

Kontaminationseffekte bei Wahlfunktionen in Mischwahlsystemen

Thomas Bräuninger, Franz Urban Pappi

1. Einleitung

In prototypischen gemischten Wahlsystemen geben die Wähler zwei Stimmen ab, eine für eine geschlossene Parteiliste in einem Wahlkreis mit mehreren Sitzen, gewöhnlich auf regionaler oder nationaler Ebene, und eine für einen Kandidaten in einem Einerwahlkreis auf lokaler Ebene. Mit der Verbreitung gemischter Wahlsysteme durch Übernahme in den Transformationsländern und auch in etablierten Demokratien wie Neuseeland, Italien oder Japan in den 1990er Jahren, schien ein kontrollierter Vergleich der Auswirkungen von Mehrheits- und Verhältniswahl möglich, da in gemischten Systemen beide in ein und derselben Wahl angewendet werden. Dabei zeigte sich allerdings schnell, dass gemischte Wahlsysteme eher eine eigene Gattung sind, da hier das Akteursverhalten in der Mehrheits- und in der Verhältniswahlkomponente vielfältig aufeinander bezogen oder „kontaminiert“ zu sein scheint. Nach Ferrara et al. (2005: 8-9) „liegt eine Kontamination auf der Mikroebene vor, wenn das Verhalten eines Wählers, einer Partei, eines Kandidaten oder eines Gesetzgebers auf einer Wahlebene nachweislich von institutionellen Regeln auf der anderen Ebene beeinflusst wird“.

Wir konzentrieren uns auf das Wählerverhalten in gemischten Systemen mit zwei Stimmen, halten jedoch die Einschränkung auf Spillover-Effekte institutioneller Regeln für zu einschränkend. Den Wählern mögen institutionelle Regeln egal sein, etwa wenn sie einfach aus Protest „Partei X“ unterstützen wollen und beispielsweise ihre Erststimme für den Kandidaten der Partei abgeben, die sie mit der Verhältniswahlstimme wählen – selbst wenn sie den Kandidaten gar nicht kennen. Wir modifizieren deshalb die Definition, indem wir nicht nur auf die Wechselwirkung institutioneller Regeln achten, sondern indem wir Motive für eine gemeinsame Zwei-Güter-Entscheidung postulieren, die zusätzlich zu den Motiven für das Verhalten in jeder Ebene auftreten können. Damit ändert sich auch die Perspektive auf den Gegenstand der Analyse. Wenn Motive für eine gemeinsame Zwei-

Güter-Entscheidung relevant sind, dann gilt es, die simultan verbundenen Entscheidungen eines Wählers in dem aus dem kartesischen Produkt von Kandidaten und Parteilisten zusammen gesetzten Entscheidungsraum zu erklären. Anstatt einer separaten Analyse von Partei- und Kandidatenstimmen betrachten wir die Stimmabgabe für Kandidaten-Partei-Paare. Dabei mag es Motive geben, die sich auf eine bestimmte Kombination von Stimmen beziehen und sich nicht auf die sicher zuvorderst relevanten Motive der Kandidaten- und Parteistimme reduzieren lassen. Kontamination umfasst auch dies.

Unser Beitrag trägt zur laufenden Diskussion über das Wählerverhalten in gemischten Wahlsystemen bei. Wir stellen ein Modell für eine Wahlentscheidung vor, das die Verbindung zwischen den zwei Ebenen berücksichtigt. Wenn die beiden Abstimmungen nicht unabhängig sind und ein „kontrollierter Vergleich“ der beiden Ebenen eines gemischten Systems zur Beurteilung der Wirkung von reinem Mehrheit- gegenüber Verhältniswahl- system nicht funktioniert, stellt sich die interessante Frage, wie genau die beiden Ebenen miteinander verbunden sind, wenn die Bürger ihre zwei Stimmen abgeben. Der Mechanismus dieser Verbindung auf individueller Ebene ist bislang weitgehend ungeklärt. Wir argumentieren, dass die Wähler einer simultanen Entscheidung oder einem Zwei-Güter-Wahlproblem gegenüberstehen, bei dem die Individuen mit $J \times K$ Alternativen konfrontiert sind und nicht nur mit J Kandidaten auf der ersten Ebene und K Parteilisten in der zweiten. Die Individuen haben möglicherweise nicht separable Präferenzen und wägen Gewinne aus der (strategischen oder aufrichtigen) Stimmabgabe in der einen Ebene gegen Verluste aus einer nicht derselben Partei zugutekommenden Stimme in der anderen Ebene ab. Hinweise für nicht separable Präferenzen können dann dazu verwendet werden, zu argumentieren, dass die Stimmen in der Tat nicht unabhängig sind und eine Kontamination vorhanden ist, die über die durch institutionelle Regeln verursachte hinausgeht.

Für die empirische Untersuchung bedeutet dies, dass wir Informationen darüber benötigen, wie die Wähler sowohl die Kandidaten als auch die Parteien bewerten – und nicht nur die Parteien. Während eine breite empirische Literatur eine Kontamination von Erst- und Zweitstimme bestätigt, beruhen die meisten Belege auf aggregierten Daten, die zwar Informationen über die ideologische Ausrichtung der Parteien liefern, aber nicht die der Kandidaten. In einem ersten Schritt verwenden wir daher eine bayesianische Entfaltung von Daten aus einer Nachwahlbefragung in zwei Wahlkreisen bei der Bundestagswahl 2013, um Schätzungen der Wähler-

wahrnehmungen von Parteien und Kandidaten zu erhalten. Insbesondere schätzen wir Positionen und Valenzen von Parteien und Kandidaten sowie die Wählerpositionen in einem gemeinsamen Raum.

Diese Informationen nutzen wir dann, um unser theoretisches Modell der Wahlentscheidung im deutschen Mischwahlsystem zu kalibrieren. Wir verwenden eine Spezifikation eines gemischten Logit-Modells, das es uns ermöglicht, Interdependenzen oder Substitutionsmuster zwischen Alternativen zu untersuchen. Die empirische Schätzung ergibt vier wichtige Ergebnisse. Erstens hat die geschätzte ideologische Nähe der Wähler zu Kandidaten und Parteien den erwarteten positiven Effekt, das Gleiche gilt für die Valenz und die Parteiidentifikation. Zweitens gibt es deutliche Hinweise darauf, dass Wähler strategisch in Bezug auf die Erststimme wählen. Drittens gibt es eine offensichtliche Tendenz der Individuen, über die zwei Ebenen hinweg kongruent zu wählen. Und viertens, andere, denkbare und durchaus vernünftige Substitutionsmuster sind von geringer Bedeutung.

2. Literatur

Wir interessieren uns dafür, wie Wähler ihre Stimme in Mischwahlsystemen abgeben, ein Forschungsgebiet, in dem wir glücklicherweise auf eine umfangreiche Literatur zurückgreifen können. In den letzten zwanzig Jahren hat die Untersuchung gemischter Wahlsysteme an Zahl und Bedeutung zugenommen, da sich diese Systeme als dritte, eigenständige Art neben reinen Verhältniswahlsystemen und reinen Mehrheitswahlsystemen etabliert haben.

Die erste Welle von Studien betrachtete gemischte Systeme als die Kombination von zwei, weitgehend unabhängigen Wahlmodi, die die Möglichkeit für kontrollierte Vergleiche von reinen Proporz- und Majorzsystemen ermöglichen sollen, was in länderübergreifenden Studien unerreichbar zu sein scheint. Stratmann und Baur (2002) beispielsweise untersuchten die Parlamentsarbeit von Bundestagsabgeordneten und stellten fest, dass direkt gewählte Abgeordnete häufiger Ausschussvorsitze in lokal ausgerichteten Politikbereichen erlangen als über Listen gewählte Abgeordnete. Moser und Scheiner (2004) betrachten die Wahlergebnisse in 15 Ländern mit Mischwahlsystemen und stellen fest, dass Verhältnis- und Mehrheitswahlsysteme „tendenziell die erwarteten Effekte approximieren, selbst wenn sie in gemischten Systemen kombiniert werden“. Andere Studien haben vielfältig Belege dafür geliefert, dass Mehrheitsregeln in gemischten Systemen

zu einer „strategischen Stimmabgabe“ führen, wenn man davon ausgeht, dass Verhältnisstimmen aufrichtig sind und eine getrennte Stimmabgabe vorgesehen ist; Wähler würden dann wie erwartet dazu neigen, ihre Kandidatenstimme nicht an wenig aussichtsreiche Kandidaten im Wahlkreis zu verschwenden (Bawn 1993; Cox 1997).

In der neueren Literatur wird die Annahme der Durchführbarkeit kontrollierter Vergleiche zunehmend skeptisch beurteilt (Ferrara et al. 2005). Bei der Erklärung des Wahlverhaltens im deutschen Mischwahlsystem haben sich Autoren allenthalben auf die Zweitstimme konzentriert, da nach der deutschen Kompensationsregel (Anrechnung der Wahlkreismandate auf die einer Partei nach der Verhältniswahl zustehenden Quote) die Zweitstimme entscheidend für die Zusammensetzung des Bundestages ist. Ausnahmen sind Analysen der strategischen Kandidatenwahl aufgrund des Duverger'schen Stimmenvergeudungsmotivs (Schoen 1999; Pappi/Thurner 2002; Gschwend 2007; Herrmann/Pappi 2008; Herrmann 2012), Arbeiten zu persönlichen Kampagnen von Wahlkreiskandidaten und deren Auswirkung auf Kandidaten- und Parteilistenstimmen (Gschwend/Zittel 2012) und experimentelle Untersuchungen zu abweichendem Wählen bei Vorliegen einer weiteren Stimme (Huber 2012). Ein innovativer Ansatz zur Entflechtung der kausalen Beziehungen zwischen den zwei Stimmabgaben wird von Rheault et al. (2020) vorgestellt. Diese Autoren fragen nach der zuerst gebildeten Verhaltensabsicht, der für einen Kandidaten oder der für eine Parteiliste. Die zuerst gebildete Absicht kann die andere Absicht beeinflussen, aber nicht umgekehrt. Da 46 Prozent der Befragten eine simultane Entscheidung angeben, scheint unser Ansatz der simultan verbundenen Zwei-Güter-Entscheidung vielversprechender zu sein.

Ein erster Indikator dafür, dass die Mehrheitskomponente von gemischten Systemen anders funktioniert als ein reines Mehrheitswahlsystem, ist die größere Anzahl von Kandidaten in den Wahlkreisen. Sie deutet darauf hin, dass diese Kandidaturen weniger durch rationale Erwartungen motiviert sind, das Mandat zu gewinnen (vgl. Herron/Nishikawa 2001; Manow 2010: 68). Vielmehr könnten die Parteien eine solche Strategie in der Hoffnung verfolgen, dass die Wähler eine aufrichtige Parteistimme mit einer Stimme für den Kandidaten derselben Partei verbinden, ungeachtet der Erfolgschancen des Kandidaten (Ferrara et al. 2005: 35). Zudem könnten sich die Wähler gezwungen fühlen, in beiden Wahlentscheidungen konsistent zu sein, indem sie für den Direktkandidaten der in der Verhältniswahlkomponente bevorzugten Partei stimmen. Eine solche Strategie würde den

psychologischen Effekt abschwächen (Duverger 1954), d.h. die strategische Erststimmabgabe zur Vermeidung einer Stimmenverschwendug, während die mechanische Wirkung dieselbe ist. Deutschland ist ein gutes Beispiel (Ferrara/Herron 2005).¹ Einerseits ist der nach Duverger zu erwartende Trend zu einem Zweiparteiensystem in der Mehrheitswahlkomponente stark ausgeprägt; er war lange Zeit sogar stärker als in Großbritannien (Shikano 2009). Andererseits stellen alle Bundestagsparteien in fast allen Wahlkreisen Kandidaten auf. Es könnte durchaus sein, dass die Anhänger kleinerer Parteien in kompensatorischen Mischwahlsystemen großzügig mit ihrer Kandidatenstimme umgehen – da die Anzahl der Sitze, die eine Partei erhält, durch den Verhältnsstimmenanteil bestimmt wird –, zumindest großzügiger als in Grabenwahlsystemen, wo sowohl der Mehrheits- als auch der Verhältniswahlerfolg die Sitzzahl bestimmen.

