

Analyse des Zusammenhangs zwischen den Wahlergebnissen und den Strukturmerkmalen auf Wahlbezirksebene in der Landeshauptstadt München

- Vorläufiger Kurzbericht -

Prof. Dr. Helmut Küchenhoff,
Dr. André Klima
M.Sc. Hannah Neppl
Dipl.-Stat. Uta Thien-Seitz
Dipl.-Stat. Sibel Aydemir-Kengeroglu
Dipl.-Stat. Ingrid Kreuzmair

Im Auftrag
des Statistischen Amtes der Landeshauptstadt München

Statistisches Beratungslabor
Institut für Statistik
Ludwig-Maximilians-Universität München

Ludwigstraße 33
80539 München

<http://www.stablab.stat.uni-muenchen.de>

München, den 19. Oktober 2017

1 Hintergrund

Die Veränderungen in der politischen Landschaft, die Verluste der Volksparteien in Metropolregionen wie München, das Erstarken der kleineren Parteien und das Entstehen von neuen Parteien erschwert zunehmend die Analyse von Wahlergebnissen. Gleichzeitig sorgen diese Veränderungen jedoch auch für ein großes öffentliches Interesse an entsprechend tiefgreifenden Analysen der Wahlen. Es zeigt sich weiterhin auch, dass bisher verwendete einfache Analyseverfahren nur unzureichend die relevanten Fragen beantworten können und klassische Verfahren, wie z.B. die Hochburgenanalyse, in der heutigen Zeit nicht mehr ausreichend sind.

Das Statistische Amt der Landeshauptstadt München als öffentlicher Dienstleister in München für Wahlanalysen hat daher begonnen den analytischen Teil der Wahlberichtserstattung zu erweitern. Bereits seit der Landtags- und der Bundestagswahl 2013 wird noch in der Wahlnacht anhand der vorläufigen amtlichen Wahlergebnisse die Wählerwanderung für München geschätzt und der Öffentlichkeit zur Verfügung gestellt.

Mit der Bundestagswahl 2017 erweitert das Statistische Amt der Landeshauptstadt München erneut den analytischen Teil. Der Schwerpunkt der neuen Analyse liegt auf dem Aspekt den Erfolg bzw. Nichterfolg von Parteien im Stadtgebiet besser verstehen zu können. Mithilfe von Regressionsverfahren soll dabei der Zusammenhang zwischen Strukturmerkmalen und den Wahlergebnissen in den Wahlbezirken ermittelt werden. Mit diesen Modellen wird alleinig der Zusammenhang zwischen den Merkmalen und den Ergebnissen geschätzt, d.h es sind anhand der Modelle Aussagen wie „In Gebieten mit vielen Älteren wird häufiger die CSU gewählt.“, nicht jedoch Schlüsse der Form „Ältere wählen häufiger die CSU.“ möglich. Bei der Modellentwicklung und Umsetzung hilft erneut das Statistische Beratungslabor der LMU.

Der Analyse werden Variablen aus dem Datenangebot des Statistischen Amtes auf Ebene der Wahlbezirkseinteilung zugrunde gelegt. Dies umfasst Informationen zur Bevölkerungsdichte, der Altersstruktur, dem Migrationshintergrund, der Veränderung des Ausländeranteils, der Kirchenzugehörigkeit, der Wohndauer in der Wohnung, den SGBII- und SGBIII-Empfängern, der Schulbildung, der Haushaltsstruktur sowie zu zugelassenen Pkws und der Kaufkraft pro Person. Diese Informationen werden auf sinnvolle Basisgrößen bezogen. Für viele ist dies die Hauptwohnsitzbevölkerung, es werden aber auch Haushalte oder z.B. alle Schulübertritte betrachtet. Bei den gehobenen Pkws wird der Anteil an allen zugelassenen Pkws betrachtet. Die Daten liegen zum Stand Dezember vor der jeweiligen Bundestagswahl vor, lediglich die Schulübertrittsquoten sind für das Jahr 2016 noch nicht vorhanden und es werden die Daten von 2015 verwendet.

Die Modelle werden auf Wahlbezirksebene in der jeweils aktuellen Einteilung geschätzt. Aufgrund von Veränderungen im Wahlbezirkzuschnitt sind diese jedoch in beiden Wahlen nicht identisch. Urnen- und Briefwahlbezirke werden gemeinsam betrachtet, die demographischen Merkmale für die Briefwahlbezirke werden als (gewichtete) Summe der jeweiligen zugehörigen Urnenwahlbezirke berechnet.

Im vorliegenden vorläufigen Bericht werden für die Wahlbeteiligung und die sechs größten Parteien sowie für die sonstigen Parteien als Summe für die Bundestagswahl 2013 und

2017 (für die Ergebnisse der Parteien auf Wahlbezirksebene siehe Abbildung 1) jeweils die drei stärksten Zusammenhänge sowie für alle Modelle der Zusammenhang mit dem Anteil der Deutschen mit Migrationshintergrund an der Hauptwohnsitzbevölkerung dargestellt. Die Stärke wird als Unterschied zwischen Minimum und Maximum ermittelt. Bei der Präsentation wird auf eine marginale Zusammenhangsdarstellung zurückgegriffen.

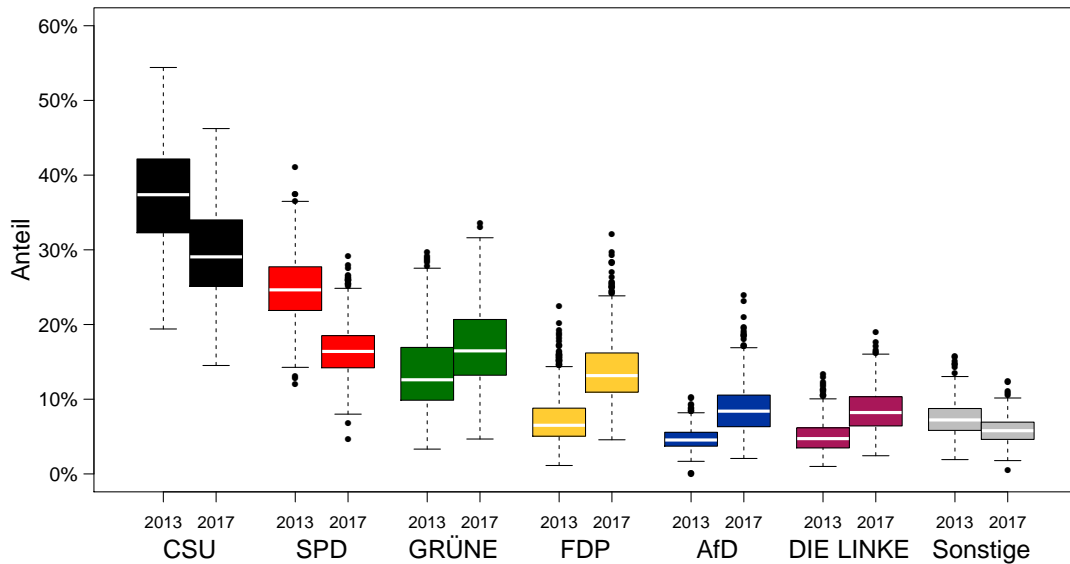


Abb. 1: Bundestagswahlergebnisse in den Wahlbezirken in München für die Wahlen 2013 und 2017, Basis sind die gültigen Stimmen. Dargestellt sind sechs ausgewählte Parteien, weitere Parteien sind unter Sonstige zusammengefasst. Der Wahlbezirkzuschnitt hat sich zwischen den Wahlen geändert.

2 Methodik

Zur Analyse des Zusammenhangs zwischen den Wahlergebnissen und den verschiedenen Strukturmerkmalen auf Wahlbezirksebene werden generalisierte additive Modelle (Fahrmeir et al.; 2009) verwendet. Die Umsetzung der Modellierung erfolgt mit der Statistik-Software R (R Core Team; 2017), primär über das R-Paket `mgcv` (Wood; 2006).

2.1 Generalisierte additive Modelle

Generalisierte additive Modelle vereinen Ansätze zur flexiblen Modellierung der Beziehung zwischen abhängigen und unabhängigen Merkmalen aus additiven Modellen und die generalisierte Anwendung dieser für nicht-normalverteilte Zielgrößen.

