

Gruppenvergleiche in linearen und logistischen Regressionsmodellen

Kühnel, Steffen M.

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Kühnel, S. M. (1996). Gruppenvergleiche in linearen und logistischen Regressionsmodellen. *ZA-Information / Zentralarchiv für Empirische Sozialforschung*, 39, 130-160. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-200402>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Gruppenvergleiche in linearen und logistischen Regressionsmodellen

von Steffen-M. Kühnel ¹

Zusammenfassung:

In Gruppenvergleichen wird geprüft, ob die Parameter eines statistischen Modells in verschiedenen (Sub-) Populationen variieren. Unterscheiden sich alle oder zumindest einige Parameter nicht, ist es sinnvoll, über die Gruppen hinweg eine gemeinsame Schätzung dieser Werte zu erhalten. Grundsätzlich lassen sich zwei alternative Strategien des Gruppenvergleichs unterscheiden. Zum einen kann die Gruppenzugehörigkeit als eine unabhängige Variable in das Modell eingeführt werden. Zum anderen können die Modellparameter jeweils in gruppenspezifischen Modellen geschätzt werden. Für beide Strategien stehen einfache Tests auf Gleichheit der Modellparameter zwischen den Gruppen zur Verfügung. In linearen Modellen bietet sich die zweite Strategie vor allem dann an, wenn zwischen den Gruppen unterschiedliche Residualvarianzen bestehen. Die generelle Vorgehensweise ist jedoch nicht auf lineare Modelle beschränkt. Anhand eines empirischen Vergleichs der Teilnehmer und Ausfälle einer Wiederholungsbefragung wird die Ähnlichkeit der Vorgehensweise in linearen und logistischen Regressionsmodellen demonstriert.

Abstract:

Group comparisons are used to investigate the stability of a statistical model in two or more (sub-) populations. If a parameter of a model is stable across groups, a common estimate of its value will be desirable. Group comparisons can be realized by two different strategies. In the first strategy "group" is treated as an additional explanatory variable. In the second strategy the statistical model is specified separately for each group. Standard tests are available for both strategies. It is demonstrated how the two strategies can be applied both to linear and non-linear regressions models. The logic of group comparison

¹ Anschrift des Autors: Prof. Dr. **Steffen-M. Kühnel**, Justus-Liebig Universität Gießen, Fachbereich Gesellschaftswissenschaften, Institut für Politikwissenschaft, Karl-Glöcker-Str. 21 E, 35394 Gießen

is illustrated by the empirical example of panel attrition. In this example it is shown, that panel attrition does not affect the relationships between voting turnout and its predictors.

In Gruppenvergleichen werden Stichproben auf Gemeinsamkeiten und Unterschiede untersucht. Ausgangspunkt ist die Frage, ob sich die Parameter eines statistischen Modells zwischen den zu vergleichenden Gruppen signifikant unterscheiden. Falls es bei einigen oder auch bei allen Parametern keine signifikanten Unterschiede gibt, stellt sich die Aufgabe einer gemeinsamen Schätzung der gleichen Parameter über die Gruppen hinweg.

Obwohl ein Gruppenvergleich im Grunde nicht mehr ist als die Anwendung von bekannten statistischen Routinen,² scheint die Logik der Vorgehensweise nicht allen Sozialforschern vertraut zu sein. Nach meiner Erfahrung gilt dies insbesondere, wenn der Gruppenvergleich bei nichtlinearen Modellen wie der logistischen Regression anzuwenden ist. Ich möchte daher in diesem Beitrag die Logik des Gruppenvergleichs an einem einfachen Beispiel verdeutlichen. Inhaltlich geht es um die Frage, ob der Panelausfall bei einer Wiederholungsbefragung die Beziehungen in einem linearen oder logistischen Regressionsmodell zur Analyse der Wahlbeteiligung verzerrt. Dazu wird die Gruppe der Teilnehmer an der Wiederholungsbefragung mit der Gruppe der nur einmal befragten Personen verglichen.

Nach einer kurzen Erläuterung des Anwendungsbeispiels werde ich zunächst auf den Gruppenvergleich im Kontext der multiplen Regression eingehen. Anschließend wird die analoge Vorgehensweise bei einer logistischen Regression angewendet. Die in beiden Situationen eingesetzte Analysestrategie besteht darin, die Gruppenzugehörigkeit als eine nominalskalierte unabhängige Variable aufzufassen und diese Variable als zusätzlichen Prädiktor in einer für alle Gruppen gemeinsame Regressionsgleichung zu berücksichtigen. Daneben gibt es noch eine zweite Strategie des Gruppenvergleichs, bei der für jede der zu vergleichenden Gruppen eine eigene Regressionsgleichung aufgestellt wird. Auch für diese Strategie gibt es einen einfachen statistischen Test auf Gleichheit der Modellparameter, der sowohl im linearen wie im logistischen Fall angewendet werden kann. Die nach Gruppen getrennte Formulierung von Regressionsgleichungen empfiehlt sich vor allem dann, wenn die Residualvarianzen zwischen den Gruppen unterschiedlich groß sind. Sollen bei einer getrennten Schätzung von Regressionsgleichungen Modellparameter über die Gruppen hinweg gleichgesetzt werden, läßt sich dies mit einem simultanen Gruppenvergleich realisieren, wie er in Programmen zur Schätzung linearer Strukturgleichungsmodelle implementiert ist.

² Ich verzichte daher im folgenden weitgehend auf Quellenangaben. Alle beschriebene Techniken sind statistische Standardanwendungen. Nähere Informationen finden sich in einschlägigen Lehrbüchern, z.B. *Greene* (1993).

