei der zweiten Wahl. Für diese wird von einer niedrigeren Response-Rate ausgegangen. Diese beträgt bei den Wählern der CSU 80% und bei den Wählern der SPD 90% der Response-Rate der anderen Parteien. Um zu vermeiden, dass Wählerwanderungen in den Individualdaten vorliegen, die nicht im Einklang mit den simulierten Wahlergebnissen sind, werden die Anpassungen nur soweit vorgenommen, wie es die Ränder der Tabellen, die simulierten Wahlergebnisse, erlauben. Dies bedeutet, dass es zum Teil gebiets- und parteispezifische Verzerrungen in den Individualdaten gibt.

Insgesamt werden daher bei der Simulation 16 Szenarien betrachtet. Für jedes dieser Szenarien werden 100 Datensätze generiert. Die Wählerwanderung wird – abhängig vom Szenario – mittels ökologischer Inferenz, einfacher Hochrechnung aus den Individualdaten und den hybriden Modellen geschätzt. Dabei wird zur Schätzung jeweils das (hybride) Multinomial-Dirichlet Modell und das (hybride) Multinomial-Normal Modell verwendet, wobei bei letztem die Schätzung der Wählerwanderung aus Zeitgründen nur für die jeweils ersten 30 Datensätze pro Szenario erfolgt. Zur Quantifizierung der Schätzgüte wird erneut der AD, vergleiche Gleichung (28), verwendet. Bei der Hochrechnung aus den Individualdaten bei den Exit-Polls werden zur Bestimmung der Schätzgüte jedoch nur die beobachtbaren Spalten berücksichtigt.

III.3.2. Festlegungen für die Simulationsstudie

Für die Simulation wird beim (hybriden) Multinomial-Dirichlet Modell ein Burn-In von 50000 und ein Thinning von 1250 gewählt. Für die Schätzung werden 1000 Ziehungen aus den Ketten herangezogen. Abgesehen vom etwas höheren Wert für das Thinning, was zu einer weiteren Reduktion der Autokorrelation der Ziehungen führt, sind diese Werte identisch mit den in Kapiteln II.4.2.2. und II.4.2.3. diskutierten Werten bei der ökologischen Inferenz. Bei einer detaillierten Evaluierung der Ketten zeigt sich auch beim hybriden Multinomial-Dirichlet Modell das bereits in Abbildung 4 gezeigte Verhalten: Bei einem kleinen Teil der Ketten kommt es zu einem kurzzeitigen Abweichen vom eigentlich erreichten stabilen Niveau. Da der AD von zwei Schätzungen mit dem Modell bei gleicher Parametrisierung und mit gleichen Daten weitestgehend unter 0,02 liegt, kann von einer Konvergenz der Ketten ausgegangen werden.

Für das (hybride) Multinomial-Normal Modell wird ein Burn-In von 3000000 und ein Thinning von 1000 gewählt. Wie beim (hybriden) Multinomial-Dirichlet Modelle werden 1000 Ziehungen aus den Ketten verwendet. Die Werte wurden gewählt, um eine ausreichende Kettenlänge bei gleichzeitig noch akzeptabler Rechenzeit zu erreichen. Es ist möglich, bei dieser Simulation von einer weitestgehenden Konvergenz der Ketten auszugehen. Der AD zwischen zwei Schätzungen mit dem Modell bei gleicher Parametrisierung und mit gleichen Daten liegt jedoch im Bereich von 0,02 und 0,04 und ist daher größer als beim hybriden Multinomial-Dirichlet Modell. Weiterhin ist auch bei dieser Simulation erneut eine Abhängigkeit der Schätzung von der Referenzpartei feststellbar, wobei diese mit ADs im Bereich von 0,03 – 0,06 eher gering ausfällt.

III.3.3. Ergebnisse der Simulationsstudie

Die Ergebnisse der Simulation sind in Abbildung 10 dargestellt. Für jedes der vier Szenarien zur Verteilung der Übergangswahrscheinlichkeiten zwischen den Wahlbezirken werden die vier Möglichkeiten für die Individualdaten betrachtet. Die oberen beiden Grafiken zeigen die Ergebnisse bei einer heterogenen Verteilung, vergleiche Gleichung (47), der Übergangswahrscheinlichkeiten ($v = 20$), die unteren beiden Grafiken die bei einer homogenen Verteilung ($v = 5$). Links sind jeweils die Ergebnisse für die Modelle ohne Aggregationsbias zu finden, während rechts die Ergebnisse mit Aggregationsbias dargestellt sind.

Die Szenarien ohne Individualdaten sind, ausgehend von Design der Simulation, in etwa vergleichbar mit Szenario Sim 2 in Kapitel II.4.3.. Ein Vergleich mit den Ergebnissen in Abbildung 8 zeigt auch, dass die Werte bei einer heterogenen Verteilung der Übergangswahrscheinlichkeiten eine vergleichbare Größenordnung aufweisen, jedoch in dieser Simulation etwas niedriger sind. Wie beim Szenario Sim 2 zeigt das Multinomial-Normal Modell auch hier eine bessere Schätzgüte im Vergleich zum Multinomial-Dirichlet Modell. Bei einer homogenen Verteilung der Übergangswahrscheinlichkeiten zwischen den Wahlbezirken ist die Schätzgüte bei beiden Modellen besser, wobei der Unterschied beim Multinomial-Dirichlet Modell deutlicher ausfällt und dieses auch die bessere Schätzgüte aufweist. Der Aggregationsbias hat in beiden Fällen einen eher vergleichsweise geringen Einfluss auf die Schätzgüte.

Für die jeweils drei Szenarien mit Individualdaten sind die Ergebnisse der einfachen Hochrechnungen und der beiden hybriden Modellen dargestellt. Bei der Telefonumfrage schlägt die einfache Hochrechnung aus den Individualdaten bei einer heterogenen

Verteilung der Übergangswahrscheinlichkeiten die beiden hybriden Modelle. Die Schätzungen mit den hybriden Multinomial-Normal Modell sind dabei etwas besser als die des hybriden Multinomial-Dirichlet Modells. Bei einer homogenen Verteilung der Übergangswahrscheinlichkeiten liegen die Schätzgüten fast immer auf einem vergleichbaren Niveau, nur ohne Aggregationsbias ist das hybride Multinomial-Normal Modell etwas besser als die einfache Hochrechnung und das hybride Multinomial-Dirichlet Modell. Jedoch ist die Schätzgüte mit allen Modellen sehr gut und die erreichten ADs liegen deutlich unter 0,1 in allen vier Szenarien. Die Schätzungen sind daher den Schätzungen mit der ökologischen Inferenz deutlich überlegen.

Evaluierung der hybriden Modelle: Simulationsstudie

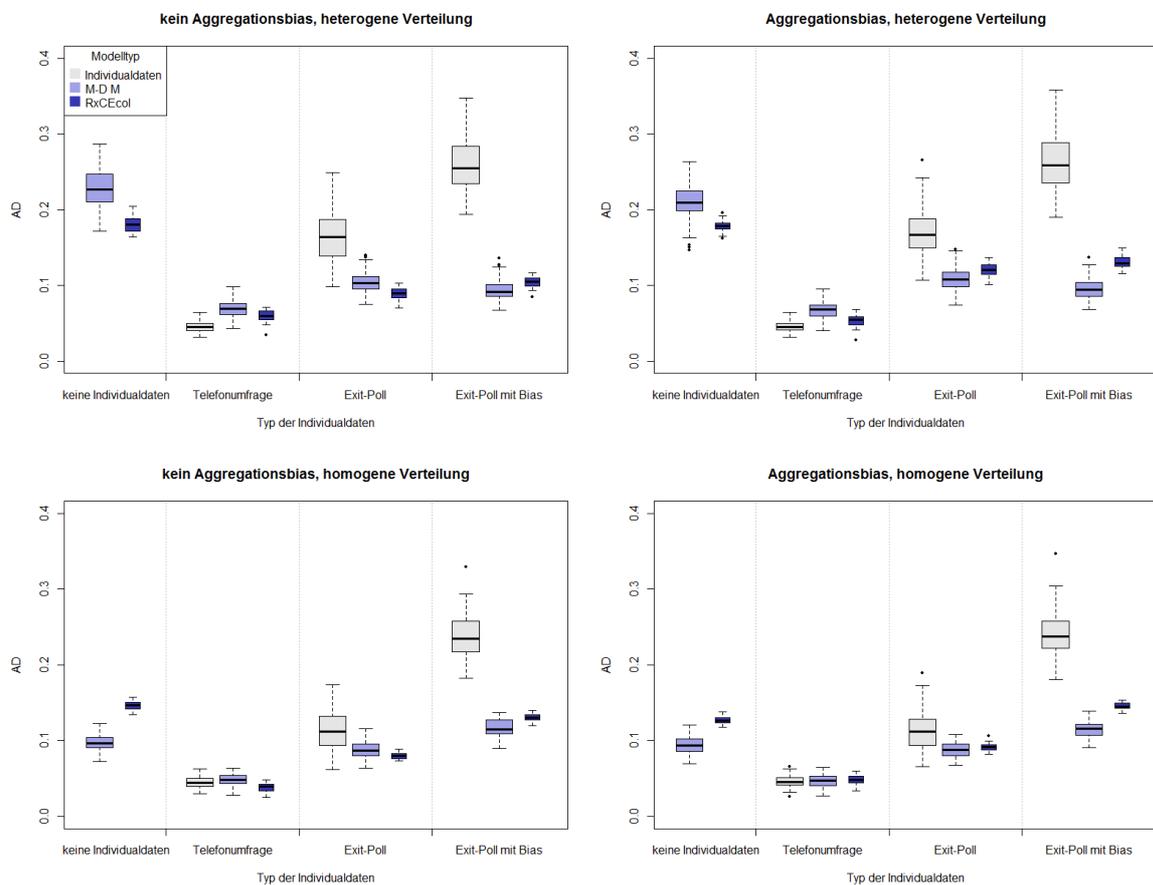


Abbildung 10: Simulation zur Evaluierung der hybriden Modelle, Ergebnisse der Simulationsstudie, dargestellt ist der AD der Schätzung mit den Modellen zu den simulierten Individualdaten. Für jedes Szenario wurden 100 Datensätze generiert. Mit dem (hybriden) Multinomial-Dirichlet Modell (M-D M) und bei der einfachen Hochrechnung aus den Individualdaten (Individualdaten) werden alle Datensätze evaluiert, mit dem (hybriden) Multinomial-Normal Modell (RxCEcol) werden nur die ersten 30 Datensätze pro Szenario evaluiert. Details zu den Szenarien finden sich in Kapitel III.3.1., die Einstellungen der Modelle werden in Kapitel III.3.2. beschrieben.

Beim Exit-Poll ohne Bias ändert sich das Bild und die einfache Hochrechnung aus den Individualdaten ist in allen vier Szenarien schlechter als die Schätzung mit den hybriden Modellen. Bei einer homogenen Verteilung der Übergangswahrscheinlichkeiten ist die Hochrechnung sogar tendenziell schlechter als die Schätzung mit dem Multinomial-Dirichlet Modell ohne Individualdaten. Während der Unterschied zwischen den hybriden Modellen und der Hochrechnung bei einer heterogenen Verteilung der Übergangswahrscheinlichkeiten noch stärker ausfällt, ist die Hochrechnung in diesen Szenarien tendenziell besser als die Schätzung mit der ökologischen Inferenz. Beim Vergleich der beiden hybriden Modelle fällt auf, dass es kein eindeutig besseres Modell gibt. Ohne Aggregationsbias schlägt das hybride Multinomial-Normal Modell das hybride Multinomial-Dirichlet Modell, während es mit Aggregationsbias genau umgekehrt ist. Die beiden hybriden Modelle sind dabei besser als die ökologische Inferenz, wobei der Unterschied bei der heterogenen Verteilung der Übergangswahrscheinlichkeiten deutlich größer ausfällt.

Für den Exit-Poll mit Bias zeigt sich, dass die Hochrechnung aus Individualdaten die schlechteste Schätzgüte aufweist. Während der Unterschied zwischen ökologischer Inferenz und Hochrechnung bei heterogener Verteilung der Übergangswahrscheinlichkeiten noch moderat ausfällt, ist der Unterschied bei einer homogenen Verteilung gravierend. Die Schätzungen mit den hybriden Modellen sind in allen vier Szenarien deutlich besser als die einfache Hochrechnung, wobei das hybride Multinomial-Dirichlet Modell in allen vier Szenarien besser als das hybride Multinomial-Normal Modell abschneidet. Bei vorliegendem Aggregationsbias fällt der Unterschied zwischen den beiden hybriden Modellen noch deutlicher zugunsten des hybriden Multinomial-Dirichlet Modells aus. Beim Vergleich der Schätzgüte zwischen den hybriden Modellen und ökologischer Inferenz ist festzuhalten, dass bei einer homogenen Verteilung der Übergangswahrscheinlichkeiten die ökologische Inferenz beim (hybriden) Multinomial-Dirichlet Modell tendenziell die bessere Schätzgüte aufweist. Beim (hybriden) Multinomial-Normal Modell ist dies nur bei vorliegendem Aggregationsbias der Fall. Bei einer heterogenen Verteilung der Übergangswahrscheinlichkeiten sind die hybriden Modelle deutlich der ökologischen Inferenz überlegen.

Beim Vergleich der Schätzungen zwischen den beiden Exit-Polls, d. h. mit und ohne Bias, mit den hybriden Modellen zeigt sich bei der homogenen Verteilung der Übergangswahrscheinlichkeiten eine relativ deutliche Verschlechterung der Schätzgüte beim Exit-Poll mit Bias. Der Unterschied bei den erreichten Schätzgüten ist bei einer

heterogenen Verteilung der Übergangswahrscheinlichkeiten deutlich geringer. Beim hybriden Multinomial-Normal Modell ist jedoch auch hier eine geringe Verschlechterung erkennbar, die bei vorliegenden Aggregationsbias stärker ausfällt. Beim hybriden Multinomial-Dirichlet Modell ist die Schätzgüte beim Exit-Poll mit Bias sogar in dieser Simulation etwas besser als beim Exit-Poll ohne Bias. Der Unterschied ist jedoch eher gering. Auch sind bei diesem Modell keine relevanten Auswirkungen des Aggregationsbias auf die Schätzgüte feststellbar.

Beim Vergleich aller drei Szenarien mit Individualdaten zeigt das Szenario mit Telefonumfrage die generell deutlich beste Schätzgüte. Dieser Fall ist jedoch auch der praktisch unwahrscheinlichste. Für die Simulation wird von einer fehlerfreien Erhebung der Individualdaten ausgegangen und alle Befragten geben wahrhaftig Auskunft. Eine solche ideale Erhebung wird es in der Praxis kaum geben. Jedoch zeigt sich auch bei diesem Szenario, dass die hybriden Modelle zumindest in die Nähe der Hochrechnung aus den Individualdaten kommen. Bei den simulierten Exit-Polls sind die hybriden Modelle bereits der Hochrechnung aus den Individualdaten überlegen. Dabei muss berücksichtigt werden, dass aufgrund der Erhebungsart die Individualdaten nicht den Modellannahmen der hybriden Modelle entsprechen. Da die Nichtwähler der zweiten Wahl bei einem Exit-Poll nicht erhoben werden können, sind auch keine Informationen über diese in den Individualdaten vorhanden. Dies impliziert auch, dass die Übergangswahrscheinlichkeiten in den Individualdaten nicht den Übergangswahrscheinlichkeiten in den Aggregatdaten entsprechen und daher die entsprechende Modellannahme verletzt wird. Beim Exit-Poll mit Bias fällt die Verletzung der Annahme noch deutlicher aus, jedoch ist auch hier mit dem hybriden Modell zumindest eine vergleichbare Schätzgüte wie mit der ökologischen Inferenz erreichbar. Im als realistischer bewerteten Fall mit heterogener Verteilung der Übergangswahrscheinlichkeiten sind die hybriden Modelle der ökologischen Inferenz sogar deutlich überlegen.

Der Vergleich der beiden hybriden Modelle ergibt keinen eindeutigen Gewinner. Vor allem wenn kein Aggregationsbias vorliegt und die Individualdaten keinen Bias aufweisen, ist die Schätzgüte mit den hybriden Multinomial-Normal Modell besser. Bei vorliegenden Aggregationsbias und/oder Individualdaten mit Bias hat hingegen fast immer das hybride Multinomial-Dirichlet Modell die besser Schätzgüte. Wenn das Szenario mit heterogener Verteilung der Übergangswahrscheinlichkeiten und Aggregationsbias als am realistischen angesehen wird, gibt es in dieser Simulation einen kleinen Vorteil für das hybride Multinomial-Dirichlet Modell.

III.4. Erweiterte Betrachtung des (hybriden) Multinomial-Dirichlet Modells

Die bisher vorgestellten Betrachtungen zeigen auf, dass mit dem (hybriden) Multinomial-Dirichlet Modell bei der Schätzung von Wählerwanderungen gute Ergebnisse erreicht werden können. Im in Kapitel II.4. durchgeführten Vergleich der Methoden für die ökologische Inferenz hat das Multinomial-Dirichlet Modell über alle Szenarien die beste Performance. Die Ergebnisse für die hybride Erweiterung des Modells in der Simulationsstudie in Kapitel III.3. zeigen, dass die Schätzgüte sich im Vergleich zum Modell für die ökologische Inferenz fast immer verbessert und auch mindestens vergleichbar mit der des hybriden Multinomial-Normal Modells ist.

Beim direkten Vergleich zwischen den (hybriden) Multinomial-Normal Modell und den (hybriden) Multinomial-Dirichlet Modell zeigt sich, dass bei letzteren im Allgemeinen deutlich kürzere Ketten für eine adäquate Schätzung notwendig sind und daher auch die notwendige Rechenzeit deutlich kürzer ausfällt. Weiterhin gilt, dass es weniger Probleme bei der Konvergenz der Ketten beim (hybriden) Multinomial-Dirichlet Modell gibt und die Schätzungen daher auch die größere Konstanz aufweisen. Beim (hybriden) Multinomial-Normal Modell kann zusätzlich noch eine Abhängigkeit der Schätzung von der an sich arbiträren Wahl einer Referenzpartei festgestellt werden. Die Wahl für eine andere Referenzpartei kann, abhängig von der Datensituation, zu geringen oder auch deutlichen Unterschieden bei der Schätzung führen.

Zusammenfassend legen die bisherigen Ergebnisse nahe, dass das (hybride) Multinomial-Dirichlet Modell für die Schätzung von Wählerwanderung mindestens eine gute Alternative darstellt. In diesem Kapitel sollen daher zwei Aspekte für dieses Modell, die bei der Anwendung von Relevanz sind, detaillierter betrachtet werden. Es handelt sich zum einen um die Abhängigkeit der Schätzgüte von den vorhanden Daten und Wissen, zum anderen soll die Güte der Kreditibilitätsintervalle für die Schätzer evaluiert werden.

III.4.1. Simulationsstudie Priori-Wissen und Umfang der Individualdaten

Die erste Detailbetrachtung für das (hybriden) Multinomial-Dirichlet Modells widmet sich der Datensituation und den erweiterten Modellfestlegungen. Dabei wird sich vor allem Fragen gewidmet, die bei der praktischen Anwendung des Modells relevant sind. Zur Evaluierung wird dabei erneut auf eine Simulationsstudie zurückgegriffen, da so die Schätzgüte direkt quantifiziert und verglichen werden kann. Die folgenden drei Aspekte werden in der Simulation dabei detaillierter evaluiert:

1. Größe der Stadt bzw. Zahl der zur Schätzung verfügbaren Wahlbezirke
2. Umfang der Individualdaten
3. Informative Wahl der Priori-Parameter bei den Hyperprioris

III.4.1.1. Szenarien für die Simulationsstudie

In den beiden bisher vorgestellten Simulationsstudien werden zwar Regionen mit einer unterschiedlichen Größe, die Großstadt München in Kapitel II.4.3. und eine angenommene kleinere Stadt mit ca. 240000 Wahlberechtigten in Kapitel III.3., betrachtet, jedoch stehen in beiden Fällen Informationen von einer vergleichbaren Zahl von Wahlbezirken, 314 bzw. 300, zur Verfügung. Die erreichbare Schätzgüte ist in beiden Fällen in etwa vergleichbar, wobei bei der kleineren Stadt mit der homogenen Zahl der Wahlberechtigten in den Wahlbezirken eine etwas besser Schätzgüte erreicht werden kann. Die Zahl der Wahlbezirke wird jetzt bei den folgenden Betrachtungen variiert. Zum einen wird erneut eine Großstadt wie München angenommen, wobei jedoch hier keine Zusammenlegung zu über die Zeit konstanten Wahlbezirken notwendig ist, so dass 600 Wahlbezirke für die Schätzung zur Verfügung stehen¹⁹⁵, zum anderen wird eine vergleichsweise kleine Großstadt mit 200 Wahlbezirken als Alternative betrachtet. Die Zahl der zur Verfügung stehenden Wahlbezirke für die Schätzung ist daher entweder deutlich größer oder kleiner als bei den bisherigen Betrachtungen. Aufgrund des geringen Einflusses von Aggregationsbias auf die Schätzgüte beim (hybriden) Multinomial-Dirichlet Modell in der Simulation in Kapitel III.3. wird bei dieser Simulation die Auswirkung von Aggregationsbias nicht mehr mit untersucht.

Weiterhin stellt sich die Frage, inwieweit die Schätzgüte durch den Umfang der Individualdaten beeinflusst wird. Führt ein mehr an Individualdaten immer zu einer besseren Schätzung oder kommt es zu einer Sättigung, ab der nur noch von einer moderaten Verbesserung ausgegangen werden kann? Diese Frage ist vor allem für die praktische Anwendung der Modelle relevant, da die Erhebung der Individualdaten bei der Schätzung von Wählerwanderungen mit der größte Kostenfaktor ist. Die Betrachtungen in Kapitel III.3. zeigen bereits, dass es eine Abhängigkeit von der Güte der Individualdaten und der Erhebungsmethode gibt. Bei den folgenden Analysen wird davon ausgegangen, dass die Individualdaten mittels Nachwahlbefragung (Exit-Poll) erhoben werden. Die Fokussierung auf diese Erhebungsmethode erfolgt, da sie für diese Art von Modellen als

¹⁹⁵ Die 314 Wahlbezirke in Kapitel II.4. resultieren aus der Notwendigkeit, über die Zeit konstante Wahlbezirke zu bilden. Zwischen den Bundestagswahlen 2005 und 2009 änderte sich der Zuschnitt der Wahlbezirke in München relativ deutlich, so dass eine starke Aggregation notwendig ist. Bei beiden Wahlen war die Stadt München in über 600 Urnenwahlbezirke unterteilt. Vgl. Kapitel II.4.1..

realistischer angesehen wird¹⁹⁶. Zusätzlich zum Fall, dass keine Individualdaten vorhanden sind, werden jeweils zwei Nachwahlbefragungen mit unterschiedlichen Umfang betrachtet.

Der letzte untersuchte Aspekt widmet sich dem Vorwissen. In Kapitel II.3.4.4. wurde bereits auf eine Besonderheit bei der Spezifikation von Priori-Wissen bei der Schätzung von Wählerwanderungen mit hierarchischen Bayesianischen Modellen eingegangen. Es ist im RxC Fall nicht wirklich möglich, ein Vorwissen zu spezifizieren, dass auf allen Ebenen als nicht-informativ betrachtet werden kann. Diese Modelleigenschaft gilt für die entsprechenden Modelle für die ökologische Inferenz, aber auch die vorgestellten hybriden Modelle. In den bisherigen Betrachtungen wurden die Standard-Priori-Parameter der jeweiligen Modelle für die Hyperprioris genutzt. Beim (hybriden) Multinomial-Dirichlet Modell bedeutet dies, dass für alle Zellen die gleiche Gamma-Verteilung¹⁹⁷ angenommen wird. Die Auswirkung einer abweichenden Wahl soll nun ebenfalls mit untersucht werden.

Für die Simulationsstudie wird die Zahl der Wahlberechtigten in den 200 bzw. 600 Wahlbezirken erneut mit einer Poisson-Verteilung simuliert.

$$N_i \sim Po(\lambda=1500), \quad \forall i \in \{1, \dots, N\} \quad (48)$$

Im Mittel wird dabei von 1500 Wahlberechtigten pro Wahlbezirk ausgegangen, so dass eine Stadt mit entweder 300000 (200 Wahlbezirke) oder 900000 Wahlberechtigten (600 Wahlbezirke) angenommen wird. Für die Simulation des Wahlergebnisses bei der ersten Wahl wird analog wie in Kapitel III.3. eine Dirichlet-Verteilung verwendet.

$$\begin{aligned} (P_{CSU,2009,i}, \dots, P_{NW,2009,i}) \sim \\ Dirichlet(\alpha_{CSU,2009}=50 * 0.236, \dots, \alpha_{NW,2009}=50 * 0.266) \end{aligned} \quad (49)$$

$$\forall i \in \{1, \dots, N\}$$

Es werden bei der ersten Wahl vier Parteien, die Sonstigen und die Nichtwähler betrachtet. Als Ausgangspunkt für die Simulation dient dabei das Wahlergebnis bei der Bundestagswahl 2009 in München.¹⁹⁸ Wie in Kapitel III.3. wird der Faktor 50 verwendet, um eine realistische Variabilität der Wahlergebnisse zwischen den Wahlbezirken zu erreichen.

Es wird bei dieser Simulation eine 6x6 Situation betrachtet. Als zweite Wahl, bei der dieselben sechs Alternativen wie bei der ersten Wahl betrachtet werden, wird die Bundestagswahl 2013 angenommen. Die für die Simulation genutzte Wählerwanderung ist

196 Es kann z. B. bei Telefonumfrage nicht davon ausgegangen werden, dass die Befragten die Wahlentscheidung und nicht die Wahlabsicht angeben. Auch ist es unwahrscheinlich, dass alle Befragten bei einer Telefonumfrage ihren Wahlbezirk exakt angeben können und somit auch einem Wahlbezirk eindeutig zugeordnet werden können.

197 Gamma-Verteilung mit shape = 4 und rate = 2.

198 CSU: 23,6%; SPD 14,2%; FDP: 12,9%;Grüne: 12,8%; Sonst: 9,9%; NW: 26,6%; vgl. Tabelle 1 und Statistisches Amt der Landeshauptstadt München, 2009.

Angenommene Wählerwanderung für die Simulation
detaillierte Evaluierung des (hybriden) Multinomial-Dirichlet Modells

	2009					
	CSU	SPD	FDP	Grüne	Sonst	NW
2013 CSU	0.80	0.02	0.03	0.03	0.05	0.07
SPD	0.04	0.79	0.02	0.03	0.06	0.06
FDP	0.37	0.06	0.37	0.01	0.16	0.03
Grüne	0.02	0.18	0.01	0.71	0.06	0.02
Sonst	0.03	0.07	0.01	0.01	0.70	0.11
NW	0.03	0.05	0.03	0.03	0.06	0.80

Abbildung 11: Detaillierte Evaluierung des (hybriden) Multinomial-Dirichlet Modells, angenommene Wählerwanderung für die Simulation, dargestellt sind die Übergangswahrscheinlichkeiten. Die Werte basieren auf der vom Statistischen Amt der Landeshauptstadt München geschätzten Wählerwanderung für die Stadt München für diese beiden Wahlen, sind jedoch zum Teil an die Bedürfnisse der Simulation angepasst. (Datenquelle: Stadt München)

in Abbildung 11 dargestellt. Ausgangspunkt dafür ist die mit dem Multinomial-Dirichlet Modell vom Statistischen Amt der Landeshauptstadt München geschätzte Wählerwanderung für München für die Bundestagswahlen 2009 und 2013¹⁹⁹. Ein Teil der Übergangswahrscheinlichkeiten ist jedoch an die Bedürfnisse der Simulation angepasst worden, so dass die Werte nicht komplett identisch sind.

Bei der Simulation der Wählerwanderung wird jedoch zusätzlich noch von drei Bevölkerungsgruppen in jeden Wahlbezirk ausgegangen. Dabei stellt jede Bevölkerungsgruppe im Mittel ein Drittel der Wahlberechtigten eines Wahlbezirkes. Die Anteile der drei Bevölkerungsgruppen werden in der Simulation mit einer Dirichlet-Verteilung generiert.

$$\begin{aligned} (P_{N_1,i}, P_{N_2,i}, P_{N_3,i}) &\sim \\ &Dirichlet(\alpha_1=1/3, \alpha_2=1/3, \alpha_3=1/3), \quad \forall i \in \{1, \dots, N\} \end{aligned} \quad (50)$$

Diese Simulation folgt damit in etwa den Überlegungen für die Sim 3 in Kapitel II.4.3.1.. Für jede der drei Bevölkerungsgruppen wird auch hier eine eigene Wählerwanderung

¹⁹⁹ Vgl. Statistisches Amt der Landeshauptstadt München. Amtliches Endergebnis der Bundestagswahl am 22.09.2013, Analyse der Wählerwanderung für die Bundestagswahlen 2009 und 2013, 2013. <https://www.muenchen.de/rathaus/Stadtinfos/Statistik/Wahlen/Bundestagswahl.html> (zuletzt aufgerufen am: 18.07.2016).

angenommen, wobei der primäre Unterschied erneut in unterschiedlich hohen Loyalitätsraten zu finden ist.

Für jede der drei Bevölkerungsgruppen wird für jede Partei der ersten Wahl eine gemeinsame Dirichlet-Verteilung für die Übergangswahrscheinlichkeiten angenommen.

$$\alpha_{r,c,k} = 5 * \beta_{r,c,k}, \quad \forall r \in \{1, \dots, 6\}, c \in \{1, \dots, 6\}, k \in \{1, 2, 3\} \quad (51)$$

Gleichung (51) zeigt die Parameter der 6 x 3 Dirichlet-Verteilung. Die Variabilität der Übergangswahrscheinlichkeiten zwischen den Wahlbezirken wird für jede Bevölkerungsgruppe als moderat angenommen. Diese Art der Modellierung der Wählerwanderung hat zur Folge, dass die Simulation von den Modellannahmen des (hybriden) Multinomial-Dirichlet Modells abweicht und die Wählerwanderung in den Wahlbezirken einer Mischung von drei Dirichlet-Verteilung mit jeweils unterschiedlichen Mischungsverhältnis entspricht.

Ausgehend von den so simulierten Individualdaten werden drei Szenarien bezüglich des Vorhandenseins von Individualdaten für die Schätzung simuliert. Im ersten Szenario wird davon ausgegangen, dass keine Individualdaten erhoben werden können. Das zweite und dritte Szenario simuliert jeweils eine Nachwahlbefragung, wobei die Zahl der untersuchten Wahlbezirke sich zwischen beiden Szenarien unterscheidet. Im zweiten Szenario wird davon ausgegangen, dass 2,5% der Wahlbezirke bei der Nachwahlbefragung berücksichtigt werden, während beim dritten Szenario die Nachwahlbefragung an 5% der Wahlbezirke durchgeführt wird. Dies bedeutet, dass angenommen wird, dass die Nachwahlbefragung bei 200 Wahlbezirken an fünf (2,5%) bzw. 10 (5%) und bei 600 Wahlbezirken an 15 (2,5%) bzw. 30 (5%) Wahlbezirken durchgeführt wird. Bei den beiden Nachwahlbefragungen wird weiter davon ausgegangen, dass die Nichtwähler der zweiten Wahl nicht erreicht werden können und dass die Response-Rate unabhängig von der Partei bei 70% liegt. Dies führt dazu, dass Individualdaten von in etwa 1,3% (2,5%) bzw. 2,6% (5%) der Wahlberechtigten vorliegen. Falsche Angaben, z. B. aufgrund fehlerhafter Erinnerung, werden nicht simuliert. Das Szenario mit der Nachwahlbefragung an 5% der Wahlbezirke ist dabei weitestgehend mit den Exit-Poll ohne Bias in der Simulation in Kapitel III.3. vergleichbar.

Beim Priori-Wissen werden jeweils zwei Szenarien evaluiert. Das erste Szenario nutzt für alle Zellen die Standard-Priori-Parameter für die Gamma-Verteilungen auf der dritten Ebene des Modells, d. h. es wird unabhängig von der betrachteten Übergangswahrscheinlichkeit die gleiche Verteilung mit den gleichen Erwartungswert angenommen. Die Parameter der 6 x 6 Gamma-Verteilungen sind in (52) dargestellt.

$$\alpha_{r,c} \sim \text{Gamma}(\text{shape}_{r,c}=4, \text{rate}_{r,c}=2), \quad \forall r \in \{1, \dots, 6\}, c \in \{1, \dots, 6\} \quad (52)$$

Das zweite Szenario geht von informativen Vorwissen aus. Dabei wird ausgenutzt, dass bei Parteien im Allgemeinen die Loyalitätsraten deutlich größer als die Übergangswahrscheinlichkeiten zu anderen Parteien sind.

