



Fußball und die Bundestagswahl 2013: Der kausale Einfluss von ursprünglich nicht-politischen Emotionen auf Wahlen

Achim Goerres · Robin Arens · Florian Rabuza

Online publiziert: 26. September 2019
© Deutsche Vereinigung für Politikwissenschaft 2019

Zusammenfassung Welchen Einfluss haben Emotionen, die ihren Ursprung außerhalb der Politik haben, auf Bundestagswahlen? Auf Wahlkreisebene werden die kausalen Effekte der Spielergebnisse der Ersten Herren-Bundesliga vom Wahlwochenende 2013 auf die Zweitstimmenergebnisse geschätzt. Die Analysen des Difference-in-Differences-Designs im Vergleich zu 2009 zeigen: Die Zuteilung von Fußballergebnissen ist quasi-zufällig und stellt ein regional differenziertes natürliches Experiment dar. Die Regierungsparteien profitierten statistisch signifikant von einem guten Abschneiden der lokalen Mannschaft, gemessen über die Tordifferenz, während Oppositionsparteien etwas verloren (nur manchmal signifikant, aber immer mit der erwarteten Richtung der Punktschätzung). Diese Effekte wurden durch die zeitliche Distanz der Spiele bis zur Schließung der Wahllokale moderiert: Je größer der zeitliche Abstand war, desto geringer waren die Effekte, die aber ebenfalls nicht immer statistisch signifikant waren. Keine Evidenz gibt es für die Moderation der Effekte durch die Wettquote als bestes Wissen über den Spielausgang vor der Partie. Insgesamt unterstützen die Befunde die aggregierten Implikationen der Mikrologik der „blinden Retrospektion“ aufgrund von undeutlichen und instabilen Schätzungen nur zum Teil.

Schlüsselwörter Wahlen · Natürliches Experiment · Fußball · Blinde Retrospektion · Difference-in-Difference-Design

Zusatzmaterial online Zusätzliche Informationen sind im Online-Anhang dieses Artikels (<https://doi.org/10.1007/s11615-019-00188-4>) enthalten.

A. Goerres (✉) · R. Arens · F. Rabuza
Duisburg, Deutschland
E-Mail: achim.goerres@uni-due.de

Football and the 2013 German Bundestag Election: the Causal Influence of Emotions Originating Outside of Politics on Elections

Abstract What influence do emotions that originate outside politics have on German federal elections? In this paper, we estimate the causal effects of football (soccer) results from the first German Bundesliga on party-list district data during the election weekend in 2013. The results from a difference-in-differences design analysis relative to the 2009 elections reveal that football results are a quasi-random treatment and thus function as a natural experiment with regional variation. Incumbent parties benefited from a positive outcome for the local football team, as measured by the difference in goals in a statistically significant result. In contrast, opposition parties did not lose clearly from these positive outcomes. The results for the opposition parties were sometimes significant, but always had the expected points estimates. All effects were moderated by the timing of the matches relative to the closing of the election: the greater the time lag, the smaller the effects, although these effects were not always statistically significant, either. There was no evidence that betting odds, as the best information available before the match about the potential result, moderated the effects. Overall, the results only partially support the aggregated implications of the micro-logic of “blind retrospection” due to imprecise and instable estimates.

Keywords Elections · Natural experiment · Football · Blind retrospection · Difference-in-Difference-Design

1 Einleitung

Am Wochenende der Bundestagswahl 2013 spielte die erste Fußball-Herrenmannschaft des Hamburger SV (HSV) gegen den SV Werder Bremen und unterlag mit 0:2. Hatte diese unverhoffte Niederlage des HSV aufgrund der Enttäuschung in der Bevölkerung irgendwelche parteipolitischen Effekte in den Hamburger Wahlkreisen? Hatte dieser Sieg der Bremer Effekte in den Bremer Wahlkreisen, weil sich Bremer Bürgerinnen und Bürger so freuten?

Deutsche Wahlen werden zunehmend durch Prozesse des De-alignments (der Abnahme von stabilem Wählerverhalten in sozio-kulturellen Milieus), der Personalisierung und der daraus folgenden Wechselwahl charakterisiert (vgl. Schoen 2003). Insbesondere in solch einem Kontext lässt sich vermuten, dass Wählerinnen und Wähler zum Teil von Emotionen in ihrer Wahlentscheidung geleitet werden. Nun gibt es Emotionen, die im politischen Prozess selbst generiert werden (wie die affektive Bewertung von Politikerinnen und Politikern) und als Entscheidungsheuristik eingesetzt werden (Marcus 2000). Darüber hinaus gibt es aber auch Emotionen, die abseits des politischen Prozesses entstehen, aber dennoch auf politische Entscheidungen einwirken.

Solche ursprünglich unpolitischen Emotionen untersuchen wir in diesem Beitrag. Der kausale Mechanismus, der bei solchen Emotionen auf individueller Ebene wirkt, wird als „blinde Retrospektion“ bezeichnet („blind retrospection“) (Achen und Bar-

tels 2012, 2016). Wählerinnen und Wähler fühlen sich aufgrund nicht-politischer Ereignisse mehr oder weniger zufrieden bzw. unzufrieden. Neben Auswirkungen des individuellen Wohlbefindens auf die Wahlbeteiligung, die wir aus Platzgründen im Weiteren vernachlässigen, kann dieser Mechanismus dazu führen, dass Amtsinhabern Verantwortung für die subjektiv empfundene Situation zugewiesen wird. Je besser sich die Bürgerinnen und Bürger fühlen, desto eher wählen sie einen Amtsinhaber; je schlechter sie sich fühlen, desto eher wählen sie einen Herausforderer.

Dieser Artikel untersucht die kausalen Effekte von Emotionen, die ihren Ursprung außerhalb der Politik haben, auf das aggregierte Wahlverhalten. Konkret nutzen wir jene Emotionen, die durch die Spielergebnisse der Ersten Fußballbundesliga der Herren am 6. Spieltag der Saison 2013/14 in Deutschland, also am Wochenende der Bundestagswahl 2013, hervorgerufen wurden. Damit versuchen wir, die Effekte auf lokale Wahlergebnisse der damaligen Amtsinhaber der Bundesregierung (CDU/CSU und FDP) im Vergleich zu den im Bundestag vertretenen Oppositionsparteien (B'90/Grüne, SPD, Linke) zu messen.

Fußball ist in Deutschland sehr populär, stellt den Inhalt vieler Alltagsgespräche dar und bindet Fans emotional an bestimmte Vereine (vgl. z.B. Porat 2010). Das hohe Maß an Popularität und emotionaler Bindung ist die Eigenschaft, die wir uns in dieser politikwissenschaftlichen Analyse zunutze machen. Dadurch, dass Spiele zumindest bei einem Nicht-Unentschieden für jede Mannschaft anders ausgehen und Fans von unterschiedlichen Mannschaften über einen Großteil Deutschlands verteilt sind, entsteht eine regionale Variation sportlich herbeigeführter Emotionen. Die emotionale Bedeutung des Fußballs kann im Bereich der Herren und der Ersten Bundesliga aufgrund der längeren Geschichte und der größten medialen „Strahlkraft“ angenommen werden. Dabei ist zu betonen, dass Ergebnisse anderer Sportarten oder Ereignisse aus ganz anderen Bereichen genauso gut als Hebel dienen könnten, wenn sie die entsprechenden Voraussetzungen, eine hohe Verbreitung und das Hervorrufen emotionaler Reaktionen, erfüllten.

Dieser Artikel ist die erste Analyse solcher Fußballdaten im Kontext einer Wahl überhaupt, nicht nur in Deutschland (vgl. für die politische Kommunikationsforschung Hagen et al. 2015; Zeh und Hagen 2006a, 2006b). Gleichwohl gibt es erste Studien zu American Football und amerikanischen Kongresswahlen (Fowler und Montagnes 2015; Healy und Malhotra 2010; Healy et al. 2010). Darin wird ein weiteres wichtiges Merkmal bezüglich der Allokation von Sportereignissen deutlich. Man kann die Realisation von Spielausgängen (Sieg, Niederlage, Unentschieden) als ein natürliches Experiment analysieren, d.h., die Zuteilung des Spielausgangs folgt einer „als-ob“-Randomisierung unter der Annahme, dass die Ursachen für den Spielausgang keine politischen Wurzeln haben. Diese Annahme kann man in Teilen testen, sie ist aber vor allem *prima facie* plausibel. Hinzu kommt für unsere Analyse: Fußball ist als Spiel aufgrund der relativ geringen Anzahl von Toren (beispielsweise im Vergleich zum Basketball) viel anfälliger für Zufälle in der Spielsituation (Reep und Benjamin 1968; Reep et al. 1971). Ein kleiner Augenblick der mangelnden Aufmerksamkeit eines Spielers kann den Spielausgang entscheidend beeinflussen, sodass beispielsweise selbst Wettquoten als beste Gesamtschätzung vor dem Spiel geringe Vorhersagekraft haben (Dixon und Coles 1997; Falter und Pérignon 2000).

Die empirische Analyse besteht aus zwei Schritten. In einem ersten Schritt wird gezeigt, dass die Spielergebnisse des 6. Spieltags ein natürliches Experiment darstellen. In einem zweiten Schritt testen wir die direkten kausalen Effekte des Ausgangs der Spiele auf die Wahlkreis-Zweitstimmenergebnisse der Bundestagswahl 2013 relativ zu 2009 für die im Bundestag vertretenen Parteien. Diese Analyse wird innerhalb eines Difference-in-Differences-Designs für Wahlkreise mit einem regionalen Erstligaverein durchgeführt. Anschließend testen wir zwei zu erwartende Moderationen bezüglich der zeitlichen Distanz zur Schließung der Wahllokale und dem Außenseiterstatus der Vereine.

Im Allgemeinen unterstützen die aufgedeckten empirischen Muster die aggregierten Implikationen der blinden Retrospektion für eine der gewählten Operationalisierungen nur zum Teil. Die Regierungsparteien CDU (die CSU schließen wir aus später dargelegten Gründen aus) und FDP profitierten von einem guten Abschneiden der Heimmannschaft, während Oppositionsparteien sehr unterschiedlich stark verloren haben. Diese Effekte wurden durch die zeitliche Distanz der Spiele relativ zur Schließung der Wahllokale moderiert. Nicht immer erreichen die inferenzstatistischen Tests der geschätzten Effekte statistische Signifikanz, sodass es keine klare Unterstützung der Logik der blinden Retrospektion durch diese Analyse gibt. Konkret schlossen die im Bund regierenden CDU und FDP zwar in den Hamburger Wahlkreisen (relativ zu 2009) tatsächlich 2,2 Prozentpunkte schlechter ab als in den Bremer Wahlkreisen, und die Oppositionsparteien (SPD, B'90/Grüne, Linke) 1,8 Prozentpunkte besser. Doch ist die Evidenz über alle Wahlkreise nicht klar genug, um darin einen kausalen Effekt für die durch das Fußballergebnis herbeigeführten Emotionen bestätigt zu finden.

2 Theoretischer Rahmen: Emotionen, natürliche Experimente und Fußball

Es gibt drei relevante Literaturstränge: (1) Retrospektives Wählen unter besonderer Berücksichtigung der Rolle von Emotionen in diesen Erklärungen, (2) natürliche Experimente in der Wahlforschung und (3) die gesellschaftlichen Auswirkungen von Fußball auf Identität und Zufriedenheit.