Für die konkrete Analyse ergibt sich für die Anteile in den binomialverteilten Zielgrößen y_i , die Wahlbeteiligung sowie die Wahlergebnisse verschiedener Parteien im Wahlbezirk i , folgende Modellgleichung, die den Zusammenhang mit den Variablen der Merkmale beschreibt ($i = 1, \dots, N$, wobei mit N die Gesamtzahl der Wahlbezirke angegeben wird):

$$g(\pi_i) = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \sum_{j=1}^q f_j(z_{i,j}).$$

Mit $P(y_i = 1) = \pi_i$ wird die Wahlbeteiligung oder das Wahlergebnis einer Partei im Wahlbezirk i angegeben, wobei $g(\cdot)$ für die logistische Linkfunktion steht. $\beta_0, \beta_1 x_{i,1}$ und $\beta_2 x_{i,2}$ beschreiben den Intercept, also die allgemeine Wahrscheinlichkeit für die Wahlbeteiligung, sowie kategoriale Effekte, die Funktionen f_1, \dots, f_q geben die flexible Modellierung, sogenannte glatte Effekte, für die nicht-linearen Effekte der q metrischen Variablen z_1, \dots, z_q an. Bei der verwendeten Prozedur wird automatisch anhand der Basisbevölkerung in den Wahlbezirken gewichtet.

2.2 Auswahl der Variablen

Der Zusammenhang verschiedener Merkmale mit der Wahlbeteiligung und den Wahlergebnissen der verschiedenen Parteien soll mithilfe der generalisierten additiven Modelle möglichst gut beschrieben werden. Die Integration aller vorhandenen Variablen der Merkmale in die Modelle kann jedoch instabile Schätzungen und damit keine sinnvolle Interpretation zur Folge haben, da einige der Variablen untereinander hoch korreliert sind. Dies gilt vor allem für Variablen, die Kategorien eines Merkmals (z.B. die Altersgruppen) darstellen, da sich deren Anteile in der Summe zu eins addieren. Relevante Korrelationen können jedoch auch zwischen Variablen unterschiedlicher Merkmale beobachtet werden.

Um ein sinnvolles Modell aufzustellen, werden daher vorab für die linearen Zusammenhänge zwischen den metrischen Variablen die Korrelationskoeffizienten nach Pearson (siehe beispielsweise Fahrmeir et al. (2016)) bestimmt, wobei für zwei größer als 0.7 miteinander korrelierte Variablen, jeweils eine nicht in das Modell mit aufgenommen wird. Bei Variablen, die Kategorien eines Merkmals sind, erfolgte vor diesem Schritt bereits eine inhaltliche Vorauswahl.

Alle Variablen, die in die Modellierung mit einbezogen werden, werden weiterhin in ihrem Wertebereich dahingehend beschränkt, dass, wenn nur sehr vereinzelte Wahlbezirke mit

sehr niedrigen oder sehr hohen Werten für die Variable existieren, der Wert für die Variable für diese Wahlbezirke auf einen näher an den Werten der anderen Wahlbezirke liegenden Wert gesetzt wird.

2.3 Bestimmen marginaler Effekte

Bei der Darstellung der geschätzten Zusammenhänge ergibt sich im Allgemeinen das Problem, dass bei einer binomialverteilten Zielgröße und der verwendeten Logit-Regression nur eine Aussage über die Veränderung der Chancen möglich ist. Mit der Bestimmung von marginalen Effekten wird versucht, den Zusammenhang zwischen den Kovariablen und der Zielgröße auf der Anteilsebene auch bei solchen Modellen interpretierbar zu machen.

Prinzipiell bestehen mehrere Möglichkeiten zur Berechnung marginaler Effekte. Der hier verwendete Ansatz ist die Betrachtung der mittleren Auswirkung einer Veränderung für eine der q betrachteten Variablen über mehrere Beobachtungen. Da mit den geschätzten Modellen eine Aussage über die Stadt getätigt werden soll, bietet es sich hier an die Information aller N Wahlbezirke zu verwenden:

$$f_{\text{marginal}}(z_k = K) = \frac{1}{N} \sum_{i=1}^N (\hat{\pi}(x_{i,1}, x_{i,2}, z_{i,1}, \dots, z_{i,q} | z_k = K)).$$

Für jeden möglichen Wert K der betrachteten k -ten Kovariable wird der Mittelwert der Prädiktion des Modells ($\hat{\pi}$) über alle Wahlbezirke gebildet. Dabei entsprechen die Ausprägungen der anderen Variablen den beobachteten Werten in den Wahlbezirken, für die k -te Kovariable wird hingegen immer der Wert $z_k = K$ angenommen. Über alle Werte für z_k ergibt sich daraus der marginale Zusammenhang zwischen der k -ten Kovariable und dem betrachteten Anteil. Die Differenz von f_{marginal} zwischen zwei z -Werten kann dann als marginaler Effekt zwischen den beiden z -Werten interpretiert werden.

Bei der Berechnung des Mittelwertes über die Prädiktionen wird noch eine Gewichtung mit der relevanten Basisbevölkerung im Wahlbezirk durchgeführt, da so die Unterschiede bei der Zahl der Wahlberechtigten bzw. Wähler besser berücksichtigt werden können. Weiterhin werden Unsicherheitsbereiche mit ausgegeben. Dabei handelt es sich jedoch nicht um statistische Konfidenzintervalle.

3 Modellergebnisse

Vorgestellt werden die stärksten Zusammenhänge der gerechneten acht Modelle für jede der beiden Bundestagswahlen. Neben der Wahlbeteiligung handelt es sich um Modelle für die großen Parteien (CSU, SPD, GRÜNE, FDP, AfD, DIE LINKE). Alle anderen Parteien werden zu Sonstigen zusammengefasst und gemeinsam betrachtet.

Die Wahlbeteiligung wird als Anteil der abgegebenen Stimmen an den Wahlberechtigten, die Parteiergebnisse als Anteil an den gültigen Stimmen berechnet. Bei den Briefwahlbezirken entspricht die Zahl der Wahlberechtigten weitestgehend der Zahl der Personen, die einen Wahlschein in den zugehörigen Urnenwahlbezirken beantragt haben und stellt somit die Grundlage der Berechnung dar.

Nach der Auswahl der Variablen ergibt sich daraus die folgende Modellformel, die Kovariablen sind bei allen acht Modellen gleich, die Zielgröße π variiert entsprechend:

$$\begin{aligned} g(\pi) = & \beta_0 + \beta_{\text{BW}}x_{\text{BW}} + \beta_{\text{BWK}}x_{\text{BWK}} + \\ & f(z_{\text{Bevölkerungsdichte}}) + f(z_{\text{Dt. mit Migrationshintergrund}}) + f(z_{\text{Veränderung Ausländeranteil}}) + \\ & f(z_{\text{18-24-Jährige}}) + f(z_{\text{75+-Jährige}}) + f(z_{\text{Keine Kirchengliederung}}) + \\ & f(z_{\text{Wohndauer 10+ Jahre}}) + f(z_{\text{Kaufkraft}}) + f(z_{\text{SGBIII-Empfänger (15-64 Jahre)}}) + \\ & f(z_{\text{Übertritte aufs Gymnasium}}) + f(z_{\text{Ohne Schulabschluss}}) + \\ & f(z_{\text{Mehrpersonenhaushalte mit Kindern}}) + f(z_{\text{Zugelassene Pkws}}) + f(z_{\text{Gehobene Pkws}}). \end{aligned}$$

Neben den vierzehn soziodemographischen Variablen, deren Zusammenhang glatt geschätzt wird, beinhaltet das Modell zusätzlich jeweils einen Indikator für die Briefwahlbezirke (BW) und die Bundestagswahlkreise (BWK). Präsentiert werden anschließend nur die drei stärksten glatten Zusammenhänge der einzelnen Modelle.

Die in den Grafiken dargestellten Zusammenhänge können wie folgt interpretiert werden: Für jeden Wert der betrachteten Kovariable wird der Anteil angegeben, der in München nach dem Modell vorhergesagt wird, wenn jeder Wahlbezirk genau diese Ausprägung bei dieser Kovariable hätte, die Werte der anderen im Modell aufgenommenen Kovariablen jedoch unverändert blieben. Interpretiert werden kann die Veränderung zwischen zwei Werten der Kovariable, indem der Verlauf der Funktion betrachtet wird. Diese Aussagen lassen dann Schlüsse über die Zusammenhänge der Kovariablen mit den Wahlergebnissen in München zu.

Beispiel:

Bei der Bundestagswahl 2013 wäre bei einem Anteil an Deutschen mit Migrationshintergrund an der Hauptwohnsitzbevölkerung von 15% in allen Wahlbezirken eine Wahlbeteiligung von knapp über 70% zu erwarten. Bei einem Anteil an Deutschen mit Migrationshintergrund an der Hauptwohnsitzbevölkerung von 25% wäre die erwartete Wahlbeteiligung mit einem vorhergesagten Wert von etwa 63% niedriger. Allgemein zeigt der gesamte Funktionsverlauf einen negativen Zusammenhang zwischen dem Anteil der Deutschen mit Migrationshintergrund an der Hauptwohnsitzbevölkerung und der Wahlbeteiligung.