Für die Simulation wird für die Entscheidung, ob informatives Vorwissen genutzt werden soll, auf ein relativ grobes Entscheidungskriterium zurückgegriffen. Die Nutzung des informativen Priori-Wissens erfolgt nicht, wenn eine Partei deutliche Verluste bei der zweiten Wahl erlitten hat und nur weniger als 60% der Stimmen der Vorwahl bei der zweiten Wahl erreichen kann, sowie, wenn es sich um eine kleine Partei mit einem Wahlergebnis von unter 5% handelt. Wenn eine Partei bei der zweiten Wahl weniger als 60% ihres Wahlergebnisses der ersten Wahl erreicht, muss die Loyalitätsrate bei dieser Partei unter 60% liegen und die Annahme einer hohen Loyalitätsrate ist a priori nicht mehr gerechtfertigt. Das zweite Kriterium, der Ausschluss von Parteien unter 5%, soll die größere Unsicherheit über und die möglicherweise höhere Wechselwilligkeit von Wählern von Kleinparteien repräsentieren.

Bei der Simulation bedeutet das Kriterium, dass für alle Parteien, mit Ausnahme der FDP, informatives Vorwissen genutzt wird. Für die FDP wird hingegen für alle C Hyperprioris die Verteilung (52) angenommen. Bei allen anderen Parteien wird für die Loyalitätsraten folgende Parameterwahl bei der Hyperpriori genutzt.

$$\alpha_{r,c} \sim \text{Gamma}(\text{shape}_{r,c}=30, \text{rate}_{r,c}=1), \quad \forall r = c \in \{CSU, \dots, NW\} \setminus FDP \quad (53)$$

Für alle anderen Zellen wird auch bei diesen Parteien für die Hyperprioris die Parametrisierung von Verteilung (52) genutzt, d. h., es wird für alle anderen Übergangswahrscheinlichkeiten das gleiche Vorwissen angenommen. Die Parameterwahl in (53) sorgt dafür, dass der Priori-Erwartungswert der Loyalitätsraten deutlich größer ist und bei sechs Spalten bei ungefähr 75% liegt.

Für jedes der sechs Datenszenarien wurden 100 Datensätze generiert. Für die Simulation wird beim (hybriden) Multinomial-Dirichlet Modell ein Burn-In von 150000 und ein Thinning von 1250 gewählt. Für die Schätzung werden 1000 Ziehungen aus den Ketten herangezogen. Der größere Burn-In muss hier gewählt werden, da bei den Modellen mit 600 Wahlbezirken bei einem Burn-In von 50000 noch nicht von einer Konvergenz der Ketten ausgegangen werden kann. Als Maß für die Schätzgüte wird erneut der AD, siehe Gleichung (28), verwendet.

III.4.1.2. Ergebnisse der Simulationsstudie

Die Ergebnisse der Simulation werden in Abbildung 12 gezeigt. Links sind die Modelle für die Stadt mit 200 Wahlbezirken, rechts die Modelle für die Stadt mit 600 Wahlbezirken dargestellt. Auf die Darstellung der einfachen Hochrechnung aus den Individualdaten wird verzichtet, so dass nur die Schätzungen mit dem (hybriden) Multinomial-Dirichlet Modell dargestellt werden. Der Farbton der Boxplots gibt dabei an, ob es sich um ein Modell ohne Vorwissen oder ein Modell mit spezifischen Vorwissen handelt. Die Modelle bei den drei Szenarien für die Individualdaten werden dabei jeweils direkt miteinander verglichen, wobei links jeweils die Modelle ohne Individualdaten, mittig die Modelle mit simulierter Nachwahlbefragung an 2,5% der Wahlbezirke und rechts die Modelle mit simulierter Nachwahlbefragung an 5% der Wahlbezirke gegenübergestellt werden.

Beim Szenario mit 200 Wahlbezirken ist die Schätzgüte mit dem Multinomial-Dirichlet Modell ohne Vorwissen am schlechtesten. Alleine durch das Hinzufügen von Vorwissen kann die Schätzgüte bei der ökologischen Inferenz jedoch bereits deutlich verbessert werden. Eine noch stärkere Verbesserung wird durch die Kombination von Individual- und Aggregatdaten beim hybriden Multinomial-Dirichlet Modell erreicht. Der AD der Modellschätzung von den simulierten Individualdaten kann dadurch mehr als halbiert

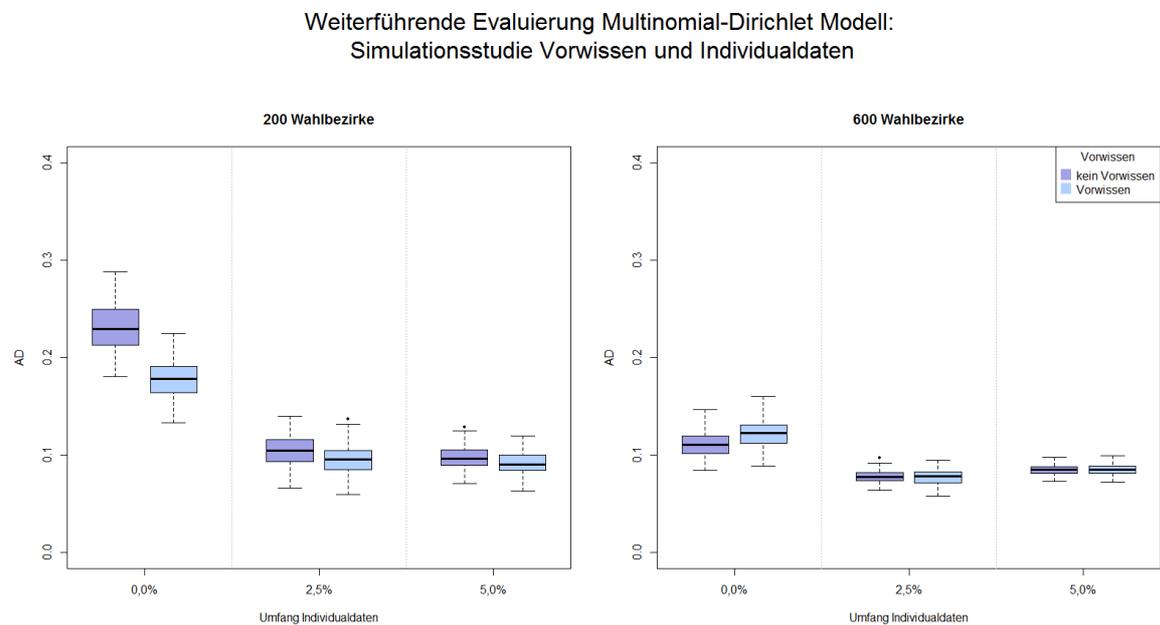


Abbildung 12: Detaillierte Evaluierung des (hybriden) Multinomial-Dirichlet Modells, Ergebnisse der Simulationsstudie, dargestellt ist der AD der Schätzung mit den Modellen zu den simulierten Individualdaten. Für jedes Szenario wurden 100 Datensätze generiert. Die Prozentzahlen auf der x-Achse geben an, an wie viel Prozent der Wahlbezirke die simulierte Nachwahlbefragung durchgeführt wurde, der Wert entspricht nicht dem Anteil an den Wahlberechtigten. Details zu den Szenarien und der Parametrisierung finden sich in Kapitel III.4.1.1..

werden. Der Unterschied zwischen den beiden Szenarien mit Nachwahlbefragung ist dabei relativ gering, die Verdoppelung der Individualdaten beim Szenario mit einer Nachwahlbefragung an 5% der Wahlbezirke führt nur noch zu einer relativ geringen Verbesserung der Schätzgüte. Beim Vergleich der Modelle mit Vorwissen und ohne Vorwissen kann jeweils eine etwas bessere Schätzgüte bei den Modellen mit Vorwissen festgestellt werden. Der Unterschied ist jedoch bei den hybriden Modellen deutlich kleiner als bei der ökologischen Inferenz.

Beim Szenario mit 600 Wahlbezirken zeigt sich ein anderes Bild. Bei der ökologischen Inferenz ist das Modell mit Vorwissen dem Modell ohne Vorwissen eindeutig unterlegen. Die Kombination von Individual- und Aggregatdaten beim hybriden Multinomial-Dirichlet Modell führt im Vergleich zur ökologischen Inferenz auch bei 600 Wahlbezirken zu einer relevanten Verbesserung der Schätzgüte. Jedoch kann die beste Schätzgüte beim Szenario mit der Nachwahlbefragung an 2,5% der Wahlbezirke erreicht werden. Die Verdoppelung der Individualdaten hat in dieser Simulation sogar einen leichten negativen Effekt und die Schätzgüte ist beim Szenario mit Nachwahlbefragung an 5% der Wahlbezirke etwas schlechter. Bei den hybriden Modellen ist kein relevanter Unterschied zwischen den Modellen mit und ohne Vorwissen identifizierbar.

Der Vergleich der Schätzgüten zwischen den Modellen bei 200 und 600 Wahlbezirken ergibt, dass das mehr an Aggregatdaten bereits zu einer deutlich besseren Schätzung der Wählerwanderung führt. Das beste hybride Modell bei 200 Wahlbezirken schlägt bei der Schätzgüte nur relativ knapp die ökologische Inferenz ohne Vorwissen bei 600 Wahlbezirken. Auch im Vergleich zu den bisherigen Simulationen, vergleiche Kapitel II.4.3.2. und III.4.1.2., sind die bei dieser Simulation bei der ökologischen Inferenz und bei 600 Wahlbezirken beobachteten ADs sehr niedrig. Der Unterschied zwischen den hybriden Modellen zwischen den Szenarien mit 200 und 600 Wahlbezirken fällt etwas niedriger aus, aber auch hier ist das beste Modell bei 600 Wahlbezirken besser als das beste Modell bei 200 Wahlbezirken.

Die bei der Simulation beobachtete Verbesserung bei den hybriden Modellen deckt sich auch mit den Erkenntnissen der Simulation in Kapitel III.4.1.2.. Die Struktur beim Szenario mit heterogener Verteilung der Übergangswahrscheinlichkeiten und dem vergleichbaren Exit-Poll ohne Bias ist fast identisch zur Struktur bei 200 Wahlbezirken in dieser Simulation. Die Simulation mit homogener Verteilung der Übergangswahrscheinlichkeiten legt bereits nahe, dass ab einem gewissen Niveau nur noch eine begrenzte Verbesserung der Schätzgüte durch das Berücksichtigen von

Individualdaten erreicht werden kann. Eine entsprechende Interpretation wird durch die eher geringer Verbesserung bei dem Szenario mit 600 Wahlbezirken in dieser Simulation, beim dem ebenfalls bereits eine relativ gute Schätzgüte mit der ökologischen Inferenz erreicht werden kann, weiter gestützt.

Die Verschlechterung der Schätzgüte beim Szenario mit 600 Wahlbezirken bei der ökologischen Inferenz mit Vorwissen kann vermutlich auf das sehr unspezifische Vorwissen zurückgeführt werden, welches nicht den zugrundeliegenden Übergangswahrscheinlichkeiten entspricht. Jedoch zeigt die deutliche Verbesserung beim Szenario mit 200 Wahlbezirken, dass auch ein solch unspezifisches Vorwissen sinnvoll sein kann und vor allem bei einer eher geringen Zahl von Wahlbezirken zu eine Verbesserung der Schätzgüte führen kann.

Das nicht besonders intuitive Ergebnis, dass ein mehr an Individualdaten beim Szenario mit 600 Wahlbezirken zu einer schlechteren Schätzung führt, ist vermutlich den spezifischen Charakter der verwendeten Individualdaten geschuldet. Wie bereits in Kapitel III.4.1.2. dargelegt, stellen die mittels Nachwahlbefragung erhobenen Individualdaten eine Annahmeverletzung des hybriden Multinomial-Dirichlet Modells dar. Für das Modell wird angenommen, dass die Übergangswahrscheinlichkeiten bei Individual- und Aggregation in einem Wahlbezirk identisch sind. Das komplette Fehlen der Nichtwähler der zweiten Wahl bei den Individualdaten stellt eine offensichtliche Verletzung dieser Annahme dar. Es liegt nahe, dass diese Verzerrung in den Individualdaten zur Verschlechterung der Schätzung in dieser Simulation geführt haben kann. Etwas Vergleichbares kann auch in der Simulation in Kapitel III.4.1.2. beobachtet werden, wo bei homogener Verteilung der Übergangswahrscheinlichkeiten der Exit-Poll mit Bias zu einer schlechteren Schätzung geführt hat, während bei heterogener Verteilung kein vergleichbarer Effekt zu beobachten ist.

III.4.2. Kreditibilitätsintervalle beim (hybriden) Multinomial-Dirichlet Modell

Bei den bisher vorgestellten Betrachtungen und den drei Simulationsstudien wird jeweils nur die Schätzgüte der Modelle evaluiert und verglichen. Ein ebenso wichtiger Aspekt bei der Schätzung ist jedoch die korrekte Quantifizierung der Unsicherheit der geschätzten Größen. Die Unsicherheit wird im Allgemeinen durch Konfidenzintervalle, oder bei Bayesianischen Schätzung durch Kreditibilitätsintervalle, spezifiziert. Die Angabe von Konfidenz- bzw. Kreditibilitätsintervallen ist bei der ökologischen Inferenz bei einem Teil der Modelle nicht ohne weiteres möglich, während bei anderen die Angabe theoretisch möglich ist, jedoch die Validität der Konfidenz- bzw. Kreditibilitätsintervalle kritisch gesehen werden muss.

Bei der ökologischen Regression nach Goodman können Schätzwerte außerhalb von $[0,1]$ beobachtet werden. Eine Korrektur dieser Schätzwerte ist möglich, jedoch ist nicht klar, wie die Konfidenzintervalle für die so korrigierten Schätzer bestimmt werden sollen.²⁰⁰ Thomsens Probit/Logit Modell selber bietet keine Schätzung der Unsicherheit an und das Modell liefert nur Punktschätzer. Park diskutiert mögliche Erweiterung des Modells und eine davon beschäftigt sich mit der Konstruktion von Konfidenzintervallen bei Thomsens Modell. Jedoch stellt Park selber fest, dass diese eine zu geringe Überdeckung aufweisen.²⁰¹ Auch beim adaptiven Modell von Andreadis und Chadjipadelis ist keine direkte Angabe der Schätzunsicherheit möglich. Dies ist vor allem der wiederholten und aufeinander aufbauend Nutzung des 2x2 Verfahrens geschuldet.²⁰²

Bei den vorgestellten hierarchischen Bayesianischen Modellen²⁰³ ist die Angabe der Schätzunsicherheit theoretisch recht einfach möglich. Die Ketten können als Posteriori-Verteilung der interessierenden Größen gegeben die Daten aufgefasst werden²⁰⁴ und es ist daher möglich, die Unsicherheit dieser Größen anhand der Ketten zu quantifizieren. Dabei können z. B., bei ausreichender Zahl an Ziehungen aus den Ketten, die Quantile dieser verwendet werden und als Quantile der interessierenden Verteilung für die Bildung des jeweiligen Kreditibilitätsintervalls aufgefasst werden. Jedoch zeigt sich bei der praktischen Anwendung, dass z. B. beim hybriden Multinomial-Normal Modell die Überdeckung der Kreditibilitätsintervalle nicht immer der Erwartung entspricht.²⁰⁵

200 Vgl. Kapitel II.3.1..

201 Vgl. Kapitel II.3.2..

202 Vgl. Kapitel II.3.3..

203 Vgl. Kapitel II.3.4. und III.2..

204 Vgl. Robert, 2007, S. 302–303.

205 Vgl. Greiner / Quinn, 2010, S. 1785.

Die Güte der Kreditibilitätsintervalle des (hybriden) Multinomial-Dirichlet Modells soll im Folgenden genauer evaluiert werden. Dies erfolgt primär anhand des Multinomial-Dirichlet Modells. Aufgrund der großen Ähnlichkeit der Modelle sind die so gewonnenen Erkenntnisse jedoch auch weitestgehend auf das hybride Modell übertragbar.

III.4.2.1. Güte der Kreditibilitätsintervalle

Die Evaluierung der Güte der Kreditibilitätsintervalle beim Multinomial-Dirichlet Modell erfolgt exemplarisch anhand der in Kapitel II.4.3. durchgeführten Simulationsstudie zum Vergleich der Methoden der ökologischen Inferenz. Da es sich um eine Simulationsstudie handelt, kann neben der Güte der Schätzung auch die Güte der Kreditibilitätsintervalle direkt spezifiziert werden. Die vier Szenarien der Simulation sind kurz in Tabelle 14 beschrieben und werden ausführlich in Kapitel II.4.3.1. vorgestellt. Zusätzlich zu dem bereits in Kapitel II.4.3. betrachteten Modell wird jedoch in diesem Kapitel für jedes Szenario auch eine alternative Schätzung mit dem Multinomial-Dirichlet Modell mit berücksichtigt.

Bisher wurden mit dem Multinomial-Dirichlet Modell nur Modelle betrachtet, bei denen die zweite Wahl in Abhängigkeit von der ersten Wahl gesehen wird. Für die zusätzlichen Modelle wird absichtlich diese Richtung umgekehrt und das Wahlergebnis der ersten Wahl wird als Ergebnis des Wahlergebnisses der zweiten Wahl betrachtet. D. h. auf der ersten Ebene des Modells wird angenommen, dass²⁰⁶:

$$(N_{CSU\ 2009,i}, N_{SPD\ 2009,i}, \dots, N_{NW\ 2009,i}) \sim MNL(\theta'_{CSU,i}, \theta'_{SPD,i}, \dots, \theta'_{NW,i}, N_i) \quad (54)$$

mit

$$\theta'_{CSU,i} = \sum_{c=1}^C \beta'_{CSU\ 2009,c,i} * P_{c,2013,i} \quad \& \quad \beta'_{CSU\ 2009,c,i} = \frac{N_{CSU\ 2009,c,i}}{N_{c,2013,i}}, \quad \forall c \in \{1, \dots, C\} \quad (55)$$

Diese Modelle werden nicht als realistische Alternative²⁰⁷ für die Schätzung der Wählerwanderung betrachtet. Sie sollen aber als eine Art negatives Extrembeispiel bei den anschließend durchgeführten Betrachtungen dienen. Um die Modellschätzungen unterscheiden zu können, wird im Folgenden der Begriff inverse (Multinomial-Dirichlet) Modelle für diese neuen Modelle verwendet.

206 Für die Darstellung wird weiterhin von einer mit Tabelle 1 vergleichbaren Tabelle ausgegangen. Die Multinomialverteilung bezieht sich daher auf den Zeilenrand und die Parameter dieser Verteilung sind jeweils die gewichtete Summe für eine Zeile über die Spalten.

207 Wenn bei der Schätzung mit dem Multinomial-Dirichlet Modell keine Abhängigkeit von der Verteilung der Übergangswahrscheinlichkeiten zwischen den Wahlbezirken bestehen würde, könnte mit beiden Modellen die gleiche Wählerwanderung geschätzt werden. Das aber eine entsprechende Abhängigkeit besteht, wurde bereits in Kapitel II.4.3. gezeigt.

Weiterführende Evaluation Multinomial-Dirichlet Modell: Ergebnisse der Simulation

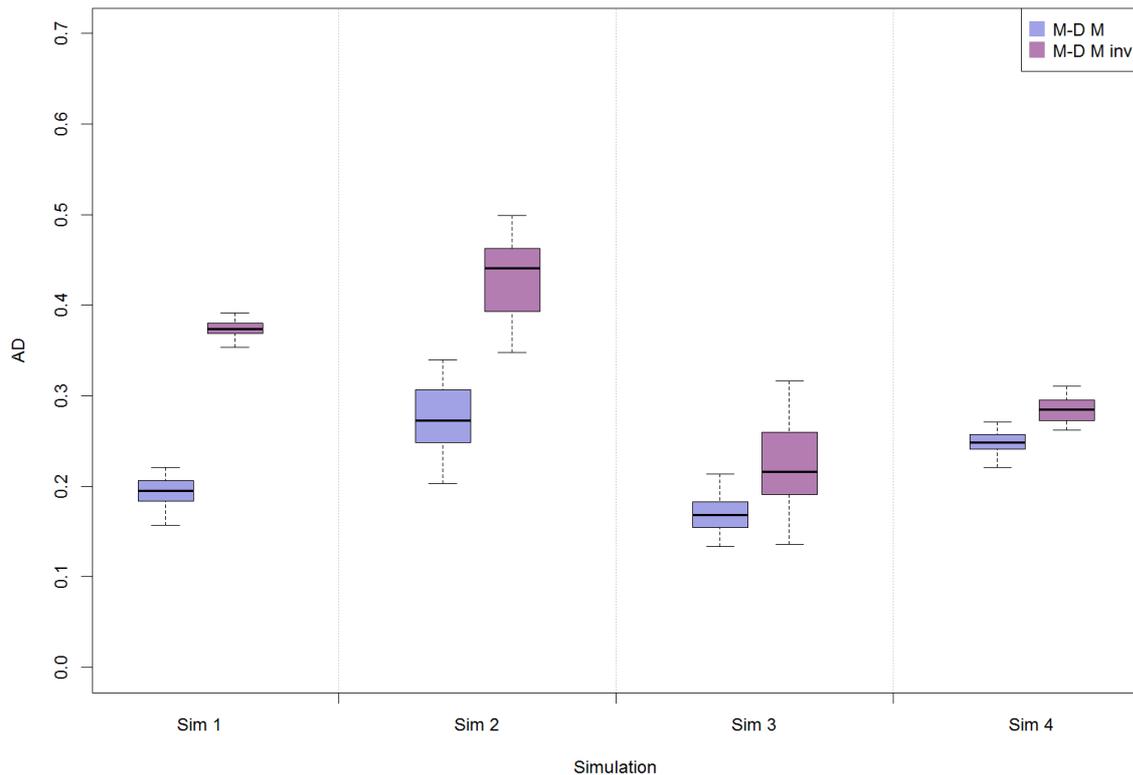


Abbildung 13: Detaillierte Evaluierung des (hybriden) Multinomial-Dirichlet Modells, Schätzgüte des Multinomial-Dirichlet Modells bei den Daten der Simulationsstudie zum Vergleich der Methoden der ökologischen Inferenz, dargestellt ist der AD der Schätzung zu den simulierten Individualdaten. Für jedes Szenario werden 30 Datensätze evaluiert. Eine Kurzdarstellung der verwendeten Szenarien befindet sich in Tabelle 14. Beim Modell M-D M wird die zweite Wahl in Abhängigkeit von der ersten Wahl betrachtet, beim Modell M-D M inv wird die erste Wahl in Abhängigkeit von der zweiten Wahl gesehen.

In Abbildung 13 werden die ADs der Modellschätzungen²⁰⁸ von den simulierten Individualdaten dargestellt. Wie bei den bisherigen Betrachtungen wird zur Quantifizierung der Schätzgüte der AD, vergleiche Gleichung (28), herangezogen. Die Schätzgüten mit den Multinomial-Dirichlet Modell und dem inversen Modell werden in der Abbildung direkt miteinander verglichen.

Es zeigt sich, dass die Schätzgüte beim inversen Modell deutlich schlechter als beim normalen Multinomial-Dirichlet Modell ist. Dies gilt vor allem für die Szenarien Sim 1 und Sim 2. Das Ergebnis betont die Relevanz der Verteilungsannahmen beim Multinomial-Dirichlet Modell für die Schätzgüte. Bei den Szenarien Sim 3 und Sim 4 ist ebenfalls eine

²⁰⁸ In dieser Simulationsstudie wird eine leicht von Kapitel II.4.3. abweichende Parametrisierung beim Multinomial-Dirichlet Modell verwendet. Für den Burn-In wird ein Wert von 100000 festgelegt, das Thinning beträgt 1000, es werden jeweils 400 Ziehungen aus den Ketten genutzt. Die Zahl der Ziehungen ist hier niedriger gewählt, um die notwendige Rechenzeit bei den späteren Betrachtungen zu begrenzen, dafür wird der Burn-In leicht erhöht. Es werden auch nur die ersten 30 Datensätze pro Szenario zur Evaluierung herangezogen.

Verschlechterung feststellbar, jedoch zeigt ein Vergleich mit den in Abbildung 8 gezeigten Ergebnissen für alle evaluierten Modelle, dass selbst die Schätzgüte der inversen Modelle mindestens vergleichbar mit der des zweitbesten Modells beim Modellvergleich ist.

Die Güte der 80%-Kredibilitätsintervalle für die mit den Modellen geschätzten $N_{r,c}$ ist in Abbildung 14 dargestellt. Die verwendeten Kredibilitätsintervalle werden über die entsprechenden Quantile der Ziehungen aus den Ketten der jeweiligen Größe gebildet. Für die Bestimmung der Güte der Kredibilitätsintervalle wird geprüft, ob der Wert in den simulierten Individualdaten in diesen liegt, er also vom Kredibilitätsintervall überdeckt wird. Die Darstellung in Abbildung 14 zeigt dabei den aggregierten Anteil über die gesamte Tabelle für einen Datensatz, d. h. ein Wert von 0,5 bedeutet, dass 50% der Kredibilitätsintervalle den Wert der simulierten Individualdaten überdeckt.

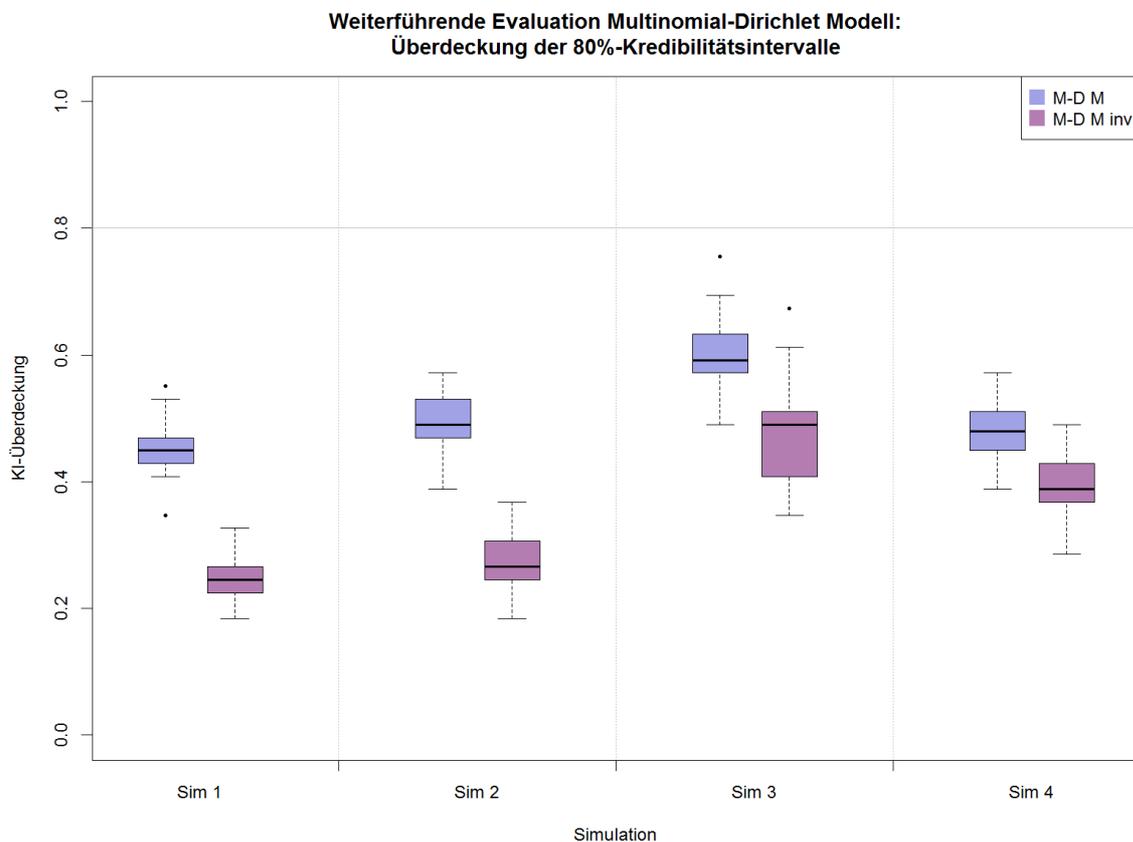


Abbildung 14: Detaillierte Evaluierung des (hybriden) Multinomial-Dirichlet Modells, Überdeckungsrate der 80%-Kredibilitätsintervalle der $N_{r,c}$ beim Multinomial-Dirichlet Modell bei den Daten der Simulationsstudie zum Vergleich der Methoden der ökologischen Inferenz, dargestellt ist für jede geschätzte Wählerwanderung der Anteil der Kredibilitätsintervalle, die den Wert in den simulierten Individualdaten überdecken. Für jedes Szenario werden 30 Datensätze evaluiert. Eine Kurzdarstellung der verwendeten Szenarien befindet sich in Tabelle 14. Beim Modell M-D M wird die zweite Wahl in Abhängigkeit von der ersten Wahl betrachtet, beim Modell M-D M inv wird die erste Wahl in Abhängigkeit von der zweiten Wahl gesehen. Die graue Linie markiert einen Anteil von 80%.

Es zeigt sich, dass die Überdeckungsraten der Kreditibilitätsintervalle weit von den angenommenen und bei validen Kreditibilitätsintervallen erwartbaren 80% entfernt sind. Der im Mittel beste Wert für das Multinomial-Dirichlet Modell kann bei Szenario Sim 3 mit ca. 60% beobachtet werden. Bei allen anderen Modellen und Szenarien sind die Überdeckungsraten – zum Teil deutlich – schlechter. Beim Vergleich der normalen Modelle mit den inversen Modellen fällt auf, dass die Überdeckungsraten bei letzteren jeweils schlechter sind. Dabei scheint die Stärke der Verschlechterung von der Schätzgüte an sich abzuhängen, da diese bei den Szenarien Sim 1 und Sim 2 deutlicher ausfällt. Beim Modell mit den schlechtesten Überdeckungsraten werden im Mittel nicht einmal 30% der Werte in den simulierten Individualdaten von den Kreditibilitätsintervallen überdeckt.

Eine genaue Betrachtung der mit dem Multinomial-Dirichlet Modell geschätzten Werte offenbart das Problem, dass vermutlich maßgeblich für die schlechten Überdeckungsraten bei den Kreditibilitätsintervallen verantwortlich ist. Die einzelnen Übergänge werden nicht Bias-frei geschätzt. In Abbildung 15 wird der Abstand der geschätzten Werte von den Individualdaten für das Szenario Sim 2 beim Multinomial-Dirichlet Modell präsentiert.

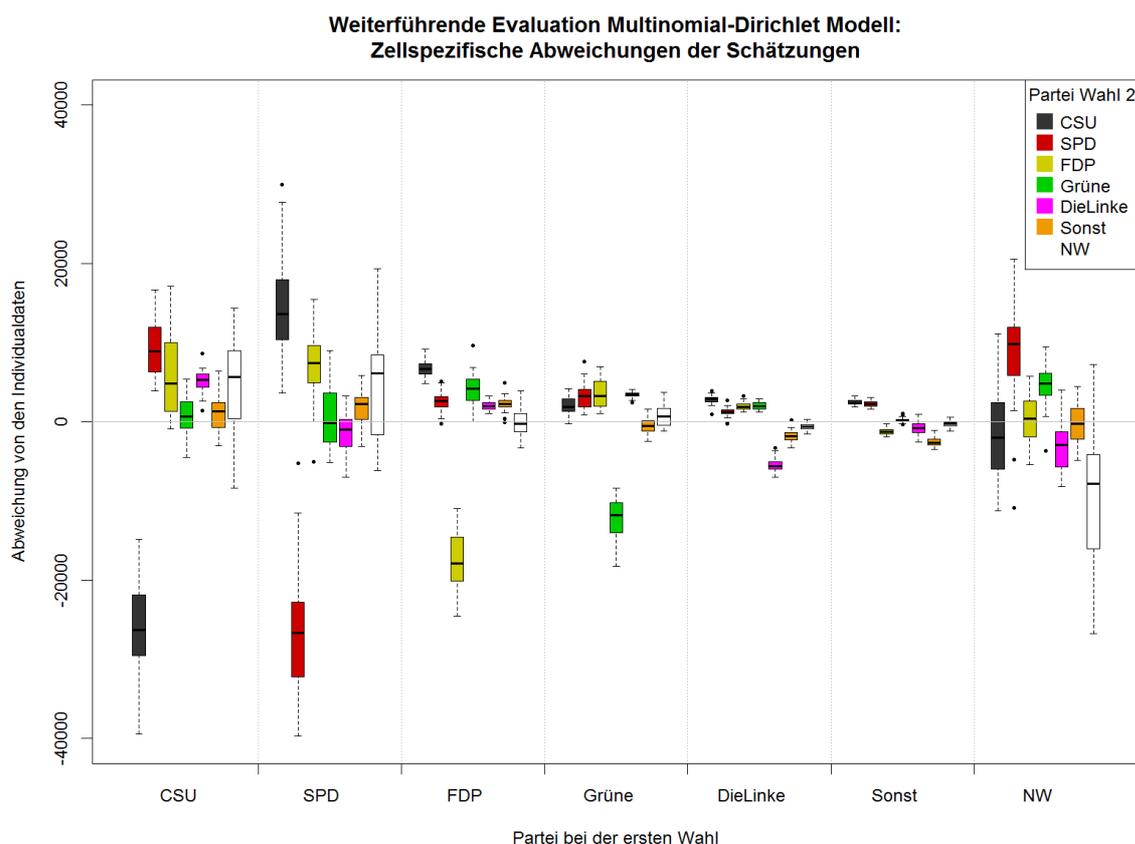


Abbildung 15. Detaillierte Evaluierung des (hybriden) Multinomial-Dirichlet Modells, Darstellung des zellspezifischen Bias bei der Schätzung mit dem Multinomial-Dirichlet Modell bei Szenario Sim 2, dargestellt ist die Abweichung in Wählern/Wahlberechtigten zwischen bei den 30 Datensätzen. Eine Kurzdarstellung des verwendeten Szenarios befindet sich in Tabelle 14.

