

Bundestagswahl 2017: Faktorenanalyse

Dimensionen der sozialen und ökonomischen Struktur und Wahlverhalten



Von Dr. Martin Jacobs

Wahlbeteiligung und Wahlentscheidung stehen mit einer Vielzahl sozialstruktureller und ökonomischer Merkmale in Verbindung. Teilweise beschreiben diese Merkmale ähnliche Sachverhalte und sind stark miteinander korreliert. Deshalb wurden diese Merkmale in einem ersten Schritt mittels einer Faktorenanalyse auf Ebene der kreisfreien Städte, verbandsfreien Gemeinden und Verbandsgemeinden zu wenigen Faktoren zusammengefasst. Diese Faktoren zeigen auf, hinsichtlich welcher Dimensionen sich die gesellschaftlichen und wirtschaftlichen Strukturen in Rheinland-Pfalz unterscheiden. Im zweiten Schritt wurden die ermittelten Faktoren als erklärende Variablen in multivariaten Regressionsmodellen eingesetzt. Diese Modelle analysieren, inwiefern sich die Unterschiede zwischen den kreisfreien Städten, verbandsfreien Gemeinden und Verbandsgemeinden bezüglich der Wahlbeteiligung und der Zweitstimmenanteile der Parteien bei der Bundestagswahl 2017 auf die ermittelten Faktoren zurückführen lassen.

Faktorenanalyse bündelt Vielzahl sozialstruktureller und ökonomischer Merkmale

Viele Merkmale stehen in Verbindung mit dem Wahlverhalten

Das Wahlverhalten der Bürgerinnen und Bürger steht mit einer Vielzahl gesellschaftlicher und wirtschaftlicher Merkmale im Zusammenhang. Für die Bundestagswahl am 24. September 2017 hat das Statistische Landesamt hierzu bereits einige Analysen vorgelegt. Die gesetzlich angeordnete Repräsentative Wahlstatistik ermöglicht durch die Auswertung von Wählerverzeichnissen und die Ausgabe markierter Stimmzettel die Untersuchung der Wahlbeteiligung und der Wahlentscheidungen nach Alter und Geschlecht. Die Aggregatdatenanalyse im Rahmen der Analyse der Wahlergebnisse in der Wahlnacht zeigte mittels einer Korrelationsanalyse statistische Zusammenhänge zwischen einer Vielzahl

sozialstruktureller und ökonomischer Merkmale in den rheinland-pfälzischen kreisfreien Städten, verbandsfreien Gemeinden und Verbandsgemeinden einerseits und der Wahlbeteiligung sowie den Zweitstimmenergebnissen der Parteien andererseits. Ein Beitrag in den Statistischen Monatsheften erweiterte diese Betrachtung um multivariate Regressionen.

Ein Modell mit einer großen Anzahl von Merkmalen als erklärende Variablen besitzt zum einen den Nachteil der Unübersichtlichkeit. Zum anderen können methodische Probleme wie Multikollinearität auftreten, wenn mehrere erklärende Variablen im Modell einen ähnlichen oder gar identischen Sachverhalt beschreiben und nicht klar ist, welcher dieser Variablen die Veränderung der zu erklärenden Größe zuzuordnen ist.

Unübersichtlichkeit und mögliche methodische Probleme bei Modellen mit vielen erklärenden Variablen

Faktorenanalyse
bietet Überblick

Mit dem Verfahren der Faktorenanalyse lassen sich viele Merkmale zu wenigen Faktoren bündeln, indem miteinander zusammenhängende Merkmale zu einem Faktor zusammengefasst werden. Die ermittelten Faktoren ermöglichen einen Überblick, in welchen Dimensionen sich die kreisfreien Städte, verbandsfreien Gemeinden und Verbandsgemeinden unterscheiden.

Die bei der Faktorenanalyse gewonnenen Faktoren, die nicht miteinander korrelieren, gehen im zweiten Schritt in übersichtlichere Regressionsmodelle ein, die die Zusammenhänge der Faktoren mit der Wahlbeteiligung sowie mit den Zweitstimmenergebnissen der sechs in Rheinland-Pfalz angetretenen Parteien, die in den 19. Deutschen Bundestag einziehen konnten, beleuchten.

28 Merkmale werden zu sieben Faktoren verdichtet

Merkmale aus
den Bereichen
Demografie,
Soziales, Bil-
dung und Er-
werbstätigkeit

Die Basis für die Faktorenanalyse bilden 28 sozialstrukturelle und ökonomische Merkmale aus den Bereichen Demografie, Soziales, Bildung und Erwerbstätigkeit (vgl. Tabelle T 1).

Daten stam-
men zu großen
Teilen aus dem
Zensus 2011

Es wurde ein möglichst aktueller Datenstand verwendet. Ein Großteil der Merkmale wurde jedoch zuletzt beim Zensus 2011 in hinreichender regionaler Gliederungstiefe erhoben. Des Weiteren wurde ein Teil der Merkmale im Rahmen der Zensus-Stichprobe erfasst, in die ausschließlich Gemeinden mit mindestens 10 000 Einwohnerinnen und Einwohnern zum Stichtag 31. Dezember 2009 einbezogen waren. In die Analyse können deshalb nur Gebietseinheiten eingehen, die diese Mindestgröße erreichten und zwischen 2011 und der Bundestagswahl 2017 von keiner Gebietsreform betroffen waren. Diese Kriterien erfüllten 139 rheinland-pfälzische Verwaltungsbezirke: die zwölf kreisfreien Städte,

26 von 30 verbandsfreien Gemeinden und 101 von 146 Verbandsgemeinden.