2.1 Retrospektives Wählen und Emotionen

Die grundlegende Annahme der Theorie des retrospektiven Wählens ist, dass Wähler die Politiker auf Grundlage der bisher geleisteten Arbeit bewerten und daraus eine Wahlentscheidung ableiten. Der klassische Ansatz geht hierbei von der Begutachtung ökonomischer Kennzahlen durch die Wählerinnen und Wähler aus. So müssten komplett rationale Wählerinnen und Wähler ihre persönliche wirtschaftliche Lage vor einer aktuellen Wahl mit der Ausgangslage vor der vorherigen Wahl vergleichen und bei einer Verbesserung der Lebensbedingungen die regierenden Politikerinnen und Politiker im Amt bestätigen (Fiorina 1981). Dieser Entscheidungsprozess kann als eine Heuristik verstanden werden, die die individuellen Informationskosten im Vorfeld einer Wahl senken. Es reicht alleine der Vergleich zwischen zwei Zeit-

punkten, um eine Wahlentscheidung treffen zu können, sodass man keine weiteren Informationen über einzelne Policies und politische Positionierungen von Parteien sammeln muss (Bartels 2010). Der Wähler oder die Wählerin entscheidet gemäß der Logik: Was in der Vergangenheit gut gewesen ist, wird auch gut für die Zukunft sein.

Healy und Malhotra (2013) unterscheiden drei theoretische Varianten des retrospektiven Wählens. Die erste Variante nimmt die Sanktionierung der politischen Amtsinhaber in den Fokus. Demnach bestrafen Wähler eine „schlechte“ Politik, was einen Anreiz für „gute“ Politik für die Politiker darstellt. Bei „schlechter“ Politik droht die Abwahl. Der zweite Ansatz ist, dass Wähler ihre Stimme danach vergeben, wie kompetent der Herausforderer in Relation zu den bisherigen Ergebnissen des Amtsinhabers wirkt. Es wird also auf Basis der vorhandenen Daten der vergangenen Legislaturperiode eine Abwägung getroffen, ob der neue Kandidat der Wählerschaft mehr „nützt“ als der bisher gewählte. Diese beiden Varianten basieren auf dem Rational-Choice-Ansatz und sind dem klassischen retrospektiven Ansatz am nächsten. Das dritte Modell geht hingegen davon aus, dass emotionale und kognitive Verzerrungen die Wahlentscheidung beeinflussen. So können Wähler ihre Stimme einer gewissen Partei oder einem bestimmten Kandidaten auch deshalb geben, weil sie in der Wahlsituation besonders positiv oder negativ gestimmt sind. Dies kann zum einen daran liegen, dass Wählerinnen und Wähler ihre emotionale Situation unbewusst auf die politischen Ereignisse der letzten Legislaturperiode zurückführen und dadurch in ihrer Entscheidung beeinflusst werden. Ferner ist es theoretisch ebenfalls möglich, dass eine emotionale Situation dazu führt, dass sich der Wähler bei der Wahlentscheidung vermehrt an positive oder negative Entscheidungen der zu wählenden Politiker erinnert und auf dieser Grundlage die eigene Stimme abgibt. In dieser dritten Variante findet also keine sachliche retrospektive Bewertung statt. Es werden stattdessen affektive Bewertungen benutzt, um eine Entscheidung zu treffen. Diese „affektive Abkürzung“ lässt sich auch in Laborexperimenten nachweisen (Huber et al. 2012).

Achen und Bartels (2016) bezweifeln grundsätzlich die rein rationale Sicht auf das Konzept des retrospektiven Wählens. Sie argumentieren, dass verschiedene Ereignisse, die keinen politischen Charakter haben, durchaus auch auf Wahlentscheidungen Einfluss nehmen können. Sie führen das darauf zurück, dass zwischen der Wahrnehmung des politischen Erfolgs und des subjektiven Wohlergehens kein bzw. nur ein zufälliger Zusammenhang besteht. So zeigen sie unter anderem, dass Haiangriffe oder Wetterphänomene, die kein Politiker beeinflussen kann, zu einem substanziellen Stimmenverlust der Amtsinhaber führen können. Die Autoren weisen in einer Fallstudie mehrerer Haiattacken im Jahr 1916 einen Stimmenverlust für den damals amtierenden US-Präsidenten Wilson in den von den Attacken betroffenen Stimmbezirken von knapp drei Prozentpunkten nach. Dieser Effekt bleibt selbst dann stabil, wenn die Modelle mit veränderten Kontrollvariablen geschätzt werden. Nun kann hier noch argumentiert werden, dass es Aufgabe amerikanischer Bundespolitik ist, die Küsten für ihre Bürgerinnen und Bürger sicher zu halten und damit auch Haiattacken zu verhindern. In einer weiteren Studie jedoch, in der die Autoren die Wetterlage über einen langen Zeitraum in Zusammenhang mit den jeweiligen Wahlen analysieren, zeigt sich, dass sowohl eine sehr trockene als auch eine sehr nasse Wit-

terung die Wahlergebnisse der bisher regierenden Politikerinnen und Politiker um knapp einen Prozentpunkt vermindert. Dieser Effekt bleibt auch bei verschiedenen Operationalisierungen der unabhängigen Variablen annähernd gleich groß und signifikant. Hier fehlt nun jegliche Möglichkeit, einzelnen Politikerinnen und Politikern Verantwortung zuzuschreiben. Daher scheint sich als Erkenntnis zusammenfassen zu lassen, dass Ereignisse, die keinen direkten Bezug zu politischen Entscheidungen aufweisen, durchaus einen statistisch signifikanten Unterschied bei Wahlausgängen hervorrufen können.

Healy et al. (2010) untersuchen weitere unpolitische Ereignisse, um deren Effekte auf Wahlergebnisse zu bestimmen. Dazu analysieren sie den Einfluss des Spielausgangs von American Football-Spielen auf College-Ebene im direkten Vorfeld von Präsidentschafts-, Senats- und Gouverneurswahlen in den USA zwischen 1964 und 2008. Ihr statistisches Modell liefert Evidenz dafür, dass der Sieg der lokalen College-Mannschaft zu einem erhöhten Stimmanteil für den Amtsinhaber oder die Amtsinhaberin führt. Allerdings finden Fowler und Montagnes (2015) mit vergleichbaren Daten keine signifikanten Zusammenhänge zwischen dem Ausgang von Football-Spielen und dem Wahlergebnis von Amtsinhabern. Sie argumentieren, dass es sich bei den Ergebnissen von Healy et al. (2010) um „false-positive“-Resultate handelt, die durch Zufall zustande gekommen sind. Ihr Resümee: Wählerinnen und Wähler sind kompetenter als auf der Basis der vorherigen Ergebnisse angenommen und nicht-politische Ereignisse haben eben keinen besonderen Einfluss auf Wahlausgänge. Neuere experimentelle Daten deuten wiederum darauf hin, dass die Bewertung von Politikerinnen und Politikern durch den Ausgang von Sportereignissen durchaus beeinflusst werden kann (Busby und Druckman 2018; Busby et al. 2017).

Es ist anzunehmen, dass die Emotionen, ausgelöst durch das Endresultat eines Spiels, sich lediglich kurzfristig auf die Wahlpräferenz auswirken. Fowler und Montagnes (2015) betrachten Spiele in einem Zeitraum von zwei Wochen vor der Wahl, wohingegen Healy et al. (2010) nur Spiele am Tag vor der Wahl in ihre Analyse mitaufnehmen. Die beiden experimentellen Untersuchungen (Busby und Druckman 2018; Busby et al. 2017) analysieren die Auswirkungen der Spiele ebenfalls im unmittelbaren zeitlichen Kontext zum Spielende (bis zu drei Tage nach der Partie).

All diese Befunde lassen zumindest auf den ersten Blick an einer rein sachlichen Bewertung politischer Zusammenhänge durch Wählerinnen und Wähler zweifeln. Das Bild des Wählers, das hier im Vordergrund steht, ist aus einer normativen Perspektive unvoreilhaft und weist auf eine unsachliche Verknüpfung von Emotionen und Wahlentscheidungen hin. Weiterhin werfen die verschiedenen Befunde über retrospektives Wählen und die Verknüpfung von nicht-politischen Erfahrungen die Frage nach der Bedeutung dieser Effekte für den deutschen Kontext mit seiner komplexeren Konstellation von Parteien auf (im Vergleich zu den Zwei-Kandidaten-Wettbewerben der USA). Die Studie über American Football und Wahlen gibt uns ferner Anlass zum Einsatz von Sportergebnissen als emotionsauslösende Ereignisse, zu der wir im Abschn. 2.3 zurückkehren.

¹ Die Null-Hypothesen werden verworfen, obwohl in der Realität kein solcher Zusammenhang besteht (Fehler 1. Art).

In dem vorliegenden Beitrag bezieht sich die empirische Analyse ebenfalls auf kurzzeitige Veränderungen der politischen Bewertung aufgrund von Emotionen. Dass der Ausgang eines Bundesliga-Spiels eine Wahlentscheidung zwei Wochen später signifikant beeinflusst, scheint unrealistisch. Daher werden nur die Ausgänge der Bundesligaspiele am Wochenende der Bundestagswahl, die schon vor dem Schließen der Wahllokale bekannt waren, in der empirischen Analyse berücksichtigt. Wir nutzen zudem die Möglichkeit, die Effekte der zeitlichen Distanz bis zur Schließung der Wahllokale über das Wochenende hinweg explizit zu testen.

2.2 Natürliche Experimente in der Wahlforschung

In den zitierten Studien von Healy et al. (2010) findet sich bereits der Umgang mit nicht-politischen Ereignissen. Die Zuteilung solcher Ereignisse folgt in der Regel der Logik eines natürlichen Experiments. Das heißt erstens, dass die Ursache des Treatments unabhängig von der abhängigen Variablen ist, also beispielsweise die Ursachen für Haiattacken kausal nicht mit politischen Wahlentscheidungen verbunden sind. Zweitens bedeutet es, dass die Variation der Zuteilung in Bezug auf die abhängige Variable quasi-zufällig (Dunning 2008) ist. Diese „as if“-Zuteilung kann empirisch unter Bezugnahme auf sehr ähnliche abhängige Variablen zu anderen Zeitpunkten oder Prädiktorvariablen der zu untersuchenden abhängigen Variablen überprüft werden.

In der Politikwissenschaft lassen sich eine ganze Reihe von natürlichen Experimenten in Bezug auf Wahlausgänge finden. So behandeln Bechtel und Hainmueller (2011) die Elbeflut im Jahr 2002 als ein natürliches Experiment, bei dem die Hochwasserschäden quasi-zufällig deutschen Wahlkreisen zugeteilt werden. Aufgrund von unterschiedlicher Aktivität der Regierung bei der Beseitigung der Schäden wurden verschiedene Wahleffekte in der Bevölkerung für die an der Regierung beteiligten Parteien hervorgerufen. In einem anderen Beispiel nutzt Stokes (2016) die nicht-politisch motivierte Zuteilung von Windparks in der kanadischen Provinz Ontario, um die tatsächliche Betroffenheit von Wählerinnen und Wählern durch die Nachteile der Windkraftwerke als natürliches Experiment zu modellieren. Kousser und Mullin (2007) operationalisieren die Zuteilung von verpflichtender Briefwahl in sehr kleinen Wahlbezirken in Kalifornien als natürliches Experiment, um den Effekt auf die Wahlbeteiligung zu messen. Lassen (2005) zeigt, dass das vorherige Informieren über ein Referendum in Kopenhagen in einigen Stadtteilen zu größerer Wahlbeteiligung führt und die Zuteilung der Bezirke quasi-zufällig erfolgte.

Obwohl natürliche Experimente in der Politikwissenschaft und insbesondere in der Wahlforschung immer häufiger eingesetzt werden, sind sie kein methodisches Allheilmittel. Sekhon und Titiunik (2012) beschreiben methodische Probleme im Umgang mit natürlichen Experimenten. Da die Zuteilung in den seltensten Fällen gänzlich bekannt ist, muss die Annahme der „as if“-Zuteilung durch die Inklusion weiterer Kovariate in der Analyse gewährleistet sein. Weiterhin muss sichergestellt werden, dass die Einheiten der Kontroll- und Treatmentgruppen wirklich vergleichbar sind. Anders als in einem Experiment, in dem diese Kontrolle ein erster Schritt des Designs ist, kann dies in natürlichen Experimenten erst im Nachhinein erfolgen. Die Autoren plädieren daher vor allem für ein sorgfältiges Verständnis des Zu-

teilungsprozesses des Treatments durch die Forscherinnen und Forscher, um diese später durch geeignete Analysen zu berücksichtigen.