Umgekehrt wird auch die Proporzkomponente in einem gemischten Wahlsystem anders funktionieren als in einem reinen Verhältniswahlsystem. Herron und Nishikawa (2001) beispielsweise zeigen für Japan und Russland, wie die schiere Existenz von Direktkandidaten den Stimmensatzteil der Partei in der Proporzkomponente erhöht (s. auch Cox/Schoppa 2002), Hainmüller et al. (2006) finden für Deutschland einen Amtsinhaberbonus des Wahlkreiskandidaten auch im Zweitstimmen-Ergebnis. Die Parteien hätten somit einen Anreiz, auch wenig aussichtsreiche Direktkandidaten aufzustellen.

Das Problem, mit dem wir uns in diesem Beitrag befassen, ist spezifischer als auch umfassender. Es ist spezifischer, da wir uns für das Verhalten der Wähler und nicht für die Strategien der Parteien interessieren. Es ist breiter angelegt, da wir ein allgemeines Kalkül der Stimmabgabe für gemischte Systeme mit zwei Stimmen konstruieren. Es beginnt mit Bewertungen von Parteien als unmittelbares Motiv für die Zweitstimme und mit Bewertungen der Direktkandidaten als unmittelbares Motiv für die Erststimme und bringt schließlich Motive für spezielle Kombinationen von Verhältnis- und Mehrheitswahl-Stimmen ein. Ob die bloße Platzierung eines Kandidaten in der Mehrheitskomponente den Erfolg der Partei in

1 Ferrara und Herron (2005) argumentieren, dass kleine Parteien Kandidaten aufstellen können, weil sie erwarten, dass die Kandidaten die Parteikampagne vorantreiben. Nicht aussichtsreiche Kandidaten haben ihrerseits Anreize zu kandidieren, wenn eine Kandidatur ihre Chancen auf einen vorderen Platz auf der geschlossenen Parteiliste erhöht. Für Deutschland weisen Schüttmeyer und Sturm (2005) darauf hin, dass eine Kandidatur im Wahlkreis oft eine informelle Voraussetzung für die Aufnahme in eine Parteiliste ist.

der Proporzkomponente erhöht, ist eine andere Frage, die wir hier nicht behandeln. Die Stimmabgabe im deutschen Mischwahlssystem bietet jedenfalls ein ideales Testfeld für unsere gemeinsame Analyse der zwei Stimmen, da alle großen Parteien in fast allen Wahlkreisen Kandidaten aufstellen. Da wir uns aus Gründen der Datenverfügbarkeit ohnehin auf die fünf Bundestagsparteien im Wahlkampf 2013 beschränken müssen, gibt es auch keine Unterschiede in Bezug auf die Platzierung oder Nicht-Platzierung von Kandidaten.²

3. Theorie

Wir argumentieren, dass die Wähler bei der Stimmabgabe vor einer simultan verbundenen Entscheidung stehen (oder es so empfinden), wenn sie ihre Erststimme für einen Kandidaten im Wahlkreis und ihre Zweitstimme für eine Parteiliste abgeben. Dies wird nicht zuletzt durch die besondere Form des Stimmzettels befördert, auf dem beide Stimmen abzugeben sind (und nicht auf getrennten Zetteln) und die Direktkandidaten und deren zugehörige Parteilisten in einer Zeile angeordnet sind. Wir betrachten daher die Wahlsituation als eine Zwei-Güter-Entscheidung, bei der die Individuen Präferenzen über die Alternativen im kartesischen Produkt der beiden Entscheidungsoptionen haben.

Wir betrachten einen allgemeinen Präferenztyp, insbesondere beschränken wir uns nicht auf separable Präferenzen. Nicht-separable Präferenzen sind dadurch gekennzeichnet, dass die optimale Entscheidung auf einer Ebene, zum Beispiel für einen Kandidaten, von der Entscheidung auf der anderen Ebene, d. h. der Partei, abhängt. Im Prinzip können nicht-separable Präferenzen viele Formen annehmen. Sie können negative Komplementaritäten haben, wie z. B. Budget-Allokationspräferenzen (wobei Ausgaben in einem Bereich gegen Ausgaben in einem anderen Bereich abgewogen werden), oder positive Komplementaritäten wie bei den öffentlichen Finanzen (wenn umsichtige Finanzminister Gesamtausgaben bevorzugen, die mit den Steuereinnahmen übereinstimmen, gegenüber nichtübereinstimmenden Ausgaben). In unserem Kontext wäre eine nicht separable Präferenzordnung über Paare (C, P) von Kandidaten und Parteien z. B.

² Kurella (2016) liefert eine überzeugende Analyse des fehlenden Zweitstimmenverstärkungseffekts von Wahlkreiskandidaturen der neu gegründeten Partei „Alternative für Deutschland“ im Jahr 2013.

gekennzeichnet durch eine erste Präferenz für (C_1, P_2) und eine zweite Präferenz für (C_2, P_1) , während die Paare (C_1, P_1) und (C_2, P_2) nachrangig wären (vgl. Kreps 2013: 40).

Während für simultan verbundene Wahlentscheidungen verschiedene Muster der Nicht-Separabilität denkbar sind, betrachten wir einen Typus, der auf konsistentem Verhalten beruht und den wir für den relevantesten halten. In praktisch allen Mischwahlsystemen – ebenso wie in reinen Mehrheitswahlsystemen – treten fast alle Wahlkreiskandidaten nicht als Unabhängige, sondern als Parteimitglieder an. Daher ist die Parteizugehörigkeit des Kandidaten das Bindeglied zwischen den beiden Komponenten, was zu einer möglichen Kontamination der beiden Entscheidungen führt: Die Wähler könnten einfach bestrebt sein, sich bei beiden Entscheidungsoptionen konsistent zu verhalten. Wir können uns zwar verschiedene Mechanismen vorstellen, die zu dieser Art von Motivation führen können, wir verwenden aber hier den Begriff „konsistentes Verhalten“ als einen allgemeinen Begriff, der dazu dienen soll, Gewinne aus (strategischer oder ehrlicher) Stimmabgabe auf einer Ebene mit Verlusten durch inkonsistentes Verhalten auf der anderen Ebene abzuwägen.³

Wir nehmen daher an, dass die Präferenzen der Wähler über die Wahlalternativen im kartesischen Produkt aus der Menge der Kandidaten und der Menge der Parteilisten durch eine Nutzenfunktion der folgenden Form dargestellt werden können:

$$U(x, y) = u_1(x) + u_2(y) + u_3(x, y)$$

wobei x und y Kandidaten- bzw. Parteilistenoptionen sind. Entscheidend ist nun, dass U möglicherweise eine nicht separable Nutzenfunktion ist, so dass $u_1(x)$ der Teil des Nutzens ist, der nur mit dem Kandidaten variiert, und $u_2(y)$ der Teil des Nutzens ist, der ausschließlich mit Parteilisten des Paares (x, y) variiert. Daher ist $u_1(x)$ nicht der Nutzen von Kandidat x und der Kandidat, der $u_1(x)$ maximiert, ist nicht notwendig der bevorzugte Kandidat. Sind die Präferenzen nicht-separabel, hat der Begriff „bevorzugter Kandidat“ keine richtige Bedeutung, da der Kandidat, bei dem U ein

³ Der Begriff Inkonsistenz bezieht sich ausschließlich auf die Nichtübereinstimmung der Stimmabgabe bei zwei Wahlgängen in Bezug auf die Parteizugehörigkeit. Er deutet nicht auf irgendeine Art von Irrationalität des Verhaltens hin. Im Gegensatz dazu geht unser Ansatz davon aus, dass solche separablen Präferenzen für Kandidaten und Parteien nicht geeignet sind, um individuelle Motivationen abzubilden; idealerweise sollten die beiden Arten von Präferenzen gemeinsam für die $J \times K$ -Paare gemessen werden (vgl. Luce/Tuckey 1964).

Maximum annimmt, mit der Parteiliste y variiert; und umgekehrt gilt dies für die bevorzugte Parteiliste.

Welche Faktoren bestimmen nun wesentlich die Wählerpräferenzen über die Kandidaten-Partei-Paare? Das ist eine konzeptionell-empirische Frage, bei der wir uns auf die umfangreiche Literatur zum Wahlverhalten stützen und annehmen, dass die Wähler einerseits an Politikergebnissen interessiert sind, andererseits auch dem Wahlakt selbst eine Bedeutung zuschreiben, also dass sowohl instrumentelle als auch expressive Motivationen vorliegen können.

3.1 Instrumentelle Wählermotivation

Der instrumentelle Nutzen beruht zunächst auf der Bereitstellung von öffentlichen Gütern, die für den Wähler wertvoll sind, durch die Gewählten. Wir folgen den Standardargumenten aus der Literatur und stellen uns die öffentliche Politik als Punkte in einem n -dimensionalen Themenraum vor und Wählerpräferenzen als Euklidische Präferenzen in diesem Raum. Dies läuft auf das Standardmodell der räumlichen Wahl hinaus, bei dem die Wähler durch ihren Idealpunkt im mehrdimensionalen Politikraum repräsentiert werden und Alternativen, die näher an ihrem Idealpunkt liegen, denen vorziehen, die weiter entfernt sind.

Eine zweite Quelle des instrumentellen Nutzens ist die Valenz von Kandidaten und Parteien. Die Valenz bezieht sich auf die Beurteilung der Wähler über die Qualität oder Kompetenz eines Kandidaten oder einer Partei (Stokes 1992). Da alle Wähler wahrscheinlich Kandidaten mit hoher gegenüber solchen mit geringer Kompetenz bevorzugen, gilt statt der Nähe-Logik mit unterschiedlichen Idealpunkten der Wähler eine Mehr-ist-besser-Logik mit demselben Wähler-Idealpunkt bei den Valenzwerten. Oder anders ausgedrückt: die instrumentelle Bewertung von Kandidaten und Parteien durch die Wähler drückt sich in zwei Arten von öffentlichen Gütern aus: Güter, bei denen sich die Idealpunkte der Wähler unterscheiden (dargestellt durch räumliche Präferenzen über Politikpositionen) und Güter (oder Übel), bei denen sich die Idealpunkte der Wähler nicht unterscheiden (hier dargestellt durch Valenz). Das Konzept der Valenz wurde in neueren, sowohl formalen als auch empirischen Studien zum Wahlverhalten und zum Parteienwettbewerb vielfach verwendet (Anscombe/Snyder 2000; Groseclose 2001; Schofield/Sened 2006). Wir folgen dieser Literatur und

fügen einen Valenzterm zur Nähekomponente des instrumentellen Nutzens hinzu.

Dabei ist zu bedenken, dass von Kandidaten und Parteien nicht erwartet werden kann, dass sie die gleichen öffentlichen Güter oder das gleiche öffentliche Gut in gleichem Maße anbieten. Erstens können die Bewertungen von Parteien und ihren Kandidaten sehr unterschiedlich ausfallen. Während eine erfolgreiche Regierungspartei mit einem beliebten und kompetenten Parteivorsitzenden einen hohen Valenzvorteil bei den Wählern haben mag, kann ein neuer Herausforderer der gleichen Partei in einem Wahlkreis, der gegen den erfahrenen lokalen Amtsinhaber der nationalen Hauptkonkurrenzpartei antritt, schlecht abschneiden. Zweitens können Wahlkreiskandidaten und ihre Parteilisten auch aufgrund der Plattformen oder Positionen, die sie zu einem Thema vertreten, unterschiedlich bewertet werden. Man könnte argumentieren, dass es gerade die Aufgabe gemischter Wahlsysteme ist sicherzustellen, dass verschiedene lokale Forderungen in einer geografisch heterogenen Bevölkerung aufgegriffen werden und in den politischen Prozess einfließen. Wenn lokale Kandidaten nur Klone ihrer Parteien wären, würde einfache Verhältniswahl ausreichen. Wir ziehen daher die Möglichkeit in Betracht, dass sich Kandidaten und ihre Listenparteien in den Positionen oder der Politik, die sie anbieten, unterscheiden.⁴

Schließlich gilt es, die strategische Reaktion der Wähler zu bedenken, die sich aus dem mechanischen Effekt des Wahlsystems ergibt (Duverger 1954). Da im Allgemeinen nicht alle Kandidaten oder Parteien mit gleicher Wahrscheinlichkeit gewählt werden und bei der Bereitstellung öffentlicher Güter effektiv sind, stellt sich für den Wähler nicht nur die Frage, wer der Kandidat oder die Partei ist, die, wenn überhaupt, die vorteilhafteste Politik bietet. Es geht auch um die Frage, welcher der Kandidaten oder Parteien überhaupt die Macht oder Chance hat, diese Güter bereitzustellen. Infolgedessen können die Wähler strategisch in dem Sinne wählen, dass sie nicht für ihren bevorzugten Kandidaten oder ihre bevorzugte Partei stimmen, sondern für die Alternative, die mit größerer Wahrscheinlichkeit das

4 Eine etwas andere Argumentation würde von der Annahme ausgehen, dass Kandidaten und Listenparteien nach unterschiedlichen Kriterien bewertet werden, da von ihnen erwartet wird, dass sie unterschiedliche Arten von öffentlichen Gütern liefern, nämlich lokal versus national zielgerichtete Güter. Dies scheint am plausibelsten in kompensatorischen gemischten Wahlsystemen wie in Deutschland, wo das Ergebnis der Wahl auf lokaler Ebene nur in Ausnahmefällen einen Einfluss auf die Gesamtverteilung der Sitze hat.

gewünschte Ergebnis liefern wird. Die wichtigste Quelle für Unterschiede in der Effektivität der eigenen Stimmabgabe wird durch institutionelle Regeln und in gewissem Maße auch durch den politischen Prozess bestimmt.