Auf Grundlage der Ausprägungen der 28 Merkmale in den 139 Gebietseinheiten wurde eine Faktorenanalyse (genauer: eine Hauptkomponentenanalyse) durchgeführt. Hierzu wurden die Merkmalsausprägungen standardisiert. Die Standardisierung bewirkt, dass für jedes Merkmal der Mittelwert über alle in die Analyse einbezogenen Gebietseinheiten null und die Standardabweichung eins beträgt. Eine Faktorenanalyse folgt keinem fixen Algorithmus, sondern erfordert eine Reihe subjektiver Entscheidungen, insbesondere bei der Wahl der Anzahl zu extrahierender Faktoren und bei der Faktorinterpretation.¹

Eine gängige Entscheidungsregel für die Bestimmung der Zahl der Faktoren ist das Kaiser-Kriterium. Es besagt, dass diejenigen Faktoren extrahiert werden sollen, deren Eigenwert größer als eins ist. Ein Eigenwert von größer als eins ist gleichbedeutend damit, dass der Faktor für die in den Daten enthaltene Varianz einen Erklärungsgehalt aufweist, der größer ist als der Erklärungsgehalt einer einzelnen standardisierten Variablen. Nach dem Kaiser-Kriterium wären bei den vorliegenden Daten acht Faktoren zu extrahieren. Die weitere Betrachtung zeigte jedoch, dass sich von diesen acht Faktoren die sieben Faktoren mit dem größten Erklärungsgehalt relativ einfach interpretieren lassen, wohingegen sich der achte Faktor einer einfachen Interpretation verschließt. Da zudem der Erklärungsgehalt des achten Faktors mit einem Eigenwert von 1,005 den einer einzelnen Variablen nur marginal übersteigt, wurde entschieden, lediglich sieben

Faktorenanalyse erfordert subjektive Entscheidungen

Kaiser-Kriterium gibt Orientierung bei Wahl der Faktorenzahl

¹ Eine ausführliche Beschreibung des Verfahrens findet sich in Backhaus, K./Erichson, B./Plinko, W./Weiber, R.: Multivariate Analysemethoden. 9. Auflage, Berlin u. a. 2000, S. 252ff.

T 1

Übersicht der in die Faktorenanalyse eingehenden Merkmale

Merkmal Kurzbezeichnung	Beschreibung	Einheit	Zeitpunkt	Datenquelle
Bevölkerung	Bevölkerungsstand absolut	Anzahl	31.12.2016	Bevölkerungsfortschreibung
Bevölkerungsdichte	Bevölkerungsstand je Gebiet	Anzahl je Quadrat-kilometer	31.12.2016	Bevölkerungsfortschreibung, Gebietsstand
Bevölkerungsentwicklung	Veränderung des Bevölkerungsstandes	Prozent	31.12.2006 / 31.12.2016	Bevölkerungsfortschreibung
Unter 18-Jährige	Anteil der unter 18-Jährigen an der Bevölkerung	Prozent	31.12.2016	Bevölkerungsfortschreibung
65-Jährige und Ältere	Anteil der 65-Jährigen und Älteren an der Bevölkerung	Prozent	31.12.2016	Bevölkerungsfortschreibung
Ausländer/-innen	Anteil der Ausländer/-innen an der Bevölkerung	Prozent	31.12.2016	Bevölkerungsfortschreibung
Arbeitslose	Anteil der Arbeitslosen an der 15- bis 64-jährigen Bevölkerung	Prozent	Durchschnitt 2016	Bundesagentur für Arbeit, Bevölkerungsfortschreibung
SGB II-Berechtigte	Anteil der Regelleistungsberechtigten in Bedarfsgemeinschaften nach dem Zweiten Sozialgesetzbuch an der unter 65-jährigen Bevölkerung	Prozent	Dezember 2015	Bundesagentur für Arbeit (geheime Werte imputiert), Bevölkerungsfortschreibung
Single-Haushalte	Anteil der Ein-Personen-Haushalte an allen Haushalten	Prozent	09.05.2011	Zensus
Haushalte ab vier Personen	Anteil der Haushalte mit vier und mehr Personen an allen Haushalten	Prozent	09.05.2011	Zensus
Katholiken	Anteil der Katholiken an der Bevölkerung	Prozent	09.05.2011	Zensus
Evangelische	Anteil der Evangelischen an der Bevölkerung	Prozent	09.05.2011	Zensus
Kein Schulabschluss	Anteil der Personen ohne Schulabschluss an der Bevölkerung ¹	Prozent	09.05.2011	Zensus-Stichprobe
Hauptschulabschluss	Anteil der Personen mit Haupt- bzw. Volksschulabschluss als höchsten Schulabschluss an der Bevölkerung ¹	Prozent	09.05.2011	Zensus-Stichprobe
Mittlere Reife	Anteil der Personen mit Mittlerer Reife als höchsten Schulabschluss an der Bevölkerung ¹	Prozent	09.05.2011	Zensus-Stichprobe
Hochschulreife	Anteil der Personen mit (Fach-) Hochschulreife an der Bevölkerung ¹	Prozent	09.05.2011	Zensus-Stichprobe
Keine Berufsausbildung	Anteil der Personen ohne Abschluss einer beruflichen Ausbildung an der Bevölkerung ¹	Prozent	09.05.2011	Zensus-Stichprobe
Berufsausbildung	Anteil der Personen mit Abschluss einer beruflichen Ausbildung als höchsten beruflichen Abschluss an der Bevölkerung ¹	Prozent	09.05.2011	Zensus-Stichprobe
Hochschulabschluss	Anteil der Personen mit Hochschulabschluss an der Bevölkerung ¹	Prozent	09.05.2011	Zensus-Stichprobe
Landwirtschaft	Anteil der Erwerbstätigen in der Land- und Fortstwirtschaft sowie in der Fischerei an allen Erwerbstätigen (Wohnort)	Prozent	09.05.2011	Zensus-Stichprobe (geheime Werte imputiert)
Produzierendes Gewerbe	Anteil der Erwerbstätigen im Produzierenden Gewerbe an allen Erwerbstätigen (Wohnort)	Prozent	09.05.2011	Zensus-Stichprobe
Dienstleistungen	Anteil der Erwerbstätigen in den Dienstleistungsbereichen an allen Erwerbstätigen (Wohnort)	Prozent	09.05.2011	Zensus-Stichprobe
Abhängig Beschäftigte	Anteil der Arbeiter/-innen sowie der Angestellten an allen Erwerbstätigen (Wohnort)	Prozent	09.05.2011	Zensus-Stichprobe
Beamte/-innen	Anteil der Beamten/-innen an allen Erwerbstätigen (Wohnort)	Prozent	09.05.2011	Zensus-Stichprobe
Selbstständige	Anteil der Selbstständigen an allen Erwerbstätigen (Wohnort)	Prozent	09.05.2011	Zensus-Stichprobe
Pendlersaldo	Anzahl der Einpendlerinnen und Einpendler (Arbeitsort) minus Anzahl der Auspendlerinnen und Auspendler (Wohnort), je 100 Einwohnerinnen und Einwohner im Alter von 15 bis 64 Jahren	Anzahl	09.05.2011	Zensus-Stichprobe, Zensus
Unternehmen mit weniger als zehn Beschäftigten	Anteil der sozialversicherungspflichtig Beschäftigten in Unternehmen mit weniger als zehn sozialversicherungspflichtig Beschäftigten an allen sozialversicherungspflichtig Beschäftigten (Ort des Unternehmenssitzes)	Prozent	Durchschnitt 2016	Unternehmensregister
Unternehmen mit 250 und mehr Beschäftigten	Anteil der sozialversicherungspflichtig Beschäftigten in Unternehmen mit 250 und mehr sozialversicherungspflichtig Beschäftigten an allen sozialversicherungspflichtig Beschäftigten (Ort des Unternehmenssitzes)	Prozent	Durchschnitt 2016	Unternehmensregister