1. Das Anwendungsbeispiel: Verzerrt der Panelausfall die Beziehung in einer Regressionsgleichung?

Das in diesem Beitrag verwendete Beispiel bezieht sich auf eine von der Fritz Thyssen Stiftung geförderten Studie von *Dieter Ohr* und mir, in der wir die Anwendbarkeit des Rational-Choice Ansatzes zur Erklärung der Wahlbeteiligung empirisch untersucht haben. Im Rahmen dieser Studie am Institut für angewandte Sozialforschung in Köln wurden von FORSA, Dortmund, kurz vor der letzten Landtagswahl in Nordrhein-Westfalen Anfang Mai 1995 landesweit insgesamt 1002 Telefoninterviews durchgeführt. Gegenstand der Interviews war das beabsichtigte Wahlverhalten und seine möglichen Determinanten. In einer kurzen Nachbefragung in der Woche nach der Wahl konnten 730 Personen der Vorbefragung zu ihrem tatsächlichen Wahlverhalten wiederbefragt werden. Wenn die in der Nachbefragung berichtete Wahlbeteiligung zur Überprüfung der Prognosekraft der Determinanten der Wahlbeteiligung herangezogen werden soll, muß angenommen werden, daß der Panelausfall von 272 Befragten keine störenden Auswirkungen auf den Zusammenhang zwischen Wahlbeteiligung und deren Determinanten hat. Ein möglicher Hinweis für die Richtigkeit dieser Annahme ergäbe sich, falls sich die Regressionsgewichte der Regression der Wahlbeteiligungsabsicht auf die Determinanten der Wahlbeteiligung in der Teilstichprobe der Panelteilnehmer nicht von den entsprechenden Koeffizienten in der Teilstichprobe derjenigen unterscheiden, die nicht an der Wiederholungsbefragung teilgenommen haben.

Tabelle 1: Resultate der logistischen Regression im Gesamtdatensatz

LR-Index P ² :	0.435	(n=973)		
LR-Test	-2lnL	df	Prob.	
Konstantenmodell	617.984			
Gesamtmodell	349.435			
Differenz	268.549	5	< 0.001	
Determinante	b	SE	t	exp(b)
Kandidatenpräferenz	-1.149	0.303	-3.79	1/3.155
Einfluß der Stimme	-1.076	0.342	-3.14	1/2.932
Soziales Umfeld	-0.407	0.092	-4.43	1/1.503
Wahlbeteiligungsnorm	-1.203	0.136	-8.86	1/3.331
Beteiligungskosten	0.550	0.179	3.08	1.734
(Konstante)	0.362	0.380	0.95	1.436

Abhängige Variable ist in diesem Beispiel die Wahlbeteiligungsabsicht. Ich verwende im folgenden zwei Operationalisierungen. In der logistischen Regression wird zwischen den beiden Ausprägungen "werde wählen" (Wert: 0) und "werde nicht wählen / bin noch unentschlossen" (Wert: 1) unterschieden (Variable: VOTE_TWO). Für das lineare Regressionsmodell wird eine sechsstufige Skala mit den Polen "werde sicher wählen" (Wert: 1) und "werde sicher nicht wählen" (Wert: 6) eingesetzt (VOTE_SIX). Beide Versionen der abhängigen Variable werden durch fünf unabhängige Variablen erklärt: Vorliegen einer Kandidatenpräferenz (KAND), Einfluß der eigenen Stimme auf den Wahlausgang (EINFL), Teilnahmeverhalten der sozialen Umgebung (UMFELD), verinnerlichte Wahlnorm (NORM) und Kosten des Wählens (KOSTEN). Das Vorliegen einer Kandidatenpräferenz und der Einfluß der eigenen Stimme auf den Wahlausgang sind 0/1-kodierte dichotome Variablen, die anderen drei Prädiktoren additive Indizes aus mehreren Items mit 9 (Umfeld), 5 (Norm) bzw. 4 (Kosten) Kategorien.³

Ausgewählt wurden die fünf Prädiktoren ursprünglich im Kontext des logistischen Regressionsmodells, wo sie zusammen eine sehr gute Prognosekraft aufweisen.⁴ Von 904 Befragten, bei denen Wahlbeteiligung prognostiziert wird, wollen tatsächlich 95.8% (866) wählen. Von den 59 Prognosen der Nichtwähler und Unentschlossenen sind 78.0% (46) zutreffend. In Tabelle 1 sind die Ergebnisse der mit der SPSS-Prozedur LOGISTIC REGRESSION berechneten ML-Schätzung der Regressionskoeffizienten wiedergegeben.⁵ Der LR-Index P^2 , der in mancher Hinsicht als Analogon zum Determinationskoeffizienten der linearen Regression aufgefaßt werden kann, weist einen sehr hohen Wert von 43.5% auf. Der zugehörige Likelihood/Ratio-Test ist mit einem Chiquadratwert von 268.549 bei fünf Freiheitsgraden auf dem 1%-Niveau signifikant. Mit gleicher Irrtumswahrscheinlichkeit von null verschieden sind auch alle fünf Regressionskoeffizienten. Da das Modell Nichtteilnahme oder Unentschlossenheit prognostiziert, weisen erwartungsgemäß die Beteiligungskosten einen positiven Effekt auf, während alle Prädiktoren, die die Wahlbeteiligung erhöhen, einen negativen Effekt haben. Wird beispielsweise der eigenen Stimme ein Einfluß auf den Wahlausgang zugesprochen, sinkt das Verhältnis (engl. "Odds") von Nichtteilnahme/Unentschlossenheit zu Teilnahme um den Faktor 0.341 bzw. $1/2.932 (=e^{-1.0756})$.

Das logistische Regressionsmodell basiert auf 973 Fällen mit gültigen Antworten bei allen beteiligten Variablen. Wird statt der dichotomen abhängigen Variable die sechsstufige Antwortskala in einem linearen Regressionsmodell verwendet, reduziert sich die Fallzahl ge-

³ Da die inhaltliche Fragestellung für dieses methodische Beispiel unwesentlich ist, kann an dieser Stelle auf eine nähere Erläuterung der verwendeten Größen verzichtet werden.

⁴ Eine Einführung in die Logik der logistischen Regression findet sich in der ZA-Information 25 (*Kühnel* u.a., 1989, S.44-75).