So sind Koalitionsregierungen in gemischten Wahlsystemen wahrscheinlich, und wenn also die Politikgestaltung von der Regierung dominiert wird, ist eine strategische Koalitionswahl in der Komponente der Verhältniswahl eine sinnvolle Strategie zur Beeinflussung der politischen Ergebnisse (Duch et al. 2010; Herrmann/Pappi 2008). Strategisches Koalitionswählen ist jedoch ziemlich komplex, da es mit der Unsicherheit zurechtkommen muss, die mit dem wahrscheinlichen Wahlergebnis und der Unsicherheit der Koalitionsverhandlungen einhergeht. Demgegenüber ist eine strategische Stimmabgabe, mit der eine „verschwendete Stimme“ für einen weniger aussichtsreichen Direktkandidaten vermieden werden soll, weniger anspruchsvoll. Die Wähler können erkennen, dass ein Kandidat eine geringe Wahrscheinlichkeit hat, gewählt zu werden, und daher den instrumentellen Nutzen, der mit dem Kandidaten einhergeht, außer Acht lassen. In diesem Papier betrachten wir die letztere Art der strategischen Wahl, nämlich die strategische Wahl in der Mehrheitskomponente, die mit dem Motiv der Vermeidung einer Stimmverschwendungsverbunden ist. Wir lassen die Koalitionswahl außer Acht (die für uns hier nicht von Interesse ist) und wir beachten auch nicht die strategische Stimmabgabe vom Typ verschwendete Stimme bei der Verhältniswahl, da die Wahrscheinlichkeit der hier betrachteten Bundestagsparteien, in der Verhältniswahlkomponente an der Fünf-Prozent-Hürde zu scheitern, gering ist und damit nicht ausreichend Variation bietet.⁵

3.2 Expressive Wählermotivation

Während sich die instrumentelle Bewertung der Wahlobjekte auf den erwarteten Nutzen im Zusammenhang mit dem Wahlausgang konzentriert, ist die expressive Motivation für das Verhalten mit dem Akt der Stimmab-

5 In der folgenden empirischen Analyse betrachten wir Kandidaten und Listen der fünf Bundestagsparteien. Wir haben keine Informationen über die Bewertungen der Wähler und die Stimmabgabe für eine der kleineren Parteien, von denen bei der Bundestagswahl 2013 zu erwarten war, dass sie die Schwelle (mehr oder weniger deutlich) verfehlten würden. Letztendlich hat die FDP mit einem in Umfragen vor der Wahl erwarteten Stimmenanteil von 5,5 Prozent die Hürde schließlich verfehlt. Doch von allen anderen Parteien wurde erwartet, dass sie die Hürde mit ziemlicher Sicherheit nehmen würden.

gabe selbst verbunden. Der Grundgedanke dabei ist, dass ein bestimmtes Verhalten die Überzeugungen, Werte, Identität oder Ideologie des Einzelnen ausdrückt und dass dieser Ausdruck wertvoll ist, unabhängig von der Tatsache, dass zum Beispiel die Stimmabgabe für einen wenig aussichtsreichen Kandidaten in einem Mehrheitswettbewerb aus instrumenteller Sicht eine verschwendete Stimme ist (für einen ausgezeichneten Überblick, siehe Hamlin/Jennings 2011).

Eine Reihe von Motivationslagen kann expressives Wählen hervorrufen. Riker und Ordeshook (1968) verwenden die Idee einer Bürgerpflicht als eine Form der expressiven Stimmabgabe, um das Paradoxon der Stimmabgabe trotz minimalen Einflusses des einzelnen Wählers auf das Endergebnis zu lösen. Schuessler (2000) betrachtet die Stimmabgabe als einen expressiven Akt, mit dem sich der Wähler mit anderen Individuen, die für die gleiche Partei stimmen, identifiziert; Brennan und Lomasky (2008) betrachten expressive Wahl als eine moralische Wahl. Wir verwenden hier eine engere Definition des expressiven Wählens, das mit dem Konzept der Parteianhängerschaft verbunden ist, und betrachten den expressiven Nutzen als einen Vorteil, der sich aus der Bestätigung der Identifizierung mit Parteien oder Kandidaten ergibt (Brennan/Hamlin 1998).

3.3 Wählerpräferenzen über Kandidaten-Parteien-Paare

Insgesamt verwenden wir eine Nutzenfunktion für Kandidat j_1 und Partei j_2 der folgenden Form:

$$\begin{aligned} U_i(j_1, j_2) &= u_{1i}(j_1) + u_{2i}(j_2) + u_{3i}(j_1, j_2) \\ u_{1i}(j_1) &= (\beta_1 + \beta_2 p_{j_1}) v_{j_1} + (\beta_3 + \beta_4 p_{j_1}) \|x_i - x_{j_1}\| + \beta_5 t_{ij_1} \\ u_{2i}(j_2) &= \beta_6 v_{j_2} + \beta_7 \|x_i - x_{j_2}\| + \beta_8 t_{ij_2} \\ u_{3i}(j_1, j_2) &= \beta_9 C_{j_1 j_2} \end{aligned}$$

wobei v_j der Valenzterm des Kandidaten oder der Partei j ist, x_i der Idealpunkt des Wählers i , x_j die Plattform von Kandidat oder Partei j und p_{j1} die Wahrscheinlichkeit, dass Kandidat j_1 aussichtsreich kandidiert. C ist eine Indikatorfunktion, die den Wert eins annimmt, wenn die Wahl „konsistent“ ist, also $j_1 = j_2$ gilt, und ansonsten Null. Schließlich nimmt t_{ij} den Wert eins an, wenn die Person i eine Parteiidentifikation mit j hat, andernfalls hat es den Wert Null.

Wir können dann die Nicht-Separabilität der Nutzenfunktion U verwenden, um auf eine Kontamination zwischen beiden Komponenten auf der Mikroebene der Wähler zu testen. In den Begriffen von Ferrara et al. (2005) ist Kontamination vorhanden, wenn das Verhalten eines Wählers in einer Komponente nachweislich von den institutionellen Regeln der anderen Komponente beeinflusst wird. Wie bereits erwähnt, modifizieren wir diese Definition, indem wir zulassen, dass nicht nur institutionelle Regeln auf die andere Komponente übergreifen, sondern auch indem wir erstens Motive für jede Ebene postulieren und zweitens Motive für eine spezielle Kombination von Kandidaten- und Parteilistenoptionen vorsehen. Unter der Annahme, dass unser Modell die Wählerpräferenzen einigermaßen gut abbildet, liegt eine Kontamination vor, wenn die Nutzenfunktionen nicht in x und y separierbar sind. Wenn hingegen die empirische Kalibrierung der Nutzenfunktionen nahelegt, dass die Präferenzen separabel sind, gibt es keine Kontamination. Mit der oben formulierten Nutzenfunktion U für die gleichzeitige Abgabe von Erst- und Zweitstimme ist dieser Test einfach. Wenn die empirische Schätzung anzeigt, dass der Nutzen aus konsistentem Verhalten, $u_3(x, y)$, Null ist, ist U separabel. Dies ist eine Implikation von Debreus (1959) *double cancellation*-Bedingung.⁶

4. Statistisches Modell

Wir versuchen, die Wahlentscheidung im deutschen Mischwahlsystem auf der Grundlage der Wählerbewertungen der Direktkandidaten und der fünf Bundestagsparteien zu analysieren. Die Menge der Wahloptionen ist definiert durch das Kreuzprodukt der Alternativen in beiden Komponenten und umfasst somit 25 Alternativen. Diese Wahlmöglichkeiten sind jedoch auf zwei nicht-hierarchischen Ebenen geclustert: Je fünf Optionen beziehen sich auf denselben Kandidaten und je fünf Optionen beziehen sich auf dieselbe Partei. Es ist daher zweifelhaft, dass die Annahme der Unabhängigkeit von irrelevanten Alternativen (IIA) des multinomialen Logit-Modells hier erfüllt ist.

⁶ Die *double cancellation*-Bedingung besagt, dass wenn für drei beliebige Bündel (x_1, x_2) , (y_1, y_2) , (z_1, z_2) und eine Präferenzrelation R $(x_1, x_2) R (y_1, y_2)$ und $(y_1, z_2) R (z_1, x_2)$ gilt, dann muss auch $(x_1, z_2) R (z_1, y_2)$ gelten. Debreu hat gezeigt, dass bei zwei Gütern die *double cancellation*-Bedingung genau dann erfüllt ist, wenn die Präferenzen durch eine additiv separable Nutzenfunktion darstellbar sind (siehe auch Luce/Tuckey 1964).

Wir verwenden daher ein gemischtes Logit-Modell, das fixe, d. h. nicht zufällige Effekte für die Kovariaten des theoretischen Modells und zufällige Null-Mittelwert-Effekte für eine Reihe von Indikatorvariablen vorsieht, die die Gruppierung der Alternativen in den beiden Komponenten erfassen. Genauer schätzen wir das folgende Modell:

$$\begin{aligned} U_i(j_1, j_2) = & (\beta_1 + \beta_2 p_{j_1}) v_{j_1} + (\beta_3 + \beta_4 p_{j_1}) \|x_i - x_{j_1}\| + \beta_5 t_{ij_1} \\ & + \beta_6 v_{j_2} + \beta_7 \|x_i - x_{j_2}\| + \beta_8 t_{ij_2} + \beta_9 C_{j_1 j_2} \\ & + \sum_{k=1}^5 \mu_{ik} I_{kj_1} + \sum_{k=1}^5 \tilde{\mu}_{ik} I_{kj_2} + \epsilon_{ij_1 j_2}. \end{aligned} \quad (1)$$

Dabei bezeichnet I eine Indikatorvariable, die den Wert 1 annimmt, wenn Kandidat j_1 (Partei j_2) der k te Kandidat (Partei) ist und ansonsten den Wert Null hat. Wir nehmen an, dass die Fehlerkomponenten μ_{ik} und $\tilde{\mu}_{ik}$ identisch und unabhängig normalverteilt verteilt sind mit Erwartungswert 0 und Varianzen σ_k bzw. $\tilde{\sigma}_k$. Wie im herkömmlichen Logit-Modell sind die $\epsilon_{ij_1 j_2}$ unabhängig und identisch Typ I extremwertverteilt.⁷

Es ist klar, dass dieses Substitutionsmuster immer noch restriktiv ist. In einem multinomialen Probit-Modell mit korrelierten Fehlern könnten wir die vollständige Kovarianzmatrix der normalisierten Fehlerdifferenzen schätzen, in unserem Fall wären dies aber $(J-1)J/2-1 = 299$ Parameter. Dies ist mit der vorliegenden Datenmenge nicht möglich und selbst wenn, so wäre aus den Fehlerdifferenzen schwerlich etwas über Substitutionsmuster zu erfahren. Unsere zweistufige Gruppierung mit zehn Parametern ermöglicht jedoch die Untersuchung der wahrscheinlichsten Korrelations- und Substitutionsmuster.