Der Übergang von der CSU zur CSU, $N_{\text{CSU,CSU}}$, wird im Mittel um deutlich mehr als 20000 Wähler bei der Schätzung mit dem Multinomial-Dirichlet Modell bei diesem Szenario unterschätzt. Eine ähnliche Struktur, d. h. eine deutliche Unterschätzung, findet sich auch bei den Zellen mit den loyalen Wählern der SPD, der FDP, der Grünen und der Partei DieLinke. Wenig deutlich ist der Bias bei den Sonstigen. Auch bei den Nichtwählern kommt es tendenziell zu einer Unterschätzung, jedoch ist bei einigen Datensätzen auch eine Überschätzung feststellbar. Der Zusammenhang der inneren Zellen einer Wählerwanderungstabelle postuliert, dass eine Unterschätzung einer Zelle zu einer Überschätzung bei einer anderen Zelle führen muss, da die Randsumme fest vorgegeben ist. Dies ist auch in Abbildung 15 zu beobachten. Tendenzuell werden die Übergänge von einer Partei zu einer anderen Partei bei der Schätzung mit dem Multinomial-Dirichlet Modell überschätzt.

Es ist jedoch nicht möglich, anhand der Ergebnisse eindeutige Aussagen über die eigentliche Ursache für den Bias bei der Schätzung zu tätigen. Eine Überschätzung der Übergänge zu anderen Parteien führt automatisch zu einer Unterschätzung der Loyalität. Umgekehrt führt eine Unterschätzung der Loyalität zu einer tendenziellen Überschätzung bei den Übergängen zu einer anderen Partei. Festgehalten werden kann nur, dass es diesen nicht vernachlässigbaren Bias bei den geschätzten Größen gibt. Da das Multinomial-Dirichlet Modell noch mit die beste Schätzgüte alle verglichenen Modelle für die ökologische Inferenz aufweist, liegt es nahe zu vermuten, dass ähnliche Probleme auch bei den anderen Modellen existieren.

Für die Kreditibilitätsintervalle bedeutet der Bias bei der Schätzung der Wählerwanderung, dass alleinig die Berücksichtigung der Variabilität der Ketten im Allgemeinen nicht ausreichend für die Konstruktion von validen Kreditibilitätsintervallen sein wird. Der Fehler bei der Schätzung mit den Multinomial-Dirichlet Modell besteht aus einer zufälligen Komponente, der Schätzunsicherheit, und aus dem Bias.

$$N'_{r,c} - N_{r,c} = \epsilon_{r,c} + \text{Bias}_{r,c}, \quad \forall r \in \{1, \dots, R\}, c \in \{1, \dots, C\} \quad (56)$$

mit

$$E(\epsilon_{r,c}) = 0, \quad \forall r \in \{1, \dots, R\}, c \in \{1, \dots, C\} \quad (57)$$

$N'_{r,c}$ stellt in Gleichung (56) die Modellschätzung dar, $N_{r,c}$ repräsentiert den Wert in Individualdaten und $\epsilon_{r,c}$ ist der zufällige Fehler der Schätzung. Für den zufälligen Fehler wird angenommen, dass der im Mittel null beträgt. Die Ketten beim Multinomial-Dirichlet

Modell berücksichtigen jedoch nur die normale Schätzunsicherheit in Gleichung (56) und nicht den strukturellen Bias bei der Schätzung. Kreditibilitätsintervalle, die auf den Ketten basieren, berücksichtigen daher nur eine der beiden Fehlerquellen.

Gleichung (56) legt auch nahe, dass eine Reduktion des strukturellen Schätzfehlers zu einer besseren Überdeckung bei den aus den Ketten gebildeten Kreditibilitätsintervalle führt. Ein Vergleich von Abbildung 13 und Abbildung 14 zeigt jedoch, dass nicht notwendigerweise eine bessere Schätzgüte zu einer besseren Überdeckung führt. Während beim Szenario Sim 3 die Schätzgüten und die Überdeckungsraten der Kreditibilitätsintervalle jeweils am besten sind, gilt beim Szenario Sim 1, dass trotz relativ guter Schätzgüte die Überdeckungsraten der Kreditibilitätsintervalle beim Multinomial-Dirichlet Modell am schlechtesten sind.

Bei der Schätzung der Wählerwanderung bei den gleichen Daten kann jedoch davon ausgegangen werden, dass jede Verbesserung der Schätzgüte beim (hybriden) Multinomial-Dirichlet Modell tendenziell auch den strukturellen Fehler in Gleichung (56) reduziert. Eine Verbesserung kann z. B. durch die Berücksichtigung von Vorwissen beim Multinomial-Dirichlet Modell erreicht werden. Eine noch deutlichere Verbesserung der Schätzgüte wird durch die Kombination von Individual- und Aggregatdaten beim hybriden Multinomial-Dirichlet Modell erreicht. In beiden Fällen kann daher auch von einer besseren Überdeckungsrate der Kreditibilitätsintervalle ausgegangen werden. Jedoch gilt auch dann, dass die Überdeckungsraten der Kreditibilitätsintervalle nicht notwendigerweise der Erwartung entsprechen müssen.

III.4.2.2. Korrektur der Kreditibilitätsintervalle

Da die direkt aus den Ketten gewonnenen Kreditibilitätsintervalle eine nur unzureichende Überdeckungsrate aufweisen, sind sie in der vorliegenden Form nicht geeignet, die Schätzunsicherheit adäquat darzustellen. Weiterhin fällt bei einer detaillierten Betrachtung der Ketten auf, dass nur ein Teil der Ketten überhaupt den Wert in den Individualdaten überdeckt. Daher sind auch die Ketten in der vorliegenden Form nicht geeignet, adäquate Kreditibilitätsintervalle zu bilden und eine weitergehende Korrektur scheint notwendig zu sein. In diesem Kapitel werden vier Verfahren zur Korrektur der Kreditibilitätsintervalle beim Multinomial-Dirichlet Modell vorgestellt und auf ihre Eignung überprüft. Alle vier Verfahren basieren lose auf dem Bootstrap-Prinzip.

Beim Bootstrap handelt es sich um ein Resampling-Verfahren, mit dem allgemein die Unsicherheit eines Schätzers oder einer Schätzstatistik bestimmt werden kann. Dafür

werden B Bootstrap-Stichproben generiert. Vereinfacht kann dabei zwischen den empirischen und den parametrischen Bootstrap unterschieden werden. Beim empirischen Bootstrap sind die Daten direkter Ausgangspunkt für das Resampling. Es wird dabei ausgenutzt, dass die Verteilung hinter den Daten durch die empirische Verteilungsfunktion der Daten approximiert wird, aus welcher dann ersatzweise gezogen werden kann. Da Stichproben mit identischer Größe benötigt werden, erfolgt das Ziehen der Bootstrap-Stichproben mit Zurücklegen. Beim parametrischen Bootstrap kann die Form der Verteilung der Daten spezifiziert werden. Daher ist es nur notwendig, die Parameter dieser Verteilung aus den Daten zu schätzen. Das Ziehen der Bootstrap-Stichproben kann jedoch direkt aus der Verteilung erfolgen. Der interessierende Schätzer bzw. die interessierende Schätzstatistik wird anschließend für alle B Bootstrap-Stichproben bestimmt und die Variabilität zwischen den B Bootstrap-Stichproben kann zur Schätzung der Unsicherheit verwendet werden.²⁰⁹

Beim ersten Korrekturverfahren wird die Idee des empirischen Bootstraps aufgegriffen. Aus den beobachteten Daten werden B neue Datensätze der gleichen Größe durch Ziehen mit Zurücklegen generiert. Ausgangspunkt für das Resampling sind dabei die Wahlbezirke. Die Wählerwanderung wird im Anschluss für jeden der B Datensätze mit dem Multinomial-Dirichlet Modell geschätzt. Die Hoffnung bei diesem Korrekturverfahren ist, dass durch das Resampling der Bias bei einem Teil der Schätzungen reduziert werden kann oder dass zumindest die Unsicherheit adäquater in den Ketten der Modelle abgebildet wird. Daher werden abweichend vom Bootstrap-Prinzip für das Korrekturverfahren nicht die Schätzer, sondern die Ketten direkt verwendet. Diese werden zu einer „Gesamt-Posteriori“ kombiniert, aus welcher dann die Kreditibilitätsintervalle gebildet werden.

Die anderen drei Korrekturverfahren greifen die Ideen des parametrischen Bootstraps und von Simulationsstudien auf. Für die drei Verfahren werden ausgehend von einer ersten Modellschätzung B Wählerwanderungen auf Individualebene simuliert. Ausgangspunkt der Simulation sind die mit dem Multinomial-Dirichlet Modell geschätzten gemeinsamen Dirichlet-Verteilungen der Übergangswahrscheinlichkeiten. Aus diesen werden für jeden Wahlbezirk neue Übergangswahrscheinlichkeiten gezogen, mit welchen im Anschluss die individuelle Wählerwanderung simuliert wird. Als parametrischen Bootstrap können die Verfahren angesehen werden, da die Verteilungsannahmen des Multinomial-Dirichlet Modells und die aus den Daten geschätzten Parameter dieser Verteilungen für die

209 Vgl. Efron, Bradley / Tibshirani, Robert: Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical science*, 1(1), 1986, S. 54–57 und Horowitz, Joel L.: The bootstrap. *Handbook of econometrics*, 5, 2001, S. 3160–3165.

Simulation der Individualdaten genutzt wird. Für das weitere Vorgehen werden aus den Individualdaten neue Aggregatdaten berechnet und für diese wird mit dem Multinomial-Dirichlet Modell die Wählerwanderung geschätzt.

Die Hoffnung bei den drei Verfahren ist, dass mit der wiederholten Schätzung und der Quantifizierung der Schätzfehler, möglich aufgrund der vorliegenden simulierten Individualdaten, eine Bestimmung der Bias-Komponente in Gleichung (56) möglich ist. Bei den Untersuchungen zeigt sich jedoch, dass eine valide Schätzung der Richtung des Bias mit dieser Methode nicht in allen Fällen gelingt. Daher wird die gewonnene Information zur Korrektur der Varianz der Ketten verwendet, so dass diese im Anschluss die Schätzunsicherheit und die zusätzliche Unsicherheit aufgrund des Bias bei der Schätzung, ohne dass dabei eine bestimmte Richtung angenommen wird, berücksichtigen. Der Unterschied zwischen den drei Verfahren ist, bei welcher Größe der Bias betrachtet wird.

Bei allen drei Verfahren erfolgt die Korrektur für die geschätzte Wählerwanderung auf Ebene der Gesamtregion. Ausgangspunkt sind die Ketten der $N'_{r,c}$. Beim zweiten Korrekturverfahren wird die relative Abweichung des mit dem Multinomial-Dirichlet Modells bestimmten Schätzwertes für die Wählerwanderung vom Wert in den Individualdaten betrachtet, d. h. der Bias stellt die prozentuale strukturelle Abweichung der Schätzung dar. Für das dritte Korrekturverfahren werden die $N'_{r,c}$ transformiert und es wird der Bias der geschätzten Übergangswahrscheinlichkeiten für die Gesamtregion in Prozentpunkten betrachtet. Beim zweiten und dritten Korrekturverfahren ist es ohne weiteres möglich, dass die korrigierten Kreditabilitätsintervalle Werte kleiner null beinhalten.

Dies wird beim vierten Korrekturverfahren ausgeschlossen. Betrachtet werden, wie beim dritten Korrekturverfahren, erneut die Übergangswahrscheinlichkeiten, jedoch wird beim vierten Korrekturverfahren der Bias der Logit-transformierten Übergangswahrscheinlichkeiten verwendet. Ein Nachteil des vierten Korrekturverfahrens ist, dass der Mittelwert der Ketten nach der Korrektur nicht mehr identisch mit den Mittelwert der Ketten vor der Korrektur ist. Eine Kurzübersicht aller vier Korrekturverfahren ist in Tabelle 21, eine formale Darstellung in Anhang X.3. dargestellt.

Die Evaluierung der vier Korrekturverfahren erfolgt mit den bereits in Kapitel III.4.2.1. betrachteten Daten und Modellen. Dies ermöglicht einen direkten Vergleich der Überdeckungsraten vor und nach der Korrektur der Kreditabilitätsintervalle. Es werden

Korrekturverfahren	Bootstrap-Prinzip	Korrekturprinzip	betrachtete Größe bei der Bestimmung des Bias
1	empirisch	Kombination von mehreren Ketten	-
2	parametrisch	Bias-Berücksichtigung in der Kettenvarianz	rel. Abweichung $N'_{r,c}$
3	parametrisch	Bias-Berücksichtigung in der Kettenvarianz	Übergangswahrscheinlichkeiten
4	parametrisch	Bias-Berücksichtigung in der Kettenvarianz	Logit-transformierte Übergangswahrscheinlichkeiten

Tabelle 21: Detaillierte Evaluierung des (hybriden) Multinomial-Dirichlet Modells, Korrektur der Kreditabilitätsintervalle, Übersicht über die vier betrachteten Korrekturverfahren für die Kreditabilitätsintervalle. Die erste Spalte gibt die in Grafiken genutzte Nummer an. Details zu den Korrekturverfahren finden sich in Kapitel III.4.2.2..

jeweils alle vier Korrekturverfahren für jedes Modell und Szenario genutzt.²¹⁰ Die Zahl der (Bootstrap-)Stichproben liegt bei allen vier Verfahren bei 10. Für die drei sich am parametrischen Bootstrap orientierenden Verfahren wird der maximale Abstand eines Schätzwertes vom Wert in den Individualdaten über alle 10 Stichproben zur Bias-Korrektur für alle Ketten verwendet. Dieser Wert wird bei der Evaluierung genutzt, da er als beste Schätzung für den maximalen Bias einer Zelle angesehen wird und spezifischere Aussagen über die Größenordnung des Bias mit den vorgeschlagenen Korrekturverfahren in der aktuellen Form nicht möglich sind. Die Überdeckungsraten der so korrigierten 80%-Kreditabilitätsintervalle sind in Abbildung 16 gezeigt. Wie bereits in Abbildung 14 wird der aggregierte Anteil über die gesamte Tabelle für einen Datensatz dargestellt.

Das erste Korrekturverfahren führt zu einer geringfügigen Verbesserung der Überdeckungsraten der Kreditabilitätsintervalle. Diese sind im Vergleich zu den Überdeckungsraten der Kreditabilitätsintervalle ohne Korrektur im Mittel um 5 Prozentpunkte höher, jedoch liegen sie bei allen Szenarien und beiden Modellen noch – meistens sehr deutlich – unter 80%. Beim zweiten Korrekturverfahren zeigt sich ein anderes Bild, in allen Szenarien und für alle Modelle sind die Überdeckungsraten deutlich größer als 80%. Es gibt keine einzige geschätzte Wählerwanderung, bei der weniger als 80% der Kreditabilitätsintervalle den Wert in den Individualdaten überdecken. Die Kreditabilitätsintervalle nach dem dritten Korrekturverfahren zeigen ebenfalls deutlich bessere Überdeckungsraten. Im Mittel überdecken bei Szenario Sim 1 und Sim 2 80% der Kreditabilitätsintervalle einer geschätzten Wählerwanderungstabelle beim normalen Multinomial-Dirichlet Modell den Wert der Individualdaten. Bei Sim 3 und Sim 4 sind es im Mittel etwas mehr als 80%. Beim inversen Modell überdecken fast immer alle

²¹⁰ Es wird für jedes gerechnete Multinomial-Dirichlet Modelle die gleiche Parametrisierung wie in Kapitel III.4.2.1. verwendet (Burn-In: 100000; Thinning: 1000; Ziehungen: 400).

Weiterführende Evaluation Multinomial-Dirichlet Modell:
Überdeckung der adjustierten 80%-Kreditabilitätsintervalle

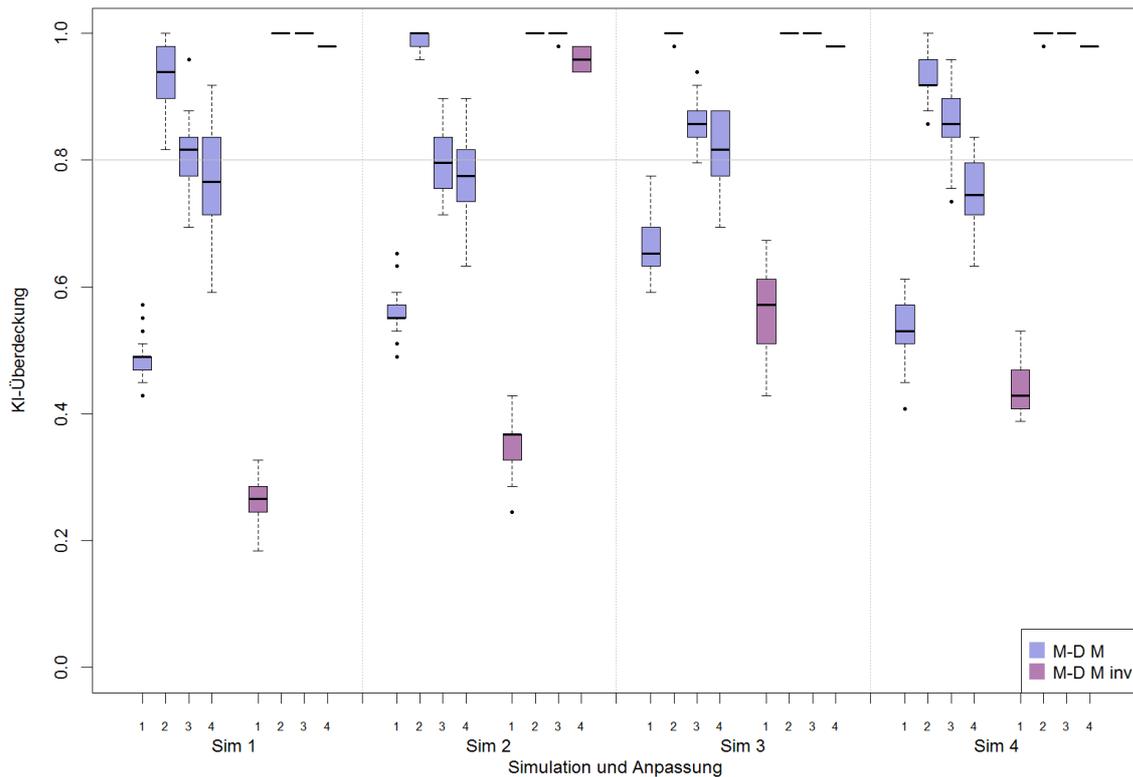


Abbildung 16: Detaillierte Evaluierung des (hybriden) Multinomial-Dirichlet Modells, Korrektur der Kreditabilitätsintervalle, Überdeckungsrate der korrigierten 80%-Kreditabilitätsintervalle bei den Daten der Simulationsstudie zum Vergleich der Methoden der ökologischen Inferenz, dargestellt ist für jedes Modell und Korrekturverfahren der Anteil der Kreditabilitätsintervalle, die den Wert in den simulierten Individualdaten überdecken. Eine Kurzdarstellung der verwendeten Szenarien befindet sich in Tabelle 14, die Korrekturverfahren sind in Tabelle 21 zusammengefasst. Beim Modell M-D M wird die zweite Wahl in Abhängigkeit von der ersten Wahl betrachtet, beim Modell M-D M inv wird die erste Wahl in Abhängigkeit von der zweiten Wahl betrachtet. Die graue Linie markiert einen Anteil von 80%. Es werden die Ergebnisse für 30 Datensätze pro Szenario gezeigt.

Kreditabilitätsintervalle den Wert der Individualdaten. Das vierte Korrekturverfahren zeigt etwas schlechtere Überdeckungsrate als das dritte Verfahren. Beim normalen Multinomial-Dirichlet Modell sind die Überdeckungsrate bei Szenario Sim 1, Sim 2 und Sim 4 im Mittel leicht unter 80%. Bei allen anderen Modellen und Szenarien sind auch bei diesem Korrekturverfahren die Überdeckungsrate größer als 80%.

Die Ergebnisse zeigen, dass das erste Korrekturverfahren nicht geeignet ist. Es kann keine relevante Verbesserung bei der Überdeckung erreicht werden. Mit dem zweiten Korrekturverfahren können zwar deutlich bessere Überdeckungsrate erreicht werden, jedoch sind die Kreditabilitätsintervalle in dieser Simulation deutlich zu konservativ, da immer mehr als 80% der 80%-Kreditabilitätsintervalle einer Wählerwanderungstabelle den Wert der Individualdaten überdecken. Das dritte und vierte Korrekturverfahren zeigen

relativ vergleichbare Ergebnisse, wobei mit dem dritten Korrekturverfahren in dieser Simulation etwas bessere Überdeckungsrate erreicht werden können. Jedoch sind auch bei diesem Korrekturverfahren beim inversen Modell die Überdeckungsrate deutlich zu groß.

Bei der Bewertung der Kreditibilitätsintervalle ist jedoch nicht alleine der Anteil der Kreditibilitätsintervalle interessant, die bei einer Wählerwanderungstabelle den Wert der Individualdaten überdecken. Für die eigentliche Güte ist die bisher nicht betrachtete Überdeckungsrate pro Zelle relevant. Bei einer detaillierten Betrachtung der jeweils 30 Kreditibilitätsintervalle pro Szenario, Modell und Zelle zeigt sich, dass auch nach der Korrektur die Überdeckungsrate bei einigen Zellen deutlich zu niedrig sind, während sie bei anderen deutlich zu groß sind. So überdecken z. B. bei Szenario Sim 2, den normalen Modell und Korrekturverfahren drei die Kreditibilitätsintervalle für $N_{CSU,CSU}$ nur in zwei von zehn Fällen den Wert in den Individualdaten. Es lässt sich daher festhalten, dass auch mit den Korrekturverfahren keine validen Kreditibilitätsintervalle konstruiert werden können. Es ist jedoch mit diesen Korrekturverfahren möglich, Kreditibilitätsintervalle zu konstruieren, die auf Ebene der Tabelle das Niveau einhalten, d. h. die Überdeckungsrate auf Ebene der Tabelle entspricht der Erwartung. Dies stellt eine deutliche Verbesserung zu den nicht korrigierten Kreditibilitätsintervalle dar.

Die nach der Korrektur erhaltenen Kreditibilitätsintervalle sind im Allgemeinen relativ groß und implizieren eine relativ große Unsicherheit der Schätzung. Abbildung 17 zeigt exemplarisch die 80%-Kreditibilitätsintervalle für Korrekturverfahren drei für den ersten Datensatz von Szenario Sim 2. Die Darstellung zeigt, dass trotz der relativ großen 80%-Kreditibilitätsintervalle diese nicht vollkommen nicht-informativ sind. So ist zum Beispiel erkennbar, dass eine fünfstellige Zahl von CSU-Wählern der ersten Wahl Nichtwähler bei der zweiten Wahl war.

Zusammenfassend lässt sich sagen, dass zwei der vier Korrekturverfahren Potential zeigen. Es handelt sich um die beiden Korrekturverfahren, bei denen der Bias der Übergangswahrscheinlichkeiten mit einem am parametrischen Bootstrap und Simulationsstudien angelehnten Verfahren geschätzt wird. Mit dieser Schätzung werden anschließend die Ketten angepasst, so dass diese die Schätzunsicherheit und den Bias berücksichtigen. Mit diesem Verfahren kann eine deutliche Verbesserung im Vergleich zur Ausgangssituation erreicht werden kann. Auf Ebene der Wählerwanderungstabelle kann bereits eine gute Überdeckungsrate der Kreditibilitätsintervalle erreicht werden.

Weiterführende Evaluation Multinomial-Dirichlet Modell: unter Grenze 80%-Kredibilitätsintervall								
2005	CSU	109921	0	6638	0	0	0	14986
	SPD	2532	54410	0	2856	0	0	18143
	FDP	0	0	35419	0	0	0	0
	Grüne	0	0	0	49604	0	0	0
	DieLinke	827	608	622	440	1487	922	663
	Sonst	747	648	534	102	377	899	592
	NW	10079	0	0	0	0	0	124777
		CSU	SPD	FDP	Grüne 2009	DieLinke	Sonst	NW
Weiterführende Evaluation Multinomial-Dirichlet Modell: Punktschätzer (Mittelwert der Ketten)								
2005	CSU	134743	13934	27609	12481	4122	7488	39576
	SPD	20711	73354	17945	21652	8221	4252	39645
	FDP	7743	6532	44463	6596	2597	3632	7064
	Grüne	4916	5714	6478	59566	3789	3879	8985
	DieLinke	3473	3320	3323	3022	4391	3655	3592
	Sonst	3162	2803	2430	1984	2202	2878	2729
	NW	33820	22891	9709	12963	13657	15372	152299
		CSU	SPD	FDP	Grüne 2009	DieLinke	Sonst	NW
Weiterführende Evaluation Multinomial-Dirichlet Modell: obere Grenze 80%-Kredibilitätsintervall								
2005	CSU	157025	39406	51199	37329	29284	31143	62330
	SPD	42106	92151	36716	40327	27477	23809	60423
	FDP	16279	15551	53726	14696	11232	11843	15634
	Grüne	14807	15428	15609	68811	13177	12796	19107
	DieLinke	6304	6195	6563	6099	7188	6451	6847
	Sonst	5504	5201	4836	4071	4453	4930	5014
	NW	58540	47919	34065	38370	37611	38772	176888
		CSU	SPD	FDP	Grüne 2009	DieLinke	Sonst	NW

Abbildung 17: Detaillierte Evaluierung des (hybriden) Multinomial-Dirichlet Modells, Korrektur der Kreditabilitätsintervalle, 80% Kreditabilitätsintervalle beim Multinomial-Dirichlet Modell nach Anwendung des Korrekturverfahrens drei bei Datensatz eins von Szenario Sim 2. Eine Kurzdarstellung des verwendeten Szenarios befindet sich in Tabelle 14, das Korrekturverfahren ist kurz in Tabelle 21 zusammengefasst. Die obere Tabelle zeigt die untere Grenze der Kreditabilitätsintervalle, die mittlere Tabelle die Mittelwerte der Ketten und die untere Tabelle die obere Grenze der Kreditabilitätsintervalle. Werte kleiner null werden auf null gesetzt.

Jedoch halten auch die auf diesem Weg konstruierten Kreditabilitätsintervalle auf Ebene der einzelnen Zellen nicht ihr Niveau. Dies veranschaulicht, dass noch weiterer Forschungsbedarf bei diesem Thema besteht. Bei den beiden Korrekturverfahren wäre eine spezifischere Schätzung des Bias eine Möglichkeit, welche das bisher noch bestehende Problem mit der unzureichenden Überdeckung der Kreditabilitätsintervalle zu lösen vermag. Dies könnte gleichzeitig zu einer Verkleinerung der Kreditabilitätsintervalle bei anderen Zellen führen, welche in der bisherigen Implementierung zu konservativ und unnötig groß

sind.

Die in diesem Kapitel untersuchten Korrekturvorschläge stellen auch nur einen Teil der möglichen Lösungen dieses Problems dar. Ein anderer Ansatz zur Konstruktion von adäquaten Konfidenzintervallen, der auf der Idee der Maximum-Entropie-Methode beruht, wurde z. B. von Elff, Gschwend und Johnston vorgeschlagen.²¹¹ Bei hybriden Modellen wären Ansätze denkbar, die die Informationen in den Individualdaten zur Schätzung des Bias oder zur Bildung von Konfidenz- bzw. Kreditabilitätsintervallen heranziehen. Da beim Multinomial-Dirichlet Modell vermutlich vor allem der Bias bei der Schätzung für die schlechten Überdeckungsraten der Kreditabilitätsintervalle verantwortlich ist, könnten auch Modellverbesserungen, die den Bias bei der Schätzung reduzieren würden, zur Lösung des Problems beitragen.

211 Vgl. Elff / Gschwend / Johnston, 2008, S. 78–82.

III.5. Hybride Modelle: Schlussbetrachtung

Mit den hybriden Modellen steht eine weitere Klasse von Modellen für die Schätzung von Wählerwanderungen zur Verfügung. Die Modelle nutzen, wie die Methoden für die ökologische Inferenz, die vorhandenen Aggregatdaten. Zusätzlich werden jedoch auch Individualdaten zur Schätzung mit herangezogen. Dabei kann zwischen zwei Arten der Verwendung bei den Modellen unterschieden werden. Ein Teil der Modelle nutzt die Individualdaten, um die Struktur der Kovariablen besser beschreiben zu können, ohne dass die Individualdaten auch für den interessierenden Zusammenhang vorliegen. Beim anderen Teil der Modelle beschreiben die Individualdaten auch den interessierenden Zusammenhang. Es handelt sich daher bei der Schätzung mit diesen Modellen nicht um eine ökologische Inferenz im strengeren Sinne.

Für die Schätzung von Wählerwanderungen sind vor allem Modelle interessant, bei denen die Individualdaten auch den interessierenden Zusammenhang beschreiben. In der Arbeit werden das hybride Multinomial-Normal Modell und das hybride Multinomial-Dirichlet Modell vorgestellt und genauer betrachtet. Beide Modelle stellen eine Erweiterung des entsprechenden in Kapitel II. vorgestellten Modells für die ökologische Inferenz dar. Für das hybride Multinomial-Normal Modell konnte bereits von Greiner und Quinn gezeigt werden, dass die Modellschätzung mit diesem Modell der ökologischen Inferenz und zum Teil auch der aus den Individualdaten überlegen ist.

Die vorgestellte Simulationsstudie für die hybriden Modelle zeigt, dass entsprechendes auch für das hybride Multinomial-Dirichlet Modell gilt. Die Schätzung mit dem hybriden Modell ist der ökologischen Inferenz und, bei den Szenarien mit Nachwahlbefragung, der Hochrechnung aus den Individualdaten überlegen. Bei einer angenommenen „perfekten“ Stichprobe ist jedoch die Hochrechnung aus den Individualdaten besser. Dies ist insofern beachtenswert, da bei der Schätzung mit den hybriden Modellen zusätzlich zu den Individualdaten eine weitere Informationsquelle mit herangezogen wird. Erwartbar wäre, dass dieses mehr an Informationen zu einer besseren Schätzung führt. Eine naheliegende Erklärung ist, dass die zusätzlichen Modellannahmen bei den hybriden Modellen ebenfalls Einfluss auf die Schätzung nehmen und dies zu dem beobachteten Ergebnis führt. Bei den simulierten Nachwahlbefragungen ist jedoch die Schätzung mit den hybriden Modellen zu bevorzugen und bei der Nachwahlbefragung mit Bias zeigt sich die Stärke der hybriden Modelle.

Der Vergleich der Schätzungsgüten der beiden hybriden Modelle in der Simulationsstudie

ergibt keinen eindeutigen Gewinner. Jedoch zeigt das hybride Multinomial-Dirichlet Modell bei den als realistischer empfundenen Szenarien die besseren Ergebnisse. Zusätzlich sind auch die Ergebnisse der Simulation in Kapitel II.4.3. zu berücksichtigen. Bei den dort untersuchten vergleichbaren Szenario zeigt ebenfalls das Multinomial-Normal Modell die beste Schätzgüte, schneidet jedoch bei den anderen Szenarien zum Teil deutlich schlechter ab. Bei der Betrachtung der Konvergenz der Ketten beim hybriden Multinomial-Normal Modell zeigen sich keine schwerwiegenden Probleme. Die Wahl der Referenzpartei nimmt jedoch weiterhin Einfluss auf die Schätzung, jedoch nur in vergleichsweise begrenzten Maße. Dieses Ergebnis könnte zum einen auf die zusätzlichen Informationen in den Individualdaten zurückgeführt werden, wahrscheinlich ist aber auch, dass die insgesamt deutlich bessere Schätzgüte mit zu den niedrigeren Werten beiträgt. Zusammenfassend ergibt sich daher beim Vergleich der Modelle ein Vorteil zugunsten des hybriden Multinomial-Dirichlet Modells.

Das (hybride) Multinomial-Dirichlet Modell wird daher anschließend detaillierter untersucht. Die erste Detailbetrachtung widmet sich dabei der Größe der betrachteten Region, dem Umfang der Individualdaten und dem Vorwissen. Die Berücksichtigung von Letzteren macht vor allem bei kleineren Regionen mit wenig Gebietseinheiten Sinn. In der Simulationsstudie ist die Schätzgüte in der großen Region mit Vorwissen sogar schlechter. Die Berücksichtigung von Individualdaten beim hybriden Modell sorgt in beiden Fällen zu einer deutlichen Verbesserung der Schätzgüte, die Verdoppelung des Umfangs hat jedoch nur einen vergleichsweise geringen Einfluss. Es scheint, dass bereits ein relativ geringer Umfang bei den Individualdaten die Schätzgüte erhöht, aber für jede weitere Verbesserung der Schätzgüte ein deutlich größerer Aufwand betrieben werden muss. Die Ergebnisse über beide Simulationsstudien mit den hybriden Modellen legen sogar nahe, dass bei einer an sich guten Schätzgüte ein mehr an verzerrten Individualdaten zu einer schlechteren Schätzung führt und weniger in diesen Fällen mehr sein kann.

