



Studienabschlussarbeiten

Fakultät für Mathematik, Informatik
und Statistik

Weiß, Andreas:

Wahlmanipulation in der Social Choice Theory:
Kritische Betrachtung und Simulation unter
verschiedenen Wahlszenarien

Masterarbeit, Sommersemester 2017

Fakultät für Mathematik, Informatik und Statistik

Ludwig-Maximilians-Universität München

<https://doi.org/10.5282/ubm/epub.41016>

Wahlmanipulation in der Social Choice Theory: Kritische Betrachtung und Simulation unter verschiedenen Wahlszenarien

**Abschlussarbeit zur Erlangung des akademischen Grades Master of
Science**

Andreas Weiß

Matrikelnr. xxxx

Sommersemester 2017

**Betreuung: Prof. Dr. Thomas Augustin, Christoph Jansen, Georg
Schollmeyer**

**Institut für Statistik: Lehrstuhl für Method(olog)ische Grundlagen der
Statistik und ihre Anwendungen
LMU München**



Zusammenfassung

Die vorliegende Masterarbeit gibt einen Überblick über verschiedene Arten des Wahlbetrugs im Kontext der Social Choice Theory (Manipulation, Bestechung, Kontrolle) und vergleicht diese anhand einer Simulationsstudie für 4 verschiedene Wahlregeln (Pluralität, Borda, Copeland, Single Transferable Vote). Dafür werden zuerst grundlegende Konzepte der Social Choice Theory eingeführt und erklärt. Anschließend werden Modelle erläutert (Impartial Culture, Impartial Anonymous Culture, Pólya-Eggenberger-Urnenmodell, räumliches Modell), mithilfe derer unterschiedliche Wahlszenarios generiert werden können. Zusätzlich werden Algorithmen und Heuristiken vorgestellt, mit denen eine Wahl effizient manipuliert werden kann. Die Simulation zeigte, dass Wahlen, die auf der Pluralitäts- oder auf der Bordaregel beruhen, leichter zu manipulieren waren, als wenn die STV- oder Copelandregel angewandt wurde. Zudem konnte beobachtet werden, dass sich Wahlen am schwersten manipulieren lassen, wenn ein räumliches Modell zugrunde liegt. Hingegen lieferte das verbreitete Impartial-Culture-Modell die einfachste Manipulierbarkeit aller Modelle.

Inhaltsverzeichnis

1	Einleitung	3
2	Social Choice Theory	3
2.1	Formalia und verschiedene Wahlsysteme	4
2.1.1	Scoring-Protokolle	6
2.1.2	Paarweise Vergleiche	10
2.1.3	Mehrstufige Wahlsysteme	12
2.2	Gütekriterien für Wahlregeln	14
2.3	Unmöglichkeitstheoreme	17
3	Modelle der Social Choice Theory	20
3.1	Stochastische Modelle	20
3.2	Räumliches Modell	24
4	Wahlbetrug	28
4.1	Manipulation	28
4.2	Bestechung	29
4.3	Kontrolle	30
4.4	Computational Social Choice	31
4.5	Algorithmen	32
5	Simulation	37
5.1	Ablauf und Konfigurationen	37
5.2	Implementation	41
5.3	Ergebnisse	43
5.3.1	Struktur der Profile	43
5.3.2	Ergebnisse I: Manipulation	47
5.3.3	Ergebnisse II: Bestechung	60
5.3.4	Ergebnisse III: Kontrolle	70
6	Fazit und Ausblick	73
7	Literaturverzeichnis	74
8	Elektronischer Anhang	77

1 Einleitung

In der heutigen Gesellschaft sind Wahlen selbstverständliche Mechanismen jeder demokratischen Abstimmung. Allgemeiner gesprochen, kommen sie immer dann zum Einsatz, wenn individuelle Vorlieben zu einer gesellschaftlichen Präferenz zusammengefasst werden sollen. Die sogenannte Social Choice Theory beschäftigt sich unter anderem damit, ob und wie diese Zusammenfassung der Wählerstimmen auf faire bzw. demokratische Weise möglich ist. Damit einher geht natürlicherweise, dass unerwünschte Formen von Wahlbetrug (das Abgeben einer unehrlichen Wahlstimme zum eigenen Vorteil) tunlichst unmöglich gemacht oder zumindest erschwert werden sollen. Ein nicht unbedeutender Teil der Literatur zur Social Choice Theory beschäftigt sich mit der (computationalen) Komplexität von Wahlen und Wahlbetrug (beispielsweise Faliszewski et al. (2006), sowie Bartholdi et al. (1989; 1992)). Ziel dieser vorliegenden Arbeit ist es, mithilfe einer Simulationsstudie die Schwierigkeit der Manipulation und Bestechung für verschiedene Wahlsysteme quantitativ zu überprüfen. Um den Aufbau der Simulationsstudie Schritt für Schritt zu erläutern, ist die vorliegende Arbeit wie folgt gegliedert: Zunächst soll in Kapitel 2 eine erweiterte Einführung in allgemeine Begriffe und Zusammenhänge der Social Choice Theory gegeben werden. Danach werden in Abschnitt 3 Modelle vorgestellt, die verwendet werden, um verschieden geartete Wähler- und Kandidatenmengen zu generieren. In Kapitel 4 werden anschließend unterschiedliche Arten des Wahlbetruges sowie Betrugsheuristiken und -algorithmen behandelt. Abschließend werden in Kapitel 5 diese Heuristiken und Algorithmen auf die zuvor generierten Wählerstimmen angewandt und schlussendlich die Ergebnisse der Simulation analysiert und interpretiert.

2 Social Choice Theory

Die *Social Choice Theory* (oder auch Sozialwahltheorie; SCT) ist ein Überbegriff für eine Vielzahl von Modellen und Resultaten, die sich mit der Aggregation individueller Präferenzen zu einer kollektiven Entscheidung befassen (vgl. List 2013, S. 1). Ihre Anfänge und Wurzeln liegen bereits im 18. Jahrhundert. Schon damals beschäftigten sich Philosophen und Mathematiker wie Marquis de Condorcet (1743-1794) oder Jean-Charles de Borda (1733-1799) mit Systematiken und Theorien hinter Wahlvorgängen.

Die Social Choice Theory findet im Allgemeinen auf beliebigen Situationen Anwendung, in denen eine Kollektiventscheidung auf Basis von mehreren (unterschiedlichen) individuellen Präferenzen getroffen werden soll.

Am anschaulichsten sind hierbei Wahlen zu nennen, da es seit Anbeginn der Demokratie im

Interesse der Menschheit ist, ein möglichst gerechtes und faires Wahlsystem bereitzustellen, das demokratische Grundzüge bestmöglich repräsentiert und den Willen der Wähler so gut wie möglich widerspiegelt.

Nachdem die Social Choice Theory für längere Zeit in Vergessenheit geriet, begannen sich Mitte des 20. Jahrhunderts auch Ökonomen mit ihr zu beschäftigen. Im Besonderen muss hier Kenneth Arrow (1921-2017) genannt werden, der auch als einer der ersten die Sozialwahltheorie mit einem mathematischen Grundgerüst ausstattete, und darauf basierend sein berühmtes Unmöglichkeitstheorem formulierte (siehe Abschnitt 2.3). Später erhielt Arrow den Wirtschaftsnobelpreis für seine Analysen zur Möglichkeit wirtschaftlicher Gleichgewichtszustände sowie seine Einführung und erfolgreiche Anwendung komplexer mathematischer Methoden in die Wirtschaftswissenschaft. Weitere bedeutende Beiträge zur Social Choice Theory im 20. Jahrhundert lieferten u.a. Duncan Black (1908-1991) und Amartya Sen (*1933).

Basierend auf dieser Formalisierung der Social Choice Theory konnten einige weitere berühmte Theoreme aufgestellt werden, z.B. von Gibbard und Satterthwaite (siehe 2.3).

Zunächst soll nun die SCT formal konkretisiert werden.

2.1 Formalia und verschiedene Wahlsysteme

Eine Sozialwahl (C, V) besteht im Allgemeinen aus einer Menge $C = \{c_1, \dots, c_m\}$ von m Kandidaten ($m \geq 2$) und einer Menge $V = \{1, \dots, n\}$ bestehend aus n Wählern, die jeweils ihre komplette Präferenzliste über alle Kandidaten angeben. Die Stimme von Wähler $i \in \{1, \dots, n\}$ wird als eine lineare Ordnung $\succsim_i \subset C \times C$ über C definiert.

Diese Ordnung ist

- **total**, $\forall a \neq b \in C : a \succsim_i b \vee b \succsim_i a$
- **transitiv**, $\forall a, b, c \in C : a \succsim_i b \wedge b \succsim_i c \Rightarrow a \succsim_i c$
- **antisymmetrisch**, $\forall a, b \in C : a \succsim_i b \wedge b \succsim_i a \Rightarrow a = b$
- **reflexiv**, $\forall a \in C : a \succsim_i a$.

Mit $\mathcal{L}(C)$ wird die Menge aller linearen Ordnungen über eine gegebene Kandidatenmenge C bezeichnet.

Definition 1 (Profil) *Ein Profil $P = (\succsim_1, \succsim_2, \dots, \succsim_n)$ spezifiziert eine lineare Ordnung als Wahlstimme für alle $i \in V$ mit $\succsim_i \in \mathcal{L}(C)$.*

Zusätzlich bezeichne $\mathcal{L}(C)^n$ die Menge aller möglichen Profile für ein gegebenes n . Außerdem sei $\mathcal{R}(C)$ die Menge aller schwachen Ordnungen über C (total, transitiv, reflexiv, aber nicht zwingend antisymmetrisch).

(vgl. Zwicker 2016, S. 26)

Eine alternative Darstellungsmöglichkeit eines Wählerprofils ist eine sogenannte Wahlsituation. Dafür reicht es anzugeben, wie oft jedes der $m!$ möglichen Rankings (Ordnungen) jeweils abgegeben wurde.

Definition 2 (Wahlsituation) *Eine Wahlsituation $S : \mathcal{L}(C) \rightarrow \mathbb{N}_0$ ist eine Funktion, die jeder möglichen linearen Ordnung der Kandidaten eine Anzahl (≥ 0) zuordnet (Zwicker 2016, S. 27).*

Ein Profil kann stets in eine Wahlsituation überführt werden, indem gezählt wird, wie oft jedes Ranking vorkommt. Die Umkehrung ist allerdings nur dann möglich, wenn die Wähler ununterscheidbar sind, d.h. wenn die Reihenfolge der Wähler im Profil keine Rolle spielt. Weiterhin werden Funktionen benötigt, die die individuellen Präferenzen geeignet aggregieren und anschließend ein Wahlergebnis des Kollektivs in Form einer nichtleeren Siegermenge oder einer kompletten gesellschaftlichen Rangliste bestimmen können.

Definition 3 (Social-Choice-Funktion) *Eine Social-Choice-Funktion (SCF) $f : \mathcal{L}(C)^n \rightarrow \mathcal{A}(C)$ ordnet jedem Präferenzprofil eine Gewinnermenge aus $\mathcal{A}(C)$ zu, wobei $\mathcal{A}(C) = 2^C \setminus \{\emptyset\}$ die Menge aller nichtleeren Teilmengen der Menge C ist.*

Definition 4 (Social-Welfare-Funktion) *Eine soziale Wohlfahrtsfunktion (SWF) ist eine Funktion $F : \mathcal{L}(C)^n \rightarrow \mathcal{R}(C)$, die ein Profil linearer Ordnungen auf eine (schwache) Ordnung R der Gesellschaft abbildet.*

Durch jede SWF F wird für ein Profil P auch eine SCF f_F induziert, mit der Festlegung:

$$f_F : \mathcal{L}(C)^n \rightarrow \mathcal{A}(C), P \mapsto \{x \in C \mid \forall y \in C : x F(P) y\} \quad (1)$$

Umgekehrt ist dies im Allgemeinen nicht der Fall.

(vgl. Zwicker 2016, S. 29).

Einige dieser Aggregations- und Wahlregeln werden im Folgenden näher erläutert und in verschiedene Gruppen unterteilt.

Diese sollen auch an einem durchgehenden Beispiel für eine Wahlsituation, das de Borda (1781, S. 660) nachempfunden ist, veranschaulicht werden (fortlaufend Beispiel 1 genannt): Es handelt sich um eine Wählerschaft bestehend aus 21 Personen und 3 Kandidaten (A,B,C). Dabei geben die Wähler die nachfolgenden Präferenzen an.

erste Präferenz	A	A	B	B	C	C
zweite Präferenz	B	C	A	C	A	B
dritte Präferenz	C	B	C	A	B	A
Präferenzordnung von x Personen	1	7	0	7	0	6

Tabelle 1: Beispielhafte Wahlsituation

2.1.1 Scoring-Protokolle

Ein sogenanntes *Scoring-Protokoll* zeichnet sich dadurch aus, dass jeder Wähler basierend auf seiner Präferenzliste durch einen Scoring-Vektor repräsentiert wird. Ein Scoring-Vektor ordnet jedem Kandidaten, abhängig von seiner Position auf der Präferenzliste, einen bestimmten Score zu.

Definition 5 (Scoring-Protokoll) *Ein Scoring-Protokoll ist eine Zuordnungsvorschrift $sp : \mathcal{L}(C) \times C \rightarrow \mathbb{R}$, mit der Eigenschaft:*

$$\forall \succ \in \mathcal{L}(C) \forall a, b \in C : a \succ b \Rightarrow sp(\succ, a) \geq sp(\succ, b).$$

Definition 6 (Scoring-Vektor) *Ein Scoring-Vektor sei eine von einem Scoring-Protokoll induzierte Abbildung $sv : \mathcal{L}(C) \rightarrow \mathbb{R}^m$, $sv_{\succ} := (sp(\succ, c_1), \dots, sp(\succ, c_m))$, die jeder linearen Ordnung einen reellen Vektor entsprechend des Protokolls zuordnet.*

Der Gesamtscore eines Kandidaten entsprechend eines festgelegten Scoring-Protokolls kann allgemein für $V = 1, \dots, n$ folgendermaßen bestimmt werden:

$$\forall j \in \{1, \dots, m\} : score_{sp}(c_j) = \sum_{i=1}^n sp(\succ_i, c_j) \quad (2)$$

Die Scoring-Protokolle für die einzelnen Wahlsysteme werden in den jeweiligen Abschnitten angegeben.

Exemplarisch wird dies nochmals konkret an Tabelle 2 verdeutlicht. Hier liegt eine Situation mit $n = 6$ Wählern und den 3 Kandidaten A, B, C vor.

	A	B	C
1	$sp(\succsim_1, A)$	$sp(\succsim_1, B)$	$sp(\succsim_1, C)$
2	$sp(\succsim_2, A)$	$sp(\succsim_2, B)$	$sp(\succsim_2, C)$
3	$sp(\succsim_3, A)$	$sp(\succsim_3, B)$	$sp(\succsim_3, C)$
4	$sp(\succsim_4, A)$	$sp(\succsim_4, B)$	$sp(\succsim_4, C)$
5	$sp(\succsim_5, A)$	$sp(\succsim_5, B)$	$sp(\succsim_5, C)$
6	$sp(\succsim_6, A)$	$sp(\succsim_6, B)$	$sp(\succsim_6, C)$
Summe	$\sum_{i=1}^6 sp(\succsim_i, A)$	$\sum_{i=1}^6 sp(\succsim_i, B)$	$\sum_{i=1}^6 sp(\succsim_i, C)$

Tabelle 2: Beispiel Scoring-Protokoll

Pluralitätsregel

Die wohl bekannteste und am meisten angewandte Wahlregel ist die Pluralitätswahl. Hierbei werden nur die Erststimmen der Wähler berücksichtigt, also ihre jeweils höchste Präferenz. Derjenige Kandidat, der die meisten Erststimmen auf sich vereinen kann, geht als Wahlsieger hervor. Der Scoring-Vektor für einen beliebigen Wähler hat die Form $s = (1, 0, \dots, 0)$, d.h. nur der erstplatzierte Kandidat der jeweiligen Präferenzliste bekommt einen Punkt zugewiesen, während die anderen Kandidaten leer ausgehen. Die Kandidaten werden abschließend nach der Summe ihrer Scores absteigend aufgelistet und derjenige mit den meisten Punkten wird zum Wahlsieger erklärt (vgl. Rothe et al. 2012, S. 124).

Das zugehörige Scoring-Protokoll sp_{pl} kann für einen beliebigen Kandidaten $a \in C$ folgendermaßen formalisiert werden:

$$sp_{pl} : \mathcal{L}(C) \times C \rightarrow \mathbb{R}$$

$$(\succsim, a) \mapsto \begin{cases} 1 & , \text{ falls } a \in \{b \in C \mid b \succsim c \forall c \in C\} \\ 0 & , \text{ sonst} \end{cases} \quad (3)$$

Der Gewinner der Wahl wird als w bezeichnet. Für ihn gilt $Score_{sp_{pl}}(w) \geq Score_{sp_{pl}}(y)$, $\forall y \in C$.

Im Beispiel 1 gewinnt A die Wahl mit 8 Stimmen vor B mit 7 und C mit 6 Stimmen. Dieses System ist aufgrund seiner Einfachheit sehr verbreitet. Diese bringt allerdings auch mit sich, dass nur sehr wenige der aus den individuellen Präferenzlisten zur Verfügung stehenden Informationen tatsächlich zur Entscheidungsfindung verwendet werden.

Veto-/Antipluralitätsregel

Wie der Name schon sagt, verhält sich diese Regel gegenteilig zur Pluralitätsregel. Die Wähler können eine Vetostimme abgeben, d.h. einem Kandidaten, von dem sie am wenigsten überzeugt sind, einen Nachteil verschaffen. Explizit passiert dies dadurch, dass jeder Wähler durch einen Scoring-Vektor der Form $s = (1, 1, \dots, 1, 0)$ repräsentiert wird. Es bekommen also alle Kandidaten außer dem Letztplatzierten einen Punkt zugewiesen. Wahlsieger ist wiederum derjenige Kandidat, der die meisten Punkte (oder in diesem Falle die wenigsten Gegenstimmen) sammeln konnte (vgl. Rothe et al. 2012, S. 124).

Formal ließe sich das Scoring-Protokoll für einen Kandidaten $a \in C$ schreiben als:

$$\begin{aligned} sp_{veto} : \mathcal{L}(C) \times C &\rightarrow \mathbb{R} \\ (\succsim, a) &\mapsto \begin{cases} 0 & , \text{ falls } a \in \{b \in C \mid c \succ b \forall c \in C\} \\ 1 & , \text{ sonst} \end{cases} \end{aligned} \quad (4)$$

Für den Wahlgewinner w gilt wiederum $Score_{sp_{veto}}(w) \geq Score_{sp_{veto}}(y), \forall y \in C$.

Im Beispiel 1 gewinnt Kandidat C mit 20 Punkten vor B mit 14 und A mit 8 Punkten.

k-Approval-Regel

Das k-Approval-Wahlsystem stellt eine Mischung aus Pluralitäts- und Vetoregel dar und verallgemeinert gleichzeitig beide Konzepte. Hierbei ist es den Wählern erlaubt, den ersten $k < m$ Einträgen ihrer Präferenzliste einen Punkt zu geben, während alle weiteren Kandidaten leer ausgehen. Der Kandidat mit dem höchsten Approval-Score, also den meisten Punkten, gewinnt die Wahl. Pluralitäts- und Vetoregel sind Spezialfälle der Approvalregel (für $k = 1$ bzw. $k = m - 1$).

Der Scoring-Vektor stellt sich wie folgt dar: $(\underbrace{1, \dots, 1}_k, 0, \dots, 0)$ (vgl. Rothe et al. 2012, S. 135f).

Das Scoring-Protokoll für Kandidat $a \in C$ lautet wie folgt:

$$\begin{aligned} sp_{appr} : \mathcal{L}(C) \times C &\rightarrow \mathbb{R} \\ (\succsim, a) &\mapsto \begin{cases} 1 & , \text{ falls } |\{b \in C \mid b \succ a\}| < k \\ 0 & , \text{ sonst} \end{cases} \end{aligned} \quad (5)$$

Im Beispiel 1 für 3 Kandidaten findet man sich für $k = 1$ bzw. $k = 2$ in den schon behandelten Spezialfällen Pluralitäts- bzw. Vetoregel wieder.

Borda-Wahl

Das nach dem einfürend genannten Jean-Charles de Borda benannte Wahlsystem ist das erste der betrachteten, das differenziertere Unterscheidungen bei der Zuweisung von Punkten an die Kandidaten, abhängig von ihrer Position auf einem individuellen Ranking, zulässt. Standardmäßig ist das Punktesystem bei einer Wahl mit m Kandidaten so vorgesehen, dass der Kandidat an der Spitze der Präferenzliste $m - 1$ Punkte erhält und die Punktzahl pro Platzierung um einen Punkt sinkt. Nur dem Letztplatzierten wird somit kein Punkt gutgeschrieben. Der Scoring-Vektor eines Wählers sieht also wie folgt aus: $(m - 1, m - 2, \dots, 1, 0)$ (vgl. Rothe et al. 2012, S. 125).

Das Scoring-Protokoll für Kandidat $a \in C$ lässt sich wie folgt darstellen:

$$\begin{aligned} sp_{borda} : \mathcal{L}(C) \times C &\rightarrow \mathbb{R} \\ (\succ, a) &\mapsto |\{b \in C \mid a \succ b\}| - 1 \end{aligned} \quad (6)$$

Im Beispiel gewinnt nach dieser Vorschrift C mit 6 Erststimmen und 14 Zweitstimmen und damit 26 Punkten vor B mit 21 (7 Erst-, 7 Zweitstimmen) und A mit 16 Punkten (8 Erststimmen).

Alternativ kann die Bordawahl auch über die sogenannte Nettopräferenz zweier Alternativen entsprechend eines Profils P (siehe Gleichung 7) motiviert werden:

$$a, b \in C : Net_P(a > b) = |\{i \in V \mid a \succ_i b\}| - |\{i \in V \mid b \succ_i a\}| \quad (7)$$

Der sogenannte symmetrische Borda-Score $Borda_P^{sym}(x)$ eines Kandidaten $a \in C$ kann dann ermittelt werden durch

$$Borda_P^{sym}(a) = \sum_{b \in C} Net_P(a > b). \quad (8)$$

Aus Gleichung 8 ergibt sich ein Scoring-Vektor der Form $(m - 1, m - 3, \dots, -(m - 1))$. Dieser kann aber durch Transformation auf die ursprüngliche, geläufige Form $(m - 1, m - 2, \dots, 0)$ – den asymmetrischen Bordascore – gebracht werden durch

$$Borda_P^{asym}(a) = n + \frac{1}{2} Borda_P^{sym}(a). \quad (9)$$

(vgl. Zwicker 2016, S. 28f)

Der Wahlsieger w wird bestimmt als $w : Borda_P^{asym}(w) = \max_{b \in C} Borda_P^{asym}(b)$.

2.1.2 Paarweise Vergleiche

Eine andere Herangehensweise ist es, statt Kandidaten nach bestimmten Scores zu rangieren, alle Kandidaten paarweise miteinander zu vergleichen. Den direkten Vergleich gewinnt derjenige Kandidat $a \in C$, den mehr Wähler einem anderen Kandidaten $b \in C$ vorziehen:

$$a >^{\mu} b, \text{ wenn } Net_P(a > b) > 0, \forall a, b \in C. \quad (10)$$

(vgl. Zwicker 2016, S. 27f)

Diese Vorgehensweise kann im Vergleich zu Scoring-Protokollen zu gravierenden Änderungen im Gesamtergebnis führen.¹

Condorcet-Wahlsystem

Die wohl einfachste, aber auch eher beschränkte Möglichkeit, einen Wahlsieger basierend auf paarweisen Vergleichen zu bestimmen, ist das Condorcet-System. Benannt nach seinem Schöpfer, spricht es einem Kandidaten genau dann den Wahlsieg zu, wenn er alle anderen Kandidaten im paarweisen Vergleich schlagen kann. Gibt es allerdings keinen sogenannten Condorcet-Sieger, kann nicht ohne weitere Maßnahmen ein Wahlsieger bestimmt werden (vgl. Rothe et al. 2012, S. 127f).

Ein Condorcet-Gewinner für ein Profil P ist eine Alternative $a \in C$, die jede Alternative im paarweisen Vergleich besiegt: $a >^{\mu} b \forall b \in C, a \neq b$. Die Condorcet-Regel erklärt den Condorcet-Sieger zum Gesamtsieger, und ist nicht definiert, wenn kein solcher existiert (vgl. Zwicker 2016, S. 34).

Der Condorcet-Sieger aus Beispiel 1 ist C , da er die direkten Vergleiche gegen A und B jeweils mit 13:8 Stimmen für sich entscheiden kann. Außerdem gewinnt B den Paarvergleich gegen A mit ebenfalls 13:8 Stimmen. Hier tritt also das Borda-Paradoxon auf, da der Condorcet-Verlierer A siegreich aus der Pluralitätswahl hervorging.

Copeland/Llull

Eine abgewandelte Form des Paarvergleichs stellt die Methode nach Copeland bzw. Llull dar. Dabei wird pro Gewinn eines direkten Vergleichs dem siegreichen Kandidaten ein Punkt gutgeschrieben. Somit vereint die Regel die Konzepte des paarweisen Vergleichs mit denen eines Scoring-Protokolls. Endet ein Vergleich unentschieden, so bekommen beide Kandida-

¹Jean Charles de Borda kam die Idee zu seinem Wahlsystem erst, als er entdeckte, dass Fälle auftreten können, in denen Kandidaten zum Sieger einer Pluralitätswahl erklärt werden, obwohl sie die paarweisen Vergleiche gegen alle anderen Kandidaten verlieren („Borda-Paradoxon“, vgl. Rothe et al. 2012, S. 150f).

ten $0 \leq \alpha \leq 1$ Punkte (allgemein *Copeland* $^\alpha$ -System) zugewiesen. Im ursprünglichen, von Copeland vorgeschlagenen Wahlsystem ist $\alpha = \frac{1}{2}$, in der Version von Llull gilt $\alpha = 1$. Als Gewinner der Wahl geht wiederum der Kandidat mit den meisten Punkten hervor (vgl. Rothe et al. 2012, S. 128ff).

In einer anderen Version des Copeland-Systems (beschrieben von Zwicker (2016, S. 28)) wird der Score gebildet aus der Anzahl der gewonnenen direkten Vergleiche abzüglich der verlorenen Vergleiche (siehe Gleichung 7). Für einen unentschiedenen Vergleich würde dann kein Kandidat Punkte erhalten bzw. abgezogen bekommen.

Der Copeland-Score eines Kandidaten $a \in C$ kann wie folgt ermittelt werden:

$$\text{Copeland}(a) = |\{b \in C | a \succ^\mu b\}| - |\{b \in C | b \succ^\mu a\}| \quad (11)$$

Die letztgenannte Variante wird auch in den Simulationen in Abschnitt 5 verwendet.

Da es im Beispiel 1 keine unentschiedenen paarweisen Vergleiche gibt – weil die Wählerzahl ungerade ist –, gewinnt C (Punktzahl in der Variante von Zwicker: 2) die Wahl vor B (0) und A (-2). In Tabelle 3 ist ein weiteres Wahlbeispiel zu sehen, bei dem der Ausgang jedes paarweisen Vergleichs in Matrixform dargestellt ist. Die erste Zeile entspricht hier exemplarisch dem Wahlgewinner A , den beispielsweise 77 von 100 Wählern vor B ranken und der insgesamt einen Score von 4 erreicht, da er alle direkten Vergleiche für sich entscheiden kann.