1 ab 15 Jahren.

Faktoren zu extrahieren. Die extrahierten Faktoren erklären gemeinsam rund 83 Prozent der Varianz in den Ausgangsdaten.²

Geometrische Darstellung von Merkmalen und Faktoren

In einer geometrischen Betrachtung stellen sich die sieben extrahierten Faktoren als zueinander im rechten Winkel stehende Vektoren bzw. als Koordinatenkreuz in einem siebendimensionalen Vektorraum dar. Die Korrelation zwischen den Faktoren und den zugrunde liegenden Merkmalen wird durch die sogenannten Faktorladungen beschrieben. Die Faktorladungen eines Merkmals auf die sieben Faktoren definieren dieses Merkmal ebenfalls als Vektor im Vektorraum. Die Ladung eines Merkmals auf einen Faktor entspricht dem Kosinus des Winkels zwischen den beiden das Merkmal und den Faktor beschreibenden Vektoren. Korrelieren ein Merkmal und ein Faktor perfekt miteinander, so liegen die beiden Vektoren auf einer gemeinsamen Linie; der Winkel beträgt 0 bzw. 180 Grad. Die Faktorladung nimmt dann einen betragsmäßigen Wert von eins an. Sind das Merkmal und der Faktor hingegen völlig unkorreliert, so beträgt die Faktorladung null und die beiden Vektoren stehen im rechten Winkel zueinander.

Rotation zur leichteren Faktorinterpretation

Um die Interpretation der einzelnen Faktoren zu erleichtern, kann das Koordinatenkreuz in seinem Ursprung rotiert werden, sodass die Faktoren nach der Rotation möglichst hohe Korrelationen mit einzelnen Merkmalen oder Merkmalsgruppen aufweisen. Die Verwendung einer orthogonalen bzw. rechtwinkligen Rotation stellt dabei sicher, dass die Faktoren bei der Rotation miteinander unkorreliert bleiben.

² Nach Faktorrotation betragen die Anteile der einzelnen Faktoren 20,9 Prozent, 17,5 Prozent, 10,2 Prozent, 9,7 Prozent, 8,9 Prozent, 8,2 Prozent und 7,7 Prozent.

Interpretation der Faktoren

Die Ladungsmatrix in Tabelle T 2 zeigt die Ladungen der einzelnen Merkmale auf die Faktoren nach der Rotation. Die Faktoren sind absteigend nach ihrem Erklärungsanteil an der Gesamtvarianz geordnet. Aus Gründen der Übersichtlichkeit werden in der Matrix nur Ladungen mit einem Betrag von mindestens 0,3 angegeben. Die Vorzeichen der Ladungen eines Faktors sind im Grunde austauschbar. Durch das Vertauschen der Vorzeichen würde der Vektor, der den Faktor beschreibt, um 180 Grad gedreht und so weiterhin auf derselben Linie im Vektorraum liegen wie zuvor.

Beim ersten Faktor zeigen zu einem diejenigen Merkmale hohe Ladungen, die städtische und ländliche Gebiete voneinander abgrenzen. Hierzu zählen die Bevölkerungszahl, die Bevölkerungsdichte, der Bevölkerungsanteil der Ausländerinnen und Ausländer sowie die beiden Variablen zur Haushaltsgröße. Zum anderen erreichen hier aber solche Variablen noch höhere Ladungen, die akademische Umfeldler von Milieus mit mittleren Bildungs- und Berufsabschlüssen trennen. Insgesamt kennzeichnet der erste Faktor die strukturellen Gegensätze zwischen städtischen Räumen mit höherem Bevölkerungswachstum und einem überdurchschnittlichen Bevölkerungsanteil höher Gebildeter auf der einen Seite und ländlichen Bereichen mit erhöhten Anteilen von Personen mit Hauptschulabschluss, Mittlerer Reife und beruflicher Ausbildung auf der anderen Seite.

Faktor I: Siedlungsstruktur und Anteil höher Gebildeter

Variablen, die die Unterschiede zwischen Stadt und Land beschreiben, weisen auch beim zweiten Faktor hohe Ladungen auf. Die höchsten Ladungen auf den zweiten Faktor finden sich allerdings bei Merkmalen, die besonders auf sozial schwache und bildungsferne Umfeldler hinweisen, wie der Arbeitslosenanteil, der Anteil der Leistungsberechtigten nach dem Zweiten

Faktor II: Siedlungsstruktur und Anteil sozial Schwacher

T 2 Matrix der Faktorladungen

Merkmal Kurzbezeichnung	Faktorladungen						
	Faktor I	Faktor II	Faktor III	Faktor IV	Faktor V	Faktor VI	Faktor VII
Bevölkerung	0,636	0,306					
Bevölkerungsdichte	0,699	0,474					
Bevölkerungsentwicklung	0,706			0,387			
Unter 18-Jährige				0,900			
65-Jährige und Ältere				-0,843			
Ausländer/-innen	0,558	0,663					
Arbeitslose		0,784				-0,300	
SGB II-Berechtigte		0,861					
Single-Haushalte	0,465	0,625		-0,395	0,320		
Haushalte ab vier Personen	-0,369	-0,499		0,619			
Katholiken							-0,945
Evangelische							0,954
Kein Schulabschluss		0,853		0,302			
Hauptschulabschluss	-0,801					-0,364	
Mittlere Reife	-0,413			0,335		0,701	
Hochschulreife	0,895		-0,320				
Keine Berufsausbildung		0,594				-0,552	
Berufsausbildung	-0,748	-0,400					
Hochschulabschluss	0,871						
Landwirtschaft		-0,426	0,320		-0,338	0,346	
Produzierendes Gewerbe	-0,532		0,761				
Dienstleistungen	0,466		-0,819				
Abhängig Beschäftigte		0,337	0,516			-0,585	
Beamte/-innen			-0,836				
Selbstständige						0,719	
Pendlersaldo		0,560			0,597		
Unternehmen mit weniger als zehn Beschäftigten						-0,887	
Unternehmen mit 250 und mehr Beschäftigten					0,875		