Die Quantifizierung der Unsicherheit ist bei der ökologischen Inferenz und bei hybriden Modellen im Allgemeinen nicht ohne weiteres und problemfrei möglich. Dies ist insofern problematisch, da die Güte der Schätzung mit den Modellen bei der praktischen Nutzung nur schwer bewertet werden kann und daher eine korrekte Spezifikation der Unsicherheit umso wichtig wäre. Die Betrachtung der Kreditabilitätsintervalle beim Multinomial-Dirichlet Modell zeigt jedoch, dass auch diese eine nur unzureichende Überdeckung aufweisen. Es zeigt sich weiterhin, dass die Schätzung der Übergänge nicht Bias-frei erfolgt. Die Vermutung ist, dass vor allem der Bias bei den Kreditabilitätsintervallen nicht

berücksichtigt wird und dies zu der zu geringen Überdeckung führt.

Zur Korrektur der Kreditibilitätsintervalle werden vier Korrekturverfahren vorgestellt und untersucht. Alle basieren auf der Idee des Bootstraps und greifen zum Teil auch die Ideen von Simulationsstudien auf. Gute Ergebnisse können vor allem mit den Korrekturverfahren erreicht werden, die den möglichen Bias versuchen zu schätzen und zu berücksichtigen. Dies legt nahe, dass der Bias wirklich ursächlich für die schlechte Güte der Kreditibilitätsintervalle ist. Die Überdeckungsrate der so korrigierten Kreditibilitätsintervalle sind deutlich besser und die Zahl der Kreditibilitätsintervalle, die bei einer geschätzten Tabelle den wahren Wert überdecken, entspricht im Mittel in der Simulationsstudie mindestens der Erwartung. Die so erhaltenen korrigierten Kreditibilitätsintervalle sind jedoch sehr groß und zum Teil auch zu groß, da auch deutlich zu gute Überdeckungsrate erreicht werden. Weiterhin kann die verbesserte Überdeckung nur auf Ebene der Tabelle, jedoch nicht für die Kreditibilitätsintervalle der einzelnen geschätzten Übergänge erreicht werden. Deren Überdeckung ist, abhängig vom betrachteten Übergang, entweder ungenügend oder zu gut. Die mit den untersuchten Methoden erfolgte Korrektur ist bisher aber nur ein sehr grobes und sehr unspezifisches Vorgehen. Bessere Kreditibilitätsintervalle könnten z. B. durch eine spezifischere Schätzung des Bias der einzelnen Übergänge erreicht werden.

IV. Praktische Fragestellungen und Anwendungsbeispiel

In den Kapiteln II. und III. werden zwei Modellklassen zur Schätzung von Wählerwanderungen, die ökologische Inferenz und die hybriden Modelle, vorgestellt und evaluiert. Die ökologische Inferenz nutzt zur Schätzung nur Aggregatdaten, während bei den hybriden Modellen Individual- und Aggregatdaten in einem gemeinsamen Modell kombiniert werden. Obwohl sich die zur Evaluierung der Modelle genutzten Szenarien an Wahlen orientiert haben, sind die spezifischen Probleme bei der Schätzung von Wählerwanderungen bisher nur am Rande angesprochen worden.

Das folgende Kapitel widmet sich daher vor allem der praktischen Anwendung der Modelle. Die Darlegungen erfolgen dabei vor allem anhand des (hybriden) Multinomial-Dirichlet Modells. Das Multinomial-Dirichlet Modell zeigt beim Methodenvergleich in Kapitel II.4. die insgesamt beste Schätzgüte aller verglichenen Methoden der ökologischen Inferenz. Die hybride Erweiterung des Modells erreicht in Kapitel III.3. eine mindestens mit dem hybriden Multinomial-Normal Modell vergleichbare Schätzgüte, ohne dessen Probleme zu teilen.

IV.1. Datenaufbereitung bei der ökologischen Inferenz²¹²

Im Allgemeinen ist bei der Schätzung von Wählerwanderungen mit der ökologischen Inferenz und mit hybriden Modelle eine Aufbereitung der vorliegenden Aggregatdaten notwendig. Die dabei getroffenen Entscheidungen gehören, neben der Wahl des Modells, mit zu den wichtigsten beim Analyseprozess. Die folgenden Betrachtungen erfolgen exemplarisch anhand des Multinomial-Dirichlet Modells, die Ergebnisse sind jedoch auch auf das hybride Modell übertragbar.

IV.1.1. Kritische Aspekte bei der Datenaufbereitung

Eine der Grundannahmen der ökologischen Inferenz ist, dass sich die beiden beobachteten Ränder mit der Wahlentscheidung auf dasselbe Gebiet und dieselbe Bevölkerung beziehen. Diese Annahme ist im Allgemeinen bei der Schätzung von Wählerwanderungen nicht gültig. Es ist üblich, dass sich die Einteilung des Wahlgebiets in Wahlbezirke im Laufe der Zeit ändert und dass daher nicht alle Wahlbezirke bei zwei

212 Die Diskussion der kritischen Aspekte und die Ergebnisse der Simulationsstudie in diesem Kapitel wurden bereits in Klima / Thurner / Molnar / Schlesinger / Küchenhoff, 2016, sowie teilweise in Klima / Küchenhoff / Thurner, 2013 veröffentlicht. Während die hier vorgestellten Betrachtungen nur die Ergebnisse für das Multinomial-Dirichlet Modell umfassen, deckt die Veröffentlichung von 2016 mit Ausnahme des Multinomial-Normal Modells alle in Tabelle 6 erwähnten Modelle ab.

Wahlen das gleiche Gebiet beschreiben.

In den meisten Fällen ist es nicht möglich, ein neu gebildetes Gebiet für die Berechnungen wieder aufzuteilen. Für eine solche Aufteilung wäre es notwendig, die Wahlberechtigten und ihre Entscheidung bei der letzten Wahl eindeutig einem Teilgebiet eines Wahlbezirkes zuordnen zu können. Entsprechend detaillierte Informationen werden jedoch nicht erhoben. Die daher am häufigsten genutzte Methode ist das Bilden von über die Zeit konstanten Wahlbezirken durch Zusammenlegen von Wahlbezirken. Eine entsprechende Datenaufbereitung erfolgte auch bei den zuvor genutzten Daten für die Bundestagswahlen 2005 und 2009, vergleiche Kapitel II.4.1., in München. Da dieser Prozess je nach Wahl und betrachteten Gebiet sehr aufwändig sein kann, stellt sich die Frage, wie genau diese Zusammenlegung sein muss und wie wichtig es ist, möglichst viele Wahlbezirke bei der Zusammenlegung zu bewahren.

Ein weiterer Aspekt betrifft die Wahlberechtigten in den Wahlbezirken. Es wird bei den Modellen davon ausgegangen, dass die beiden Ränder eine Partitionierung der gleichen Bevölkerung darstellen. Dies geht bereits aus Tabelle 1 hervor, bei der sich beide Ränder zu eins aufsummieren. Diese Annahme ist jedoch im Allgemeinen bei der Schätzung von Wählerwanderungen verletzt. Zwischen den beiden betrachteten Wahlen liegt oft ein längerer Zeitraum, zwischen zwei Bundestagswahlen sind es im Allgemeinen z. B. vier Jahre. Innerhalb von vier Jahren können Wahlberechtigte jedoch versterben, oder Heranwachsende bzw. Eingebürgerte wahlberechtigt werden.

Ein zahlenmäßig und inhaltlich daher noch größeres Problem stellen aber Umzüge von Wahlberechtigten dar. Umzüge innerhalb eines Wahlbezirks sind für die Modelle dabei relativ problemfrei. Umzüge zwischen den Wahlbezirken sorgen jedoch dafür, dass die Wahlentscheidungen eines Wahlberechtigten bei den beiden Wahlen in zwei unterschiedlichen Wahlbezirken erfasst werden. Dies bedeutet bei den Modellen, dass die Wahlentscheidung bei der ersten Wahl im Zeilenrand der Wählerwanderungstabelle des Wahlbezirks vor dem Umzug erfasst wird, während die Wahlentscheidung bei der zweiten Wahl entsprechend im Spaltenrand der Wählerwanderungstabelle des Wahlbezirks nach dem Umzug beinhaltet ist. Für diese Wahlberechtigten ist es daher eigentlich nicht möglich, eine Wählerwanderung mit den Modellen zu schätzen, mit Abstrichen kann jedoch noch argumentiert werden, dass auf Ebene der Gesamtregion eine valide Wählerwanderungstabelle entsteht. Neben diesen Regions-internen Umzügen ist es auch möglich, die betrachtete Region komplett zu verlassen oder ihr zuzuziehen. Die Größenordnung dieser Wanderungsbewegungen hängt dabei von der betrachteten Region

	CSU 2013	...	NW 2013	Alt 2013	Weg 2013	Wahl 2009
CSU 2009	$\beta_{CSU2009,CSU2013}$...	$\beta_{CSU2009,NW2013}$			23,56 %
...
NW 2009	$\beta_{NW2009,CSU013}$...	$\beta_{NW2009,NW2013}$			26,60 %
Neu 2009		...		0	0	
Zu 2009		...		0	0	
Wahl 2013	26,91 %	...	28,80 %			

Tabelle 22: Darstellung der Datensituation bei der ökologischen Inferenz, erweiterte Darstellung im Vergleich zu Tabelle 1. Nichtwähler werden mit NW abgekürzt, neu hinzugekommen sind die Spalten für die nicht mehr wahlberechtigten Wähler [Tod oder Entzug des Wahlrechts, Alt], weggezogene Wähler (Weg) und die Zeilen für die erstmals Wahlberechtigten (Neu) und die zugezogenen Wahlberechtigten (Zu). Aus inhaltlichen Gründen muss ein Teil der inneren Zellen null sein und ist daher bekannt, während die Randinformation zu den ergänzten Spalten nicht Teil der offiziellen Wahlstatistiken ist. Die angegebenen Wahlergebnisse entsprechen dem Ergebnis der Bundestagswahl 2009, respektive 2013 in München, Nichtwähler werden mit NW abgekürzt, ein Teil der Tabelle wird nicht dargestellt. (Datenquelle: Stadt München)

ab. Offensichtlich sind Wahlberechtigte, die die Region verlassen oder ihr zuziehen nur bei einer Wahl in den amtlichen Wahlergebnissen in der Region erfasst, eine Wählerwanderung kann daher für diese Wahlberechtigten eigentlich nicht geschätzt werden. Tabelle 22 zeigt eine etwas realistischere Darstellung einer Wählerwanderungstabelle, bei der die gerade erwähnten Aspekte mit berücksichtigt werden.

Es ist im Allgemeinen bei der ökologischen Inferenz nicht möglich, eine Wählerwanderung für Tabelle 22 zu schätzen, es wird fast immer eine mit Tabelle 1 vergleichbare Situation betrachtet. Die dabei nicht berücksichtigten Bevölkerungsveränderungen haben jedoch eine inhaltliche Implikation. Es ist fraglich, wie die mit den Modellen geschätzte Wählerwanderung zu interpretieren ist. Es kann nicht mehr notwendigerweise davon ausgegangen werden, dass die Wählerwanderung von Individuen beschrieben wird. Dafür müsste angenommen werden, dass ein Wahlberechtigter im Allgemeinen durch einem Wahlberechtigten mit ähnlicher politischer Haltung und Wahlverhalten ersetzt wird. Eine alternative Betrachtungsweise wäre, sich von der Individualebene zu lösen und die geschätzte Wählerwanderung als eine Wählerwanderung des Gebietes an sich aufzufassen. Eine Wählerwanderung von einer Partei zu einer anderen Partei muss dann nicht notwendigerweise auf das veränderte Wahlverhalten eines Individuums zurückgeführt werden, sie kann auch Ergebnis einer geänderten Zusammensetzung der Wählerschaft in dem Gebiet sein.

Die Bevölkerungsveränderungen im Laufe der Zeit haben neben der inhaltlichen Implikation aber noch eine weitere, praktisch relevante Implikation. Die Bevölkerungszahlen in den Wahlbezirken müssen bei beiden Wahlen nicht identisch sein. Da aber für beide Wahlen eine identische Bevölkerung angenommen wird, muss dieser

möglicherweise vorhandene Unterschied bei der Datenaufbereitung aufgefangen werden. Eine Möglichkeit ist, für beide Wahlen eine identische Bevölkerung anzunehmen. Die Zahl der Wähler für die Parteien kann dann – unter Beibehaltung der relativen Häufigkeiten – angepasst werden.²¹³ Diese Art der Datenaufbereitung impliziert die Annahme, dass die veränderte Zahl der Wahlberechtigten zu keiner strukturell anderen Wählerstruktur im Wahlbezirk führt. Eine Alternative dazu ist, die größere Bevölkerungszahl für beide Wahlen zu nutzen. Der absolute Bevölkerungsunterschied zwischen beiden Wahlen wird jedoch bei dieser Alternative gesondert oder als Teil einer der betrachteten Möglichkeiten, z. B. der Nichtwähler, erfasst und bei der Schätzung mit berücksichtigt.

Weitere Schwierigkeiten bei der Datenaufbereitung können durch die Ausgestaltung des Wahlrechts verursacht werden. Falls zum Beispiel mehrere Stimmen pro Wahlberechtigten abgegeben werden können, stellt sich die Frage, welche Größe bei der Schätzung betrachtet werden soll. Die Wählerwanderung kann in solchen Fällen theoretisch auf Ebene der einzelnen Stimmen oder auf Ebene der Wahlberechtigten geschätzt werden. Aus inhaltlichen Gründen bietet es sich meist an, zur Schätzung der Wählerwanderung zwischen Parteien auch die, falls vorhanden, Parteienstimme bei der Wahl zu nutzen. Falls die Wählerwanderung der Wahlberechtigten geschätzt werden soll, mehrere Stimmen für Parteien pro Wahlberechtigten vergeben werden können und die Möglichkeit zum Panaschieren besteht, ist es notwendig, eine Definition eines loyalen bzw. eines Wechselwählers aufzustellen. Aufgrund der möglichen Aufteilung der Stimmen auf mehrere Parteien kann ein Wahlberechtigter bei der zweiten Wahl für die gleiche und für eine andere Partei stimmen. Meist ist es in diesen Fällen jedoch nicht möglich, ohne Abstriche die Wählerwanderung zu schätzen.

Ebenfalls relevant ist die im deutschen Wahlrecht verankerte Möglichkeit der Briefwahl. Das Wahlgebiet ist in Wahlbezirke eingeteilt und jeder Wahlberechtigte ist aufgrund seines Wohnortes eindeutig einem Wahlbezirk zugeordnet. Wahlberechtigte, die Briefwahl beantragen²¹⁴, gehören weiterhin zu ihrem Wahlbezirk, werden aber in den amtlichen Ergebnissen in gesonderten Briefwahlbezirken ausgewiesen. Eine naheliegende Möglichkeit wäre, die Briefwahlbezirke bei der Schätzung als Wahlbezirke mitzuberechnen. Dieses Vorgehen ist aus zwei Gründen problematisch. Briefwahl muss keine über die Zeit konstante Eigenschaft eines Wahlberechtigten sein, ein Wahlberechtigter kann bei einer Wahl Briefwähler sein und bei der zweiten Wahl in seinem

213 Es handelt sich um ein 1969 von A.G. Hawkes vorgeschlagenes Vorgehen, vgl. Kellermann, 2011, S. 35.

214 In Wahlstatistiken „Wahlberechtigte mit Wahlschein“ genannt.

Urnenwahlbezirk seine Stimme abgeben. Der zweite Grund ist aus methodischen Gründen schwerwiegender und basiert auf der gesonderten Erfassung der Wahlberechtigten, die Briefwahl beantragt haben. Die Wahlbeteiligung unter dieser Personengruppe ist im Allgemeinen deutlich höher als unter den Urnenwählern. Der Unterschied ist aber nur auf die gesonderte Erfassung zurückzuführen. Ein Briefwähler kann nur dann Nichtwähler werden, wenn er zuvor die Briefwahl aktiv beantragt hat und dann anschließend seine Stimme nicht abgibt, d. h. wenn er sich nach einer aktiven Entscheidung für eine Beteiligung an der Wahl doch noch dagegen entscheidet. Dieser strukturelle Unterschied stellt aber eine Verletzung der Modellannahmen dar, bei denen von einer Ähnlichkeit der Übergangswahrscheinlichkeiten zwischen den Wahlbezirken ausgegangen wird. Da die Übergangswahrscheinlichkeiten zu den Nichtwählern bei den Briefwählern strukturell niedriger sind, sind die anderen Übergangswahrscheinlichkeiten im Allgemeinen größer.

Aus diesen Gründen wird die Möglichkeit der gesonderten Aufnahme nicht als eine adäquate Berücksichtigung der Briefwähler betrachtet. Um die Briefwähler bei der Schätzung dennoch handhaben zu können, werden zwei grundlegende Prinzipien in Betracht gezogen. Die einfachere Lösung stellt das Ignorieren der Briefwähler dar. Bei diesen Berechnungen werden nur die Urnenwahlbezirke zur Schätzung der Wählerwanderung herangezogen und es wird davon ausgegangen, dass die so geschätzte Wählerwanderung auf alle Wahlberechtigten übertragbar ist. Als Alternative dazu wird die Verteilung der Briefwähler auf die zugehörigen Urnenwahlbezirke in Betracht gezogen. Die Genauigkeit der Verteilung hängt dabei von den vorliegenden Informationen und der Bildung der Briefwahlbezirke ab. Bei der Verteilung der Briefwähler wird zusätzlich davon ausgegangen, dass sich zwischen den Urnenwahlbezirken, die einen Briefwahlbezirk bilden, keine strukturellen Unterschiede zwischen den Wahlberechtigten finden lassen, die Briefwahl beantragt haben. Bei diesem Ansatz hat das Modell implizit den Anspruch, die Wählerwanderung für alle Wahlberechtigten zu erklären. Eine Verteilung der Briefwähler erfolgte auch bei den zuvor genutzten Daten für die Bundestagswahlen 2005 und 2009 in Kapitel II.4.1. und II.4.2..

IV.1.2. Simulationsstudie zur Auswirkung der Datenaufbereitung

Die Auswirkungen der notwendigen Entscheidungen bei der Datenaufbereitung auf die Schätzgüte werden mit einer Simulationsstudie evaluiert. Die Betrachtungen konzentrieren sich dabei auf vier Aspekte: Notwendigkeit einer möglichst geringen Aggregation, zusätzliche Partei bei der zweiten Wahl, Umgang mit Briefwählern, sowie

Bevölkerungsveränderungen in den Wahlbezirken. Ausgangspunkt für die Simulationsstudie sind die bereits in Kapitel II.4.3. betrachteten Szenarien.

Insgesamt werden vier Hauptszenarien, eines für jeden betrachteten Aspekt, in der Simulation evaluiert. Für jedes Hauptszenario gibt es dabei jeweils mehrere Unterszenarien. Im ersten Szenario (II.1) wird für die vier Szenarien Sim 1 – Sim 4, siehe Tabelle 14, die Auswirkung einer stärkeren Aggregation überprüft. Die ursprünglich gebildeten 314 über die Zeit konstanten Wahlbezirke werden durch weitere Aggregation auf 60 Gebietseinheiten reduziert. Der Grund für diese Betrachtungen ist, dass eine Bewahrung von möglichst vielen Wahlbezirken zeitaufwändig sein kann und eine gröbere Darstellung mit stärkerer Aggregation leichter und schneller realisiert werden kann. Beim zweiten Szenario (II.2) wird für Szenario Sim 2 und Sim 4 die Auswirkung einer neuen Partei auf die Schätzgüte überprüft. Dafür wird in der Simulation bei der zweiten Wahl eine neue Kleinpartei aufgenommen, so dass anstelle einer 7x7-Tabelle eine 7x8-Tabelle betrachtet wird. Mit diesem Szenario soll untersucht werden, ob das Auftauchen einer neuen Partei zu Problemen bei der Schätzung der Wählerwanderung führen kann.

Szenario II.3 widmet sich den Briefwählern. Es wird für dieses Szenario angenommen, dass ein Teil der simulierten Wähler Briefwahl beantragt hat und deren Wahlergebnisse in gesonderten Briefwahlbezirken erfasst werden. Ausgangspunkt für die Simulation der Briefwähler stellt dabei Sim 3 dar. Beim Anteil der Briefwähler wird sich an den in München beobachteten Anteilen orientiert. Weiterhin wird angenommen, dass ein Briefwähler mit ~50% Wahrscheinlichkeit bei beiden Wahlen Briefwähler ist und Briefwähler mit höherer Wahrscheinlichkeit aus der Bevölkerungsgruppe mit höherer Loyalität stammen. Die drei Unterszenarien repräsentieren drei Methoden, wie die Briefwähler bei der Schätzung mit dem Multinomial-Dirichlet Modell berücksichtigt werden können. Die erste betrachtete Alternative stellt das Ignorieren der Briefwähler dar und die Schätzung mit dem Modell erfolgt nur mit den Urnenwählern. Die beiden anderen überprüften Alternativen unterscheiden sich primär in der Genauigkeit der Zuordnung der Briefwähler zu ihren Urnenwahlbezirk. Im zweiten betrachteten Szenario ist nur eine relativ grobe Zuordnung möglich, während beim dritten Szenario eine genauere Zuordnung umsetzbar ist.

Im Szenario II.4 wird auf Bevölkerungsveränderungen eingegangen. Ausgangspunkt für die Simulation der Bevölkerungsveränderungen stellt erneut Sim 3 dar. Es werden dabei primär zwei Szenarien betrachtet. Im Ersten gibt es nur relativ geringe Bevölkerungsveränderungen und nur jeweils ungefähr 4% der Bevölkerung sind nur bei

einer Wahl wahlberechtigt. Um ein unterschiedliches Wahlverhalten zu simulieren, werden die Wahlberechtigten, die nur bei der ersten Wahl wahlberechtigt sind, aus der Bevölkerungsgruppe mit hoher Loyalität gezogen. Die Wahlberechtigten, die nur bei der zweiten Wahl wahlberechtigt sind, werden neu zu 75% aus der Bevölkerungsgruppe mit niedriger Loyalität und zu 25% aus der Bevölkerungsgruppe mit mittlerer Loyalität simuliert. Für die Schätzung der Modelle werden nur die bei einer Wahl in den amtlichen Wahlergebnissen vorhandenen Informationen genutzt. Das zweite Szenario erweitert das erste Szenario und berücksichtigt zusätzlich auch Veränderungen durch Umzüge. Dafür wird angenommen, dass im Mittel 25% der Wahlberechtigten eines Wahlbezirkes innerhalb der Stadt umziehen. Diese Wahlberechtigten werden auf alle Wahlbezirke verteilt. Zusätzlich wird angenommen, dass weitere 25% der Wahlberechtigten eines Wahlbezirks im Mittel die Region verlassen, während 35% von außerhalb zuziehen. Die zuziehenden Wahlberechtigten werden dabei aus Sim 2 gezogen und besitzen daher eine gewisse Ähnlichkeit zu den Wahlberechtigten in der Region. Der höhere Anteil der zuziehenden Wahlberechtigten sorgt dafür, dass es im Mittel mehr Wahlberechtigte bei der zweiten Wahl gibt. Die veränderte Bevölkerung wird dabei entweder durch eine proportionale Anpassung der Wahlergebnisse an die größere Bevölkerung, zweites Unterszenario, oder durch Aufnahme bei den Nichtwähler, drittes Unterszenario, berücksichtigt.

Eine Kurzzusammenfassung der vier Hauptszenarien befindet sich in Tabelle 23. Für jedes Szenario werden bei dieser Simulationsstudie 30 Simulationsdatensätze betrachtet. Dabei handelt es sich entweder um die ersten 30 Datensätze der Szenarien aus Kapitel II.4.3. oder um Datensätze, die auf diesen aufbauen. Für das Multinomial-Dirichlet Modell werden die gleichen Einstellungen²¹⁵ wie in Kapitel II.4., siehe auch Kapitel II.4.2.3., verwendet. Wie bei den bisherigen Simulationsstudien wird als Maß für die Schätzgüte erneut der AD, siehe Gleichung (28), genutzt.

Szenario	Name	#Unterszenarien	Verteilung zwischen den Gebieten
II.1	Aggregation60	4	Stärke Aggregation auf 60 Wahlbezirke (von 314)
II.2	Neue Partei	2	Zusätzliche Partei bei der zweiten Wahl, 7x8 Tabelle
II.3	Briefwähler	3	Berücksichtigung von Briefwählern
II.4	Bevölkerungsveränderungen	3	Berücksichtigung von Bevölkerungsveränderungen

Tabelle 23: Simulation zur Datenaufbereitung bei der ökologischen Inferenz, Übersicht über die vier in der Simulation betrachteten Datenszenarien, dargestellt sind die Schlüsseigenschaften der Szenarien. Details zu den einzelnen Szenarien finden sich in Kapitel IV.1.2.

²¹⁵ Burn-In: 50000; Thinning: 1000; Ziehungen: 1000. Nicht-informative Wahl für Priori-Parameter der Hyperprioris wie in Kapitel II.3.4.4. beschrieben.

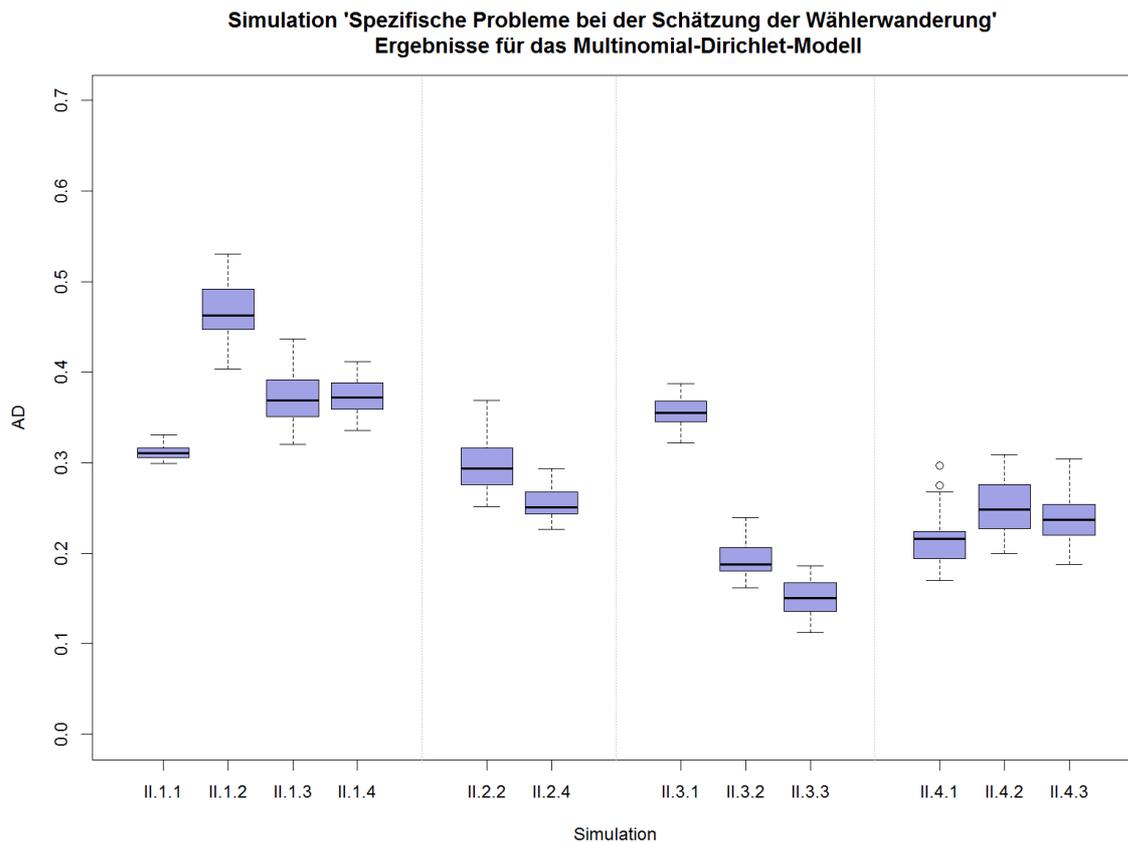


Abbildung 18: Simulation zur Datenaufbereitung bei der ökologischen Inferenz, Ergebnisse der Simulationsstudie für das Multinomial-Dirichlet Modell, dargestellt ist der AD der Schätzung mit den Modellen zu den simulierten Individualdaten. Für jedes Szenario wurden 30 Datensätze generiert und evaluiert. Ein Überblick der Szenarien befindet sich in Tabelle 23, die letzte Ziffer bei Szenario II.1 und II.2 bezieht sich dabei auf das betrachtete Szenario in Tabelle 14, bei Szenario II.3 und II.4 auf das Unterszenario, vgl. Kapitel IV.1.2..

Die Ergebnisse für das Multinomial-Dirichlet Modell werden in Abbildung 18 präsentiert. Es zeigt sich deutlich bei einem Vergleich mit Abbildung 8, dass die stärkere Aggregation zu einer deutlichen Verschlechterung der Schätzgüte beim Multinomial-Dirichlet Modell führt. Dies betrifft alle vier Szenarien, wobei die Verschlechterung unterschiedlich stark ausfällt. Eine Betrachtung der geschätzten Wählerwanderung zeigt weiterhin, dass die Loyalitätsraten tendenziell zu niedrig, während die anderen Übergangswahrscheinlichkeiten tendenziell zu hoch geschätzt werden. Das Ergebnis dieser Simulation deckt sich auch mit den Ergebnissen in Kapitel III.4.1., wo bereits eine Abhängigkeit von der Zahl der Wahlbezirke festgestellt werden konnte. Es liegt daher nahe anzunehmen, dass die Zahl der zur Verfügung stehenden Wahlbezirke ein kritisches Merkmal bei der Schätzung mit dem Multinomial-Dirichlet Modell ist.

Die Hinzunahme einer neuen Partei hat keinen merklichen Einfluss auf die Schätzgüte und die Werte sind mit dem jeweiligen Ausgangsszenario vergleichbar. Jedoch kann nicht

ausgeschlossen werden, dass eine deutliche Erhöhung der Komplexität zu einer schlechteren Schätzung führt.

Das Ignorieren der Briefwähler führt im betrachteten Szenario zu einer deutlichen Verschlechterung der Schätzgüte und stellt die schlechteste der drei betrachteten Vorgehensweisen dar. Der AD ist ungefähr um 0,2 höher im Vergleich zum Ausgangsszenario. Ein Grund für den deutlichen Unterschied dürfte die Annahme in der Simulation sein, dass Briefwähler tendenziell eine höhere Loyalität besitzen. Bereits mit einer relativ groben Zuordnung der Briefwähler kann eine deutlich bessere Schätzgüte erreicht werden. Die erreichte Schätzgüte mit Briefwählern ist in dieser Simulation etwas schlechter im Vergleich zum Ausgangsszenario. Eine weitere deutliche Verbesserung der Schätzgüte kann für das Szenario mit genauerer Zuordnung und Verteilung der Briefwähler festgestellt werden. Die Schätzgüte in diesem Fall erreicht das Niveau des Ausgangsszenarios.

Bei der Betrachtung der Bevölkerungsveränderungen ist bei dem Szenario mit nur geringen Bevölkerungsveränderungen, Szenario II.4.1, keine Verschlechterung der Schätzgüte feststellbar. Die Güte ist in etwa vergleichbar mit dem Ausgangsszenario. Die Evaluierung der anderen beiden Szenarien ist durch die Art der vorliegenden Individualdaten erschwert. Für die Schätzung werden nur die auch in amtlichen Wahlergebnissen vorliegenden Informationen, vergleiche Tabelle 1, genutzt, die Wählerwanderungstabelle auf Individualebene ist aber mit Tabelle 22 vergleichbar. Ein direkter Vergleich der geschätzten Wählerwanderung mit den simulierten Individualdaten ist daher nicht möglich. Für den Vergleich werden aus diesem Grund nur die inneren Zellen, d. h. die Übergänge zwischen den Parteien, betrachtet und alle anderen Zeilen und Spalten ignoriert. Die so bestimmte Schätzgüte der Modelle ist geringfügig schlechter als die des Ausgangsszenarios. Der Vergleich der beiden in Betracht gezogenen Möglichkeiten Bevölkerungsdifferenzen zwischen den Wahlen zu berücksichtigen ergibt nur relativ geringe Unterschiede, wobei bei dieser Simulation die Aufnahme bei den Nichtwählern (II.4.3) zu einer etwas besseren Schätzgüte geführt hat.

Die Betrachtungen dokumentieren, dass die Datenaufbereitung relevanten Einfluss auf die Schätzgüte haben kann, der Einfluss jedoch je nach untersuchten Aspekt unterschiedlich stark ausfällt. Für das Multinomial-Dirichlet Modell werden vor allem die Zahl der zur Verfügung stehenden Wahlbezirke und die Handhabung der Briefwähler als relevant bewertet. Es sollten möglichst viele Wahlbezirke bewahrt werden und die Briefwähler sollten so genau wie möglich auf ihre Urnenwahlbezirke verteilt werden. Bei

einer entsprechenden Datenaufbereitung ist die Schätzung vergleichbar oder nur marginal schlechter als beim Ausgangsszenario, bei dem keine entsprechende Datenaufbereitung notwendig ist.