	A	B	C	D	E	score
A	–	77	96	52	70	4
D	48	74	83	–	60	2
E	30	64	78	40	–	0
B	23	–	71	26	36	-2
C	4	29	–	17	22	-4

Tabelle 3: Darstellung einer paarweisen Wahl inklusive Copeland-Score

Dodgson

Auch das Wahlsystem nach Dodgson, der auch und vor allem unter dem Pseudonym Lewis Carroll (Autor von u.a. „Alice im Wunderland“) bekannt ist, ist erwähnenswert und beruht ebenfalls auf paarweisen Vergleichen. Hierbei wird betrachtet, wie viele Vertauschungen zweier benachbarter Kandidaten in individuellen Präferenzlisten mindestens vorgenommen werden müssten, sodass ein ausgewählter Kandidat zum Condorcet-Sieger werden würde. Derjenige Kandidat, der am wenigsten Vertauschungen benötigt, geht als Sieger aus der Wahl hervor. Das bedeutet offensichtlich auch, dass, wenn von Anfang an ein Condorcet-

Sieger existiert, dieser auch die Wahl nach Dodgson auf jeden Fall gewinnt (vgl. Rothe et al. 2012, S. 130ff).

Daher gewinnt C als Condorcet-Sieger im Beispiel 1 auch die Dodgson-Wahl vor B und A , für die 3 bzw. 6 Vertauschungen nötig wären.

2.1.3 Mehrstufige Wahlsysteme

Im Folgenden werden Wahlsysteme beschrieben, die im Gegensatz zu den vorhergehenden nicht direkt eine Entscheidung treffen, sondern (potenziell) mehrere Durchgänge für eine Wahlentscheidung benötigen.

Pluralitätswahl mit Stichwahl

Hierbei wird die normale Pluralitätswahl erweitert, sodass jeweils die 2 Kandidaten mit den meisten Erststimmen im ersten Wahldurchgang nochmal zu einer Stichwahl antreten müssen (sofern nicht der erstplatzierte Kandidat bereits eine absolute Mehrheit erringen konnte; in diesem Falle würde er die Stichwahl aber in jedem Fall auch gewinnen). Durch die Stichwahl können mehr als nur die Informationen über die Erstplatzierten der Präferenzlisten für die endgültige Wahlentscheidung verwendet werden, da Wähler, deren erstplatzierte Kandidat bereits vor der Stichwahl ausgeschieden ist, nochmal anhand ihrer weiteren Liste eine neue Stimme vergeben müssen. Zusätzlich garantiert eine Stichwahl eine absolute Mehrheit für einen Kandidaten (außer beide Kandidaten können genau gleich viele Stimmen auf sich vereinen) (vgl. Rothe et al. 2012, S. 137).

Im Beispiel 1 ginge B als Sieger der Wahl hervor, nachdem A zuvor in der Pluralitätswahl mit 8 Stimmen keine absolute Mehrheit erreichen konnte. B schlägt A in der Stichwahl mit 13:8 Stimmen.

Single Transferable Vote (Übertragbare Einzelstimmgebung)

Die STV-Methode (oder auch Instant-Runoff-Wahlsystem) bestimmt den Wahlsieger in maximal $m - 1$ Runden (bei m zur Wahl stehenden Kandidaten). Sobald ein Kandidat eine absolute Mehrheit der Erststimmen erreicht, geht er als Gewinner der Wahl hervor. Bekommt in einem Durchgang keiner der Kandidaten eine absolute Mehrheit, so wird derjenige mit den wenigsten Erststimmen von den weiteren Runden ausgeschlossen und scheidet somit aus. Wähler, die dem ausgeschlossenen Kandidaten ihre Erststimme gegeben haben, übertragen ihre Erststimme dann an den davor Zweitplatzierten in ihrer Rangliste. Diese Prozedur wird so lange wiederholt, bis ein Kandidat die absolute Mehrheit der Erststimmen erreicht. Für 3 Kandidaten gleicht STV der Pluralitätswahl mit Stichwahl (vgl. Rothe et al. 2012, S.

138).

Man startet also mit der kompletten Kandidatenmenge $C_1 := C$ und führt folgende zwei Schritte (potenziell mehrmals) durch:

für $k = 1, \dots, m - 1$

1. berechne $Score_{sp_{pl}}(a) = \sum_{i=1}^n sp_{pl}(\succ_i, a)$ (siehe Formel 3 und 2) für jeden Kandidaten $a \in C_k$. Wenn $\exists a \in C_k : Score_{sp_{pl}}(a) > \lceil \frac{n}{2} \rceil$ so ist die Wahl nach der k-ten Runde abgeschlossen und a wird zum Sieger erklärt ($\lceil \cdot \rceil$ beschreibt das Aufrunden auf die nächste ganze Zahl).
2. Ist dies nicht der Fall, so sucht man $b \in C_k : Score_{sp_{pl}}(b) = \min_{a \in C_k} Score_{sp_{pl}}(a)$ und bildet $C_{k+1} := C_k \setminus \{b\}$. Mit der Teilmenge $C_{k+1} \subset C$ beginnt man dann wieder bei Schritt 1.

Der Ausgang der Wahl stimmt im Beispiel 1 mit dem der Pluralitätswahl mit Stichwahl überein und findet daher nach 2 Runden einen Sieger.

Bucklin

Das Wahlsystem nach Bucklin sucht ebenfalls in potenziell mehreren Runden nach Kandidaten mit absoluter Stimmenmehrheit. Dabei stellt $maj(V) = \lfloor |V|/2 \rfloor + 1$ den Schwellenwert zur absoluten Mehrheit dar, wobei $\lfloor \cdot \rfloor$ das Abrunden auf die nächste ganze Zahl beschreibt. In der ersten Runde wird überprüft, ob ein Kandidat die Mehrheit aller Erststimmen erreicht hat, wenn nicht, geht man weiter in die zweite Runde. Hier wird geprüft, ob ein Kandidat (oder mehrere) den Schwellenwert $maj(V)$ in den addierten Erst- und Zweitstimmen erreichen kann. Wenn ja, geht der einzelne Kandidat oder derjenige, der den Schwellenwert am deutlichsten überschreitet, als Sieger hervor. Wenn nein, wird analog weiter verfahren, indem Drittstimme, etc. zusätzlich betrachtet werden (vgl. Rothe et al. 2012, S. 140ff).

Im Beispiel 1 wird C zum Wahlsieger erklärt, nachdem in der zweiten Wahlrunde sowohl C als auch B $maj(V) = 11$ überschreiten, aber der Score von C (20) größer als der von B (14) ist. A überschreitet in keiner der beiden Runden den Schwellenwert und wird damit Drittplatzierter.

Es zeigt sich, dass es keine unbedeutende Rolle spielt, welches Wahlsystem verwendet wird. Saari (2001, S. 13f) warnt sogar davor, er könne, nachdem er die individuellen Präferenzen der Wähler kennenlerne, die Wahl einer jeden Organisation in Richtung eines gewünschten Kandidaten lenken bzw. diesen sogar zum Sieger machen, indem er nur ein passendes, offenbar demokratisches Wahlsystem auswählt. Obwohl diese Warnung nicht komplett ernst

genommen werden sollte, zeigt sie doch ebenfalls, wie abhängig eine Wahl vom Wahlsystem ist.

2.2 Gütekriterien für Wahlregeln

Im Folgenden werden verschiedene Gütekriterien vorgestellt, die ein zentrales Gerüst von Fairness und Gerechtigkeit um ein Wahlsystem konstruieren sollen. Genauer gesagt verlangen die Gütekriterien bestimmte Eigenschaften von SWFs oder SCFs.

Nicht-Diktatur

Das wohl wichtigste Kriterium ist die Nicht-Diktatur, da ein Diktator jedem demokratischen Grundgedanken widerstrebt. Es darf also keinen einzelnen Wähler geben, dessen Präferenzen automatisch zum kollektiven Wahlergebnis werden. Eine SWF F ist diktatorisch, wenn ein Wähler i als Diktator fungiert, d.h. $F(P)$ ist für jedes Profil P das Ranking von i .

$$\nexists i \in V : \forall P \in \mathcal{L}(C)^n : F(P) = \succsim_i$$

Nicht-Diktatur ist eine schwache Form von Anonymität. Eine SWF F ist anonym, wenn die Rollen jedes Paares von Wählern austauschbar sind:

$F(P) = F(P^*)$, wenn man P^* aus P erhält, indem man die Stimmen von beliebigen Wählern i und j vertauscht: $\succsim_i^* = \succsim_j$; $\succsim_j^* = \succsim_i$; $\succsim_k^* = \succsim_k \forall k \in V \setminus \{i, j\}$ (vgl. Zwicker 2016, S. 31).

Pareto-Konsistenz

Die Pareto-Konsistenz definiert eine Einstimmigkeitseigenschaft: Wenn alle Wähler eine bestimmte Alternative a einer anderen Alternative b vorziehen, so sollte auch auf der Kollektivebene a vor bzw. auf keinen Fall hinter b stehen (siehe French 1986, S. 286).

Wenn also $\forall i \in V : a \succsim_i b$ gilt, dann soll auch $a \succsim_R b$ gelten ($a, b \in C$), mit der Gesellschaftsordnung \succsim_R .

Unabhängigkeit von irrelevanten Alternativen

Die Unabhängigkeit von irrelevanten Alternativen (Independence of irrelevant Alternatives/ IIA) ist im Gegensatz zu den beiden zuvor genannten Kriterien ein viel diskutiertes Gütekriterium. Es besagt, dass beim Vergleich zweier Alternativen auf Gemeinschaftsebene nur das Verhältnis dieser beiden Alternativen auf Individualebene betrachtet werden soll. Anders gesagt soll es keine Rolle spielen, wie die Beziehung zu anderen (dritten/irrelevanten) Alternativen ist.

Eine SWF ist unabhängig von irrelevanten Alternativen, wenn für zwei Profile $P, P' \in \mathcal{L}(C)^n$ und zwei Alternativen $a, b \in C$ gilt: für alle i gilt entweder $a \succsim_i b \Leftrightarrow a \succsim'_i b$ oder $b \succsim_i a \Leftrightarrow b \succsim'_i a$. Dann gilt auch $a \succsim_R b \Leftrightarrow a \succsim'_R b$ oder $b \succsim_R a \Leftrightarrow b \succsim'_R a$, wobei \succsim_R bzw. \succsim'_R den von $F(P)$ bzw. $F(P')$ definierten Relationen der (schwachen) Gesellschaftsordnungen entsprechen (siehe French 1986, S. 285).

Die Bordawahl verstößt beispielsweise gegen dieses Kriterium, da es im Allgemeinen nicht möglich ist, zwei Kandidaten zu vergleichen, ohne andere Alternativen mit einzubeziehen. Vielmehr hängt die Punktzahl der beiden betrachteten Kandidaten direkt davon ab, ob bzw. wie viele andere Kandidaten dazwischen eingereicht werden.

Strategiesicherheit

Einen zentralen Punkt dieser Arbeit wird im weiteren Verlauf das Kriterium der Strategiesicherheit darstellen. Die Frage, die sich hier stellt, ist, ob die Möglichkeit besteht, durch gezielte Manipulation, also der Abgabe einer unehrlichen Stimme, einen Vorteil zu ziehen. Da in dieser Arbeit Wahlen betrachtet werden, bei denen nur *ein* Gewinner gewählt werden soll (im Gegensatz z.B. zu einem ganzen Gremium) muss das Wahlergebnis nicht zwingend ein komplettes gesellschaftliches Ranking liefern, sondern es ist ausreichend, Strategiesicherheit für Social Choice Funktionen zu betrachten (einzelne Gewinner ergeben sich als Spezialfall, indem nur einelementige Siegermengen zugelassen werden).

Eine SCF f heißt manipulierbar (durch einen Wähler) wenn für ein Paar $P, P' \in \mathcal{L}(C)^n$ von Profilen, und einen unehrlichen Wähler $i \in V$ gilt: $\succsim'_j = \succsim_j \forall j \in V \setminus \{i\}$ und $f(P') \succsim_i f(P)$, d.h. Wähler i präferiert den Gewinner der manipulierten Wahl gegenüber dem der Wahl mit ehrlichen Präferenzen, nachdem er als einziger die Möglichkeit hatte, eine andere (unehrliche) Präferenzordnung als im ursprünglichen Profil anzugeben (vgl. Zwicker 2016, S. 30).

Im Zusammenhang mit SCFs ist auch von optimistischer und pessimistischer Manipulierbarkeit die Rede. Man spricht von optimistischer Manipulierbarkeit, wenn ein Profil $P = (\succsim_1, \dots, \succsim_n)$ und eine unehrliche Wahlstimme \succsim'_i des Wählers i existieren, sodass die Gewinner des Profils $P' = (\succsim_1, \dots, \succsim_{i-1}, \succsim'_i, \succsim_{i+1}, \dots, \succsim_n)$ von Wähler i (bzgl. seiner wahren Präferenzen \succsim_i) gegenüber *allen* Gewinnern des Profils P präferiert werden. Bei pessimistischer Manipulation reicht es hingegen, wenn die „neuen“ Gewinner gegenüber *mindestens einem* „alten“ Gewinner von Wähler i präferiert werden (vgl. Taylor 2002, S. 324).

Eine SCF f heißt strategiesicher, wenn sie nicht manipulierbar durch einen Wähler ist.

Analog kann auch eine Teilmenge $M \subset V$, bestehend aus $0 < h < n$ Wählern, als manipulative Koalition definiert werden. Dabei soll gelten, dass $\succsim_k = \succsim_l \forall k, l \in M$ ist, d.h. alle

Mitglieder der Manipulatoren haben ursprünglich die selbe ehrliche Präferenzordnung. Nun kann wiederum ein Paar P, P' von Profilen definiert werden mit $\succsim'_j = \succsim_j \forall j \in V \setminus M$, sodass $f(P') \succsim_M f(P)$ gilt, wobei \succsim_M der ehrlichen Präferenzordnung jedes Mitglieds der manipulativen Koalition entsprechen soll. Zusätzlich sei noch angemerkt, dass $\succsim'_k = \succsim'_l \forall k, l \in M$ im Allgemeinen nicht gelten muss, d.h. die Manipulatoren können durchaus verschiedene unehrliche Rankings abgeben.

Neutralität

Wird ein Wahlsystem als neutral bezeichnet, so bedeutet dies, dass alle Kandidaten bei der Auswertung gleich behandelt werden. Das heißt, dass für jeden Kandidaten nur die Anzahl seiner erhaltenen Stimmen bzw. sein Score ausschlaggebend sein soll und keine weiteren Aspekte. Eine SWF F ist neutral, wenn jedes Paar von Alternativen auf folgende Weise austauschbar ist: kann ein Profil P^* aus P erhalten werden ($P, P^* \in \mathcal{L}(C)^n$), indem man zwei Alternativen $a, b \in C$ in jeder Wählerstimme vertauscht, dann kann $F(P^*)$ ebenso durch eine Vertauschung von a und b im Gesellschaftsranking erhalten werden (Zwicker 2016, S. 31f).

Neutralität kann zusammen mit Anonymität als eine Art Symmetrieeigenschaft von Wahlsystemen angesehen werden (vgl. Kelly 1993, S. 169ff).

Resoluthheit

Diese Eigenschaft fordert von einem Wahlsystem, dass stets nur genau ein Gewinner ausgewählt wird. Mehrere oder auch gar kein Gewinner sollen nicht zugelassen werden (siehe Rothe et al. 2012, S. 155).

Für SWFs kann Resoluthheit so definiert werden, dass der Erstplatzierte in der Gesellschaftsordnung stets eindeutig dem oder den Zweitplatzierten vorgezogen wird (keine Indifferenz, und wenn doch, muss ein Mechanismus definiert werden, der die Unentschiedenheit auflöst).

Souveränität der Bürger

Dieses Kriterium beschreibt die Möglichkeit der Bürger bzw. Wähler, frei nach ihrem Willen wählen zu können. Eine SWF F heißt auferlegt (imposed), wenn es ein Paar $a, b \in C$ von Alternativen gibt ($a \neq b$), sodass $a \succsim_R b$ für jedes Profil $P \in \mathcal{L}(C)^n$ gilt, wobei \succsim_R der Gesellschaftsordnung entspricht. Das heißt also, dass von vornherein keine Präferenzen zwischen beliebigen Alternativen ausgeschlossen oder verboten werden sollten. Stattdessen sollten die Gesellschaftspräferenzen nur und ausschließlich von den Präferenzen der Wähler abhängen. Ist die SWF nicht auferlegt, gilt das Kriterium der Souveränität der Bürger als

erfüllt (vgl. Arrow 1963, S. 28f).

2.3 Unmöglichkeitstheoreme

Basierend auf den im vorangegangenen Abschnitt vorgestellten Gütekriterien wurden Anstrengungen unternommen, Wahlsysteme zu finden, die möglichst viele der Kriterien erfüllen und damit den Wahlsieger oder die Gesellschaftspräferenzordnung so fair wie möglich ermitteln. Im Zuge dessen ergaben sich allerdings erstaunliche Unvereinbarkeiten, was vor allem in den folgenden beiden berühmten Unmöglichkeitstheoremen gipfelte.

Arrow

Das wohl berühmteste Unmöglichkeitstheorem wurde von Kenneth Arrow (1963; von ihm selbst als „General Possibility Theorem“ benannt) aufgestellt und bewiesen. Es besagt, dass es, sofern mindestens drei Kandidaten ($m \geq 3$) zur Wahl stehen, kein präferenzbasiertes² Wahlsystem (bzw. keine SWF) gibt, das gleichzeitig die Gütekriterien Pareto-Konsistenz³, Unabhängigkeit von irrelevanten Alternativen und Nicht-Diktatur erfüllt (Rothe et al. 2012, S. 154).

Das bedeutet allerdings nicht, dass bei einer gegebenen Wahlregel (die ordinale Präferenzen als Input erhält) für jedes erdenkliche Profil von individuellen Präferenzen ein Kriterium verletzt wird, sondern lediglich, dass sich für jede Regel mindestens ein Profil finden lässt, sodass gegen ein Kriterium verstoßen wird. In seiner ursprünglichen Form betrachtet das Theorem soziale Wohlfahrtsfunktionen und es wird gefordert, dass deren Bild – also das gesellschaftliche Ranking – zusätzlich transitiv und vollständig sein soll. Alle in Abschnitt 2.1 genannten Wahlsysteme erfüllen die Kriterien Nicht-Diktatur und Pareto-Konsistenz. Das bedeutet, dass sie (nach Arrow) das Kriterium der Unabhängigkeit von irrelevanten Alternativen (in mindestens einem Profil) verletzen müssen (vgl. Rothe et al. 2012, S. 155).

Eine Ausnahme stellen die auf paarweisen Vergleichen beruhenden Systeme dar. Diese verletzen allerdings die Annahme der „universal domain“ (also die Annahme, dass die SWF jedem

²Genauer gesagt beruhend auf *ordinalen* Präferenzen, d.h. es können nur qualitative Aussagen über die Präferenzen getroffen werden (z.B. a wird gegenüber b präferiert) aber keine quantitativen Aussagen (z.B. die Präferenz von a gegenüber b ist doppelt so stark wie die Präferenz von c gegenüber d).

³Arrow selbst (1963, S. 59) drückte die Pareto-Konsistenz mithilfe der zwei Kriterien Souveränität der Bürger und Monotonie aus. Monotonie bedeutet hierbei, dass für zwei Profile P und P' und daraus resultierenden schwachen Ordnungen R und R' , wobei für jeden Wähler i gilt: für Alternativen x' und y' ungleich einer gegebenen Alternative x : $x' \succsim'_i y' \Leftrightarrow x' \succsim_i y'$; für alle $y' : x \succsim_i y' \Rightarrow x \succsim'_i y'$. Dann gilt: $x \succsim_R y \Rightarrow x \succsim_{R'} y$, d.h. x darf nicht sozial schlechter gestellt werden, wenn es individuell gleich gut oder besser gestellt wird.

Profil, das als Input verwendet wird, eine transitive und vollständige Ordnung zuordnet), da hier Fälle auftreten können, bei denen die gesellschaftliche Ordnung intransitive Zyklen enthält.

Für das Theorem gibt es zahlreiche Beweise, darunter u.a. von Arrow (1963) selbst, drei kurze Beweise von Geanakoplos (2005) oder von French (1986, S. 286ff).

Trotzdem, oder gerade deswegen, gibt es diverse Ansätze, das Theorem zu umgehen, indem die Annahmen gelockert werden sollen, von denen im Folgenden einige kurz umrissen werden.

Die Axiome Nicht-Diktatur und Pareto-Konsistenz bleiben weitgehend von Kritik unberührt, da man einen Diktator sowie eine Verletzung des Einstimmigkeitsprinzips nicht mit demokratischen Konzepten vereinigen kann.

Kontroverser diskutiert werden vor allem die Annahmen eines universellen Definitionsbereichs der SWF sowie die Unabhängigkeit von irrelevanten Alternativen. Bei ersterer Annahme fragt man sich, ob es wirklich notwendig ist, alle möglichen individuellen Präferenzordnungen zuzulassen, unabhängig davon, wie unrealistisch sie auch sein mögen. Einen möglichen Ausweg aus der Unmöglichkeit bietet eine Einschränkung auf sogenannte *single-peaked* Präferenzen. Hierbei wird davon ausgegangen, dass sich die zur Wahl stehenden Alternativen gewissermaßen in eine natürliche (also nicht präferenzabhängige) Ordnung bringen lassen. Als Beispiel wäre eine Sortierung von Kandidaten anhand ihrer politischen Ausrichtung von links nach rechts im politischen Spektrum denkbar. Darauf basierend bieten *Single-Peaked*-Präferenzen eine sinnvolle Interpretation und decken den größten Teil aller realistischen Präferenzordnungen ab. Formell bedeutet das, dass in jedem Tripel von Alternativen (a_1, a_2, a_3) eine der drei Alternativen von keinem Individuum auf dem 3. Platz gerankt wird⁴. Es lässt sich außerdem zumindest für den Fall einer ungeraden Anzahl von Wählern zeigen, dass eine paarweise Wahl beruhend auf *Single-Peaked*-Präferenzen alle sonstigen Axiome erfüllt (Sen 1966, S. 495f).

Außerdem wird vielfach diskutiert, ob für den Vergleich zweier Alternativen tatsächlich eine Unabhängigkeit von anderen Alternativen gegeben sein muss bzw. ob diese überhaupt als irrelevant angesehen werden können. Vielmehr wird argumentiert, dass es durchaus von Wichtigkeit ist, ob andere Alternativen zur Wahl stehen (und welche das sind). Im Speziellen ermöglicht das Kriterium bis auf weiteres keine Berücksichtigung von Präferenzstärke

⁴Sen (1966) verallgemeinert *Single-Peaked*-Präferenzen und die konzeptionell ähnlichen *Single-Caved*-Präferenzen (für jedes Tripel von Alternativen wird eine Alternative von keinem Individuum vor den beiden anderen präferiert) sowie das Analogon für Zweitplatzierungen („Separabilität in zwei Gruppen für jedes Tripel“) zu sogenannten „triple-wise value restrictions“ als geeignete Einschränkung des Wertebereichs der SWF, d.h. dadurch können zyklische Gesellschaftspräferenzen ausgeschlossen werden.

(kardinalen Präferenzen) im Aggregationsprozess (French 1986, S. 293).

Es lässt sich allerdings zeigen, dass mit der Einführung von sogenannten Value-Difference-Funktionen $v_i : \mathcal{L}(C) \rightarrow \mathbb{R}^m$, die einer linearen Ordnung einen reellen Vektor zuordnen (damit wird es den Wählern ermöglicht, ihre Präferenzen kardinal auszudrücken), und einer leicht veränderten Arrowschen Axiomatik⁵ die Unmöglichkeit umgangen werden kann. Da allerdings die Value-Difference-Funktionen nur bis auf positiv affine Abbildungen eindeutig sind, bleibt die Frage offen, wie kardinale Präferenzen verschiedener Individuen vergleichbar gemacht werden können (siehe French 1986, S. 294ff).

Gibbard-Satterthwaite

Das folgende Theorem wurde 1973 von Allan Gibbard und 1975 von Mark Satterthwaite unabhängig voneinander bewiesen und besagt folgendes:

Für eine Wahl (C, V) mit $|V| \geq 3$ gibt es kein präferenzbasiertes Wahlsystem (keine SCF), das die Eigenschaften Nicht-Diktatur, Resoltheit, Souveränität der Bürger und Strategiesicherheit erfüllt (Rothe et al. 2012, S. 156).

Das bedeutet somit anschaulich gesprochen, dass alle Wahlsysteme, die strikt immer nur genau einen Wahlsieger ermitteln und dabei nicht diktatorisch vorgehen, sondern alle Macht den Wählern überlassen, nicht strategiesicher und damit manipulierbar (in mindestens einem Profil) sind. Ein weiterer Beweis des Theorems, neben denen der namensgebenden Autoren, ist in Svensson (1999) zu finden. Des Weiteren liefern Duggan und Schwartz (2000) eine Erweiterung des Gibbard-Satterthwaite-Theorems, das auch nichtleere Mengen von Siegern anstatt nur einen einzigen Gewinner zulässt. Wenn eine anonyme und nicht-auferlegte (also im Sinne von Souveränität der Bürger definierte) SCF für mindestens drei Alternativen nicht (optimistisch oder pessimistisch) manipulierbar ist, dann ist jeder Wähler ein Nominator, d.h. die Siegermenge besteht aus allen Alternativen, die mindestens bei einem Wähler an erster Stelle der Präferenzordnung stehen (vgl. Zwicker 2016, S.49f).

Dieses besorgniserregende Resultat liefert die Motivation für die in Abschnitt 5 durchgeführte Simulationsstudie. In dieser soll ermittelt werden, wie präsent die Problematik des Wahlbetrugs wirklich ist und wie sich die relativen Häufigkeiten, wie oft Wahlbetrug von Erfolg gekrönt ist, darstellen. Zunächst werden allerdings im nächsten Kapitel die für die Simulation benötigten Modelle vorgestellt.

⁵Das IIA-Axiom wird so abgewandelt, dass zwei zur Wahl stehende Alternativen nur noch unabhängig von anderen zur Wahl stehenden Alternativen bewertet werden sollen, wenn die Werte für die beiden Alternativen in den individuellen value-difference-functions jeweils gleich sind.

3 Modelle der Social Choice Theory

In diesem Abschnitt wird ein weiterer Blickwinkel auf Wahlen im Kontext der Social Choice Theory vorgestellt: die Modellierung von (realistischen) Wahlen und Wahlverhalten. Zudem wird in der sogenannten Behavioural Social Choice Theory (für einen Überblick siehe Regenwetter et al. 2008) empirisch untersucht, wie ähnlich sich Ergebnisse echter Wahlen unter verschiedenen Wahlregeln sind oder wie oft sich Paradoxa bei der Aggregation der Wählerstimmen (z.B. im Zuge von paarweisen Vergleichen⁶) beobachten lassen. Im Fokus dieser Arbeit soll aber die Modellierung von individuellen Präferenzen liegen, um eine theoretische Grundlage für die später folgenden Simulationen von Wahlbetrug schaffen zu können. Zuerst werden einige (einfachere) stochastische Modelle betrachtet, die in unterschiedlichem Maße Homogenität zwischen den Wählern berücksichtigen können. Des Weiteren wird ein räumliches Modell vorgestellt, das auf Attributen von Wählern und Kandidaten beruht.