Sozialgesetzbuch (sogenannte „Hartz IV“-Leistungen) sowie die Bevölkerungsanteile ohne Schulabschluss und ohne Abschluss einer beruflichen Ausbildung. Der zweite Faktor stellt mithin auf die Verschiedenheit von dichter besiedelten Gebieten mit erhöhten Anteilen sozial schwacher und bildungsferner Personen gegenüber dünner besiedelten Räumen mit einer unterdurchschnittlichen Ausprägung problematischer Strukturen ab.

Auf den dritten Faktor laden besonders Merkmale, die die Wirtschaftsbereiche, in denen Erwerbstätige arbeiten, sowie die Art der

Erwerbstätigkeit betreffen. Auf diese Weise werden Verwaltungsbezirke, in denen ein überdurchschnittlicher Teil der Erwerbstätigen im Produzierenden Gewerbe arbeitet und abhängig beschäftigt ist, von Gebieten mit höheren Anteilen von Beamtinnen und Beamten sowie von Erwerbstätigen in den Dienstleistungsreichen abgegrenzt. Der Anteil der Selbstständigen an den Erwerbstätigen erreicht hingegen keine hohe Faktorladung und ist somit nicht Bestandteil dieser Unterscheidung.

Der vierte Faktor bündelt Merkmale zur Altersstruktur der Bevölkerung sowie zur

Faktor III:
Wirtschafts-
struktur
und Art der
Erwerbstätigkeit

Faktor IV:
Alters- und
Familienstruktur

Haushaltsgröße. Er kennzeichnet die Unterschiede zwischen Gebieten mit einem erhöhten Anteil Minderjähriger und einem überdurchschnittlichen Anteil größerer Haushalte, d. h. Gebiete mit relativ vielen Familien, einerseits und Räumen mit einem erhöhten Bevölkerungsanteil der Seniorinnen und Senioren sowie relativ vielen Single-Haushalten andererseits.

Faktor V:
Unternehmens-
größenstruktur

Hohe Ladungen auf den fünften Faktor weisen vor allem die beiden Variablen zur Unternehmensgröße und des Weiteren auch der Pendlersaldo auf. Der fünfte Faktor deutet auf die Differenzen zwischen Gebietseinheiten, die Sitz großer Unternehmen sind und Einpendlerüberschüsse verzeichnen, und Verwaltungsbezirken mit einer kleinteiligen Unternehmensstruktur und Auspendlerüberschüssen.

Faktor VI:
Selbst-
ständigkeit

Die höchste Ladung mit positivem Vorzeichen auf den sechsten Faktor zeigt der Anteil der Selbstständigen an den Erwerbstätigen, während der Anteil der abhängig Beschäftigten die größte negative Faktorladung aufweist. Gleichzeitig finden sich aber auch bei einigen Merkmalen zu Bildung und Berufsqualifikation hohe Ladungen. So lädt der Bevölkerungsanteil der Personen mit Mittlerer Reife positiv und der Anteil der Personen mit Hauptschulabschluss sowie der Anteil ohne abgeschlossene Berufsausbildung negativ auf den Faktor.

Faktor VII:
Konfessions-
struktur

Der siebte Faktor ist aufgrund seiner Ladungsstruktur sehr einfach zu interpretieren. Lediglich die beiden Variablen zur Konfessionszugehörigkeit verfügen über sehr hohe Ladungen auf den siebten Faktor. Dieser Faktor beschreibt also den Gegensatz zwischen evangelisch und katholisch geprägten Regionen.

Faktorwerte zeigen Positionierung der Gebietseinheiten

Für die 139 in die Faktorenanalyse einbezogenen Gebietseinheiten lassen sich sogenannte Faktorwerte angeben. Der Faktorwert misst die Ausprägung einer Einheit hinsichtlich der Dimension, die der Faktor beschreibt. Mittels ihrer sieben Faktorwerte für die sieben Faktoren kann jede Gebietseinheit als Punkt im Vektorraum dargestellt werden. Wie die der Faktorenanalyse zugrunde gelegten Merkmale sind auch die Faktorwerte standardisiert. Das bedeutet, für jeden Faktor beträgt der Durchschnitt der Faktorwerte über alle Gebietseinheiten null und die Standardabweichung eins.

Tabelle T 3 zeigt, welche der untersuchten Verwaltungsbezirke bei den sieben einzelnen Faktoren die höchsten bzw. niedrigsten Faktorwerte aufweisen. Außerdem lässt sich anhand der Faktorwerte für alle sieben Faktoren bestimmen, wie nahe die sozialstrukturellen und ökonomischen Verhältnisse in den einzelnen Gebietseinheiten dem rheinland-pfälzischen Durchschnitt (d. h. dem Durchschnitt aller Verwaltungsbezirke, die in die Analyse einbezogen werden konnten) insgesamt kommen. Da durchschnittliche Ausprägungen durch einen Faktorwert von null gekennzeichnet sind, misst der Abstand vom Ursprung zum Punkt im Vektorraum, der eine Gebietseinheit repräsentiert, die Abweichung vom Durchschnitt. Dieser Abstand kann aufgrund der Unabhängigkeit der Faktoren voneinander mit dem Satz des Pythagoras ermittelt werden. Die Verwaltungsbezirke, die in ihrer gesellschaftlichen und wirtschaftlichen Struktur am stärksten und am geringsten vom Landesdurchschnitt abweichen, sind ebenfalls der Tabelle T 3 zu entnehmen.