⁵ Die von SPSS nicht ausgedruckten Werte des LR-Index P^2 sind ebenso wie die t-Werte mit der Hand berechnet. SPSS verwendet statt der t-Werte die Wald-Statistik. Wie in der linearen Regression berechnen sich die t-Werte als Verhältnis des geschätzten Regressionskoeffizienten durch seinen Standardfehler. Die Wald-Statistik eines Koeffizienten ist das Quadrat des t-Wertes.

ringfügig auf 970 Fälle.⁶ In dem linearen Modell wird 34% der Varianz der sechsstufigen Wahlbeteiligungsskala erklärt. Die Erklärungskraft kann deutlich erhöht werden, wenn zusätzliche Interaktionseffekte zwischen der Kandidatenpräferenz und dem Einfluß der eigenen Stimme sowie zwischen dem Teilnahmeverhalten der sozialen Umgebung und der Wahlnorm spezifiziert werden. Da beide Interaktionseffekte die Wirkung der jeweils sehr starken Haupteffekte reduzieren, dürften sie Folge der Begrenzung des Wertebereichs der abhängigen Variable sein (Ceiling-Effekt). Im nichtlinearen logistischen Modell sind entsprechend beide Interaktionseffekte auf dem 5%-Niveau nicht signifikant. Um in der Darstellung des Gruppenvergleichs das logistische Modell und das lineare Modell möglichst analog zu halten, werden die beiden Interaktionseffekte auch im linearen Modell ausgelassen. Die Ergebnisse der Kleinstquadratschätzung der Regressionskoeffizienten zeigt Tabelle 2. Mit Ausnahme der Beteiligungskosten, die nur auf dem 5%-Niveau signifikant sind, sind auch im linearen Modell die Regressionsgewichte der unabhängigen Variablen auf dem 1%-Niveau signifikant von null verschieden.

Tabelle 2: Resultate der linearen Regression im Gesamtdatensatz

Erklärte Varianz R ² :	0.340	(n=970)		
F-Test	Quadrate	df	Prob.	
Regression	436.46529	5		
Residuen	847.10275	964		
F-Statistik	99.338	< 0.001		
Determinante	b	SE	t	∃
Kandidatenpräferenz	-0.419	0.077	-5.43	-0.1474
Einfluß der Stimme	-0.718	0.106	-6.78	-0.1882
Soziales Umfeld	-0.163	0.024	-6.84	-0.1887
Wahlbeteiligungsnorm	-0.369	0.031	-12.02	-0.3384
Beteiligungskosten	0.105	0.045	2.316	0.0625
(Konstante)	3.107	0.116	26.83	

6 Ursache sind drei zusätzliche ungültige Antworten bei der Nachfrage nach der Sicherheit bzw. Wahrscheinlichkeit der Wahlbeteiligung oder Nichtteilnahme, die zusätzlich zu der Ausgangsfrage nach Wahlbeteiligung, Nichtteilnahme oder Unentschlossenheit für die Erstellung der sechsstufigen Skala herangezogen wurde.

Die in Tabelle 1 und 2 aufgeführten Ergebnisse beziehen sich auf den Gesamtdatensatz. Im Gruppenvergleich wird geprüft, ob die beiden Modellgleichungen tatsächlich in gleicher Weise sowohl für die Teilnehmer an der Wiederholungsbefragung als auch für diejenigen gelten, die nicht an der Wiederholungsbefragung teilgenommen haben. Befragte der ersten Gruppe werden im folgenden als 'Panelteilnehmer' bezeichnet, Befragte der zweiten Gruppe als 'Panelausfälle'. Voraussetzung für die Anwendung des Gruppenvergleichs ist die Unabhängigkeit der zu vergleichenden Stichproben. Da alle Befragten unabhängig voneinander ausgewählt worden sind, ist diese Bedingung erfüllt.

2. Die Logik des Gruppenvergleich im linearen Regressionsmodell

Im Kontext des linearen Regressionsmodells können sich Unterschiede zwischen den Panelteilnehmern und Panelausfällen auf verschiedene Weise auswirken. So könnte man die Vermutung haben, daß die Wahlbeteiligungsneigung bei den Panelteilnehmern höher ist als bei den Panelausfällen. Falls dies nicht allein eine Folge unterschiedlicher Ausprägungen bei den im Regressionsmodell betrachteten unabhängigen Variablen ist, bedeutet dies, daß sich auch bei Kontrolle dieser Variablen Mittelwertunterschiede zwischen den beiden Gruppen zeigen. Im Regressionsmodell müßten sich in diesem Fall die Regressionskonstanten zwischen den Gruppen unterscheiden. Unterschiede können auch zu gruppenspezifischen Beziehungen zwischen abhängiger Variable und unabhängigen Variablen führen. Im Regressionsmodell würde dann das Regressionsgewicht einer unabhängigen Variable zwischen den Gruppen variieren. Unterschiede zwischen den Gruppen können auch die Residualvarianzen betreffen. Im Beispiel des Panelausfalls könnte etwa die Residualvarianz der Panelausfälle deutlich höher sein als die Residualvarianz der Panelteilnehmer. Denkbar ist schließlich auch eine Kombination der drei aufgeführten Auswirkungen der Gruppenzugehörigkeit.

Wenn für den Gruppenvergleich die Strategie der Spezifikation einer gemeinsamen Gleichung für die Gruppen eingesetzt wird und die Koeffizienten der gemeinsamen Regressionsgleichung mit der gewöhnlichen Kleinstquadratschätzung berechnet werden, wird unterstellt, daß sich die Residualvarianzen zwischen den Gruppen nicht unterscheiden. Statistiker sprechen davon, daß es keine Heteroskedastizität zwischen den Gruppen gibt. Diese aus der Varianzanalyse bekannte Forderung soll zunächst als erfüllt gelten. Im übernächsten Abschnitt wird diese Annahme explizit geprüft und gezeigt, wie ein Gruppenvergleich auch bei Heteroskedastizität zwischen den Gruppen durchgeführt werden kann.

Unterscheiden sich allein die Regressionskonstanten zwischen den Gruppen, handelt es sich varianzanalytisch gesehen um einem Haupteffekt des Faktors 'Gruppenzugehörigkeit'. Variieren die Regressionsgewichte zwischen den Gruppen, bestehen Interaktionseffekte zwischen der Gruppenzugehörigkeit und den eigentlichen Prädiktoren des Regressionsmodells.