Ausgangspunkt ist hier wiederum eine Wahl mit m Kandidaten und n Wählern bzw. eine in Definition 2 definierte *Wahlsituation*, die jedem der $m!$ möglichen individuellen Rankings eine Anzahl n_r zwischen 0 und n zuordnet, sodass die Summe dieser Anzahlen $\sum_{r=1}^{m!} n_r = n$ ergibt. Außerdem sei N_r eine diskrete Zufallsvariable, die die Wahrscheinlichkeiten von $n_r = 0, \dots, n$ für ein Ranking r beschreibt. Sei weiterhin N ein Vektor der Länge $m!$, der eine zufällige Realisation von $N_1, \dots, N_{m!}$ darstellt (vgl. Tideman und Plassmann 2012, S. 219).

3.1 Stochastische Modelle

Die im Anschluss vorgestellten Modelle unterscheiden sich in der Art, wie die Wahrscheinlichkeitsmasse auf die verschiedenen Rankings oder Profile aufgeteilt wird.

Das wohl bekannteste Modell ist das *Impartial Culture (IC)*-Modell. Bei diesem wird die einerseits stark vereinfachende, aber andererseits auch sehr restriktive Annahme getroffen, dass die Wahrscheinlichkeiten für jedes Ranking gleich sind, nämlich $\frac{1}{m!}$, und dass jeder Wähler unabhängig von allen anderen Wählern ein Ranking abgibt (siehe Tideman und Plassmann 2012, S. 224).

Das *Impartial Anonymous Culture (IAC)*-Modell geht hingegen davon aus, dass alle Wahrscheinlichkeitsverteilungen über die Rankings und damit auch alle Wahlsituationen gleich wahrscheinlich sind (vgl. Tideman und Plassmann 2012, S. 220-223).

In anderen Worten beschreiben Gehrlein und Fishburn (1976, S. 1ff) das IAC-Modell mit sogenannten „A-Profilen“, die in der hier vorliegenden Arbeit in Abschnitt 2.1 als *Wahlsi-*

⁶In den meisten Fällen wird hier das Condorcet-Paradoxon betrachtet, das dann auftritt, wenn ein Profil aus transitiven Individualpräferenzen zu einer intransitiven Gesellschaftsordnung aggregiert wird.

tuationen definiert wurden, und jeder möglichen linearen Ordnung eine Anzahl (natürliche Zahl) zuordnen. Unter IAC tritt jede Wahlsituation mit gleicher Wahrscheinlichkeit ein. Für die Profile bedeutet dies, dass jene Profile, die aufgrund der Ununterscheidbarkeit (Anonymität) der Wähler zur selben Wahlsituation führen, auch jeweils innerhalb dieser sogenannten anonymen Äquivalenzklasse gleich wahrscheinlich sein sollen. Eine Gleichverteilung der Wahlsituationen bedeutet also im Allgemeinen nicht, dass die Profile auch gleich verteilt sind. Dies erlaubt im Speziellen auch, dass manche Rankings öfter oder seltener auftreten können, als dies unter der Gleichverteilungsannahme bei IC der Fall wäre.

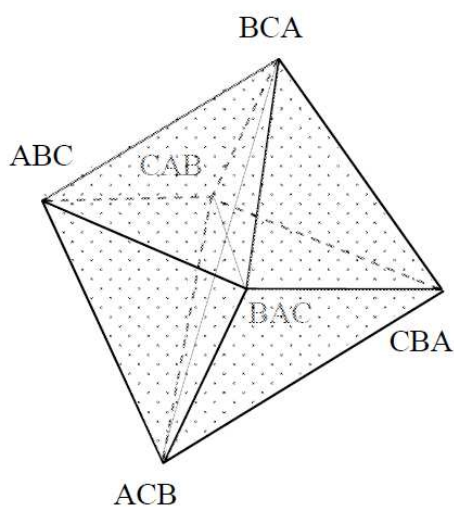


Fig. 9.1a Three-dimensional representation of the support of IAC and variations

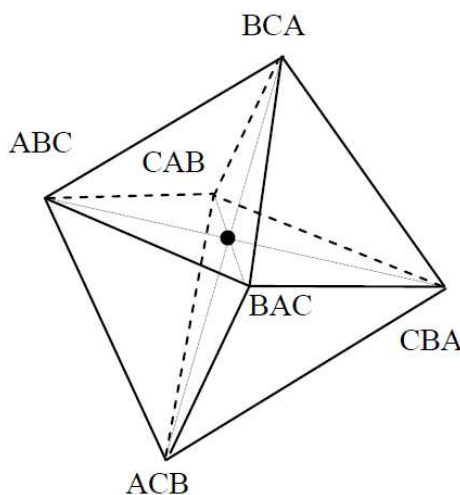


Fig. 9.1b Three-dimensional representation of the support of IC

Abbildung 1: Graphische Darstellung von IAC- und IC-Modell (aus Tideman und Plassmann 2012, S. 222)

Eine graphische Veranschaulichung von den genannten und weiteren unterschiedlichen Modellen für $m = 3$ Kandidaten ist bei Tideman und Plassmann (siehe Abbildung 1) zu finden, wobei der eigentliche 5-dimensionale Simplex (da sich eine Wahrscheinlichkeit jeweils aus den 5 anderen bestimmen lässt, ist die Dimension $3! - 1 = 5$) hier vereinfacht als dreidimensionaler Oktaeder dargestellt wird. Die Bezeichnungen ABC, ACB, etc. stellen hier Rankings der Kandidaten A , B und C dar. ABC entspricht beispielsweise dem Ranking $A \succ B \succ C$. Jeder Eckpunkt des Oktaeders entspricht somit derjenigen Wahlsituation, in

der alle Wähler mit Wahrscheinlichkeit 1 das nebenstehende Ranking abgeben. Die Punkte innerhalb der Oktaeder beschreiben mögliche Wahrscheinlichkeitsverteilungen der Rankings im zugehörigen Modell. Entsprechend ist für das IC-Modell nur in der Mitte, wo alle Rankings gleich wahrscheinlich sind, ein Punkt eingetragen, da dies die einzige vom Modell zugelassene Verteilung ist. Hingegen ist im IAC-Modell jeder Punkt innerhalb des Oktaeders eine prinzipiell mögliche Wahrscheinlichkeitsverteilung. Auch Punkte auf den Kanten (hier sind jeweils Wahrscheinlichkeiten für 4 der Rankings gleich 0, die übrigen beiden summieren sich zu 1) sowie auf den Flächen (je 3 Wahrscheinlichkeiten gleich 0) sind möglich (siehe Tideman und Plassmann 2012, S.220ff).

Ein drittes Modell, das eine weitere Verallgemeinerung des IC- und des IAC-Modells darstellt, ist das multivariate **Pólya-Eggenberger-Modell**, beruhend auf der gleichnamigen Verteilung. Dieses modelliert ebenfalls Wahrscheinlichkeiten für die Anzahlen, wie oft die verschiedenen Rankings von den Wählern abgegeben werden. Diese werden basierend auf Berg (1985, S. 380) folgendermaßen definiert:

$$P(N = x) = \frac{n!}{\prod_{i=1}^{m!} x_i!} \frac{\prod_{i=1}^{m!} A_i^{[\sigma, x_i]}}{A^{[\sigma, n]}}, \quad (12)$$

wobei $x = (x_1, \dots, x_{m!})$, $\sum_{i=1}^{m!} x_i = n$, $A_1, \dots, A_{m!} > 0$ beliebig aber fest,

$\sigma \geq 0$ und $A_1 + A_2 + \dots + A_{m!} = A$

und die generalisierten aufsteigenden Fakultäten $A^{[\sigma, n]} = A(A + \sigma) \cdots (A + (n - 1)\sigma)$ sind.

Man kann sehen, dass für den Fall, in dem $\sigma = 0$ und $p_i := A_i/A$ der Spezialfall einer Multinomial-Verteilung resultiert mit:

$$P(N = x) = \frac{n!}{\prod_{i=1}^{m!} x_i!} \prod_{i=1}^{m!} p_i^{x_i} \quad (13)$$

Sind außerdem alle $p_i = \frac{1}{m!}$, erhält man das Setting einer Impartial Culture (vgl. Berg 1985, S. 380f).

Setzt man hingegen $\sigma = 1$ und alle $A_i = 1$, erhält man den Fall der Impartial Anonymous Culture. Da $1^{[1,x]} = x!$ gilt, lässt sich außerdem zeigen, dass in diesem Fall $P(N = x)$ auf folgende Konstante reduziert werden kann:

$$P(N = x) = \frac{1}{\binom{(m!-1)+n}{n}} \quad (14)$$

wobei $\sum_{i=1}^{m!} x_i = n$ und $\binom{n}{m} = \frac{m!}{n!(m-n)!}$ gilt. Dies ist eine Verallgemeinerung von Berg (1985, S. 380f), der dieses Modell lediglich für 3 Kandidaten beschrieben hat.

Eine alternative Darstellung/Parametrisierung des Modells orientiert sich an dessen anschaulicher Interpretation. Man stelle sich eine Urne vor, in der sich anfangs c_i Kugeln der Farben \mathcal{C}_i , ($i = 1, \dots, k$) befinden (k entspricht $m!$, also der Anzahl möglicher Rankings bei m Kandidaten). Nun werden n Kugeln mit Zurücklegen gezogen, wobei zusätzlich nach jedem Zug $s > 0$ Kugeln der zuletzt gezogenen Farbe der Urne hinzugefügt werden. Somit fungiert s genauso wie σ als Homogenitätsparameter. Die Dichte nach n Ziehungen entspricht dann:

$$f(N_1 = x_1, \dots, N_k = x_k) = \frac{n!}{\prod_{i=1}^k x_i!} \frac{\prod_{i=1}^k a_i^{[x_i]}}{a^{[n]}}, \quad (15)$$

wobei $a_i = c_i/s$, $a = \sum_{i=1}^k a_i$ und $n = \sum_{i=1}^k x_i$ (vgl. Johnson und Kotz 1977, S. 194f).

Außerdem entspricht $a^{[x]}$ hier dem sogenannten Pochhammer-Symbol (auch r -te steigende Faktorielle genannt), wobei für $x \geq 0$ gilt: $a^{[0]} = 1$ und $a^{[x]} = a(a+1)(a+2) \cdots (a+x-1) = \frac{\Gamma(a+x)}{\Gamma(a)}$ mit der Gamma-Funktion $\Gamma(\cdot)$ (vgl. Davis 1972, S. 256).

Das Modell hängt also zum einen von der Anfangsverteilung der Kugeln in der Urne und außerdem – je nach Größe von s – unterschiedlich stark von jenem Homogenitätsparameter ab. Bei kleinem s nähert sich die Verteilung einer Multinomialverteilung mit den Anteilen $\frac{c_i}{n}$ an. Je größer s wird, desto abhängiger wird das Ergebnis nach n Ziehungen vom Ergebnis des ersten Zugs.

Die Pólya-Verteilung wurde mithilfe der Statistiksoftware *R* (R Core Team 2016) simuliert. Dafür wurde zuerst eine Matrix mit n Zeilen (Anzahl der Ziehungen, entspricht in der Anwendung der Anzahl der Wähler) und $m!$ Spalten (entspricht den möglichen Rankings) erzeugt. Für die Rankings muss außerdem eine Startverteilung („prob“) angegeben werden, die der anfänglichen, farblichen Zusammensetzung des Inhalts der Urne entspricht. Diese Verteilung

wird dann innerhalb der Funktion *rmultinom* aus dem R-Paket *stats* in eine Wahrscheinlichkeitsverteilung umgewandelt und mithilfe dieser Multinomialverteilung wird eine Kugel gezogen. Der Eintrag der Matrix in der ersten Zeile ist eine 1 in der Spalte, die der Farbe der gezogenen Kugel entspricht, und 0 sonst. Im Anschluss wird die Startverteilung mit dem Produkt des Eintrags aus der ersten Zeile und *c* addiert, um zu simulieren, dass *c* Kugeln der gezogenen Farbe der Urne hinzugefügt werden. Die nächste Ziehung wird dann mit der geupdatedeten Verteilung durchgeführt. Diese Vorgehensweise wird rekursiv fortgesetzt. Wurden alle Ziehungen durchgeführt, werden zuletzt noch die Spaltensummen der Matrix gebildet, um die Anzahl der Ziehungen pro Kugelfarbe (Ranking) zu erhalten. Das Verfahren kann folgendermaßen als R-Code ausgeführt werden:

```
#Simulation einer Polya-Urne
polya <- function(n, prob, c) {
  # prob=startanzahl, c= anzahl kugeln hinzufuegen, n= Anzahl Ziehungen
  p <- matrix(nr = n, nc = length(prob), 0)
  for (i in 1:n) {
    p[i, ] <- t(rmultinom(1, 1, prob))
    prob <- prob + c * p[i, ]
  }
  return(colSums(p))
}
```

3.2 Räumliches Modell

Ein räumliches Modell (spatial model) für die Social Choice Theory wurde sowohl von Good und Tideman (1976) als auch von Chamberlin und Cohen (1978) vorgeschlagen. Die Motivation zur Einbeziehung dieses Modells in die vorliegende Arbeit liefern Tideman und Plassmann, die zeigen konnten, dass das räumliche Modell die Variation der Wähler (ihre abgegebenen Rankings betreffend) fast vollständig und damit am besten von allen von ihnen betrachteten Modellen erklären konnte (siehe Tideman und Plassmann 2012, S. 236f). Somit kann es getrost als empirisch realistisches Modell beschrieben werden. In diesem Modell wird angenommen, dass individuelle Wahlentscheidungen bzw. die dafür stehenden Rankings abhängig von Ausprägungen verschiedener Attribute (beliebiger Anzahl) der Kandidaten sowie der Wähler sind. Im politischen Sinne könnte man unter Attributen beispielsweise die Einstellung zu verschiedenen politisch relevanten Fragestellungen verstehen. Qualitativ gesprochen bedeutet dies, dass ein Wähler denjenigen Kandidaten favorisieren wird, dessen Einstellungen den seinen am ähnlichsten sind. In einer graphischen Visualisierung des

Modells (Abbildung 2) spannen die gegebenen Attribute eine (Hyper-)Ebene auf, die auch Attributebene genannt wird. Jeder Kandidat und Wähler kann dann durch einen Punkt in der Ebene dargestellt werden. Der Punkt eines Wählers befindet sich dort, wo sich der nach seinen Ansichten ideale Kandidat befinden würde.

Chamberlin und Cohen (1978, S. 1347) simulieren die Einstellungen gegenüber den $q \in \mathbb{N}$ Attributen voneinander abhängig und rekursiv mithilfe von (Standard-)Normalverteilungen. Dazu wird die Ausprägung des ersten Attributs als standardnormalverteilte Zufallsvariable simuliert und alle weiteren Attribute als der Wert des vorhergehenden Attributs mit addiertem normalverteiltem Rauschen mit Erwartungswert 0 und Standardabweichung 1.4 (oder Varianz $1.4^2 = 1.96$). Für die Attribute $A_j^{(i)}$, $j = 1, \dots, q$, $i = 1, \dots, n$ des i -ten Individuums soll also gelten:

$$\begin{aligned} A_1^{(i)} &\sim N(0, 1) \\ A_j^{(i)} &\sim N(A_{j-1}^{(i)}, (1.4)^2 = 1.96) \quad \forall j = 2, \dots, q. \end{aligned} \tag{16}$$

Alternativ können die Attribute aber auch unabhängig voneinander (jeweils standardnormalverteilt) gezogen werden.

Was die wählbaren Kandidaten betrifft, werden diese als Stichprobe – mit fester Größe m – aus der Wählerschaft (der Größe n) gezogen, wobei jeder Kandidat auch weiterhin noch ebenso ein Wähler ist. Für die Zusammensetzung der Stichprobe werden drei verschiedene Fälle betrachtet. Hierbei wird die Extremität der Kandidaten variiert, wobei diese im Sinne der euklidischen Distanz zu verstehen ist (dabei ist ein Individuum umso extremer, je weiter sein Ideal vom Nullpunkt der Attributebene entfernt ist).

1. Die Kandidaten haben eine ähnliche Extremitätsstruktur wie die Wähler und können daher auch direkt als Abbild der Wähler angesehen werden.
2. Die Kandidaten sind ihre Ansichten (Attributausprägungen) betreffend extremer als die Wähler, das heißt, sie besitzen tendenziell extremere/kontroversere und eventuell auch eher entgegengesetzte Meinungen.
3. Sie sind in ihren Ansichten gemäßiger als die Wähler, d.h. ihre Attributausprägungen sind weniger extrem.

(vgl. Chamberlin und Cohen 1978, S. 1347f)

Diese Unterscheidung wirkt sich auf den Prozess aus, wie die Kandidaten in der Simulation generiert werden. So werden jeweils m Kandidaten ohne Zurücklegen direkt aus der Menge der n Wähler gezogen, jedoch ist die Gewichtung unterschiedlich: in Fall 1) (ähnliche Struktur) ist die Ziehungswahrscheinlichkeit für jeden Wähler gleich, d.h. jeder Wähler hat die gleiche Chance, ein Kandidat zu werden. In Fall 2) und 3) hingegen werden die Ziehungswahrscheinlichkeiten abhängig von der Extremität eines Individuums bestimmt. In Fall 2) steigt die Ziehungswahrscheinlichkeit proportional mit der Extremität, während sie in Fall 3) proportional fällt; das bedeutet: hier bekommen Individuen die höchsten Wahrscheinlichkeiten, als Kandidat ausgewählt zu werden, zugewiesen, die sich mit ihrer Einstellung zu den Attributen nahe dem Nullpunkt befinden.

Sei $A^{(i)} = (A_1^{(i)}, \dots, A_q^{(i)}) \in \mathbb{R}^q$ der Attributvektor von Wähler $i \in V$ und $d(x, y) = \sqrt{(x_1 - y_1)^2 + \dots + (x_q - y_q)^2}$ die euklidische Distanz für $x, y \in \mathbb{R}^q$.

$$\begin{aligned}
\text{Fall 1: P('i wird Kandidat')} &\equiv \frac{\binom{n-1}{m-1}}{\binom{n}{m}} = \frac{m}{n} \\
\text{Fall 2: P('i wird Kandidat')} &\propto d(A^{(i)}, \mathbf{0}) \\
\text{Fall 3: P('i wird Kandidat')} &\propto \frac{1}{d(A^{(i)}, \mathbf{0})} \\
\mathbf{0} &= (0, 0, \dots, 0) \in \mathbb{R}^q
\end{aligned} \tag{17}$$

In der Implementation kann dies verwirklicht werden, indem die euklidischen Distanzen für alle Wähler so normiert werden, dass sie in Summe 1 ergeben, und somit als Ziehungswahrscheinlichkeiten interpretiert werden können.

Nachdem die Kandidaten generiert wurden, muss ermittelt werden, welche Präferenzen die Wähler über der Menge der Kandidaten besitzen. Um die individuellen Rankings der Wähler über die Kandidaten bilden zu können, misst man mithilfe eines geeigneten Distanzmaßes (hier: Euklidische Distanz) für jeden Wähler den Abstand von seinem Idealkandidaten (also dem Punkt, der ihn in der Ebene repräsentiert) zu allen zur Wahl stehenden Kandidaten. Je kleiner die Distanz ist, desto höher wird ein Kandidat im individuellen Ranking gelistet sein. Zur Veranschaulichung wird in Abbildung 2 ein Wähler V und drei Kandidaten A , B und C in einer Attributebene mit 2 Attributen dargestellt, sowie der (gerundete) euklidische Abstand zwischen dem Wähler und den drei Kandidaten. Beruhend auf der Minimierung

der Distanzen wäre das individuelle Ranking des Wählers V : $B \succ A \succ C$.

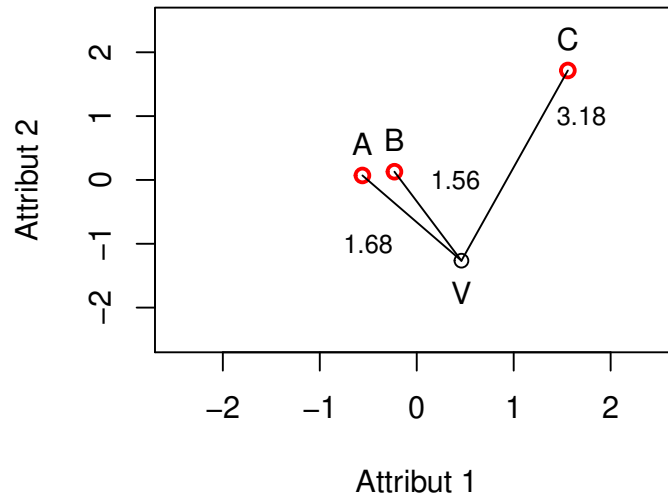


Abbildung 2: Visualisierung des räumlichen Modells: Wähler V , Kandidaten A , B , C und euklidische Distanzen in einer zweidimensionalen Attributebene

Good und Tideman (1976, S. 375) definieren zusätzlich direkt das kollektive Ranking basierend auf den addierten Distanzen der Wähler zu den Kandidaten. In diesem Falle wäre also der Kandidat an der Spitze des Gesellschaftsrankings, dem die Wähler basierend auf dem euklidischen Distanzmaß in Summe am nächsten sind. Dieses Vorgehen wird in der vorliegenden Arbeit allerdings nicht übernommen. Stattdessen fällt die Wahlentscheidung erst im Anschluss auf Basis der verschiedenen Wahlsysteme.

4 Wahlbetrug

In diesem Abschnitt wird nun näher auf Varianten des Wahlbetrugs eingegangen und somit die Grundlage für die später folgenden Simulationsstudien geschaffen. Dabei wird zwischen drei Arten von Wahlbetrug unterschieden: Manipulation, Bestechung und Kontrolle. Allen dreien ist gemeinsam, dass durch gezielte Änderungen des Profils der Ausgang der Wahl bzw. das Gesellschaftsranking zugunsten des Wahlbetrügers verändert werden soll. Außerdem soll versucht werden, für verschiedene Wahlregeln einen Algorithmus oder eine Heuristik anzugeben, die (effizient) Erfolg bei der Manipulation versprechen.

4.1 Manipulation

Der Begriff der Manipulation beschreibt die gezielte Einflussnahme eines Manipulators auf den Ausgang einer Wahl, indem dieser eine strategische Präferenzordnung als eigene Wahlstimme abgibt (statt seiner ehrlichen Präferenz), um den Ausgang der Wahl für sich persönlich (seine wahren Präferenzen) zu verbessern oder, in einer strengeren Variante, einen ausgezeichneten Kandidaten (seinen Favoriten) zum Sieger zu machen. Eine grundsätzliche Annahme ist hierbei, dass der Manipulator die individuellen Rankings der anderen Wähler, die ehrliche Stimmen abgeben, genau kennt und sich diese Information zu Nutzen machen kann (complete information setting, vgl. Conitzer et al. 2007, S. 6).

Es soll folgende Fragestellung beantwortet werden:

Manipulation. *Existiert für eine Kandidatenmenge C und eine Menge ehrlicher Wähler V eine lineare Präferenzordnung $s \in \mathcal{L}(C)$, sodass ein ausgezeichnete Kandidat $c \in C$ die Wahl $(C, V \cup \{s\})$ unter einer vorgegebenen Wahlregel gewinnen kann? (Rothe et al. 2012, S. 179)*

Eine Verallgemeinerung dessen besteht darin, dass es nicht nur einen Manipulator gibt, sondern eine Koalition von Manipulatoren M (mit $k \in \mathbb{N}$ Mitgliedern), die das gemeinsame Ziel verfolgen, den Wahlausgang in ihrem Sinne zu verändern. Dabei müssen nicht gezwungenermaßen alle Manipulatoren das gleiche Ranking abgeben, sondern es kann so variiert werden, dass es dem gewünschten Wahlausgang am nächsten kommt. Ein einzelner Manipulator entspricht also einem Spezialfall des koalitionären Manipulierens, in dem die Koalition nur aus einer Person besteht (d.h. $M = \{s\}$).

Koalitionale Manipulation. *Existiert eine Koalition von manipulativen Wählern M , $|M| \leq k$, sodass ein festgelegter Kandidat $c \in C$ eine Wahl $(C, V \cup M)$ unter einer gegebenen Wahlregel gewinnen kann (oder wie groß muss k dafür mindestens sein)?*

4.2 Bestechung

Anders als bei der Manipulation steht hier als Ausgangspunkt ein komplettes Profil, zu dem keine unehrlichen individuellen Rankings mehr hinzugefügt werden können. Allerdings besteht die Möglichkeit für den Wahlbetrüger, Wähler zu bestechen, damit sie ihr ehrlich abgegebenes Ranking in seinem Sinne abändern. Um ein Limit zu setzen, wie viele Wähler maximal bestochen werden können, sind verschiedene Ansätze denkbar. Zum einen kann von vornherein eine Höchstzahl an Wählern, die bestochen werden dürfen, gegeben sein (vgl. Faliszewski et al. 2006, S. 15), zum anderen gibt es die Möglichkeit, die Bestechlichkeit in Geldeinheiten zu messen („\$-Bribery“). Hierbei kann jeder Wähler angeben, wie viel der Wahlbetrüger bezahlen müsste, um sein Ranking in beliebiger Weise verändern zu können. Diesem stünde dann ein zuvor definiertes Budget zur Verfügung, um ein für ihn bestmögliches Ergebnis zu erzielen (vgl. Faliszewski et al. 2006, S. 15f). Die zuvor beschriebene Bestechung kann also auch als Spezialfall der \$-Bribery gesehen werden, bei der jeder Wähler den Preis 1 angibt. Dann sind ein einzuhaltendes Budget und eine Höchstzahl zu bestechender Wähler äquivalent.

Weiterhin denkbar ist, dass jeder Wähler eine variable Kostenfunktion angibt, wobei die Kosten sich dahingehend unterscheiden, wie weit zwei auf Anordnung des Wahlbetrügers zu vertauschende Kandidaten auf der ursprünglichen Präferenzliste auseinanderliegen. Das bedeutet im Normalfall: je weiter zwei Kandidaten auseinander liegen, desto kostspieliger wird es für den Wahlbetrüger. Mit dem zur Verfügung stehenden Budget kann dieser dann versuchen, auf den individuellen Ranglisten Kandidaten in effizienter Weise paarweise zu vertauschen. Effizient bedeutet hierbei, dass von zwei Vertauschungen, die denselben Effekt für das Gesamtranking erzielen, die günstigere bevorzugt werden sollte.

Rothe et al. (2012, S. 213) definieren eine variable Kostenfunktion für jedes Kandidatenpaar und erlauben des Weiteren nur Vertauschungen von benachbarten Kandidaten („swap bribery“), d.h. sollen zwei bestimmte Kandidaten auf einer individuellen Präferenzliste miteinander vertauscht werden, ist dies nur möglich indem $x+1$ paarweise Vertauschungen durchgeführt werden, wobei x die Anzahl der Kandidaten ist, die zwischen den beiden ausgezeichneten Kandidaten rangieren. Die Kosten für diese Vertauschung ergeben sich dann aus der Summe der einzelnen Vertauschungen.