Gebietseinheiten als Punkte im Vektorraum

Abstand zum Ursprung misst Abweichung vom Durchschnitt

T 3

Verwaltungsbezirke mit den höchsten und niedrigsten Faktorwerten sowie mit den höchsten und niedrigsten Gesamtabweichungen vom Durchschnitt

Faktor	Höchste Faktorwerte		Niedrigste Faktorwerte	
	Verwaltungsbezirk	Faktorwert	Verwaltungsbezirk	Faktorwert
Faktor I Siedlungsstruktur und Anteil höher Gebildeter	Mainz, kfr. St.	5,55	VG Altenglan	-1,82
	Trier, kfr. St.	3,58	VG Hamm (Sieg)	-1,45
	Limburgerhof, vfr.	2,24	VG Rockenhausen	-1,37
Faktor II Siedlungsstruktur und Anteil sozial Schwacher	Germersheim, vfr.	4,54	VG Nieder-Olm	-1,91
	Ludwigshafen am Rhein, kfr. St.	3,10	VG Offenbach an der Queich	-1,68
	Pirmasens, kfr. St.	2,68	VG Grünstadt-Land	-1,47
Faktor III Wirtschaftsstruktur und Art der Erwerbstätigkeit	Morbach, vfr.	2,77	Grafschaft, vfr.	-2,57
	VG Prüm	2,50	Koblenz, kfr. St.	-2,37
	Ludwigshafen am Rhein, kfr. St.	2,23	VG Ulmen	-2,14
Faktor IV Alters- und Familienstruktur	Germersheim, vfr.	2,44	Bad Neuenahr-Ahrweiler, vfr.	-2,94
	VG Lingenfeld	1,90	Idar-Oberstein, vfr.	-2,45
	VG Ransbach-Baumbach	1,64	Pirmasens, kfr. St.	-2,16
Faktor V Unternehmensgrößenstruktur	VG Offenbach an der Queich	2,41	VG Kaiserslautern-Süd	-2,14
	Ingelheim am Rhein, vfr.	2,30	VG Landau-Land	-2,13
	VG Nassau	1,98	VG Freinsheim	-2,02
Faktor VI Selbstständigkeit	VG Freinsheim	2,18	Trier, kfr. St.	-2,75
	VG Diez	2,07	Mainz, kfr. St.	-2,48
	VG Deidesheim	2,05	Pirmasens, kfr. St.	-2,35
Faktor VII Konfessionsstruktur	VG Kusel	1,83	VG Saarburg	-2,09
	Idar-Oberstein, vfr.	1,83	VG Prüm	-2,01
	VG Herrstein	1,70	VG Hermeskeil	-1,86
	Höchste Abweichungen vom Durchschnitt		Geringste Abweichungen vom Durchschnitt	
	Verwaltungsbezirk	Abweichung	Verwaltungsbezirk	Abweichung
	Mainz, kfr. St.	6,20	VG Annweiler am Trifels	1,04
	Germersheim, vfr.	5,65	VG Langenlonsheim	1,38
Trier, kfr. St.	4,93	VG Westerburg	1,40	

Faktoren erklären regionale Unterschiede bei Wahlbeteiligung und Wahlergebnissen

Regressionsmodelle mit den extrahierten Faktoren als erklärende Variablen

Die bei der Faktorenanalyse ermittelten sieben Faktoren, die die Dimensionen der sozialen und ökonomischen Unterschiede in Rheinland-Pfalz beschreiben, werden nun verwendet, um die Zusammenhänge zwischen diesen Dimensionen und den Wahlergebnissen der Bundestagswahl 2017 zu analysieren. Hierzu werden multivariate lineare Regressionsmodelle aufgestellt, in die die sieben Faktoren sowie eine Konstante als erklärende Variablen eingehen. Es werden insgesamt sieben Regressionsmo-

delle für die sieben zu erklärenden Variablen – die Wahlbeteiligung sowie die Zweitstimmenergebnisse derjenigen sechs in Rheinland-Pfalz zur Wahl angetretenen Parteien, die im 19. Deutschen Bundestag vertreten sind – geschätzt.

Die Ergebnisse der Regressionen sind wie folgt zu interpretieren: Der Regressionskoeffizient der Konstante entspricht der geschätzten Wahlbeteiligung bzw. dem geschätzten Zweitstimmenergebnis der betreffenden Partei in einem hypothetischen Verwaltungsbezirk, für den die Faktorwerte aller sieben Faktoren

Interpretation der Regressionskoeffizienten

gleich null sind. Aufgrund der Standardisierung der Variablen geben diese Werte gleichzeitig die geschätzte durchschnittliche Wahlbeteiligung bzw. das geschätzte durchschnittliche Zweitstimmenergebnis in den 139 in die Analyse einbezogenen Verwaltungseinheiten an, wobei jedoch abweichend vom realen Berechnungsverfahren alle Gebietseinheiten ungeachtet der Zahl der Wahlberechtigten bzw. der Wählerinnen und Wähler mit gleicher Gewichtung eingehen. Der Koeffizient eines Faktors zeigt die geschätzte Veränderung der zu erklärenden Variable in Prozentpunkten, wenn sich der Faktorwert des betreffenden Faktors um eins erhöht. Diese Erhöhung entspricht einer Standardabweichung in der Verteilung der Faktorwerte der 139 Gebietseinheiten.

Siedlungsstruktur und Anteil sozial Schwacher hat größten Effekt auf die Wahlbeteiligung

Fünf der sieben Faktoren stehen in signifikantem Zusammenhang mit der Wahlbetei-

ligung. Kein signifikanter Einfluss findet sich bei den Faktoren zur Alters- und Familienstruktur (Faktor IV) sowie zur Konfessionsstruktur (Faktor VII).

Den größten Effekt hat Faktor II, bei dem Gebietseinheiten mit städtischer Prägung und überdurchschnittlichen Anteilen sozial schwacher und bildungsferner Personen an der Bevölkerung hohe Faktorwerte erreichen. Eine Erhöhung des Faktorwerts um eins geht mit einem Rückgang der Wahlbeteiligung um 2,8 Prozentpunkte einher.

Faktor V zur Unternehmensgrößenstruktur, bei dem durch Großunternehmen geprägte Verwaltungsbezirke hohe Faktorwerte aufweisen, zeigt den zweitstärksten Effekt: Mit einem um eins höheren Faktorwert reduziert sich die Wahlbeteiligung um 1,5 Prozentpunkte.