Bei der Spezifikation von Effekten der Gruppenzugehörigkeit gibt es verschiedene Möglichkeiten. Ich werde auf zwei Spezifikationen eingehen. Die vielleicht naheliegendste Methode besteht darin, die Gruppenzugehörigkeit durch dichotome 0/1-kodierte Dummyvariablen abzubilden. Wenn - wie im Anwendungsbeispiel - nur zwei Gruppen verglichen werden, reicht es aus, eine Dummyvariable (PANEL) zu spezifizieren, die z.B. bei den Panelteilnehmern den Wert "1" und bei den Panelausfällen den Wert "0" aufweist. Würde man drei Gruppen vergleichen und z.B. bei den Panelausfällen zwischen den expliziten Verweigerern und den nicht Erreichbaren unterscheiden, könnten die drei Gruppen über zwei Dummyvariablen erfaßt werden. Zusätzlich zu der Variable für die Panelteilnehmer könnte die zweite 0/1-kodierte Variable (NOPANEL) den Wert "1" aufweisen, wenn ein Befragter explizit nicht an der Wiederholungsbefragung teilnehmen will. Panelteilnehmer haben dann bei der ersten Variable den Wert "1" und bei der zweiten den Wert "0". Bei expliziten Verweigerern ist es gerade umgekehrt. Nicht Erreichbare hätten schließlich bei beiden Variablen den Wert "0". Generell gilt, daß bei insgesamt g Gruppen $g-1$ 0/1-kodierte Dummyvariablen gebildet werden. Die Gruppe, die bei allen Dummyvariablen den Wert "0" aufweist, ist die sogenannte Referenzgruppe für den Gruppenvergleich.

Interaktionseffekte erfassen Unterschiede zwischen den Regressionsgewichten in den Gruppen. Dazu werden zusätzliche Produktvariablen aus den unabhängigen Variablen und den Dummyvariablen der Gruppenzugehörigkeit gebildet. Wenn 'PANEL' die Dummyvariable ist, die die Panelteilnahme anzeigt, und 'KAND', 'EINFL', 'UMFELD', 'NORM' und 'KOSTEN' die Namen der fünf Prädiktoren bezeichnen, würden in SPSS die Produktvariable durch folgende fünf COMPUTE-Anweisungen erzeugt werden:

```
COMPUTE KAND_P=KAND*PANEL .  
COMPUTE EINFL_P=EINFL*PANEL .  
COMPUTE UMFELD_P=UMFELD*PANEL .  
COMPUTE NORM_P=NORM*PANEL .  
COMPUTE KOSTEN_P=KOSTEN*PANEL .
```

Sind mehr als zwei Gruppen zu vergleichen, sind entsprechende Produktvariablen für alle 0/1-kodierte Dummyvariablen zu bilden.

Tabelle 3: Gruppenvergleich im linearen Regressionsmodell durch schrittweise Regression

	Schritt 1		Schritt 2		Schritt 3	
R ² :	0.340		0.340		0.343	
F-Test:	99.338	p<0.001	82.727	p<0.001	45.483	p<0.001
R ² -Zuwachs:			< 0.001		0.003	
F-Test:			0.12086	p=0.728	0.861	p=0.507
Quadratsummen:						
Regression	436.46529	df=5	436.57160	df=6	440.36170	df=11
Residuen	847.10275	df=964	846.99645	df=963	843.20634	df=958
	b	t	b	t	b	t
KAND	-0.419	-5.4	-0.420	-5.4	-0.492	-3.0
EINFL	-0.718	-6.8	-0.718	-6.8	-0.980	-5.1
UMFELD	-0.163	-6.8	-0.163	-6.8	-0.182	-4.1
NORM	-0.369	-12.0	-0.368	-11.9	-0.342	-6.2
KOSTEN	0.105	2.3	0.105	2.3	0.085	1.0
(Konstante)	3.107	26.8	3.123	25.0	3.440	16.2
PANEL	-		-0.024	-0.3	-0.480	-1.9
KAND_P	-		-		0.096	0.5
EINFL_P	-		-		0.384	1.7
UMFELD_P	-		-		0.028	0.5
NORM_P	-		-		-0.037	-0.6
KOSTEN_P	-		-		0.024	0.2

Der Gruppenvergleich kann dann durch eine schrittweise Regression realisiert werden. Im ersten Schritt wird das ursprüngliche Modell mit gleichen Koeffizienten in allen Gruppen geschätzt. Im zweiten Schritt kann der Effekt der Gruppenzugehörigkeit auf die Regressionskonstante durch die zusätzliche Aufnahme der Dummyvariablen für die Gruppenzugehörigkeit untersucht werden. Im dritten Schritt werden schließlich durch Aufnahme der Produktvariablen die Auswirkung der Gruppenzugehörigkeit auf die Regressionsgewichte getestet. Für mein Anwendungsbeispiel würde in SPSS folgende Anweisung zur schrittweisen Regression gegeben:

```
REGRESSION /STATISTICS DEFAULT CHA
/DEPENDENT=VOTE_SIX
/ENTER KAND, EINFL, UMFELD, NORM, KOSTEN
/ENTER PANEL
/ENTER KAND_P, EINFL_P, UMFELD_P, NORM_P, KOSTEN_P.
```

Das Schlüsselwort "CHA" in der Option "STATISTICS" fordert den für den Gruppenvergleich relevanten F-Test auf Veränderungen im Determinationskoeffizienten an. Bei dem Anwendungsbeispiel ergeben sich die in Tabelle 3 wiedergegebenen Ergebnisse. Für jeden der drei schrittweise geschätzten Regressionsmodelle ist der Determinationskoeffizient R^2 und dessen F-Statistik, die Quadratsummen der Vorhersagewerte (Regression) und der Residuen sowie die unstandardisierten Regressionskoeffizienten und die zugehörigen t-Werte aufgeführt. Ab dem zweiten Schritt ist zudem der Anstieg des Determinationskoeffizienten (R^2 -Zuwachs) gegenüber dem vorherigen Schritt und das Resultat eines F-Tests dieses Anstiegs angegeben.