Da sich Preise bzw. Preisfunktionen und zugehörige Budgetbeschränkungen im Allgemeinen nur sehr willkürlich bestimmen lassen, wird sich diese Arbeit im Weiteren auf den Standard-

fall der Bestechung (bzw. \$-Bribery mit Preis 1) beschränken, formal ausgedrückt ist also das folgende Problem zu lösen:

Bestechung. *Ist es möglich höchstens $k \leq |V|$ Stimmen in V so zu ändern, dass ein ausgezeichnete Kandidat $c \in C$ eine Wahl (C, V) unter einer bestimmten Wahlregel gewinnen kann (oder wie groß muss k minimal sein, damit das gelingt)?* (vgl. Rothe et al. 2012, S. 211)

4.3 Kontrolle

Die Wahlkontrolle hat mit Manipulation und Bestechung gemeinsam, dass es einen externen Manipulator gibt, der allerdings weder ein eigenes Ranking abgeben noch andere Wähler bestechen kann. Stattdessen agiert er als Wahlleiter und kann damit die im folgenden beschriebenen Arten von Veränderungen in der Zusammensetzung der Kandidaten und Wählerschaft vornehmen.

Zum einen ist es möglich, Kandidaten hinzuzufügen, zu entfernen oder sie beliebig disjunkt zu partitionieren. Letzteres ist einerseits denkbar als sequentielle Wahl, das heißt, die Wähler stimmen zunächst über eine Teilmenge $C_1 \subset C$ von Kandidaten ab und anschließend über den/die gewählten Sieger aus C_1 vereinigt mit der zweiten Teilmenge $C_2 \subset C$, also über $f((C_1, V)) \cup C_2$. Andererseits kann eine Partition mit Stichwahl durchgeführt werden; dabei stimmen zunächst alle Wähler separat über die Teilmengen $C_1 \subset C$ und $C_2 \subset C$ ab und im Anschluss noch einmal über die Menge der Sieger der beiden Wahlen $f((C_1, V)) \cup f((C_2, V))$ (vgl. Bartholdi et al. 1992, S. 31f).

Eine beispielhafte Situation, in der es sinnvoll sein kann, einen weiteren Kandidaten hinzuzufügen, wird von Tideman (1987, S. 185) beschrieben. Hierbei handelt es sich um eine Pluralitätswahl mit zunächst zwei Kandidaten A und B. Letztendlich kann sich A den Sieg nur dadurch sichern, indem ein weiterer Kandidat C nominiert wird, der Kandidat B sehr ähnlich ist, sodass sich die Wählerstimmen, die eigentlich B gegolten hätten, gleichmäßig auf B und C verteilen. Dieses Beispiel beruht auf einem weiteren, hier nicht näher betrachteten Gütekriterium, das Tideman im genannten Paper einführt, nämlich der Unabhängigkeit von Klonen. Als „Klon“ eines Kandidaten B wird hierbei ein zusätzlicher Kandidat bezeichnet, wenn er auf jeder individuellen Präferenzliste entweder direkt über oder direkt unter B gerankt wird. Offensichtlich verletzt die Pluralitätswahl dieses Kriterium, da sich das Wahlergebnis nach Entfernen des Klons zugunsten von Kandidat B verändern würde (vgl. Tideman 1987, S. 186).

Ein beliebtes Beispiel, in dem ein Entfernen eines Kandidaten zum Erfolg führen kann, wird oft mit der US-Präsidentschaftswahl 2000 angegeben, als sich der Republikaner George W. Bush am Ende knapp gegen seinen demokratischen Konkurrenten Al Gore durchsetzen konnte. Gore hatte zwar am Ende insgesamt mehr Stimmen gesammelt, verlor allerdings die Wahl wegen des Wahlmännersystems der USA trotzdem. Entscheidend war hierfür der Ausgang im Staat Florida, in dem Bush sich nur mit etwa 500 Stimmen Vorsprung die 25 Wahlmännerstimmen sicherte. Dem Kandidaten der Grünen Partei, Ralph Nader, der insgesamt nur einen Anteil von etwa 3 % der Stimmen sammeln konnte (und keinen einzigen Wahlmann) kann hier als entscheidender Faktor des Wahlausganges gesehen werden, da angenommen wird, dass seine Unterstützer größtenteils Gore gegenüber Bush bevorzugten. Wäre Nader also gar nicht zur Wahl angetreten (oder von einem Wahlleiter von der Wahl ausgeschlossen worden) hätte höchstwahrscheinlich Gore dank der über 97.000 Nader-Wähler die Wahlmänner in Florida für sich gewinnen können und wäre statt Bush Präsident geworden (vgl. Rothe et al. 2012, S. 191).

Des Weiteren können auch Wähler hinzugefügt, entfernt oder partitioniert werden. Partitionierung bedeutet hierbei, dass zuerst zwei disjunkte Teilmengen von Wählern $V_1, V_2 \subset V$, $V_1 \cap V_2 = \emptyset$ jeweils über die gesamte Kandidatenmenge abstimmen und anschließend die gesamte Wählerschaft über die vereinigten Siegermengen der beiden Wahlen nochmals abstimmt, um den Gesamtsieger zu ermitteln, d.h. das Ergebnis der Wahl $(f((C, V_1)) \cup f((C, V_2)), V)$ gebildet wird (vgl. Bartholdi et al. 1992, S. 33).

In dieser Arbeit wird stets nur der Fall der konstruktiven Kontrolle betrachtet, also diejenige, bei der versucht wird, einen favorisierten Kandidaten zum Sieger zu machen (im Gegensatz zur destruktiven Kontrolle, die auch als erfolgreich angesehen wird, wenn einem vom Wahlleiter niedriger gerankten Kandidaten der Wahlsieg streitig gemacht werden kann).

4.4 Computational Social Choice

Angestoßen vor allem durch das Theorem von Gibbard und Satterthwaite war man auf der Suche nach einer Lösung, um die scheinbar allgegenwärtige Manipulierbarkeit von Wahlen umgehen zu können. Die ersten, die versuchten, Konzepte aus der Informatik (Komplexitätstheorie) auf die Social Choice Theory zu übertragen, waren Bartholdi, Tovey und Trick mit ihrem Paper „The computational difficulty of manipulating an election“ (1989). Ihre Idee war es, die Manipulation so computationally schwer wie möglich zu machen. Unter computationaler Schwere versteht man, dass ein Computer dieses Manipulationsproblem (vermutlich) nicht in

polynomieller Zeit (polynomiell abhängig von der Länge der Eingabe) mit einem effizienten Algorithmus lösen kann, also das Problem nicht in der Komplexitätsklasse P liegt.⁷

Gleichzeitig soll es nicht zu schwer werden, den Gewinner der Wahl zu bestimmen. Das Problem der „Gewinnerermittlung“ sollte also in P liegen. Aus dieser Idee heraus entwickelte sich die Fachrichtung der *Computational Social Choice Theory (CSCT)* und die Motivation, Wahlsysteme zu kreieren, die computationally resistent gegenüber Wahlbetrug sind. Inspiriert von der CSCT werden für die Algorithmen im folgenden Abschnitt, falls vorhanden, solche verwendet, die bekanntermaßen in der Komplexitätsklasse P liegen, und somit effiziente Lösungen von Manipulationsproblemen sind. Intuitiv liegt die Erwartung nahe, dass für Wahlsysteme, für die keine effizienten Manipulationsalgorithmen bekannt sind, die Erfolgsaussichten für Manipulatoren geringer sein werden.

4.5 Algorithmen

Im Folgenden soll erörtert werden, wie ein Wahlbetrüger, der manipulieren oder bestechen will, abhängig vom vorliegenden Wahlsystem vorgehen muss, um sein Manipulationsziel bestmöglich zu verwirklichen. Genauer gesagt bedeutet dies, dass, wenn es nicht gelingt mit der beschriebenen Vorgehensweise einen ausgezeichneten Kandidaten $c \in C$ zum Wahlsieger zu machen, es prinzipiell für die vorliegende Wahlsituation, eine bestimmte Wahlregel und die gegebenen Beschränkungen des Manipulators (Anzahl zu manipulierender Wähler/Budget) nicht möglich ist, c zum Sieger zu machen (insbesondere nicht durch eine andere Vorgehensweise).

Pluralitätswahl

Dadurch, dass bei der Pluralitätswahl von jedem Wähler nur die erste Präferenz berücksichtigt wird, wird auch eine klare Vorgehensweise ersichtlich. Die einzig sinnvolle Strategie für *Manipulation* ist es, ein oder mehrere Rankings hinzuzufügen, die den zum Gewinner zu machenden Kandidaten (c) an die erste Stelle setzen. Die restlichen Präferenzen können willkürlich gesetzt werden, da sie für das Wahlergebnis keine Rolle spielen. Hieraus ergibt sich auch, dass die manipulativen Wahlstimmen sich unabhängig von den Stimmen der ehrlichen Wähler ergeben. Das heißt der Manipulator müsste sich nicht im in der Praxis sehr unrealistischen complete-information-setting befinden, um erfolgreich manipulieren zu können. Um

⁷Vorausgesetzt, dass $P \neq NP$. Dies stellt eines der größten, ungelösten Probleme der theoretischen Informatik dar. Es ist allerdings bekannt, dass $P \subseteq NP$. Dabei steht NP für die Komplexitätsklasse, die Probleme enthält, die nichtdeterministisch in polynomieller Zeit gelöst werden können. Für eine Übersicht über die Komplexitätstheorie siehe beispielsweise Homer und Selman (2011).

die benötigte Größe der kleinsten erfolgreichen manipulativen Koalition bestimmen zu können, muss lediglich das ursprüngliche Endergebnis der ehrlichen Wähler bekannt sein. Aus diesem kann die benötigte Anzahl abgelesen werden. Es muss hierfür nur die Differenz aus der Stimmenzahl des aktuellen Gewinners und der Stimmenzahl des ausgezeichneten Kandidaten (jeweils auf Basis der Stimmen der ehrlichen Wähler) gebildet werden. Sollte jedoch ein Gleichstand (Tie) zwischen diesen beiden Kandidaten bestehen, muss dieser zuerst aufgelöst werden (dies geschieht zufällig, siehe Abschnitt 5.2). Geht der Tiebreak zugunsten des aktuellen Gewinners aus, so ist ein weiterer Manipulator nötig, um c zum Sieger zu machen. Sei also w die Stimmenanzahl des Gewinners und s die Stimmenzahl des ausgezeichneten Kandidaten, so ist die gesuchte Größe der Koalition mit gleicher Wahrscheinlichkeit $w - s$ oder $w - s + 1$.

Im Sinne von *Bestechung* ist die gleiche Strategie zu verfolgen, wobei zuerst die Wähler, die den momentanen Gewinner auf Platz 1 ranken, bestochen werden müssen, um gleichzeitig den eigenen Kandidaten zu stärken und seinen stärksten Konkurrenten zu schwächen. Dafür ist allerdings im Gegensatz zur Manipulation Voraussetzung, dass man die Stimmen der ehrlichen Wähler im Einzelnen kennt, um die minimale Anzahl nötiger Bestechungen herausfinden zu können. Diese kann dann ähnlich wie bei der Manipulation direkt aus dem Ergebnis der ehrlichen Wähler abgelesen werden.

Seien wiederum w und s die Stimmenzahlen des Gewinners und ausgezeichneten Kandidaten. Ist $(w - s)$ geradzahlig, so beträgt die minimale Anzahl mit gleicher Wahrscheinlichkeit $\frac{w-s}{2}$ oder $\frac{w-s}{2} + 1$. Für den Fall, dass $(w - s)$ ungerade ist, kann es unmöglich nach Bestechungen zu einem Tie mit dem aktuellen Sieger kommen, weswegen die Anzahl auf $\left\lfloor \frac{w-s}{2} \right\rfloor + 1$ gesetzt werden kann.

Borda

Die Vorgehensweise für die Manipulation eines Bordawahlszenarios beruht auf einem Algorithmus von Zuckerman et al. (2008, S. 281). Hierbei setzt zuerst jedes Mitglied der manipulierenden Koalition den ausgezeichneten Kandidaten c als erste Präferenz. Im Weiteren bilden die Manipulatoren ihre restliche Präferenzliste nacheinander. Dazu geht jeder Manipulator so vor, dass er den Stand im Gesamtranking vor seinem Eingreifen in umgekehrter Reihenfolge auf den zweiten bis letzten Platz seiner Liste platziert. Das bedeutet, er setzt denjenigen Kandidaten, der vor seiner Manipulation den höchsten Borda-Score besitzt, auf den letzten Platz, den mit dem zweithöchsten Score auf den vorletzten Platz und so weiter. Kandidat c wird hierbei übersprungen, da er bereits auf Platz 1 der Präferenzliste gesetzt

wurde. Nach jedem Hinzufügen einer manipulierten Wahlstimme werden mit dem aktualisierten Profil die Borda-Scores aller Kandidaten neu berechnet und überprüft, ob c bereits der Gewinner ist. Falls nicht, geht der nächste Manipulator analog basierend auf dem aktualisierten Borda-Ranking vor.

Auch im Falle der Bestechung kann analog vorgegangen werden, wobei hier darauf geachtet werden sollte, dass ein größerer Effekt zu erzielen ist, wenn zuerst die Wähler bestochen werden, die den ausgezeichneten Kandidaten auf dem letzten Platz ranken (was für jenen eine größtmögliche Verbesserung liefert). Sind keine solchen Wähler mehr übrig, aber c noch nicht der Sieger, sollten im weiteren Vorgehen Wähler bestochen werden, die jeweils den Sieger nach aktuellem Stand auf Platz 1 ihres Rankings führen (falls nötig auch noch Platz 2 usw.). Diese Vorgehensweise ist allerdings lediglich als Heuristik zu verstehen, d.h. sie liefert nicht zwingend das bestmögliche Ergebnis.

Für eine festgelegte Maximalanzahl k von Manipulatoren (oder zu bestechenden Wähler) funktioniert der Algorithmus folgendermaßen (Pseudocode):

für $i = 1, \dots, k$

- Ist c Sieger unter vorgegebenem Wahlsystem? Wenn ja, fertig (gebe $i - 1$ als minimal erfolgreiche Koalition zurück).
- Sonst: Setze c an die Spitze der Präferenzliste von Manipulator i .
- Betrachte Gesellschaftsranking nach $i - 1$ Manipulationen (ohne c) und bilde in umgekehrter Reihenfolge die weitere Präferenzliste von i (Letzte (m -te) Präferenz der Gesellschaft ist zweite Präferenz von i , etc.)
- Manipulation: Füge i zur bestehenden Wählerschaft hinzu ($i \cup (V \cup \{1, \dots, i - 1\})$).
- Bestechung: Ändere zuerst die Präferenzlisten derjenigen Wähler, die den ausgezeichneten Kandidaten als letzte Präferenz genannt haben. Sind diese erschöpft, ändere die Rankings derer, die den derzeitigen Sieger als 1. Präferenz führen (wenn nötig auch noch 2./3./4. Präferenz).

Paarweise Vergleiche

Für paarweise Vergleiche bzw. im Speziellen die Methode nach Copeland ist die sinnvollste Strategie, wie schon bei der Pluralitätswahl, den ausgewählten Kandidaten auf Platz 1 der manipulativen Rankings aufzuführen und ihm somit eine zusätzliche Stimme in allen direk-

ten Vergleichen, in die er involviert ist, zu geben, während die Reihenfolge der sonstigen Kandidaten willkürlich gewählt werden kann, da sie keinen weiteren Einfluss auf das Wahlergebnis des ausgewählten Kandidaten hat. Wie auch bei der Pluralitätswahl, sind hier die Stimmen der ehrlichen Wähler für die manipulativen Stimmen nicht von weiterer Bedeutung. Hierbei muss folgendermaßen vorgegangen werden, um den gesuchten Schwellenwert direkt aus dem Endergebnis herauszufinden: man identifiziere zuerst den minimalen Wert (genannt min_c) in der entsprechenden Zeile des Endwahlergebnisses, die dem ausgezeichneten Kandidaten zuzuordnen ist. Die Spalte, in der sich dieser Wert befindet, entspricht demjenigen Kandidaten, zu dem das Stimmendefizit am größten ist (in Tabelle 3 auf Seite 11 ist das für den ausgezeichneten Kandidaten „C“ der Wert $min_c = 4$ in Zeile „C“ und Spalte „A“). Daher muss die Koalition mindestens so viele Manipulatoren enthalten, um dieses Defizit ausgleichen zu können. Dies ist gleichbedeutend damit, dass mit gleicher Wahrscheinlichkeit entweder $n - 2min_c$ oder $n - 2min_c + 1$ Manipulatoren benötigt werden. Letztere Anzahl wird benötigt, falls ein Unentschieden noch gegen den ausgezeichneten Kandidaten aufgelöst wird.

Im Falle der Bestechung muss außerdem noch beachtet werden, ob die Anzahl der Wähler geradzahlig ist oder nicht: Ist n geradzahlig, so beträgt die Anzahl der zu bestechenden Wähler mit gleicher Wahrscheinlichkeit $\frac{n}{2} - min_c$ und $\frac{n}{2} - min_c + 1$, wobei min_c wiederum das Minimum in der oben beschriebenen Zeile bezeichnet (hier werden erneut die zwei verschiedenen Ausgänge bei einem Tie berücksichtigt). Ist hingegen n ungerade, muss die Anzahl immer $\lfloor \frac{n}{2} \rfloor + 1 - min_c$ betragen, da dann keine Ties in direkten Vergleichen möglich sind. Hierbei spielt es eine untergeordnete Rolle, welche Wähler zuerst bestochen werden, solange es sich nur um solche handelt, die den ausgezeichneten Kandidaten hinter demjenigen Kandidaten, zu dem das maximale Defizit besteht, ranken.

Es ist anzumerken, dass der Borda-Algorithmus ebenfalls für Pluralitäts- und Copelandwahlen optimal wäre. Nach sorgfältiger Studie der Simulationsergebnisse, die der Algorithmus für Pluralitäts- und Copelandwahlen lieferte, fiel auf, dass diese auch in einfacher Weise aus dem Endergebnis der Wahl abgelesen werden können. Dem Autor der vorliegenden Arbeit ist keine Quelle aus der Literatur bekannt, die diese Methoden benutzt. Aus computationalen Gründen werden daher für die weiteren Simulationen die weniger rechenaufwendigen, zuvor beschriebenen Ablesetechniken verwendet, die die gleichen Ergebnisse liefern.

Single Transferable Vote

Im Gegensatz zu den drei zuvor genannten Wahlregeln, ist für STV keine effiziente Vorge-

hensweise oder Algorithmus bekannt. Vielmehr konnte sogar von Bartholdi und Orlin (1991) gezeigt werden, dass STV computational resistant gegenüber strategischem Wählen ist, das heißt ein Manipulator müsste ein – sogar im complete-information-setting – NP-schweres Problem lösen, um die Wahl erfolgreich manipulieren zu können. Um Vergleichbarkeit zwischen den Wahlsystemen in den späteren Simulationsergebnissen gewährleisten zu können, wird also (auch mangels bekannter Alternativen) als Heuristik derselbe Algorithmus wie bei der Bordawahl eingesetzt.

5 Simulation

Im folgenden Abschnitt sollen die in den vorhergehenden Kapiteln dargestellten Konzepte aus der (Behavioural) Social Choice Theory unter dem Dach einer Simulationsstudie zusammengeführt werden. Dazu werden im ersten Schritt Wahlsituationen/Profile aus unterschiedlichen Modellen bzw. Modellkonfigurationen (siehe Abschnitt 3) simuliert und mit vier verschiedenen Wahlregeln ausgewertet. Für die erhaltenen Wahlergebnisse werden anschließend die Manipulationsalgorithmen, die in Abschnitt 4.5 vorgestellt wurden, angewandt. Hierbei ist von besonderem Interesse, wie die Erfolgsquote der Manipulation und Bestechung und die Anzahl der nötigen Manipulatoren sich abhängig von Manipulations- und Wahlparametern (wie Kandidaten- oder Wählerzahl) sowie zugrunde liegendem Wahlsystem verhält. Bevor die Ergebnisse dargestellt werden, wird zuerst noch auf den Aufbau der Studie sowie die Implementation eingegangen.

5.1 Ablauf und Konfigurationen

Zunächst werden in der Simulation für jedes der Modelle (IC, IAC, Pólya, Spatial) 100.000 zufällige Wahlsituationen/Profile mit $m = 5$ Kandidaten und $n = 1000$ Wählern gezogen. Hierbei werden also genauer gesagt die Präferenzlisten der Wähler zufällig (modellbasiert) bestimmt.

Es werden folgende (Ausgangs-)Modelle (siehe Abschnitt 3) näher betrachtet:

1. IC-Modell (keine Homogenität)
2. IAC-Modell (leichte Homogenität)
3. Pólya-Eggenberger-Modell mit Homogenitätsparameter 10 (starke Homogenität)
4. Spatial-Modell mit 4 unabhängig generierten, standardnormalverteilten Attributen, wobei die Kandidaten in ungewichteter Weise aus der Menge der Wähler gezogen werden.

Die ersten drei Modelle lassen unterschiedlich starke Homogenität der Wähler zu. Die Homogenität im räumlichen Modell wird nicht direkt durch einen Parameter gesteuert und soll später noch genauer untersucht werden.

Um den Einfluss der verschiedenen Konfigurationen im räumlichen Modell veranschaulichen zu können, wird jeweils eine Eigenschaft variiert, während die restlichen konstant auf der

Ausgangsausprägung verweilen. Für das veränderte Setting werden dann 100.000 neue Profile simuliert. Hierzu werden die folgenden alternativen Ausprägungen festgelegt:

- 10 Attribute,
- abhängige Ziehung der Attribute (siehe Abschnitt 3.2, Gleichung 16) und
- gemäßigte bzw. extremere Kandidaten, d.h. Ziehung von Kandidaten gewichtet nach deren Extremität.

Außerdem werden für jede der 4 Modellklassen noch weitere 4 Settings erzeugt. Die ersten beiden betrachten eine Verringerung bzw. Erhöhung der Wählerzahl (auf 100 bzw. 10.000). Analog wird bei den restlichen beiden Modellen die Anzahl der Kandidaten auf 3 bzw. 7 variiert.

In Tabelle 4 und 5 werden nochmals alle 24 betrachteten Settings in übersichtlicher Weise zusammengefasst.

Abkürzung	Modell	Kandidaten	Wähler
ic1000	IC	5	1000
ic100	IC	5	100
ic10000	IC	5	10000
ic3	IC	3	1000
ic7	IC	7	1000
iac1000	IAC	5	1000
iac100	IAC	5	100
iac10000	IAC	5	10000
iac3	IAC	3	1000
iac7	IAC	7	1000
pol1000	Pólya	5	1000
pol100	Pólya	5	100
pol10000	Pólya	5	10000
pol3	Pólya	3	1000
pol7	Pólya	7	1000

Tabelle 4: Überblick über betrachtete Konfigurationen von stochastischen Modellen (siehe Abschnitt 3.1)

Abkürzung	Wähler	Kandidaten	Kandidaten-Struktur	Attribute	Attribute unabhängig?
sp1000	1000	5	neutral	4	ja
sp100	100	5	neutral	4	ja
sp10000	10000	5	neutral	4	ja
sp1000homogen	1000	5	gemäßigt	4	ja
sp1000extrem	1000	5	extrem	4	ja
sp1000ni	1000	5	neutral	4	nein
sp1000ten	1000	5	neutral	10	ja
sp3	1000	3	neutral	4	ja
sp7	1000	7	neutral	4	ja

Tabelle 5: Überblick über betrachtete Konfigurationen des räumlichen Modells (siehe Abschnitt 3.2)

Nachdem die Profile simuliert wurden, können auf ihnen die Algorithmen und Heuristiken zur Manipulation und Bestechung aus Abschnitt 4.5 angewandt werden. Im Speziellen sollen Pluralitäts-, Borda-, Copeland- und STV-Wahlen manipuliert und bestochen werden. Damit wird jede Obergruppe (Scoring-Protokolle, Paarweise Vergleiche und Mehrstufige Wahlsysteme) mindestens einmal näher betrachtet. Da für STV kein effizienter Algorithmus bekannt ist, wird hier als Heuristik der gleiche Algorithmus verwendet wie für Bordawahlen. Zur besseren Vergleichbarkeit sollte normalerweise auch für Pluralitäts- und Copelandwahlen der gleiche Algorithmus verwendet werden, allerdings liefern die beschriebenen einfacheren Abzählmethoden das gleiche Ergebnis für den zum Sieger zu machenden Kandidaten und sind zudem computational weniger aufwendig, da die Präferenzliste der Manipulatoren nicht in potenziell jedem Schritt verändert werden muss, sondern für alle Manipulatoren die gleiche ist. Zusätzlich ist die Manipulation – wie schon erwähnt – mit der vereinfachten Methode nicht abhängig von den Wahlstimmen der ehrlichen Wähler, und somit auch in der Praxis deutlich leichter zu handhaben. Für die Bestechungsheuristiken bei Borda- und STV-Wahl ist allerdings implizit vorausgesetzt, dass die Wählerstimmen im Einzelnen bekannt sind, da die Reihenfolge der Bestechungen von den ursprünglich abgegebenen (ehrlichen) Rankings abhängig ist.

Eine weitere Größe von Interesse ist die relative Erfolgshäufigkeit der Manipulation/Bestechung für ein festes Betrugskontingent. Dieses wird im Vorhinein auf einen realistisch erscheinenden Schwellenwert gesetzt. Für Manipulation und Bestechung soll dieser bei $0.05 \cdot |V|$ liegen, also 5 % der Größe der Wählerschaft. Konkret bedeutet dies, dass z.B. für 1000 Wähler bei Manipulation bis zu 50 Manipulatoren als Wähler hinzugefügt werden können bzw. dass

bei Bestechung maximal 50 ehrliche Wähler bestochen werden können. Außerdem können hier auch die relativen Häufigkeiten der konstruktiven Wahlkontrolle in Form von Hinzufügen eines (Klon-)Kandidaten sowie Entfernen des aktuellen Gewinners betrachtet werden (siehe Abschnitt 4.3). Ein Klon wird hierbei so generiert, dass er auf jeder individuellen Präferenzliste zufällig entweder genau eine Position über oder unter dem Gesamtsieger steht. Schlussendlich werden die Ergebnisse der Wahlkontrolle in 3 Kategorien unterteilt:

1. Fälle, in denen der ausgezeichnete Kandidat schon von vornherein Sieger war (und die Kontrollmaßnahme nicht angewandt werden musste),
2. Fälle, in denen die Kontrollmaßnahme angewandt wurde und erfolgreich war,
3. Fälle, in denen die Kontrollmaßnahme angewandt wurde und nicht erfolgreich war.