Um dies zu erkennen, betrachten wir die Kovarianz der Zufallsvariablen $\eta = \mu + \tilde{\mu} + \epsilon$ zweier Alternativen (j_1, j_2) und (j_1, j'_2) , die sich auf denselben Kandidaten, aber unterschiedliche Parteien beziehen:

$$\text{Cov}(\eta_{j_1 j_2}, \eta_{j_1 j'_2}) = E[(\mu_{j_1} + \tilde{\mu}_{j_2} + \epsilon_{j_1 j_2})(\mu_{j_1} + \tilde{\mu}_{j'_2} + \epsilon_{j_1 j'_2})] = \sigma_{j_1}$$

wobei wir uns die Tatsache zunutze machen, dass der Erwartungswert der angenommenen Extremwertverteilung gleich Null ist. Analog ist die Kovarianz zwischen zwei Alternativen (j'_1, j_2) und (j_1, j_2) , die dieselbe Partei, aber unterschiedliche Kandidaten haben, $\text{Cov}(\eta_{j_1 j_2}, \eta_{j'_1 j_2}) = \tilde{\sigma}_{j_2}$. Weiterhin ist die Kovarianz von zwei Optionen, die sich auf denselben Kandidaten und dieselbe Partei beziehen, einfach die Varianz für die Alternative (j_1, j_2) :

⁷ Genauer nehmen wir einen Skalenparameter von 1 und einen Lageparameter von $-\gamma$, der negativen Euler-Mascheroni-Konstanten, an.

$$\text{Var}(\eta_{j_1, j_2}) = E[(\mu_{j_1} + \tilde{\mu}_{j_2} + \epsilon_{j_1, j_2})^2] = \sigma_{j_1} + \tilde{\sigma}_{j_2} + \pi^2/6.$$

Die Korrelation zwischen zwei beliebigen Alternativen (j_1, j_2) und (j'_1, j'_2) ist somit

$$\text{Corr}((j_1, j_2), (j'_1, j'_2)) = \begin{cases} 1 & \text{falls } j_1 = j'_1 \text{ und } j_2 = j'_2 \\ \frac{\sigma_{j_1}}{\sqrt{\sigma_{j_1} + \tilde{\sigma}_{j_2}} \sqrt{\sigma_{j'_1} + \tilde{\sigma}_{j'_2}}} & \text{falls } j_1 = j'_1 \text{ und } j_2 \neq j'_2 \\ \frac{\tilde{\sigma}_{j_2}}{\sqrt{\sigma_{j_1} + \tilde{\sigma}_{j_2}} \sqrt{\sigma_{j'_1} + \tilde{\sigma}_{j'_2}}} & \text{falls } j_1 \neq j'_1 \text{ und } j_2 = j'_2 \\ 0 & \text{falls } j_1 \neq j'_1 \text{ und } j_2 \neq j'_2 \end{cases} \quad (2)$$

Wir verwenden Markov-Chain-Monte-Carlo-Simulation in einem bayesianischen Ansatz, um Parameterschätzungen für das gemischte Logit-Modell zu erhalten. Als Ausgangsverteilungen für die Modellparameter wählen wir $\beta_k \sim N(0, 100)$, und $1/\sigma_k \sim \Gamma(0,01, 0,01)$. Die Simulation wird in JAGS (Plummer 2003) implementiert, mit zwei Markov-Ketten mit jeweils 5000 Iterationen (ausgedünnt um den Faktor 10; nach einem Burn-in von 15.000 Iterationen). Als Startwerte für die β -Koeffizienten verwenden wir in der einen Kette die Schätzungen eines konventionellen konditionalen Logit-Modells ohne den μ -Term und in der anderen Kette die gleichen Schätzungen plus zufälliges Rauschen. Die Standardabweichungen σ_k werden mit Zufallszahlen aus einer uniformen Verteilung über dem Intervall $[0, 2]$ initialisiert. Zur Überprüfung der Konvergenz verwenden wir den Ansatz von Gelman und Rubin (1992). Die vollständigen Replikationsunterlagen sind auf Anfrage bei den Autoren erhältlich.

5. Daten

Komponentenübergreifende Effekte in Mischwahlsystemen zu untersuchen erfordert auf empirischer Ebene relevante Informationen über Wählerinnen und Alternativen. Das größte Hindernis für eine empirische Analyse der Präferenzen von Wählern auf der Mikroebene ist der Mangel an Daten über die Wahlkreiskandidaten. Während viele Wahlumfragen Informationen über Wähler und Parteien liefern, sind Daten über die wahrgenommenen Merkmale der Kandidaten, die möglicherweise für die Wahl relevant

sind, aus offensichtlichen Gründen selten. In Deutschland hat eine prototypische nationale Wahlumfrage etwa 2000 Befragte, so dass bei 299 Wahlkreisen die Zahl der Befragten, die die Bewertungen eines einzelnen lokalen Kandidaten abgeben könnten, gering ist.

Wir wählen einen anderen Ansatz und führen eine Fallstudie über eine kleine Anzahl von Wahlkreisen durch, so dass die Zahl der Befragten, die sowohl Parteien als auch Direktkandidaten bewerten, hinreichend groß ist. Da wir vermuten, dass es für den Einzelnen schwieriger ist, Kandidaten zu bewerten als Parteien, haben wir zwei Wahlkreise mit allgemein bekannteren Direktkandidaten ausgewählt. Da wir uns außerdem für die Eigenschaften der Kandidaten und nicht für das Ansehen ihrer Kreispartei interessieren, versuchen wir, die Varianz des lokalen Parteiensystems zu minimieren, die nicht auf die einzelnen Kandidaten zurückzuführen ist. Wir haben daher zwei benachbarte Wahlkreise in Stuttgart ausgewählt, die Wahlbezirke 258 und 259 (im Folgenden S1 und S2) der Bundestagswahl 2013, die am 22. September stattfand. In S1 trat der bisherige Inhaber des Direktmandats Stefan Kaufmann (CDU) wieder an, herausgefordert von der ehemaligen Landesvorsitzenden der SPD, Ute Vogt, einem der damaligen Bundesvorsitzenden der Grünen, Cem Özdemir, und einer bei der Vorwahl erfolgreichen Listenabgeordneten der FDP, Judith Skudelny. In S2 kandidierte die amtierende Wahlkreisabgeordnete Karin Maag (CDU) gegen eine Abgeordnete der Grünen, Brigitte Bender, und zwei politische Neulinge, den Sozialdemokraten Nicolas Schäfstoß und den Freidemokraten Matthias Werwigk. Wir berücksichtigen auch die lokalen Kandidaten der sozialistischen Partei 'Die LINKE' (Christina Frank bzw. Marta Aparicio de Eckelmann) und somit alle damals im Bundestag vertretenen Parteien.

In 20 zufällig ausgewählten Stimmbezirken wurde eine Wahltagsbefragung durchgeführt, bei der jeder fünfte Wähler, der das Wahllokal verließ, gebeten wurde, einen zweiseitigen Fragebogen auszufüllen. Die Zielgruppe ist also die Gruppe der Wahllokal-Wähler. Insgesamt haben 1252 bzw. 839 Wähler an der Umfrage in den Wahlkreisen S1 bzw. S2 teilgenommen. Die Unterschiede zwischen den aggregierten Umfrage- und den tatsächlichen Ergebnissen bei den Wahllokal-Wählern sind gering (s. Anhang A).

Mit der Umfrage wurde die Erst- und Zweitstimmenentscheidung erfasst (unter Verwendung der oben genannten Wahlmöglichkeiten plus „Andere“), Skalometer-Bewertungen der fünf Direktkandidaten und der fünf Bundestagsparteien auf einer von „halte sehr viel“ über „ziemlich viel“ „teils/teils“ „ziemlich wenig“ bis „gar nichts“ laufenden Skala, eine Angabe zur Parteianhängerschaft („Welcher Partei fühlen Sie sich im Allgemeinen

zugehörig?“) sowie eine offene Frage zu dem wahrscheinlichen Wahlkreisgewinner. Diese Fragen verwenden wir zur Modellierung der Wählerpräferenzen über den Entscheidungsraum. Mit dem Ausschluss von Befragten, die angaben einen Kandidaten oder eine Partei jenseits der fünf Bundestagsparteien gewählt zu haben (n=264) oder keine bzw. unvollständige Angaben zur Wahlentscheidung oder beim Skalometer machten (n=559), reduziert sich die Fallzahl für die Modellschätzung auf 1268.

5.1 Instrumentelle Bewertung von Kandidaten und Parteien

Wir betrachten Politiknähe und Valenz als die beiden Quellen des instrumentellen Nutzens, die Wähler mit Kandidaten und Parteien verbinden. Der gängige Ansatz zur Erfassung der Nähe der Wähler zu Parteien ist die Befragten zu bitten, sich selbst und die Parteien auf einer oder mehreren ideologischen oder inhaltlichen Politikdimensionen zu verorten und dann die Nähe durch eine Form des Abstands zwischen der Selbsteinordnung der Wähler und dem Mittelwert der wahrgenommenen Parteipositionen zu operationalisieren. Die Verwendung des Mittelwerts der Parteipositionen ist zwar in der Literatur mindestens schon seit Markus und Converse (1979) bekannt, ihre weite Verbreitung auch in länderübergreifenden Umfragen (z. B. CSES) hat diesen Ansatz zunehmend populär gemacht (siehe z. B. Alvarez/Nagler 2004; Bawn/Somer-Topcu 2012). Allerdings haben Wählerbefragungen zur Schätzung von Parteipositionen einige offensichtliche Nachteile. Der wichtigste Nachteil ist, dass es nicht unwahrscheinlich ist, dass die Befragten die Skalen und ihre Endpunkte unterschiedlich interpretieren. Zur direkten Erfassung von Valenzmerkmalen gibt es überhaupt keine etablierte Methode. Viele empirische Studien, die das Konzept der Valenz verwenden, versuchen nicht, diese *ex ante* zu messen, sondern interpretieren alternativenspezifische Konstanten in z. B. multinomialen Regressionsmodellen als *ex post*-Proxy der Valenz.

Wir greifen deshalb auf den Bewertungsskalometer, wie in der Umfrage verwendet, zurück. Derartige Skalometer liefern ja gerade eine Gesamtbeurteilung von Objekten, die potenziell sowohl räumliche Nähe als auch Valenzmerkmale widerspiegelt. Genau auf diese angenommene Doppelinformation baut unser Ansatz. Wir nutzen die Methode der statistischen Entfaltung der Skalometerwerte der Befragten, um Positionen und Valenzen von Kandidaten und Parteien sowie die Positionen der Befragten aus den Daten abzuleiten. Wir verwenden dann diese entfalteten Daten und

nicht die Skalometerwerte selbst, da wir nicht ausschließen können, dass Wähler bei der Bewertung von Kandidaten und Parteien ideologische Nähe und Valenz in unterschiedlicher Weise gewichten.

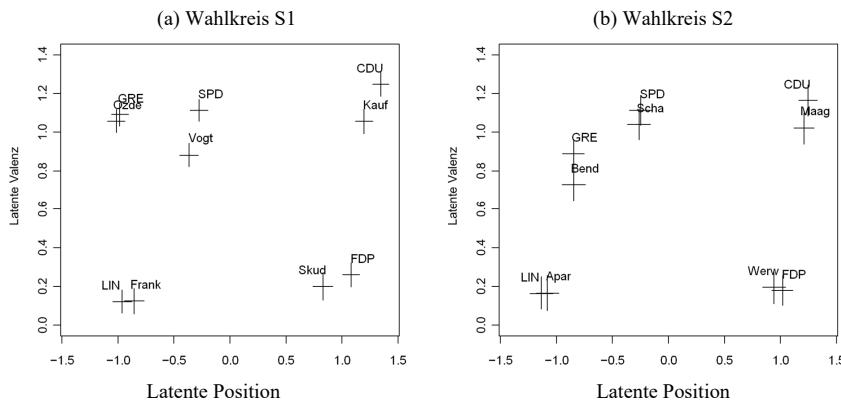
Konkret wenden wir eine Variante von Coombs (1964) Entfaltungsme- thode an (siehe auch Enelow/Hinich 1984; Brady 1990). Der Skalometer- wert s der Person i für das Objekt (Partei oder Kandidat) j wird als lineare Funktion der Valenz des Objekts v_j und des Abstands des Idealpunkts der Person von dem Bewertungsobjekt $\|x_i - x_j\|$ angenommen:

$$s(i, j) \sim N(v_j + \alpha \|x_i - x_j\|, \sigma)$$

wobei α die Gewichtung der Nähe relativ zur Valenz ist, N für die Normalverteilung mit Varianz σ steht, und s im Bereich von -2 und +2 liegt. Alle Parameter auf der rechten Seite werden aus den Daten bayesianisch geschätzt.⁸ Es ist zu beachten, dass wir die Makrovalenz der Parteien und Kandidaten schätzen und nicht individuelle Werte der Befragten. Individuelle Abweichungen von der Makrovalenz sind vermutlich auf die Logik der Nähe bei der Beantwortung der Bewertungsfragen zurückzuführen. Daher halten wir es für gerechtfertigt, unser Maß für die Makrovalenz als Allgemeinwissen in dem jeweiligen Wahlkreis zu interpretieren. Abbildung 1 zeigt die latenten Positionen und Valenzschätzungen für die Kandidaten und Parteien in den beiden Wahlkreisen.