Wahlbeteiligung

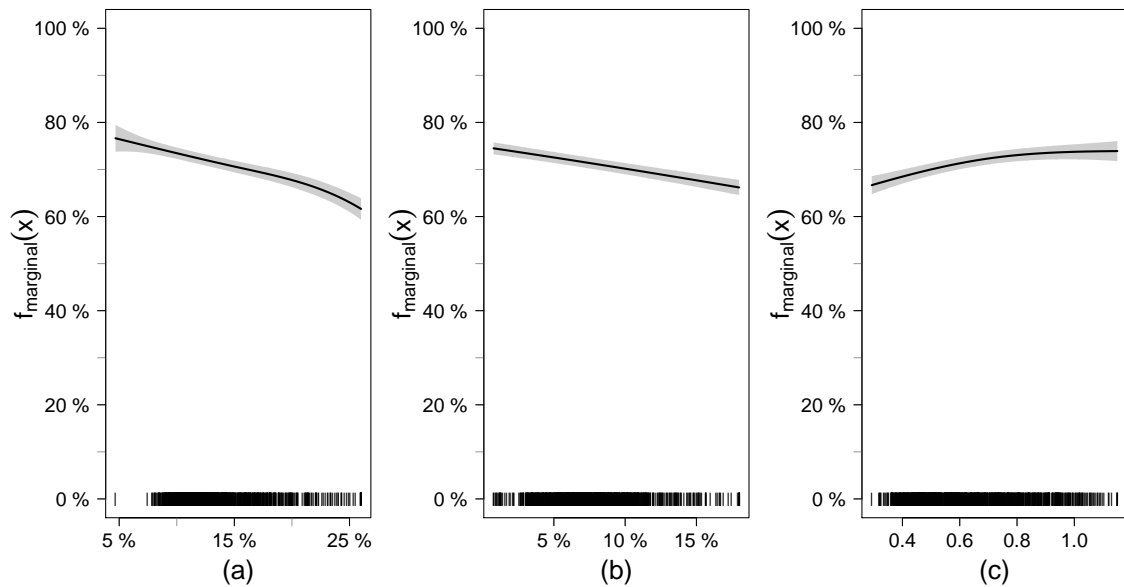


Abb. 2: Die drei stärksten Zusammenhänge zwischen der Wahlbeteiligung bei der Bundestagswahl 2013 und (a) dem Anteil der Deutschen mit Migrationshintergrund an der Hauptwohnsitzbevölkerung, (b) dem Anteil der Altersgruppe 75 Jahre und älter an der Hauptwohnsitzbevölkerung und (c) der Anzahl der zugelassenen Pkws pro Haushalt

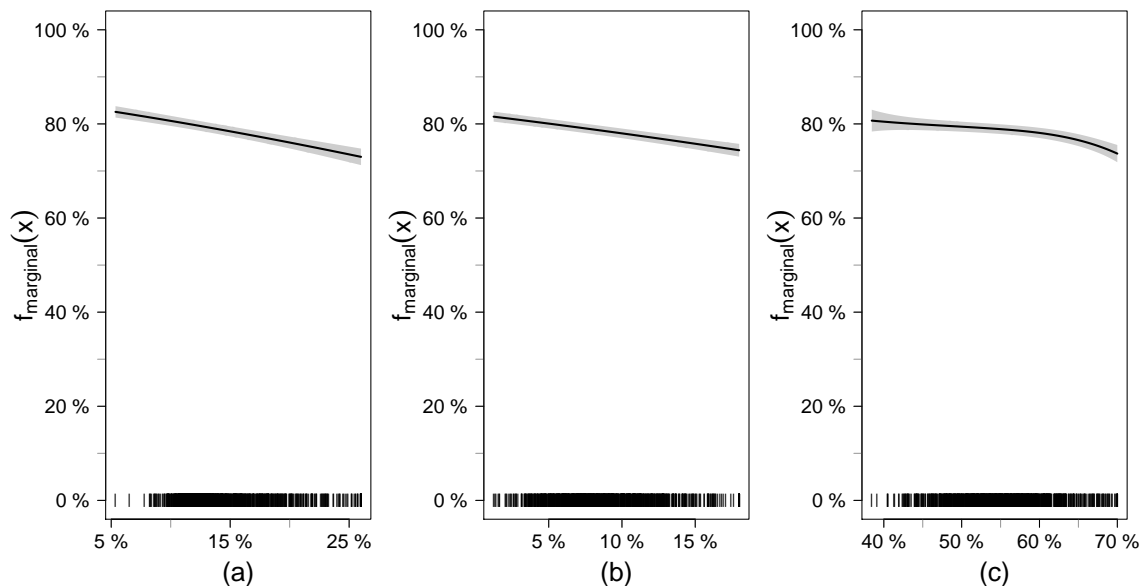


Abb. 3: Die drei stärksten Zusammenhänge zwischen der Wahlbeteiligung bei der Bundestagswahl 2017 und (a) dem Anteil der Deutschen mit Migrationshintergrund an der Hauptwohnsitzbevölkerung, (b) dem Anteil der Altersgruppe 75 Jahre und älter an der Hauptwohnsitzbevölkerung und (c) dem Anteil der Personen ohne Kirchenzugehörigkeit an der Hauptwohnsitzbevölkerung

CSU

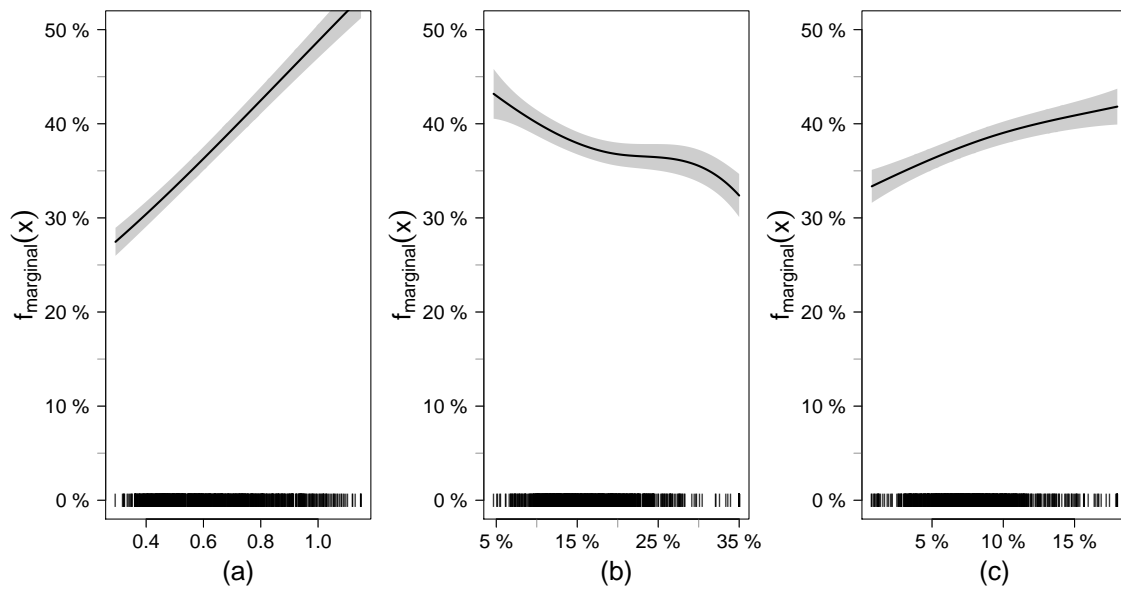


Abb. 4: Die drei stärksten Zusammenhänge zwischen den Wahlergebnissen der CSU bei der Bundestagswahl 2013 und (a) der Anzahl der zugelassenen Pkws pro Haushalt, (b) dem Anteil der Mehrpersonenhaushalte mit Kindern an den Haushalten und (c) dem Anteil der Altersgruppe 75 Jahre und älter an der Hauptwohnsitzbevölkerung