5.2 Implementation

Durchgeführt wurden die Simulationen wiederum mithilfe der Statistik-Software *R*. Dazu wurden zuerst Funktionen für die verschiedenen Modelle geschrieben, mit denen Wahlsituationen („freqs“) bzw. Profile („prefs“) zufällig mit den verschiedenen oben genannten Eigenschaften generiert werden können. Hierbei werden zunächst mithilfe der Funktion *permutations* aus dem R-Paket *gtools* alle möglichen Permutationen der Zahlen 1 bis m (Kandidatenzahl) bestimmt, welche wiederum allen möglichen Rankings entsprechen. Aus diesen wird dann basierend auf den Wahrscheinlichkeiten, die die verschiedenen Modelle zugrunde legen, zufällig mit Zurücklegen gezogen. Im IC-Modell werden die Rankings aller Wähler unabhängig voneinander mit gleicher Wahrscheinlichkeit gezogen, für das IAC-Modell wird zuvor zufällig eine Wahrscheinlichkeitsverteilung aus einer Gleichverteilung über alle Wahrscheinlichkeitsverteilungen gezogen, bevor bezüglich dieser jedem Wähler ein Ranking zugewiesen wird. Im Pólya-Modell liegt zunächst eine Gleichverteilung über die Rankings vor, die allerdings nach jeder Ziehung entsprechend des Ergebnisses der Ziehung angepasst wird. Für das räumliche Modell wurden zunächst (je nach Konfiguration unabhängig oder abhängig) die Attribute simuliert und entsprechend der Strukturvorgaben die Kandidaten gezogen. Im Anschluss wurde für jeden Wähler entsprechend der euklidischen Distanzen zu den Kandidaten das Ranking bestimmt und diese Rankings zu Profilen bzw. Wahlsituationen zusammengefasst.

Ein Profil wird hier als eine Matrix mit m Spalten (benannt nach den ersten m Großbuchstaben des Alphabets) und n Zeilen dargestellt. Jede Zeile enthält die Zahlen 1 bis m , die für die vergebenen Ränge stehen. Rang 1 bezeichnet hierbei, entgegen der Konvention, den höchsten Rang (also den 1. Platz). Eine Wahlsituation entspricht hingegen einer Matrix mit $m!$ Zeilen (ein Ranking pro Zeile) und $m+1$ Spalten. Dabei stellt die letzte Spalte die Anzahl dar, wie oft das zugehörige Ranking aus dieser Zeile abgegeben wurde.

Des Weiteren wurden die verschiedenen Wahlregeln implementiert, die ein erzeugtes Profil als Input nehmen und die Einzelstimmen zu einer Rangliste aggregieren und ausgeben. Eine hierbei zu beachtende Tatsache ist, dass, sollten Ties (Unentschieden) im Wahlergebnis auftreten, diese ursprünglich immer von der Implementation automatisch alphabetisch sortiert werden. Das bedeutet, dass Kandidat *A* eine geringfügig höhere Gewinnchance hat als die anderen Kandidaten, da für ihn auch schon ein Tie gegen beliebige andere Kandidaten zum Sieg reichen würde. Diese leichte Verzerrung kann umgangen werden, indem Ties stattdessen zufällig aufgelöst werden, was in gleichen Gewinnchancen für jeden Kandidaten resultiert. Somit muss auch nicht jede Manipulation für jeden Kandidaten separat berechnet werden, sondern es reicht, einen beliebigen aber festen, exemplarischen Kandidaten auszu-

wählen.

Zusätzlich wurden auch die Manipulations- und Bestechungsalgorithmen programmiert, die auf die Funktionen für die Wahlregeln zurückgreifen und die Zusammensetzung des Profils so verändern (Wähler hinzufügen oder vorhandene Wähler bestechen), dass der als Input gegebene ausgezeichnete Kandidat mit minimalem Aufwand (minimaler Manipulatorzahl) die Wahl gewinnt, also an erster Stelle des aggregierten Rankings steht. Außerdem wurden auch für das Hinzufügen eines Klonkandidaten (der in jeder Präferenzliste zufällig genau vor oder genau hinter dem ursprünglichen Gewinners eingefügt wird) bzw. das Löschen des aktuellen Gewinners (Wahlkontrolle) kurze Programme geschrieben, die jeweils ausgeben, ob die jeweilige Methode den ausgezeichneten Kandidaten erfolgreich zum Wahlsieger machen konnte oder nicht, oder ob er gar schon von vornherein der Wahlsieger war und die Methode deshalb nicht angewandt werden musste.

Anschließend wurden die Algorithmen auf jedes der je 100.000 Profile pro Modell(konfiguration) angewandt (für die computational deutlich aufwendigere STV-Wahl wurden nur die ersten 25.000 Profile betrachtet) und anschließend die jeweils benötigten Anzahlen in einem Vektor zusammengefasst. Im Anschluss daran wurden diese Vektoren als *.rds*-Dateien gespeichert, die in *R* mit der Funktion *readRDS* eingelesen werden können.

Diese Dateien sind nach folgendem Schema benannt: Befindet sich am Anfang des Namens ein „br“, so steht dies für Bribery, also Bestechung. Das Präfix „a“ (add) steht hingegen für das Hinzufügen eines Klonkandidaten und „d“ (delete) für das Entfernen des ursprünglichen Siegers. Findet sich keine dieser Vorsilben, handelt es sich um Manipulation. Der zweite Bestandteil des Namens ist die verwendete Wahlregel. Dabei steht „p“ für Pluralität, „c“ für Copeland, „b“ für Borda und „s“ für STV. Der letzte Teil des Namens bezeichnet das zugrunde liegende Modell, wie in den Tabellen 4 und 5 angegeben. Somit bezieht sich beispielsweise die Datei *brbsp1000.rds* auf Borda-Bestechung im räumlichen Modell mit 1000 Wählern und 5 Kandidaten, während z.B. *cic3.rds* die Werte für Copeland-Manipulation im IC-Modell mit 3 Kandidaten und 1000 Wählern beinhaltet.

Ebenso wurden die generierten Profile jeweils in einem *.rds*-File gespeichert, benannt nach dem entsprechenden Modell (z.B. *iac1000.rds*, etc.). Hier wurden in den meisten Fällen nur die Wahlsituationen gespeichert, da diese weniger Speicherplatz benötigen (für 1000 Kandidaten und 5 Wähler: $5! \cdot (5 + 1) = 720$ Einträge statt $5 \cdot 1000 = 5000$ Einträge für das zugehörige Profil). Die Ausnahme bilden die Modelle mit 7 Kandidaten; hierfür sind die Profile weniger speicherintensiv. Mithilfe der programmierten Funktion *freqs2prefs* kann eine Wahlsituation aber einfach in ein zugehöriges Profil transformiert werden (da die Wähler als ununterscheidbar angenommen werden). Mit der Funktion *prefs2freqs* kann das Gegenteil

erreicht werden. Die Berechnungen der Manipulationen und Bestechungen, sowie die Generation der Profile wurden größtenteils parallelisiert mithilfe der R-Pakete *doSNOW* (Windows) und *parallel* (Linux) auf verschiedenen Rechnern durchgeführt, um den zeitlichen Aufwand zu verringern.

5.3 Ergebnisse

Im folgenden Abschnitt werden die Ergebnisse der Simulationsstudie vorgestellt. Dabei wird zuerst auf die Struktur der generierten Profile eingegangen. Hierbei ist vor allem interessant, wie sich Unterschiede in der Homogenität der Wähler auswirken. Daher wird zum einen untersucht, wie viele der möglichen Rankings von den Wählern überhaupt tatsächlich abgegeben werden, und zum anderen, wie ähnlich sich diese Rankings sind. Für die räumlichen Modelle, die keinen Homogenitätsparameter besitzen, wird zusätzlich noch gesondert betrachtet, wie sich die Homogenität bei verschiedenen Modellkonfigurationen verhält. Danach werden Resultate für Manipulation, Bestechung und Kontrolle betrachtet.

5.3.1 Struktur der Profile

Um den Effekt der Homogenität der Wählerschaft beobachten zu können, sollen zunächst für jedes Profil die Anzahl der unterschiedlichen abgegebenen Rankings betrachtet werden. Im Modell *ic1000* werden im Schnitt alle $120 (= 5!)$ Rankings angenommen, bei *iac1000* etwa 113. Für *pol1000* und *sp1000* liegen die durchschnittlichen Werte deutlich niedriger, bei 43 bzw. 67 (siehe Abbildung 3). Bei diesen beiden Modellen steigen auch bei zehnfacher Wählerzahl (*pol10000* bzw. *sp10000*) die Anzahlen nicht beträchtlich (auf 59 bzw. 85).

Es fiel allerdings bei den Simulationen von Manipulation und Bestechung (siehe Abschnitt 5.3.2 und 5.3.3) auf, dass Profile, die aus dem räumlichen Modell generiert wurden, unabhängig von der verwendeten Wahlregel, tendenziell schwerer zu manipulieren sind (d.h. es sind mehr Manipulatoren notwendig) als solche, die auf dem Pólya-Modell beruhen. Daher muss noch eine zweite Struktureigenschaft betrachtet werden, nämlich die Ähnlichkeit der abgegebenen Rankings.

Dafür wird für jedes Profil Kendalls Konkordanzkoeffizient („Kendalls W “) berechnet. Sei hierfür R_{ij} der Rang, auf dem Kandidat j im Ranking von Wähler i rangiert. Dann wird zuerst für jeden Kandidaten die Summe aller erhaltenen Ränge bestimmt ($T_j = \sum_{i=1}^n R_{ij}$). Der durchschnittliche Wert der Rangsummen beträgt für m Kandidaten und n Wähler $\bar{T} = n(m+1)/2$. Die Idee hinter Kendalls W ist, dass, je ähnlicher sich die Rangsummen sind, desto verschiedener sind die abgegebenen Rankings der Wähler. Kendalls W berechnet

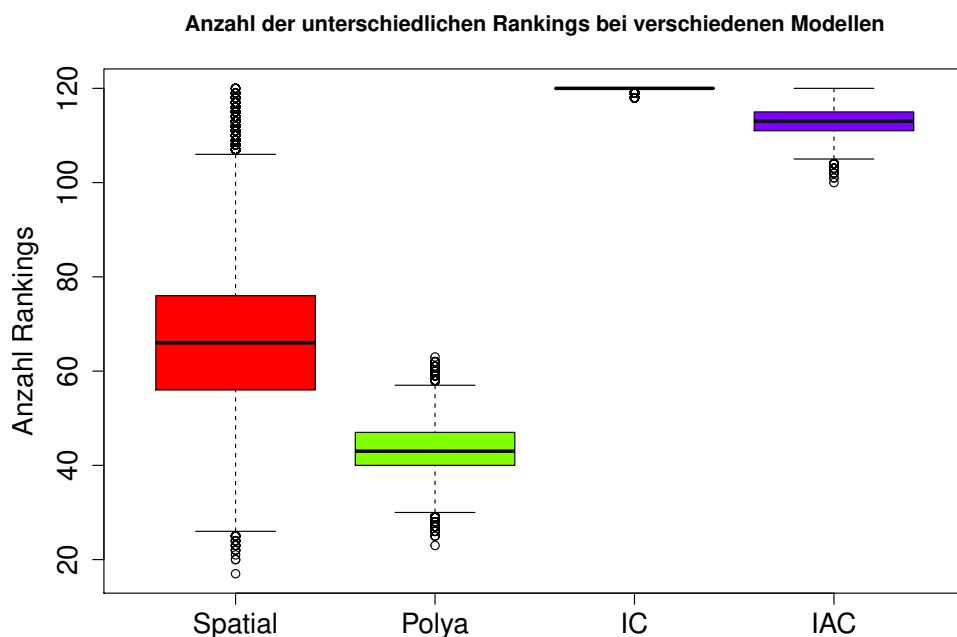


Abbildung 3: Boxplot der Anzahl abgegebener Rankings für 1000 Wähler und 5 Kandidaten

sich schließlich wie folgt:

$$W = \frac{12 \sum_{j=1}^m (T_j - \bar{T})^2}{n^2(m^3 - m)}. \quad (18)$$

W kann somit maximal 1 (wenn alle Wähler das gleiche Ranking abgeben) und minimal 0 werden (wenn alle $T_j = \bar{T}$ sind).

(vgl. Kendall und Smith 1939, S. 276)

Somit erhält man einen Wert dafür, wie ähnlich oder unähnlich sich die Rankings der Wähler sind. Abschließend werden die errechneten W 's aller Profile gemittelt, um verlässliche Werte zu erhalten.

Hierbei ist auffällig, dass das W für räumliche Profile deutlich größer ist als für Pólya-Modelle (sowie IC- und IAC-Modelle). Bei den drei letztgenannten hat das durchschnittliche W einen Wert von 0.077 für Pólya-Profilen und 0.001 bzw. 0.004 bei IC und IAC. Beim räumlichen Modell liegt der Wert hingegen bei etwa 0.260.⁸ Dies ist auf die spezielle räumliche Struktur des Modells zurückzuführen, die dazu führt, dass Wähler, die in der Attributebene nahe beieinander liegen, auch gleiche bzw. sehr ähnliche Rankings besitzen. Des Weiteren gilt

⁸Basierend auf den jeweils ersten 10000 Profilen.

für Kandidaten, die weiter am Rand der Hyperebene liegen, dass sie tendenziell seltener auf Platz 1 und relativ häufig auf dem letzten Platz gerankt werden, da die Mehrzahl der Wähler weit von ihnen entfernt ist. Somit kann dieser „Randkandidat“ zumeist keinen hohen Score erreichen bzw. wird beim paarweisen Vergleich von einem oder mehreren anderen Kandidaten teils sehr klar dominiert (d.h. die meisten Wähler ranken diesen „Randkandidaten“ hinter gewissen anderen Kandidaten). Unabhängig vom Wahlsystem ist dann eine hohe Anzahl von Manipulatoren oder bestechlichen Wählern nötig, um den Kandidaten zum Sieger zu machen.

Bei den betrachteten Profilen aus Pólya-Modellen liegt hingegen keine räumliche Struktur zugrunde. Hier treten zwar allgemein nur wenige Rankings auf, da aufgrund der Homogenität die Wahrscheinlichkeiten für bereits angenommene Rankings steigt, allerdings können diese Rankings mitunter sehr verschieden sein. Die bloße Anzahl der angenommenen Rankings ist also, für sich genommen, kein Indiz für schwere Manipulierbarkeit, vielmehr ist es entscheidend, ob diese wenigen Rankings sich auch untereinander ähnlich sind. Für das räumliche Modell ist diesbezüglich ein deutlich stärkerer negativer Zusammenhang zwischen der Anzahl der Rankings und dem dazugehörigen W erkennbar (siehe Abbildung 4).

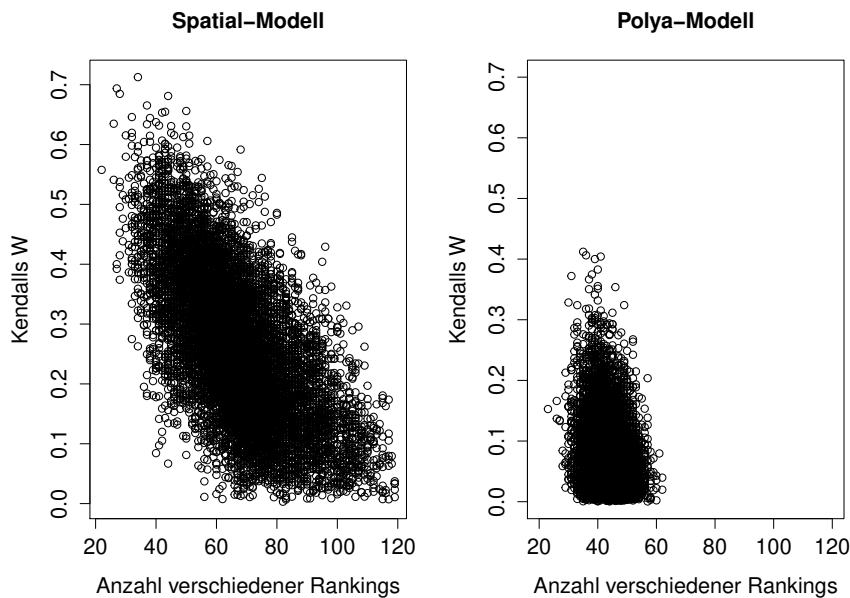


Abbildung 4: Scatterplot des Zusammenhangs der Anzahl unterschiedlicher Rankings und Kendalls W im räumlichen und im Pólya-Modell

Eine weitere Beobachtung ist, dass eine Erhöhung der Attributzahl auf 10 (*sp1000ten*) einerseits das oben beschriebene W nur leicht erhöht (Ähnlichkeit der Rankings nimmt zu),

auf etwa 0.266. Allerdings erhöht sich auch die Anzahl der unterschiedlichen abgegebenen Rankings auf ca. 89. Für besonders gemäßigte Kandidaten (*sp1000homo*) sowie für extremere Kandidaten (*sp1000hetero*) ist das W etwas größer (jeweils 0.269). Die Anzahl der einzigartigen Rankings gleicht aber im Schnitt bei beiden Modellen der des Ausgangsmodells (67). Werden die Attribute nicht unabhängig generiert (*sp1000ni*), sinkt die Anzahl der realisierten Rankings allerdings drastisch auf ca. 32. Jedoch besitzen diese Rankings auch im Schnitt ein etwas niedrigeres W von etwa 0.23, das heißt sie sind einander unähnlicher. Dies ist eine Folge der Abhängigkeitsstruktur, wie auch auf Abbildung 5 für 2 Attribute (im Vergleich zu Abbildung 6) ersichtlich wird.⁹

Im Speziellen sind die Wähler (graue Punkte) hier ellipsenförmig angeordnet, sodass es v.a. wenig Ausreißer in Regionen gibt, wo ein Attribut stark positiv ausgeprägt ist und das andere stark negativ. Stattdessen befinden sich die Wähler entlang der großen Halbachse der Ellipse verteilt. Befinden sich allgemein die Kandidaten (rote Punkte) nahe beieinander im Zentrum der Ebene, so resultiert das in einem sehr kleinen W . Treten hingegen deutliche Ausreißer unter den Kandidaten auf oder sind die Kandidaten abseits des Zentrums konzentriert, liefert das W hohe Werte.

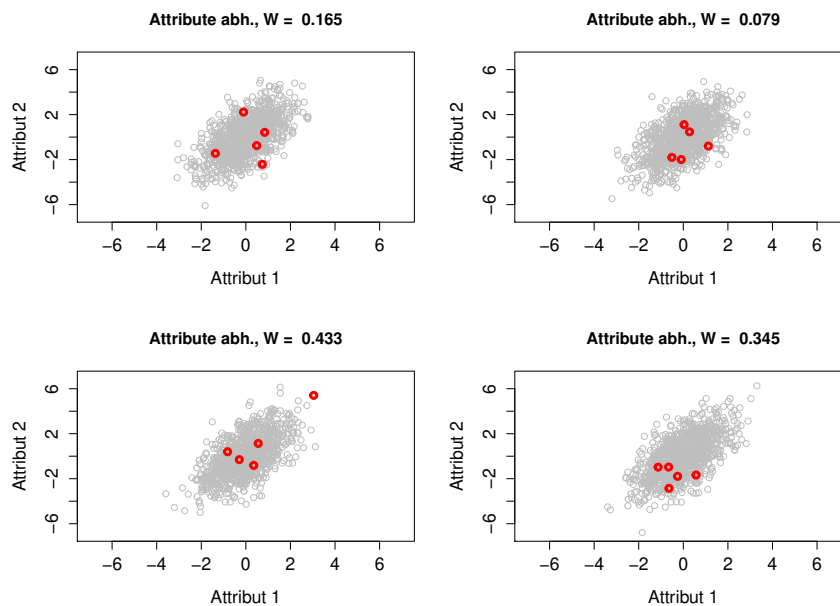


Abbildung 5: Verteilung der Wähler (grau) und Kandidaten (rot) für 2 nicht unabhängig generierte Attribute und Kendalls W der zugehörigen Rankings

⁹Hierbei handelt es sich um je 4 zufällig erzeugte Profile mit 1000 Wählern und 5 Kandidaten mit abhängiger bzw. unabhängiger Ziehung der Kandidaten aus der Wählermenge.

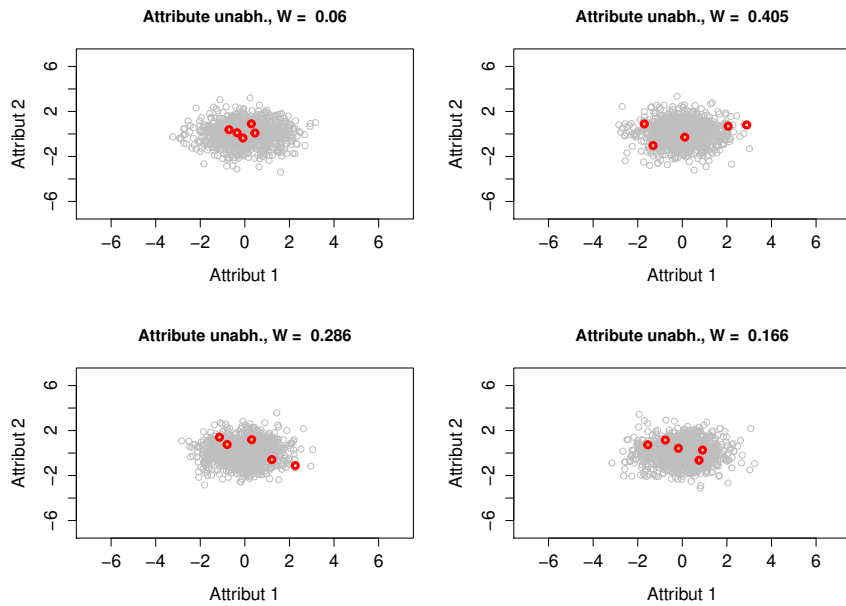


Abbildung 6: Verteilung der Wähler (grau) und Kandidaten (rot) für 2 unabhängig generierte Attribute und Kendalls W der zugehörigen Rankings

Zuletzt kann noch betrachtet werden, wie sehr sich die Implementierung von gemäßigten (*sp1000homo*) bzw. extremen Kandidaten (*sp1000hetero*) tatsächlich auf die Variabilität der Attributwerte der Kandidaten auswirkt. Hierfür werden die durchschnittlichen Varianzen der 4 Attribute über die 5 Kandidaten betrachtet. Eine kleine Varianz entspricht hier einer Ähnlichkeit der Kandidaten bzgl. eines Attributs, eine große Varianz hingegen, dass die Kandidaten, ein Attribut betreffend, sehr unterschiedlich eingestellt sind. Hier ergeben sich Werte von etwa 0.74 für gemäßigte Kandidaten und 1.26 für extreme Kandidaten. Im Vergleich dazu stehen bei neutraler Ziehung (nicht abhängig von Extremität) der Kandidaten Varianzen von ca. 0.99 zu Buche.¹⁰

5.3.2 Ergebnisse I: Manipulation

Für die anschließenden Vergleiche der Ergebnisse werden Density-Plots verwendet. Diese schätzen auf Basis eines Histogramms die Dichte der Daten mithilfe einer Kernfunktion. Dies erleichtert den Vergleich der jeweiligen geschätzten Dichten, da sie in einer Grafik darstellbar sind und man nicht verschiedene Histogramme miteinander vergleichen muss. In *R* wurde hierfür die Funktion *density* aus dem Paket *stats* verwendet. Als Kernfunktionen wurden Gaußkerne benutzt. Des Weiteren wurde als Bandweite die Standardeinstellung „nrd0“

¹⁰Werte beruhend auf 1000 zufälligen Profilen der jeweiligen Konfiguration.

gewählt, die die Daumenregel von Silverman (1986, S. 48) implementiert. Außerdem wird für die Darstellung der Dichten für Fälle mit 100 Wählern die Anzahl der äquidistanten Punkte, an denen die Kerndichte geschätzt wird, auf Werte zwischen 20 und 35 heruntersgesetzt, um einen bestmöglichen Tradeoff zwischen Glattheit der Schätzung und Anpassung an die Daten zu erreichen. Für die Fälle mit 1000 oder 10000 Wählern wird jeweils die Standardeinstellung von 512 Punkten verwendet.

Zunächst wird, getrennt nach den verwendeten Modellen zur Profilerzeugung, dargestellt, wie die Anzahl der Manipulatoren für verschiedene Wahlregeln verteilt ist. Hierfür werden jeweils die Settings mit 1000 Wählern und 5 Kandidaten betrachtet.¹¹ Außerdem wurden Fälle ausgeschlossen, in denen der ausgezeichnete Kandidat bereits der ursprüngliche Wahlsieger war. Die Schwierigkeit der Manipulation kann mithilfe der Dichtekurven gut beschrieben werden. Nimmt die Dichte vor allem für kleine Manipulationszahlen hohe Werte an, so bedeutet dies, dass die Manipulation tendenziell mit wenig Aufwand (also wenigen Manipulatoren) durchgeführt werden kann. Schwere Manipulierbarkeit zeichnet sich hingegen dadurch aus, dass die Dichte kleine Werte für kleine Anzahlen annimmt (und dementsprechend große Werte für große Anzahlen).

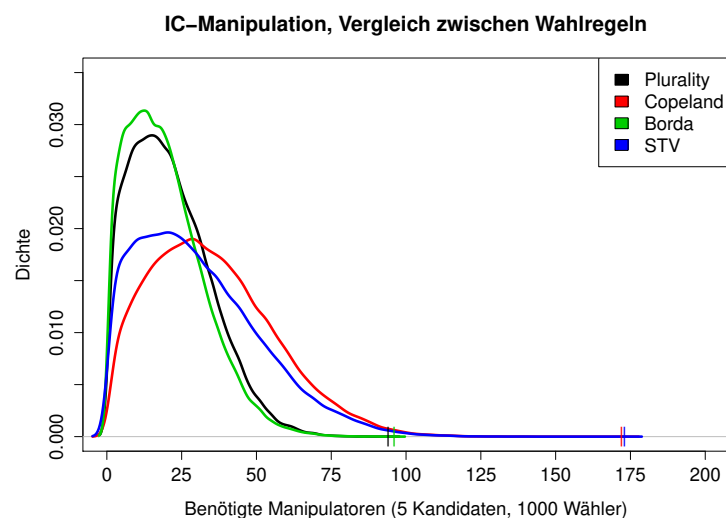


Abbildung 7: Manipulation für verschiedene Wahlregeln beruhend auf dem IC-Modell

In Abbildung 7 werden Dichten der Manipulationen im *ic1000*-Modell dargestellt. Die klei-

¹¹In allen Grafiken, in denen auch die STV-Wahl betrachtet wird, werden für alle 4 Kurven nur die ersten 25000 Einträge der entsprechenden Vektoren betrachtet, da aufgrund des computationalen Aufwands für STV nur jeweils Ergebnisse für eben diese zugehörigen Profile berechnet wurden.

nen Striche auf der x-Achse in den jeweiligen Farben entsprechen der maximalen Anzahl an Manipulatoren, die bei der Wahlregel mit der entsprechenden Farbe benötigt wurde. Zudem wirken sie dem Overplotting-Effekt entgegen, sodass man erkennen kann, wo Linien, die sich überschneiden, enden. Es fällt auf, dass vor allem bei der Pluralitäts- und Bordaregel vermehrt kleine Werte auftreten. Für beide ergeben sich maximale Werte, die jeweils unter 100 Manipulatoren liegen. Ausgehend von der Grafik identifiziert man unter der Annahme eines unterliegenden IC-Modells die Borda-Regel als am leichtesten zu manipulierende Wahlregel. Sie erreicht ihre maximale Dichte schon bei etwa 12 Manipulatoren und fällt dann ab 18 Manipulatoren sehr steil ab (im Mittel werden 15.58 Manipulatoren benötigt). Für Copeland- und STV-Wahlen steigen die Kurven weniger steil an und erreichen erst bei 28 bzw. 20 Manipulatoren ihre maximale Dichte (Durchschnittswerte 28.03 und 24.52). Zudem treten auch einige Fälle auf, bei denen deutlich über 100 Manipulatoren benötigt werden.