⁸ Hier sind zwei Anmerkungen zu machen. Erstens ist die Bewertungsskala in {-2, -1, 0, 1, 2} kodiert, so dass wir tatsächlich eine zensierte Normalverteilung mit den Grenzen -2 und +2 verwenden. Zweitens haben Käppner und Shikano (2015) gezeigt, dass die Behandlung solcher Skalometerwerte als ordinale Multinomialverteilung ana- loge Ergebnisse liefert, sodass wir trotz der vorliegenden 5er Skala bei der einfacheren Variante bleiben.

Abbildung 1: Aus den Bewertungsskalen rekonstruierte Positionen und Valenzen von Kandidaten und Parteien



Anmerkung: Die Abbildung zeigt die Regionen mit 95% höchster Dichte. Entfaltung mit latenten Positionen der Parteien und Kandidaten, normalisiert auf einen Mittelwert von 0 und eine Standardabweichung von 1. Die Schätzungen basieren auf den Bewertungsskalometern von $N = 959$ Personen in S1 bzw. $N = 570$ in S2 nach Ausschluss von Befragten mit 5 oder mehr ungültigen oder Nicht-Antworten auf die 10 Stimuli (fünf Parteien und fünf Kandidaten) ($n = 293$ und $n = 269$).

Auf der Nähe-Dimension erscheinen die Parteien in der Reihenfolge, in der wir sie auf einer ideologischen Links-Rechts-Dimension erwarten würden: Die Partei „Die Linke“ stehen am weitesten links, gefolgt von Grünen und Sozialdemokraten, während Liberale und Christdemokraten die rechte Seite des Spektrums besetzen. Wir bezeichnen daher die Positionen auf der Nähe-Dimension als ideologische Links-Rechts-Positionen. Auf der zweiten, der Valenzdimension, sind die Parteien mit der höchsten Valenz CDU und SPD, dicht gefolgt von den Grünen, während FDP und Linke dahinter zurückbleiben. Drei Dinge sind in diesem Gesamtbild erwähnenswert. Erstens sind die Positions- und Valenzeinschätzungen der Parteien in den beiden Bezirken sehr ähnlich, was darauf hindeutet, dass der Stimulus „Partei“ einen Bezug zu einem Objekt auslöst, der über den Wahlkreis hinausgreift. Zweitens sind die Positionen von Kandidaten und Parteien oft fast identisch, was darauf schließen lässt, dass die Kandidaten entweder nicht viel Spielraum zum Aufbau eines persönlichen Profils auf der Grundlage von Politik haben oder aber sie haben diesen, die Befragten sind jedoch nicht in der Lage, diese Unterschiede wahrzunehmen. Drittens

unterscheiden sich Kandidaten und Parteien häufiger erheblich in Bezug auf die zugeschriebene Valenz. Dabei schneiden Parteien fast immer besser ab als Kandidaten. Dies alles deutet darauf hin, dass Valenzmerkmale und Positionen tatsächlich unterschiedliche Objektattribute sind. In den folgenden Analysen verwenden wir diese Valenz von Kandidaten und Parteien sowie die räumliche Distanz von Kandidaten und Parteien zu den Wählern.

5.2 Psychologische Wirkung der Wahlformel

Erst- und Zweitstimme unterscheiden sich nicht nur in Bezug auf das zu wählende Objekt, sondern auch in der Wahlformel. In der Verhältniswahl-Komponente werden die Parteien im Wesentlichen proportional zu ihren Stimmenanteilen vertreten. In der Mehrheits-Komponente gewinnt der Kandidat mit der relativen Mehrheit, was zu einer strategischen Stimmabgabe motivieren kann, um die eigene Stimme nicht an aussichtslose Kandidaten zu verschwenden. Wenn Individuen also zukunftsorientiert sind, werden sie den instrumentellen Nutzen eines Kandidaten mit dessen Gewinnchancen abwägen. Während strategisches Wählen tatsächlich ein spieltheoretisches Problem ist, kann man sich in erster Näherung die Gewinnchancen eines Kandidaten vor der Wahl als Diskontierungsfaktor für den instrumentellen Nutzen vorstellen. Theil (1969) und King (1990) schlugen vor, die erwartete Gewinnwahrscheinlichkeit in einem Mehrparteiensystem durch das Verhältnis der Stimmenanteile zu operationalisieren, die mit einer Konstante ρ potenziert werden, was wir hier aus dem einfachen Grund nicht tun können, weil uns für unsere beiden Wahlkreise keine Vorwahlumfragen oder andere Informationen über die erwarteten Stimmenanteile zur Verfügung stehen.

Wir können uns jedoch die Tatsache zunutze machen, dass sich die lokalen Parteiorganisationen der Sozialdemokraten und Grünen während des Wahlkampfs koordiniert und öffentlich empfohlen haben, für den grünen Kandidaten in S1 und den SPD-Kandidaten in S2 zu stimmen. Angesichts des Siegs der Amtsinhaber bei der letzten Wahl kann man annehmen, dass die aussichtsreichsten Kandidaten in den beiden Wahlkreisen die Kandidaten der CDU und der Grünen im ersten und der CDU und SPD im zweiten Wahlkreis sind. Auch wenn sich daraus kein Wahrscheinlichkeitsmaß ergibt, können wir eine einfache Dummy-Variable verwenden, um zwischen nicht aussichtsreichen Kandidaten (für die wir einen Malus erwarten) und aussichtsreichen Kandidaten (für die wir keinen Malus erwarten) zu un-

terscheiden. Schließlich machen wir Gebrauch von einer Frage, die offen nach dem „wahrscheinlichsten Gewinner“ der Mehrheitswahlen fragte. Nur wenn die Befragten in der Lage waren, einen der in Frage kommenden Kandidaten zu nennen, betrachten wir diesen als möglichen Empfänger strategischer Stimmen. Dafür ist die Variable „aussichtsreich“ gedacht.

5.3 Expressive Motivation zur Kandidaten- und Parteiwahl

Die Wahltagsbefragung enthielt auch die Frage, welcher Partei man sich im Allgemeinen „nahe fühlt“. Sie erfasst nicht die instrumentelle, sondern die expressive Motivation für die Wahlentscheidung. Mit der Frage wird weniger die Parteiidentifikation im Sinn einer langfristigen und stabilen affektiven Orientierung an einer Partei erfasst, die der Einzelne in seiner politischen Sozialisation entwickelt hat. In europäischen Mehrparteiensystemen war es vor allem die vermeintliche Langfristigkeit des ursprünglichen Konzepts der Parteiidentifikation (Campbell et al. 1954, 1960), das sich als Hindernis für die Anwendung des Konzeptes herausgestellt hat (Pappi 2011). Wir denken, dass die Frageformulierung „sich einer Partei nahe fühlen“ eine affektive Bindung jenseits kurzfristiger Kampagneneffekte erfasst, die gleichzeitig weniger anspruchsvoll ist in Bezug auf die Langlebigkeit als das ursprüngliche Konzept der Parteiidentifikation.

6. Ergebnisse

Die ersten drei Spalten von Tabelle 1 enthalten eine Zusammenfassung der A-posteriori-Werte der fixen, nicht zufälligen sowie der zufälligen Parameter des statistischen Modells auf der Grundlage von Gleichung 1. Wir geben die Mittelwerte und die Spannweite der Regionen mit der höchsten 95-prozentigen Dichte der A-posteriori-Werte an. Die Ergebnisse sind recht beeindruckend, da sie insgesamt darauf hindeuten, dass die Wahlentscheidung bei der Zwei-Komponenten-Wahl durch Faktoren in allen drei Bereichen des theoretischen Modells erheblich beeinflusst wird. Erstens hängt die Wahlentscheidung zwischen den 25 Optionen von der Nähe des Einzelnen und der Valenz der Partei auf die erwartete Weise ab. Der negative Mittelwert des geschätzten Werts für die *Distanz zur Partei* (-0,924) deutet darauf hin, dass die Positionen der Partei und des Wählers umso näher beieinanderliegen, je wahrscheinlicher es ist, dass die Person für die Partei stimmen

wird. Der Mittelwert der *Parteivalenz* ist positiv bei 0,260, was darauf hindeutet, dass Parteien, denen eine hohe Valenz zugeschrieben wird, mit größerer Wahrscheinlichkeit eine Zweitstimme erhalten. Wir stellen auch fest, dass die „*Parteiidentifikation*“ bei der Stimmabgabe für die Parteiliste von Bedeutung ist (mittlerer Schätzwert von 2,144).

Tabelle 1: Parameterschätzungen des gemischten konditionalen Logit-Modells der Wahlentscheidung (95-prozentige Wahrscheinlichkeitsbereiche)

	Modell 1			Modell 2		
	2,5%	Mittelwert	97,5%	2,5%	Mittelwert	97,5%
Räumliche Distanz: Kandidat	-1,229	-0,967	-0,732	-1,116	-0,917	-0,715
Valenz: Kandidat	1,181	1,623	2,069	1,495	1,984	2,507
Räumliche Distanz: Partei	-1,206	-0,924	-0,684	-0,934	-0,798	-0,665
Valenz: Partei	-0,107	0,260	0,590	0,245	0,478	0,719
Parteiidentifikation: Kandidat	1,193	1,467	1,762	1,128	1,313	1,506
Parteiidentifikation: Partei	1,571	2,144	2,632	1,475	1,643	1,816
Distanz Kandidat \times aussichtsreich	-0,518	-0,242	0,029	-0,485	-0,235	0,022
Valenz Kandidat \times aussichtsreich	1,045	1,454	1,849	0,917	1,290	1,654
Kongruente Wahl	0,546	0,826	1,112			
Kongruente Wahl: CDU				0,748	1,012	1,290
Kongruente Wahl: SPD				0,999	1,241	1,510
Kongruente Wahl: FDP				-0,419	0,416	1,117
Kongruente Wahl: Grüne				-0,337	-0,080	0,179
Kongruente Wahl: LINKE				1,645	2,174	2,708
σ_1	0,100	0,804	1,741	0,056	0,133	0,373
σ_2	0,096	0,397	0,829	0,049	0,117	0,250
σ_3	0,095	0,248	0,541	0,052	0,106	0,217
σ_4	0,102	0,428	1,015	0,049	0,117	0,225
σ_5	0,085	0,540	1,435	0,044	0,114	0,226
$\tilde{\sigma}_1$	0,123	1,240	2,800	0,048	0,125	0,261
$\tilde{\sigma}_2$	0,087	1,468	3,267	0,044	0,091	0,166
$\tilde{\sigma}_3$	0,089	0,517	1,642	0,045	0,127	0,336
$\tilde{\sigma}_4$	0,137	1,006	2,024	0,052	0,104	0,204
$\tilde{\sigma}_5$	0,079	0,481	1,136	0,050	0,112	0,203
N	1268			1268		

Zweitens, und das ist am interessantesten, haben wir Hinweise auf einen analogen Mechanismus, der sich bei der Erststimme abspielt. Tabelle 1 zeigt, dass die Distanz zu den Kandidaten negativ mit -0,967 die Kandidatenwahl beeinflusst und der Effekt wiederum mit hoher Sicherheit die

erwartete Richtung hat. Die Kandidatenvalenz wirkt sich positiv auf die Stimmabgabe für diesen Kandidaten aus, wie auch die Parteidentifikation. Am bemerkenswertesten ist wohl die starke Evidenz, die wir für strategisches Wählen mit der Erststimme finden. Die Parameterschätzung für die Interaktionsterme des bekannten aussichtsreichen Kandidaten (die wir in Ermangelung besserer Daten als einfache Dummy kodiert haben) und der Kandidatendistanz und -valenz, die die instrumentelle Motivation der Individuen erfassen, haben die erwartete negative bzw. positive Wirkung. Genauer gesagt, die Kandidatenwahl von Personen, die zu einer strategischen Wahlentscheidung neigen (stellvertretend für die korrekte Wahrnehmung der in Frage kommenden Kandidaten), wird stärker durch die Nähe des Kandidaten und seiner Valenz bestimmt als die einer vermutlich nicht-strategischen Person. Tatsächlich ist der Effekt für die Kandidatenvalenz fast doppelt so groß ($1,623 + 1,454 = 3,077$ am Mittelwert).