Zusammengefasst zeigt dieser Abschnitt, dass natürliche Experimente in der Wahlforschung zunehmend verwendet werden und ihre Analyse – anders als bei klassischen Experimenten – eine große Umsicht bei der Interpretation des kausalen Effektes verlangt.

2.3 Die gesellschaftlichen Auswirkungen von Fußball auf Identität und Zufriedenheit

Die gesellschaftliche Bedeutung des Fußballs in Deutschland ist außerordentlich. So hatten die 18 Bundesliga-Clubs am Ende des Jahres 2015 zusammen etwas mehr als 900.000 Mitglieder², für den DFB wies die Statistik knapp 6,9 Mio. Mitglieder aus.³ Ferner haben in der Spielzeit 2014/2015 ungefähr 13,3 Mio. Fans die Spiele live im Stadion verfolgt.⁴ Auch bei den Konsumausgaben im Sportbereich rangiert der Fußball in Deutschland in herausgehobener Stellung. So geben allein 20 % der mindestens sechszehnjährigen Deutschen, die sich für Sport interessieren, Geld für Fußball aus. Handball als zweitbeliebteste Sportart in diesem Ranking erreicht hingegen gerade mal einen Anteil von 4,5 % in dieser Zielgruppe. Zeitgleich ist Fußball deutschlandweit die am häufigsten betriebene Mannschaftssportart (Preuß und Alfs 2013). Die Bedeutung des Fußballs in der deutschen Gesellschaft zeigt sich ebenfalls daran, dass signifikante Teile der Bevölkerung mit Fußballteams als Fans und/oder als Zuschauerinnen und Zuschauern verbunden sind. Die Bedeutung wird weiter dadurch deutlich, dass zwischen 2008 und 2015 das Programm mit der höchsten Einschaltquote pro Jahr jeweils eine Fußballübertragung war. Betrachtet man die Top 10 der meistgesehenen Sendungen über die acht Jahre, so kommt man auf insgesamt 62 Fußballübertragungen bei 80 Sendungen insgesamt.⁵

Die Gruppe von Zuschauerinnen und Zuschauern ist zu groß, um bei ihr eine einheitliche Leidenschaft für Fußball zu vermuten. Für die emotionale Berührung durch Fußball ist die Gruppe der Fans bedeutsamer. Hunt et al. (1999) definieren Fans als eine Gruppe von Interessierten, die auf einem gewissen Mindestniveau mit einem Verein oder einer Sportart verbunden sind. Ferner wird argumentiert, dass die Motivation, Fan eines Vereins zu sein, nicht nur mit den sportlichen Erfolgen zusammenhängt, sondern auch durch geografische und psychologische Merkmale hervorgerufen wird. Das „Fansein“ kann eine prägende Rolle für die Identität einer Person spielen und eine sehr stabile Komponente sein (Porat 2010).

Diese starke Verbundenheit zwischen Fan und Verein schlägt sich in Emotionen nieder (Porat 2010). So zeigt sich bei einem empirischen Vergleich von Fans sieg-

² <http://de.statista.com/statistik/daten/studie/29723/umfrage/anzahl-der-mitglieder-ausgewaehler-vereine-der-bundesliga/>. Zugegriffen: 14. Mai 2019.

³ <http://de.statista.com/statistik/daten/studie/184918/umfrage/sportarten-in-deutschland-nach-anzahl-der-mitglieder/>. Zugegriffen: 14. Mai 2019.

⁴ <http://de.statista.com/statistik/daten/studie/1626/umfrage/entwicklung-der-zuschauerzahl-in-der-bundesliga-seit-1963/>. Zugegriffen: 14. Mai 2019.

⁵ https://www.agf.de/daten/tvdaten/hitliste/?name=hitliste_20jahre. Zugegriffen: 05. Mai 2019.

reicher Mannschaften, dass diese deutlich angenehmere Emotionen erleben als die Fans des unterlegenen Teams (Kerr et al. 2005). Depetris-Chauvin et al. (2018) weisen nach, dass Siege der nationalen Mannschaft zu einer Stärkung nationaler und einer Schwächung sub-nationaler ethnischer Identität in Subsahara-Afrika führen. Fußballerfahrungen haben also identifikative Konsequenzen jenseits der Fußballwelt. Schwarz et al. (1987) zeigen auf, dass ein Sieg der favorisierten Mannschaft zu einer erhöhten Lebenszufriedenheit führt, da man in einer positiven Stimmung die subjektive Situation wohlwollender betrachtet. Diese Ergebnisse stützen unsere Proposition, dass die Wahlentscheidung von Fußballfans durch das Abschneiden ihrer favorisierten Mannschaft am gleichen Wochenende beeinflusst werden könnte. Diese kurzfristige Veränderung in der Lebenszufriedenheit wird durch das Resultat der eigenen Mannschaft maßgeblich beeinflusst. So ist davon auszugehen, dass Fans mit einer höheren Lebenszufriedenheit eher die regierende Koalition „belohnen“, wohingegen die Fans mit einer geringeren Zufriedenheit eher für die Opposition stimmen. Politikerinnen und Politikern kann sachlich gesehen jedoch keine Verantwortung für den Ausgang von Fußballspielen zugeschrieben werden.

3 Fußball, Lebenszufriedenheit und die Belohnung oder Bestrafung der Parteien an der Wahlurne

Stammwählerschaft verliert in Deutschland immer mehr an Bedeutung (Blumenstiel und Wiegand 2014). Im Kontext des De-alignments erfüllen soziostrukturelle Merkmale somit kaum noch eine Push-Funktion bezüglich des Wahlverhaltens. Immer mehr Wählerinnen und Wähler entscheiden sich kurzfristig vor einer Wahl. In diesen kurzfristigen Prozessen spielen Emotionen und die Lebenszufriedenheit eine wichtige Rolle. Je zufriedener die Menschen sind, desto eher belohnen sie den Amtsinhaber. Je unzufriedener die Menschen sind, desto eher bestrafen sie den Amtsinhaber. Es findet eine bewusste oder unbewusste Zuweisung von Verantwortung für die persönliche Zufriedenheit auf den Amtsinhaber statt. Das ist der Kern dessen, was Achen und Bartels (2016) „Blinde Retrospektion“ nennen.

Wenn die Verbindung zwischen allgemeiner Zufriedenheit und politischer Verantwortungszuweisung eine gute Beschreibung der Wirklichkeit ist, müssten sich allgemeine Schocks der durchschnittlichen Zufriedenheit der Bevölkerung in einem veränderten Wahlverhalten niederschlagen. Nun interessieren wir uns vor allem für die Verantwortungszuweisung von Schocks auf die allgemeine Zufriedenheit, die ihren Ursprung eindeutig in der nicht-politischen Welt hat.

In der vorliegenden Studie nutzen wir die Ergebnisse der ersten Fußballbundesliga der Herren am Wahlwochenende der Bundestagswahl 2013 (20.–22.09.2013). Die Gründe hierfür sind erstens, dass Fußball der Sport mit der größten gesellschaftlichen Bedeutung in Deutschland ist. Zweitens zeigen sportsoziologische Studien, dass der Spielausgang der favorisierten Mannschaft die Zufriedenheit von Fans maßgeblich beeinflussen kann. Dabei hängt vermutlich das Ausmaß, mit dem jemand persönlich betroffen ist, mit der Intensität zusammen, in der sich jemand mit einem Fußballverein identifiziert. Drittens haben Fußballergebnisse den politikwissenschaftlich interessanten Vorteil, dass sie praktisch omnipräsent sind und damit vielfältig als kausale

Komponente in der Analyse von politischem Verhalten eingesetzt werden können. Fußballspiele, die große allgemeine Aufmerksamkeit bekommen, finden anders als natürliche Katastrophen sehr häufig statt. Viertens kann man für die Zuteilung des Spielergebnisses testen, ob es sich um ein natürliches Experiment handelt: Man kann die Erwartung testen, dass das Spielergebnis zwar nicht kontrolliert zufällig, aber quasi-zufällig zugeteilt ist (vgl. auch Bauer 2015). Damit hätte man einen kausalen Mechanismus identifiziert, der das Wahlergebnis beeinflusst, dessen Ursprung jedoch komplett außerhalb der politischen Sphäre liegt. Das Ergebnis eines Fußballspiels der Ersten Bundesliga ist abhängig vom Talent und der Tagesform der eingesetzten Spieler, vom Glück und von weiteren Variablen. Keiner dieser Faktoren hat etwas mit Politik zu tun. Weitere Unterschiede, wie die Fähigkeiten der beiden Teams relativ zueinander, werden über die Wettquoten des Tages gemessen.

Die kausalen Mechanismen blinder Retrospektion spielen sich dabei in erster Linie auf der Mikroebene ab, die in unseren Daten hier nicht abgebildet sind. Stattdessen beschränken wir uns auf die Tests der Implikationen auf der Makroebene. In einer optimalen Datenwelt sollten wir Panel-Daten von Individuen haben, die zeitlich verteilt vor und nach der eigentlichen Wahl erhoben worden sind, Informationen über die Zufriedenheit einer Person im Zeitverlauf bereitstellen und uns Informationen über die „Fußballidentifikation“ eines Wählers bzw. einer Wählerin geben. So könnten wir untersuchen, ob das Spielergebnis einen Effekt auf die Zufriedenheit und damit auf die Wahlentscheidung von Individuen hat, die sich in einer gewissen Intensität mit einem bestimmten Verein identifizieren. Da wir aber über solche Daten nicht verfügen und uns auch nicht bekannt ist, dass sie je erhoben worden sind, wenden wir uns den Aggregatdaten auf Wahlkreisebene zu. Hier konzentrieren wir uns auf die Zweitstimmenanteile, weil wir hier eine plausiblere Zuschreibung zur Regierungsverantwortung als bei den Erststimmen für die Wahlkreiskandidatinnen und -kandidaten erwarteten. Es ist also insgesamt ein empirisch sehr konservativer Test mit einem kleinen Datensatz, der nur sehr deutliche Effekte aggregiert nachzuweisen vermag.

Im ersten Schritt unserer Argumentation erwarten wir, dass die Zuteilung der Spielergebnisse innerhalb der Wahlkreise wie ein natürliches Experiment funktioniert, d. h. nicht mit soziostrukturellen Merkmalen ko-variiert. Diese Erwartung kann natürlich nur dann überprüft werden, wenn ein Erstligaverein vor Ort vorhanden ist, der gewinnen oder verlieren kann. Somit müsste gelten: Die Spielergebnisse des Wahlwochenendes haben einen Nullzusammenhang mit strukturellen Prädiktoren des Wahlergebnisses 2013 wie Arbeitslosigkeit, Bevölkerungsdichte, Anteil bestimmter Bildungsgruppen, demografische Zusammensetzung oder dem Wahlergebnis von 2009.⁶

Gemäß einer retrospektiven Wahllogik führt eine Niederlage des lokalen Bundesligavereins zu einer geringeren Zufriedenheit unter Fußballinteressierten. Die Wahlberechtigten darunter übertragen zum Teil diese emotionale Reaktion auf ihre

⁶ 2013 existierten Bundesligavereine nur in den alten Bundesländern und tendenziell eher in städtischen Regionen. Deswegen ist die externe Validität des natürlichen Experiments auf die Teilgebiete der Bundesrepublik beschränkt, in denen ein Erstligaverein verortet ist.

Wahlentscheidung und strafen die Regierungsparteien ab, indem sie entweder eine der anderen Parteien wählen oder gar nicht zur Wahl gehen.

