

Zusatzkapitel 3: Variablenselektion zur Identifikation relevanter Anreize in den Analysen der Wahlteilnahme^{*}

Ein aussagekräftiger Test der MFS-Hypothesen ist nur für statistisch erklärungsstarke Anreizvariablen möglich, da andernfalls keine hinreichende Teststärke vorliegt (siehe dazu das ebenfalls auf den Verlagsseiten hinterlegte Zusatzkapitel zur statistischen Prüfbarkeit der MFS-Hypothesen). Im Buch wurde die Variablenselektion zur Identifikation der teilnahmerelevanten Anreize aus Platzgründen nur teilweise dargestellt. In diesem Anhang werden die zusätzlichen Informationen bereitgestellt.

Ausgangspunkt der Variablenselektion ist ein RC-Modell, das möglichst viele Anreizvariablen enthält, von denen eine Wirkung zu erwarten ist. Dieses wird auf Basis von Likelihood-Ratio-Tests schrittweise um Variablen reduziert, von denen offensichtlich kein bedeutender Einfluss ausgeht. Bei einem derartigen Vorgehen besteht die Gefahr, dass ein Modell resultiert, das überangepasst an die Stichprobe ist (Fox 1997: 359; Hosmer und Lemeshow 2000: 92). Daher wird der Selektionsprozess *nicht* so weit geführt, dass nur noch Prädiktoren mit signifikanten Koeffizienten im Modell verbleiben. Dies wäre auch deshalb problematisch, weil die MFS-Hypothese 1 erwarten lässt, dass der Effekt einer Reihe von Anreizvariablen möglicherweise erst sichtbar wird, wenn man den Interaktionseffekt mit dem Internalisierungsgrad der Wahlnorm berücksichtigt.

Die Tabellen 1 und 2 enthalten die RC-Ausgangsmodelle für die beiden verwendeten Datensätze. Dieses sind logistische Regressionen, in die zunächst alle untersuchten Anreiz- und Kontrollvariablen eingehen und die in ihrer Spezifikation den RC-Theorien der Wahlteilnahme folgen. In beiden

^{*} Dieses Zusatzkapitel kann wie folgt zitiert werden: Kroneberg, Clemens (2011): Zusatzkapitel 3 zu „Die Erklärung sozialen Handelns“: Variablenselektion zur Identifikation relevanter Anreize in den Analysen der Wahlteilnahme. URL: <http://vs-verlag.de/tu/Kroneberg-Erklaerung>. Stand: 01.06.2011

Datensätzen ergeben sich keinerlei Anzeichen für Interaktionseffekte zwischen den Einflussereignissen bzw. der wahrgenommenen Knappheit des Wahlausgangs auf der einen und den auf den Wahlausgang bezogenen Nutzenvariablen auf der anderen Seite (siehe bereits Andreß, Hagenaars und Kühnel 1997: 411; Blais und Young 1999: 54). Im NRW-Datensatz weisen drei der vier Produktterme ein negatives Vorzeichen auf anstatt des vorhergesagten positiven Vorzeichens. Der vierte Produktterm ist von geringer Größe und statistisch insignifikant. Letzteres gilt auch für die Produktterme auf nationaler und Wahlkreis-Ebene im Studenten-Datensatz. Die instrumentellen RC-Theorien der Wahlteilnahme lassen sich also nicht bestätigen.¹

¹ Um die Relevanz instrumenteller Anzeizeffekte nicht vorschnell auszuschließen, wurden einige Anschlussanalysen durchgeführt. In diesen Analysen wurden erstens die verschiedenen Interaktionseffekte einzeln getestet, um etwaige Multikollinearitätsprobleme zu umgehen. Zweitens wurden die Nutzenindikatoren und/oder die Erwartungsindikatoren zu additiven Indizes zusammengefasst, um den möglichen Einfluss von Messfehlern zu reduzieren und um über erklärungskräftigere Komponentenvariablen zu verfügen. Schließlich wurde drittens unter Vorwegnahme der MFS-Hypothese 1 betrachtet, ob sich die instrumentellen Anzeizeffekte aufzeigen lassen, wenn man nur die Teilstichprobe der Befragten mit relativ niedrigem Internalisierungsgrad der Wahlnorm betrachtet (analog zum Vorgehen in Kroneberg, Heintze und Mehlkop 2010). In keiner dieser Analysen ließen sich instrumentelle Anzeizeffekte, wie sie die instrumentellen RC-Theorien der Wahlteilnahme vorhersagen, nachweisen. Möglicherweise besteht hier ebenfalls das Problem einer zu geringen Teststärke. Dennoch ist für die weiteren Analysen der Schluss zu ziehen, dass die in den Daten aufzeigbaren Anzeizeffekte keine Interaktionseffekte zwischen Erwartungs- und Nutzenindikatoren beinhalten.