Auch Faktor I zeigt einen deutlichen Zusammenhang mit der Neigung, vom Wahlrecht Gebrauch zu machen. Gebietseinheiten

Signifikante Korrelation der Wahlbeteiligung mit fünf Faktoren

Siedlungsstruktur und Anteil sozial Schwacher zeigt größten Erklärungseffekt für die Wahlbeteiligung

Zweitstärkster Zusammenhang mit der Unternehmensgrößenstruktur

T 4 Zusammenhang zwischen den in der Faktorenanalyse ermittelten Faktoren und dem Wahlverhalten bei der Bundestagswahl 2017: Ergebnisse der Regressionsanalysen							
Variable	Wahlbeteiligung	CDU	SPD	AfD	FDP	GRÜNE	DIE LINKE
Konstante	78,138***	36,418***	24,278***	11,717***	10,395***	6,867***	6,409***
Faktor I Siedlungsstruktur und Anteil höher Gebildeter	0,989***	-0,916***	-1,160***	-0,430***	0,590***	1,585***	0,332***
Faktor II Siedlungsstruktur und Anteil sozial Schwacher	-2,825***	-1,348***	0,271	1,000***	-0,244**	-0,164**	0,457***
Faktor III Wirtschaftsstruktur und Art der Erwerbstätigkeit	-0,317**	-0,474**	0,117	0,930***	-0,199**	-0,388***	-0,211***
Faktor IV Alters- und Familienstruktur	0,041	-0,439**	0,096	0,422***	-0,182*	0,053	-0,008
Faktor V Unternehmensgrößenstruktur	-1,467***	-0,171	0,179	-0,412***	-0,040	0,220***	0,272***
Faktor VI Selbstständigkeit	0,662	0,710***	-0,772***	-0,491***	0,923***	0,353***	-0,493***
Faktor VII Konfessionsstruktur	-0,161	-3,944***	2,094***	1,255***	-0,209**	0,111	0,280***
Bestimmtheitsmaß (R ²)	0,788	0,773	0,525	0,580	0,545	0,773	0,559

***/**/*: signifikant zum 1-/5-/10-Prozent-Niveau.

mit hohen Faktorwerten bei diesem Faktor sind durch eine dichte Siedlungsstruktur und die Ausprägung akademischer Milieus gekennzeichnet; die höchsten Faktorwerte erreichen die Universitätsstädte Mainz und Trier. Eine Erhöhung des Faktorwerts um eins bedeutet hier eine Steigerung der Wahlbeteiligung um einen Prozentpunkt.

79 Prozent der Unterschiede bei der Wahlbeteiligung werden erklärt

Zusammen können die sieben Faktoren etwa 79 Prozent der Varianz in den Daten zur Wahlbeteiligung in den 139 Verwaltungsbezirken erklären.

CDU-Ergebnis steht vor allem mit Konfessionsstruktur in Zusammenhang

Mit Ausnahme der Unternehmensgrößenstruktur (Faktor V) stehen alle Faktoren in signifikanter Korrelation mit dem Zweitstimmenanteil, den die CDU in den verschiedenen Verwaltungsbezirken erreichen konnte.

Deutlicher Effekt der Konfessionsstruktur auf den Zweitstimmenanteil der CDU

Faktor VII, die Konfessionsstruktur, zeigt den mit Abstand größten Effekt auf das CDU-Ergebnis. Hohe Faktorwerte erhalten hier besonders evangelisch geprägte Gebietseinheiten; dementsprechend weisen katholisch geprägte Verwaltungsbezirke niedrige Faktorwerte auf. Mit einer Erhöhung des Faktorwerts um eins sinkt der Anteil der Zweitstimmen, der auf die Christdemokraten entfällt, um 3,9 Prozentpunkte.

Zusammenhänge mit den Faktoren I und II zeigen schlechteres CDU-Ergebnis in Städten

Weitere größere Effekte auf das CDU-Ergebnis finden sich bei den Faktoren zur Siedlungsstruktur und dem Anteil sozial Schwacher (Faktor II) sowie zur Siedlungsstruktur und dem Anteil höher Gebildeter (Faktor I). Mit einem um eins höheren Faktorwert geht bei beiden Faktoren jeweils ein Rückgang des CDU-Zweitstimmenanteils einher: Bei Faktor II beträgt dieser Rückgang 1,3 Prozent-

punkte und bei Faktor I 0,9 Prozentpunkte. Hierin spiegelt sich das schlechtere Abschneiden der Partei in Städten wider, da bei beiden Faktoren städtisch geprägte Gebietseinheiten tendenziell höhere Faktorwerte erzielen.

Das Regressionsmodell für die Zweitstimmenergebnisse der CDU kann rund 77 Prozent der Unterschiede zwischen den betrachteten Gebietseinheiten erklären.

Erfolg der SPD korreliert ebenfalls mit Konfessionsstruktur

Wie beim CDU-Ergebnis hat auch bei der SPD die Konfessionsstruktur (Faktor VII) den größten Effekt auf das Abschneiden, allerdings mit umgekehrten Vorzeichen. Überdurchschnittliche Zweitstimmenanteile erreicht die SPD in Verwaltungsbezirken, die evangelisch geprägt sind und dementsprechend hohe Faktorwerte aufweisen. Eine Erhöhung des Faktorwertes um eins geht mit einer Steigerung des SPD-Stimmenanteils um 2,1 Prozentpunkte einher.

Den zweitgrößten Effekt auf das SPD-Ergebnis hat Faktor I, der auf die Siedlungsstruktur und den Anteil höher Gebildeter abstellt. In Gebietseinheiten mit höheren Faktorwerten schneidet die SPD schlechter ab; bei einem um eins höheren Faktorwert reduziert sich ihr Stimmenanteil um 1,2 Prozentpunkte.

SPD in städtischen Gebieten mit akademischem Umfeld weniger erfolgreich

Auch Faktor VI, welcher Merkmale zur Selbstständigkeit bündelt, zeigt einen signifikanten Zusammenhang mit dem Resultat der Partei. Verwaltungsbezirke mit einem überdurchschnittlichen Selbstständigenganteil an den Erwerbstätigen erzielen bei diesem Faktor hohe Faktorwerte. Mit einer Steigerung des Faktorwertes um eins sinkt der Zweitstimmenanteil der SPD um 0,8 Prozentpunkte.