Weiterhin ist es möglich zu zeigen, dass auch bei relativ großen Bevölkerungsveränderungen zwischen den Wahlen eine valide Schätzung der Wählerwanderung mit dem Multinomial-Dirichlet Modell möglich ist. Bei der Interpretation der Ergebnisse muss jedoch berücksichtigt werden, dass für diese meist zusätzliche Annahmen notwendig sind und je nach Annahme unterschiedliche inhaltliche Aussagen getroffen werden können.

IV.2. Münchner Nachwahlbefragung

Während beim Multinomial-Dirichlet Modell nur eine Datenaufbereitung der Aggregatdaten notwendig ist, werden für die Schätzung mit dem hybriden Multinomial-Dirichlet Modell zusätzlich auch Individualdaten benötigt. Für die Nutzung der Individualdaten im Modell ist es weiterhin notwendig, diese eindeutig einem Wahlbezirk²¹⁶ zuordnen zu können.²¹⁷ Die Gelegenheit entsprechende Individualdaten zu erheben, und dabei auch auftretende Probleme kennen zu lernen, eröffnete sich im Jahr 2013.

Die Erhebung der Daten erfolgte im Rahmen einer Forschungs- und Lehrkooperation des Lehrstuhls für empirische Politikforschung, Prof. Paul W. Thurner, und des Statistischen Beratungslabors der Ludwig-Maximilians-Universität München (LMU), Prof. Helmut Küchenhoff. Das Jahr 2013 bot sich insofern an, da die Landtagswahlen in Bayern am 15.09.2013 und die Wahl zum Deutschen Bundestag eine Woche später am 22.09.2013 stattfanden. Diese seltene Konstellation ermöglichte es, bei zwei Wahlen in enger zeitlicher Nähe eine Erhebung von Individualdaten durchführen. Vor allem für die Schätzung der Wählerwanderung zwischen Landtags- und Bundestagswahl 2013 bedeutet dies auch, dass viele der üblichen mit den Daten verbundenen Schwierigkeiten bzw. Unsicherheiten bei der Schätzung dieser Wählerwanderung nur im geringen Maße zu erwarten sind.

Aufgrund der räumlichen Nähe und einer vorhergehenden Zusammenarbeit mit dem Statistischen Amt der Landeshauptstadt München bot es sich an, die Erhebung der Individualdaten in München durchzuführen. Bei der Landtagswahl 2013 war München auch aus inhaltlichen Gründen besonders attraktiv, da der langjährige Münchner Oberbürgermeister Christian Ude für die SPD als Spitzenkandidat bei der Landtagswahl antrat. Aufgrund von praktischen Überlegungen, mit den erhobenen Individualdaten soll z. B. die Wählerwanderung geschätzt werden, erfolgte die Erhebung im Rahmen einer Nachwahlbefragung.

In den folgenden beiden Kapiteln werden kurz die Überlegungen hinter der Erhebung der Individualdaten und ein paar Ergebnisse und gewonnenen Erkenntnisse aus der Münchner Nachwahlbefragung vorgestellt. Dabei wird vor allem auf die für die Schätzung von Wählerwanderungen mit hybriden Modellen relevante Punkte eingegangen. Die Erhebung der Individualdaten wurde jedoch nicht ausschließlich für die Nutzung in

216 Es wird vereinfachend nur der Begriff Wahlbezirk genutzt, obwohl dieser Benennung nur bei den betrachteten Bundestagswahlen korrekt ist und bei den Landtagswahlen richtigerweise von Stimmbezirken gesprochen werden müsste.

217 Vgl. Kapitel III.1.2. und III.2.2..

hybriden Modellen konzipiert. Auf andere Überlegungen und Designentscheidungen bei der Konzipierung der Erhebung, die nicht mit hybriden Modellen direkt in Verbindung stehen, wird in dieser Arbeit nur sehr rudimentär eingegangen.

IV.2.1. Planung der Nachwahlbefragung

Die Datenerhebung erfolgte im Rahmen einer Nachwahlbefragung aus vor allem zwei praktischen Überlegungen: Es war die Intention mit der Befragung die Wahlentscheidung bei der aktuellen Wahl zu erheben. Dies erfolgt am besten direkt nach dem eigentlichen Wahlvorgang. Bei einer Telefonumfrage kann, wenn sie vor der Wahl stattfindet, im Allgemeinen nur die Wahlabsicht abgefragt werden. Bei einer Telefonumfrage nach der Wahl erfolgt die Erhebung im Allgemeinen einige Zeit nach der Wahlentscheidung, so dass Störfaktoren²¹⁸ die Angabe bereits wieder beeinflussen können.

Zusätzlich ist für die Nutzung der Individualdaten in hybriden Modellen die Kenntnis des Gebietes notwendig, in welchen diese erhoben worden sind. Bei einer Nachwahlbefragung fällt diese Informationen weitestgehend automatisch an, da die Befragung vor dem Wahllokal²¹⁹ bzw. -gebäude²²⁰ stattfindet, das Gebiet des befragten Wahlberechtigten also bei der Erhebung bekannt ist. Bei einer Telefonbefragung müsste diese Gebietsinformation mit erhoben werden. Während vor der Wahl für den Befragten die Möglichkeit besteht, die Information auf der Wahlbenachrichtigung anzugeben, müsste er sich bei einer Befragung nach der Wahl daran erinnern oder eine alternative Information, wie z. B. Straße und Hausnummern, angeben, aus der die Zuordnung dann hergeleitet werden könnte.

Während die Nachwahlbefragung bei der Erhebung der Daten und bei der zu erwartenden Datengüte Vorteile hat, verfügt sie aber auch über einen relevanten Nachteil. Es ist bei einer Nachwahlbefragung nicht möglich, Informationen von allen Wahlberechtigten zu erheben. Weder die Briefwähler, noch die Nichtwähler bei der zweiten Wahl können vor den Wahllokalen befragt werden. Während das Fehlen der Nichtwähler der zweiten Wahl für die Schätzung des Wahlergebnisses weniger bedeutsam ist, stellt ihr Fehlen für die Schätzung der Wählerwanderung ein grundlegendes Problem dar. Die Übergangswahrscheinlichkeiten zwischen Individual- und Aggregatdaten unterscheiden sich strukturell und die entsprechende Modellannahme der hybriden Modelle ist daher verletzt.

218 Die Kenntnis des Wahlergebnisses kann z. B. als solch ein Störfaktor betrachtet werden.

219 Unter Wahllokal ist der Abstimmungsort für einen Wahlbezirk zu verstehen.

220 Ein Wahlgebäude bezeichnet das Gebäude in dem ein oder mehrere Wahllokale untergebracht sind. Es kann sich dabei z. B. um eine Schule handeln.

Nach der Entscheidung für die Erhebungsmethode erfolgte die Auswahl der Orte, an denen die Nachwahlbefragung durchgeführt werden sollte. Im ersten Schritt war es notwendig, die Erhebungsebene festzulegen. Da bei der Planung der Studie davon ausgegangen wurde, dass die Nachwahlbefragung nur vor den Wahlgebäuden, aber nicht in den Wahlgebäuden vor den jeweiligen Wahllokalen stattfinden kann, wurden anfangs die Wahlgebäude als primäre Erhebungsebene für die Nachwahlbefragung definiert.

Im ersten Schritt erfolgte die Bildung von über die Zeit konstanten Gebieten. Dieser Schritt ist notwendig, da bei der Schätzung der Wählerwanderung mit der ökologischen Inferenz oder mit hybriden Modellen Gebiete betrachtet werden müssen, die bei beiden Wahlen in etwa identisch sind. Da es aber zwischen der Landtagswahl 2008 und der Bundestagswahl 2009 zu einer kleinen Änderung, sowie zwischen Bundestagswahl 2009 und den Wahlen 2013²²¹ zu etwas größeren Änderungen beim Zuschnitt der Wahlbezirke kam, kann die Wahlbezirkseinteilung dieser Wahlen nicht direkt verwendet werden.

Bei den Wahlen im Jahr 2013 war die Stadt München in 702 Urnenwahlbezirke unterteilt²²², die sich wiederum auf insgesamt 198 Wahlgebäude verteilten. Auf Basis von Information, die das Statistische Amt der Landeshauptstadt München zur Verfügung stellte, wurden Wahlbezirke, deren Zuschnitt sich zwischen zwei Wahlen verändert hatte, so zusammengefasst, dass möglichst konstante Gebiete über beide Wahlen betrachtet werden können. So wurden insgesamt 649 bzw. 650 über die Zeit konstante Wahlbezirke gebildet. Zum Teil war auch eine Zusammenlegung auf der Ebene der Wahlgebäude notwendig, da zwischen den beiden Wahlen aufgeteilte Wahlbezirke bei den 2013er Wahlen unterschiedlichen Wahlgebäuden zugeordnet gewesen sind. Dies reduziert die Zahl der betrachteten Wahlgebäude von 198 auf 188.

Eine Vorgabe der Wahlleitung war, dass an jedem Wahlgebäude nur eine Nachwahlbefragung stattfinden soll. Da bei den Wahlen auch Nachwahlbefragungen von z. B. Infratest dimap oder der Forschungsgruppe Wahlen durchgeführt wurden und diese bereits ihre Wahlgebäude ausgewählt hatten, reduzierte sich die Zahl der für die Nachwahlbefragung zur Verfügung stehenden Wahlgebäude auf 131. Die Ziehung der Wahlgebäude für die Nachwahlbefragung erfolgte aus diesen 131 über die Zeit konstanten und noch freien Wahlgebäuden mittels Zufallsauswahl²²³. Es erfolgte nur eine Ziehung für

221 Bei der Landtags- und Bundestagswahl 2013 war der Zuschnitt der Wahlbezirke identisch.

222 Für die Briefwähler aus diesen 702 Wahlbezirken wurden 237 Briefwahlbezirke gebildet. Die Briefwähler konnten jedoch im Rahmen einer Nachwahlbefragung nicht mit erhoben werden.

223 Da auch eine Schätzung der Wählerwanderung mit den Individualdaten geplant war, erfolgte die Ziehung nicht mittels einer einfachen Zufallsauswahl und ein komplexeres Stichprobendesign wurde gewählt. So sollten mögliche Unterschiede zwischen der Stadt bei der Ziehung mit berücksichtigt

die Landtags- und Bundestagswahl, so dass bei beiden Wahlen an den gleichen Wahlgebäuden die Nachwahlbefragung stattfand.

Aufgrund der vorhandenen personellen Kapazitäten und der Ressourcenausstattung konnten insgesamt 14 Wahlgebäude ausgewählt werden. Erst nach Auswahl der Wahlgebäude und wenige Tage vor der Landtagswahl wurde uns von der Wahlleitung mitgeteilt, dass eine Erhebung auch innerhalb der Wahlgebäude direkt vor den Wahllokalen der einzelnen Wahlbezirke stattfinden kann. Deshalb wurden aus den bereits gewählten Wahlgebäuden erneut mittels Zufallsauswahl einzelne Wahlbezirke ausgewählt. Die Auswahl der gezogenen Wahlbezirke erfolgte dabei in Abhängigkeit von der Anzahl der Wahlbezirke im Wahlgebäude. Es wurden zwei Wahlbezirke gezogen, wenn mehr als fünf Wahlbezirke im Wahlgebäude untergebracht waren, ansonsten wurde nur ein Wahlbezirk gezogen. Insgesamt wurden so 17 über die Zeit konstante Wahlbezirke für die Nachwahlbefragung ausgewählt. Diese letzte Änderung verlagerte die Erhebungsebene der Individualdaten von den Wahlgebäuden auf die Wahlbezirke.

Vor den ausgewählten Wahlbezirken wurde schließlich eine Vollerhebung durchgeführt, d. h. alle Wähler wurden – nachdem sie den Abstimmungsraum verlassen hatten – angesprochen und gebeten an der Umfrage teilzunehmen. Durchgeführt wurde die Erhebung von Studierenden der LMU, die durch ein T-Shirt mit dem LMU-Logo und ein Namensschild uniform gekleidet waren. Diese Maßnahme sollte helfen, mögliche Interviewer-Effekte zu reduzieren. Insgesamt erfolgte der Hinweis auf die LMU mehrmals, um den guten Ruf der Universität zu nutzen und auch um den Befragungshintergrund offen erkennbar zu machen. Zur Vorbereitung der Befragung wurden aller Interviewer in einer mehrstündigen Schulung, in der unter anderem das professionelle Ansprechen der Wähler geübt wurde, auf die Befragung vorbereitet.

Die Erhebung der Individualdaten erfolgte dabei über einen Fragebogen und nicht in Form eines Interviews. Hintergrund für diese Entscheidung war die Hoffnung, dass so die Zahl der Antwortverweigerungen reduziert werden kann. Aus diesem Grund wurde bei der Nachwahlbefragung auch der eigentliche Wahlvorgang möglichst genau kopiert. So sollten z. B. die ausgefüllten Fragebögen von den Befragten selber in eine versiegelte Urne eingeworfen werden, um die (gefühlte) Anonymität bei der Befragung zu erhöhen. Die verwendeten Fragebögen zur Erhebung des Wahlverhaltens wurden im Rahmen eines interdisziplinären Seminars zur Wahlforschung von Studierenden entwickelt. Als

werden. Für die Nutzung der Daten im hybriden Modell ist das genaue Stichprobendesign jedoch nicht relevant und es wird daher nicht detailliert vorgestellt.

Anhaltspunkte dienten etablierte Frageformulierungen und Fragebogensequenzen unter anderem von Infratest dimap und aus der German Longitudinal Election Study.

Für beide Wahlen wurde jeweils ein eigener Fragebogen konzipiert, wobei der Fragebogen für die Bundestagswahl umfangreicher ausfiel, da auch das Wahlverhalten bei den letzten beiden Landtagswahlen (2008 und 2013) mit erhoben wurde. Bei beiden Fragebögen wurde die aktuelle Wahlentscheidung direkt abgefragt, wobei der Schwerpunkt auf der Parteienstimme (Zweitstimme) lag. Die für die Schätzung der Wählerwanderung notwendige Erhebung der Wahlentscheidung bei der Vorwahl erfolgte über eine Erinnerungsfrage. Zusätzlich zur Wahlentscheidung wurde das Alter und Geschlecht des Befragten erhoben und mit einer offenen Frage ist der Grund für einen möglichen Wechsel abgefragt worden. Um eine Zuordnung der Fragebögen zum jeweiligen Wahlbezirk zu ermöglichen, wurde auf jeden Fragebogen am unteren Rand der entsprechende Wahlbezirk vermerkt. Die Kopfzeile beider Fragebögen beinhaltete das Logo und Siegel der LMU, sowie einen Hinweis auf die durchführenden Institute und deren Lehrstuhlinhaber. Die beiden Fragebögen sind im Anhang X.4. dargestellt.

IV.2.2. Ergebnisse der Nachwahlbefragung

Insgesamt konnten 13644 Befragungen realisiert werden, davon entfallen 6522 auf die Landtagswahl und 7122 auf die Bundestagswahl. Die Response-Rate ist bei dieser Erhebung exakt bestimmbar, da die Zahl der Wähler in einem Wahlbezirk aus den amtlichen Wahlergebnissen²²⁴ hervorgeht und es sich um eine Vollerhebung handelt. Im Mittel nahmen ungefähr 70% der Wähler an der Befragung teil. Eine Kurzzusammenfassung ist in Tabelle 24 dargestellt.

Beim Vergleich fällt auf, dass bei der Landtagswahl die Response-Raten noch etwas höher als bei der Bundestagswahl waren. Dies könnte auf einen leichten Ermüdungseffekt

Befragung	Minimum	1. Quantil	Median	Mittelwert	3. Quantil	Maximum
Landtagswahl	0,38	0,70	0,73	0,71	0,75	0,82
Bundestagswahl	0,54	0,65	0,68	0,68	0,74	0,83

Tabelle 24: Münchner Nachwahlbefragung, Kurzzusammenfassung der Response-Raten in den 17 Wahlbezirken mit Nachwahlbefragung bei der Landtags- bzw. Bundestagswahl 2013. Die Response-Rate ist dabei die Zahl der ausgefüllten Fragebögen geteilt durch die Zahl der Wähler in der amtlichen Wahlstatistiken. (Datenquellen: Münchner Nachwahlbefragung, Stadt München)

224 Statistisches Amt der Landeshauptstadt München. Amtliches Endergebnis der Landtagswahl am 15.09.2013, Zweitstimmenergebnisse der Landeshauptstadt München, 2013. <https://www.muenchen.de/rathaus/Stadtfinfos/Statistik/Wahlen/Landtagswahl.html> und Statistisches Amt der Landeshauptstadt München. Amtliches Endergebnis der Bundestagswahl am 22.09.2013, Erst- und Zweitstimmenergebnisse der Landeshauptstadt München, 2013 (beide zuletzt aufgerufen am: 25.07.2016).

hindeuten, da die Befragung bei der Bundestagswahl nur eine Woche später in denselben Wahlbezirken stattgefunden hat. Das niedrige Minimum bei der Landtagswahl stellt dabei einen Ausreißer dar und ist vermutlich auf Interviewer-Effekte zurückzuführen. Bei der Bundestagswahl konnte mit anderen Interviewern in diesem Wahlbezirk eine deutlich höhere Response-Rate erreicht werden.

Ein weiterer Vorteil der Vollerhebung und eine Besonderheit bei einer Nachwahlbefragung ist, dass nicht nur die allgemeinen Response-Raten bestimmt werden können. Da alle Wähler nach ihrer aktuellen Wahlentscheidung befragt worden sind, kann auch eine Partei-spezifische Response-Rate berechnet werden. Es handelt sich dabei um den Anteil der Wähler einer Partei, die an der Befragung teilgenommen haben. Die Zahl der Wähler einer Partei in einem Wahlbezirk kann den amtlichen Wahlstatistiken entnommen werden. Die Partei-spezifische Response-Rate für die Befragung bei der Landtagswahl ist in Abbildung 19 dargestellt.

Es fällt auf, dass zwischen den Parteien relativ große Unterschiede bestehen. Die

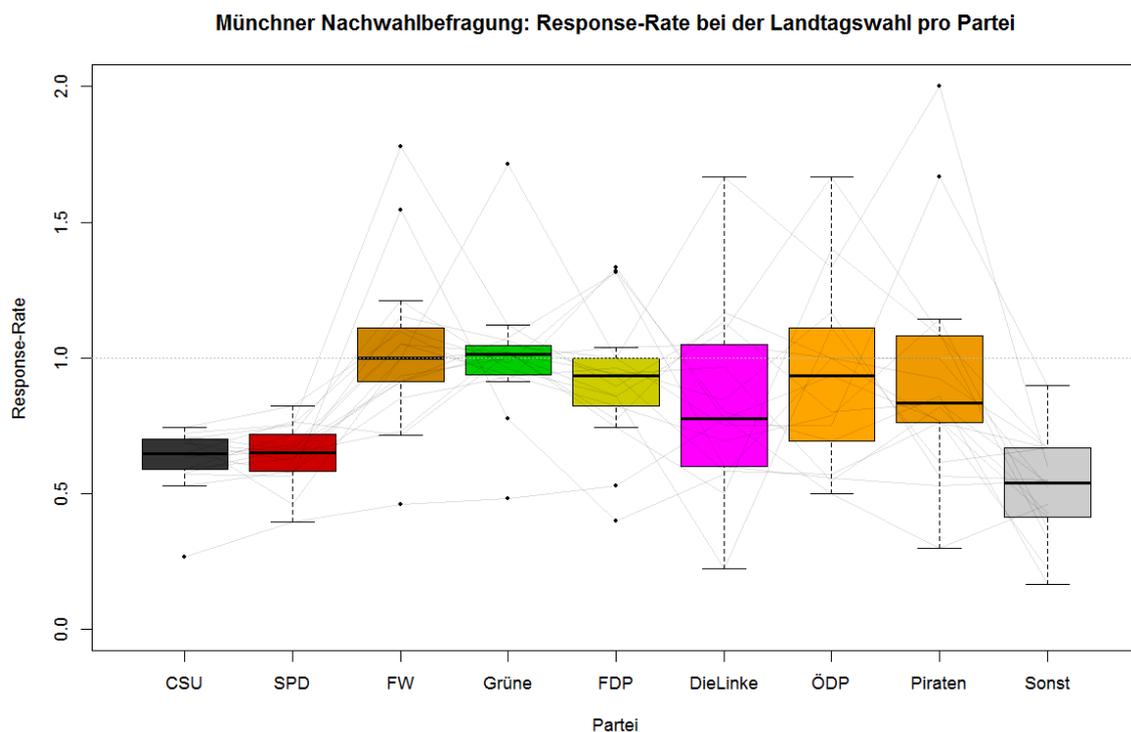


Abbildung 19: Münchener Nachwahlbefragung, Partei-spezifische Response-Raten in den 17 Wahlbezirken mit Nachwahlbefragung bei der Landtagswahl. Die jeweilige Response-Rate ist der Quotient von Wählern einer Partei in der Nachwahlbefragung und dem Wahlergebnis der Partei in den Wahlbezirk in den amtlichen Wahlstatistiken. Die einzelnen Linien stellen das Profil der Wahlbezirke dar. Parteienamen wie in Kapitel II.4.1., zusätzlich: Freie Wähler (FW), Ökologisch-Demokratische Partei (ÖDP) und Piratenpartei Deutschland (Piraten). (Datenquellen: Münchener Nachwahlbefragung, Stadt München)

niedrigste Beteiligung an der Nachwahlbefragung kann für die Wähler der CSU und der SPD, sowie den nicht gesondert betrachteten Kleinparteien in Sonstigen (Sonst) festgestellt werden. Deutlich noch problematischer sind jedoch die inhaltlich nicht zu erklärenden Response-Raten von über 100%. Entsprechende Response-Raten können für die Grünen und die FDP, sowie für fast alle kleineren Parteien in einigen Wahlbezirken identifiziert werden. Es ist auch nicht möglich, diese Response-Raten über ungültige Stimmen zu erklären, da die Zahl der ungültigen Stimmen in den amtlichen Wahlergebnissen praktisch nicht ausreicht um alle Mehrstimmen für Parteien abzudecken.

Bei der Befragung zur Bundestagswahl 2013, siehe Abbildung 20, taucht dieses Problem an sich nicht auf. Nur in einem einzigen Wahlbezirk gab es einmal für eine Partei in der Nachwahlbefragung mehr Wähler als in den amtlichen Wahlergebnissen. In diesen Fall würden theoretisch jedoch ungültige Stimmen dieses Ergebnis erklären können. Insgesamt sind die Response-Raten für die Parteien bei der Nachwahlbefragung bei der Bundestagswahl sich deutlich ähnlicher als bei der Landtagswahl. Auffällig sind jedoch

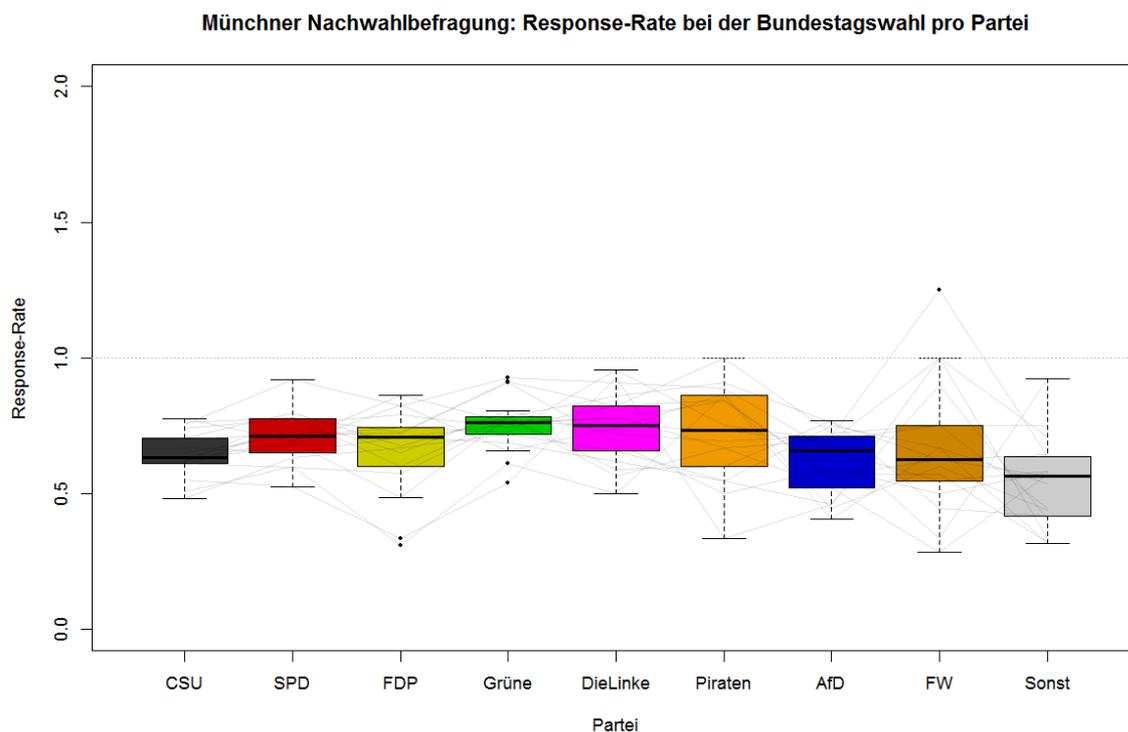


Abbildung 20: Münchner Nachwahlbefragung, Partei-spezifische Response-Raten in den 17 Wahlbezirken mit Nachwahlbefragung bei der Bundestagswahl. Die jeweilige Response-Rate ist der Quotient von Wählern einer Partei in der Nachwahlbefragung und dem Wahlergebnis der Partei in den Wahlbezirk in den amtlichen Wahlstatistiken. Die einzelnen Linien stellen das Profil der Wahlbezirke dar. Parteienamen wie in Kapitel II.4.1., zusätzlich: Piratenpartei Deutschland (Piraten), Alternative für Deutschland (AfD) und Freie Wähler (FW). (Datenquellen: Münchner Nachwahlbefragung, Stadt München)

erneut die relativ niedrigen Werte für die CSU und den nicht gesondert betrachteten Kleinparteien in Sonstigen (Sonst), sowie bei der Alternative für Deutschland (AfD). Im Mittel vergleichsweise höhere Response-Raten sind bei den Grünen, den Linken und der Piratenpartei Deutschland (Piraten) feststellbar.

Der Unterschied zwischen Landtags- und Bundestagswahl lässt vermuten, dass eine Besonderheit bei der Landtagswahl mit zu diesen Ergebnissen beigetragen hat. Möglicherweise kann das Problem mit den zu hohen Response-Raten auf das Wahlsystem bei der Landtagswahl zurückgeführt werden, bei dem keine strikte Trennung zwischen Partei- und Kandidatenstimme möglich ist. Beide Stimmen werden addiert und zählen zum Parteiergebnis, bei der Nachwahlbefragung wurde jedoch explizit nur die Zweitstimme erfragt. Ein zusätzlicher Störfaktor kann auch die parallel zur Landtagswahl durchgeführte Bezirkswahl gewesen sein.

Die unterschiedlichen Response-Raten pro Partei deuten weiterhin darauf hin, dass es möglicherweise allgemeine Effekte gibt, die die Wahrscheinlichkeit an der Erhebung teilzunehmen beeinflussen können. Falls sich eine positive bzw. negative Einstellung gegenüber der Erhebungen des Wahlverhaltens mit der politischen Überzeugung überlagert, werden tendenziell Wähler von bestimmten Parteien über- bzw. unterrepräsentiert sein. Es ist weiterhin möglich, dass der universitäre Hintergrund der Erhebung und die studentischen Interviewer diese Unterschiede mit verursacht oder entsprechende Tendenzen verstärkt haben.

Bei den bisherigen Betrachtungen standen nur die aktuellen Wahlentscheidungen im Fokus. Bei diesen sollte es die geringsten Unterschiede zwischen den erhobenen Individualdaten und den amtlichen Wahlergebnissen geben. Für die Schätzung der Wählerwanderung ist jedoch auch das Wahlergebnis bei der Vorwahl und die Angaben zum Wahlverhalten bei dieser Wahl in der Nachwahlbefragung relevant. Aufgrund der in Kapitel IV.1.1. diskutierten Punkte und den nicht erfassten Nichtwählern bei der zweiten Wahl, ist ein direkter Vergleich jedoch schwierig. Festgehalten werden kann jedoch, dass die Abweichungen und Schwankungen zwischen amtlichen Wahlergebnis und Erinnerungsfrage größer als die Unterschiede bei der aktuellen Wahl sind.

Die Ergebnisse der Münchner Nachwahlbefragung belegen weiterhin, dass es zu fehlerhaften Angaben des Wahlverhaltens bei der Vorwahl kam. Bei der Erinnerungsfrage wurde das Wahlverhalten bei der letzten Wahl zweistufig abgefragt. Zuerst wurde gefragt, ob eine andere Partei gewählt wurde und nur falls eine andere Partei gewählt wurde, sollte

diese auch explizit angegeben werden²²⁵. Ein Teil der Befragten nutzte die Möglichkeit anzugeben, dass er bei der letzten Wahl die gleiche Partei gewählt hatte, obwohl er bei der aktuellen Wahl für eine Partei gestimmt hat, die bei der letzten Wahl nicht angetreten ist. Bei der Bundestagswahl 2009 traten zum Beispiel die Freien Wähler (FW) und die AfD noch nicht an. Jedoch gaben insgesamt 24 Befragte, die bei der Bundestagswahl 2013 für die FW bzw. die AfD gestimmt haben wollen, an, auch bei der letzten Wahl bereits für diese Parteien gestimmt zu haben. Diese fehlerhaften Angaben können entweder bewusst getätigt worden sein, z. B. da der Befragte die Nachwahlbefragung nicht ernst genommen hat, oder auf ein fehlerhaftes Verständnis des Fragebogens bzw. auf eine fehlerhafte Erinnerung zurückführt werden. Direkt angegeben, sich nicht mehr an die Wahlentscheidung bei der Vorwahl erinnern zu können, haben bei der Befragung bei der Landtagswahl 290 (4,45%) und bei der Bundestagswahl 304 (4,27%) Personen.

Zusammenfassend lässt sich sagen, dass die Münchner Nachwahlbefragung aufzeigt, dass auch Individualdaten, die zur Schätzung von Wählerwanderungen erhoben werden, problembehaftet sein können. Während die Response-Raten bei der Nachwahlbefragung mit jeweils ungefähr 70% relativ hoch sind, zeigt sich, dass die Partei-spezifischen Response-Raten deutliche Unterschiede aufweisen können. Aufgrund der durchgeführten Vollerhebung ist sogar die Güte der erhobenen Individualdaten direkt verifizierbar. Vor allem bei der Landtagswahl zeigen sich dabei problematische Ergebnisse. Die Güte der für die Schätzung der Wählerwanderung notwendigen Information, das Wahlverhalten bei beiden betrachteten Wahlen, ist auch mit diesen Daten nicht überprüfbar. Offensichtlich fehlerhafte Angaben legen jedoch nahe, dass es auch dort Probleme – in unbekannter Größenordnung – gibt.

225 Es sei darauf hingewiesen, dass diese Konstruktion von nur einen Teil der Befragten verstanden wurde und es vermehrt zur Angabe von gleichen Parteien bei der Befragung kam.

IV.3. Schätzung der Wählerwanderung in München: Bundestagswahlen 2009/2013

In diesem Kapitel erfolgt die Schätzung der Wählerwanderung zwischen den Bundestagswahlen 2009 und 2013 in München. Dabei wird auf die in dieser Arbeit vorgestellten Erkenntnisse zurückgegriffen. Zur Schätzung wird das hybride Multinomial-Dirichlet Modell verwendet. Die Wahlergebnisse bei den Wahlen für die gebildeten 650 über die Zeit konstante Wahlbezirke sind in Abbildung 21 dargestellt.

Beim Vergleich der Wahlergebnisse in den Wahlbezirken zeigt sich, dass die CSU und die SPD im Vergleich zur Wahl im Jahr 2009 in fast allen Wahlbezirken bei Wahl im Jahr 2013 ihr Ergebnis verbessern konnte. Für die FDP, die Grünen und die Linken sind Verluste feststellbar, wobei vor allem die FDP 2013 deutlich schlechtere Ergebnisse erreichte. Die Alternative für Deutschland (AfD), die erstmals bei der Bundestagswahl 2013 angetreten ist, konnte ein mit den Linken vergleichbares Ergebnis in München erreichen. Leichte Zugewinne sind auch bei den sonstigen Parteien (Sonst) feststellbar. Da die Zahl der Nichtwähler bei der Wahl 2013 größer ist, sank die Wahlbeteiligung in den betrachteten Gebieten.