H1a In Wahlkreisen, in denen ein Bundesligaverein beheimatet ist, schneiden die Amtsinhaberparteien (CDU/CSU und FDP) im Schnitt umso besser ab, je besser das Ergebnis der heimischen Mannschaft ist.

H1b In Wahlkreisen, in denen ein Bundesligaverein beheimatet ist, schneiden die Oppositionsparteien (SPD, B'90/Grüne, Linke) im Schnitt umso schlechter ab, je besser das Ergebnis der heimischen Mannschaft ist.

Die Bestrafungs- und Belohnungslogik sollte durch die zeitliche Distanz des Spiels relativ zur Schließung der Wahllokale in einer Interaktion moderiert werden. Die hervorgerufenen Emotionen haben keine hohe Halbwertszeit, sodass die Effekte bei weiter wegliegenden Spielen gegen Null abgemildert werden sollten.

H2a Der positive Effekt eines guten Abschneidens der Heimmannschaft für die Regierungsparteien ist umso stärker, je näher das Spiel vor Schließung der Wahllokale endet.

H2b Der negative Effekt eines guten Abschneidens der Heimmannschaft für die Oppositionsparteien ist umso stärker, je näher das Spiel vor Schließung der Wahllokale endet.

Die durchschnittliche emotionale Reaktion der wahlberechtigten Fußballinteressierten ist abhängig von den allgemeinen Erwartungen bezüglich des Spiels. Wenn der Tabellenletzte gegen den Tabellenersten gewinnt, dann wird dies vermutlich zu einer stärkeren Reaktion führen, als wenn der Erste gegen den Letzten gewinnen würde. Diese allgemeinen Erwartungen schlagen sich in den Wettquoten nieder, sodass wir diesen Zusammenhang in Form einer Interaktion messen können.

H3a Je höher die Wettquote auf das erzielte Ergebnis war (d. h. je unwahrscheinlicher der „Markt“ dieses Ergebnis vorher einschätzte), desto stärker gewinnen die Regierungsparteien durch ein gutes Abschneiden der Heimmannschaft.

H3b Je höher die Wettquote auf das erzielte Ergebnis war (d. h. je unwahrscheinlicher der „Markt“ dieses Ergebnis vorher einschätzte), desto stärker verlieren die Oppositionsparteien durch ein gutes Abschneiden der Heimmannschaft.

4 Daten, Methoden und Design

4.1 Der Kontext der Bundestagswahl 2013

Bundestagswahlen generell und die Bundestagswahl 2013 speziell sind höchst relevante Kontexte für unser Design. Zum einen gab es eine relativ hohe Anzahl an Wechselwählern aufgrund des fortgeschrittenen De-alignments (Dalton 2014). Nie war es vorher aufgrund der zunehmenden Individualisierung einfacher, kurzfristig

Wähler für sich zu gewinnen als bei der Bundestagswahl 2013. Zum anderen war das Parteiensystem deutlich komplexer als das Zweiparteiensystem der USA, für welches die meisten Studien durchgeführt worden sind. Wir erwarten, dass die Effekte der Verantwortungszuschreibung aufgrund multipler Möglichkeiten beispielsweise in einer Koalition deswegen deutlich schwieriger nachzuweisen sind. Opposition und Regierung wurden 2013 jeweils von mindestens zwei Parteien verkörpert. Um eine Regierung zu belohnen, hatte man seit 1990 immer die Wahl zwischen zwei Regierungsparteien. Somit ist möglich, dass sich einfache Effekte von Belohnung und Bestrafung des Amtsinhabers in Deutschland nur schwer nachweisen lassen, weil sie sich auf zwei potenzielle Empfänger aufteilen.

4.2 Datenquellen und Operationalisierungen

Als Wahldaten werden die amtlichen Wahlergebnisse des Bundeswahlleiters in allen 299 Wahlkreisen (Brief- und Urnenergebnisse zusammen) verwendet. Ein Wahlkreis hatte bei der Wahl 2013 im Schnitt etwa 207.000 Wahlberechtigte mit einem Minimum von etwa 155.000 und einem Maximum von etwa 257.000. Für die Zweitstimmenergebnisse der Parteien wurde jeweils der relative Prozentsatz mit der Anzahl der tatsächlichen Wähler als Basis berechnet. Gleichzeitig nutzen wir die Ergebnisse der Änderungen der Wahlkreise, um Differenzen zwischen 2009 und 2013 zu berechnen. Wir unterscheiden nicht nach Brief- und Urnenwahlergebnissen, weil Briefwahlunterlagen beispielsweise von Kranken auch noch am Wahltag abgegeben werden können (und somit Emotionen, herbeigeführt durch die Fußballergebnisse, diese auch beeinflussen können) und weil der Anteil der Briefwähler im Zeitverlauf ansteigt, so dass nur eine gemeinsame Berücksichtigung beider Ergebnisse die steigende Zahl an Brief- und sinkende Zahl an Urnenwählern zwischen 2009 und 2013 neutralisiert.

Für die uns interessierenden Wahlkreisdaten liegen weiterhin ausgewählte sozio-ökonomische Strukturdaten vor, die vom Statistischen Bundesamt für die Wahlkreisebene zur Verfügung gestellt worden sind: Die Arbeitslosenquote, die Verteilung von Bildungsabschlüssen als Indikator sozialer Deprivation, die Bevölkerungsdichte als Indikator für den Stadt-Land-Unterschied und den Anteil an Personen im Alter 60+ für die demographische Struktur. Wir nehmen dabei die vier Indikatoren, bei denen wir eine große Korrelation mit den Wahlergebnissen im Querschnitt kennen.

Diesen Daten haben wir anschließend Details der Spiele der Fußballbundesliga zugeordnet. Die Daten für die Bundesligaergebnisse sind mithilfe des Online-Archivs der Deutschen Bundesliga ermittelt worden. Zur Absicherung wurden diese Ergebnisse noch mit den Daten des RSSF-Archivs (Rec. Sport. Soccer Statistics Foundation) verglichen. Sowohl die Primär- als auch die Kontrollquelle sind sehr verlässliche Bezugsmöglichkeiten für diesen speziellen Datentyp. Ferner weisen beide Quellen für alle Spiele die gleichen Ergebnisse aus. Diese Daten wurden händisch übertragen und codiert.

Wir führen die Analysen nur für Wahlkreise durch, in denen ein Erstligaclub beheimatet ist oder die an einen Wahlkreis angrenzen, für den das zutrifft. So ist zum Beispiel Borussia Dortmund in den beiden Dortmunder Wahlkreisen beheimatet und grenzt an verschiedene weitere Wahlkreise an, wie zum Beispiel an den

Wahlkreis Hagen-Ennepe-Ruhr-Kreis I. Bei Braunschweig und Wolfsburg gibt es einige Wahlkreise mit Doppelbelegungen, deren Ausschluss aber die Ergebnisse nicht verändert.⁷ Von den 112 Wahlkreisen dieser Definition schließen wir 24 Wahlkreise aus, weil das Fußballspiel am Wahlwochenende erst nach Schließung der Wahllokale endet und somit keinen Effekt haben kann. Somit bleiben 88 Wahlkreise, davon 13 in Bayern. Diese 13 bayrischen Wahlkreise müssen wir für die CDU-Analyse ausschließen, da die Inklusion der bayrischen Wahlkreise in einer übergreifenden CDU/CSU-Analyse deutliche Heterogenität der Effekte für die zwei Untergruppen (bayrisch versus nicht-bayrisch) aufweist.

Insgesamt ist die empirische Analyse ein konservativer Test: Er geht davon aus, dass für alle Parteien die regionale Reaktion auf das Fußballergebnis messbar und im Schnitt anders als für Wahlkreise mit einem anderen Ergebnis ist. Es gab allerdings Fußballclubs an diesem Spielwochenende, die eine stärker geografisch diffundierte Fanmitgliedschaft hatten als andere. Dies macht es also umso schwerer, lokal realisierte, emotional bedingte Emotionen im Wahlkreisergebnis nachzuweisen.⁸ Das natürlich experimentelle Treatment (relatives Abschneiden im Vergleich zu 2009) ist also in seinem Effekt auf das Zweitstimmenergebnis eingeschränkt, weil die emotionalen Reaktionen geografisch unterschiedlich verteilt sind.

Alle unabhängigen Variablen sind im Online-Anhang, Tab. 1, zu finden. Die Operationalisierung der Treatmentvariable „Fußballergebnis“ ist im Kontext eines First-Difference-Modells mit einigen Schwierigkeiten behaftet, da man die Veränderungen zwischen den Messzeitpunkten in das Modell aufnehmen muss. Es wurden dabei zwei Operationalisierungen für das Fußballergebnis gewählt. Als erste Operationalisierung des Spielergebnisses wurde eine metrische unabhängige Variable in Form der Differenz der Tordifferenzen zwischen den Spielergebnissen 2013 und 2009 der Mannschaft gebildet. Diese nimmt z. B. dann positive Werte an, wenn die Tordifferenz im Jahre 2013 im Vergleich zu 2009 besser geworden ist. Zweitens wurde eine kategoriale Variable gebildet, die folgende Ausprägungen annimmt: Das Resultat ist 2013 im Vergleich zu 2009 (a) besser, (b) gleich oder (c) schlechter geworden im Hinblick auf die drei möglichen Outcomes Sieg, Unentschieden oder Niederlage. Hat eine Mannschaft 2009 beispielsweise verloren und 2013 unentschieden gespielt oder sogar gewonnen, so ist das Ergebnis besser geworden. Siegt eine Mannschaft 2009 und 2013, spielt beide Male unentschieden oder verliert beide Partien, so bleibt das Ergebnis der Tendenz nach gleich. In Summe kann diese Variable die Werte von -2 bis $+2$ annehmen.

Für die Interaktionsanalyse wurde in Stunden gemessen, wie lange vor der Schließung der Wahllokale das Bundesligaspiel angepfiffen wurde (zeitliche Distanz). Als Indikatorvariable für die Bewertung der Ausgangslage einer Mannschaft bei Spielantritt als Außenseiter oder Favorit haben wir die Differenz der Wettquoten für die

⁷ In den gezeigten Analysen werden diesen Wahlkreisen Wolfsburger Ergebnisse zugeteilt.

⁸ Ein Gutachter/eine Gutachterin schlug vor, die großen Vereine wie Dortmund und Bayern München von der Analyse auszuschließen. Dies würde aber nichts an der geografischen Konzentration als Problem ändern, da dies ja nicht nur die „großen“ Vereine betrifft. In einer perfekten Datenwelt könnte man die Analyse umdrehen und für alle n Partien unter Berücksichtigung der regionalen Fanbasis die Ergebnisse auf das lokale Zweitstimmenergebnis in n Analysen schätzen, was leider nicht durchführbar ist.

Tab. 1 BL-Tabelle vor dem 6. Spieltag 2013/2014. (Quelle: Online-Archiv der Deutschen Bundesliga und RSSSF-Archiv. Eigene Darstellung)

Platz	Verein	Spiele	Siege	Unent-schieden	Nieder-lagen	Torver-hältnis	Tordif-ferenz
1	Borussia Dortmund	5	5	0	0	15:4	11
2	Bayern München	5	4	1	0	9:2	7
3	Bayer Leverkusen	5	4	0	1	11:6	5
4	Hannover 96	5	3	0	2	8:7	1
5	Mainz 05	5	3	0	2	8:8	0
6	FC Augsburg	5	3	0	2	5:7	-2
7	1899 Hoffenheim	5	2	2	1	14:13	1
8	Hertha BSC Berlin	5	2	1	2	9:6	3
9	Schalke 04	5	2	1	2	7:9	-2
10	VfB Stuttgart	5	2	0	3	10:8	2
11	Borussia Mönchengladbach	5	2	0	3	11:10	1
12	VfL Wolfsburg	5	2	0	3	7:7	0
13	Eintracht Frankfurt	5	2	0	3	7:9	-2
14	Werder Bremen	5	2	0	3	3:8	-5
15	Hamburger SV	5	1	1	3	10:15	-5
16	1. FC Nürnberg	5	0	3	2	5:8	-3
17	SC Freiburg	5	0	2	3	7:11	-4
18	Eintracht Braunschweig	5	0	1	1	2:10	-8

einzelnen Spiele erhoben. Die Wettquoten für die einzelnen Spiele wurden aus einer Online-Quelle händisch eingearbeitet.⁹ Wenn ein Verein 2009 nicht in der Ersten Bundesliga war, wurden für die Differenzbildung die jeweiligen Ergebnisse in der zweiten oder dritten Bundesliga herangezogen.