Tabelle 1: RC-Ausgangsmodell im NRW-Datensatz (Logistische Regressionen der intendierten und berichteten Wahlteilnahme)

	RC-Ausgangsmodell			
	Teilnahmeabsicht		Berichtete Teilnahme	
Lösungskompetenz	0.23	(0.36)	1.21**	(0.45)
Interessenvertretung	0.90**	(0.31)	0.57	(0.35)
Leistungsdifferenz	0.66	(0.50)	-0.90	(0.62)
Kandidatenpräferenz	0.33	(0.28)	0.00	(0.33)
Bedeutung der eigenen Stimme	1.24**	(0.46)	0.27	(0.53)
Bedeutung x Lösungskompetenz	-0.24	(0.49)	-1.42**	(0.55)
Bedeutung x Interessenvertretung	-1.00*	(0.47)	-0.86+	(0.51)
Bedeutung x Leistungsdifferenz	-0.88	(0.75)	0.66	(0.83)
Bedeutung x Kandidatenpräferenz	0.00	(0.43)	0.32	(0.47)
Parteiidentifikation	0.09	(0.13)	0.44**	(0.14)
Demokratiezufriedenheit	0.25	(0.30)	0.32	(0.31)
Opportunitätskosten	-0.18	(0.20)	0.16	(0.27)
Alter (in Jahren)	0.01*	(0.00)	0.02**	(0.01)
Abitur	-0.01	(0.16)	0.47**	(0.18)
Realschule	-0.05	(0.16)	0.13	(0.16)
Frau	-0.04	(0.13)	-0.25+	(0.13)
Internalisierungsgrad der Wahlnorm (IW)	0.88**	(0.21)	0.46*	(0.21)
Politisches Interesse	0.38	(0.32)	-0.02	(0.31)
Beteiligung im sozialen Umfeld	1.52**	(0.33)	1.85**	(0.38)
Konstante	-2.55**	(0.43)	-2.19**	(0.48)
σ_{y^*} (Std.abw. der Disposition)	2.50		2.44	
Pseudo R^2	0.375		0.279	
-2*Log-Likelihood	329.01		318.81	
N	855		623	

Anmerkungen: Logistische Regressionen der intendierten und berichteten Wahlteilnahme: y^* -standardisierte Logit-Koeffizienten (β^{Sy*}); Standardfehler in Klammern
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$ (zweiseitige Tests)

Tabelle 2: RC-Ausgangsmodell im Studenten-Datensatz (Logistische Regressionen der berichteten Wahlteilnahme)

RC-Ausgangsmodell		
Wichtigkeit (National)	0.41	(0.47)
Knappheit (National)	-0.61	(0.59)
Wichtigkeit x Knappheit (National)	0.45	(0.74)
Wichtigkeit (Wahlkreis)	0.04	(0.33)
Knappheit (Wahlkreis)	0.16	(0.38)
Wichtigkeit x Knappheit (Wahlkreis)	0.22	(0.57)
Parteineigung	-0.02	(0.10)
Kosten (Index)	-0.88**	(0.19)
Internalisierungsgrad der Wahlnorm (IW)	1.20**	(0.27)
Politisches Interesse	0.37+	(0.21)
Freunde/Familie gegen Nichtwahl	0.47*	(0.22)
Western Ontario	-0.61**	(0.12)
Präsentation Wählerparadoxon	-0.07	(0.12)
Panelteilnahme	0.07	(0.11)
Älter als 23 Jahre	0.23	(0.19)
Konstante	-0.43	(0.42)
σ_y^* (Std.abw. der Disposition)	2.36	
Pseudo R^2	0.252	
-2*Log-Likelihood	525.29	
N	628	

Anmerkungen: Logistische Regressionen der berichteten Wahlteilnahme: y^* -standardisierte Logit-Koeffizienten (β^{Sy*}); Standardfehler in Klammern
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$ (zweiseitige Tests)

Die schrittweise Variablenselektion bzw. Reduktion der RC-Ausgangsmodelle ist in den Tabellen 3 und 4 zusammengefasst. Dabei wurden die RC-Ausgangsmodelle sukzessive um Variablen reduziert, deren Koeffizienten statistisch insignifikant *und* ausgesprochen gering waren oder deren Vorzeichen den theoretischen Erwartungen der RC-Theorien widersprach. Mit Hilfe des Likelihood-Ratio-Tests (LR-Test) wurde überprüft, dass die ausgeschlossenen Variablen nicht signifikant zur Devianzreduktion beitragen (vgl.