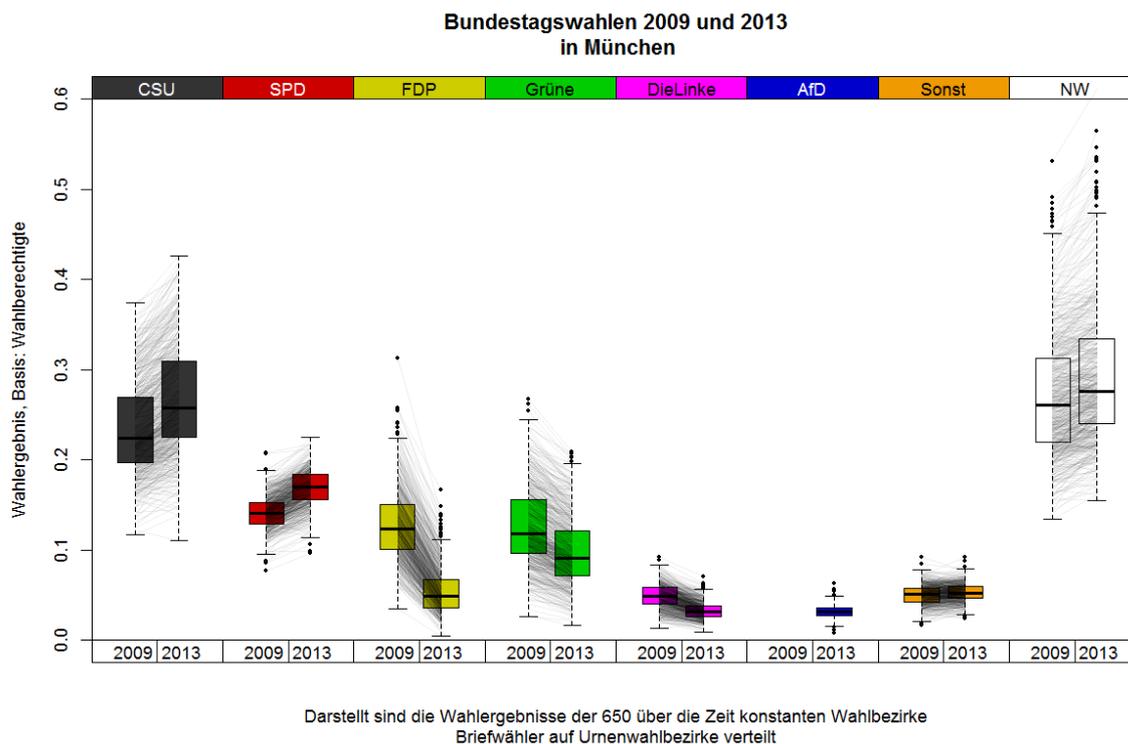


Abbildung 21: Wahlergebnisse (Zweitstimme) bei den Bundestagswahlen 2009 und 2013 in München für die gebildeten 650 über die Zeit konstanten Wahlbezirke. Briefwähler sind auf die Urnenwahlbezirke verteilt. Sonstige (Sonst) umfasst alle nicht explizit betrachteten Parteien, Nichtwähler (NW) beinhalten die Nichtwähler und ungültige Stimmen, Bevölkerungsveränderung wurden bei Beibehaltung der relativen Wahlergebnisse angepasst. Jede Linie repräsentiert bei jeder Partei jeweils einen Wahlbezirk. (Datenquelle: Stadt München)

Beim Vergleich des Wahlergebnisses 2013 mit den Ergebnissen bei den Bundestagswahlen 2005 und 2009, siehe Abbildung 3, zeigt sich, dass die CSU 2013 in etwa wieder das Niveau von 2005 erreichen konnte. Die SPD kann die Verluste bei der Wahl 2009 im Jahr 2013 nur zum Teil wiedergutmachen und liegt auch 2013 noch unter dem Niveau von 2005. Die FDP, die Grünen und die Linken verlieren 2013 wieder alle 2009 erlangten Zugewinne und schneiden 2013 zum Teil auch schlechter als 2005 ab.

IV.3.1. Datengrundlage und Datenaufbereitung

Bei der Schätzung der Wählerwanderung sollen bei der Bundestagswahl 2009 die CSU, die SPD, die FDP, die Grünen und die Linke betrachtet werden. Alle anderen Parteien werden zu Sonstigen zusammengefasst. Wie in Kapitel II.4. werden auch hier die Nichtwähler als Alternative mit aufgeführt. Bei der Bundestagswahl 2013 werden die gleichen sieben Alternativen und zusätzlich auch die AfD betrachtet. Bei der zu schätzenden Wählerwanderung handelt es sich daher um eine 7x8 Tabelle. Die Wählerwanderung wird dabei nur mit den Zweitstimmen, den Parteienstimmen bei der Bundestagswahl, geschätzt.

Datengrundlage für die Schätzung der Wählerwanderung mit dem hybriden Multinomial-Dirichlet Modell sind zum einen die amtlichen Wahlergebnisse der Bundestagswahlen 2009²²⁶ und 2013²²⁷ für die Stadt München. Es handelt sich dabei um die aggregierten Wahlentscheidungen auf Wahlbezirksebene. Die Daten umfassen jeweils die Informationen für die Urnen- und Briefwähler in gesonderten Wahlbezirken. Bei der Bundestagswahl 2009 gab es 664 Urnenwahlbezirke und 236 Briefwahlbezirk, bei der Bundestagswahl 2013 702 Urnenwahlbezirke und 237 Briefwahlbezirke.

Im ersten Schritt wurden, wie in Kapitel IV.1. dargelegt, bei jeder Wahl die Briefwähler auf die zugehörigen Urnenwahlbezirke verteilt. Dabei konnte auf Informationen vom Statistischen Amt der Landeshauptstadt München zurückgegriffen werden. Nach der Verteilung waren alle Wahlberechtigten den 664 bzw. 702 Urnenwahlbezirken zugeordnet. Die unterschiedliche Anzahl an Urnenwahlbezirken impliziert bereits, dass die Einteilung der Stadt München in Wahlbezirke bei beiden Wahlen nicht identisch war. Aus diesem Grund ist es für die Analyse notwendig, über die Zeit konstante Wahlbezirke zu erstellen. Insgesamt können 650 über die Zeit konstante Wahlbezirke bei diesen beiden Wahlen gebildet werden. Die dabei notwendige Aggregation fiel somit deutlich niedriger als noch

226 Statistisches Amt der Landeshauptstadt München. Amtliches Endergebnis der Bundestagswahl am 27.09.2009, Erst- und Zweitstimmenergebnisse der Landeshauptstadt München, 2009.

227 Statistisches Amt der Landeshauptstadt München. Amtliches Endergebnis der Bundestagswahl am 22.09.2013, Erst- und Zweitstimmenergebnisse der Landeshauptstadt München, 2013.

bei den Bundestagswahlen 2005 und 2009 aus, wo vergleichsweise nur 314 über die Zeit konstante Wahlbezirke gebildet werden konnten. Vorhandene Unterschiede in der Zahl der Wahlberechtigten in den Wahlbezirken werden durch eine proportionale Anpassung der Wahlergebnisse der Bundestagswahl 2009 an die Zahl der Wahlberechtigten bei der Bundestagswahl 2013 aufgelöst.

Die zweite Datenquelle für die Schätzung mit dem hybriden Modell stellen die Individualdaten dar. Diese stammen aus der Münchner Nachwahlbefragung. Von den 7122 ausgefüllten Fragebögen ist jedoch nur ein Teil für die Schätzung der Wählerwanderung verwendbar. Es ist notwendig, dass für beide Wahlen plausible Informationen vorliegen. Dies bedeutet, dass die Fragebögen mit fehlenden Angaben bei mindestens einer der beiden Wahlen oder mit nicht plausiblen Angaben nicht für die Schätzung herangezogen werden können. Eine nicht plausible Angabe ist dabei z. B. die Angabe eines AfD-Wählers bei der letzten Bundestagswahl die gleiche Partei gewählt zu haben. Insgesamt können 6566 Fragebögen genutzt werden.

Da es sich um eine Nachwahlbefragung handelt, sind weitestgehend die Nichtwähler bei der zweiten Wahl nicht erfasst wurden. Eine Ausnahme davon stellen die Wähler dar, die angegeben haben, nur ihre Erststimme abgegeben zu haben. Zusätzlich werden auch die Wähler, die angegeben haben ungültig gewählt zu haben, den Nichtwählern zugeordnet. Die gesamten Individualdaten über alle 17 Wahlbezirke sind in Abbildung 22 dargestellt. Hohe Werte sind vor allem bei den loyalen Wählern zu identifizieren, jedoch zeigt sich auch eine relativ starke Wählerwanderung von der FDP zur CSU und von den Grünen zur SPD.

Individualdaten der Münchner Nachwahlbefragung: Bundestagswahl

2009	CSU	1508	101	65	29	15	64	59	2
	SPD	115	1139	19	94	62	40	50	3
	FDP	358	57	309	33	18	82	45	3
	Grüne	87	281	16	762	82	37	73	1
	DieLinke	4	18	1	10	123	11	9	1
	Sonst	27	19	2	9	6	32	83	2
	NW	152	139	42	101	40	35	74	17
			CSU	SPD	FDP	Grüne	DieLinke	AfD	Sonst
		2013							

Abbildung 22: Individualdaten der Münchner Nachwahlbefragung bei der Bundestagswahl 2013 für die Bundestagswahlen 2009 und 2013 in München, die Zahlen repräsentieren die Anzahl der Wähler, die Angaben eine bestimmte Kombination gewählt zu haben. Es handelt sich um die Summe über alle 17 Wahlbezirke der Nachwahlbefragung. Als Nichtwähler (NW) bei der zweiten Wahl werden Personen betrachtet, die z. B. angegeben haben ungültig gewählt oder nur ihre Erststimme abgegeben zu haben (Datenquelle: Münchner Nachwahlbefragung)

IV.3.2. Modellparameter und -evaluierung

Für die Schätzung mit dem hybriden Multinomial-Dirichlet Modell wird als Burn-In ein Wert von 200000 und ein Thinning von 1250 als geeignet betrachtet. Insgesamt sollen 2000 Ziehungen aus der Kette betrachtet werden. Durch die größere Zahl der Ziehungen aus den Ketten besteht auch noch die Möglichkeit, den Burn-In weiter anzupassen, falls der gewählte Wert 200000 sich als unzureichend herausstellen sollte. Bei der Wahl der Parameter werden dabei die in Kapitel III.4.1. gemachten Erfahrungen berücksichtigt. Auch bei dieser Simulation ist ein größerer Burn-In aufgrund des Szenarios mit 600 Wahlbezirken notwendig. Ebenfalls aufgrund des Simulationsergebnisses werden nicht-informative Prioris, vergleiche Kapitel II.3.4.4., für die Schätzung verwendet. Mit diesen kann beim Szenario mit 600 Wahlbezirken eine leicht bessere Schätzgüte erreicht werden. Zur Evaluierung der Stabilität der Schätzung wurde das gleiche Modell mit identischer Parametrisierung dreimal gerechnet.

Im ersten Schritt wird überprüft, ob der Burn-In ausreichend groß gewählt wurde und die Ketten konvergiert sind. Dafür werden drei Teile der Ketten miteinander verglichen. Es handelt sich dabei um die ersten 200 Ziehungen, die mittleren 200 Ziehung und die letzten 200 Ziehungen. Bei ausreichenden Burn-In sind nur geringe Unterschiede zwischen den Ketten in allen drei Bereichen zu erwarten. Für den Vergleich werden jedoch nicht die einzelnen Ketten, sondern die mit diesen Teilketten geschätzten Wählerwanderungen genutzt. Es wird dabei als ausreichend für die Konvergenz betrachtet, wenn die Modelle gegen eine Wählerwanderungstabelle konvergieren. Für den Vergleich wird erneut der AD, vergleiche Gleichung (28), verwendet.

Das Ergebnis dieses Vergleiches ist in Tabelle 25 dargestellt. Für jedes der drei Modelle werden die drei Teilketten miteinander verglichen. Der AD zwischen den aus den Teilketten geschätzten Wählerwanderungen ist in allen Fällen sehr niedrig und fast immer unter 0,010. Über alle drei Modelle ist keine eindeutige Struktur zu erkennen, die nahelegt,

Mod1	1	2	3	Mod2	1	2	3	Mod3	1	2	3
1	0,000	0,005	0,009	1	0,000	0,009	0,011	1	0,000	0,008	0,009
2	0,005	0,000	0,008	2	0,009	0,000	0,009	2	0,008	0,000	0,006
3	0,009	0,008	0,000	3	0,011	0,009	0,000	3	0,009	0,006	0,000

Tabelle 25: Schätzung der Wählerwanderung zwischen den Bundestagswahlen 2009 und 2013 in München, Evaluierung des hybriden Multinomial-Dirichlet Modells, Konvergenz der Ketten. Die Schätzung erfolgt dreimal mit identischer Parametrisierung (Mod1, Mod2, Mod3). Für jedes Modell werden jeweils drei Teilbereiche der Ketten miteinander verglichen: 1-200 (1), 901-1100 (2) und 1801-2000 (3). Verglichen wird der AD der mit diesen Teilketten geschätzten Wählerwanderung.

dass es Probleme mit der Konvergenz bei dem jeweiligen Modell gegeben hat.

Beim Vergleich der drei Modellschätzungen untereinander, bei Nutzung aller 2000 Ziehungen, sind ebenfalls keine großen Unterschiede zu identifizieren. Der AD zwischen

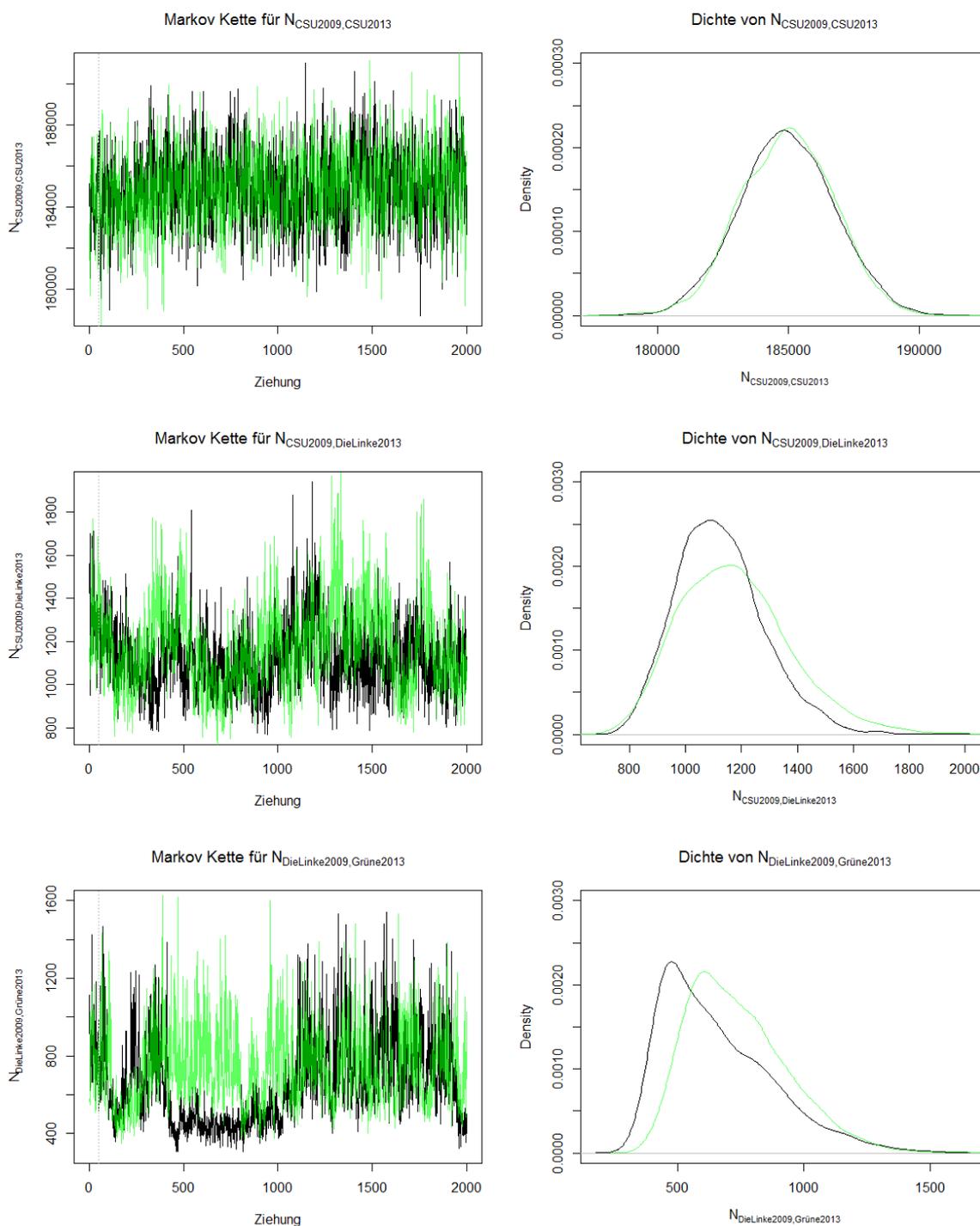


Abbildung 23: Schätzung der Wählerwanderung zwischen den Bundestagswahlen 2009 und 2013 in München, Evaluierung des hybriden Multinomial-Dirichlet Modells, Konvergenz der Ketten. Vergleich von zwei Modellschätzungen, Modell 1 (Mod1): schwarz, Modell 2 (Mod2): grün. Die linken Grafiken zeigen die Markov Ketten für drei ausgewählten Übergänge, die rechten Grafiken die zugehörigen Dichten

zwei Schätzungen beträgt maximal 0,004. Die Ergebnisse dokumentieren, dass der Burn-In ausreichend groß gewählt wurde und die Modelle konvergiert sind. Jedoch zeigen sich bei einer detaillierten Betrachtung der Ketten, siehe Abbildung 23, erneut die bereits angesprochen Probleme. Bei einigen der bereits konvergierten Ketten kann es zu einer kurzfristigen Abweichung vom eigentlich erreichten stabilen Niveau kommen. In Abbildung 23 ist dies sehr gut im Bereich der 500sten bis 1000sten Ziehung bei der Kette von Modell 1 (Mod1, schwarz) für $N_{\text{DieLinke2009,Grüne2013}}$ zu erkennen. Primär trifft dieses Phänomen aber vor allem bei relativ kleinen Übergängen auf und es hat daher, wie auch die sehr niedrige ADs zwischen den Modellschätzungen belegen, einen nur sehr geringen Einfluss auf die Schätzung.

IV.3.3. Modellergebnisse

Für die Schätzung der Wählerwanderung zwischen den Bundestagswahlen 2009 und 2013 in München werden die Ergebnisse von Modell 1 (Mod1) verwendet. Die geschätzte Wählerwanderung und die geschätzten Übergangswahrscheinlichkeiten sind in Abbildung 24 dargestellt. Bei den gezeigten Übergangswahrscheinlichkeiten handelt es sich um die Werte für die gesamte Stadt München.

Geschätzte Wählerwanderung mit dem hybriden Multinomial-Dirichlet Modell:
Bundestagswahlen 2009 und 2013 in München

2009	CSU	184933	7513	2384	1562	1124	5205	4765	7240
	SPD	6374	105121	998	3180	2920	4116	4743	1433
	FDP	41487	7194	43129	2806	1288	9905	4536	7752
	Grüne	4287	20897	1255	78528	5098	1969	4761	947
	DieLinke	656	2945	410	657	16770	1728	2368	19250
	Sonst	2196	2816	447	708	924	4641	23107	11137
	NW	6243	9283	1598	3970	1858	1891	4156	217853
			CSU	SPD	FDP	Grüne	DieLinke	AfD	Sonst
2013									

Geschätzte Übergangswahrscheinlichkeiten mit dem hybriden Multinomial-Dirichlet Modell:
Bundestagswahlen 2009 und 2013 in München

2009	CSU	0.86	0.03	0.01	0.01	0.01	0.02	0.02	0.03
	SPD	0.05	0.82	0.01	0.02	0.02	0.03	0.04	0.01
	FDP	0.35	0.06	0.37	0.02	0.01	0.08	0.04	0.07
	Grüne	0.04	0.18	0.01	0.67	0.04	0.02	0.04	0.01
	DieLinke	0.01	0.07	0.01	0.01	0.37	0.04	0.05	0.43
	Sonst	0.05	0.06	0.01	0.02	0.02	0.10	0.50	0.24
	NW	0.03	0.04	0.01	0.02	0.01	0.01	0.02	0.88
			CSU	SPD	FDP	Grüne	DieLinke	AfD	Sonst
2013									

Abbildung 24: Schätzung der Wählerwanderung zwischen den Bundestagswahlen 2009 und 2013 in München, hybrides Multinomial-Dirichlet Modell, geschätzte Wählerwanderung (oben) und Übergangswahrscheinlichkeiten (unten), präsentiert werden die Ergebnisse von Modell 1 (Mod1).

Für die Wähler der CSU und der SPD wird mit dem Modell eine hohe Loyalitätsrate geschätzt. 2013 entschieden sich demnach 86% (CSU) bzw. 82% (SPD) erneut für ihre Partei. Der höchste Wert ist jedoch mit 88% bei den Nichtwähler bei beiden Wahlen zu finden. Für allen anderen Parteien werden deutlich niedrigere Loyalitätsraten ermittelt, nur die Grünen haben mit 67% noch einen Wert, der deutlich über 50% liegt. Für die FDP und die Linke wird jeweils eine Loyalitätsrate von 37% geschätzt. Relevante Wählerbewegungen gibt es von der FDP zur CSU, den Grünen zur SPD, den Linken und den sonstigen Parteien (Sonst) zu den Nichtwählern. Die AfD konnte von fast allen Parteien Wähler gewinnen, die höchsten Übergangsraten werden dabei von den Sonstigen zur AfD und der FDP zur AfD geschätzt.

In Abbildung 25 werden die saldierten Übergänge zwischen den Parteien dargestellt. Zusätzlich wird auch angegeben, in welchen Maße die saldierten Abwanderungen an bzw. Gewinne von anderen Parteien das Wahlergebnis der jeweiligen Partei beeinflussen. Der absolut größte Wählerstrom kann von der FDP zur CSU identifiziert werden. Ungefähr 50% aller saldierten Stimmverluste der FDP gehen an die CSU. Dieser Wählerstrom

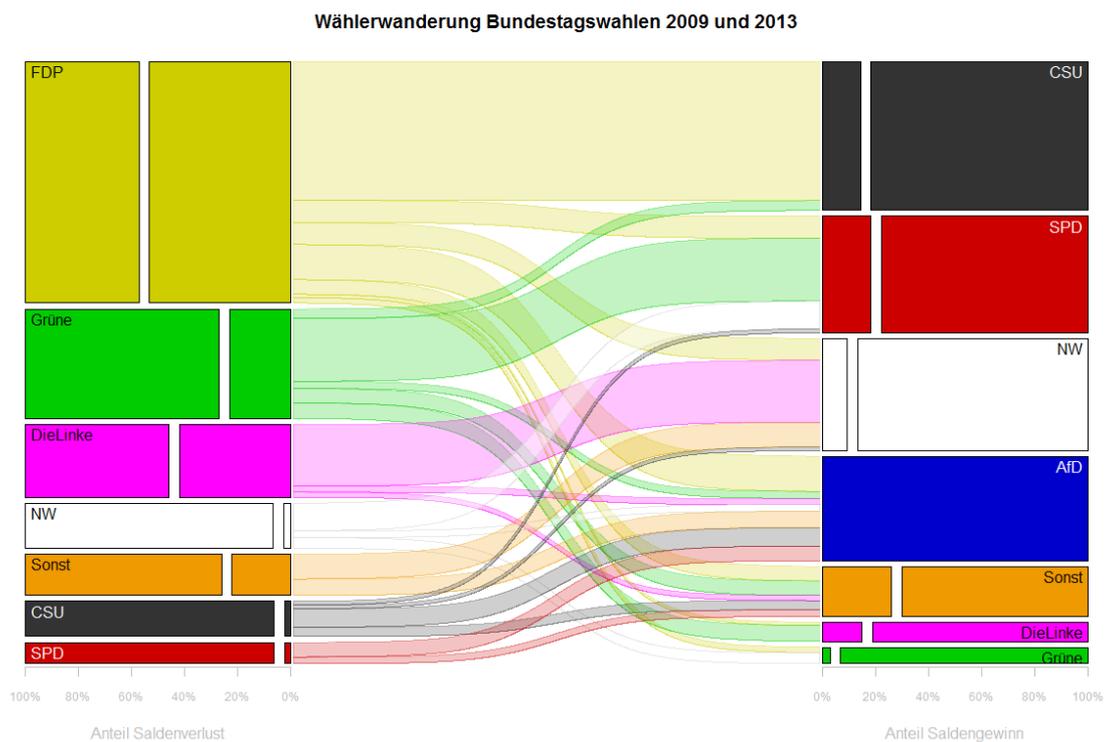


Abbildung 25: Schätzung der Wählerwanderung zwischen den Bundestagswahlen 2009 und 2013 in München, hybrides Multinomial-Dirichlet Modell, Darstellung der saldierten Wanderungsströme. Die Darstellungen an den Seiten geben jeweils an, welchen Anteil die saldierten Verluste (links) bzw. Gewinne (rechts) im Bezug auf die gesamten saldierten Wählerbewegungen (y-Achse) und im Bezug auf die Gesamtwählerschaft der Partei (x-Achse) darstellen.

repräsentiert auch ungefähr 20% aller saldierten Wanderungsbewegungen zwischen den Parteien. Weitere nennenswerte Verluste der FDP gehen an die AfD, die SPD und die Nichtwähler. Die starken Wählerbewegungen von der FDP zu anderen Parteien sorgen dafür, dass die FDP mehr als 50% ihrer Wählerschaft der Wahl 2009 an andere Parteien bei der Wahl 2013 verloren hat. Ebenfalls einen relevanten Anteil ihrer Wählerschaft haben die Linken, die Grünen und die sonstigen Parteien verloren. Die Verluste der Linken gehen dabei vor allem an die Nichtwähler, die der sonstigen Parteien an die Nichtwähler und die AfD und die der Grünen an die SPD.

Der Zugewinne der CSU von der FDP stellen fast die gesamten saldierten Stimmgewinne der CSU von anderen Parteien dar. Der Wählerstrom von den Grünen zur CSU ist weitestgehend vernachlässigbar. Auch wenn die CSU die absolut größten Zugewinne aller Parteien erlangte, bilden diese jedoch nur einen vergleichsweise kleinen Teil ihrer Wählerschaft bei der Wahl 2013. Die Stimmgewinne der SPD bei der Bundestagswahl 2013 sind vor allem auf Gewinne von den Grünen, der FDP und den Nichtwählern zurückzuführen. Die Stimmgewinne stellen dabei bei der SPD ungefähr 16% ihrer Gesamtstimmen bei der Bundestagswahl 2013 dar. Die Wählerschaft der AfD setzt sich vor allem aus ehemaligen Wählern der FDP und in schwächerem Umfang der CSU, SPD und der sonstigen Parteien zusammen.

Beim Vergleich mit der vom Statistischen Amt der Landeshauptstadt München mit dem Multinomial-Dirichlet Modell geschätzten Wählerwanderung für diese Wahlen²²⁸ fällt auf, dass die Werte weitestgehend eine vergleichbare Größenordnung haben. Unterschiede gibt es bei den geschätzten Loyalitätsraten, die beim hybriden Modell für fast alle Alternativen niedriger geschätzt werden. Nur für die FDP wird mit 37% ein fast identischer Wert geschätzt. Weitere relevante Unterschiede gibt es bei den geschätzten Übergängen von den Linken und den sonstigen Parteien (Sonst) zu den Nichtwählern. Diese werden beim Multinomial-Dirichlet Modell der Stadt München deutlich niedriger als beim hybriden Modell geschätzt.

Ein direkter Vergleich mit den Individualdaten in Abbildung 22 ist aufgrund der fehlenden Informationen über die Nichtwähler bei der zweiten Wahl schwierig. Dies veranschaulicht jedoch auch direkt eine der Stärken der hybriden Modellierung, da mit dem Modell die gesamte Wählerwanderung geschätzt werden kann. Bei einem Vergleich der geschätzten bzw. beobachteten Übergangsraten ohne die Nichtwähler bei der zweiten

228 Statistisches Amt der Landeshauptstadt München. Amtliches Endergebnis der Bundestagswahl am 22.09.2013, Analyse der Wählerwanderung für die Bundestagswahlen 2009 und 2013, 2013.

Wahl zeigt sich jedoch, dass die Loyalitätsraten für die CSU, die SPD, die Grünen und die sonstigen Parteien in den Individualdaten niedriger als beim hybriden Modell sind, während die Übergangsrate von den Sonstigen zur CSU in den Individualdaten mehr als doppelt so groß wie beim hybriden Modell ist.

Aufgrund der Ergebnisse in Kapitel III.3. und III.4.1. wird die Schätzung mit dem hybriden Multinomial-Dirichlet Modell als beste Schätzung der Wählerwanderung für die Bundestagswahlen 2009 und 2013 in München betrachtet. Bei der Betrachtung der Unterschiede zwischen dem Multinomial-Dirichlet Modell, dem hybriden Multinomial-Dirichlet Modell und den Individualdaten zeigt sich auch, dass die Schätzung mit dem hybriden Modell tendenziell zwischen der Schätzung mit der ökologischen Inferenz und den Individualdaten liegt, was als Indiz für die Mittlung der Informationen in den beiden genutzten Datenquellen beim hybriden Modell verstanden werden kann.

V. Schlussbemerkungen

In dieser Arbeit werden Methoden zur Schätzung von Wählerwanderungen in Mehrparteiensystemen vorgestellt und evaluiert. Gemeinsam haben die betrachteten Modelle und Verfahren, dass Aggregatdaten zur Schätzung herangezogen werden. Dabei wird zwischen der ökologischen Inferenz, bei der nur Aggregatdaten genutzt werden, und hybriden Modellen, bei denen Individual- und Aggregatdaten gemeinsam verwendet werden, unterschieden. Die vorgestellten Untersuchungen zeigen, dass mit beiden Modellklassen eine Schätzung der Wählerwanderung möglich ist.

Der Methodenvergleich bei der ökologischen Inferenz zeigt jedoch, dass nicht alle Verfahren im gleichen Maße geeignet sind. Das Beste der in dieser Arbeit verglichenen Modelle für die ökologische Inferenz ist das Multinomial-Dirichlet Modell. Dieses zeigt die besten Modelleigenschaften und die insgesamt beste Schätzgüte. Andere Modelle zeigen hingegen Eigenschaften, die sie als ungeeignet für die Schätzung von Wählerwanderungen erscheinen lassen. Dies gilt für die ökologische Inferenz nach Goodman und Thomsens Probit/Logit Modell. Das Multinomial-Normal Modell besitzt ebenfalls problematische Modelleigenschaften, die zumindest bei der praktischen Anwendung berücksichtigt werden müssen und unter Umständen auch zu einer Nichteignung führen können.

Mit der neueren Modellklasse der hybriden Modelle kann im Allgemeinen eine deutlich bessere Schätzgüte als mit der ökologischen Inferenz erreicht werden. Zum Teil kann der Fehler bei der Schätzung mehr als halbiert werden. Bei vorliegenden Individualdaten sind daher die hybriden Modelle der ökologischen Inferenz gegenüber zu bevorzugen. Der Nachteil dieser Modelle ist jedoch, dass zusätzlich zu den amtlichen Wahlergebnissen auch Individualdaten zur Verfügung stehen müssen. Der Aufwand bei der Datengewinnung ist daher deutlich größer und auch nicht immer können Individualdaten erhoben werden. Es zeigt sich aber auch, dass die Schätzung mit den hybriden Modellen der Hochrechnung aus den Individualdaten überlegen sein kann. Dies kann vor allem bei den realistischeren Szenarien festgestellt werden. Beim Vergleich der beiden hybriden Modelle schneidet das hybride Multinomial-Dirichlet Modell besser ab, auch wenn dieses nicht in allen Fällen die beste Schätzgüte erreicht.

Die Kombination von Individual- und Aggregatdaten erfolgt bei den hybriden Modellen vor dem Hintergrund, dass bei der Schätzung von Wählerwanderungen die Schätzung jeweils alleine aus einer der beiden Datenquellen kritisch gesehen werden kann. Zwar ist

eine Schätzung von Wählerwanderungen mit der ökologischen Inferenz möglich, jedoch ist die Güte der Modellschätzung bei der praktischen Anwendung nicht ohne weiteres bestimmbar. Weiterhin zeigen die vorgestellten Untersuchungen, dass die Annahmen der Modelle Einfluss auf die Schätzung nehmen. Bei Individualdaten wiederum sind mehrere Faktoren bekannt, die zu falschen Angaben oder Verzerrungen in den Daten führen können. Entsprechendes kann auch bei der Münchner Nachwahlbefragung festgestellt werden. Es gibt falsche Angaben, Erinnerungsprobleme und die Wahrscheinlichkeit an der Nachwahlbefragung teilzunehmen hing von der Wahlentscheidung ab. Mit der Kombination von Individual- und Aggregatdaten bei den hybriden Modellen wird versucht, die Schwächen der jeweiligen Datenquelle durch die andere Datenquelle abzumildern. Dieser Effekt ist auch bei der Schätzung der Wählerwanderung zwischen den Bundestagswahlen 2009 und 2013 feststellbar. Die Schätzwerte des hybriden Modells liegen tendenziell zwischen denen der ökologischen Inferenz und den Individualdaten.