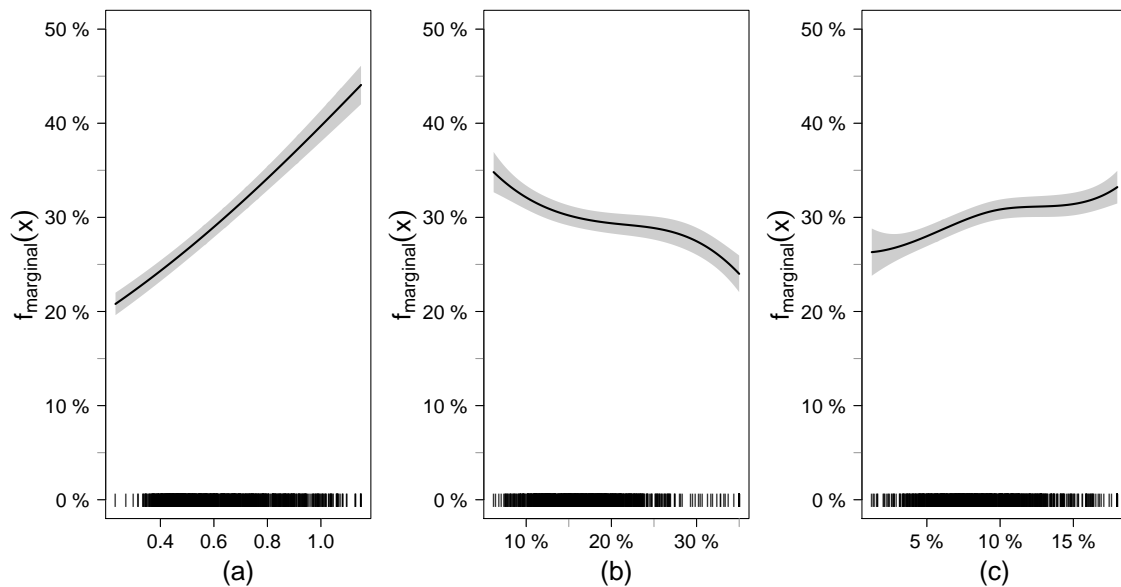


Abb. 5: Die drei stärksten Zusammenhänge zwischen den Wahlergebnissen der CSU bei der Bundestagswahl 2017 und (a) der Anzahl der zugelassenen Pkws pro Haushalt, (b) dem Anteil der Mehrpersonenhaushalte mit Kindern an den Haushalten und (c) dem Anteil der Altersgruppe 75 Jahre und älter an der Hauptwohnsitzbevölkerung

SPD

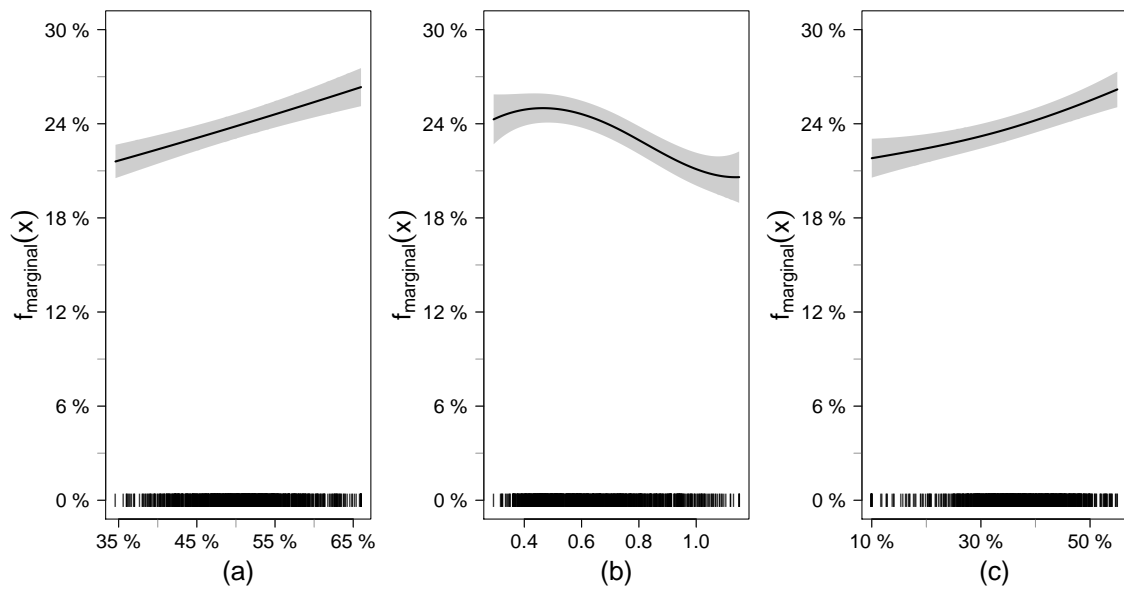


Abb. 6: Die drei stärksten Zusammenhänge zwischen den Wahlergebnissen der SPD bei der Bundestagswahl 2013 und (a) dem Anteil der Personen ohne Kirchenzugehörigkeit an der Hauptwohnsitzbevölkerung, (b) der Anzahl der zugelassenen Pkws pro Haushalt und (c) dem Anteil der Personen mit langer Wohndauer (10+ Jahre) an der Hauptwohnsitzbevölkerung

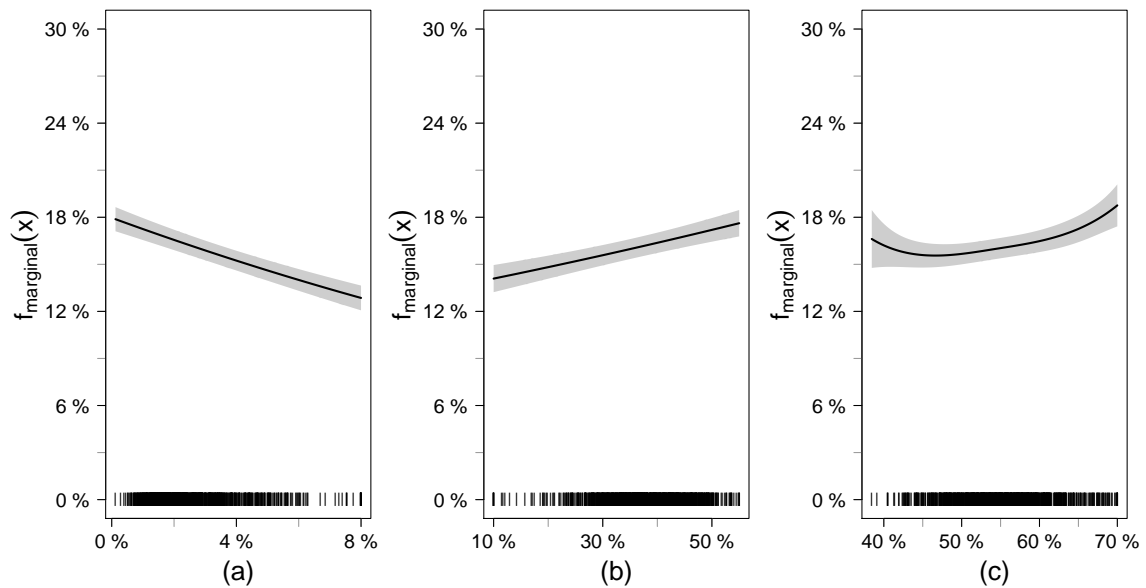


Abb. 7: Die drei stärksten Zusammenhänge zwischen den Wahlergebnissen der SPD bei der Bundestagswahl 2017 und (a) dem Anteil der gehobenen Pkws an den zugelassenen Pkws, (b) dem Anteil der Personen mit langer Wohndauer (10+ Jahre) an der Hauptwohnsitzbevölkerung und (c) dem Anteil der Personen ohne Kirchenzugehörigkeit an der Hauptwohnsitzbevölkerung