Das im ersten Schritt wiedergegebene Modell ist mit dem aus Tabelle 2 identisch. Die zusätzliche Berücksichtigung der Gruppenzugehörigkeit durch die 0/1-kodierte Dummyvariable PANEL erhöht die Erklärungskraft um weniger als 0.1%. Der F-Wert des Zuwachses von 0.121 ergibt ein empirisches Signifikanzniveau von nur 0.728: Die Wahrscheinlichkeit, daß der Zuwachs des Determinationskoeffizienten größer oder gleich dem beobachteten Wert ist, wenn die zusätzlichen Prädiktoren in der Grundgesamtheit gar keinen Effekt haben, beträgt also 72.8%. Die Gruppenzugehörigkeit hat also vermutlich keinen Effekt auf die Regressionskonstante. Der gleiche Schluß wird durch den sehr geringen t-Wert von 0.3 des Regressionsgewichts der Dummyvariablen PANEL nahegelegt. Die Berücksichtigung der Produktvariablen im dritten Schritt erhöht die Erklärungskraft um 0.3%. Relativ zum Modell aus Schritt 2 ist auch dieser Zuwachs nicht signifikant: Der F-Test ergibt einen Wert von 0.861, was zu einem empirischen Signifikanzniveau von 0.507 führt.

Betrachtet man die Werte der einzelnen Regressionskoeffizienten, so sind diese bei den Modellen aus Schritt 1 und 2 praktisch unverändert, während es bei dem Modell aus dem letzten Schritt merkbare Änderungen gibt. Für das Verständnis des Gruppenvergleichs mag es hilfreich sein, für dieses letzte Modell die Vorhersagegleichungen für jede Gruppe getrennt zu betrachten.

Da die 0/1-kodierte Dummyvariable PANEL bei den Panelausfällen den Wert "0" aufweist, haben auch alle Produkte dieser Dummyvariable mit den ursprünglichen Modellvariablen in der Gruppe der Panelausfälle den Wert "0". Hier gilt somit:

$$\begin{aligned} \hat{Y} &= 3.440 - 0.492 \times \text{KAND} - 0.980 \times \text{EINFL} - 0.182 \times \text{UMFELD} - 0.342 \times \text{NORM} \\ &\quad + 0.085 \times \text{KOSTEN} - 0.480 \times 0 + 0.096 \times 0 + 0.384 \times 0 + 0.028 \times 0 - 0.037 \times 0 \\ &\quad + 0.024 \times 0 \\ &= 3.440 - 0.492 \times \text{KAND} - 0.980 \times \text{EINFL} - 0.182 \times \text{UMFELD} - 0.342 \times \text{NORM} \\ &\quad + 0.085 \times \text{KOSTEN}. \end{aligned}$$

Bei den Panelteilnehmern hat die Variable PANEL den Wert "1". Infolgedessen sind die Produktvariablen KAND_P bis KOSTEN_P in dieser Gruppe nicht Null, sondern sind mit

den Werten der entsprechenden Ausgangsvariablen KAND bis KOSTEN identisch. Für die Vorhersagegleichung gilt also:

$$\begin{aligned}
 \hat{Y} &= 3.440 - 0.492 \times \text{KAND} - 0.980 \times \text{EINFL} - 0.182 \times \text{UMFELD} - 0.342 \times \text{NORM} \\
 &\quad + 0.085 \times \text{KOSTEN} - 0.480 \times 1 + 0.096 \times \text{KAND} + 0.384 \times \text{EINFL} \\
 &\quad + 0.028 \times \text{UMFELD} - 0.037 \times \text{NORM} + 0.024 \times \text{KOSTEN} \\
 &= (3.440 - 0.480) + (-0.492 + 0.096) \times \text{KAND} + (-0.980 + 0.384) \times \text{EINFL} \\
 &\quad + (-0.182 + 0.028) \times \text{UMFELD} + (-0.342 - 0.037) \times \text{NORM} \\
 &\quad + (0.085 + 0.024) \times \text{KOSTEN} \\
 &= 2.960 - 0.396 \times \text{KAND} - 0.596 \times \text{EINFL} - 0.154 \times \text{UMFELD} - 0.379 \times \text{NORM} \\
 &\quad + 0.104 \times \text{KOSTEN}.
 \end{aligned}$$

Den beiden Vorhersagegleichungen ist zu entnehmen, daß es in der Gruppe der Panelteilnehmer eine geringere Regressionskonstante gibt, die Panelteilnehmer also im Durchschnitt eine etwas höhere Teilnahmebereitschaft aufweisen. Außerdem sind die Effekte des Vorhandenseins einer Kandidatenpräferenz, des wahrgenommenen Einflusses der eigenen Stimme sowie der Wahlteilnahme des sozialen Umfelds bei den Panelteilnehmern geringer und die Effekte der verinnerlichteten Wahlnorm und der Teilnahmekosten stärker. Alle Unterschiede ergeben sich durch die Dummyvariable für die Gruppenzugehörigkeit und die Produktvariablen. Diese Variablen messen also die Differenz der Effekte zwischen der Gruppe der Panelausfälle (Referenzgruppe) und der Gruppe der Panelteilnehmer. Sie können somit direkt als Gruppenzugehörigkeitseffekte interpretiert werden.

Obwohl es bei der numerischen Betrachtung der Effekte sichtbare Unterschiede zwischen den beiden Gruppen gibt, besagt der F-Test, daß diese Unterschiede selbst auf dem 10%-Niveau nicht signifikant sind, also auch durch die Zufälligkeiten der Stichprobenziehung generiert sein können. Allerdings beziehen sich die in Tabelle 3 aufgeführten Tests auf R^2 -Zuwachs immer nur relativ zum vorherigen Modell. Es fehlt noch ein Test zwischen dem ersten und dem letzten Modell. Mit diesem Test wird die Nullhypothese geprüft, daß die Gruppenzugehörigkeit sowohl keinen Einfluß auf die Regressionskonstante wie auf die Regressionsgewichte hat. Aus den Angaben in Tabelle 3 läßt sich die Teststatistik leicht berechnen. Wenn sich nämlich zwei Regressionsmodelle nur dadurch unterscheiden, daß ein Modell ein Spezialfall des anderen Modells ist, dann kann mit einem F-Test geprüft werden, ob der Spezialfall signifikant vom allgemeineren Modell abweicht. Wenn $SS(E_1)$ die Quadratsumme der Residuen aus dem allgemeineren Modell bezeichnet, $SS(E_0)$ die Qua-