Weiterhin besteht die Hoffnung, dass so die Schätzung von Wählerwanderungen in mehr Fällen möglich ist. Es ist aufgrund des damit verbundenen Aufwands nicht immer möglich, Individualdaten im für eine Hochrechnung ausreichenden Maße und Güte zu erheben. Mit hybriden Modellen kann ggf. eine Schätzung der Wählerwanderung auch dann erfolgen, wenn die Individualdaten als nicht ausreichend für eine valide Hochrechnung betrachtet werden. Ebenfalls zeigen die durchgeführten Betrachtungen, dass die Schätzung mit der ökologischen Inferenz in kleineren Regionen mit eher wenig Wahlbezirken schlechter ist. Die Kombination von Individual- und Aggregatdaten könnte jedoch auch in diesen Fällen eine Schätzung der Wählerwanderung mit überschaubarem Aufwand erlauben.

Zusätzlich zur Frage der theoretischen Eignung der Modelle bestehen bei der Schätzung von Wählerwanderungen auch praktische Probleme bei der Anwendung. Bei den Modellen wird im Allgemeinen angenommen, dass bei beiden Wahlen die gleiche Bevölkerung betrachtet wird. Dies ist in der Praxis nicht der Fall. Die Wahlberechtigten eines Wahlbezirks sind bei zwei aufeinanderfolgenden Wahlen nur zum Teil identisch. Ein Teil der Wahlberechtigten ist verstorben oder weggezogen, andere werden erstmals wahlberechtigt oder sind zugezogen. Dieser Aspekt hat vor allem Auswirkungen auf die Interpretation der geschätzten Wählerwanderung, im geringeren Maße auch auf die Datenaufbereitung. Zusätzlich muss ebenfalls sichergestellt werden, dass ein Wahlbezirk in beiden Wahlen in etwa das gleiche Gebiet beschreibt. Da sich der Zuschnitt und die Benennung der Wahlbezirke im Laufe der Zeit jedoch ändert, ist die Schaffung von

vergleichbaren Gebieten, von über die Zeit konstanten Wahlbezirken, unter Umständen mit einem erheblichen Aufwand verbunden. Jedoch ist aufgrund der Abhängigkeit der Schätzgüte von der Zahl der zur Verfügung stehenden Gebiete in der betrachteten Region eine möglichst feine Unterteilung zu bevorzugen. Relevant ist auch der Umgang mit den Briefwählern, die bei Wahlen in gesonderten Briefwahlbezirken erfasst werden. Die durchgeführten Betrachtungen zeigen, dass eine Verteilung der Briefwähler auf die Urnenwahlbezirke sinnvoll ist und die Schätzgüte so verbessert werden kann.

Problematisch beim Modell mit der besten Schätzgüte, dem (hybriden) Multinomial-Dirichlet Modell, ist, dass die Kreditabilitätsintervalle nur unzureichend die Unsicherheit der Schätzung beschreiben. Es handelt sich dabei nicht um ein spezifisches Problem dieses Modells, vergleichbare Schwierigkeiten bestehen auch bei anderen Modellen und nicht bei jedem Modell ist die Angabe der Schätzunsicherheit ohne weiteres überhaupt möglich. Als vermutliche Ursache für die schlechte Überdeckung wird der Bias der Schätzung identifiziert. Mit zwei der untersuchten Korrekturverfahren kann die Überdeckung der Kreditabilitätsintervalle deutlich verbessert werden, wobei jedoch zum Teil auch zu gute Überdeckungsrate erreicht werden. Weiterhin ist auch nach Korrektur nicht sichergestellt, dass die Überdeckungsrate der Kreditabilitätsintervalle für die einzelnen Übergänge der Erwartung entsprechen.

Weiterer Forschungsbedarf besteht daher bei der Bestimmung der Unsicherheit der Schätzer. Da im Allgemeinen die Schätzgüte bei der ökologischen Inferenz und bei hybriden Modellen nicht bestimmt werden kann, ist eine korrekte Angabe der Unsicherheit der Schätzung für eine sinnvolle Interpretation der Ergebnisse umso wichtiger. Eine Möglichkeit wäre zu versuchen, aufbauend auf den in dieser Arbeit vorgestellten Ideen, den Bias für die einzelnen Übergänge spezifischer zu schätzen und so die Güte der Kreditabilitätsintervalle zu verbessern. Es ist jedoch offen, ob dies mit den im Allgemeinen vorliegenden Informationen ohne weiteres möglich ist. Bei hybriden Modellen stehen zusätzlich die Individualdaten zur Verfügung, welche mit zur Bestimmung der Unsicherheit herangezogen werden können und welche auch weitere Möglichkeiten zur Korrektur eröffnen.

Ebenfalls noch weiter untersucht werden muss der für eine sinnvolle Schätzung notwendige Umfang der Individualdaten bei den hybriden Modellen. Die vorgestellten Ergebnisse legen nahe, dass eine Nachwahlbefragung an 2,5% der Wahlbezirke ausreicht. Offen ist jedoch, ob eine weitere Reduktion möglich ist und ob dieser Wert auch für relativ kleine Regionen gilt, bei denen 2,5% der Wahlbezirke ein oder zwei Wahlbezirken

entspricht. Weiterhin ist zu untersuchen, welcher Umfang in den einzelnen Wahlbezirken zu erheben ist. In den Simulationsstudien wird fast immer von einer Vollerhebung mit relativ guter Response-Rate ausgegangen. Eine kleinere Stichprobe würde jedoch den Erhebungsaufwand deutlich reduzieren.

Während sich die in dieser Arbeit vorgestellten Ergebnisse primär auf die Schätzung von Wählerwanderungen und auch auf die damit verbundene Datensituation beziehen, ist die Nutzung nicht aller vorgestellten Modelle alleinig auf diesen Anwendungsfall beschränkt. Beim (hybriden) Multinomial-Dirichlet Modell handelt es sich z. B. um ein allgemeines Modell für die ökologische Inferenz bzw. hybride Analysen. Auch die vorgestellten Ergebnisse sind, zumindest teilweise, auf andere Anwendungsgebiete übertragbar. Notwendig ist jedoch, dass eine in etwa vergleichbare Struktur beim zugrundeliegenden Schätzproblem angenommen werden kann. Kritisch wird eine Übertragung der Ergebnisse, wenn der zugrundeliegende Prozess deutlich von den hier vorgestellten Sachverhalt abweicht.

VI. Quellen

VI.1. Literaturverzeichnis

- Achen, Christopher H. / Shively, W. Phillips: Cross-level inference, Chicago, 1995.
- Alkaya, Levent / Bracher, Johannes / Seibold, Heidi / Shao, Shuai: Analyse von Wahlen in München, Praktikumsbericht, 2012.
- Ambühl, Mathias: Methoden zur Rekonstruktion von Wählerströmen aus Aggregatdaten. Bundesamt für Statistik, Neuenburg, 2003.
- Andreadis, Ioannis / Chadjipadelis, Theodore: A Method for the Estimation of Voter Transition Rates. *Journal of Elections, Public Opinion and Parties*, 19(2), 2009, S. 203–218.
- Anselin, Luc / Cho, Wendy K. Tam: Spatial Effects and Ecological Inference. *Political Analysis*, 10(3), 2002, S. 276–297.
- Brown, Philip J. / Payne, Clive D.: Aggregate Data, Ecological Regression, and Voting Transitions. *Journal of the American Statistical Association*, Volumen 81, Nummer 394, 1986, S. 452–460.
- Cho, Wendy K. Tam: Iff the Assumption Fits...: A Comment on the King Ecological Inference Solution. *Political Analysis*, 7(1), 1998, S. 143–163.
- Cho, Wendy K. Tam / Gaines, Brian J.: The Limits of Ecological Inference. The Case of Split-Ticket Voting, in: *American Journal of Political Science*, Volumen 48, Nummer 1, 2004, S. 152–171.
- Cho, Wendy K. Tam / Manski, Charles F.: Cross – Level / Ecological Inference, in: Brady, Henry E. / Collier, David / Box-Steffensmeier, Janet M.: *Oxford Handbook of Political Methodology*, Oxford, 2008, S. 547–569.
- Deming, W. Edwards / Stephan, Frederick F.: On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *Ann. Math. Stat.*, 11(4), 1940, S. 427–444.
- Duncan, Otis Dudley / Davis, Beverly: An alternative to ecological correlation. *American Sociological Review*, 18(6), 1953, S. 665–666.
- Efron, Bradley / Tibshirani, Robert: Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical science*, 1(1), 1986, S. 54–75.

- Elff, Martin / Gschwend, Thomas / Johnston, Ron J.: Ignoramus, ignorabimus? On uncertainty in ecological inference. *Political Analysis*, 16(1), 2008, S. 70–92.
- Ferree, Karen E.: Iterative Approaches to R x C Ecological Inference Problems. Where They Can Go Wrong and One Quick Fix. *Political Analysis*, 12(2), 2004, S. 143 – 159.
- Freedman, David A.: Ecological Inference and the Ecological Fallacy, Technical Report No. 549, Berkely, 1999.
- Greiner, D. James / Quinn, Kevin M.: RxC ecological inference: bounds, correlations, flexibility and transparency of assumptions. *Journal of the Royal Statistical Society*, 172(1), 2009, S. 67–81.
- Greiner, D. James / Quinn, Kevin M.: Exit Polling and Racial Bloc Voting: Combining Individual-Level and R x C Ecological Data. *The Annals of Applied Statistics*, 4(4), 2010, S. 1774–1796.
- Greiner, D. James / Quinn, Kevin M.: Long Live the Exit Poll. *Daedalus*, 141(4), 2012, S. 9–22.
- Goodman, Leo A.: Ecological Regressions and Behavior of Individuals. *American Sociological Review*, 18(6), 1953, S. 663–664.
- Grofman, Bernard / Merrill, Samuel: Ecological Regression and Ecological Inference. In: King, Gary / Rosen, Ori / Tanner, Martin A. (Editoren): *Ecological Inference: New Methodological Strategies*, Cambridge, 2004, S. 123–143.
- Gschwend, Thomas: Ökologische Inferenz. In: Behnke, Joachim / Gschwend, Thomas / Schindler, Delia / Schnapp, Kai-Uwe (Editoren): *Methoden der Politikwissenschaft: neuere qualitative und quantitative Analyseverfahren.*, Baden-Baden, 2006, S. 227–237.
- Horowitz, Joel L.: The bootstrap. *Handbook of econometrics*, 5, 2001, S. 3159–3228.
- Kellermann, Thomas: Vom Wahlergebnis zur Wählerwanderung: Welche Wähler wechselten wie ihre Entscheidung, *Stadtforschung und Statistik*, 2011(1), 2011, S. 34–40.
- King, Gary: *A Solution to the Ecological Inference Problem. Reconstructing Individual Behavior from Aggregate Data*, Princeton, 1997.
- King, Gary / Rosen, Ori / Tanner, Martin A.: Binomial-Beta Hierarchical Models for Ecological Inference. *Sociological Methods & Research*, 28(1), 1999, S. 61–90.
- King, Gary / Rosen, Ori / Tanner, Martin A./ Wagner, Alexander F.: Ordinary economic voting behavior in the extraordinary election of Adolf Hitler. *Journal of Economic History*, 68(04), 2008, S. 951–996.

- Klima, André / Küchenhoff, Helmut / Thurner, Paul W.: Analysis of voter transition using ecological data: Comparison of different approaches for Munich election data. Proceedings of the 28th International Workshop on Statistical Modelling (Palermo, 08.07.-12.07.2013), 1, 2013, S. 203–207.
- Klima, André / Thurner, Paul W. / Molnar, Christoph / Schlesinger, Thomas / Küchenhoff, Helmut: Estimation of Voter Transitions Based on Ecological Inference: An Empirical Assessment of Different Approaches. *Advances Statistical Analysis*, 100(2), 2016, S. 133–159.
- Klima, André / Schlesinger, Thomas / Küchenhoff, Helmut / Thurner Paul W.: Combining aggregate data and exit-polls for the estimation of voter transitions, unveröffentlichtes Manuskript, 2016.
- Hilmer, Richard / Kunert, Michael: Wählerwanderung: Das Modell von Infratest dimap. In: *Wahlen und Wähler*. VS Verlag für Sozialwissenschaften, 2005, S. 134–156.
- Lau, Olivia / Moore, Ryan .T. / Kellermann, Micheal: eiPack: RxC ecological inference and higher-dimension data management. *R News*, 7(2), 2007, S. 43–47.
- Neu, Viola: Bundestagswahl in Deutschland am 27. September 2009, Konrad-Adenauer-Stiftung, 2009.
- Park, Won-ho: *Ecological Inference and Aggregate Analysis of Elections*, The University of Michigan, Dissertation, 2008.
- Robert, Christian: *The Bayesian choice: from decision-theoretic foundations to computational implementation*. Springer Science & Business Media, 2007.
- Robinson, William S.: Ecological correlations and the behavior of individuals. *American Sociological Review*, 15(3), 1950, S. 351–357.
- Rosen, Ori / Jiang, Wenxin / King, Gary / Tanner, Martin A.: Bayesian and frequentist inference for ecological inference: the R x C case. *Statistica Neerlandica*, 55(2), 2001, S. 134–156.
- Schlesinger, Thomas: *Kombination von Aggregat- und Individualdaten bei der Analyse von RxC-Tafeln*. Neue Implementierungen in R, Masterarbeit, 2013.
- Salway, Ruth / Wakefield, Jon: A hybrid model for reducing ecological bias. *Biostatistics*, 9(1), 2008, S. 1–17.
- Smoot, Elizabeth / Haneuse, Sebastien: On the Analysis of Hybrid Designs that Combine Group- and Individual-Level Data, *Biometrics*, 71(1), 2015, S. 227–236.

Sora Institut: SORA Wählerstromanalysen, online:

<http://www.sora.at/themen/wahlverhalten/wahlanalysen/waehlerstromanalysen.html>

(zuletzt aufgerufen am: 13.04.2016, Archiv: <http://archive.is/NP6Jf>)

Turner, Paul W. / Klima, André / Küchenhoff, Helmut: Agricultural structure and the rise of the Nazi Party reconsidered. *Political Geography* 44, 2015, S. 50–63.

Thomsen, Søren R.: Danish elections 1920–79: A Logit Approach to Ecological Analysis and Inference, Aarhus, 1987.

Wakefield, Jon: Ecological inference for 2x2 tables, *Journal of Royal Statistical Society A*, 167(3), 2004, S. 385–445.

VI.2. Datenverzeichnis

Lehrstuhls für empirische Politikforschung, Prof. Paul W. Thurner, und Statistisches Beratungslabors, Prof. Helmut Küchenhoff (beide Ludwig-Maximilians-Universität München), Münchner Nachwahlbefragung. Nachwahlbefragung bei der Landtagswahl am 15.09.2013 und der Bundestagswahl am 22.09.2013 in München, 2013.

Statistisches Amt der Landeshauptstadt München. Amtliches Endergebnis der Bundestagswahl am 18.09.2005, Erst- und Zweitstimmenergebnisse der Landeshauptstadt München, 2005. <http://www.mstatistik-muenchen.de/themen/wahlen/wahlberichterstattung/btw2005/index.html>. (zuletzt aufgerufen am: 08.06.2016)

Statistisches Amt der Landeshauptstadt München. Amtliches Endergebnis der Bundestagswahl am 27.09.2009, Erst- und Zweitstimmenergebnisse der Landeshauptstadt München, 2009. <http://www.mstatistik-muenchen.de/themen/wahlen/wahlberichterstattung/btw2009/index.html>. (zuletzt aufgerufen am: 03.07.2016)

Statistisches Amt der Landeshauptstadt München. Amtliches Endergebnis der Landtagswahl am 15.09.2013, Analyse der Wählerwanderung für die Landtagswahlen 2008 und 2013, 2013. <https://www.muenchen.de/rathaus/Stadtfinfos/Statistik/Wahlen/Landtagswahl.html>. (zuletzt aufgerufen am: 17.07.2016)

Statistisches Amt der Landeshauptstadt München. Amtliches Endergebnis der Landtagswahl am 15.09.2013, Zweitstimmenergebnisse der Landeshauptstadt München, 2013. <https://www.muenchen.de/rathaus/Stadtfinfos/Statistik/Wahlen/Landtagswahl.html>. (zuletzt aufgerufen am: 25.07.2016)

Statistisches Amt der Landeshauptstadt München. Amtliches Endergebnis der Bundestagswahl am 22.09.2013, Analyse der Wählerwanderung für die Bundestagswahlen 2009 und 2013, 2013. <https://www.muenchen.de/rathaus/Stadtfinfos/Statistik/Wahlen/Bundestagswahl.html>. (zuletzt aufgerufen am: 18.07.2016)

Statistisches Amt der Landeshauptstadt München. Amtliches Endergebnis der Bundestagswahl am 22.09.2013, Erst- und Zweitstimmenergebnisse der Landeshauptstadt München, 2013. <https://www.muenchen.de/rathaus/Stadtfinfos/Statistik/Wahlen/Bundestagswahl.html>. (zuletzt aufgerufen am: 25.07.2016)

VI.3. Software / Pakete

Greiner, D. James / Baines, Paul / Quinn, Kevin M.: RxCEcolInf. R x C Ecological Inference With Optional Incorporation of Survey Information. R package version 0.1-3., 2013, <http://CRAN.R-project.org/package=RxCEcolInf>.

Harrell Jr, Frank E / with contributions from Charles Dupont and many others: Hmisc: Harrell Miscellaneous. R package version 3.15-0, 2015, <http://CRAN.R-project.org/package=Hmisc>.

Lau, Olivia / Moore, Ryan T. / Kellermann, Michael: eiPack: Ecological Inference and Higher-Dimension Data Management, R package version 0.1-7, 2012. <http://CRAN.R-project.org/package=eiPack>.

Martin, Andrew D. / Quinn, Kevin M. / Park, Jong Hee: MCMCpack: Markov Chain Monte Carlo in R. Journal of Statistical Software. 42(9), 2011, S. 1–21, URL <http://www.jstatsoft.org/v42/i09/>.

R Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2013, URL: <http://www.R-project.org/>.

Revolution Analytics / Weston, Steve: doSNOW: Foreach parallel adaptor for the snow package. R package version 1.0.12, 2014, <http://CRAN.R-project.org/package=doSNOW>.

Revolution Analytics / Weston, Steve: foreach: Foreach looping construct for R. R package version 1.4.2, 2014, <http://CRAN.R-project.org/package=foreach>.

Schlesinger, Thomas: eiwild: Ecological Inference with individual and aggregate data. R package version 0.6.7, 2014. <http://CRAN.R-project.org/package=eiwild>.

Thomsen, Søren R. / Frandsen, Annie G. / Kristmar, Thomas / Lauritsen, Per / Sørensen, Michael B.: Ecol: Version 3, Aarhus University, 1995.

Wickham, Hadley: ggplot2: elegant graphics for data analysis. Springer New York, 2009.

Wickham, Hadley: Reshaping Data with the reshape Package. Journal of Statistical Software, 21(12), 2007, S. 1–20. URL <http://www.jstatsoft.org/v21/i12/>.

VII. Abbildungsverzeichnis

Abbildung 1	Wählerströme von bzw. zur Union (CDU, CSU) bei der Bundestagswahl 2013 im Vergleich zur Bundestagswahl 2009. Dargestellt sind die saldierten Wählerwanderungen, d. h. die absoluten Gewinne bzw. Verluste der Union.	9
Abbildung 2	Darstellung der Auswirkung der Wahl der Parameter bei den Hyperprioris beim Multinomial-Dirichlet Modell; die erste Zeile zeigt eine für allen Spalten identische Wahl, die zweite Zeile zeigt für die erste Spalte eine abweichende Wahl, so dass für diesen Parameter eher höhere Übergangswahrscheinlichkeiten a priori erwartet werden.	42
Abbildung 3	Wahlergebnisse (Zweitstimme) bei den Bundestagswahlen 2005 und 2009 in München für die gebildeten 314 über die Zeit konstanten Wahlbezirke. Briefwähler sind auf die Urnenwahlbezirke verteilt.	45
Abbildung 4	Evaluierung der Methoden für die ökologische Inferenz, Multinomial-Dirichlet Modell, Schätzung der Wählerwanderung bei den Bundestagswahlen 2005 und 2009 in München, Vergleich von zwei Modellschätzungen. Die Parameter bei beiden Modellen sind identisch. Die linken Grafiken zeigen die Markov Ketten für drei ausgewählten Übergänge, die rechten Grafiken die zugehörigen Dichten.	52
Abbildung 5	Evaluierung der Methoden für die ökologische Inferenz, Multinomial-Normal Modell, Schätzung der Wählerwanderung bei den Bundestagswahlen 2005 und 2009 in München, Vergleich von zwei Modellschätzungen. Die Parameter bei beiden Modellen sind identisch. Die linken Grafiken zeigen die Markov Ketten für drei ausgewählten Übergänge, die rechten Grafiken die zugehörigen Dichten.	54
Abbildung 6	Evaluierung der Methoden für die ökologische Inferenz, Schätzung der Wählerwanderung bei den Bundestagswahlen 2005 und 2009 in München, dargestellt sind die geschätzten absoluten Zahlen für jede Kombination. Unterschiede vom Wahlergebnis sind Algorithmus bedingt. oben: Multinomial-Dirichlet Modell (Burn-In: 50000, Thinning: 1000, Ziehungen: 1000), unten: Thomsens Probit/Logit Modell mit Originalalgorithmus.	59
Abbildung 7	Evaluierung der Methoden für die ökologische Inferenz, angenommene Wählerwanderung für München bei der Simulation, dargestellt sind die Übergangswahrscheinlichkeiten.	61
Abbildung 8	Evaluierung der Methoden für die ökologische Inferenz, Ergebnisse der Simulationsstudie, dargestellt ist der AD der Schätzung mit den Modellen zu den simulierten Individualdaten.	65

Abbildung 9	Simulation zur Evaluierung der hybriden Modelle, angenommene Wählerwanderung für die Simulation, dargestellt sind die Übergangswahrscheinlichkeiten.	86
Abbildung 10	Simulation zur Evaluierung der hybriden Modelle, Ergebnisse der Simulationsstudie, dargestellt ist der AD der Schätzung mit den Modellen zu den simulierten Individualdaten.	91
Abbildung 11	Detaillierte Evaluierung des (hybriden) Multinomial-Dirichlet Modells, angenommene Wählerwanderung für die Simulation, dargestellt sind die Übergangswahrscheinlichkeiten.	97
Abbildung 12	Detaillierte Evaluierung des (hybriden) Multinomial-Dirichlet Modells, Ergebnisse der Simulationsstudie, dargestellt ist der AD der Schätzung mit den Modellen zu den simulierten Individualdaten.	100
Abbildung 13	Detaillierte Evaluierung des (hybriden) Multinomial-Dirichlet Modells, Schätzgüte des Multinomial-Dirichlet Modells bei den Daten der Simulationsstudie zum Vergleich der Methoden der ökologischen Inferenz, dargestellt ist der AD der Schätzung zu den simulierten Individualdaten.	105
Abbildung 14	Detaillierte Evaluierung des (hybriden) Multinomial-Dirichlet Modells, Überdeckungsrate der 80%-Kreditabilitätsintervalle der $N_{r,c}$ beim Multinomial-Dirichlet Modell bei den Daten der Simulationsstudie zum Vergleich der Methoden der ökologischen Inferenz, dargestellt ist für jede geschätzte Wählerwanderung der Anteil der Kreditabilitätsintervalle, die den Wert in den simulierten Individualdaten überdecken.	106
Abbildung 15	Detaillierte Evaluierung des (hybriden) Multinomial-Dirichlet Modells, Darstellung des zellspezifischen Bias bei der Schätzung mit dem Multinomial-Dirichlet Modell bei Szenario Sim 2, dargestellt ist die Abweichung in Wählern/Wahlberechtigten zwischen bei den 30 Datensätzen.	107
Abbildung 16	Detaillierte Evaluierung des (hybriden) Multinomial-Dirichlet Modells, Korrektur der Kreditabilitätsintervalle, Überdeckungsrate der korrigierten 80%-Kreditabilitätsintervalle bei den Daten der Simulationsstudie zum Vergleich der Methoden der ökologischen Inferenz, dargestellt ist für jedes Modell und Korrekturverfahren der Anteil der Kreditabilitätsintervalle, die den Wert in den simulierten Individualdaten überdecken.	113
Abbildung 17	Detaillierte Evaluierung des (hybriden) Multinomial-Dirichlet Modells, Korrektur der Kreditabilitätsintervalle, 80% Kreditabilitätsintervalle beim Multinomial-Dirichlet Modell nach Anwendung des Korrekturverfahrens drei bei Datensatz eins von Szenario Sim 2. Untere Grenze der Kreditabilitätsintervalle, Mittelwerte der Ketten und obere Grenze der Kreditabilitätsintervalle.	115

- Abbildung 18 Simulation zur Datenaufbereitung bei der ökologischen Inferenz, 127
Ergebnisse der Simulationsstudie für das Multinomial-Dirichlet Modell, dargestellt ist der AD der Schätzung mit den Modellen zu den simulierten Individualdaten.
- Abbildung 19 Münchner Nachwahlbefragung, Partei-spezifische Response-Raten in 135
den 17 Wahlbezirken mit Nachwahlbefragung bei der Landtagswahl. Die jeweilige Response-Rate ist der Quotient von Wählern einer Partei in der Nachwahlbefragung und dem Wahlergebnis der Partei in den Wahlbezirk in den amtlichen Wahlstatistiken.
- Abbildung 20 Münchner Nachwahlbefragung, Partei-spezifische Response-Raten in 136
den 17 Wahlbezirken mit Nachwahlbefragung bei der Bundestagswahl. Die jeweilige Response-Rate ist der Quotient von Wählern einer Partei in der Nachwahlbefragung und dem Wahlergebnis der Partei in den Wahlbezirk in den amtlichen Wahlstatistiken.
- Abbildung 21 Wahlergebnisse (Zweitstimme) bei den Bundestagswahlen 2009 und 139
2013 in München für die gebildeten 650 über die Zeit konstanten Wahlbezirke. Briefwähler sind auf die Urnenwahlbezirke verteilt.
- Abbildung 22 Individualdaten der Münchner Nachwahlbefragung bei der 141
Bundestagswahl 2013 für die Bundestagswahlen 2009 und 2013 in München, Anzahl der Wähler, die Angaben eine bestimmte Kombination gewählt zu haben. Es handelt sich um die Summe über alle 17 Wahlbezirke der Nachwahlbefragung.
- Abbildung 23 Schätzung der Wählerwanderung zwischen den Bundestagswahlen 143
2009 und 2013 in München, Evaluierung des hybriden Multinomial-Dirichlet Modells, Konvergenz der Ketten. Vergleich von zwei Modellschätzungen. Die linken Grafiken zeigen die Markov Ketten für drei ausgewählten Übergänge, die rechten Grafiken die zugehörigen Dichten.
- Abbildung 24 Schätzung der Wählerwanderung zwischen den Bundestagswahlen 144
2009 und 2013 in München, hybrides Multinomial-Dirichlet Modell, geschätzte Wählerwanderung (oben) und Übergangswahrscheinlichkeiten (unten).
- Abbildung 25 Schätzung der Wählerwanderung zwischen den Bundestagswahlen 145
2009 und 2013 in München, hybrides Multinomial-Dirichlet Modell, Darstellung der saldierten Wanderungsströme.
- Abbildung 26 Evaluierung der Methoden für die ökologische Inferenz, Ergebnisse 170
der Simulationsstudie, weiterführende Evaluierung des Multinomial-Normal Modells, Vergleich von zwei Modellschätzungen. Die Parameter bei beiden Modellen sind identisch. Die linken Grafiken zeigen die Markov Ketten für drei ausgewählten Übergänge, die rechten Grafiken die zugehörigen Dichten.

- Abbildung 27 Münchner Nachwahlbefragung, Fragebogen bei der Münchner 174
Nachwahlbefragung für die Landtagswahl am 15.09.2013.
- Abbildung 28 Münchner Nachwahlbefragung, Fragebogen bei der Münchner 175
Nachwahlbefragung für die Bundestagswahl am 22.09.2013,
Vorderseite.
- Abbildung 29 Münchner Nachwahlbefragung, Fragebogen bei der Münchner 176
Nachwahlbefragung für die Bundestagswahl am 22.09.2013,
Rückseite.

VIII. Tabellenverzeichnis

Tabelle 1	Darstellung der Datensituation bei der ökologischen Inferenz, die hellgrau markierten Ränder sind bekannt, die weißen inneren Zellen unbekannt. Die angegebenen Wahlergebnisse entsprechen dem Ergebnis der Bundestagswahl 2009 respektive 2013 in München.	10
Tabelle 2	Darstellung einer möglichen Wählerwanderung in München zwischen den Wahlen 2009 und 2013, angenommen werden hohe Loyalitätsraten.	11
Tabelle 3	Darstellung einer möglichen Wählerwanderung in München zwischen den Wahlen 2009 und 2013. Angenommen werden niedrige Loyalitätsraten.	12
Tabelle 4	Darstellung der Datensituation bei der ökologischen Inferenz, die hellgrau markierten Ränder sind bekannt, die weißen inneren Zellen unbekannt. In den inneren Zellen sind die jeweiligen Übergangsraten ergänzt. Es handelt sich um eine vereinfachte Darstellung, da sich in den inneren Zellen nur die Übergangsraten befinden. Der eigentliche Zellwert ergibt sich aus: $\beta_{r,c} \times P_{r,2009}$.	13
Tabelle 5	Darstellung der Datensituation bei der ökologischen Inferenz im 2x3 Fall, die hellgrau markierten Ränder sind bekannt, die weißen inneren Zellen unbekannt. Es handelt sich um eine kollabierte Version von Tabelle 1.	29
Tabelle 6	Evaluierung der Methoden für die ökologische Inferenz, Übersicht über die weiter betrachteten Modell in Kapitel II.4..	43
Tabelle 7	Evaluierung der Methoden für die ökologische Inferenz, ökologische Regression nach Goodman, Schätzung der Wählerwanderung bei den Bundestagswahlen 2005 und 2009 in München, Unterschied zwischen den Schätzungen bei möglichen Festlegungen, dargestellt ist der AD zwischen den geschätzten Wählerwanderungstabellen.	47
Tabelle 8	Evaluierung der Methoden für die ökologische Inferenz, Thomsens Probit/Logit Modell, Schätzung der Wählerwanderung bei den Bundestagswahlen 2005 und 2009 in München, Unterschied zwischen den Schätzungen für mögliche Wahlen der Referenzparteien, dargestellt ist der AD zwischen den geschätzten Wählerwanderungstabellen. Bei beiden Wahlen wird jeweils die gleiche Referenzpartei genutzt, Park steht für Parks Vorschlag ohne Referenzpartei.	48

Tabelle 9	Evaluierung der Methoden für die ökologische Inferenz, Thomsens Probit/Logit Modell, Schätzung der Wählerwanderung bei den Bundestagswahlen 2005 und 2009 in München, Unterschied zwischen den Schätzungen bei den weiteren Festlegungen in der Software, dargestellt ist der AD zwischen den geschätzten Wählerwanderungstabellen.	49
Tabelle 10	Evaluierung der Methoden für die ökologische Inferenz, Andreadis und Chadjipadelis rekursives Modell, Schätzung der Wählerwanderung bei den Bundestagswahlen 2005 und 2009 in München, Unterschied zwischen den Schätzungen bei möglichen Festlegungen, dargestellt ist der AD zwischen den geschätzten Wählerwanderungstabellen.	50
Tabelle 11	Evaluierung der Methoden für die ökologische Inferenz, Multinomial-Dirichlet Modell, Schätzung der Wählerwanderung bei den Bundestagswahlen 2005 und 2009 in München, Unterschied zwischen den Schätzungen bei unterschiedlicher Parameterwahl, dargestellt ist der AD zwischen den geschätzten Wählerwanderungstabellen.	51
Tabelle 12	Evaluierung der Methoden für die ökologische Inferenz, Multinomial-Normal Modell, Schätzung der Wählerwanderung bei den Bundestagswahlen 2005 und 2009 in München, Unterschied zwischen den Schätzungen bei unterschiedlicher Wahl für die Referenzpartei, dargestellt ist der AD zwischen den geschätzten Wählerwanderungstabellen.	56
Tabelle 13	Evaluierung der Methoden für die ökologische Inferenz, Schätzung der Wählerwanderung bei den Bundestagswahlen 2005 und 2009 in München, Unterschied zwischen den Schätzungen mit den Modellen, dargestellt ist der AD zwischen den geschätzten Wählerwanderungstabellen.	58
Tabelle 14	Evaluierung der Methoden für die ökologische Inferenz, Übersicht der betrachteten vier Datenszenarien in der Simulation zum Methodenvergleich der Verfahren für die ökologischen Inferenz.	64
Tabelle 15	Evaluierung der Methoden für die ökologische Inferenz, Ergebnisse der Simulationsstudie, weiterführende Evaluierung des Multinomial-Normal Modells, Konvergenz der Ketten, dargestellt ist der AD zwischen den Schätzungen, die letzte Spalte gibt die Schätzgüte an.	67
Tabelle 16	Evaluierung der Methoden für die ökologische Inferenz, Ergebnisse der Simulationsstudie, weiterführende Evaluierung des Multinomial-Normal Modells, Abhängigkeit von der Referenzpartei, dargestellt ist der AD zwischen den Schätzungen, die letzte Spalte gibt die Schätzgüte an.	67

Tabelle 17	Darstellung der Datensituation bei der Anwendung der hybriden Modelle bei der Schätzung von Wählerwanderungen, gezeigt werden die Aggregatdaten im Gebiet i , die Ränder entsprechen den Wählern bei der jeweiligen Wahl, die hellgrau markierten Ränder sind bekannt, die weißen inneren Zellen unbekannt.	75
Tabelle 18	Darstellung der Datensituation bei der Anwendung der hybriden Modelle bei der Schätzung von Wählerwanderungen, gezeigt werden die Individualdaten im Gebiet i , die Zahlen entsprechen den befragten Wählern mit der jeweiligen Kombination/Partei, die hellgrau markierten inneren Zellen sind bekannt.	75
Tabelle 19	Datenaufbereitung beim hybriden Modell nach Wakefield in einem Gebiet i mit Individualdaten (links) und Aggregatdaten (rechts), die hellgrau markierten Zellen sind bekannt oder können mit den vorhandenen Informationen berechnet werden. Darstellung angepasst auf den Anwendungsfall Wählerwanderung in Mehrparteiensystemen (RxC Tabellen).	76
Tabelle 20	Darstellung der Datensituation bei der Anwendung der hybriden Modelle bei der Schätzung von Wählerwanderungen, gezeigt werden die Aggregatdaten nach der Datenaufbereitung im Gebiet i , die Ränder entsprechen den Wählern bei der jeweiligen Wahl, die nicht befragt wurden, die hellgrau markierten Ränder sind bekannt, die weißen inneren Zellen unbekannt.	77
Tabelle 21	Detaillierte Evaluierung des (hybriden) Multinomial-Dirichlet Modells, Korrektur der Kreditabilitätsintervalle, Übersicht über die vier betrachteten Korrekturverfahren für die Kreditabilitätsintervalle.	112
Tabelle 22	Darstellung der Datensituation bei der ökologischen Inferenz, erweiterte Darstellung im Vergleich zu Tabelle 1. Nichtwähler werden mit NW abgekürzt, neu hinzugekommen sind die Spalten für die nicht mehr wahlberechtigten Wähler [Tod oder Entzug des Wahlrechts, Alt], weggezogene Wähler (Weg) und die Zeilen für die erstmals Wahlberechtigten (Neu) und die zugezogenen Wahlberechtigten (Zu). Aus inhaltlichen Gründen muss ein Teil der inneren Zellen null sein und ist daher bekannt, während die Randinformation zu den ergänzten Spalten nicht Teil der offiziellen Wahlstatistiken ist.	122
Tabelle 23	Simulation zur Datenaufbereitung bei der ökologischen Inferenz, Übersicht über die vier in der Simulation betrachteten Datenszenarien.	126

Tabelle 24 Münchner Nachwahlbefragung, Kurzzusammenfassung der 134
Response-Raten in den 17 Wahlbezirken mit Nachwahlbefragung bei
der Landtags- bzw. Bundestagswahl 2013. Die Response-Rate ist
dabei die Zahl der ausgefüllten Fragebögen geteilt durch die Zahl der
Wähler in der amtlichen Wahlstatistiken.