Betrachtet man die Tabelle vom Spieltag vor dem Wahlwochenende (Tab. 1), in diesem Fall die Abschlusstabelle des 5. Spieltags der Saison 2013/2014, zeigt sich, dass die meisten Mannschaften bei der Anzahl der zu diesem Zeitpunkt gesammelten Punkte noch sehr nah beieinanderlagen. So betrug zu diesem Zeitpunkt die Punktedifferenz zwischen dem Tabellenführer und dem Tabellendritten drei Punkte, wohingegen am Ende der Saison die Differenz 26 Punkte betrug. Zwischen dem Tabellenführer und dem Tabellenletzten betrug der Unterschied nach dem 5. Spieltag 14 Punkte, am Ende der Saison waren es 65 Punkte. Ferner weist diese Tabelle ein weiteres Charakteristikum früher Tabellenstände auf, nämlich, dass einige Mannschaften im Verhältnis zu ihrer allgemeinen Stärke noch keine besonders aussagekräftige Tabellenposition belegen. So sind Vereine wie Schalke, Wolfsburg und Mönchengladbach nach dem 5. Spieltag noch im Mittelfeld der Tabelle, wohingegen sie sich am Ende der Saison alle für die europäischen Wettbewerbe qualifizieren konnten und daher über die komplette Saison gesehen zu den stärkeren Teams gehörten.

Sieht man die Spieltagergebnisse des 6. Spieltags kombiniert mit den Wettquoten (Tab. 2), so fällt auf, dass die Quoten der Wettanbieter das Ergebnis nur bei

⁹ <http://www.football-data.co.uk/>. Zugegriffen: 17. Mai 2019.

Tab. 2 Spielpaarungen am 6. Spieltag 2013/14 inklusive Wettquoten. (Quelle: Online-Archiv der Deutschen Bundesliga und RSSSF-Archiv. Eigene Darstellung)

Datum	Anstoß	Heim	Auswärts	Resultat	Quote ^a Heim	Quote Remis	Quote Aus- wärts
Freitag, 20.09.2013	20:30 Uhr	Borussia Mönchen- gladbach	Eintracht Braun- schweig	4:1	1,38	4,74	8,19
Samstag, 21.09.2013	15:30 Uhr	Hannover 96	FC Augsburg	2:1	2,42	3,37	2,93
Samstag, 21.09.2013	15:30 Uhr	VfL Wolfs- burg	1899 Hof- fenheim	2:1	1,73	3,81	4,66
Samstag, 21.09.2013	15:30 Uhr	1. FC Nürn- berg	Borussia Dortmund	1:1	8,44	5,12	1,35
Samstag, 21.09.2013	15:30 Uhr	FSV Mainz 05	Bayer 04 Leverkusen	1:4	3,44	3,44	2,10
Samstag, 21.09.2013	15:30 Uhr	Hamburger SV	SV Werder Bremen	0:2	1,98	3,53	3,71
Samstag, 21.09.2013	18:30 Uhr	Schalke 04	FC Bayern München	0:4	5,77	4,20	1,55
Sonntag, 22.09.2013	15:30 Uhr	SC Freiburg	Hertha BSC Berlin	1:1	2,49	3,36	2,81
Sonntag, 22.09.2013	17:30 Uhr	VfB Stuttgart	Eintracht Frankfurt	1:1	2,17	3,48	3,21

^aDie Quoten sind jeweils arithmetische Mittel aller Wettquoten unseres Datensatzes

fünf Begegnungen zutreffend vorhersagen – ein Hinweis auf die große Rolle des Zufalls beim Fußball sowie andere kurzfristige Faktoren wie Tagesform, Taktik und Verletzungen. Bei immerhin vier Begegnungen hat sich nicht der favorisierte Verein durchgesetzt, wobei von diesen vier Paarungen drei Spiele unentschieden ausgegangen sind; in nur einer Partie (Hamburger SV gegen Werder Bremen, siehe Einleitung) gewann die Auswärtsmannschaft, obwohl die Heimmannschaft die bessere Quote aufwies. Daneben fällt vor allem das Remis zwischen Nürnberg und Dortmund auf, für das im Vorfeld Dortmund als klarer Favorit gehandelt wurde. Außerdem sticht heraus, dass drei Mannschaften gleich vier Tore erzielten und zwei davon (Bayer Leverkusen und Bayern München) dies auswärts gelang. Vier Tore für eine Mannschaft ist daher ein relativ hoher Wert, da im Durchschnitt während dieser Saison eine Mannschaft etwa 1,5 Tore pro Spiel erzielte.

4.3 Analysestrategie

In der Analyse werden die in Prozentpunkten berechneten Veränderungen in den Ergebnissen zwischen 2009 und 2013 genutzt (siehe Tab. 3). Wenn sich also eine Partei von 25,0 auf 25,5 % oder von 5,0 auf 5,5 % verbessert, enthält die abhängige Variable die Ausprägung +0,5 Prozentpunkte. Die Datenstruktur hat eine First-Difference-Form mit nur einer Differenz. Bei den OLS-Regressionsanalysen interessiert uns als zentrale unabhängige Variable das Resultat der einzelnen Bundesligapaarungen, die mit Beteiligung einer Mannschaft im Wahlkreis oder im Nachbarwahlkreis

Tab. 3 Übersicht über die Zweitstimmenveränderungen zwischen den Bundestagswahlen 2009 und 2013 in 75 (CDU) und 88 (andere Parteien) Wahlkreisen. (Quelle: Eigene Darstellung)

Variable	Mittelwert	Std.-Abw	Min	Max
<i>Amts inhaber</i>				
CDU	7,2	2,1	3,4	12,6
FDP	-9,8	2,0	-13,5	-6,1
<i>Opposition</i>				
SPD	3,5	1,3	-0,3	5,9
B'90/Grüne	-2,4	0,7	-3,8	-1,2
Linke	-3,0	1,5	-9,3	-0,9

stattgefunden haben. Hierbei werden der Logik der First-Difference-Schätzungen folgend Veränderungen des Spielergebnisses im Vergleich zum Jahre 2009 als Treatmentvariablen verwendet. Als Kontrollvariablen werden zudem Veränderungsraten sozioökonomischer Strukturvariablen der untersuchten Wahlkreise zwischen 2009 und 2013 herangezogen. Bei der Analyse der CDU klammern wir die CSU und die bayrischen Wahlkreise aus. Somit bezieht sich die Analyse auf 75 (CDU) oder 88 (alle anderen Parteien) Wahlkreise.

Die OLS-Regressionsanalyse enthält robuste Standardfehler, um Heteroskedastizität zu berücksichtigen, die aufgrund des X-zentrierten Designs mit einem sehr sparsamen Gesamtmodell entsteht. Gerichtete Hypothesen werden mit einseitigen Tests getestet, deren p -Wert explizit gezeigt wird. Obwohl die vorliegenden Daten eine Vollerhebung für das Universum der Veränderungen der Wahlkreise zwischen den Bundestagswahlen 2009 und 2013 darstellen, können die inferenzstatistischen Tests für die Fußballergebnisse unter der Annahme der natürlich-experimentellen Zuteilung als Wahrscheinlichkeit interpretiert werden, mit der solch ein kausaler Effekt in einer zufälligen Realisation der Werte der jeweiligen Variablen diese Richtung und Größe hat. Beispielsweise bedeutet ein p -Wert von 0,04, dass es eine vierprozentige Wahrscheinlichkeit gibt, eine Manifestation des experimentellen Effektes vorzufinden, der nicht in der postulierten Richtung liegt. Wir berichten die präzisen p -Werte und nutzen das 10%-Niveau für die Interpretation aufgrund der kleinen Stichprobe von maximal 88 Wahlkreisen.

Es werden für beide Operationalisierungen des Fußballergebnisses ein Modell jeweils ohne (bivariate Regression) und mit Strukturindikatoren (multiple Regression) auf Wahlkreisebene geschätzt.

5 Empirische Analyse

5.1 Überprüfung der Spielergebnisse als quasi-zufällige Zuteilung in einem natürlichen Experiment

Grundlage unserer Argumentation ist die testbare Überlegung, dass Fußballergebnisse als natürliche Experimente, die den Wahlausgang beeinflussen, aufgefasst werden können. Deshalb müssen wir an dieser Stelle zuallererst die Vorbedingung nachwei-

sen, dass tatsächlich ein natürliches Experiment vorliegt. Dabei ist zu beachten, dass die Existenz eines lokalen Erstligaver eins vor Ort nicht Teil des natürlichen Experiments ist, sondern nur die konkreten Spielergebnisse ergeben die Präsenz eines solchen Vereins. Wir können keine empirischen Aussagen über Wahlkreise treffen, die keinen Erstligaver ein beheimaten oder an einen solchen Wahlkreis angrenzen.

In einem ersten Schritt wurde innerhalb der Wahlkreise, die Bundesligamannschaften beherbergen bzw. an einen „Bundesligawahlkreis“ angrenzen, überprüft, ob sich Wahlkreise mit Mannschaften, die am untersuchten Spieltag vor der Bundestagswahl 2013 gewonnen haben, systematisch von Wahlkreisen unterscheiden, deren Heimmannschaft verloren bzw. unentschieden gespielt hat. Hier wurden zum einen Aspekte des Wahlergebnisses (Zweitstimmenanteile im Jahr 2009) und zum anderen relevante Strukturindikatoren verglichen. Außerdem wurde überprüft, ob der Favoriten- oder Außenseiterstatus einer Mannschaft systematisch mit dem Wahlergebnis oder Strukturindikatoren zusammenhängt.

Bei der Frage, ob sich Wahlkreise, in denen gewonnen wurde, systematisch von solchen unterscheiden, in denen verloren wurde, zeigt sich ein deutliches Bild. Da wir drei Gruppen von Wahlkreisen (Sieg, Niederlage, Unentschieden) untersuchen, wurden zum Vergleich der Mittelwerte ANOVAs durchgeführt. Hier ist nur für einen Strukturindikator ein signifikanter Unterschied gefunden worden. Der Anteil niedrig gebildeter Wählerinnen und Wähler in Wahlkreisen, in denen die Heimmannschaft verloren hat, ist um etwa drei Prozentpunkte höher als in Wahlkreisen, in denen die Heimmannschaft gewonnen hat. Alle Regressionsanalysen wurden deshalb auch noch unter Kontrolle des Anteils niedrig Gebildeter geschätzt, wobei die Ergebnisse unter Inklusion dieser Variable qualitativ identisch bleiben.

Wahlkreise, die Favoriten auf den Sieg beherbergen, unterscheiden sich strukturell ebenfalls wenig von „Außenseiter“-Wahlkreisen. Dies zeigen bivariate Korrelationen (Pearsons r) zwischen Wettquoten und dem Wahlergebnis bzw. den Strukturindikatoren. Dabei hat kein Korrelationskoeffizient ein konventionelles Maß statistischer Signifikanz erreichen können, sodass wir die Annahme bestätigt sehen, dass die Favoriten des betrachteten Spieltages nicht in systematisch unterschiedlicheren Wahlkreisen beheimatet waren als die Außenseiter. Diese Abwesenheit systematischer Zusammenhänge konnte für Wettquoten verschiedener Wettanbieter nachgewiesen werden.