Hosmer und Lemeshow 2000: 92ff.). Zudem wurden in beiden Datensätzen zwei Nutzenkomponenten zu einem Index zusammengefasst, um zu erklärunskräftigeren Anreizvariablen und damit zu einer höheren Teststärke zu gelangen. Dass die Indexbildung statistisch gerechtfertigt ist, wird durch ähnliche Koeffizienten im vorherigen Modell nahe gelegt und wird auch von den LR-Tests bestätigt. Letztere zeigen, dass der Zusammenschluss der Variablen zu keiner statistisch signifikanten Erhöhung der Devianz führt.

Tabelle 3: Variablenselektion auf Basis des RC-Ausgangsmodells im NRW-Datensatz (Logistische Regressionen der intendierten und berichteten Wahlteilnahme)

Modell	Veränderung gegenüber dem vorherigen Modell	Grund für Veränderung des vorherigen Modells	Anzahl Parameter	D(M _B)	Pseudo R ²
Intendierte Wahlteilnahme (N = 855):					
M1			19	329.01	0.3754
M2	ohne <i>Bedeutung</i> x (<i>Lösungskompetenz</i> , <i>Interessenvertretung</i> , <i>Leistungsdifferenz</i> , <i>Kandidatenpräferenz</i>)	größtenteils falsche Vorzeichen	15	337.57	0.3592
M3	ohne <i>Lösungskompetenz</i>	Nullkoeffizient	14	337.59	0.3591
M4	Indexbildung: <i>Präferenzindex</i> = <i>Interessenvertretung</i> + <i>Kandidatenpräferenz</i>	hinreichend ähnliche Koeffizienten	13	337.64	0.3590
M5	ohne <i>Leistungsdifferenz</i>	sehr geringer Koeffizient (selbst in Teilstichprobe)	12	338.36	0.3577
Berichtete Wahlteilnahme (N = 628):					
M5			12	334.10	0.2531
M6	ohne <i>Bedeutung</i> , <i>Opportunitätskosten</i>	falsche Vorzeichen	10	340.01	0.2586

Anmerkungen: D(M_B) bezeichnet die Devianz: -2*Log-Likelihood

Tabelle 4: Variablenselektion auf Basis des RC-Ausgangsmodells im Studenten-Datensatz (Logistische Regressionen der berichteten Wahlteilnahme)

Modell	Veränderung gegenüber dem vorherigen Modell	Grund für Veränderung des vorherigen Modells	Anzahl Parameter	D(M _B)	R ² _{McF}
M1			15	525.29	0.2516
M2	ohne <i>Wichtigkeit</i> x <i>Knappheit</i> (<i>National</i>), <i>Wichtigkeit</i> x <i>Knappheit</i> (<i>Wahlkreis</i>)	falsches VZ bzw. sehr geringer Koeffizient	13	526.01	0.2506
M3	ohne <i>Knappheit</i> (<i>National</i>)	falsches VZ	12	527.94	0.2478
M4	ohne <i>Parteineigung</i>	Nullkoeffizient	11	528.00	0.2477
M5	Indexbildung: <i>Wichtigkeit</i> (<i>Index</i>) = <i>Wichtigkeit</i> (<i>National</i>) + <i>Wichtigkeit</i> (<i>Wahlkreis</i>)	hinreichend ähnliche Koeffizienten	10	529.71	0.2453

Anmerkungen: N = 628. D(M_B) bezeichnet die Devianz: -2*Log-Likelihood

Literatur

- Andreß, Hans-Jürgen, Jacques A. Hagenaars und Steffen Kühnel, 1997: Analyse von Tabellen und kategorialen Daten. Log-lineare Modelle, latente Klassenanalyse, logistische Regression und GSK-Ansatz. Berlin: Springer.
- Blais, André und Robert Young, 1999: Why do People Vote? An Experiment in Rationality. *Public Choice* 99: 39-55.
- Fox, John, 1997: Applied regression analysis, linear models, and related models. Thousand Oaks, CA: Sage.
- Hosmer, David W. und Stanley Lemeshow, 2000: *Applied Logistic Regression*. New York: Wiley.
- Kroneberg, Clemens, Isolde Heintze und Guido Mehlkop, 2010: The Interplay of Moral Norms and Instrumental Incentives in Crime Causation. *Criminology* 48: 259-294.

Die Erklärung sozialen Handelns
Grundlagen und Anwendung einer integrativen Theorie
Kroneberg, C.
2011, X, 348 S., Softcover
ISBN: 978-3-531-17389-4