Tabelle 25 Schätzung der Wählerwanderung zwischen den Bundestagswahlen 142
2009 und 2013 in München, Evaluierung des hybriden Multinomial-
Dirichlet Modells, Konvergenz der Ketten. Die Schätzung erfolgt
dreimal mit identischer Parametrisierung. Für jedes Modell werden
jeweils drei Teilbereiche der Ketten miteinander verglichen: 1-200
(1), 901-1100 (2) und 1801-2000 (3). Verglichen wird der AD der mit
diesen Teilketten geschätzten Wählerwanderung.

IX. Berücksichtigte Manuskripte und Forschungsergebnisse

Diese Arbeit entstand eingebettet im Rahmen einer langfristigen Forschungs- und Lehrkooperation des Lehrstuhls für empirische Politikforschung, Prof. Paul W. Thurner, und des Statistischen Beratungslabors der LMU, Prof. Helmut Küchenhoff. Im Rahmen der Kooperation wurden z. B. eine gemeinsame interdisziplinäre Seminarreihe und die Münchner Nachwahlbefragung durchgeführt. Mehrere größere Teilprojekte wurden ebenfalls im Rahmen von Abschlussarbeiten oder Praktika bearbeitet.

Die Darlegungen in Kapitel II. basieren zum Teil auf meiner Masterarbeit „Analysen von Wahlergebnissen in Deutschland 1924 – 1933: Räumlich – zeitliche Analyse und ökologische Inferenz“ (2011). Aufgrund der teilweise ähnlichen Inhalte sind bei den allgemeineren Beschreibungen inhaltliche Überschneidungen nicht vermeidbar. Dies betrifft vor allem die Geschichte der ökologischen Analysen (Kapitel II.1.1.), die Vorstellung der ökologischen Regression nach Goodman (Kapitel II.3.1.), sowie die allgemeine Beschreibung von iterativen bzw. adaptiven Verfahren (Kapitel II.3.3.1.) und von hierarchischen Modellen (Kapitel II.3.4.1.). Die Vorstellung der Modelle (Kapitel II.3.) und der präsentierte Modellvergleich / Evaluierung (Kapitel II.4.) wurde bereits vorab zum Teil in:

Klima, André / Thurner, Paul W. / Molnar, Christoph / Schlesinger, Thomas / Küchenhoff, Helmut: Estimation of Voter Transitions Based on Ecological Inference: An Empirical Assessment of Different Approaches. *Advances Statistical Analysis*, 100(2), 2016, S. 133–159.

und

Klima, André / Küchenhoff, Helmut / Thurner, Paul W.: Analysis of voter transition using ecological data: Comparison of different approaches for Munich election data. *Proceedings of the 28th International Workshop on Statistical Modelling* (Palermo, 08.07.-12.07.2013), 1, 2013, S. 203–207.

veröffentlicht. Bei den vorgestellten Betrachtungen und Analysen handelt es sich um die Ergebnisse eines für das Statistische Amt der Landeshauptstadt München durchgeführten Forschungsprojektes. Die Planung der jeweils für die Modelle durchgeführten Betrachtungen, sowie die Planung der Simulation und auch der Szenarien für die Simulation, erfolgte durch mich. Unterstützung bei den Berechnungen erhielt ich von Christoph Molnar (Multinomial-Dirichlet Modell) und Thomas Schlesinger (rekursiver

Algorithmus von Andreadis und Chadjipadelis und Kellermann, Multinomial-Dirichlet Modell). Die in dieser Arbeit genutzten Grafiken zur Darstellung von Wählerwanderungstabellen basieren auf Code, der im Rahmen eines Praktikumsprojektes erstellt wurde und später von Christoph Molnar übernommen wurde. Die Ergebnisse des Forschungsprojektes wurden auch bereits mehrmals auf Tagungen der Städtestatistiker präsentiert. Für diese Arbeit wurde zusätzlich das Multinomial-Normal Modell erneut evaluiert und in der Simulationsstudie mit aufgenommen. Die vorgestellten Überlegungen zum Vorwissen (Kapitel II.3.4.4.) bei hierarchischen Bayesianischen Modellen sind ähnlich den Überlegungen in einem noch nicht veröffentlichten Manuskript:

Klima, André / Schlesinger, Thomas / Küchenhoff, Helmut / Thurner Paul W.: Combining aggregate data and exit-polls for the estimation of voter transitions, unveröffentlichtes Manuskript, 2016.

welches zum großen Teil wiederum auf der von mir mitbetreuten Masterarbeit von Thomas Schlesinger basiert:

Schlesinger, Thomas: Kombination von Aggregat- und Individualdaten bei der Analyse von RxC-Tafeln. Neue Implementierungen in R, Masterarbeit, 2013.

Die Diskussion des zellspezifischen Vorwissens in der Arbeit erfolgte auf mein Bestreben.

Bei der allgemeinen Beschreibung der hybriden Modelle (Kapitel III.1.2.) in Kapitel III. ist erneut eine inhaltliche Überschneidung mit meiner Masterarbeit vorhanden. Weiterhin basieren große Teile des Kapitels (Kapitel III.1.2., III.2.2., III.3. und III.4.1.) inhaltlich auf dem noch nicht veröffentlichten Manuskript (Klima / Schlesinger / Küchenhoff / Thurner, 2016). Das im Kapitel III.2.2. vorgestellte hybride Multinomial-Dirichlet Modell wurde von Thomas Schlesinger im Rahmen seiner Masterarbeit (Schlesinger, 2013) implementiert. Die Betreuung der Arbeit erfolgte auch mit durch mich. Die Grundidee für die Erweiterung entstammt der Forschungs- und Lehrkooperation, die Struktur des Modells wurde im Rahmen der Betreuung diskutiert. Die Aufstellung der notwendigen Likelihoods und die technische Umsetzung erfolgte durch Thomas Schlesinger. Teil der Masterarbeit von Thomas Schlesinger ist eine Evaluierung des hybriden Multinomial-Dirichlet Modells. Auf dieser Evaluierung basiert die in Kapitel III.3. vorgestellte Simulationsstudie. Das Szenario „Exit-Poll mit Bias“ wurde für das oben erwähnte unveröffentlichte Manuskript (Klima / Schlesinger / Küchenhoff / Thurner, 2016) von mir hinzugefügt, die Berücksichtigung des hybriden Multinomial-Normal Modells ist eine Erweiterung für

diese Arbeit. Die Simulation in Kapitel III.4.1. wurde unter meiner Leitung zusammen mit Thomas Schlesinger geplant, die Durchführung der Simulation erfolgte durch ihn. Die Betrachtungen zu den Kreditabilitätsintervallen und den Korrekturverfahren in Kapitel III.4.2. wurden für diese Arbeit durchgeführt.

Die Überlegungen und Darstellungen zur Datenaufbereitung bei der ökologischen Inferenz (Kapitel IV.1.) in Kapitel IV. wurde bereits vorab publiziert (Klima / Thurner / Molnar / Schlesinger / Küchenhoff, 2016, sowie Klima / Küchenhoff / Thurner, 2013). Sie entstammen wie der Modellvergleich / Evaluierung (Kapitel II.4.) dem für das Statistische Amt der Landeshauptstadt München durchgeführten Forschungsprojekt, bei dem mich Christoph Molnar und Thomas Schlesinger bei den Berechnungen unterstützt haben. In dieser Arbeit wird jedoch nur ein Auszug, die Teilergebnisse für das Multinomial-Dirichlet Modell, präsentiert. Bei der Münchner Nachwahlbefragung (Kapitel IV.2.) handelt es sich um ein Gemeinschaftsprojekt, bei dem mehrere Doktoranden und Studierende bei der Planung mit involviert waren. Die Planung und Umsetzung der Ziehung der Wahlbezirke erfolgte durch mich, der Fragebogen wurde von Studierenden der Politikwissenschaft entworfen, aber gemeinsam diskutiert. Die Datenaufbereitung erfolgte zum Teil durch mich. Die allgemeinen Erläuterungen im Kapitel zur Planung der Nachwahlbefragung (Kapitel IV.2.1.) basieren auf einer gemeinsam mit Mirjam Selzer und Laura Mittermeier erstellten Dokumentation. Die gezeigten Ergebnisse aus der Nachwahlbefragung in Kapitel IV.2.2. wurden für diese Arbeit erstellt, es wird jedoch inhaltliche Überschneidungen zu einer Seminararbeit aus dem interdisziplinären Seminar mit diesem Thema geben. Die Berechnung der Wählerwanderung zwischen den Bundestagswahlen 2009 und 2013 in München und Darstellung der Ergebnisse erfolgte für diese Arbeit. Zum Teil wurde dafür Code zur Datenaufbereitung von Thomas Schlesinger mit benutzt.

Der Inhalt des bereits veröffentlichten Artikels (Klima / Thurner / Molnar / Schlesinger / Küchenhoff, 2016), des unveröffentlichten Manuskripts (Klima / Schlesinger / Küchenhoff / Thurner, 2016), der Dokumentation zur Nachwahlbefragung und der Seminararbeit zur Datengüte bei der Nachwahlbefragung sind auch Teil eines Buchmanuskripts (Arbeitstitel: Die Münchner Exit Poll Studie 2013: Ein neuer Ansatz zur Modellierung von Wählerwanderungen) über die Münchner Nachwahlbefragung.

Bei inhaltlichen Übernahmen aus den hier erwähnten Veröffentlichungen und dem unveröffentlichten Manuskript erfolgt weitestgehend keine Referenzierung der entsprechenden Quelle in dieser Arbeit. Eine Zitierung erfolgt jedoch, wenn auf Inhalt verwiesen wird, der maßgeblich von einem Koautor stammt oder auf dessen Arbeit basiert.

X. Anhang

X.1. Beschreibung des Modells von Grofman und Merrill

Bernard Grofman und Samuel Merrill schlugen in ihren 2004 veröffentlichten Artikel drei Modelle für die ökologische Inferenz vor. Alle drei Modelle sind nur für die Schätzung im 2x2 Fall geeignet. Bei dem von Ioannis Andreadis und Theodore Chadjipadelis²²⁹ genutzten Verfahren handelt es sich um die erste von Grofman und Merrill vorgeschlagene Methode.²³⁰

Dieses Modell nutzt eine besondere Eigenschaft bzw. Darstellung des Schätzproblems im 2x2 Fall. Es ist möglich, für jede betrachtete Gebietseinheit eine sogenannte „Tomografielinie“²³¹ zu erstellen. Es handelt sich dabei um einen linearen Zusammenhang zwischen den Übergangswahrscheinlichkeiten im 2x2 Fall, die sich direkt aus dem in Gleichung (2) dargestellten linearen Zusammenhang zwischen beiden Wahlergebnissen ergibt. Es gilt z. B.:

$$\beta_{\overline{CSU2009}, CSU2013, i} = -\frac{P_{CSU, 2009, i}}{P_{CSU, 2009, i}} * \beta_{CSU2009, CSU2013, i} + \frac{P_{CSU, 2013, i}}{P_{CSU, 2009, i}} \quad (\text{A. 1})$$

$\beta_{\overline{CSU2009}, CSU2013, i}$ beschreibt die Übergangswahrscheinlichkeit der Nicht-CSU-Wähler im Jahr 2009 zur CSU im Jahr 2013 im Gebiet i. Diese hängt alleinig von den Wahlergebnissen, $P_{CSU, 2009, i}$ steht für das Ergebnis der CSU im Jahr 2009 im Gebiet i, und der Übergangswahrscheinlichkeit $\beta_{CSU2009, CSU2013, i}$ ab. Ein vergleichbarer Zusammenhang lässt sich für alle Gebiete, aber auch für die betrachtete Gesamtregion aufstellen.²³²

Grofman und Merrill betrachten bei ihrem Modell die Lösung als optimal, bei der der Abstand der Tomografielinien der N Gebietseinheiten minimal zur Tomografielinie der Gesamtregion wird. Dabei wird der Abstand zwischen den Linien immer senkrecht von den Tomografielinien der Gebietseinheiten ausgehend bestimmt und als Abstandsmaß wird die quadrierte Entfernung genutzt. Motivieren lässt sich dieser Ansatz und das Optimierungskriterium über einen Spezialfall bei der ökologischen Inferenz. Bei identischen Übergangswahrscheinlichkeiten würden sich alle Tomografielinien, die gebietsspezifischen und die der Gesamtregion, in einem Punkt schneiden.²³³

229 Siehe Kapitel II.3.3.2. für Details zur diesem Verfahren.

230 Vgl. Andreadis / Chadjipadelis, 2009, S. 211.

231 Vgl. King, 1997, S. 81–82.

232 Vgl. Grofman / Merrill, 2004, S. 129.

233 Vgl. Grofman / Merrill, 2004, S. 129.

X.2. Multinomial-Normal Modell: Konvergenz der Ketten

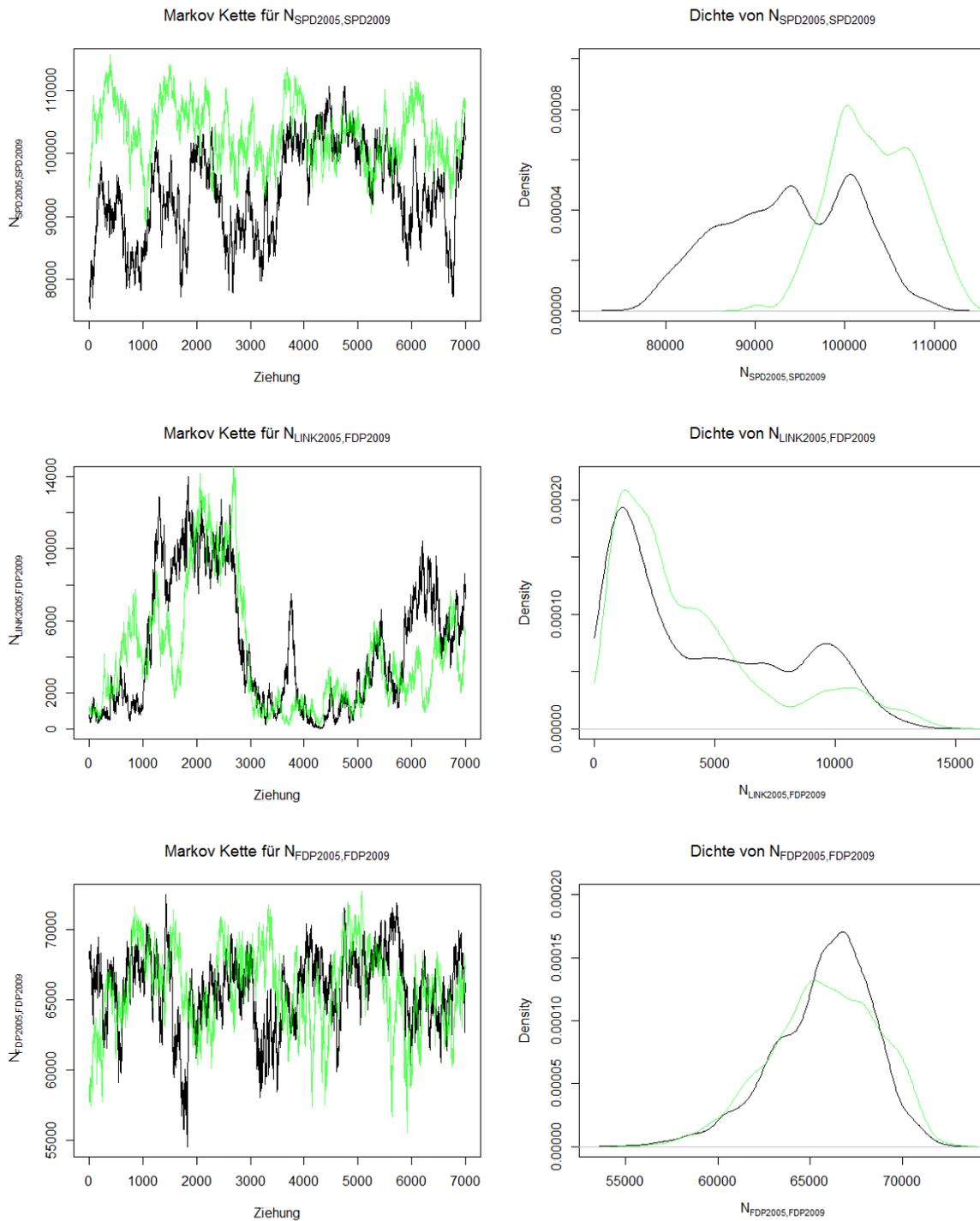


Abbildung 26: Evaluierung der Methoden für die ökologische Inferenz, Ergebnisse der Simulationsstudie, weiterführende Evaluierung des Multinomial-Normal Modells mit dem ersten Datensatz von Sim 2 (Simulationsstudie II.4.3.), Vergleich von zwei Modellschätzungen, Modell 1: schwarz (1), Modell 2: grün (6), die Nummern in Klammern beziehen sich auf Tabelle 15. Die Parameter bei beiden Modellen sind identisch, Burn-In: 3000000, Thinning: 1000, Ziehungen: 7000. Die linken Grafiken zeigen die Markov Ketten für drei ausgewählte Übergänge, die rechten Grafiken die zugehörigen Dichten.

X.3. Multinomial-Dirichlet Modell: Korrekturverfahren Kredititätsintervalle

Eine Kurzübersicht der Korrekturverfahren findet sich in Tabelle 21, ausführlich werden sie in Kapitel III.4.2.2. vorgestellt.

Korrekturverfahren 1

Für jede Bootstrap-Stichprobe wird die geschätzte Wählerwanderung als vergleichbar betrachtet. Da jedoch die Zahl der Wahlberechtigten N_{total} unterschiedlich sein kann, wird bei der Kombination der Ketten das jeweilige $N_{\text{total},k}$ der Bootstrap-Stichprobe berücksichtigt.

$$N''_{r,c,m,k} = \frac{N'_{r,c,m,k}}{N_{\text{total},k}}, \quad \forall r \in \{1, \dots, R\}, c \in \{1, \dots, C\}, m \in \{1, \dots, M\}, k \in \{1, \dots, B\} \quad (\text{A. 2})$$

$N'_{r,c,m,k}$ stellt in Gleichung (A. 2) die m-te Ziehungen aus der Kette des Multinomial-Dirichlet Modells für die Wählerwanderung einer Zelle bei der k-ten Bootstrap-Stichprobe dar. Für die angepasste Version $N''_{r,c,m,k}$ wird der Wert durch die Zahl der Wahlberechtigten $N_{\text{total},k}$ der k-ten Bootstrap-Stichprobe geteilt. Die für die Kredititätsintervalle betrachtete Kette ist die Vereinigung aller B Ketten der mit den B Bootstrap-Stichproben gerechneten Modelle.

$$\underline{N''_{r,c}} = \underline{N''_{r,c,1}} \cup \dots \cup \underline{N''_{r,c,B}}, \quad \forall r \in \{1, \dots, R\}, c \in \{1, \dots, C\} \quad (\text{A. 3})$$

$\underline{N''_{r,c}}$ bezeichnet dabei die vereinigte Gesamtkette und $\underline{N''_{r,c,k}}$ die Ketten der Schätzung einer Bootstrap-Stichproben. Für die Bildung der Kredititätsintervalle werden die Werte aus (A. 3) noch mit der Zahl der Wahlberechtigten N_{total} der Gesamtregion multipliziert.

$$\underline{N'''_{r,c}} = \underline{N''_{r,c}} * N_{\text{total}}, \quad \forall r \in \{1, \dots, R\}, c \in \{1, \dots, C\} \quad (\text{A. 3})$$

Korrekturverfahren 2

Bei diesem Korrekturverfahren wird die betragsmäßige relative Abweichung $B_{r,c,k}$ des Schätzwertes $N'_{r,c,k}$ vom Wert $N_{r,c,k}$ in den simulierten Individualdaten der k-ten Bootstrap-Stichprobe betrachtet.

$$B_{r,c,k} = \left| \frac{N'_{r,c,k} - N_{r,c,k}}{N_{r,c,k}} \right|, \quad \forall r \in \{1, \dots, R\}, c \in \{1, \dots, C\}, k \in \{1, \dots, B\} \quad (\text{A. 4})$$

Für die Korrektur der Ketten wird das Maximum aller relativen Abweichungen $B_{r,c,k}$ verwendet.

$$B_{max} = \underset{(r,c,k)}{argmax}(B_{r,c,k}) \quad (A. 5)$$

Mit diesem Wert wird die Varianz der Ketten der Modellschätzung mit dem Multinomial-Dirichlet Modell angepasst, so dass die Varianz nach der Anpassung die Summe der Varianz vor der Anpassung und der quadrierten maximalen Abweichung ist. Es gilt:

$$Var\left(\frac{N''_{r,c,m} - N'_{r,c}}{N'_{r,c}}\right) = Var\left(\frac{N'_{r,c,m} - N'_{r,c}}{N'_{r,c}}\right) + B_{max}^2, \quad (A. 6)$$

$$\forall r \in \{1, \dots, R\}, c \in \{1, \dots, C\}, m \in \{1, \dots, M\}$$

mit

$$N'_{r,c} = \frac{1}{M} \sum_{m=1}^M N'_{r,c,m} \quad (A. 7)$$

Für die Anpassung (A. 6) erfolgt im ersten Schritt eine Transformation der Ketten. Bei den transformierten Ketten stellt jede Ziehung die relative Abweichung vom jeweiligen Kettenmittelwert, siehe (A. 7), dar. Die Varianz der transformierten Kette wird anschließend vergrößert und jeder Kettenwert entsprechend angepasst. Im letzten Schritt werden die so angepassten Kettenwerte wieder zurück transformiert.

Korrekturverfahren 3

Bei diesem Korrekturverfahren wird die betragsmäßige Abweichung $B_{r,c,k}$ der geschätzten Übergangswahrscheinlichkeiten $\beta'_{r,c,k}$ von den Übergangsraten $\beta_{r,c,k}$ in den simulierten Individualdaten der k -ten Bootstrap-Stichprobe betrachtet.

$$B_{r,c,k} = |\beta'_{r,c,k} - \beta_{r,c,k}|, \quad \forall r \in \{1, \dots, R\}, c \in \{1, \dots, C\}, k \in \{1, \dots, B\} \quad (A. 8)$$

mit

$$\beta'_{r,c,k} = \frac{N'_{r,c,k}}{N_{r,t=1,k}}, \quad \forall r \in \{1, \dots, R\}, c \in \{1, \dots, C\}, k \in \{1, \dots, B\} \quad (A. 9)$$

Für die Korrektur der Ketten wird erneut das Maximum aller relativen Abweichungen $B_{r,c,k}$ verwendet.

$$B_{max} = \underset{(r,c,k)}{argmax}(B_{r,c,k}) \quad (A. 10)$$

Die Varianz der Ketten der Modellschätzung mit dem Multinomial-Dirichlet Modell wird mit diesem Wert angepasst. Dies erfolgt analog zu Korrekturverfahren 2 und den in (A. 6) dargestellten Prinzip, mit dem Unterschied, dass anstatt der relativen Abweichung vom Kettenwert die Abweichungen der Übergangswahrscheinlichkeiten von den mittleren

Übergangswahrscheinlichkeiten der Ketten betrachtet und angepasst werden.

Korrekturverfahren 4

Korrekturverfahren 4 entspricht weitestgehend Korrekturverfahren 3. Erneut wird die betragsmäßige Abweichung $B_{r,c,k}$ der geschätzten Übergangswahrscheinlichkeiten $\beta'_{r,c,k}$ von den Übergangsraten $\beta_{r,c,k}$ in den simulierten Individualdaten der k -ten Bootstrap-Stichprobe betrachtet. Im Unterschied zum dritten Korrekturverfahren werden bei diesem Korrekturverfahren jedoch transformierte Übergangswahrscheinlichkeiten für die Bestimmung des Bias genutzt.

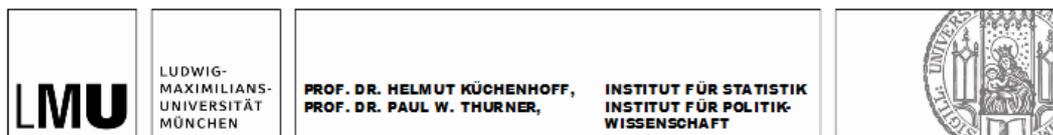
$$B_{r,c,k} = \left| \gamma'_{r,c,k} - \gamma_{r,c,k} \right|, \quad \forall r \in \{1, \dots, R\}, c \in \{1, \dots, C\}, k \in \{1, \dots, B\} \quad (\text{A. 11})$$

mit z. B.:

$$\gamma'_{r,c,k} = \log \left(\frac{\beta'_{r,c,k}}{1 - \beta'_{r,c,k}} \right), \quad \forall r \in \{1, \dots, R\}, c \in \{1, \dots, C\}, k \in \{1, \dots, B\} \quad (\text{A. 12})$$

Die Übergangswahrscheinlichkeiten in (A. 12) werden dabei nach (A. 9) berechnet. Für die Korrektur der Ketten wird erneut das Maximum aller relativen Abweichungen $B_{r,c,k}$, siehe (A. 10), genutzt. Die anschließende Anpassung der Varianz der Ketten erfolgt analog zum vorher beschriebenen Vorgehen unter Berücksichtigung der hier verwendeten Transformation und Betrachtungsebene.

X.4. Fragebogen bei der Münchner Nachwahlbefragung



Landtagswahl Bayern 2013

Wahlstudie

Bitte Zutreffendes ankreuzen!

Alle Angaben sind
vollständig
ANONYM!

1. Welcher Partei haben Sie bei der Landtagswahl heute Ihre Zweitstimme gegeben?

- CSU
 SPD
 FREIE WÄHLER
 GRÜNE
 FDP
 DIE LINKE
 ödp
 PIRATEN
 SONSTIGE, und zwar: _____
 Habe nicht gewählt.
 Habe ungültig gewählt.

2. Wie war das bei der letzten Landtagswahl im September 2008: Haben Sie damals mit Ihrer Zweitstimme eine andere Partei gewählt?

- Ja, und zwar:
 CSU
 SPD
 GRÜNE
 FREIE WÄHLER
 FDP

 DIE LINKE
 ödp
 SONSTIGE, und zwar: _____
 Nein, gleich gewählt.
 War noch nicht wahlberechtigt.
 Habe nicht gewählt.
 War nicht in Bayern wahlberechtigt.
 Kann mich nicht mehr erinnern.

3. Falls Sie sich heute anders entschieden haben als 2008: warum?

4. Ihr Alter

- unter 21
 21 – 24
 25 – 29
 30 – 34
 35 – 39
 40 – 44
 45 – 49
 50 – 59
 60 – 69
 70 oder älter

5. Ihr Geschlecht

- männlich
 weiblich

110

Abbildung 27: Münchner Nachwahlbefragung, Fragebogen bei der Münchner Nachwahlbefragung für die Landtagswahl am 15.09.2013.



Bundestagswahl 2013

Wahlstudie

Bitte Zutreffendes ankreuzen!

Alle Angaben sind
vollständig
ANONYM!

1. Welcher Partei haben Sie bei der Bundestagswahl heute Ihre Zweitstimme gegeben?

- CSU
 SPD
 FDP
 DIE LINKE
 GRÜNE
 PIRATEN
 AfD
 FREIE WÄHLER
 SONSTIGE, und zwar: _____
 Habe nur die Erststimme vergeben.
 Habe ungültig gewählt.

2. Wie war das bei der letzten Bundestagswahl im September 2009: Haben Sie damals mit Ihrer Zweitstimme eine andere Partei gewählt?

- Ja.
 Nein.
 Habe nicht gewählt.
 War nicht wahlberechtigt.
 Weiß nicht mehr.

Falls ja, welche Partei haben Sie 2009 gewählt?

- CSU
 SPD
 FDP
 GRÜNE
 DIE LINKE
 PIRATEN
 SONSTIGE, und zwar: _____

Warum haben Sie sich bei der heutigen Bundestagswahl für eine andere Partei entschieden als 2009?

BITTE UM BLÄTTERN!

Abbildung 28: Münchner Nachwahlbefragung, Fragebogen bei der Münchner Nachwahlbefragung für die Bundestagswahl am 22.09.2013, Vorderseite.

3. Wie haben Sie Ihre **Erststimme** bei der **heutigen Bundestagswahl** vergeben?

Habe mit meiner Erststimme den Kandidaten der Partei _____ gewählt.

4. Noch kurz zur Landtagswahl: Welcher Partei haben Sie bei der **Landtagswahl am vergangenen Sonntag** Ihre **Zweitstimme** gegeben?

- CSU SPD FREIE WÄHLER GRÜNE FDP DIE LINKE
- ödp PIRATEN SONSTIGE, und zwar: _____
- Habe nicht gewählt. Habe ungültig gewählt.

5. Wie war das bei der **Landtagswahl im September 2008**: Welche Partei haben Sie damals mit Ihrer **Zweitstimme** gewählt?

- CSU SPD GRÜNE FREIE WÄHLER FDP DIE LINKE
- ödp SONSTIGE, und zwar: _____ Habe nicht gewählt. Habe ungültig gewählt.
- War noch nicht wahlberechtigt.
- War nicht in Bayern wahlberechtigt.
- Kann mich nicht mehr erinnern.

6. Ihr Alter

- unter 21 21 – 24 25 – 29 30 – 34 35 – 39
- 40 – 44 45 – 49 50 – 59 60 – 69 70 oder älter

7. Ihr Geschlecht

- männlich
- weiblich

110

Abbildung 29: Münchner Nachwahlbefragung, Fragebogen bei der Münchner Nachwahlbefragung für die Bundestagswahl am 22.09.2013, Rückseite.

Eidesstattliche Versicherung

(gemäß Promotionsordnung vom 12.07.11, § 8, Abs. 2 Pkt. .5.)

Hiermit erkläre ich an Eidesstatt, dass die Dissertation von mir selbstständig, ohne unerlaubte Beihilfe angefertigt ist.

München, den 01.08.2016

(André Klima)