In der Summe legen die Ergebnisse nahe, dass wir den Charakter des natürlichen Experiments bis auf eine Variable nachweisen können. Selbst bei einer Vielzahl von Vergleichen zwischen Wahlkreisen mit unterschiedlichen natürlichen Treatments finden wir nur eine signifikante und vor allem keine substanz iell bedeutsamen Unterschiede zwischen Wahlkreisen. Damit können wir uns im Folgenden dem eigentlichen Kernstück der Analyse widmen, nämlich der Frage, ob und in welcher Art und Weise sich Fußballresultate auf das Wahlergebnis auswirken.

5.2 Multiple Regressionsanalyse der kausalen Effekte der Spielergebnisse auf Zweitstimmenanteile

Dieser Abschnitt dient der Diskussion der Regressionsanalyseergebnisse über den Effekt von Fußballresultaten auf Wahlergebnisse. In einem ersten Schritt werden die

Tab. 4 First-Difference OLS-Regressionen Zweitstimmenanteile 2009–2013 mit *Differenz der Tordifferenzen* als zentraler Experimentalvariable. (Quelle: Eigene Darstellung)

	Amtsinhaber				Opposition					
	CDU		FDP		SPD		B'90/ GRUENE		LINKE	
	b	p	b	p	b	p	b	p	b	p
<i>Biv. Reg</i>										
Tordifferenz Ver. 2009 zu 2013	0,19	0,04	0,17	0,01	-0,02	0,39	-0,02	0,21	-0,14	0,03
<i>Multip. Reg</i>										
Tordifferenz Ver. 2009 zu 2013	0,21	0,02	0,12	0,03	-0,04	0,27	-0,02	0,17	-0,10	0,02
Konst. bivariat	7,39	0,00	-9,53	0,00	3,53	0,00	-2,41	0,00	-3,13	0,00
Konst. multivariat	8,04	0,00	-10,67	0,00	2,96	0,00	-2,26	0,00	-3,20	0,00
N	75	-	88	-	88	-	88	-	88	-
R ² bivariat	0,044	-	0,046	-	-0,011	-	-0,002	-	0,059	-
R ² multivariat	0,226	-	0,272	-	0,151	-	0,265	-	0,318	-

Einseitige Hypothesentests für Tordifferenz, multiple Regressionen beinhalten Kontrollvariablen Bevölkerungsdichte, Arbeitslosenquote, Anteil 60+, Anteil gering Gebildeter

direkten Effekte diskutiert, während der zweite Schritt die Betrachtung moderierter Effekte darstellt. Tab. 4 enthält die geschätzten Effekte der Differenz der Tordifferenzen auf die Zweitstimmenanteile der fünf im Bundestag vertretenen Parteien (CDU, FDP, SPD, B'90/Grüne, Linke), einmal ohne und einmal mit Kontrollvariablen. Das wichtigste Ergebnis ist, dass die Punktschätzungen für die beiden Amtsinhaberparteien positiv sind, während sie für die Oppositionsparteien negativ sind. Dies bestätigt im Kern die These, dass das gute Abschneiden der eigenen Mannschaft relativ zur „Geschichte“ der Ergebnisse (im relativen Vergleich zu 2009) einen positiven Effekt auf Amtsinhaberinnen und Amtsinhaber hat. Ein zusätzliches Tor in der Tordifferenz relativ zu 2009 hat einen geschätzten Effekt von 0,19 Prozentpunkten für die CDU ($p=0,04$) und von 0,17 für die FDP ($p=0,01$) in der einfachen bivariaten Regression. Die geschätzten Effekte für die Regressionen mit Kontrollvariablen sind fast identisch, da, wie oben gezeigt, fast kein Zusammenhang zwischen Strukturindikatoren und Fußballergebnissen vorliegt. Bei den Oppositionsparteien SPD, B'90/Grüne und Linke liegen die Punktschätzungen zwar in der hypothetisierten Richtung: Ein Anstieg in der Tordifferenz geht laut den Modellen mit dem Rückgang des jeweiligen Zweitstimmenergebnisses einher. Jedoch offenbaren die p -Werte nur für die Linke signifikante Effekte ($-0,14$, $p=0,03$ im bivariaten Modell, $-0,10$, $p=0,02$ im multiplen Regressionsmodell).¹⁰ Bezüglich der Oppositionsparteien kann man

¹⁰ Grundsätzlich fällt bei den Modellen eine geringe Passgüte beim R^2 auf, das zum Teil sogar negativ ist. Diese geringen Werte beinhalten, dass die Varianz unserer jeweiligen abhängigen Variablen nicht gut erklärt wird. Dies ist aber in einem x -zentrierten Design, das unserem experimentellen Design unterliegt, nachrangig, solange die jeweiligen Effekte des Treatments klar geschätzt werden können.

Tab. 5 First-Difference OLS-Regressionen Zweitstimmenanteile 2009–2013 mit *Ergebnisdifferenz* als zentraler Experimentalvariable. (Quelle: Eigene Darstellung)

	Amtsinhaber				Opposition						
	CDU		FDP		SPD		B'90/ GRUENE		LINKE		
	b	p	b	p	b	p	b	p	b	p	
<i>Biv. Reg</i>											
Ergebnis- differenz	0,35	0,05	-0,56	0,99	-0,35	0,004	0,20	0,999	0,10	0,83	
<i>Multip. Reg</i>											
Ergebnis- differenz	0,08	0,36	-0,28	0,99	-0,28	0,02	0,19	0,999	0,09	0,76	
Konst. biva- riat	7,17	0	-9,60	0,00	3,6	0,00	-2,42	0,00	-3,04	0,00	
Konst. multi- vari- at	7,82	0	-	-	-	-	-2,36	-	-3,16	0	
N	75	-	88	-	88	-	88	-	88	-	
R ² bivariat	0,028	-	0,08	-	0,070	-	0,088	-	0,005	-	
R ² multivari- at	0,211	-	0,306	-	0,232	-	0,372	-	0,326	-	

Einseitige Hypothesentests für Tordifferenz, multiple Regressionen beinhalten Kontrollvariablen Bevölkerungsdichte, Arbeitslosenquote, Anteil 60+, Anteil gering Gebildeter

spekulieren, dass sich die Abstrafung der Regierungsparteien unterschiedlich in der Wahl einer bestimmten Oppositionspartei widerspiegelt. Beispielsweise könnte es sein, dass die Linke als die deutlichste Anti-Establishment-Partei wahrgenommen wurde, sodass sich die Unzufriedenheit aufgrund des relativ schlechten Abschneidens der lokalen Mannschaft eher in der Wahl der Linken niederschlug. Unsere Hypothesen 1a und 1b finden also über alle aufgedeckten empirischen Muster hinweg Unterstützung für die Tordifferenz als Experimentalvariable, wenngleich die Effekte nicht immer signifikant sind.

Tab. 5 zeigt die analogen Ergebnisse der Modelle mit der Ergebnisdifferenz als Experimentalvariable. Hier wird gemessen, ob eine qualitative Veränderung des Ergebnisses von 2013 gegenüber 2009 einen kausalen Effekt auf die Veränderungen der Zweitstimmenanteile hat. Die Variable nimmt den Wert 0 an, wenn die Mannschaft mit dem gleichen Ergebnis abschneidet wie vier Jahre zuvor. Wenn sie sich von einer Niederlage im Jahr 2009 (-1) zu einem Sieg 2013 (+1) verbessert, hat sie den Wert +2. Alle anderen Veränderungen liegen dazwischen. Das beobachtete Minimum ist -1.

Hier zeigt sich nun ein deutlich gemischteres Bild, weshalb die Hypothesen 1a und 1b mit dieser Operationalisierung insgesamt eher zurückgewiesen werden. Nur CDU und SPD zeigen die erwarteten Punktschätzungen. Wenn sich die Heimmannschaft um eine Ergebnisstufe verbessert, steigt das CDU-Zweitsergebnis geschätzt um 0,35 Prozentpunkte an. Für die SPD sinkt das Zweitstimmenergebnis um den gleichen Wert. Für die SPD finden wir auch, dass die Inklusion der Kontrollvariablen den Signifikanzwert des Experimentaleffekts unter das 10-Prozent-Niveau senkt. Für

die anderen Parteien FDP, B'90/Grüne und Linke müssen wir die Hypothesen 1a und 1b in dieser Operationalisierung klar zurückweisen.

Wir können also festhalten, dass die Hypothesen 1a und 1b in der Operationalisierung mit der Differenz der Tordifferenzen im Großen und Ganzen Unterstützung finden, sich das erwartete Muster bei der Ergebnisdifferenz aber nur für CDU und SPD herausarbeiten lässt. Trotzdem sind diese ersten Ergebnisse bemerkenswert, weil sie relativ flüchtige Effekte auf Individualniveau im Aggregat zumindest in Teilen nachweisen.

Der nächste Schritt der Analyse befasst sich mit der Überprüfung von Interaktionseffekten für die Hypothesen 2a und 2b (zeitliche Distanz bis zur Schließung der Wahllokale) sowie 3a und 3b (Wettquoten). Zuerst schauen wir, inwieweit gegebenenfalls mit der Zeit schwächer werdende Emotionen sich durch kleinere Effekte auf das Zweitstimmenergebnis nachweisen lassen (H2a, b). Da wir uns im Differenz-in-Differenzen-Design bewegen, schauen wir dabei die relative Veränderung der zeitlichen Distanz des Spiels zum Wahltag 2013 im Vergleich zu 2009 an. Wir konzentrieren uns auf die Tordifferenz und fügen die Kontrollvariablen hinzu.

Für die Differenz des Abstandes zwischen Schließungszeitpunkt und Wahl zeigt Tab. 6, dass nur in zwei Fällen eine signifikante Modellverbesserung durch die Moderation vorliegt (CDU und B'90/Grüne). Für alle Parteien erwarten wir, dass die marginalen Effekte der Tordifferenz sich umso stärker von Null entfernen, je näher das Spiel der Heimmannschaft am Wahlzeitpunkt liegt. Dementsprechend erwarten wir gemäß Hypothese 2a für die Regierungsparteien CDU und FDP, dass die marginalen Effekte für negative Werte der Moderatorvariable stark positiv sind und sich dann für positive Werte der Moderatorvariable der Null annähern. Dieses Muster können wir auch beobachten, wobei es für die FDP nicht deutlich genug ist, um das Modell insgesamt zu verbessern ($p=0,351$). Trotzdem gilt laut den Modellen: Je näher das Spiel der Heimmannschaft relativ zu 2009 an der Schließung der Wahllokale lag, desto besser ist der positive Effekt eines weiteren Tores in der Differenz der Tordifferenzen für die Regierungsparteien (aber nur signifikant für die CDU). Daraus folgt eine mäßige Unterstützung der Hypothese 2a.

Für die drei Oppositionsparteien finden wir in den Punktschätzungen eindeutig das erwartete Muster: Je kleiner der Wert der Moderatorvariable ist, desto schlechter ist der vorhergesagte Wert für diese Parteien bei einem weiteren Tor auf der Tordifferenzvariable. Ein positives Abschneiden in näher an der Wahl liegenden Spielen ist schlechter für die Oppositionsparteien als zeitlich weiter zurückliegende Spiele. Doch sind diese Muster nur klar genug für B'90/Grüne, um das Modell statistisch zu verbessern, was Hypothese 2b zwar in der Tendenz, aber nicht eindeutig mit statistischer Signifikanz unterstützt.