GRÜNE

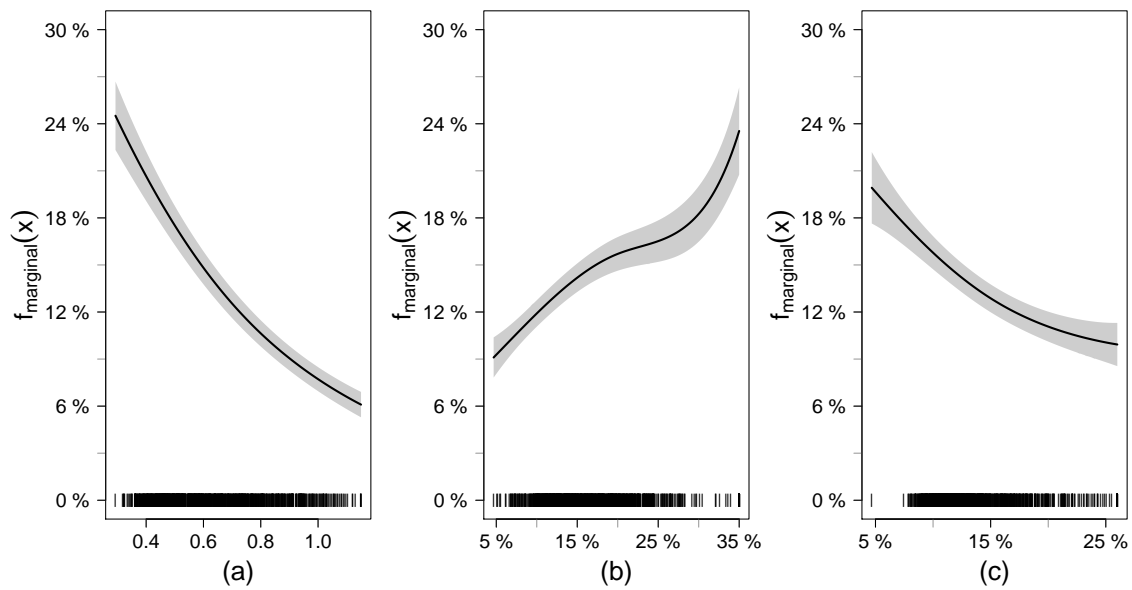


Abb. 8: Die drei stärksten Zusammenhänge zwischen den Wahlergebnissen der GRÜNEN bei der Bundestagswahl 2013 und (a) der Anzahl der zugelassenen Pkws pro Haushalt, (b) dem Anteil der Mehrpersonenhaushalte mit Kindern an den Haushalten und (c) dem Anteil der Deutschen mit Migrationshintergrund an der Hauptwohnsitzbevölkerung

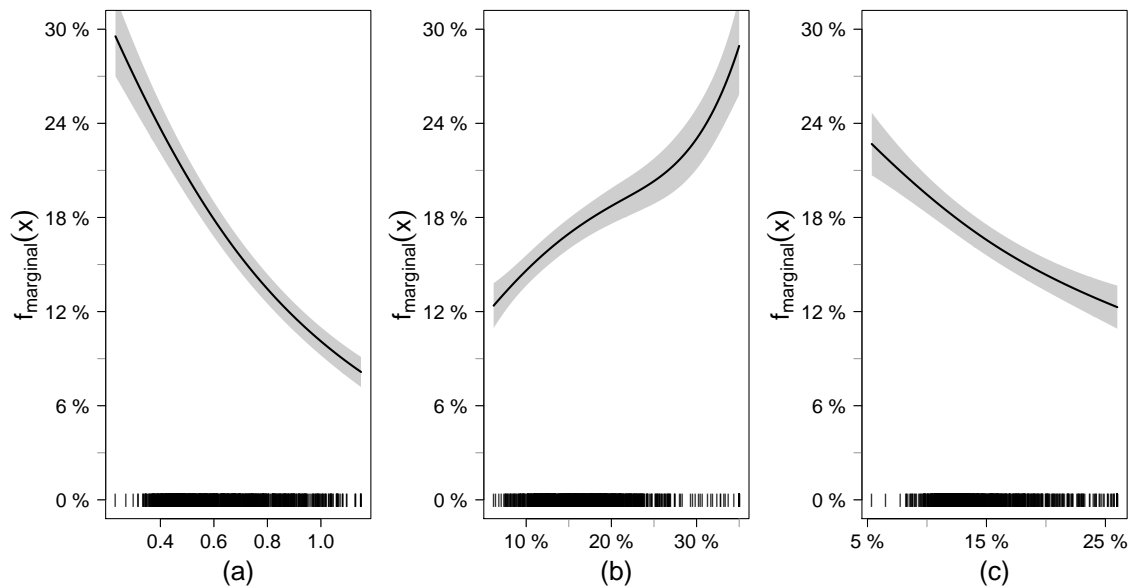


Abb. 9: Die drei stärksten Zusammenhänge zwischen den Wahlergebnissen der GRÜNEN bei der Bundestagswahl 2017 und (a) der Anzahl der zugelassenen Pkws pro Haushalt, (b) dem Anteil der Mehrpersonenhaushalte mit Kindern an den Haushalten und (c) dem Anteil der Deutschen mit Migrationshintergrund an der Hauptwohnsitzbevölkerung

FDP

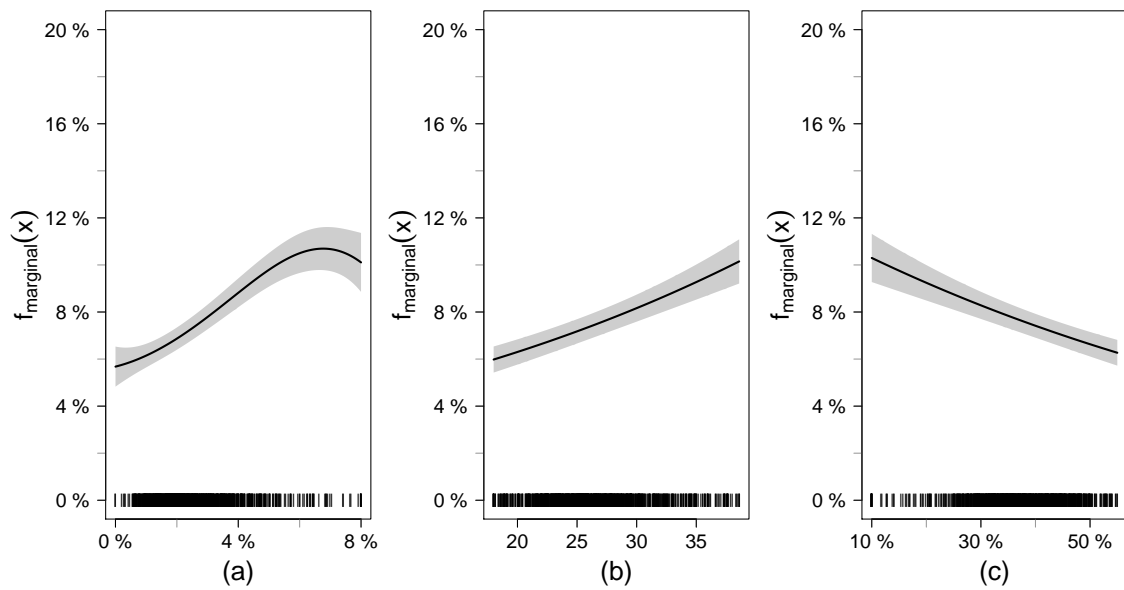


Abb. 10: Die drei stärksten Zusammenhänge zwischen den Wahlergebnissen der FDP bei der Bundestagswahl 2013 und (a) dem Anteil der gehobenen Pkws an den zugelassenen Pkws, (b) der durchschnittlichen Kaufkraft pro Person (in 1000€/Jahr) und (c) dem Anteil der Personen mit langer Wohndauer (10+ Jahre) an der Hauptwohnsitzbevölkerung

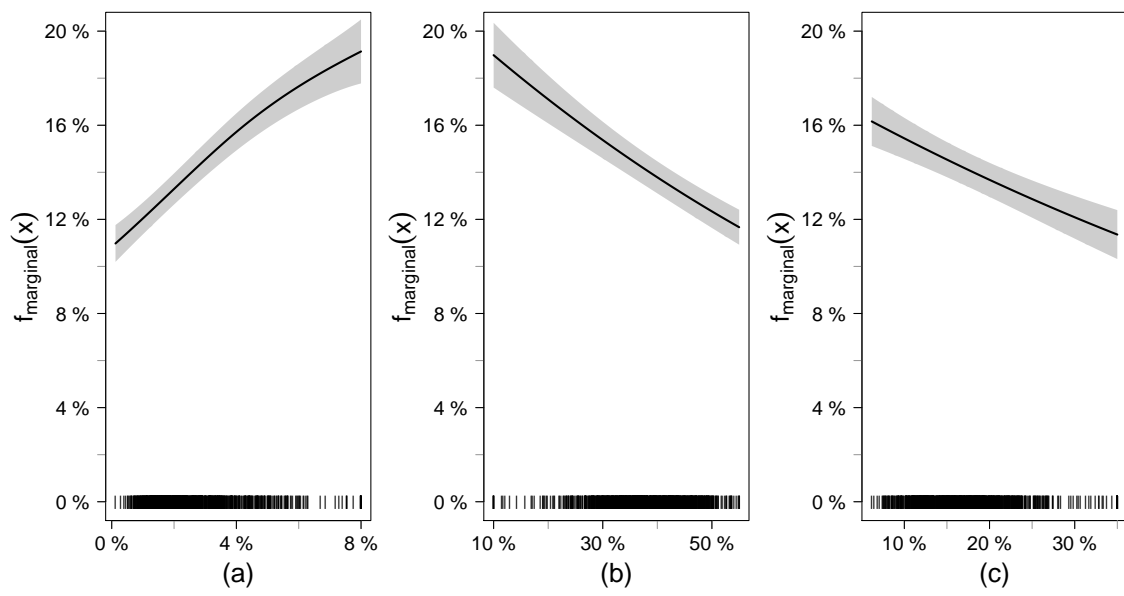


Abb. 11: Die drei stärksten Zusammenhänge zwischen den Wahlergebnissen der FDP bei der Bundestagswahl 2017 und (a) dem Anteil der gehobenen Pkws an den zugelassenen Pkws, (b) dem Anteil der Personen mit langer Wohndauer (10+ Jahre) an der Hauptwohnsitzbevölkerung und (c) dem Anteil der Mehrpersonenhaushalte mit Kindern an den Haushalten