Drittens finden wir Belege für die erwartete Motivation, sich konsistent über die zwei Stimmabgaben hinweg zu verhalten, und damit für die Nicht-Unabhängigkeit der Erst- und Zweitstimme. Der geschätzte Koeffizient der Dummy-Variablen, die eine kongruente Abstimmung anzeigt, d. h. die Wahl eines Kandidaten derselben Partei, die man mit der Zweitstimme wählt, ist positiv, was darauf hindeutet, dass kongruente Wahloptionen *ceteris paribus* häufiger attraktiver sind als nicht-kongruente. Wir stellen fest, dass fast alle diese Parameter sehr hoch „signifikant“ in dem Sinne sind, dass wir fast sicher sein können, dass der Effekt negativ (bzw. positiv) ist, da die *A-posteriori*-Werte fast vollständig im negativen (bzw. positiven) Halbraum liegen. Ob diese Effekte auch substantiell in dem Sinne sind, dass die Wahrscheinlichkeit der Stimmabgabe für diese oder jene Option in erheblichem Maße mit den Kovariaten variiert, ist eine andere Frage. Wir kommen auf diese Frage zurück, wenn wir die Vorhersagewahrscheinlichkeiten diskutieren.

An dieser Stelle scheint es angebracht, eine Pause einzulegen, um die Relevanz dieser auf den ersten Blick recht plausiblen und einfach anmutenden Ergebnisse sowie einige Einschränkungen zu bedenken. Einerseits könnte man die oben dargelegten Ergebnisse als Beweis dafür interpretieren, dass Erst- und Zweitstimmen, Kandidaten- und Parteilistenstimmen von denselben Faktoren beeinflusst werden. Was wir über die Wahl der Parteilisten wissen – die im Mittelpunkt früherer Studien über gemischte Wahlsysteme standen – scheint sich gut auf die Kandidatenstimme übertragen zu lassen. Man sollte aber auch betonen, dass die Parameterschätzungen für Kandidaten und Parteiliste in signifikanter Weise variieren: Die Valenz der Kandi-

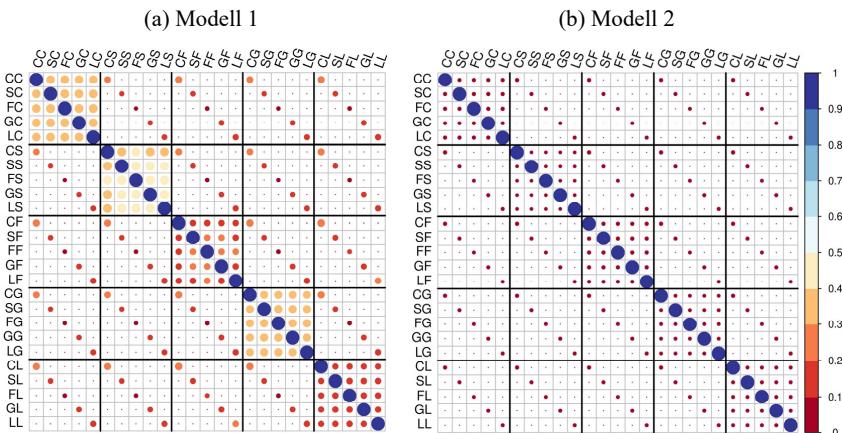
daten hat einen größeren Einfluss auf die Erststimme als die Parteivalenz (während sich die Wirkung der Distanz nicht unterscheidet). Umgekehrt ist der Effekt der Parteiidentifikation doppelt so groß für die Parteiliste wie der für die Kandidaten. Dies alles deutet darauf hin, dass die Erststimme tatsächlich durch eine Personalisierung der Stimme gekennzeichnet ist. Für den lokalen Kandidaten sind persönliche Eigenschaften wie Integrität, Kompetenz und was auch immer eine Valenz aufbaut, vergleichsweise wichtiger als seine oder ihre Parteizugehörigkeit. Wenn sich jemand einer Partei verbunden fühlt, hat er eine starke Motivation, für eben diese Parteiliste zu stimmen, aber nicht in gleichem Maße für den Kandidaten dieser Partei.

Andererseits sollten wir einen Hinweis zur Vorsicht hinzufügen. Wie wir aus den obigen Ausführungen wissen, unterscheiden sich die ideologischen Positionen der Parteien und ihrer Kandidaten nicht sehr stark. Größere Unterschiede gibt es bei der Valenz, die den Kandidaten und Parteien zugeschrieben wird, aber auch hier sind die Unterschiede eher gering. Was wäre, wenn die Wähler die Parteinähe und Parteivalenz auf die lokalen Kandidaten der Parteien übertragen, so dass die Kandidaten eher Klone als Persönlichkeiten sind und die vermeintlichen Wirkungen der Kandidateneigenschaften in Wirklichkeit Auswirkungen der Parteieigenschaften sind? Wir denken, dass dies weniger plausibel ist, wenn wir die geschätzten Effektgrößen berücksichtigen. Wenn die Valenz der Kandidaten ein bescheidener Ersatz für die Parteivalenz wäre, würde man einen größeren Effekt der Partei- als der Kandidatenvalenz beobachten (was nicht der Fall ist). Wir würden auch erwarten, dass der Effekt der Parteiidentifikation praktisch der gleiche auf beiden Ebenen ist (was hier ebenfalls nicht der Fall ist). Ein direkter Test würde eine erneute Schätzung des Modells, nachdem ein Kandidatenmerkmal nach dem anderen durch ein Merkmal der Partei des Kandidaten ersetzt wurde, beinhalten. Wenn die Parteieigenschaften einfach übertragen würden, würde die Varianz zwischen den Kandidaten- und Parteieigenschaften nur weißes Rauschen bedeuten, und wir würden erwarten, dass das so geschätzte Modell besser zu den Daten passt als das ursprüngliche Modell. Anhang B enthält die Ergebnisse dreier multinomialer Logit-Modellen, in denen wir einen denkbaren Transfer von Parteiwahrnehmungen untersuchen. Die Ergebnisse sind nicht völlig eindeutig. Sie deuten aber darauf hin, dass die Wahl der Kandidaten besser durch die Kandidatenvalenz „erklärt“ wird als durch die Parteivalenz.

6.1 Nicht-modellierte Substitutionsmuster

Als nächstes wenden wir uns den Zufallstermen μ zu, die zur Modellierung von Korrelationen zwischen Wahloptionen verwendet werden, um geeignete oder wahrscheinliche Substitutionsmuster zu entdecken, die nicht mit den Kovariaten erfasst werden. Die Zufallsterme wurden als normalverteilt mit Mittelwert Null angenommen. Zugegebenermaßen können die Varianzschätzungen für jeden einzelnen Cluster nur begrenzt interpretiert werden. Intuitiver ist die Korrelation der Residualnutzen für die Wahloptionen, die sich aus diesen Schätzungen mit Hilfe der Gleichung 2 ergeben.

Abbildung 2: Korrelationen der Residualnutzen der 25 Wahloptionen des gemischten konditionalen Logit-Modells (Matrizen nach der Zweitstimme sortiert)



Anmerkung: Die Felder zeigen die geschätzten Korrelationen der Fehlerkomponente von Gleichung 1 auf der Grundlage der A-posteriori-Mittelwerte der Varianzen σ und $\tilde{\sigma}$ für die Modelle (1) und (2) in Tabelle 1.

Panel (a) von Abbildung 2 zeigt die Korrelationsmatrix der paarweise betrachteten Fehler der 25 Optionen. Die Zeilen und Spalten sind nach der Zweitstimme sortiert, beginnend mit der Option (CDU, CDU), gekennzeichnet mit CC, dann SC, FC, GC und so weiter bis LL für das Paar (Linke, Linke) als 25. Zeile bzw. Spalte. Die Korrelationen in der Hauptdiagonalen sind definitionsgemäß 1, die Korrelationen in den Blockmatrizen außerhalb derer entlang der Hauptdiagonalen sind nur in den jeweiligen Diagonalen empirisch gewonnen, außerhalb aber annahmegemäß 0 (vgl.

Gleichung 2). Letztere haben weder eine Erst- noch eine Zweitstimmenwahl gemeinsam. Die empirisch gewonnenen Korrelationen deuten darauf hin, dass die Fehler zwischen den Wahloptionen eher gering korrelieren; die Koeffizienten reichen bis 0,4. Wenn zwei Wahloptionen weder bezüglich der Kandidaten- noch der Parteiwahl übereinstimmen, gehen wir nicht davon aus, dass die eine Alternative die andere ersetzen könnte. Im Gegensatz dazu teilen sich z. B. die Optionen SS und CS die Zweitstimme – die sozialdemokratische Parteiliste –, was es wahrscheinlicher macht, dass ihre Fehler korreliert sind. In dem Maße, in dem wir Stimmen für die SPD-Parteiliste nicht erklären können, leisten wir sowohl bei SS als auch bei CS eine schlechte Arbeit.

Die nach der Zweitstimme sortierten Blockmatrizen entlang der Hauptdiagonalen weisen insgesamt höhere Korrelationen auf, insbesondere für Optionen, die Listenwahlen für die beiden größten Parteien, die Christdemokraten und die Sozialdemokraten, beinhalten. Offensichtlich ist unser Modell einer Zwei-Güter-Wahl am wenigsten aussagekräftig für Wahlmöglichkeiten innerhalb dieser beiden Cluster. Mit anderen Worten, wir tragen vergleichsweise wenig zur Erklärung von Listenwahlen der CDU und SPD bei, was zu großen Fehlervarianzen für diese Wahloptionen führt. Die gute Nachricht hier ist, dass die Korrelationsmuster darauf hindeuten, dass wir die Erststimmen bemerkenswert gut erklären können. Dies ist in der Tat bemerkenswert, da die Kandidatenwahl eine strategische Wahl beinhaltet, die den Wahlmechanismus komplexer macht.

Panel (b) von Abbildung 2 präsentiert die Korrelationen eines erweiterten Modells, das eine mögliche Variation des Effekts einer kongruenten Stimmabgabe berücksichtigt.⁹ Die letzten drei Spalten von Tabelle 1 zeigen die Parameterschätzungen für dieses zweite Modell. Die Ergebnisse sind einfach, aber auch aufschlussreich. Erstens finden wir in Bezug auf die

9 Wir haben auch ein erweitertes Modell betrachtet, das separate Effekte für den kongruenten Wahlanreiz $C_{j_1j_2}$ für Personen mit und ohne angegebene Parteidentifikation schätzt. Da ein stark parteigebundener Wähler der Partei, der er nahesteht, seine beiden Stimmen geben würde, könnte der bedingte Effekt der kongruenten Stimmabgabe für Parteianhänger kleiner sein als für Nicht-Parteianhänger. Die empirischen Ergebnisse sind nicht schlüssig. Wenn wir den Term $\beta_{10}t_{ij_2}C_{j_1j_2}$ hinzufügen (wobei t_{ij_2} gleich 1 ist, wenn i sich mit der Partei j_2 identifiziert, und andernfalls gleich Null), dann ist der Posterior-Wert von β_{10} nur mit einer Wahrscheinlichkeit von etwa 60 Prozent positiv. Während jedoch die meisten Posterior-Werte in Bezug auf Skala und Streuung robust sind, steigt der Effekt von C , β_9 , auf das 15-fache und der Effekt der Parteivalenz wird negativ.