Zweitens wenden wir uns nun der Frage zu, ob sich ein überraschender Sieg oder eine überraschende Niederlage der Heimmannschaft stärker auf das Wahlergebnis auswirken. Die Wettquoten sind dabei die bestmögliche Messung des Außenseiter- bzw. Favoritenstatus der jeweiligen Mannschaft, denn sie beinhalten alle zum Zeitpunkt der Ansetzung verfügbaren Informationen wie Tabellenplatz, Form der Mannschaft und weitere Variablen. Beispielsweise war die Quote auf einen Sieg von Schalke 04 gegen Bayern München etwa 5,5, während man für einen Sieg von Bayern München für jeden Euro nur 1,53 € zurückbekommen hat. Die Quote auf

Tab. 6 Interaktionsanalyse der Moderation der Differenz im Timing des Spiels. (Quelle: Eigene Darstellung)

		Marginale Effekte (Punktschätzung und <i>p</i> -Wert in Klammern) der Tordifferenzen-Differenz für Minimum, Mittelwert und Maximum der Moderatorvariable zeitliche Distanz bis zur Schließung der Wahllokale		
<i>p</i> -Wert des F-tests zur Inklusion des Produktterms		Min = -24 Spiel 24 h näher an Wahl '13 als '09	MW = 2,1	Max = 43 Spiel 43 h weiter weg von Wahl '13 als '09
<i>Amtsinhaber</i>				
CDU	0,039	0,6 (0,015)	0,21 (0,043)	-0,49 (0,122)
FDP	0,351	0,4 (0,107)	0,17 (0,013)	-0,22 (0,61)
<i>Opposition</i>				
SPD	0,140	-0,21 (0,171)	-0,04 (0,559)	0,26 (0,167)
B'90/ Grüne	0,001	-0,18 (0,001)	-0,03 (0,17)	0,23 (0,005)
Linke	0,120	-0,37 (0,025)	-0,14 (0,015)	0,27 (0,306)

Kontrollvariablen nicht gezeigt, $N = 75$ (CDU) und 88 (andere Parteien)

Robuste Standardfehler in First-Differenz-OLS-Regressionen der Zweitstimmenveränderung 2009–2013

einen Sieg des 1. FC Nürnberg gegen Borussia Dortmund betrug sogar 8, während die Buchmacher für einen Sieg von Dortmund nur eine Quote von 1,33 errechnet hatten. Aufgrund des Differenz-in-Differenzen-Designs nutzen wir die Veränderung der Quote der Heimmannschaft relativ zu 2009.

Tab. 7 enthält alle Schätzungen für die Wettquoten-Differenz als Moderatorvariable. Gemäß unserer Hypothese 3 müssen zwei notwendige Bedingungen empirisch erfüllt werden. Erstens muss die Berücksichtigung der Interaktion eine signifikante Modellverbesserung mit sich bringen. Dies ist laut F-Test in der ersten Spalte nur für die CDU und die SPD der Fall, nicht für FDP, B'90/Grüne oder Linke. Das bedeutet, dass hier die Moderation in drei Fällen keine signifikante Verbesserung der Erklärungskraft bringt. Zweitens erwarten wir, dass die Amtsinhaberparteien besonders stark von einer positiven Tordifferenz der Heimmannschaft profitieren, wenn diese relativ zu 2009 einen besonders großen Außenseiterstatus hat; für Oppositionsparteien erwarten wir genau das Gegenteil. Jedoch zeigen die Schätzungen der marginalen Effekte dieses Muster nicht. Die marginalen Effekte für CDU und FDP müssten bei einem großen Wert der Moderatorvariable ganz rechts beim Maximum positiv und besonders groß sein und dann für kleine Werte der Moderatorvariable gegen Null gehen, aber nicht ins Negative. Das gefundene Muster für die CDU ist der Erwartung entgegengesetzt, für die FDP konnte gar kein Muster identifiziert werden. Das Modell sagt stattdessen beispielsweise vorher, dass in Wahlkreisen, in denen die Heimmannschaft relativ zu 2009 besonders stark in die Außenseiterposition gerutscht ist, eine positive Tordifferenz mit einem starken negativen Effekt assoziiert ist. Für die FDP zeigen sich keine Moderationseffekte (was wir anhand des *p*-Werts des F-Tests bereits erkennen). In Summe zeigt keine der Parteien das

Tab. 7 Interaktionsanalyse der Moderation durch die Wettquotendifferenz. (Quelle: Eigene Darstellung)

		Marginale Effekte (Punktschätzung und <i>p</i> -Wert in Klammern) der Tordifferenz-Differenz für Minimum, Mittelwert und Maximum der Moderatorvariable Wettquotendifferenz		
<i>p</i> -Wert des F-tests zur Inklusion des Produktterms		Min = -1,57 Heimmannschaft wurden '13 bessere Gewinnchance eingeräumt als '09	MW = 0,98	Max = 6,64 Heimmannschaft wurden '13 schlechtere Gewinnchance eingeräumt als '09
<i>Amtsinhaber</i>				
CDU	0,003	0,79 (0,001)	0,41 (0,001)	-0,47 (0,032)
FDP	0,787	0,12 (0,153)	0,11 (0,074)	0,08 (0,513)
<i>Opposition</i>				
SPD	0,089	-0,14 (0,152)	-0,05 (0,415)	0,15 (0,174)
B'90/ Grüne	0,927	-0,03 (0,344)	-0,03 (0,174)	-0,04 (0,476)
Linke	0,218	-0,17 (0,053)	-0,11 (0,044)	0,03 (0,776)

Kontrollvariablen nicht gezeigt, $N = 74$ (CDU) und 87 (andere Parteien)

Robuste Standardfehler in First-Differenz-OLS-Regressionen der Zweitstimmenveränderung 2009–2013
Wettquotendifferenz für eine Beobachtung nicht berechenbar, weil Verein aus der Dritten in die Erste Bundesliga aufgestiegen war

erwartete Muster. Hier müssen wir konstatieren: Es gibt keinerlei Unterstützung der Außenseiterlogik mithilfe der vorliegenden Daten und gewählten Operationalisierungen.

So lässt sich zusammenfassend festhalten: Die Fußballergebnisse am Wochenende der Bundestagswahl 2013 folgten dem Muster eines natürlichen Experiments und können deswegen für sozialwissenschaftliche Kausaluntersuchungen bezüglich ihrer gesellschaftlichen Konsequenzen herangezogen werden. Die damaligen Regierungsparteien CDU und FDP profitierten von Gewinnen der Heimmannschaft, weshalb Hypothese 1a bestätigt werden kann. Die geschätzten Effekte sind eher gering (0,19 Prozentpunkte für die CDU und 0,17 Prozentpunkte für die FDP für ein Tor in der Tordifferenz), können aber im realen Wahlergebnis nachgewiesen werden. Für die Oppositionsparteien SPD, B'90/Grüne und Linke deuten die Muster in Richtung Hypothese 1b, sind aber nur für die Linken signifikant: Oppositionsparteien verlieren Anteile, wenn die Heimmannschaft schlecht abscheidet. Die Linke verliert pro Tor in der Tordifferenz 0,14 Prozentpunkte. Die gefundenen Muster folgen gemäß Hypothesen 2a und 2b bezüglich der Punktschätzungen, aber nicht der *p*-Werte der Signifikanztests, auch der erwarteten Moderation durch zeitliche Distanz: Je näher die Wahl, desto größer die Effekte. Doch ist diese Moderation nur für die CDU und B'90/Grüne signifikant. Keine Unterstützung finden wir für die Logik der Außenseiter gemäß den Hypothesen 3a und 3b. Amtsinhaber profitieren nicht umso stärker, je unwahrscheinlicher das gute Abschneiden der Heimmannschaft war. In Summe findet also nur Hypothese 1a klare Unterstützung.

5.3 Robustheitstests

In diesem Abschnitt werden die Ergebnisse der verschiedenen Robustheitsanalysen dargestellt, die testeten, ob die gefundenen Ergebnisse zufällig zustande gekommen sind und ob die Regressionen durch andere Spezifikationen andere Ergebnisse liefern. Zusätzlich wurde getestet, ob weitere fußballbezogene Variablen die Ergebnisse beeinflussen bzw. verändern.

Im ersten Schritt wurde ein Placebo-Test durchgeführt. Dabei wurde in allen Modellen als abhängige Variable das Wahlergebnis der Bundestagswahl 2009 verwendet. Falls unsere Modelle auch hier erklärungskräftig wären, so spräche das für zufällige Ergebnisse und gegen einen genuine Fußballereffekt. Die Ergebnisse sind allerdings eindeutig. In keinem Modell sind die Effekte des Fußballeregebnisses statistisch signifikant.

Des Weiteren wurde getestet, ob es einen Zusammenhang zwischen der lokalen Größe der Fanbasis¹¹ eines Bundesligavereins und dem Fußballereffekt gibt. Auch hier finden wir keine statistisch signifikanten Effekte. Schließlich wurde der Unterschied getestet, ob die Mannschaft zuhause oder auswärts spielt. Auch hier gibt es keinen signifikanten Effekt. Die Ergebnisse der Robustheitstests sind beispielhaft für die CDU/CSU dem Online-Anhang (Tab. 2) zu entnehmen.

6 Schlussfolgerungen

Dieser Beitrag begann mit der Überlegung, dass nicht-politische Ereignisse Emotionen bei Wählerinnen und Wählern auslösen. Diese Emotionen, beispielsweise Zufriedenheit und Euphorie oder aber Unzufriedenheit und Frustration, beeinflussen politische Wahlentscheidungen. Aus dieser Perspektive werden die Regierungsparteien an der Urne belohnt oder abgestraft, weil ihnen die Verantwortung für positive oder negative Emotionen zugeschrieben werden.

Diese sogenannte „blind retrospection“ der Mikroebene testeten wir auf ihre aggregierten Implikationen im Kontext der Bundestagswahl 2013. Dabei ist die zentrale Proposition, dass Bundesligaergebnisse Realisierungen eines natürlichen Experiments darstellen und damit die Schätzung eines kausalen Effekts auf das Wahlergebnis in den Wahlkreisen zulassen, die eine Bundesligamannschaft beheimaten.

Die Ergebnisse unserer empirischen Analysen unterstützen teilweise unsere Überlegungen, werfen aber zugleich eine Reihe von Widersprüchen und neuen Fragen auf. Zuerst zeigen die Analysen, dass sich Wahlkreise, in denen gewonnen wurde, nicht systematisch bezüglich sozioökonomischer Strukturvariablen von Wahlkreisen unterschieden haben, in denen der Bundesligist am betrachteten Spieltag vor der Bundestagswahl 2013 verlor. Damit weisen die konkreten Fußballergeb-

¹¹ 18 Fußballvereine haben die örtliche Verteilung ihrer Mitglieder nach Postleitzahlen der Berliner Morgenpost mitgeteilt, sodass sich eine räumliche Verteilung von Mitgliedern sehen lässt. Unter der Postleitzahl 47057 (Duisburg-Neudorf) gibt es beispielsweise folgende Verteilung: 53 Schalke, 4 HSV, 3 FC Köln, 2 Hertha BSC Berlin, 21 Gladbach, siehe: <http://interaktiv.morgenpost.de/beta-fussballkarte/#11/51.4525/6.8678>. Zugegriffen: 17. Mai 2019.

nisse die quasi-zufällige Zuteilung eines natürlichen Experiments auf. Weiterhin haben Fußballergebnisse elektorale Auswirkungen. Allerdings sind sie nicht immer so, wie man sie erwarten würde. Positive Spielausgänge gemessen anhand der Differenz der Tordifferenzen von 2009 und 2013 wirken sich positiv und signifikant auf das Wahlergebnis der Regierungsparteien CDU/CSU und FDP aus und negativ für die Oppositionsparteien SPD, B'90/Grüne und Linke, jedoch nur für die Linke mit einem signifikanten Schätzergebnis. Für die alternative Operationalisierung mit der Veränderung des Spielergebnisses (verloren, unentschieden, gewonnen) sind die Befunde deutlich widersprüchlicher.