Nur drei Faktoren korrelieren signifikant mit SPD-Ergebnis

Nur diese drei Faktoren stehen in signifikanter Verbindung mit dem Abschneiden der Sozialdemokraten. Die Koeffizienten der übrigen vier Faktoren bleiben insignifikant. Von den sieben Regressionsmodellen zeigt dasjenige zur Erklärung des SPD-Zweitstimmenanteils die schlechteste Anpassung. Lediglich etwa 52 Prozent der Unterschiede zwischen den SPD-Ergebnissen in den betrachteten Gebietseinheiten können mit den sieben Faktoren in Zusammenhang gebracht werden.

Auch das Abschneiden der AfD hängt am stärksten mit der Konfessionsstruktur zusammen

Alle sieben Faktoren haben einen signifikanten Erklärungsgehalt für das Zweitstimmenergebnis der AfD. Die Faktoren erklären gemeinsam 58 Prozent der Varianz in den Daten.

Wie bereits bei CDU und SPD hat auch bei der AfD Faktor VII zur Konfessionsstruktur den größten Effekt. Höhere Faktorwerte, die tendenziell mit einem hohen Evangelischenanteil und einem niedrigen Katholikenanteil verbunden sind, gehen mit höheren AfD-Zweitstimmenanteilen einher. Der Effekt geht somit in die gleiche Richtung wie bei der SPD und ist dem bei der CDU entgegengesetzt. Mit einer Erhöhung des Faktorwertes um eins verbessert sich das AfD-Ergebnis um 1,3 Prozentpunkte.

AfD in städtischen Gebieten mit schwachen sozialen Strukturen erfolgreich

Faktor II zeigt die zweistärkste Korrelation mit dem Resultat der AfD. Bei diesem Faktor weisen städtisch geprägte Verwaltungsbezirke mit überdurchschnittlicher Ausprägung sozial schwacher und bildungsferner Milieus hohe Faktorwerte auf. Erhöht sich der Faktorwert um eins, nimmt der Stimmenanteil der Partei um einen Prozentpunkt zu.

Auch zwischen Faktor III und dem AfD-Ergebnis findet sich ein starker Zusammenhang. In Faktor III gehen Merkmale zur Wirtschaftsstruktur und zur Art der Erwerbstätigkeit ein. Gebietseinheiten mit überdurchschnittlichen Anteilen von Erwerbstätigen im Produzierenden Gewerbe und abhängig Beschäftigten erhalten hier hohe Faktorwerte, während niedrige Faktorwerte auf eine Prägung des Verwaltungsbezirks durch die Dienstleistungsbereiche sowie einen hohen Anteil von Beamtinnen und Beamten hindeuten. Mit einem um eins höheren Faktorwert erhöht sich der Stimmenanteil der AfD um 0,9 Prozentpunkte.

Besseres AfD-Ergebnis in vom Produzierenden Gewerbe und abhängiger Beschäftigung geprägten Gebieten

Selbstständigkeit zeigt größten Effekt auf FDP-Ergebnis

Faktor VI, der Merkmale zur Selbstständigkeit bündelt, weist den stärksten Zusammenhang mit dem Wahlerfolg der FDP auf. In Gebietseinheiten mit hohen Faktorwerten, die tendenziell von einem überdurchschnittlichen Anteil der Selbstständigen an den Erwerbstätigen gekennzeichnet sind, ist die FDP erfolgreicher. Mit einer Erhöhung des Faktorwertes um eins geht eine Steigerung des Zweitstimmenanteils der Partei um 0,9 Prozentpunkte einher.

FDP schneidet in von selbstständigen Tätigkeiten geprägten Gebieten besser ab

Auch in städtischen Verwaltungsbezirken mit vielen höher Gebildeten kann die FDP überdurchschnittliche Stimmenanteile realisieren, wie der Koeffizient von Faktor I zeigt. Eine Erhöhung dieses Faktors um eins bedeutet ein um 0,6 Prozentpunkte besseres Ergebnis für die FDP.

Überdurchschnittliche FDP-Ergebnisse in städtischen Räumen mit akademischen Milieus

Unterdurchschnittlich schneidet die Partei hingegen in verdichteten Räumen mit vielen sozial Schwachen ab. Ein um eins höherer Faktorwert bei Faktor II reduziert das Zweitstimmenergebnis der Liberalen um 0,2 Prozentpunkte.

Sechs der sieben Faktoren stehen in Zusammenhang mit dem Ergebnis der FDP; lediglich bei der Unternehmensgrößenstruktur (Faktor V) lässt sich keine signifikante Korrelation nachweisen. Die in das Modell eingehenden Faktoren erklären knapp 55 Prozent der Unterschiede bei den FDP-Zweitstimmenanteilen.

Sehr großer Effekt der Siedlungsstruktur und des Anteils höher Gebildeter auf Zweitstimmenanteil der GRÜNEN

Regressionsmodell zur Erklärung des Stimmenanteils der GRÜNEN zeigt gute Anpassung

Bei dem Regressionsmodell für die GRÜNEN haben fünf der sieben Faktoren einen signifikanten Erklärungsgehalt. Insignifikant sind die Faktoren zur Alters- und Familienstruktur (Faktor IV) sowie zur Konfessionsstruktur (Faktor VII). Die beiden Modelle für die CDU und die GRÜNEN zeigen von den sechs Regressionsmodellen zur Erklärung der Stimmenanteile der Parteien gemeinsam die beste Anpassung. Das Bestimmtheitsmaß beträgt in beiden Modellen rund 77 Prozent. Das bedeutet, von den Unterschieden bei den Wahlergebnissen der beiden Parteien können jeweils 77 Prozent durch die Faktoren erklärt werden.

Bei den GRÜNEN mit Abstand größter relativer Effekt eines einzelnen Faktors auf das Ergebnis

Den mit deutlichem Abstand größten Erklärungsgehalt für den Stimmenanteil der GRÜNEN hat Faktor I. Das Ergebnis der Regression zeigt, dass die GRÜNEN in akademisch geprägten städtischen Verwaltungsbezirken deutlich überdurchschnittlich abschneiden. Mit einer Erhöhung des Faktorwertes um eins steigt der GRÜNEN-Zweitstimmenanteil um 1,6 Prozentpunkte. Relativ zum landesdurchschnittlichen Ergebnis der Partei bedeutet dies einen Anstieg um mehr als 20 Prozent. In den betrachteten Modellen handelt es sich hierbei um den mit Abstand stärksten relativen Effekt eines einzelnen Faktors auf den Stimmenanteil einer Partei.