Determinanten der Erst- und Zweitstimme praktisch die gleichen Effekte wie in Modell 1. Zweitens unterscheidet sich die Bedeutung der kongruenten Stimme tatsächlich zwischen Parteien: Die Ergebnisse zeigen, dass die Abgabe einer kongruenten Stimme besonders wichtig ist, wenn es um die Bewertung der Parteien und Kandidaten von CDU und SPD geht. Dies könnte die Tatsache widerspiegeln, dass der CDU-Kandidat in beiden Bezirken aussichtsreich ist, und während der SPD-Kandidat nur in einem Bezirk aussichtsreich ist, werden einige (langjährige) SPD-Parteimitglieder eine kongruente Abstimmung für ihre Partei als die natürlichste Option ansehen – eine Reminiszenz an das alte Dreiparteiensystem, in dem der SPD-Kandidat immer ein aussichtsreicher Kandidat war. Anders ist dies bei den Liberalen und den Grünen, für die strategisches „split-ticket“-Wählen schon immer eine wichtige Überlegung war. Am interessantesten und möglicherweise überraschend ist, dass wir den größten positiven Effekt für die Die LINKE finden. Dies deutet darauf hin, dass ihre Anhänger am ehesten geneigt sind, eine kongruente Stimme abzugeben, weil eine strategische, getrennte Stimmabgabe nicht attraktiv ist (d. h. die aussichtsreichen Kandidaten sind nicht attraktiv, sodass sie von den Positions- und Valenzmerkmalen, die im Modell enthalten sind, nicht gut erfasst werden), möglicherweise weil linke Parteianhänger stärker ideologisch orientiert sind. Drittens, die Zufallseffekte (σ) sind viel kleiner und gleichmäßiger verteilt, so dass die resultierende Korrelationsmatrix von Panel (b) durchgängig nur sehr geringe Werte aufweist. Insgesamt deutet dies alles darauf hin, dass Modell 2 die Logik der strategischen Stimmabgabe im deutschen Mischwahlsystem sehr gut erfasst.

6.2 Vorhergesagte Stimmabgabe

Tabelle 2 zeigt eine Kreuztabelle der vorhergesagten und der beobachteten Wahlentscheidungen. Insgesamt ist die Anzahl der korrekt vorhergesagten Wahlentscheidungen 767, was 60,5 Prozent entspricht. Nach den Daten verzeichnet das Paar (CDU, CDU) mit 435 Stimmen oder 36,3 Prozent die größte Häufigkeit, gefolgt von den kongruenten Paaren (SPD, SPD) mit 226 Stimmen (18,8 Prozent) und (Grüne, Grüne) mit 165 Stimmen (13,8 Prozent). Alle anderen Optionen erhalten deutlich weniger Stimmen. Unser Modell bildet dieses Gesamt muster sehr gut ab. Interessanter aber ist, wo wir scheitern. Ein Punkt ist die kongruente Stimmabgabe für die Grünen. Das Modell prognostiziert 290 Stimmen für (GRÜ, GRÜ), was fast das

Doppelte der tatsächlichen Zahl ist. In gewisser Weise ist dies nicht wirklich überraschend angesichts der der grünen Partei und den Kandidaten in beiden Wahlkreisen zugeschriebenen hohen Valenz, und der Tatsache, dass viele Befragte angaben, den Grünen nahe zu stehen (insgesamt 306 im Vergleich zu 323 für die SPD). Dies deutet darauf hin, dass es einen „Fehlanreiz“ für eine kohärente Grünen-Stimme und einen Anreiz für eine getrennte Stimmabgabe für einen grünen Kandidaten oder die grüne Partei gibt, die wir nicht erfassen.

Tabelle 2: Vorhergesagte (Spalten) und beobachtete (Zeilen) Wahlentscheidung auf Basis der A-posteriori-Mittelwerte

	CC	CF	SS	SG	FF	GS	GG	LL	Gesamt
CC	406	4	18	0	1	0	5	1	435
CS	10	0	10	1	0	0	2	0	23
CF	58	6	0	0	0	0	0	0	64
CG	5	0	2	0	0	0	8	0	15
CL	3	0	0	0	0	0	0	3	6
SC	10	0	8	0	0	0	0	0	18
SS	15	0	185	3	0	0	23	0	226
SF	0	1	0	0	0	0	0	0	1
SG	0	0	20	3	0	0	22	0	45
SL	1	0	10	0	0	0	6	6	23
FC	7	0	0	0	0	0	0	0	7
FF	9	1	0	0	1	0	0	0	11
FG	0	0	0	0	0	0	1	0	1
GC	15	0	2	1	0	0	10	1	29
GS	4	0	46	0	0	0	41	1	92
GF	3	4	1	0	0	0	1	0	9
GG	4	0	15	6	0	1	134	5	165
GL	0	0	5	0	0	0	25	10	40
LS	0	0	3	0	0	0	1	0	4
LG	0	0	1	0	0	0	1	2	4
LL	1	0	7	0	0	0	10	32	50
Gesamt	551	16	333	14	2	1	290	61	1268

Diese Interpretation wird durch eine zweite Beobachtung gestützt. Eine beträchtliche Anzahl von Befragten gibt an, dass sie für den grünen Kandidaten, aber die sozialdemokratische Liste ($N = 92$) oder umgekehrt für den SPD-Kandidaten und die grüne Liste ($N = 45$) gestimmt haben. Dies ist jedoch genau das, was wir erwarten sollten, wenn SPD- und Grünen-Parteianhänger auf ihre lokalen Parteigliederungen gehört haben, die öffentlich empfohlen hatten, die aussichtsreichsten linken Kandidaten im

jeweiligen Wahlkreis zu unterstützen. Unser Modell ist gegenüber dieser Art von Wahlverhalten unempfindlich. Unser Modell sieht diese Art von Wahlkoordinierung bei der Erststimme nicht vor, so dass wir beide Optionen unterschätzen. Schließlich unterschätzen wir auch die Wahl CF des CDU-Kandidaten und der FDP-Liste (beobachtet $N = 64$, vorhergesagt $N = 16$). Ein Grund hierfür könnte sein, dass einige konservative Wähler, die sonst die CDU wählen würden, die liberale Liste unterstützen, um sicherzustellen, dass die Partei die Fünf-Prozent-Hürde überschreitet. Da wir die strategische Stimmabgabe im Zusammenhang mit der Fünf-Prozent-Hürde nicht berücksichtigen, haben wir die CF-Option erwartungsgemäß unterschätzt.

7. Schlussfolgerungen

Unser Aufsatz liefert einen Beitrag zur laufenden Diskussion über das Wahlverhalten in gemischten Wahlsystemen. Wenn es eine Kontamination auf der Ebene der Wähler gibt, was genau verbindet dann die zwei Stimmen, die Mehrheitswahl- und die Verhältniswahl-Stimme? Wir argumentieren, dass die Entscheidungssituation tatsächlich eine simultan verbundene Entscheidung zwischen Parteilisten und Kandidaten darstellt. Infolgedessen schlagen wir vor, die Präferenzen der Wähler über die gesamte Menge aller Kombinationen von Parteilisten und Kandidaten zu modellieren, so dass split-ticket-Wählen die Regel und nicht die Ausnahme oder der abweichende Fall ist, der erklärt werden muss.

Obwohl unsere Modellierungsstrategie einfach ist, weicht sie von früheren Annäherungen an Spillover-Effekte in wichtiger und vorteilhafter Weise ab. Erstens unterscheiden wir zwischen der Bewertung von Wahlmöglichkeiten und der tatsächlichen Wahl. Wenn eine Kontamination vorliegt, wird sie sich auf die individuellen Präferenzen gegenüber den Wahlmöglichkeiten auswirken. Wir schlagen daher vor, die Kontamination auf der Ebene der Präferenzen zu betrachten. Verändert die Bewertung von Optionen in der einen Komponente die Bewertung der Optionen in der anderen Komponente? Zweitens ermöglicht uns die Betrachtung der verbundenen Entscheidung als ein Wahlproblem mit zwei Gütern es uns, einen eindeutigen Test auf Kontamination zu entwickeln. Wir argumentieren, dass eine Kontamination dann und nur dann vorliegt, wenn die Präferenzen nicht separabel sind. Für zwei Güter bedeutet Separabilität sowohl notwendig als auch hinreichend, dass die Nutzenfunktion additiv separabel ist. Drittens

erfordert dies auf der empirischen Ebene Informationen darüber, wie Wähler sowohl Kandidaten als auch Parteien bewerten. In der Literatur gibt es zwar zahlreiche Belege für die Kontamination von Erst- und Zweitstimmen, aber die meisten Belege beruhen auf aggregierten Daten, die in den meisten Fällen zwar Informationen über die ideologische Ausrichtung der Parteien liefern, aber nicht die der Kandidaten.

Aus diesem Grund ist auch unsere empirische Strategie anders. Wir führen eine Fallstudie zu zwei Wahlkreisen der Bundestagswahl 2013 durch, so dass die Menge der Befragten, die Auskunft über ihre Bewertungen sowohl der Parteien als auch der Kandidaten geben, groß ist. Wir nutzen diese Informationen, um unser theoretisches Modell der Wahlentscheidung zu kalibrieren und zusätzlich eine Zufallskomponentenspezifikation eines gemischten Logit-Modells einzuführen, die es uns ermöglicht, Interdependenzen oder Substitutionsmuster zwischen Alternativen zu untersuchen.

Die empirische Schätzung führt zu mehreren wichtigen Ergebnissen. Erstens, die politische Distanz der Wähler zu Kandidaten und Parteien hat den erwarteten negativen Effekt, und Valenz und Parteianhängerschaft wirken sich erwartungsgemäß positiv aus. Zweitens, und in Übereinstimmung mit der theoretischen Erwartung, sind Valenz-Eigenschaften für die Bewertung von Kandidaten wichtiger als für die Bewertung von Parteien, während die Parteianhängerschaft für die Verhältniswahl wichtiger ist als für die Mehrheitswahl. Drittens gibt es deutliche Hinweise darauf, dass Individuen strategisch wählen in Bezug auf die Kandidatenwahl. Viertens gibt es eine offensichtliche Tendenz der Individuen ihre zwei Stimmen kongruent abzugeben, während andere sinnvolle Substitutionsmuster von geringerer Bedeutung sind.

Diese Ergebnisse unterstreichen unser anfängliches Argument, dass die Konzentration auf Spillover-Effekte von institutionellen Regeln wahrscheinlich zu einschränkend ist, wenn man Nicht-Unabhängigkeit oder Kontamination im Wahlverhalten in gemischten Wahlsystemen betrachtet. Unser Argument über den Anreiz zur kongruenten Stimmabgabe hängt nicht mit den spezifischen institutionellen Regeln zusammen, die in der einen oder der anderen Komponente gelten. Die Ergebnisse deuten auf einen psychologischen Mechanismus der Dissonanzvermeidung hin, wenn Kandidaten-Parteilisten-Paare der gleichen Partei mehr präferiert werden als es die individuellen Eigenschaften der Kandidaten und Parteien vermuten lassen. Bemerkenswert ist, dass es sich dabei nicht um eine einfache Folge der Parteianhängerschaft (die sich direkt auf die Präferenzen von Kandidaten und Parteilisten auswirkt) handelt.