dratsumme der Residuen aus dem spezielleren Modell und df_1 und df_2 die Freiheitsgrade dieser Residuen,⁷ dann ergibt sich die F-Statistik nach:

$$F = \frac{(SS(E_0) - SS(E_1)) / (df_0 - df_1)}{SS(E_1) / df_1}$$

Soll also im Beispiel das Ausgangsmodell aus Schritt 1 gegen das letzte Modell (aus Schritt 3) getestet werden, berechnet sich aus den Werten in Tabelle 3 für die F-Statistik der Wert

$$F = \frac{(847.10275 - 843.20634) / (964 - 958)}{843.20634 / 958} = 0.738$$

Falls die Nullhypothese, daß der Spezialfall zutrifft, richtig ist, hier also das Ausgangsmodell zutrifft, dann ist die F-Statistik unter den üblichen Annahmen des Regressionsmodells F-verteilt mit $(df_1 - df_2)$ und df_2 Freiheitsgraden.⁸ Im Beispiel führt der F-Wert von 0.738 bei 6 und 958 Freiheitsgraden zu einem empirischen Signifikanzniveau von 0.619. Das in Schritt 3 geschätzte Regressionsmodell weicht also nicht signifikant vom ersten Modell aus Schritt 1 ab.

Voraussetzung für den F-Test ist, daß das eine Modell ein Spezialfall des anderen ist. Die beiden Modelle sind dann hierarchisch geschachtelt (engl.: nested). Im Beispiel ist diese Bedingung gegeben, da das erste Modell sich vom letzten Modell nur dadurch unterscheidet, daß die Regressionskoeffizienten der Dummyvariablen PANEL und der Produktvariablen jeweils auf den Wert null gesetzt sind. Auch die vom SPSS ausgegebenen F-Werte für den R^2 -Zuwachs sind nach dieser Formel berechnet.⁹ Soll die Teststatistik von SPSS be-

⁷ Die Freiheitsgrade der Residuen ergibt sich aus der Differenz $n - k$ der Fallzahl minus der Anzahl der im Modell geschätzten Regressionskoeffizienten, einschließlich einer evtl. vorhandenen Regressionskonstante.

⁸ Diese Annahmen fordern, (1) daß zumindest das allgemeinere Modell korrekt spezifiziert ist, die bedingten Mittelwerte der abhängigen Variable also eine lineare Funktion der Werte der unabhängigen Variablen sind, (2) daß die Residuen homoskedastisch sind, also für alle Ausprägungskombinationen der unabhängigen Variablen gleiche Varianzen aufweisen, und (3) daß die Residuen (in etwa) normalverteilt sind.

⁹ Der F-Test wird auch als R^2 -Change-Test bezeichnet, da die Teststatistik alternativ über die Formel:

$$F = \frac{(R_1^2 - R_0^2) / (k_1 - k_0)}{(1 - R_1^2) / (n - k_1)}$$

rechnet werden, kann dies dadurch geschehen, daß bei der schrittweisen Regression in einem vierten Schritt wieder das Ausgangsmodell berechnet wird. Dazu werden mit der Option REMOVE alle Effekte der Gruppenzugehörigkeit entfernt:

```
REGRESSION /STATISTICS DEFAULT CHA
  /DEPENDENT=VOTE_SIX
  /ENTER KAND, EINFL, UMFELD, NORM, KOSTEN
  /ENTER PANEL
  /ENTER KAND_P, EINFL_P, UMFELD_P, NORM_P, KOSTEN_P
  /REMOVE PANEL, KAND_P, EINFL_P, UMFELD_P, NORM_P, KOSTEN_P.
```

Für das Beispiel könnte damit der Gruppenvergleich abgeschlossen werden. Die Tests haben ergeben, daß die Gruppenzugehörigkeit weder einen signifikanten Haupteffekt noch signifikante Interaktionseffekte bewirken. Das Ausgangsmodell (aus Schritt 1) gilt somit für beide Gruppen. Dies gilt, obwohl sich rein numerisch zumindest bei der Konstante und dem Effekt der Beurteilung der Einflusses der eigenen Stimme größere Differenzen zwischen den Gruppen ergeben. Die entsprechenden Änderungsgrößen, also die Effekte der Prädiktoren PANEL und EINFL_P, sind mit t-Werten von -1.9 und 1.7 auf dem 5%-Niveau nicht signifikant. Die t-Werte sind jedoch nahe an dieser Signifikanzschwelle. Es scheint daher sinnvoll, ein weiteres Modell zu spezifizieren, in dem nur diese beiden Effekte zusätzlich zum Ausgangsmodell spezifiziert sind:

```
REGRESSION /STATISTICS DEFAULT CHA
  /DEPENDENT=VOTE_SIX
  /ENTER KAND, EINFL, UMFELD, NORM, KOSTEN
  /ENTER PANEL, EINFL_P.
```

Aber auch dieses Modell ist gegenüber dem strengeren Ausgangsmodell, daß gar keine Unterschiede zwischen den Gruppen zuläßt, nicht signifikant besser. Der von SPSS berechnete F-Test der Änderung ergibt eine Teststatistik von 1.812, die bei 2 und 962 Freiheitsgraden selbst auf dem 10%-Niveau nicht signifikant ist (Prob.: 0.164).