Die nächstgrößten Effekte weisen die Faktoren zu Wirtschaftsstruktur und Art der Erwerbstätigkeit (Faktor III) sowie zur Selbstständigkeit (Faktor VI) auf. Hohe Faktorwerte gehen bei Faktor III mit höheren und bei Faktor VI mit niedrigeren Stimmenanteilen der GRÜNEN einher. Eine Erhöhung des Faktorwertes um eins bewirkt jeweils eine Erhöhung bzw. Verringerung um 0,4 Prozentpunkte.

Stimmenanteil der LINKEN steht am stärksten mit Selbstständigkeit in Zusammenhang

Faktor VI zur Selbstständigkeit besitzt die stärkste Verbindung mit dem Ergebnis der Partei DIE LINKE. Der Effekt des Faktors ist demjenigen, der bei der FDP zu beobachten ist, entgegengesetzt. In Gebietseinheiten mit hohen Faktorwerten schneidet DIE LINKE schlechter ab. Eine Erhöhung des Faktorwertes um eins bedeutet einen um 0,5 Prozentpunkte niedrigeren Zweitstimmenanteil für die Partei.

DIE LINKE weniger erfolgreich in Gebieten, in denen selbstständige Tätigkeit eine größere Rolle spielt

An zweiter und dritter Stelle bei den signifikanten Effekten auf das Ergebnis der LINKEN stehen die Faktoren II und I. Bei beiden Faktoren erreicht die Partei in Verwaltungsbezirken mit höheren Faktorwerten bessere Ergebnisse. Dies verdeutlicht, dass DIE LINKE allgemein in städtischen Räumen überdurchschnittliche Stimmenanteile erzielt, und zwar sowohl in sozial schwachen als auch in akademisch geprägten Gebieten. Bei Faktor II beträgt die Steigerung des Stimmenanteils, die mit einem um eins höheren Faktorwert einhergeht, 0,5 Prozentpunkte, und bei Faktor I sind es 0,3 Prozentpunkte.

Korrelationen mit den Faktoren II und I machen besseres Ergebnis der LINKEN in Städten deutlich

Sechs der sieben Faktoren stehen mit dem Zweitstimmenergebnis der LINKEN in Zusammenhang. Einzig die Alters- und Famili-

enstruktur (Faktor IV) zeigt keine signifikante Korrelation. Der Anteil der durch die Faktoren erklärten Varianz beträgt etwa 56 Prozent.

Fazit

Sieben Faktoren beschreiben die sozialen und ökonomischen Strukturen in Rheinland-Pfalz

Mit der Faktorenanalyse kann die Vielzahl gesellschaftlicher und wirtschaftlicher Variablen, von denen das Wahlverhalten abhängig ist, systematisiert und es können Zusammenhänge aufgezeigt werden. Die Datengrundlage für die Faktorenanalyse zur Bundestagswahl 2017 bilden 28 sozialstrukturelle und ökonomische Merkmale in 139 rheinland-pfälzischen kreisfreien Städten, verbandsfreien Gemeinden und Verbandsgemeinden. Aus den 28 Merkmalen wurden sieben Faktoren extrahiert, die die grundlegenden Dimensionen der sozialen und ökonomischen Struktur abbilden, hinsichtlich derer sich die betrachteten Verwaltungsbezirke unterscheiden. Die sieben Faktoren wurden interpretiert und wie folgt benannt: Siedlungsstruktur und Anteil höher Gebildeter, Siedlungsstruktur und Anteil sozial Schwacher, Wirtschaftsstruktur und Art der Erwerbstätigkeit, Alters- und Familienstruktur, Unternehmensgrößenstruktur, Selbstständigkeit sowie Konfessionsstruktur.

Faktoren erklären Unterschiede bei den Wahlergebnissen

Im zweiten Schritt dienten die extrahierten Faktoren als erklärende Variablen in Regressionsmodellen, die die Zusammenhänge zwischen den Faktoren und der Wahlbeteiligung bzw. dem Zweitstimmenergebnis einer Partei bei der Bundestagswahl 2017 beleuchten. Die in die Modelle eingehenden Faktoren können gemeinsam zwischen 52 und 79 Prozent der Unterschiede bei der zu erklärenden Variablen aufklären. Am bes-

ten ist die Modellanpassung bei der Regression zur Wahlbeteiligung; am schlechtesten lassen sich die Unterschiede zwischen den SPD-Ergebnissen in den betrachteten Verwaltungsbezirken erklären.

Die Wahlbeteiligung steht am stärksten mit der Siedlungsstruktur und dem Anteil sozial Schwacher in Zusammenhang. In städtischen Gebieten mit hohen Anteilen sozial schwacher und bildungsferner Personen ist die Wahlbeteiligung geringer als in ländlichen Räumen ohne ausgeprägte soziale Problematik. Die Ergebnisse der drei Parteien, die bei der Bundestagswahl 2017 in Rheinland-Pfalz die meisten Zweitstimmen erzielen konnten, stehen allesamt am stärksten mit der Konfessionsstruktur in Verbindung. Während die CDU in evangelisch geprägten Gebieten unterdurchschnittlich und in katholisch geprägten Gebieten überdurchschnittlich abschneidet, verhält es sich bei SPD und AfD umgekehrt. Auf die Zweitstimmenanteile der FDP und der Partei DIE LINKE hat der Faktor Selbstständigkeit den größten Effekt. In Gebieten, in denen selbstständige Tätigkeiten eine überdurchschnittlich große Rolle spielen, ist die FDP stark und DIE LINKE schwach. Das Ergebnis der GRÜNEN korreliert am stärksten mit der Siedlungsstruktur und dem Anteil höher Gebildeter. Die Partei ist in verdichteten Räumen mit hohem Akademikeranteil deutlich erfolgreicher.

Konfessionsstruktur hat größten Effekt auf die Stimmenanteile von CDU, SPD und AfD

Dr. Martin Jacobs ist Referent im Referat „Analysen, Auftragsarbeiten, FDZ“.