Referenzen

- Alvarez, R. M., und J. Nagler (2004). Party system compactness: Measurement and consequences. *Political Analysis* 12(1), 46–62.
- Ansolabehere, S., und J. M. Snyder (2000). Valence politics and equilibrium in spatial election models. *Public Choice* 103(3/4), 327–336.
- Bawn, K. (1993). The logic of institutional preferences: German electoral law as a social choice outcome. *American Journal of Political Science* 37(4), 965–989.
- Bawn, K., und Z. Somer-Topcu (2012). Government versus opposition at the polls: how governing status affects the impact of policy positions. *American Journal of Political Science* 56(2), 433–446.
- Brady, H. E. (1990). Traits versus issues: Factor versus ideal-point analysis of candidate thermometer ratings. *Political Analysis* 2(1), 97–129.
- Brennan, G., und L. E. Lomasky (2008). *Democracy and decision: the pure theory of electoral preference*. Cambridge: Cambridge University Press.
- Brennan, G. H., und A. P. Hamlin (1998). Expressive voting and electoral equilibrium. *Public Choice* 95(1), 149–175.
- Campbell, A., P. E. Converse, W. E. Miller und D. E. Stokes (1960). *The American voter*. New York: Wiley.
- Campbell, A., G. Gurin und W. E. Miller (1954). *The voter decides*. Evanston, Ill.: Row, Peterson.
- Coombs, C. H. (1964). *A theory of data*. New York: Wiley.
- Cox, G. W. (1997). Making votes count: strategic coordination in the world's electoral systems. Cambridge: Cambridge University Press.
- Cox, K. E., und L. J. Schoppa (2002). Interaction effects in mixed-member electoral systems: theory and evidence from Germany, Japan, and Italy. *Comparative Political Studies* 35(9): 1027–1053.
- Debreu, G. (1959). Theory of value: An axiomatic analysis of economic equilibrium. New York: Wiley.
- Duch, R. M., J. May und D. Armstrong (2010). Coalition-directed voting in multi-party democracies. *American Political Science Review* 104(4), 698–719.
- Duverger, M. (1954). *Political parties: their organization and activity in the modern state*. London: Methuen.
- Enelow, J. M., und M. J. Hinich (1984). *The spatial theory of voting: An introduction*. Cambridge: Cambridge University Press.
- Ferrara, F., und E. S. Herron (2005). Going it alone? Strategic entry under mixed electoral rules. *American Journal of Political Science* 49(1), 16–31.
- Ferrara, F., E. S. Herron und M. Nishikawa (2005). *Mixed electoral systems: Contamination and its consequences*. Basingstoke: Palgrave Macmillan.
- Gelman, A., und D. B. Rubin (1992). Inference from iterative simulation using multiple sequences. *Statistical Sciences* 7, 457–511.
- Groseclose, T. (2001). A model of candidate location when one candidate has a valence advantage. *American Journal of Political Science* 45(4), 862–886.

- Gschwend, T. (2007). Ticket-splitting and strategic voting under mixed electoral rules: Evidence from Germany. *European Journal of Political Research* 46(1), 1–23.
- Gschwend, T., und T. Zittel (2012). Machen Wahlkreiskandidaten einen Unterschied? Die Persönlichkeitswahl als interaktiver Prozess. In R. Schmitt-Beck (Hrsg.), *Wählen in Deutschland*, S. 371–392. Baden-Baden: Nomos.
- Hainmueller, J., H. Kern und M. Bechtel (2006). Wahlkreisarbeit zahlt sich doppelt aus. Zur Wirkung des Amtsinhaberstatus einer Partei auf ihren Zweitstimmenanteil bei den Bundestagswahlen 1949 bis 1998. In T. Bräuninger und J. Behnke (Hrsg.), *Jahrbuch für Handlungs- und Entscheidungstheorie*, S. 11–45. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Hamlin, A., und C. Jennings (2011). Expressive political behaviour: Foundations, scope and implications. *British Journal of Political Science* 41, 645–670.
- Herrmann, M. (2012). Voter uncertainty and failure of Duverger's law: An empirical analysis. *Public Choice* 151(1), 63–90.
- Herrmann, M., und F. U. Pappi (2008). Strategic voting in German constituencies. *Electoral Studies* 27(2), 228–244.
- Herron, E. S., und M. Nishikawa (2001). Contamination effects and the number of parties in mixed superposition electoral systems. *Electoral Studies* 20(1), 63–86.
- Huber, Sascha (2012). Strukturen des politischen Kontexts und die demokratische Kompetenz der Wähler. Baden-Baden: Nomos.
- King, G. (1990). Electoral responsiveness and partisan bias in multiparty democracies. *Legislative Studies Quarterly* 15(2), 159–181.
- Kreps, D. M. (2013). *Microeconomic foundations. 1. Choice and competitive markets*. Princeton, NJ: Princeton University Press.
- Kurella, A.-S. (2016). Hätten Direktkandidaten der AfD über die 5%-Hürde verholfen? Eine Untersuchung des Kontaminationseffekts im Mischwahlsystem. In H. Schoen und B. Weßels (Hrsg.), *Wählen und Wähler: Analysen aus Anlass der Bundestagswahl 2013*, S. 202–222. Wiesbaden: Springer VS.
- Käppner, K., und S. Shikano (2015). A polytomous IRT unfolding model for the extraction of ideological space and valence factors from feeling thermometer ratings. Paper prepared for the MPSA conference, Chicago 2015.
- Luce, R. D., und J. W. Tuckey (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology* 1(1), 1–27.
- Manow, P. (2010). Dimensionen der Disproportionalität – Erststimmen und Direktmandate in den Bundestagswahlen von 1953 bis 2009. *Politische Vierteljahrsschrift* 51(3), 433–455.
- Markus, G. B., und P. E. Converse (1979). A dynamic simultaneous equation model of electoral choice. *The American Political Science Review* 73(4), 1055–1070.
- Moser, R. G., und E. Scheiner (2004). Mixed electoral systems and electoral system effects: controlled comparison and cross-national analysis. *Electoral Studies* 23(4), 575–599.
- Pappi, F. U. (2011). Party identification. In B. Badie, D. Berg-Schlosser und L. Morlino (Hrsg.), *International encyclopedia of political science*, Bd. 6, S. 1806–1811. Los Angeles: Sage.

- Pappi, F. U., und P. W. Thurner (2002). Electoral behaviour in a two-vote system: Incentives for ticket splitting in German Bundestag elections. *European Journal of Political Research* 41(2), 207–232.
- Plummer, M. (2003). Jags: A program for analysis of Bayesian graphical models using Gibbs sampling. <http://mcmc-jags.sourceforge.net/>.
- Rheault, L., A. Blais, J. Aldrich und T. Gschwend (2020). Understanding people's choice when they have two votes. *Journal of Elections, Public Opinion and Parties* 30(4), 466–483.
- Riker, W. H., und P. C. Ordeshook (1968). A theory of the calculus of voting. *The American Political Science Review* 62(1), 25–42.
- Schoen, H. (1999). Split-ticket voting in German federal elections, 1953–90: An example of sophisticated balloting? *Electoral Studies* 18(4), 473–496.
- Schofield, N., und I. Sened (2006). *Multiparty democracy: Elections and legislative politics*. Cambridge: Cambridge University Press.
- Schuessler, A. A. (2000). *A logic of expressive choice*. Princeton, NJ: Princeton University Press.
- Schüttmeyer, S. S., und R. Sturm (2005). Der Kandidat – das (fast) unbekannte Wesen: Befunde und Überlegungen zur Aufstellung der Bewerber zum Deutschen Bundestag. *Zeitschrift für Parlamentsfragen* 36(3), 539–553.
- Shikano, S. (2009). Simulating party competition and vote decision under mixed member electoral systems. *Czech Economic Review* 3(3), 270–291.
- Stokes, D. E. (1992). Valence politics. In D. Kavanagh (Hrsg.), *Electoral Politics*, S. 161–164. Oxford: Clarendon Press.
- Stratmann, T., und M. Baur (2002). Plurality rule, proportional representation, and the German Bundestag: How incentives to pork-barrel differ across electoral systems. *American Journal of Political Science* 46(3), 506–514.
- Theil, H. (1969). The desired political entropy. *American Political Science Review* 63(2), 521–525.

Anhang

A Amtliche Ergebnisse und Nachwahlbefragung

Tabelle 3: Erst- und Zweitstimmen bei der Bundestagswahl 2013 in den Wahlkreisen Stuttgart I und II (in Prozent)

Kandidat	Erststimmen			Partei	Zweitstimmen		
	Gesamt ^a	Wahllokal-Wähler ^b	Nachwahl-Befragung ^c		Gesamt ^a	Wahllokal-Wähler ^b	Nachwahl-Befragung ^c
Wahlkreis Stuttgart I							
Kaufmann	42,0	44,4	43,5	CDU	37,5	40,8	39,9
Vogt	16,6	18,7	21,0	SPD	21,0	24,7	26,0
Skudelny	2,6	2,3	1,3	FDP	8,3	8,2	7,3
Özdemir	27,5	29,9	30,0	Grüne	17,5	18,6	18,0
Frank	3,9	4,7	4,2	Linke	6,2	7,8	8,8
Sonstige	7,6			Sonstige	9,5		
N	151.248	93.076	1.130	N	151.500	90.932	1.083
Wahlkreis Stuttgart II							
Maag	43,8	47,4	46,8	CDU	39,3	43,6	43,2
Schäfstoß	26,1	30,4	33,0	SPD	22,9	27,0	28,1
Werwigk	2,7	2,4	1,2	FDP	6,7	6,8	5,3
Bender	13,9	13,8	14,1	Grüne	13,8	14,4	14,7
Aparicio	5,0	6,0	5,0	Linke	6,7	8,2	8,8
Sonstige	8,4			Sonstige	10,6		
N	132.835	87.872	740	N	133.041	85.593	702

Anmerkungen: a) Gültige Stimmen im Wahlkreis, b) Gültige Stimmen abzgl. Briefwahlstimmen, Prozentuierung nach Gesamtzahl der Stimmen für die fünf Kandidaten bzw. Listen, c) Befragte mit vollständigen Angaben (aus 1252 bzw. 839 Befragten).

Quelle: Landeshauptstadt Stuttgart 2013, S. 62-67 (Gesamt/Wahllokalwähler), eigene Berechnung (Nachwahlbefragung).

B Verwendung von Parteidistanzen und Partivalenzen als Stellvertreter für Kandidateneigenschaften

Um zu testen, ob die Wähler Parteidistanzen und Partivalenzen auf die lokalen Kandidaten der Parteien übertragen, so dass die vermuteten Effekte der Kandidateneigenschaften tatsächlich Effekte der Parteidistanzen sind, schätzen wir Modell 1 erneut, mit Parteidistanz (Modell 2) und Partivalenz (Modell 3) als Proxies für Kandidatendistanz bzw. -valenz. Die nachstehende Tabelle zeigt die Ergebnisse von multinomialen Logit-Modellen (d. h. Modelle, die Substitutionsmuster außer Acht lassen), angepasst an

die Daten von $N = 1268$ Befragten. Die Ergebnisse sind nicht schlüssig. Das Basismodell in der ersten Spalte (bei dem wir sowohl kandidaten- als auch parteispezifische Informationen verwenden, wie aus der Entfaltung der Bewertungsskalen gegeben) liefert Schätzungen für die β -Koeffizienten, die in Bezug auf Größe und Richtung denjenigen des bayesianischen gemischten Logit-Modells im Haupttext ziemlich ähnlich sind. Ein wesentlicher Unterschied zwischen den Modellen besteht in der Wirkung der Interaktion von Kandidatendistanz und aussichtsreicher Kandidatur, die im bayesianischen Modell höchstwahrscheinlich negativ ist, sich aber im hier dargestellten Basismodell nicht signifikant von Null unterscheidet.

Tabelle 4: Multinomiale Logit-Modelle der Wahlentscheidung bei Ersetzung der Kandidaten durch Parteieigenschaften

	(1) Basismodell		(2) Partei- statt Kandidatendistanz		(3) Partei- statt Kandidatenvalenz	
	β	SE	β	SE	β	SE
Räumliche Distanz: Kandidat	-0,966	0,087	-0,951	0,086	-0,999	0,088
Valenz: Kandidat	2,191	0,166	2,254	0,167	1,884	0,150
Räumliche Distanz: Partei	-0,695	0,064	-0,693	0,065	-0,694	0,064
Valenz: Partei	0,161	0,136	0,155	0,135	0,130	0,136
Parteiidentifikation: Kandidat	1,155	0,091	1,158	0,091	1,109	0,090
Parteiidentifikation: Partei	1,557	0,083	1,558	0,083	1,556	0,083
Distanz: Kandidat \times aussichtsreich	0,056	0,195	0,093	0,194	0,071	0,196
Valenz: Kandidat \times aussichtsreich	0,352	0,278	0,341	0,277	0,369	0,238
Kongruente Wahl	0,878	0,093	0,878	0,093	0,879	0,092
Log-likelihood	-1809,228		-1807,589		-1819,194	
N	1.268		1.268		1.268	

Wenn wir die Kandidatendistanz durch die Parteidistanz ersetzen (Modell 2), ändern sich die Ergebnisse wenig. Das Gleiche gilt für Modell 3, bei dem die Kandidaten- durch die Parteivalenz ersetzt wurde, wenngleich die Effektgröße der nun „falsch“ gemessenen Valenz geringer ist. Weiterhin kann man feststellen, dass die Log-likelihood der Modelle 1 und 2 kleiner ist als die von Modell 3. Insgesamt deutet dies darauf hin, dass die Parteivalenz ein suboptimaler Ersatz für die Kandidatenvalenz ist.