AfD

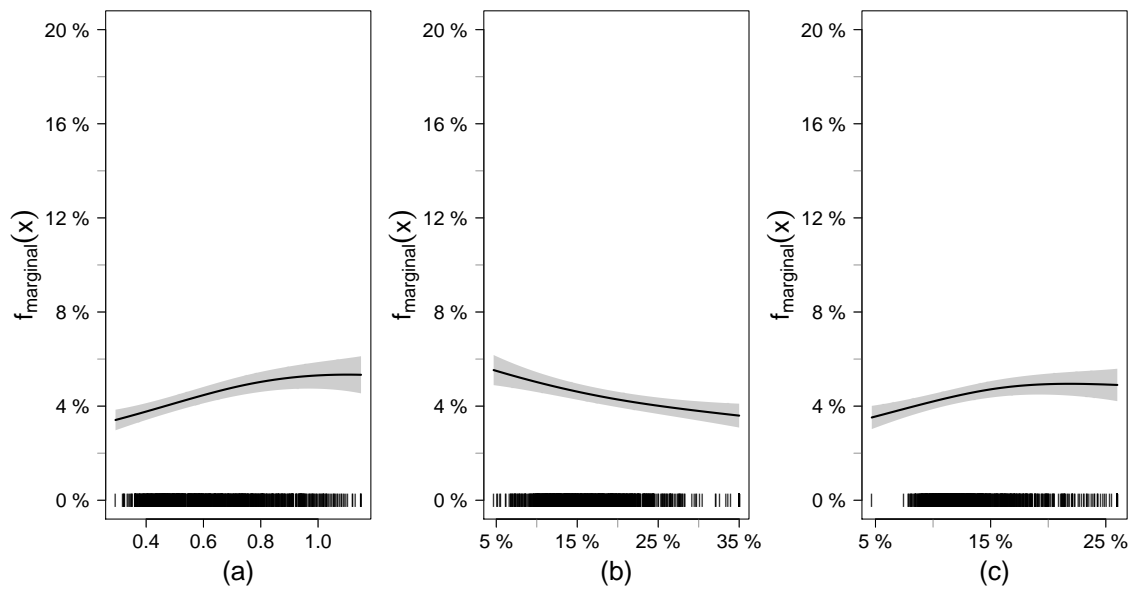


Abb. 12: Die drei stärksten Zusammenhänge zwischen den Wahlergebnissen der AfD bei der Bundestagswahl 2013 und (a) der Anzahl der zugelassenen Pkws pro Haushalt, (b) dem Anteil der Mehrpersonenhaushalte mit Kindern an den Haushalten und (c) dem Anteil der Deutschen mit Migrationshintergrund an der Hauptwohnsitzbevölkerung

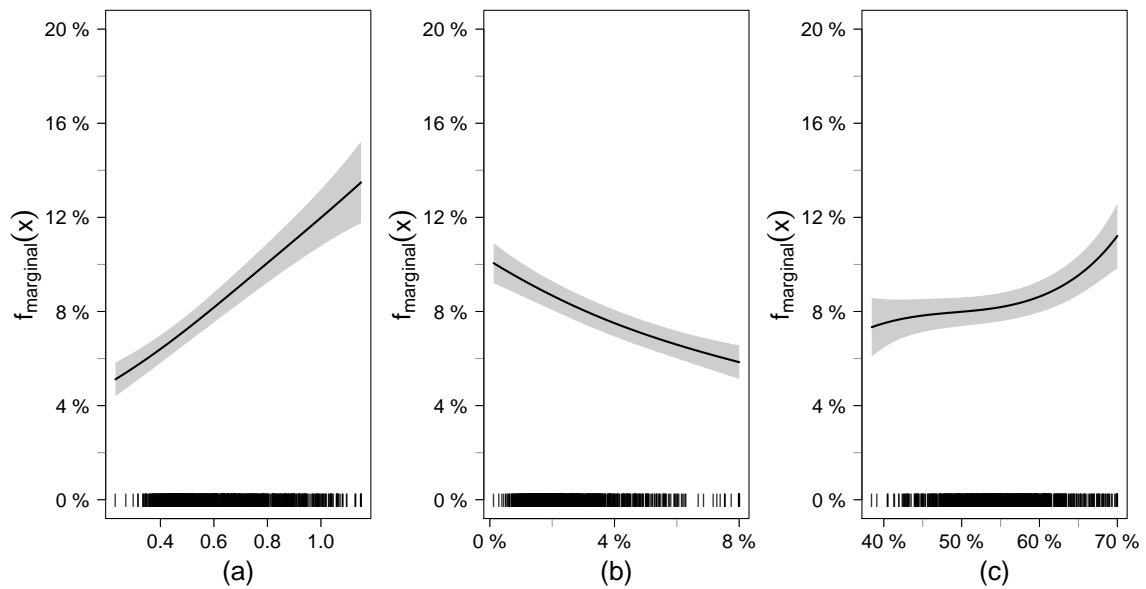


Abb. 13: Die drei stärksten Zusammenhänge zwischen den Wahlergebnissen der AfD bei der Bundestagswahl 2017 und (a) der Anzahl der zugelassenen Pkws pro Haushalt, (b) dem Anteil der gehobenen Pkws an den zugelassenen Pkws und (c) dem Anteil der Personen ohne Kirchenzugehörigkeit an der Hauptwohnsitzbevölkerung

DIE LINKE

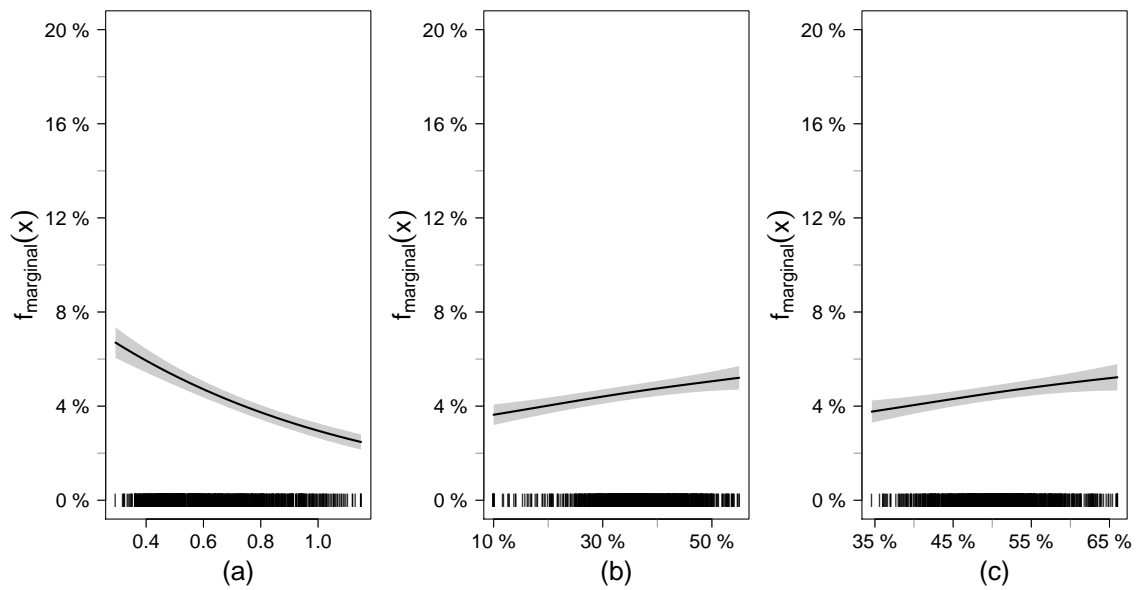


Abb. 14: Die drei stärksten Zusammenhänge zwischen den Wahlergebnissen der DIE LINKE bei der Bundestagswahl 2013 und (a) der Anzahl der zugelassenen Pkws pro Haushalt, (b) dem Anteil der Personen mit langer Wohndauer (10+ Jahre) an der Hauptwohnsitzbevölkerung und (c) dem Anteil der Personen ohne Kirchengliederung an der Hauptwohnsitzbevölkerung