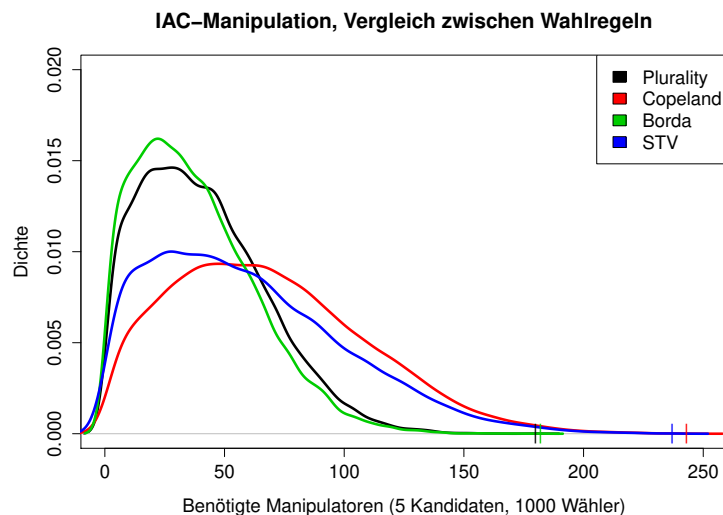


Abbildung 8: Manipulation für verschiedene Wahlregeln beruhend auf dem IAC-Modell

Ein qualitativ ähnliches Bild zeigt sich bei der Manipulation mit unterliegendem IAC-Modell (Abbildung 8). Allerdings ist hier zu beachten, dass die y-Achse kürzer sowie die x-Achse länger ist. Das bedeutet, während kleine Werte hier allgemein seltener auftreten als unter dem IC-Modell, nimmt die Dichte für größere Werte zu. Daher steigt auch die Zahl der im Durchschnitt benötigten Manipulatoren an. Die durchschnittlich benötigten Manipulatoren können der Tabelle 6 entnommen werden. Weiterhin benötigt die Copeland-Regel die im Mittel größte Anzahl von Manipulatoren und die Borda-Regel die geringste.

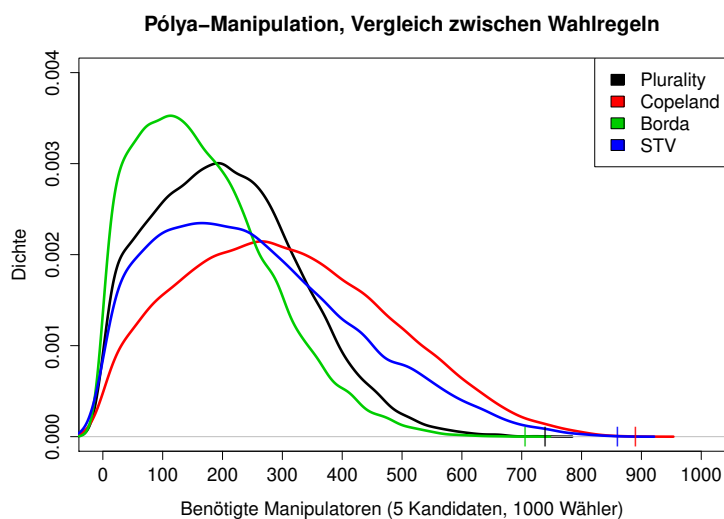


Abbildung 9: Manipulation für verschiedene Wahlregeln beruhend auf dem Pólya-Modell

Für Wahlen, die beruhend auf dem Pólya-Modell durchgeführt werden, ergibt sich wiederum ein ähnliches Bild (Abbildung 9; man beachte wiederum die geänderte Skalierung der Achsen), allerdings verstärkt sich durch die zusätzliche Homogenität der Effekt, der sich schon zuvor für das IAC-Modell gezeigt hat: Die Dichten für kleine Manipulatoranzahlen nehmen weiter ab, während tendenziell immer mehr Manipulatoren benötigt werden (bei Pluralität und Borda-Wahl maximal über 700, bei STV- und Copeland-Wahl knapp 900). Im qualitativen Vergleich verhalten sich die Wahlsysteme weiterhin wie schon unter IC- und IAC-Modellen.

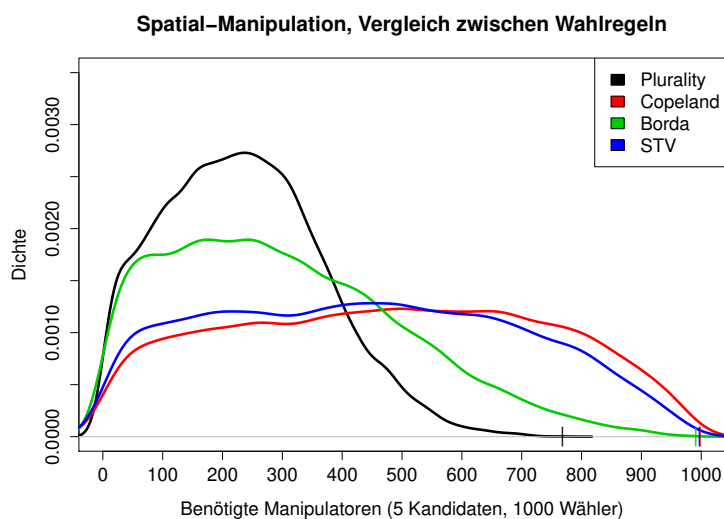


Abbildung 10: Manipulation für verschiedene Wahlregeln beruhend auf dem Spatial-Modell

Für das räumliche Modell (Abbildung 10) ergeben sich im Vergleich zum Pólya-Modell in der Grafik, die Dichten betreffend (also auf der y-Achse), keine weiteren enormen Größenunterschiede. Die Dichte der Pluralitätswahl verhält sich sehr ähnlich zum Pólya-Modell, schneidet hier aber insgesamt am schlechtesten ab. Im Gegensatz zu den anderen Modellen ist nicht die Bordawahl mit der einfachsten Manipulierbarkeit verbunden. Im Maximum werden hier in etwa genauso viele Manipulatoren gebraucht wie für STV- und Copeland-Manipulation. Die Ergebnisse der beiden letztgenannten Wahlregeln nähern sich unter Annahme des räumlichen Modells einander an, wobei für Copeland auch hier leicht bessere Ergebnisse erreicht werden. In der folgenden Tabelle werden zusätzlich die jeweils für ein bestimmtes Modell und eine bestimmte Wahlregel durchschnittliche benötigte Anzahlen an Manipulatoren angegeben. Auch hier sieht man, dass die Copelandregel modellübergreifend die höchsten Werte aufweist.

	IC	IAC	Pólya	Spatial
Plurality	16.94	32.92	166.6	187.1
Copeland	28.03	54.64	244.4	379.3
Borda	15.58	30.27	134.7	243.5
STV	24.52	47.76	203.7	350.8

Tabelle 6: Durchschnittliche Anzahlen von benötigten Manipulatoren für Wahlregel-Modell-Kombinationen

Des Weiteren ist von Interesse, wie sich weitere Parameter der Wahlen auf die Manipulati-

onsergebnisse auswirken.

a) Wählerzahl

Die folgenden Grafiken stellen zunächst am Beispiel der Copelandregel die Manipulation für die 4 unterschiedlichen Modelle und jeweils für 100, 1000 und 10000 Wähler dar (bei je 5 Kandidaten). Dabei ist auf den x-Achsen die Anzahl der Manipulatoren relativ zur Gesamtwählerzahl dargestellt, ein Wert von 0.2 bedeutet also beispielsweise, dass bei 100 Wählern 20 Manipulatoren benötigt werden, bei 1000 Wählern 200 Manipulatoren, etc. Auf den y-Achsen werden weiterhin die Dichten abgetragen.

Für das IC- und IAC-Modell ist zu sehen, dass die Manipulierbarkeit mit steigender Wählerzahl relativ immer einfacher wird. Der Grund dafür ist, dass unter diesen beiden Modellen alle Rankings in etwa gleich oft auftreten; somit treten auch im Mittel keine größeren Discrepanzen in den erhaltenen Stimmen/Punkten auf. Dadurch werden auch mit steigender Wählerzahl die paarweisen Vergleiche immer ausgeglichener und damit leichter zu manipulieren. Mit sinkender Anzahl abgegebener Rankings (von IC über IAC zu Pólya) nähern sich die Dichten für die verschiedenen Wählerzahlen immer weiter einander an. Liegen zusätzlich (im räumlichen Modell) sich ähnelnde Rankings vor, unterscheiden sich die Kurven kaum noch (siehe Abbildung 11).

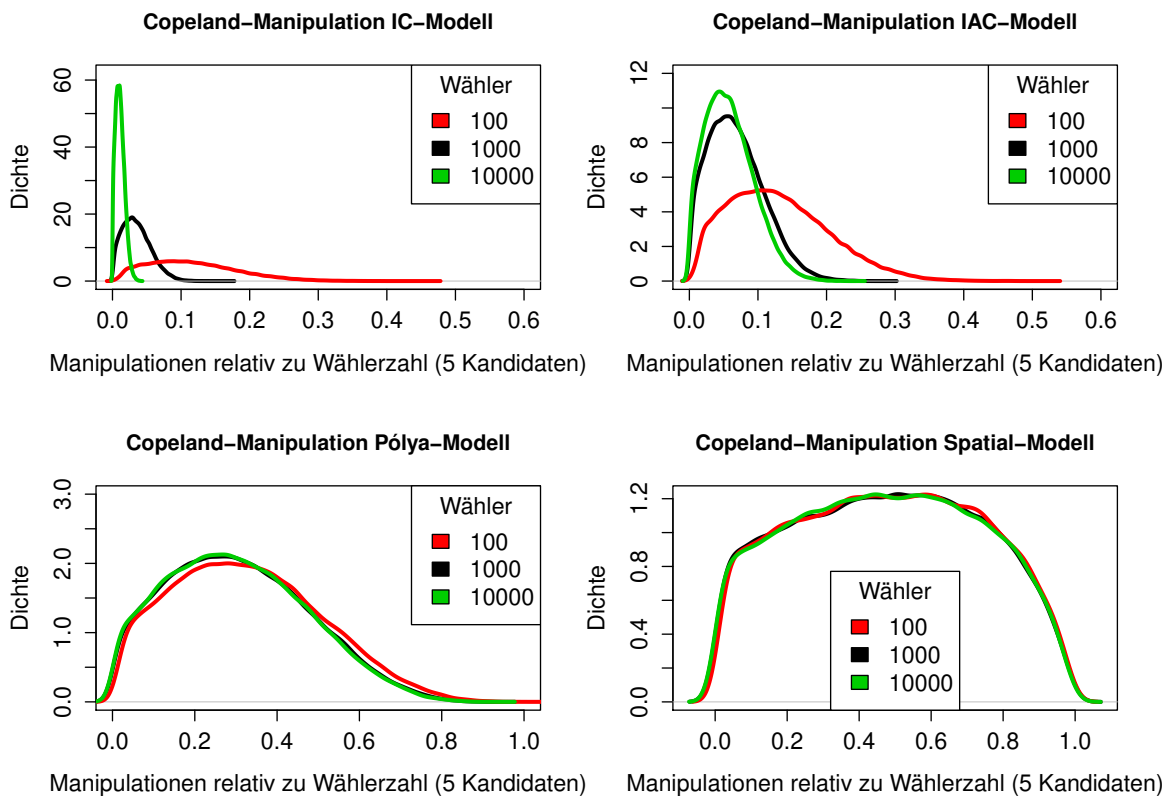


Abbildung 11: Copeland-Manipulation für verschiedene Wählerzahlen

b) Kandidatenzahl

Betrachtet man die Ergebnisse nun für die verschiedenen Modelle und jeweils für 3, 5 und 7 Kandidaten (bei je 1000 Wählern) können folgende Erkenntnisse gewonnen werden. Während sich für 7 Kandidaten keine großen Unterschiede zwischen den ersten drei Modellen ergeben, verändern sich die Dichten für 3 und 5 Kandidaten im Verlauf sehr drastisch. Ist es im IC-Modell noch am einfachsten, Wahlen mit 3 Kandidaten zu manipulieren, so ist schon im IAC-Modell das Gegenteil der Fall. Im Pólya-Modell ergibt sich ein sehr extremes Ergebnis, bei dem im Großteil der Fälle über 900 Manipulatoren nötig sind. Dies liegt daran, dass hier, aufgrund der hohen Homogenität und den für 3 Kandidaten nur 6 möglichen Rankings, vermehrt Profile auftreten, bei denen ausschließlich ein einziges Ranking abgegeben wird (und höchstens vereinzelt andere Rankings). Auch für 5 Kandidaten wird es mit abnehmender Anzahl an Rankings (von IC bis Pólya) immer schwerer erfolgreich zu manipulieren. Die Struktur der Ergebnisse für das räumliche Modell ähnelt zumindest qualitativ der des IC-Modells. Man beachte allerdings, dass hier die y-Achse um den Faktor 10 kleiner dargestellt ist. Aufgrund der tendenziell ähnlicheren Rankings im räumlichen Modell, die große

Unterschiede in paarweisen Vergleichen hervorbringen können, ergeben sich dort allgemein über alle Kandidatenzahlen höhere Manipulatorenzahlen (Abbildung 12).

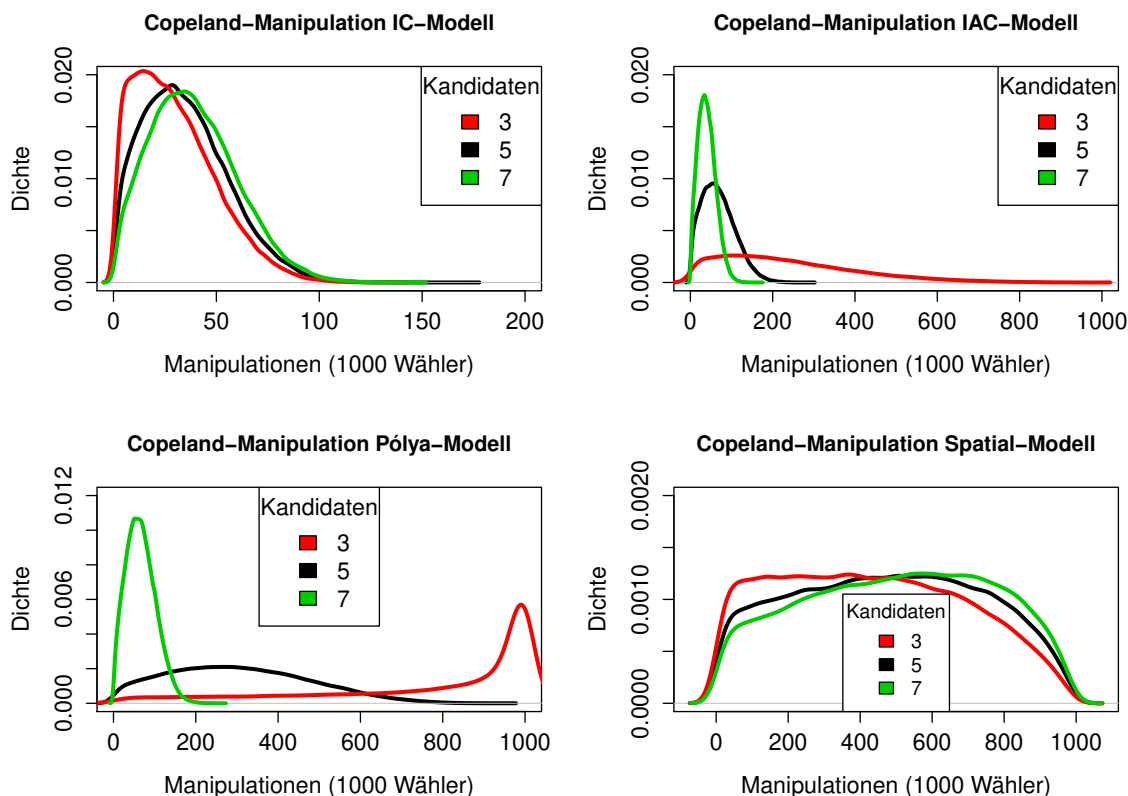


Abbildung 12: Copeland-Manipulation für verschiedene Kandidatenzahlen

c) Kandidatenstruktur (räumliches Modell)

Abbildung 13 zeigt die Manipulationsergebnisse für Copelandwahlen im räumlichen Modell für die verschiedenen Kandidatenstrukturen Neutral, Extrem und Gemäßigt. Die Dichtekurven liegen bis zu einem Wert von etwa 50 Manipulatoren noch auf gleichem Niveau, zwischen 50 und 800 Manipulatoren übersteigt die Dichte beim neutralen Modell die beiden anderen Dichtekurven fast durchgängig. Ab 800 Wählern ist das Gegenteil der Fall. Die Kurven der extremen und gemäßigten Strukturen unterscheiden sich nicht auffallend in eine Richtung. Allgemein kann aber festgehalten werden, dass für diese beiden alternativen Kandidatenstrukturen die Manipulation etwas schwerer fällt als für neutrale, nicht auf Extremität beruhende Ziehung der Kandidaten. Die Anzahl durchschnittlich benötigter Manipulatoren beträgt 388.4 für den Fall der gemäßigten und 386 für den Fall der extremen

Kandidatenstruktur (im Vergleich zu 379.3 bei neutraler Ziehung).

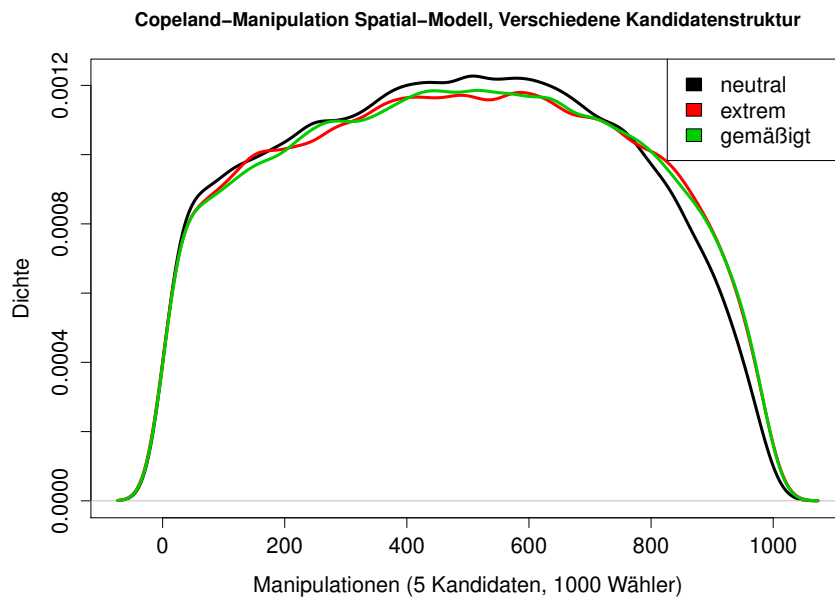


Abbildung 13: Copeland-Manipulation für verschiedene Kandidatenstrukturen im räumlichen Modell

d) Anzahl und Abhängigkeitsstruktur der Attribute (räumliches Modell)

Die Abbildungen 14 und 15 geben den Verlauf der Dichtekurven für unterschiedliche Anzahl bzw. Abhängigkeitsstruktur von Attributen wieder. Im ersten Plot kann man erkennen, dass Copeland-Manipulation für 10 unabhängige Attribute etwas schwerer fällt als für 4 Attribute, der Mittelwert der Manipulatorenanzahl beträgt hier 382.7. Die zweite Grafik verrät, dass die Copeland-Manipulation für das Modell, das auf abhängig gezogenen Attributen beruht, einfacher ist als beim Modell mit unabhängigen Attributen. Bei abhängig generierten Attributen werden im Schnitt 364.5 Manipulatoren benötigt.

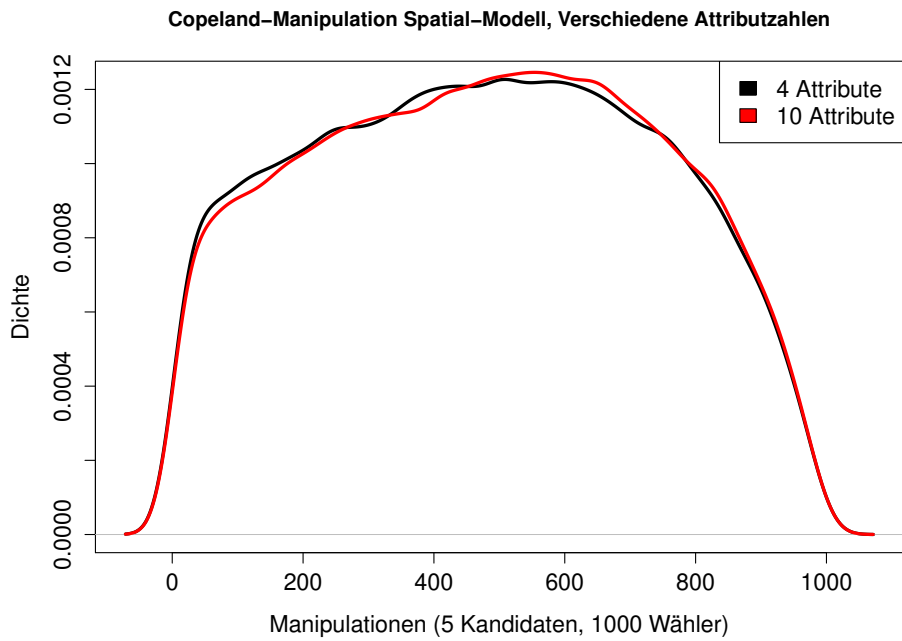


Abbildung 14: Copeland-Manipulation für verschiedene Anzahlen von unabhängigen Attributen im räumlichen Modell

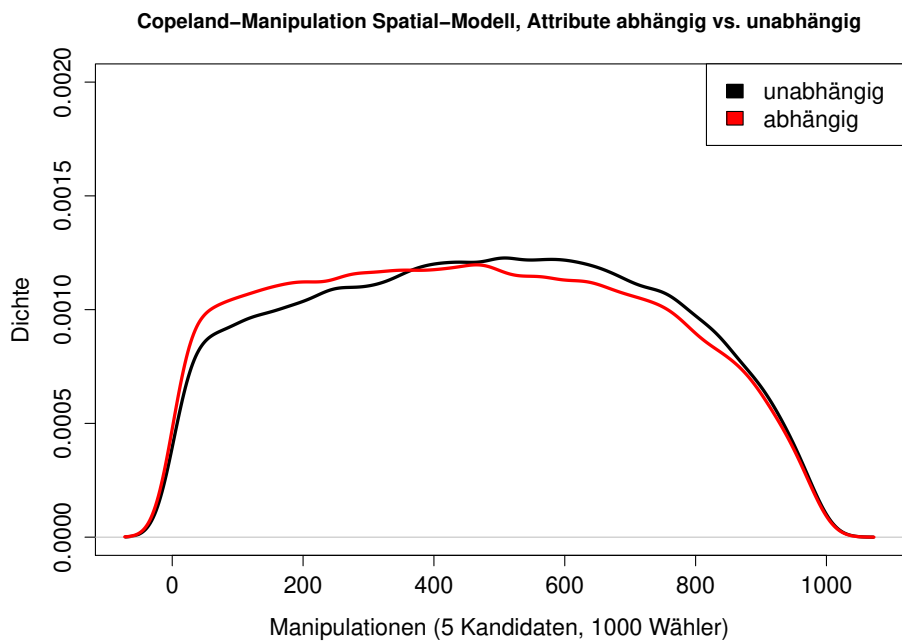


Abbildung 15: Copeland-Manipulation für abhängige und unabhängige generierte Attribute im räumlichen Modell

Interessanterweise ergeben sich beim Vergleich über verschiedene Kandidatenzahlen unterschiedlich geartete Ergebnisse für die Wahlregeln (im Gegensatz zur Variation der Wählerzahlen, für die sich keine nennenswerten Unterschiede ergeben). Für die in Abbildung 16 dargestellte Pluralitätsregel sind im Gegensatz zur Copeland-Regel bei zugrunde liegendem IC-Modell die Wahlen mit nur 3 Kandidaten am schwierigsten zu manipulieren, und nicht am leichtesten. Auch im räumlichen Modell ist der gleiche Trend erkennbar, während die Ergebnisse für IAC- und Pólya-Modell (im Vergleich zu Abbildung 12) einander ähneln.

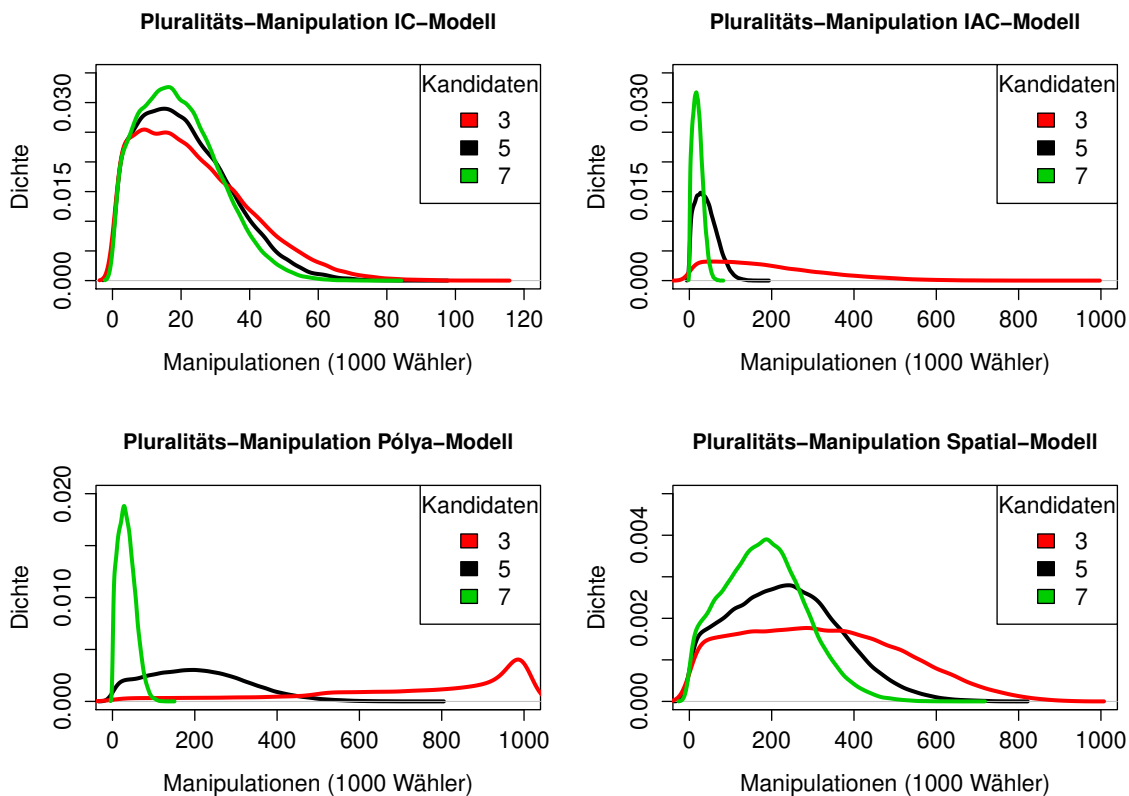


Abbildung 16: Pluralitäts-Manipulation für verschiedene Kandidatenzahlen

Auch für die verwandte Bordaregel (ebenfalls ein Scoring-Protokoll) ergeben sich ähnliche Ergebnisse (Abbildung 17). Erkennbare Unterschiede zur Pluralitätsregel liegen zum einen im räumlichen Modell, wo die Kurven näher beieinander liegen. Zum anderen ist im Pólya-Modell deutlich eine zweite Erhebung bei einer Anzahl von etwa 500 erkennbar. In diesem Bereich befinden sich die Fälle, bei denen (fast) ausschließlich ein einziges Ranking abgegeben wird, auf dem der ausgezeichnete Kandidat den zweiten Platz einnimmt. Die zweite Erhebung bei etwa 1000 entspricht analog den Fällen, wo ein einziges Ranking abgegeben

wird, auf dem der Kandidat den letzten Platz einnimmt. Die Pluralitätsregel macht zwischen diesen beiden Möglichkeiten hingegen keinen Unterschied, dort werden jeweils ca. 1000 Manipulatoren benötigt.