Bezüglich der erwarteten Moderation ist die zeitliche Distanz des Spiels relativ zur Schließung der Wahllokale für die Effekte wichtig, wobei auch hier nicht alle geschätzten Effekte signifikant waren. Als überwiegendes Muster finden wir: Je näher das Spiel 2013 relativ zum Termin des Spiels 2009 lag, desto deutlicher sind die postulierten Effekte. Die Amtsinhaber profitierten vor allem dann von einem Sieg der lokalen Mannschaft, wenn das Spiel zeitlich nah an der Wahl lag, wobei nur die CDU ein signifikantes Ergebnis bei der Moderation aufweist. Oppositionsparteien verloren entsprechend, wobei das Muster nur bei B'90/Grüne signifikant war. Das unterstützt die Idee, dass diese kurzfristig hervorgerufenen Emotionen keine lange Halbwertszeit haben. Keine Unterstützung findet eine Moderation durch die Wettquotenveränderung relativ zu 2009. Ob eine Mannschaft gewann, der man in den Wettbüros wenig zutraute, spielt empirisch keine nachweisbare Rolle.

Nichtsdestoweniger sind wir in der Lage, einige bemerkenswerte Zusammenhänge zwischen Bundesligaergebnissen und Wahlverhalten zu identifizieren. Wir haben uns ausschließlich mit der Bundestagswahl 2013 bzw. den Veränderungen im Wahlergebnis zwischen 2009 und 2013 beschäftigt. Weitere Forschung kann versuchen, (1) diese Effekte auf individueller Ebene mithilfe von Umfragedaten nachzuzeichnen,¹² (2) die Effekte auch auf niedrigeren politischen Ebenen wie Landtagswahlen nachzuzeichnen und (3) die Daten um weitere Bundestagswahlen zu ergänzen. Zukünftige Forschung sollte auch durch Methoden wie Feld- oder Laborexperimente versuchen, den kausalen Mechanismen auf die Spur zu kommen.

Die eindeutigen Ergebnisse zur Eigenschaft eines natürlichen Experiments der Spielergebnisse der Ersten Fußballbundesliga der Herren lassen weiterhin die Option offen, Spielergebnissen Umfragedaten zuzuspielen, die andere abhängige Variablen wie z. B. die individuelle Unterstützung für die Regierung beinhalten. Solange man einen lokalen Identifier der Respondenten hat, der es erlaubt, Bundesligavereine den Befragten örtlich zuzuteilen, kann die hohe Frequenz von Fußballspielen und die leichte Verfügbarkeit relevanter „Fußball-Daten“ dazu genutzt werden, die Spielergebnisse systematisch als kausale Hebel für Emotionen, die ihren Ursprung außerhalb der Politik haben, miteinzubeziehen.

Die Ergebnisse zeigen, dass der Effekt nicht-politischer Ereignisse auf politisches Verhalten vorhanden ist und dass wir erst beginnen zu verstehen, wie politisches Verhalten mit nicht-politischen Ereignissen interagiert. Menschen stellen Bezüge

¹² Hier besteht das praktische Problem, dass um eine Bundestagswahl herum deutsche Umfrageinstitute typischerweise aufgrund einer Selbstverpflichtung nicht im Feld sind.

zwischen sozialen Ereignissen und der Politik her, die objektiv betrachtet Zweifel an Wählerationalität im Sinne von Sachbezogenheit nähren.

Danksagung Achim Goerres dankt Christian Breunig für Gespräche über den Zusammenhang zwischen Fußball und Wahlverhalten im internationalen Vergleich. Alle Autoren danken Lukas Fiege, Anne-Kathrin Fischer, Hayfat Hamidou, Sebastian Krause, Sabrina J. Mayer und Erik Wenker für Verbesserungsvorschläge. Alle Autoren danken den Gutachterinnen und Gutachtern und insbesondere dem betreuenden Redaktionsmitglied Kai-Uwe Schnapp für hilfreiche Hinweise.

Author Contribution Achim Goerres hat das Projekt koordiniert sowie die originären theoretischen Ausführungen, Methodenteil und Einleitung erstformuliert. Er ist hauptverantwortlich für alle Überarbeitungen nach den Gutachterhinweisen. Robin Arens hat die Daten gesammelt, die Literatur recherchiert, erste Voranalysen durchgeführt und die erste Version des Literaturberichts geschrieben. Florian Rabuza hat die empirische Analyse durchgeführt und die empirische Interpretation sowie die Conclusio erstformuliert. Alle Autoren haben die Endfassung zusammen redigiert. Eine Pilotstudie hat Robin Arens im Rahmen seiner BA-Arbeit an der Universität Duisburg-Essen unter Supervision von Achim Goerres durchgeführt. Florian Rabuza und Robin Arens waren zum Zeitpunkt der Bearbeitung Mitglieder der Universität Duisburg-Essen.

Literatur

- Achen, Christopher H., und Larry M. Bartels. 2012. Blind Retrospection. Why Shark Attacks Are Bad for Democracy. https://www.vanderbilt.edu/csdi/research/CSDI_WP_05-2013.pdf. Zugegriffen: 10. Juli 2019.
- Achen, Christopher H., und Larry M. Bartels. 2016. *Democracy for Realists. Why Elections Do Not Produce Responsive Government*. Princeton: Princeton University Press.
- Bartels, Larry M. 2010. The Study of Electoral Behavior. In *The Oxford Handbook of American Elections and Political Behavior*, Hrsg. Jan E. Leighley, 239–261. New York: Oxford University Press.
- Bauer, Gerrit. 2015. Natürliche Experimente in den Sozialwissenschaften. Ein Überblicksartikel mit ausgewählten Beispielen. *Soziale Welt* 22:268–288. Sonderband.
- Bechtel, Michael M., und Jens Hainmueller. 2011. How Lasting Is Voter Gratitude? An Analysis of the Short- and Long-Term Electoral Returns to Beneficial Policy. *American Journal of Political Science* 55:852–868. <https://doi.org/10.1111/j.1540-5907.2011.00533.x>.
- Blumenstiel, Jan E., und Elena Wiegand. 2014. Wechselwähler. In *Zwischen Fragmentierung und Konzentration. Die Bundestagswahl 2013*, Hrsg. Rüdiger Schmitt-Beck, 133–144. Wiesbaden: Nomos.
- Busby, Ethan C., und James N. Druckman. 2018. Football and Public Opinion. A Partial Replication and Extension. *Journal of Experimental Political Science* 5:4–10.
- Busby, Ethan C., James N. Druckman, und Alexandria Fredendall. 2017. The Political Relevance of Irrelevant Events. *The Journal of Politics* 79:346–350. <https://doi.org/10.1086/688585>.
- Dalton, Russell J. 2014. Interpreting Partisan Dealignment in Germany. *German Politics* 23:134–144. <https://doi.org/10.1080/09644008.2013.853040>.
- Depetris-Chauvin, Emilio, Ruben Duante, und Filipe R. Campante. 2018. *Building nations through shared experiences. Evidence from African football*. NBER Working Papers, Bd. 24666. Cambridge: National Bureau of Economic Research, Inc.
- Dixon, Mark J., und Stuart G. Coles. 1997. Modelling Association Football Scores and Inefficiencies in the Football Betting Market. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 46:265–280. <https://doi.org/10.1111/1467-9876.00065>.
- Dunning, Thad. 2008. Improving Causal Inference. Strengths and Limitations of Natural Experiments. *Political Research Quarterly* 61:282–293.
- Falter, Jean-Marc, und Christophe Pérignon. 2000. Demand for Football and Intramatch Winning Probability. An Essay on the Glorious Uncertainty of Sports. *Applied Economics* 32:1757–1765.
- Fiorina, Morris P. 1981. *Retrospective Voting in American National Elections*. New Haven: Yale University Press.
- Fowler, Anthony, und B. Pablo Montagnes. 2015. College Football, Elections, and False-positive Results in Observational Research. *Proceedings of the National Academy of Sciences* 112:13800–13804.
- Hagen, Lutz M., Reimar Zeh, und Maike Müller-Klier. 2015. Soccer in the Media, Public Mood and Government Popularity. *Zeszyty Prasoznawcze* 56:111–129.

- Healy, Andrew, und Neil Malhotra. 2010. Random Events, Economic Losses, and Retrospective Voting. Implications for Democratic Competence. *Quarterly Journal of Political Science* 5:193–208.
- Healy, Andrew, und Neil Malhotra. 2013. Retrospective Voting Reconsidered. *Annual Review of Political Science* 16:285–306.
- Healy, Andrew, Neil Malhotra, und Cecilia Hyunjung Mo. 2010. Irrelevant Events Affect Voters' Evaluations of Government Performance. *Proceedings of the National Academy of Sciences* 107:12804–12809.
- Huber, Gregory A., Seth J. Hill, und Gabriel S. Lenz. 2012. Sources of Bias in Retrospective Decision Making. Experimental Evidence on Voters' Limitations in Controlling Incumbents. *American Political Science Review* 106:720–741.
- Hunt, Kenneth A., Terry Bristol, und R. Edward Bashaw. 1999. A Conceptual Approach to Classifying Sports Fans. *Journal of Services Marketing* 13:439–452.
- Kerr, John H., George V. Wilson, Isamu Nakamura, und Yoshiko Sudo. 2005. Emotional Dynamics of Soccer Fans at Winning and Losing Games. *Personality and Individual Differences* 38:1855–1866.
- Kousser, Thad, und Megan Mullin. 2007. Does Voting by Mail Increase Participation? Using Matching to Analyze a Natural Experiment. *Political Analysis* 15:428–445.
- Lassen, David D. 2005. The Effect of Information on Voter Turnout. Evidence from a Natural Experiment. *American Journal of Political Science* 49:103–118. <https://doi.org/10.1111/j.0092-5853.2005.00113.x>.
- Marcus, George E. 2000. Emotions in Politics. *Annual Review of Political Science* 3:221–250.
- Porat, Amir B. 2010. Football Fandom. A Bounded Identification. *Soccer & Society* 11:277–290.
- Preuß, Holger, und Christian Alfs. 2013. Wirtschaftliche Bedeutung des Sportkonsums in Deutschland. *Sportwissenschaft* 43:239–252.
- Reep, C., und B. Benjamin. 1968. Skill and Chance in Association Football. *Journal of the Royal Statistical Society. Series A (General)* 131:581–585.
- Reep, C., R. Pollard, und B. Benjamin. 1971. Skill and Chance in Ball Games. *Journal of the Royal Statistical Society. Series A (General)* 134:623–629.
- Schoen, Harald. 2003. *Wählerwandel und Wechselwahl. Eine vergleichende Untersuchung*. Wiesbaden: Westdeutscher Verlag.
- Schwarz, Norbert, Fritz Strack, Detlev Kommer, und Dirk Wagner. 1987. Soccer, Rooms, and the Quality of Life. Mood Effects on Judgements of Satisfaction with Life in General and with Specific Domains. *European Journal of Social Psychology* 17:69–79.
- Sekhon, Jasjeet S., und Rocío Titunik. 2012. When Natural Experiments Are Neither Natural nor Experiments. *American Political Science Review* 106:35–57. <https://doi.org/10.1017/S0003055411000542>.
- Stokes, Leah C. 2016. Electoral Backlash against Climate Policy. A Natural Experiment on Retrospective Voting and Local Resistance to Public Policy. *American Journal of Political Science* 60:958–974. <https://doi.org/10.1111/ajps.12220>.
- Zeh, Reimar, und Lutz M. Hagen. 2006a. Fußball als Wahlentscheider? Wie die deutsche Nationalmannschaft politische Popularität beeinflusst. In *Fußball – Fernsehen – Politik*, Hrsg. Christina Holtz-Bacha, 193–213. Wiesbaden: VS.
- Zeh, Reimar, und Lutz M. Hagen. 2006b. Stimmungen und Wählerstimmen. Was die Papstwahl mit der Bundestagswahl zutun hat (und mit Fußball). In *Die Massenmedien im Wahlkampf. Die Bundestagswahl 2005*, Hrsg. Christina Holtz-Bacha, 338–355. Wiesbaden: VS.