Ich habe oben erwähnt, daß es verschiedene Möglichkeiten gibt, die Gruppenzugehörigkeit im Regressionsmodell zu berücksichtigen. Neben der verwendeten Methode der Spezifikation von Unterschiede relativ zu einer Referenzgruppe besteht eine zweite Möglichkeit darin, für jede Gruppe die gruppenspezifischen Regressionskoeffizienten zu schätzen. Dazu ist es notwendig, für alle Gruppen 0/1-kodierte Dummyvariablen und Produktvariablen zu bilden. Im Anwendungsbeispiel müssen also entsprechende Variablen für die Gruppe der Panelausfälle generiert werden. Dies könnte durch folgende SPSS-Anweisungen geschehen:

berechnet werden kann, wobei R und R die Determinationskoeffizienten des allgemeineren bzw. des Spezialfalls bezeichnen, k_1 und k_0 die Anzahl der Regressionskoeffizienten in den beiden Modellen und n die Fallzahl. Aus Gründen der Rechengenauigkeit sollte diese Formel allerdings möglichst nicht verwendet werden.

```

COMPUTE NOPANEL=1-PANEL.
COMPUTE KAND_N=KAND*NOPANEL.
COMPUTE EINFL_N=EINFL*NOPANEL.
COMPUTE UMFELD_N=UMFELD*NOPANEL.
COMPUTE NORM_N=NORM*NOPANEL.
COMPUTE KOSTEN_N=KOSTEN*NOPANEL.

```

In der ersten Anweisung wird eine 0/1-kodierte Dummyvariable erzeugt, die den Wert "0" bei Panelteilnehmern und den Wert "1" bei Panelausfällen aufweist. Mit den anschließenden Anweisungen werden die Produktvariablen für die Panelausfälle generiert. In dieser Gruppe haben diese Variablen den Wert der Ausgangsvariablen KAND, EINFL, UMFELD, NORM bzw. KOSTEN. in der Gruppe der Panelteilnehmer haben die Produktvariablen dagegen stets die Ausprägung "0".

Im Unterschied zur ersten Vorgehensweise gibt es bei dieser Methode keine Referenzkategorie. Es ist dann möglich, ein Regressionsmodell zu schätzen, das die Regressionskoeffizienten für beide Gruppen getrennt ausgibt. Dabei ist allerdings zu beachten, daß das Modell keine gemeinsame Regressionskonstante für beide Gruppen enthalten darf. Dies wird dadurch erreicht, daß ein Regressionsmodell ohne Konstante geschätzt wird (Regression durch den Ursprung). In SPSS muß dazu vor der Angabe der abhängigen Variablen die Option ORIGIN angegeben werden:

```

REGRESSION /ORIGIN /DEPENDENT=VOTE_SIX
           /ENTER PANEL KAND_P EINFL_P UMFELD_P NORM_P KOSTEN_P
           NOPANEL KAND_N EINFL_N UMFELD_N NORM_N KOSTEN_N.

```

Das Ergebnis der Schätzung ist in Tabelle 4 wiedergegeben. Die in den beiden letzten Spalten angegebenen Regressionskoeffizienten und t-Werte beziehen sich auf die Effekte für die Gruppe der Panelausfälle, sind also die Koeffizienten für die Variablen KAND_N, EINFL_N, UMFELD_N, NORM_N, KOSTEN_N und NOPANEL (=Konstante für Panelausfälle). Entsprechend gelten die Werte der beiden ersten Spalten für die Panelteilnehmer, beziehen sich also auf die Variablen KAND_P, EINFL_P, UMFELD_P, NORM_P, KOSTEN_P und PANEL (=Konstante für Panelteilnehmer).

Da bei der Schätzung ein Regressionsmodell ohne gemeinsame Regressionskonstante berechnet wird (Regression durch den Ursprung), ist zu beachten, daß die von SPSS ausgegebenen Werte für den Determinationskoeffizienten und den F-Test i.a. nicht angemessen sind. Der F-Test bezieht auch die gruppenspezifischen Regressionskonstanten ein, im Beispiel aus Tabelle 4 also die Effekte der Dummyvariablen NOPANEL und PANEL. Der ausgedruckte Determinationskoeffizient basiert zudem auf der durchschnittlichen Rohproduktsumme der abhängigen Variablen und nicht auf deren Varianz.

Tabelle 4: Getrennte Koeffizientenschätzung der Gruppen mit einer Regressionsgleichung

Quadratsummen:

Regression	2299.79366	df=12
Residuen	843.20634	df=958

	Panelteilnehmer		Panelausfälle	
	b	t	b	t
Kandidatenpräferenz	-0.396	-4.5	-0.492	-3.0
Einfluß der Stimme	-0.596	-4.7	-0.980	-5.1
Soziales Umfeld	-0.154	-5.4	-0.182	-4.1
Wahlbeteiligungsnorm	-0.379	-10.2	-0.342	-6.2
Beteiligungskosten (Konstante)	0.109 2.960	2.0 21.2	0.085 3.440	1.0 16.2

Vergleicht man die Werte aus Tabelle 4 mit den Koeffizienten aus dem letzten Modell aus Tabelle 3 (Schritt 3), fällt auf, daß die Regressionskoeffizienten und t-Werte für die Gruppe der Panelausfälle mit den entsprechenden Koeffizienten aus Tabelle 3 exakt übereinstimmen. Tatsächlich wurde ja anhand der Vorhersagegleichung gezeigt, daß diese Koeffizienten die Vorhersage für die Panelausfälle wiedergeben. Die Regressionskoeffizienten für die andere Gruppe (Panelteilnehmer) stimmen mit den Werten überein, die oben aus den Koeffizienten des letzten Modells der schrittweisen Regression berechnet worden sind. Beide Modelle ergeben also die gleichen Vorhersagewerte. Daher ist auch die Summe der quadrierten Residuen in beiden Modellen identisch (Wert: 843.20634 bei 958 Freiheitsgraden). Genau dies ist gemeint, wenn von unterschiedlichen Parametrisierungen bzw. Reparametrisierungen die Rede ist. Man spricht auch davon, daß die Modelle empirisch äquivalent sind. Wenn anstelle der Haupt- und Interaktionseffekte wie bei Tabelle 4 in der Modellgleichung für jede Gruppe eigene Koeffizienten berechnet werden, werden diese Koeffizienten auch als konditionale Haupteffekte bezeichnet.