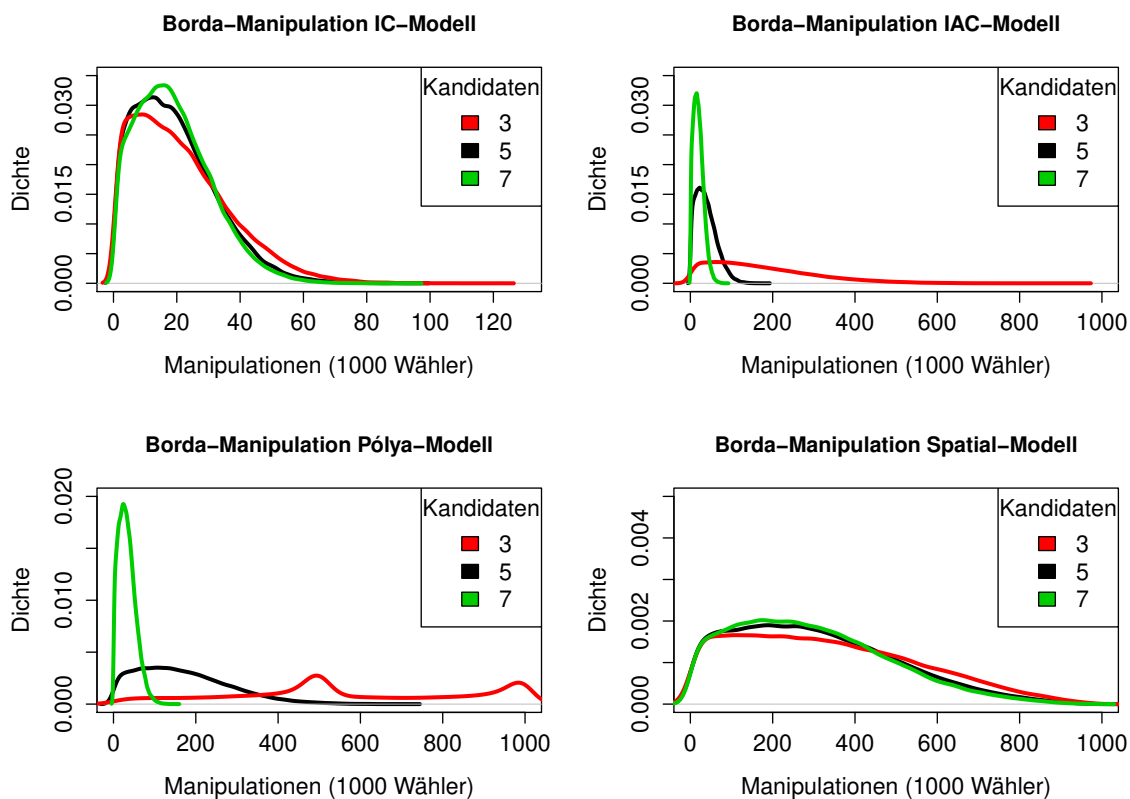


Abbildung 17: Borda-Manipulation für verschiedene Kandidatenzahlen

Die Ergebnisse der STV-Manipulation weisen wiederum qualitative Ähnlichkeit mit denen der Copeland-Manipulation auf. Hier ist erneut im IC-Modell die Wahl mit 3 Kandidaten am einfachsten zu manipulieren (im Gegensatz zu den Scoring-Protokollen). Es fällt weiterhin auf, dass die Dichte im Pólya-Modell um den Wert 1000 nicht so stark erhöht ist, außerdem unterschieden sich die Kurven im räumlichen Modell weiterhin nur minimal, wobei Wahlen für 7 Kandidaten geringfügig schwerer zu manipulieren sind (Mittelwert 457.6) als Wahlen für 3 oder 5 Kandidaten (Mittelwerte 421.3 bzw. 407).

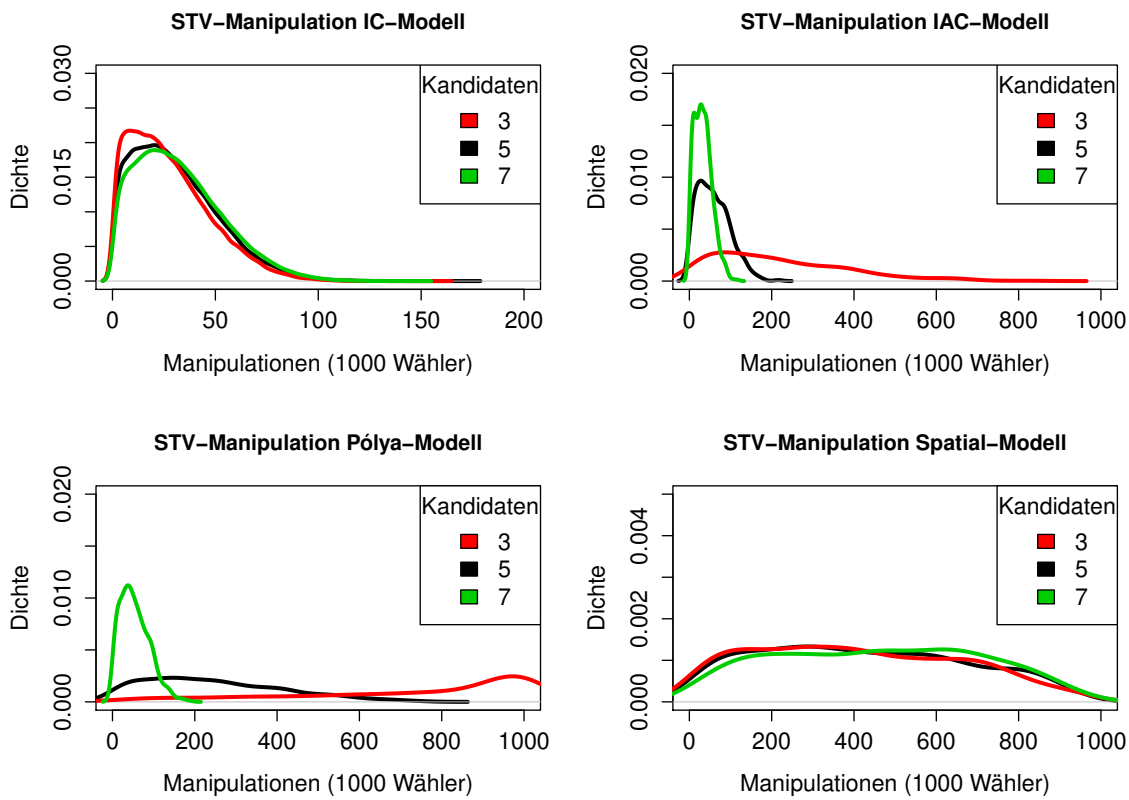


Abbildung 18: STV-Manipulation für verschiedene Kandidatenzahlen¹²

¹²Bei IAC-, Pólya- und Spatial-Modell beruht die Schätzung auf lediglich 1000 Werten (wegen hohem computationalem Aufwand der STV-Wahl)

5.3.3 Ergebnisse II: Bestechung

Die Bestechung liefert qualitativ ähnliche Ergebnisse wie die Manipulation. Allerdings sind hier die erhaltenen Werte deutlich niedriger, da die Größe der Wählerschaft im Gegensatz zur Manipulation konstant bleibt und nicht stetig mit der Anzahl der Manipulatoren mitwächst.

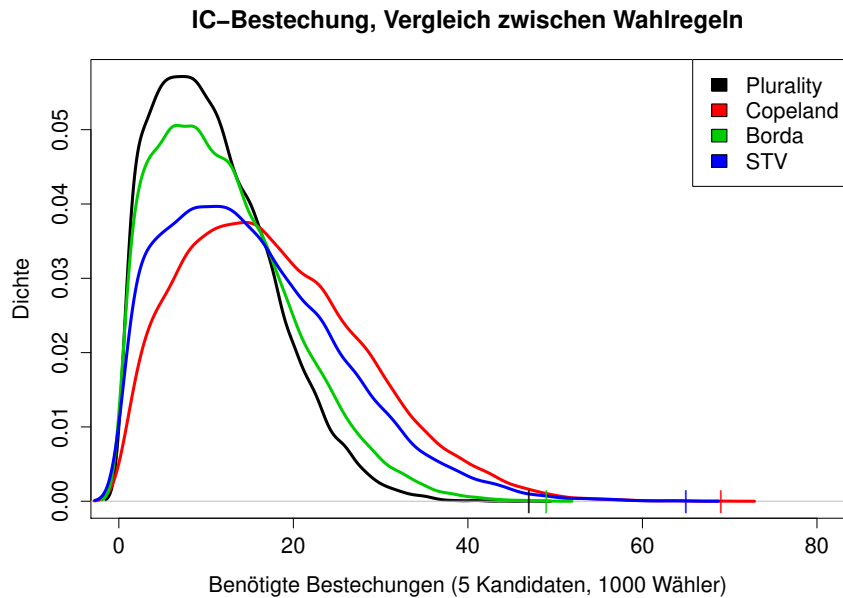


Abbildung 19: Bestechung für verschiedene Wahlregeln beruhend auf dem IC-Modell

Im Unterschied zur analogen Situation mit Manipulation, zeigt sich für Bestechung im IC-Modell (Abbildung 19) die Pluralitätswahl als am anfälligsten. Copeland ist etwas sicherer als STV (erfordert mehr Bestechungen).

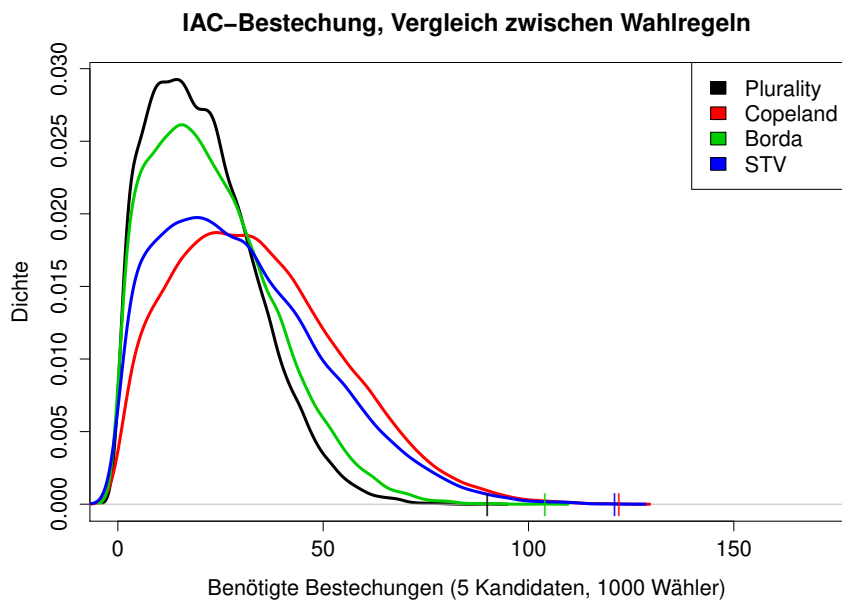


Abbildung 20: Bestechung für verschiedene Wahlregeln beruhend auf dem IAC-Modell

Ein sehr ähnliches Bild ergibt sich beim IAC-Modell (Abbildung 20). Auch hier schneidet die Pluralitätswahl hinter der Bordawahl am schlechtesten ab, während STV weitgehend von Copeland dominiert wird. Allgemein werden im IAC-Modell erneut tendenziell mehr Bestechungen benötigt als im IC-Modell.

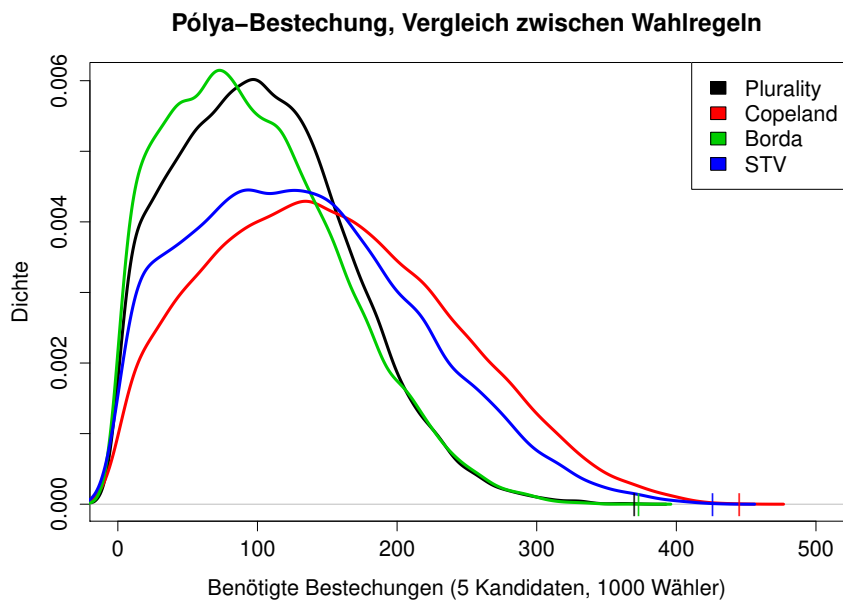


Abbildung 21: Bestechung für verschiedene Wahlregeln beruhend auf dem Pólya-Modell

Im Pólya-Modell (Abbildung 21) ist der Unterschied zwischen Pluralitäts- und Bordawahl nicht mehr so eindeutig, allerdings schneidet hier Borda etwas schlechter ab als Pluralität. Copeland erweist sich ebenfalls wieder als unempfindlichstes Wahlsystem mit bis zu 445 nötigen Bestechungen.

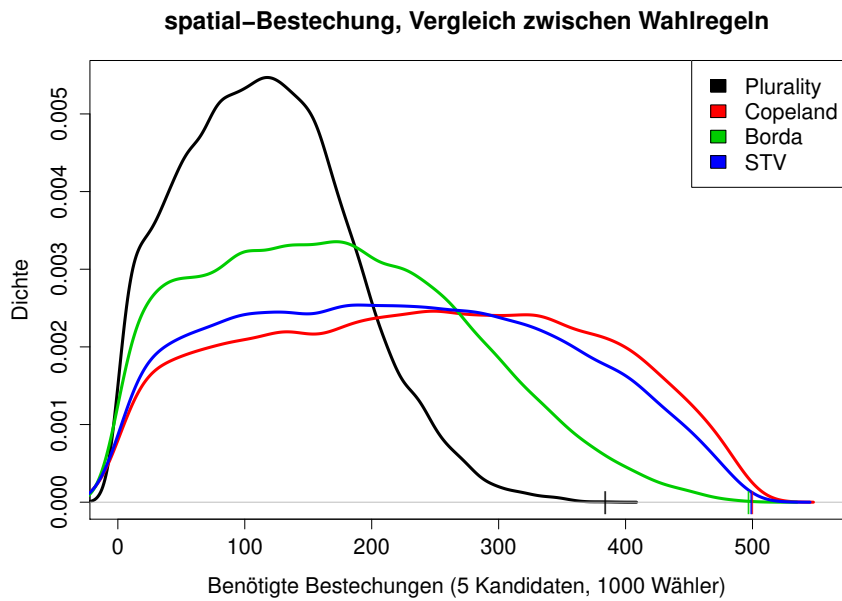


Abbildung 22: Bestechung für verschiedene Wahlregeln beruhend auf dem Spatial-Modell

Das eindeutigste Ergebnis (relativ gesehen) ergibt sich erneut im räumlichen Modell (Abbildung 22), wo der Pluralitätswahl am einfachsten mit Bestechung beizukommen ist. Des Weiteren schneidet die Bordawahl wiederum schlechter ab als STV- und Copelandwahlen. In der folgenden Tabelle 7 kann auch für das Bestechungsproblem entnommen werden, wie viele Bestechungen jeweils im Schnitt für alle Kombinationen von Modellen und Wahlregeln benötigt werden.

	IC	IAC	Pólya	Spatial
Plurality	8.667	16.66	83.48	93.74
Copeland	14.2	27.54	122.4	189.8
Borda	9.798	18.67	78.62	137.7
STV	12.41	24.46	106.8	176.6

Tabelle 7: Durchschnittliche Anzahlen von benötigten Bestechungen für Wahlregel-Modell-Kombinationen

Auch für das Bestechungsproblem können im Weiteren die Auswirkungen von Änderungen in den Modellsettings betrachtet werden. Dabei werden wiederum die Ergebnisse für die Copeland-Wahlen herangezogen, die sich auch am resistantesten gegenüber Bestechung erwiesen haben.

a) Wählerzahl

In Abbildung 23 ist zu erkennen, dass in allen 4 Modellen ein Zusammenhang zwischen der Anzahl der Wähler und der Einfachheit der Bestechung besteht. Je mehr Wähler, desto geringer ist der Aufwand der Bestechung. Dementsprechend sind die im Mittel nötigen Bestechungen (relativ zur Wählerzahl) für 10000 Wähler am geringsten, allerdings sind die Unterschiede im Pólya- und Spatial-Modell nur geringfügig.

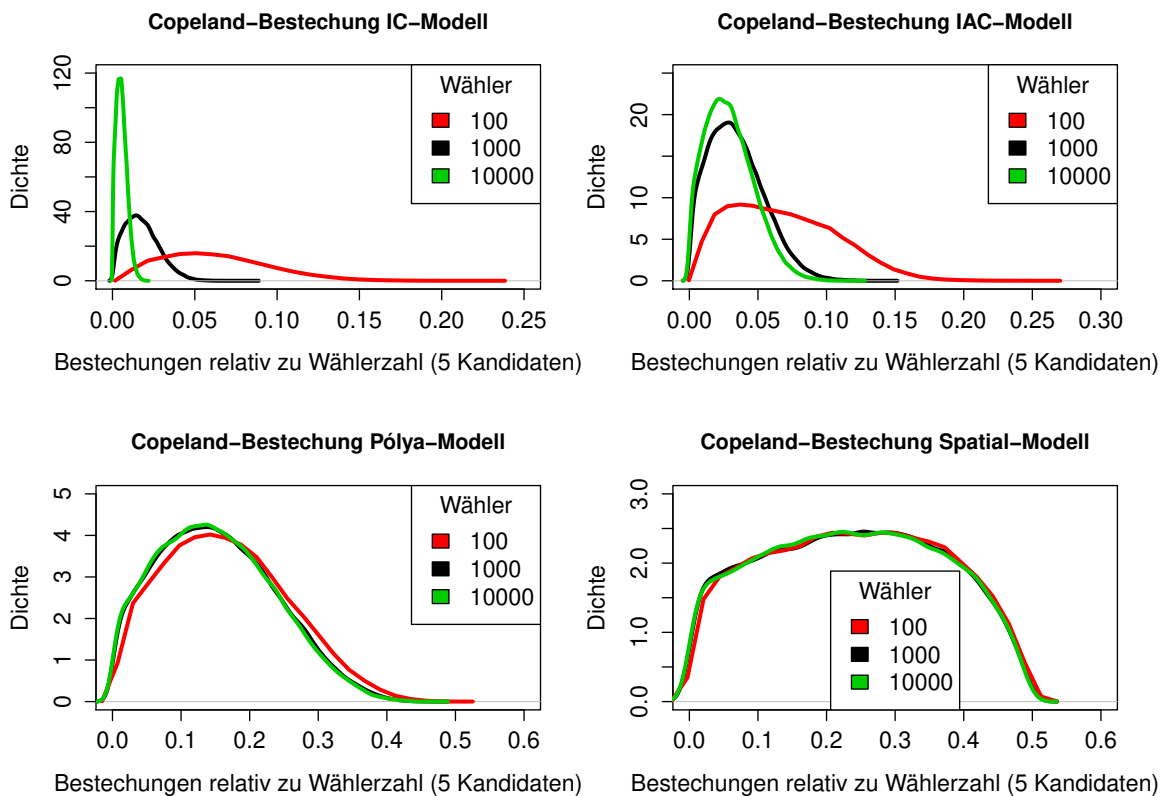


Abbildung 23: Copeland-Bestechung für verschiedene Wählerzahlen

b) Kandidatenzahl

Aus Abbildung 24 lassen sich die gleichen Erkenntnisse ziehen wie schon im Manipulationsfall aus Abbildung 12.

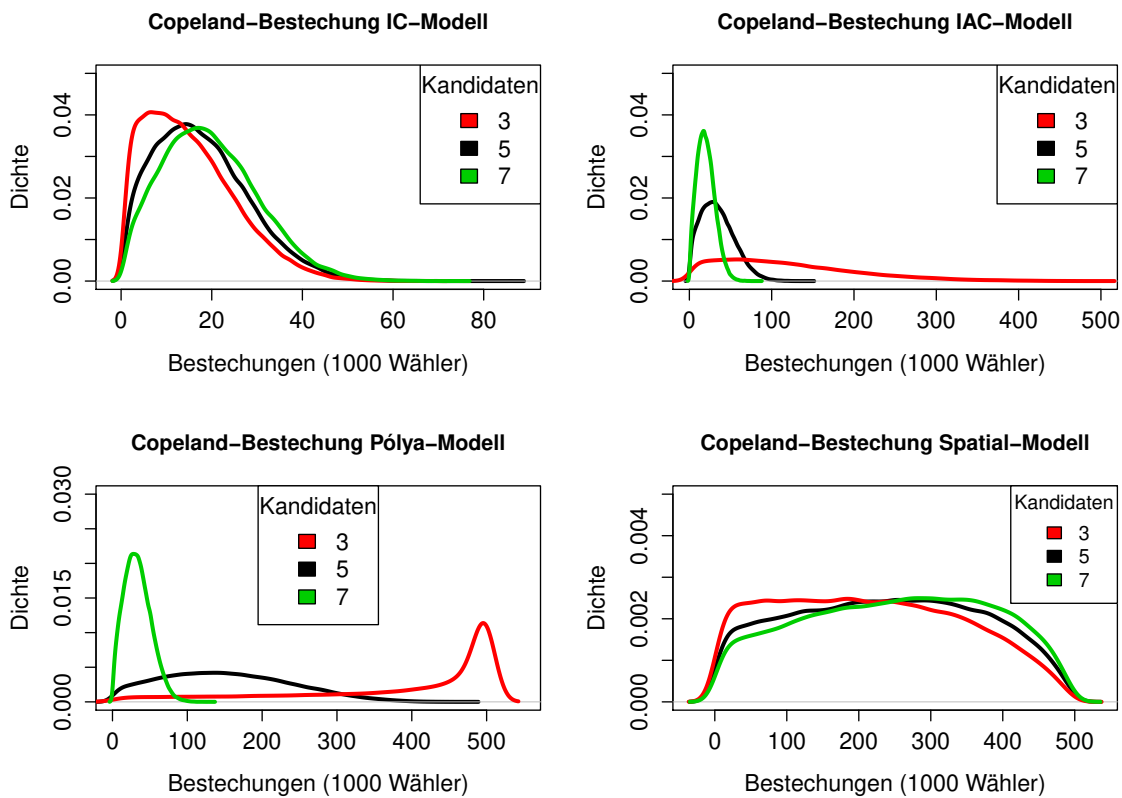


Abbildung 24: Copeland-Bestechung für verschiedene Kandidatenzahlen

c) Kandidatenstruktur (räumliches Modell)

Für die verschiedenen Kandidatenstrukturen, dargestellt in Abbildung 25, ergibt sich ebenfalls ein ähnliches Bild wie im Fall der Manipulation. Allerdings sind die resultierenden Mittelwerte deutlich geringer, da im Höchstfall nur 501 Wähler bestochen werden müssen, im Gegensatz zu 1001 maximal nötigen Manipulatoren. Qualitativ unterscheiden sich die Dichtekurven aber kaum. Die Mittelwerte betragen 189.8 für die neutrale Kandidatenstruktur, sowie 193.2 für extremere Kandidaten und 194.4 für die gemäßigeren Kandidaten.

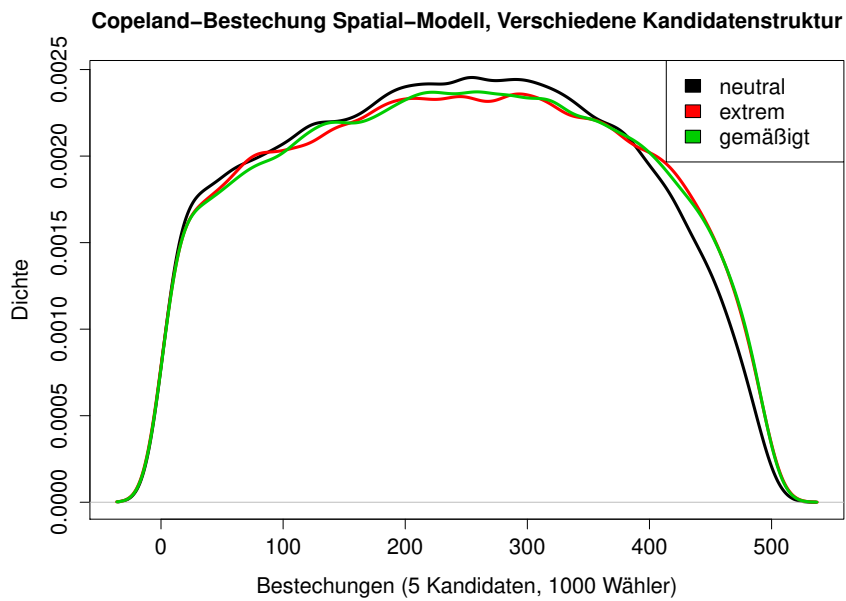


Abbildung 25: Copeland-Bestechung für verschiedene Kandidatenstrukturen im räumlichen Modell

d) Anzahl und Abhängigkeitsstruktur der Attribute (räumliches Modell)

Auch für die Anzahl und Abhängigkeitsstruktur der Attribute (Abbildungen 26 und 27) ergeben sich qualitativ ähnliche Grafiken wie für die Manipulation. Die Mittelwerte liegen bei 191.6 für das Modell mit 10 Attributen sowie bei 182.5 für nicht unabhängig generierte Attribute (im Vergleich zu 189.8 für 4 Attribute bzw. unabhängige Attribute).