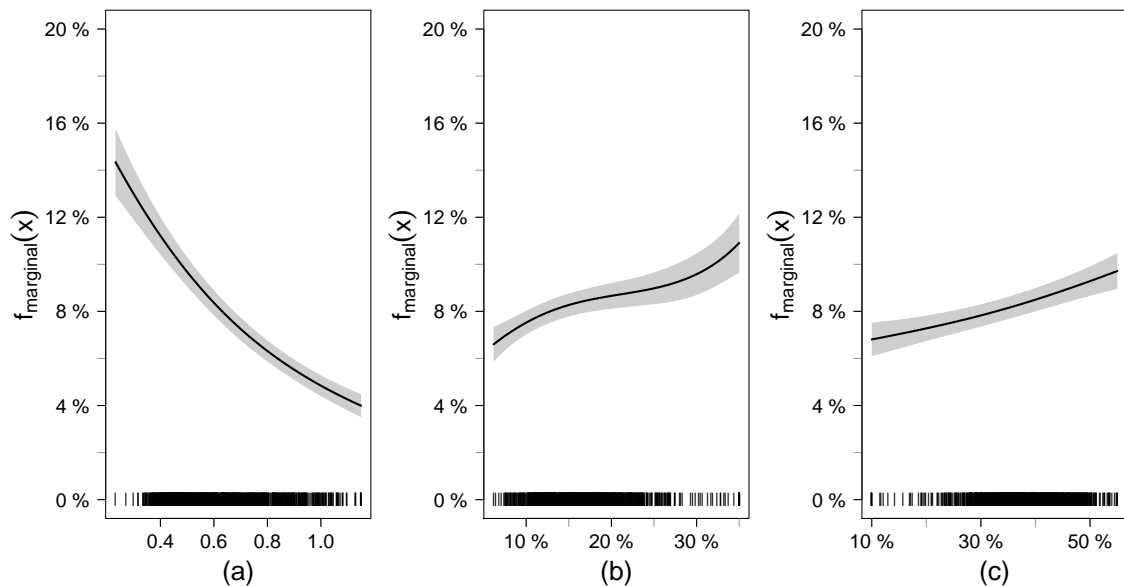


Abb. 15: Die drei stärksten Zusammenhänge zwischen den Wahlergebnissen der DIE LINKE bei der Bundestagswahl 2017 und (a) der Anzahl der zugelassenen Pkws pro Haushalt, (b) dem Anteil der Mehrpersonenhaushalte mit Kindern an den Haushalten und (c) dem Anteil der Personen mit langer Wohndauer (10+ Jahre) an der Hauptwohnsitzbevölkerung

Sonstige

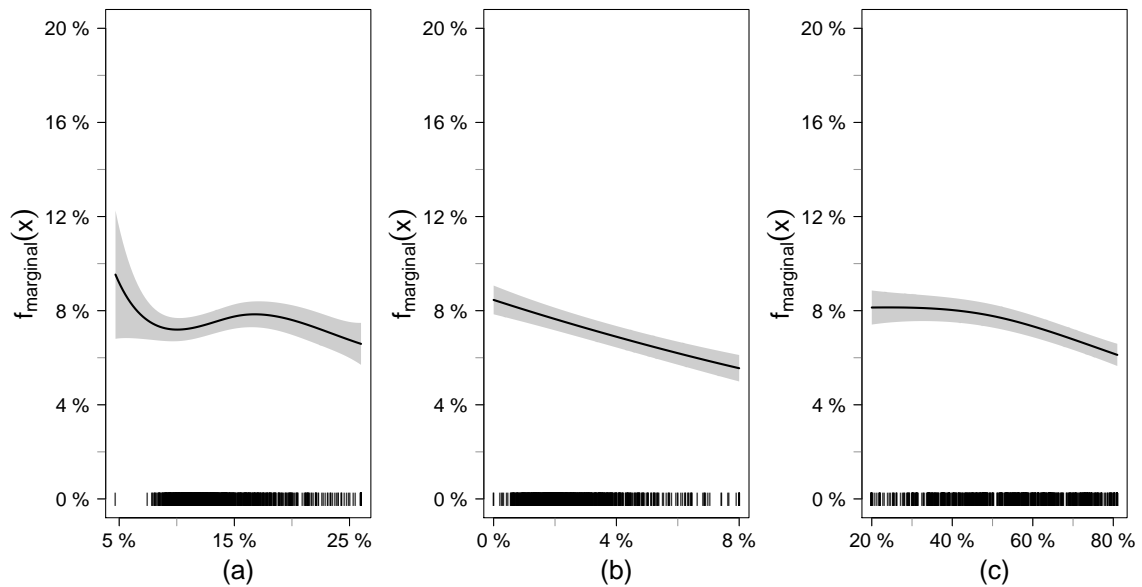


Abb. 16: Die drei stärksten Zusammenhänge zwischen den Wahlergebnissen der sonstigen Parteien bei der Bundestagswahl 2013 und (a) dem Anteil der Deutschen mit Migrationshintergrund an der Hauptwohnsitzbevölkerung, (b) dem Anteil der gehobenen Pkws an den zugelassenen Pkws und (c) dem Anteil der Personen mit Übertritt auf das Gymnasium an allen Schulübertritten

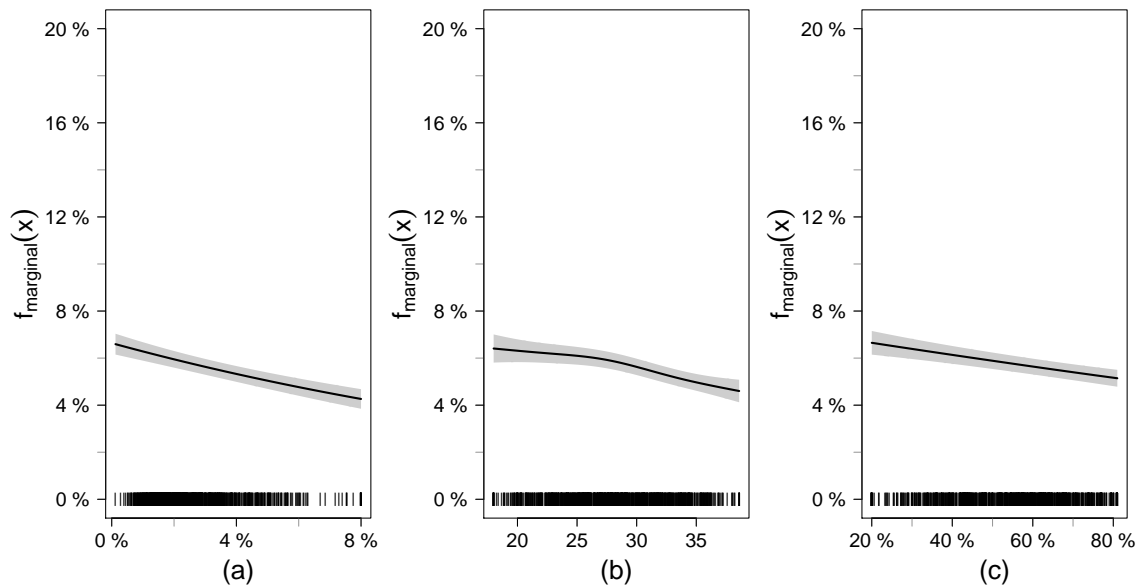


Abb. 17: Die drei stärksten Zusammenhänge zwischen den Wahlergebnissen der sonstigen Parteien bei der Bundestagswahl 2017 und (a) dem Anteil der gehobenen Pkws an den zugelassenen Pkws, (b) der durchschnittlichen Kaufkraft pro Person (in 1000€/Jahr) und (c) dem Anteil Personen mit Übertritt auf das Gymnasium an allen Schulübertritten