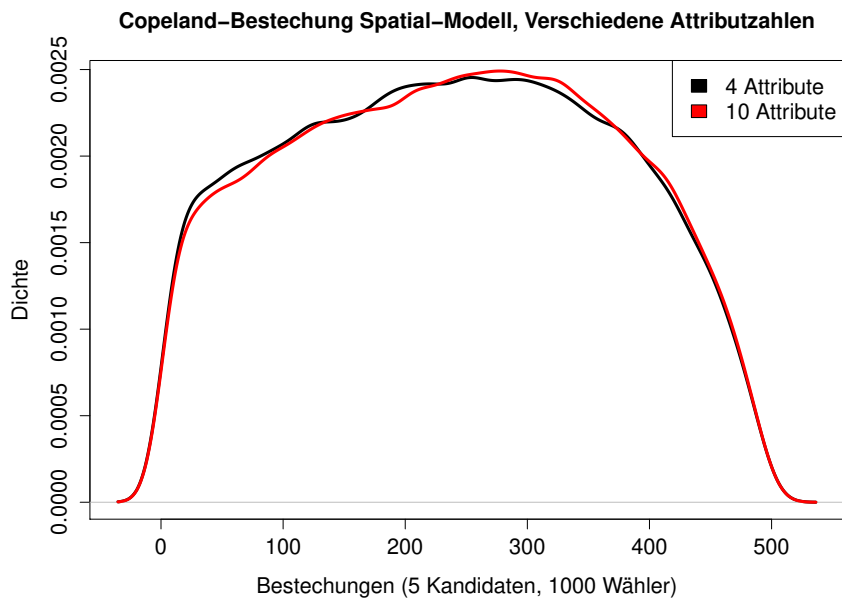


Abbildung 26: Copeland-Bestechung für verschiedene Anzahlen von unabhängigen Attributen im räumlichen Modell

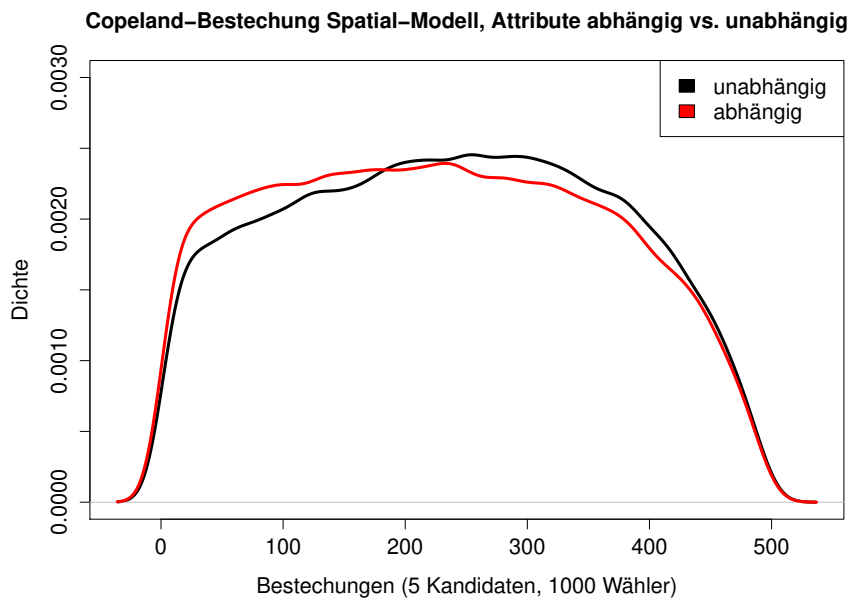


Abbildung 27: Copeland-Bestechung für abhängige und unabhängige generierte Attribute im räumlichen Modell

Zuletzt können auch hier analog zur Manipulation unterschiedliche Verläufe für die verschie-

denen Wahlsysteme beobachtet werden.

Für Bestechungen mit unterliegendem Pluralitäts-Wahlsystem sind die erhaltenen Resultate denen für das analoge Manipulationsproblem qualitativ sehr ähnlich. Ebenfalls können hier im Vergleich zur Copeland-Bestechung im IC- und im räumlichen Modell gegenläufige Verläufe der Kurven beobachtet werden (siehe Abbildung 28).

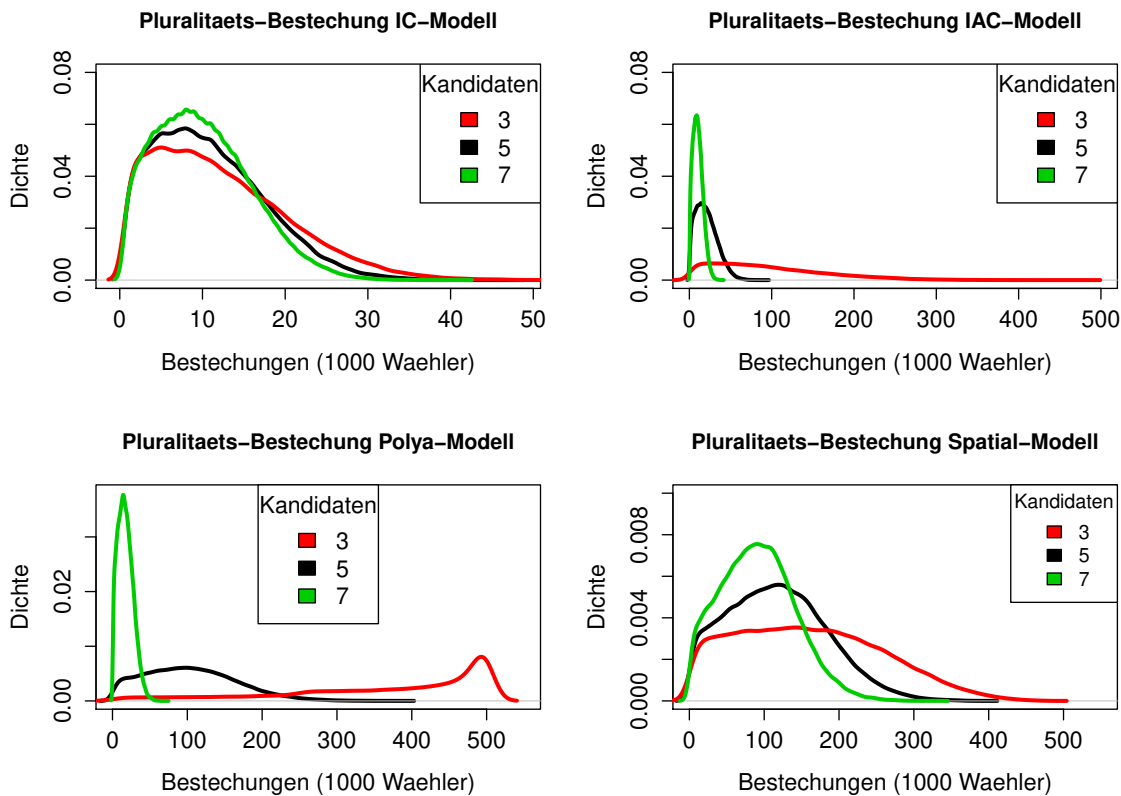


Abbildung 28: Pluralitäts-Bestechung für verschiedene Kandidatenzahlen

Bei der Bordwahl (siehe Abbildung 29) scheint die Bestechungsheuristik am besten für Wahlen mit 3 Kandidaten im IC- und Spatial-Modell zu funktionieren. Diese waren tendenziell am schwersten zu manipulieren, hingegen ist ihnen aber am einfachsten mit Bestechung beizukommen. Somit ähneln die Ergebnisse qualitativ denen der Copelandwahl. Die Ausnahme bildet hierbei die schon bei der Manipulation beobachtete zusätzliche Erhebung der Kurve für 3-Kandidaten-Wahlen im Pólya-Modell. Diese liegt im Bestechungsfall bei einem Wert von etwa 300.

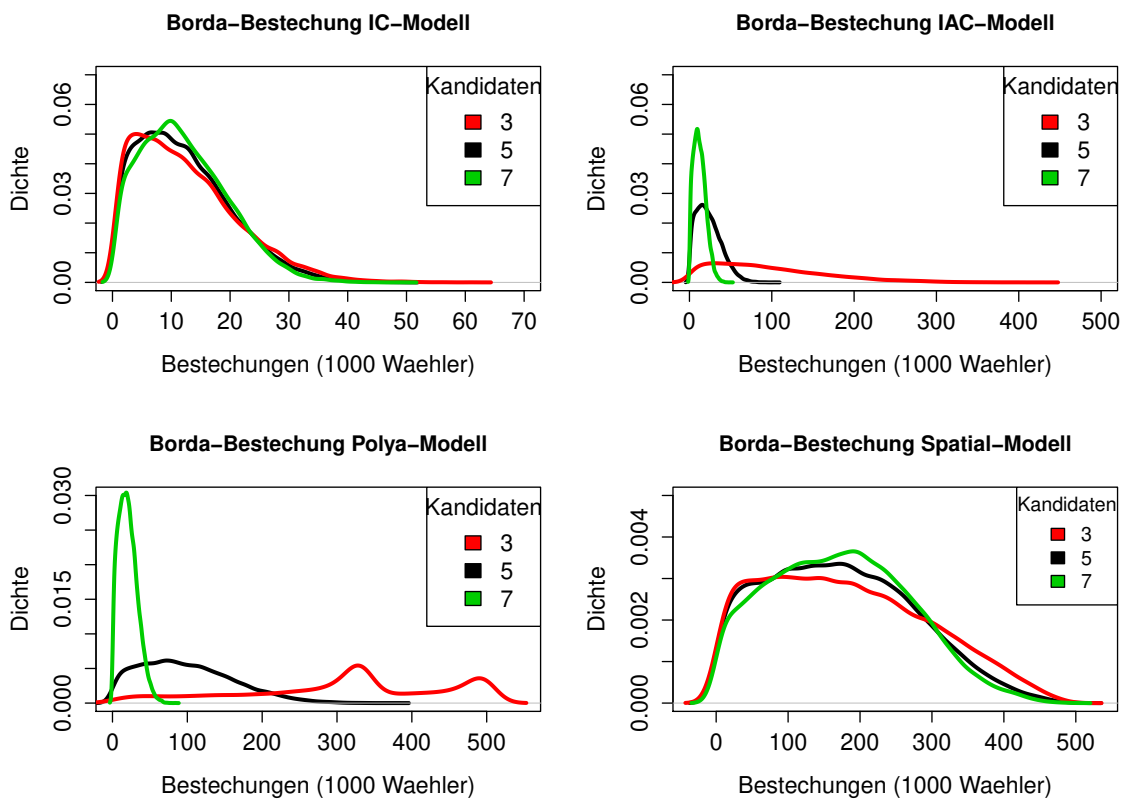


Abbildung 29: Borda-Bestechung für verschiedene Kandidatenzahlen

Im Falle der STV-Wahlen (Abbildung 30) verlaufen die Kurven qualitativ ebenfalls sehr ähnlich zur Copelandwahl (vgl. Abbildung 24).

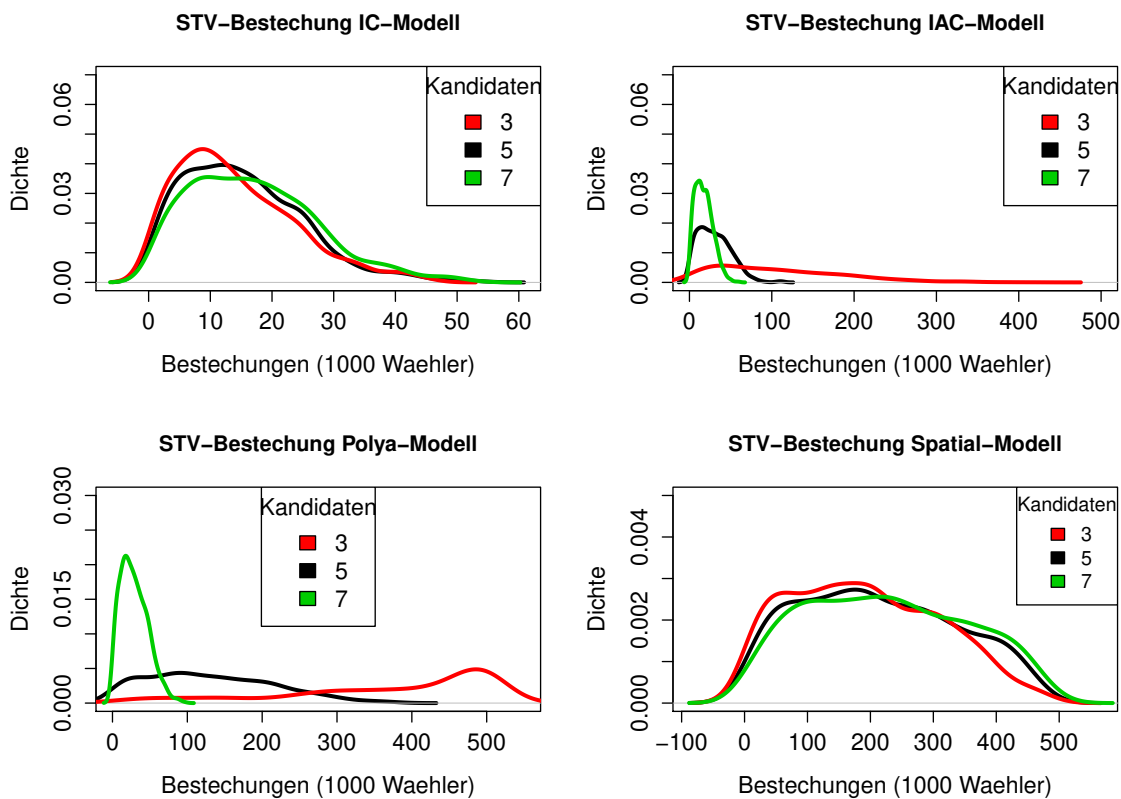


Abbildung 30: STV-Bestechung für verschiedene Kandidatenzahlen¹³

5.3.4 Ergebnisse III: Kontrolle

In diesem Abschnitt sollen nun zuletzt noch die Ergebnisse der Wahlkontrolle in Form des Hinzufügens eines „Klonkandidaten“ oder des Entfernens des ursprünglichen Wahlgewinners betrachtet werden. Zusätzlich wird überprüft, wie diese Kontrollmaßnahmen im Vergleich mit Bestechung und Manipulation mit vorgegebenem Kontingent von Manipulatoren/bestechlichen Wählern abschneiden. Dieses Kontingent beträgt, wie schon in Abschnitt 5.1 erwähnt, 5 % der Größe der Wählerschaft. Bei den hier betrachteten Profilen handelt es sich um diejenigen mit 1000 Wählern und 5 Kandidaten; daher entspricht das Kontingent 50 Personen.

In der folgenden Tabelle 8 sind für alle Modell-Wahlregel-Kombinationen die relativen Häufigkeiten angegeben, mit der ein Wahlbetrug mit der entsprechenden Strategie von Erfolg gekrönt ist (ausgezeichneter Kandidat kann zum Sieger gemacht werden). Dabei steht „add“

¹³Kerndichteschätzungen beruhend auf den jeweils ersten 1000 Profilen (wegen computablem Aufwand von STV)

für das Hinzufügen eines Kandidaten und „delete“ für das Entfernen des Siegers. Wie schon zuvor werden für Bestechung und Manipulation nur diejenigen Fälle betrachtet, bei denen der ausgezeichnete Kandidat nicht schon der ursprüngliche Sieger ist. Auch für das Hinzufügen und Löschen von Kandidaten werden nur solche Fälle herangezogen.

		IC	IAC	Pólya	Spatial
Plurality	Manipulation	0.969	0.658	0.010	0.080
	Bestechung	1	0.974	0.212	0.175
	add	0.251	0.249	0.203	0.195
	delete	0.252	0.252	0.249	0.248
Copeland	Manipulation	0.766	0.364	0.051	0.040
	Bestechung	0.996	0.778	0.120	0.085
	add	0.016	0.018	0.015	<0.001
	delete	0.250	0.249	0.249	0.250
Borda	Manipulation	0.976	0.711	0.147	0.077
	Bestechung	1	0.946	0.258	0.135
	add	0.003	0.006	0.011	0.002
	delete	0.252	0.250	0.251	0.249
STV	Manipulation	0.820	0.466	0.088	0.047
	Bestechung	0.996	0.866	0.345	0.095
	add	0.006	0.003	0.001	<0.001
	delete	0.251	0.249	0.249	0.250

Tabelle 8: Relative Häufigkeiten des erfolgreichen Wahlbetrugs mit verschiedenen Methoden für alle Modell-Wahlregel-Kombinationen (Manipulation und Bestechung jeweils mit Kontingent von 50)

Es lassen sich folgende Erkenntnisse aus der Tabelle gewinnen:

- Bestechung ist bei gleichem Kontingent immer (deutlich) effektiver als Manipulation.
- Der Erfolg von „add“ hängt nur vom Wahlsystem ab (Pluralitätswahl erfüllt nicht das Kriterium der Unabhängigkeit von Klonen (siehe Tideman 1987, S. 189) und ist daher anfällig). Dort ist diese Methode in etwa einem Viertel der Fälle erfolgreich, da der ausgezeichnete Kandidat einer von 4 Kandidaten ist, die nicht Wahlsieger sind. Eine etwas niedrigere Quote wird bei Pólya und dem räumlichen Modell erreicht, da dort teilweise eindeutigere Wahlergebnisse vorliegen; somit treten auch mehr Fälle auf,

in denen der ursprüngliche Sieger oder dessen Klon trotzdem noch gewinnen können, obwohl die Stimmen des Siegers auf ihn und seinen Klon aufgeteilt werden.

- „delete“ hat für alle Wahlregel-Modell-Kombinationen ungefähr die gleiche Erfolgsquote. Diese beträgt jeweils wiederum etwa ein Viertel, da 4 Kandidaten übrig bleiben.
- Im Pólya-Modell sowie im räumlichen Modell ist „delete“ sogar die effektivste Vorgehensweise, noch vor Bestechung.
- „add“ ist über alle Wahlsysteme hinweg die schlechteste Vorgehensweise.

6 Fazit und Ausblick

Zusammengefasst lässt sich festhalten, dass sich v.a. die eher geläufigen Borda- und Pluralitätswahlregeln in den meisten Fällen als leicht zu manipulieren erwiesen haben. Hingegen schneiden die seltener angewandten Regeln (die auch teilweise aufwendiger auszuwerten sind) Copeland und STV besser ab. Die intuitive Erwartung, dass Wahlen mit der STV-Regel, für die das Manipulationsproblem bekanntermaßen in der Komplexitätsklasse NP liegt, am schwierigsten zu manipulieren sind, kann teilweise bestätigt werden. In allen betrachteten Modellen und Szenarien erwies sich Copeland als schwerer manipulier- bzw. bestechbar. Der Erfolg der betrachteten Wahlkontrollmaßnahmen (Hinzufügen bzw. Entfernen eines Kandidaten) erwies sich als weitgehend unabhängig vom unterliegenden Modell. Allerdings konnte für das Hinzufügen eines Kandidaten erkannt werden, dass es außer für Pluralitätswahlen kaum Erfolg verspricht.

Des Weiteren sind die hier enthaltenen Ergebnisse noch nicht allumfassend. Als ein erster weitergehender Ansatz gibt es noch viele weitere Wahlregeln und Modellkonfigurationen, die betrachtet werden können. Beispielsweise könnten auch Ergebnisse für verschiedene Homogenitätsparameter im Pólya-Modell verglichen werden.

Eine zweite Anregung für weitergehende Untersuchungen betrifft die STV-Regel. Da mit der vorliegenden Arbeit nicht abschließend zu klären ist, wie gut der verwendete Borda-Algorithmus tatsächlich auch für die Manipulation von STV-Wahlen geeignet ist, könnten hier durch Vergleichsstudien eventuell noch bessere Algorithmen bzw. Heuristiken entdeckt werden.

Um herauszufinden, wie relevant die Ergebnisse tatsächlich für die Realität sind (und um zusätzlich erneut zu überprüfen, wie empirisch realistisch das räumliche Modell wirklich ist), könnten zusätzlich die Auswertungen an realen Wahldaten, die individuelle Rankings beinhalten, durchgeführt werden, indem die Struktur von real abgegebenen Rankings wie in Abschnitt 5.3.1 untersucht und mit der Struktur der Rankings aus den Modellen verglichen wird.

Als letzten Punkt könnte man auch das zum Thema Wahlkontrolle erwähnte Partitionieren von Wählern und Kandidaten noch genauer beleuchten und versuchen, Heuristiken dafür zu finden, wie man die Wähler oder Kandidaten am besten partitionieren sollte, um die Erfolgswahrscheinlichkeiten für den Wahlbetrug zu maximieren.

7 Literaturverzeichnis

- Arrow, K. (1963), „*Social Choice and Individual Values*“. John Wiley & Sons, New York, London, Sydney.
- Bartholdi, J., Orlin, J. (1991), „*Single Transferable vote resists strategic voting*“. Social Choice and Welfare 8: S.341-354.
- Bartholdi, J., Tovey, C., Trick, M. (1989), „*The Computational Difficulty of Manipulating an Election*“. Social Choice and Welfare 6: S. 227-241.
- Bartholdi, J., Tovey, C., Trick, M. (1992), „*How hard is it to control an election?*“. Mathematical Computational Modeling 16: S. 27-40.
- Berg, S. (1985), „*Paradox of Voting under an Urn Model: The Effect of Homogeneity*“. Public Choice 47/2: S. 377-387.
- Chamberlin, J., Cohen, M. (1978), „*Toward applicable social choice theory: a comparison of social choice functions under spatial model assumptions*“. The American Political Science Review 72/4: S. 1341-1356.
- Conitzer, V., Sandholm, T., Lang, J. (2007), „*When are elections with few candidates hard to manipulate?*“. Journal of the ACM 54/3: Article 14
- Davis, P. (1972), „*Gamma Function and Related Functions*“. Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables (Abramowitz, M., Stegun, I. (Hrsg.)): S. 253-266.
- de Borda, J. (1781), „*Mémoire sur les élections au scrutin*“. Mémoires de l'académie royale des sciences: S. 657-664.
- Duggan, J., Schwartz, T. (2000), „*Strategic manipulability without resoluteness or shared beliefs: Gibbard-Satterthwaite generalized*“. Social Choice and Welfare 17: S. 85-93.
- Faliszewski, P., Hemaspaandra, E., Hemaspaandra, L., Rothe, J. (2006), „*A richer understanding of the complexity of election systems*“.
- French, S. (1986), „*Decision Theory: an Introduction to the Mathematics of Rationality*“. S. 280-298.
- Gehrlein, W., Fishburn, P. (1974), „*Condorcet's Paradox and anonymous preference profiles*“. Public Choice 26: S. 1-18.
- Geanakoplos, J. (2005), „*Three brief proofs of Arrow's Impossibility Theorem*“. Economic Theory 26: S. 211-215.

- Gibbard, A. (1973), „*Manipulation of voting schemes: A general result*“. *Econometrica* 41/4: S. 587-601.
- Good, I., Tideman, T. (1976), „*From individual to collective ordering through multi-dimensional attribute space*“. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* 347/1650: S. 371-385.
- Homer, S., Selman, A. (2011), „*Computability and complexity theory*“. Springer Science & Business Media.
- Johnson, N., Kotz, S. (1977), „*Urn Models and Their Application*“. John Wiley & Sons, New York, Chichester, Brisbane, Toronto.
- Kelly, J. (1993), „*Almost all social choice rules are highly manipulable, but a few aren't*“. *Social Choice and Welfare* 10/2: S. 161-175.
- Kendall, M., Smith, B. (1939), „*The problem of m rankings*“. *The annals of mathematical statistics* 10/3, S. 275-287.
- List, C. (2013), „*Social Choice Theory*“. *The Stanford Encyclopedia of Philosophy* (Zalta, E. (Hrsg.)).
- R Core Team (2016), „*R: A language and environment for statistical computing*“. R Foundation for Statistical Computing, Wien. URL <http://www.r-project.org/>.
- Regenwetter, M., Grofman, A., Popova, A., Messner, W., Davis-Stober, C., Cavagnaro, D. (2008), „*Behavioural social choice: a status report*“. *Philosophical Transactions of the royal society B* 364: S. 833-843.
- Rothe, J., Baumeister, D., Lindner, C., Rothe, I. (2012), „*Einführung in Computational Social Choice*“. Spektrum-Verlag Heidelberg.
- Saari, D. (2001), „*Decisions and elections: explaining the unexpected*“. Cambridge University Press.
- Satterthwaite, M. (1975), „*Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions*“. *Journal of Economic Theory* 10: S. 187-217.
- Sen, A. (1966), „*A Possibility Theorem on Majority Decisions*“. *Econometrica* 34/2: S. 491-499.
- Silverman, B. (1986), „*Density Estimation*“. Chapman and Hall, London.
- Svensson, L. (1999), „*The Proof of the Gibbard-Satterthwaite Theorem Revisited*“.

- Taylor, A. (2002), „*The manipulability of voting systems*“. The American Mathematical Monthly 109/4: S. 321-337.
- Tideman, T. (1987), „*Independence of Clones as a Criterion for Voting Rules*“. Social Choice and Welfare 4: S. 185-206.
- Tideman, T., Plassmann, F. (2012), „*Modeling the outcomes of vote-casting in actual elections*“. Electoral Systems: Paradoxes, Assumptions and Procedures (Felsenthal, D., Machover, M. (Hrsg.)): S. 217-251.
- Zuckerman, M., Procaccia, A., Rosenschein, J. (2008), „*Algorithms for the coalitional manipulation problem*“. Proceedings of the nineteenth annual ACM-SIAM symposium on Discrete Algorithms: S. 277-286.
- Zwicker, W. (2016), „*Introduction to the theory of voting*“. Handbook of Computational Social Choice: S. 23-56.

8 Elektronischer Anhang

Der Anhang auf dem beiliegenden Speichermedium ist folgendermaßen strukturiert:

- Ordner Codes:
enthält alle R-Skripte sowie eine `readme.txt`-Datei, die die verschiedenen Code-Files genauer beschreibt
- Ordner Simulation:
enthält die simulierten Daten in den Unterordnern „*Profile*“, „*Manipulation*“, „*Bestechung*“ und „*Kontrolle*“.
- Ordner Grafiken:
enthält alle Grafiken die mit dem R-Skript `plots.r` aus den simulierten Daten aus dem Ordner *Simulation* erzeugt werden können.
- Zusätzlich ist die Masterarbeit selbst im `.pdf`-Format enthalten.

Eidesstattliche Erklärung

Der Verfasser erklärt, dass er die vorliegende Arbeit selbständig, ohne fremde Hilfe und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt hat. Die aus fremden Quellen (einschließlich elektronischer Quellen) direkt oder indirekt übernommenen Gedanken sind ausnahmslos als solche kenntlich gemacht. Die Arbeit ist in gleicher oder ähnlicher Form oder auszugsweise im Rahmen einer anderen Prüfung noch nicht vorgelegt worden.

München, den 24. Mai 2017

Andreas Weiß