Gemeinsame Betrachtung der Modelle

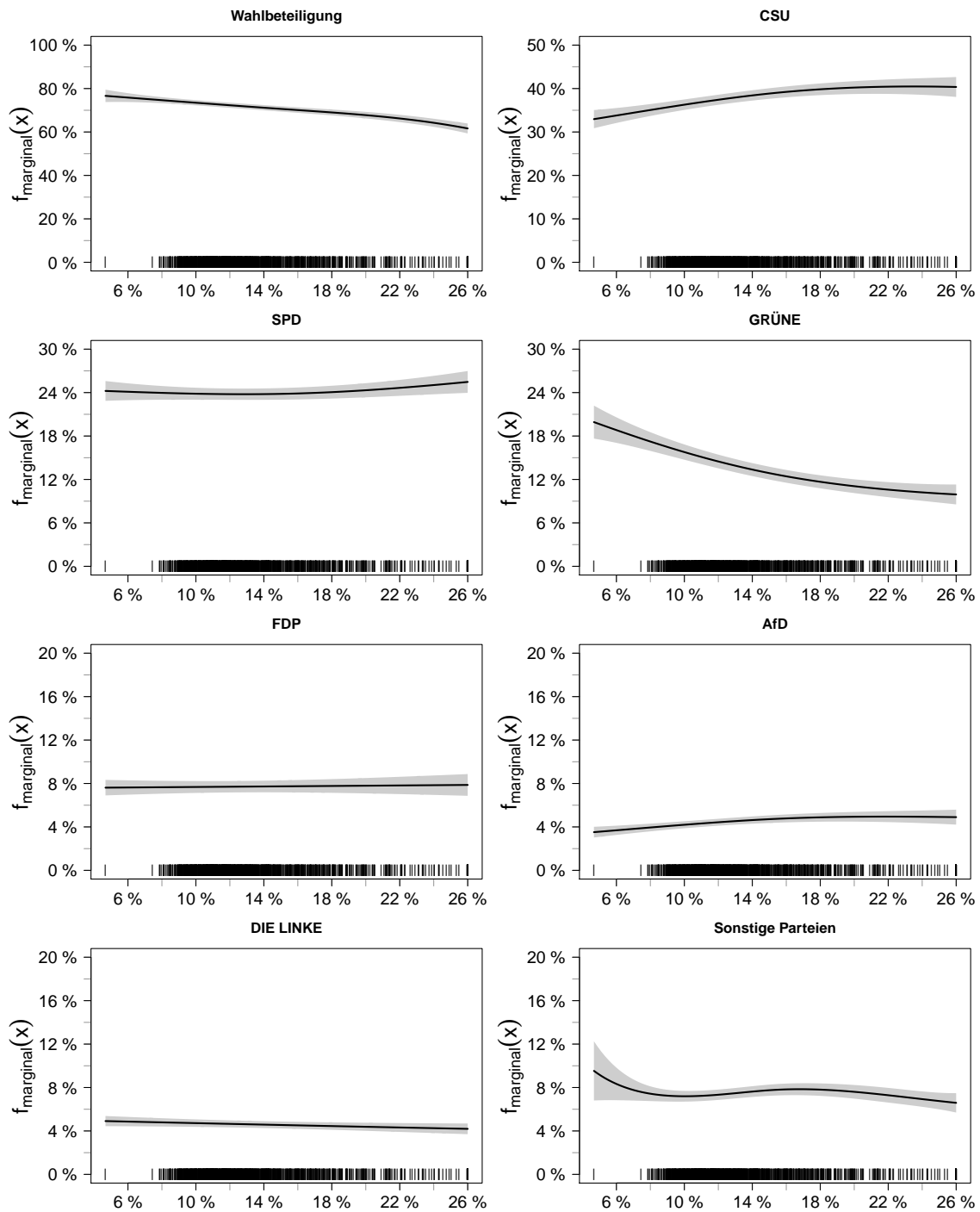


Abb. 18: Zusammenhang zwischen den Wahlergebnissen bei der Bundestagswahl 2013 und dem Anteil der Deutschen mit Migrationshintergrund an der Hauptwohnsitzbevölkerung in den Wahlbezirken. Über der Grafik ist zur Kennzeichnung des Modells jeweils die mit dem Anteil betrachtete Zielgröße genannt.

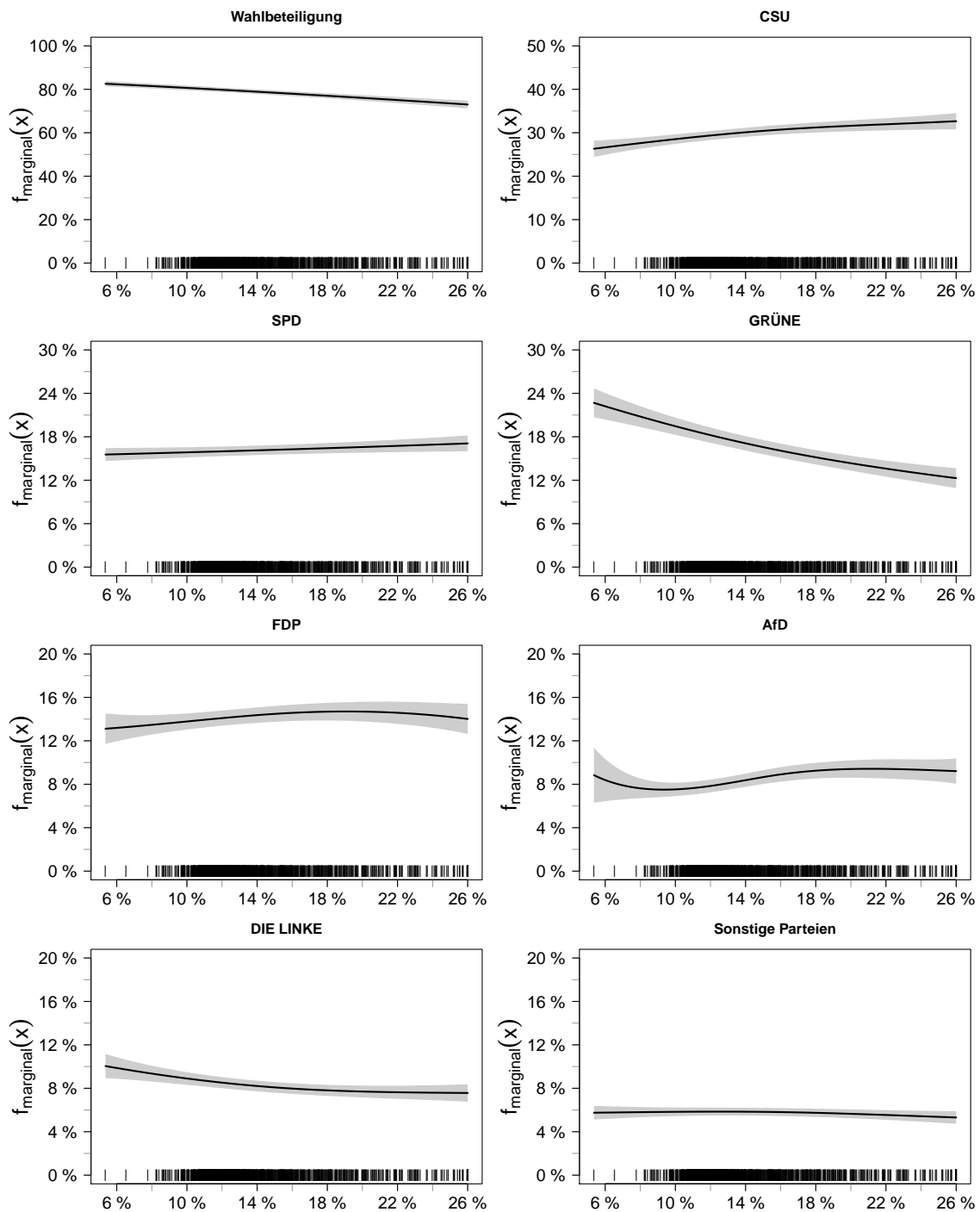


Abb. 19: Zusammenhang zwischen den Wahlergebnissen bei der Bundestagswahl 2017 und dem Anteil der Deutschen mit Migrationshintergrund an der Hauptwohnsitzbevölkerung in den Wahlbezirken. Über der Grafik ist zur Kennzeichnung des Modells jeweils die mit dem Anteil betrachtete Zielgröße genannt.

Literatur

Fahrmeir, L., Heumann, C., Künstler, R., Pigeot, I. und Tutz, G. (2016). *Statistik: Der Weg zur Datenanalyse*, Springer-Lehrbuch, Springer Berlin Heidelberg.

Fahrmeir, L., Kneib, T. und Lang, S. (2009). *Regression: Modelle, Methoden und Anwendungen*, Statistik und ihre Anwendungen, Springer-Verlag Berlin Heidelberg.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. Version: 3.4.0 **URL:** <https://www.R-project.org/>.

Wood, S. (2006). *Generalized Additive Models: An Introduction with R*, Chapman and Hall/CRC. mgcv-Package Version.