Im Prinzip ist es beliebig, ob das letzte Modell aus Tabelle 3 mit allen Interaktionseffekten oder das Modell mit konditionalen Haupteffekten aus Tabelle 4 berechnet wird. Die Vorgehensweise in Tabelle 3 ist zu empfehlen, wenn die Interaktionseffekte direkt getestet werden sollen und damit zu rechnen ist, daß es vermutlich keine Gruppenunterschiede gibt. Wenn dagegen mit Unterschieden zwischen den Gruppen gerechnet wird, kann das Modell der getrennten Koeffizienten (Tabelle 4) von Vorteil sein, da hier sofort für jede Gruppe die Koeffizienten und dessen Standardfehler (in der Tabelle nicht aufgeführt) bzw. t-Werte ersichtlich sind. Die Spezifikation konditionaler Haupteffekte dürfte vor allem dann ange-

bracht sein, wenn mehr als zwei Gruppen zu vergleichen sind. Bei der Verwendung von Interaktionseffekten beziehen sich diese nämlich stets auf die Differenz zur Referenzkategorie.

Bei der Verwendung konditionaler Haupteffekte können die einander entsprechenden Koeffizienten direkt zwischen allen Gruppen verglichen werden. Bei spezifischen Differenzen kann dann gezielt geprüft werden, ob diese signifikant sind. Zum Testen von Unterschieden zwischen den Gruppen ist das Modell mit den konditionalen Haupteffekten mit einem Modell zu vergleichen, bei dem einige oder alle Koeffizienten zwischen den Gruppen gleichgesetzt sind. Für den Test aller Koeffizienten auf Gleichheit ist das Vergleichsmodell das Ausgangsmodell aus Tabelle 2. Geprüft werden die Unterschiede wieder mit dem F-Test des Vergleichs der Residuen. Da die Quadratsumme der Residuen und deren Freiheitsgrade aus Tabelle 4 mit den entsprechenden Werten aus dem letzten Modell (Schritt 3) in Tabelle 3 übereinstimmen und das Ausgangsmodell in Tabelle 3 (Schritt 1) das Modell aus Tabelle 2 ist, ist dieser F-Test mit dem oben bereits durchgeführten Test identisch.¹⁰ Mit Hilfe der Daten aus Tabelle 2 kann auch ein korrekter Determinationskoeffizient für das Modell aus Tabelle 4 berechnet werden. Die Variation der abhängigen Variable ergibt sich aus der Summe der Quadratsummen für die Residuen und die Regressionsgeraden in Tabelle 2, also als 1283.5604 (=436.46529+847.10275). Der Determinationskoeffizient ist dann 1 minus dem Quotienten aus der Quadratsumme der Residuen des Modells aus Tabelle 4 und diesem Wert:

$$R^2 = 1 - 843.20634/1283.5604 = 0.343.$$

Der F-Test zwischen den Residuen aus Tabelle 2 und 4 und die Berechnung des Determinationskoeffizienten setzt voraus, daß das Modell aus Tabelle 2 tatsächlich ein Spezialfall des Modells aus Tabelle 4 ist. Tatsächlich unterscheiden sich die beiden Modelle nur dadurch, daß in Tabelle 4 die korrespondierenden Regressionskoeffizienten zwischen den Gruppen unterschiedliche Werte annehmen können, während in Tabelle 2 diese Werte zwischen den Gruppen gleichgesetzt sind.

Bei der Spezifikation konditionaler Haupteffekte ist es möglich, gezielt Koeffizienten gleichzusetzen. Dazu ist jeweils eine neue Variable zu bilden und anstelle der konditionalen Haupteffekte in das Modell einzusetzen. Angenommen, X_1 , X_2 und X_3 bilden die konditionalen Haupteffekte einer unabhängigen Variable X in drei Gruppen ab. Die drei Variablen sind also Produkte einer in allen Gruppen erfaßten Variable X mit den drei gruppenspezifischen 0/1-kodierten Dummyvariablen. Es kann nun beispielsweise getestet werden,

¹⁰ Da die Modelle aus Tabelle 2 und 4 getrennt berechnet werden, ist darauf zu achten, daß die Analyse jeweils auf den gleichen Fällen beruht.

ob die Effekte (Regressionsgewichte) von X in der zweiten und dritten Gruppe gleich sind. Dazu wird eine neue Variable X_23 gebildet, die bei Fällen aus der zweiten Gruppe die Werte von X_2 und bei Fällen aus der dritten Gruppe die Werte von X_3 enthält. Bei Fällen aus der ersten Gruppe hat die X_23 den Wert Null. Wenn D_2 und D_3 die 0/1-kodierten Dummyvariablen sind, die bei Fällen aus der zweiten bzw. dritten Gruppe den Wert eins aufweisen, könnte X_23 in SPSS z.B. folgendermaßen kreiert werden:

```
COMPUTE X_23=0.
IF (D_2 = 1) X_23=X_2.
IF (D_3 = 1) X_23=X_3.
```

In dem Regressionsmodell mit gleichgesetzten Effekten wird dann anstelle von X_1, X_2 und X_3 nur X_1 und X_23 spezifiziert. Der F-Test vergleicht dann die Residuen dieses Modells mit den Residuen des Modells, in denen alle drei Variablen X_1, X_2 und X_3 vorkommen. Wenn der F-Test zu einem signifikanten Ergebnis führt, sind die Effekte von X in der zweiten und dritten Gruppe signifikant verschieden.

Tabelle 5: Gruppenvergleich mit Interaktionseffekt und konditionalen Haupteffekt

Erklärte Varianz R ² :	0.343	(n=970)	
F-Test	Quadrate	df	Prob.
Regression	439.64412	7	
Residuen	843.92392	962	
F-Statistik	71.594		< 0.001
Determinante	b	SE	t
Kandidatenpräferenz	-0.418	0.037	-5.42
Einfluß (Panelausfälle)	-0.997	0.183	-5.45
Einfluß (Panelteilnehmer)	-0.593	0.125	-4.73
Soziales Umfeld	-0.163	0.024	-6.81
Wahlbeteiligungsnorm	-0.368	0.031	-11.97
Beteiligungskosten	0.106	0.045	2.33
Konstante bei Panelausfällen	3.366	0.180	18.69
Änderung bei Panelteilnehmern	-0.383	0.204	-1.88

Wenn es aus inhaltlichen Gründen angezeigt ist, können konditionale Haupteffekte auch mit Interaktionseffekten gemischt werden. In meinem Beispiel könnte z.B. der Unterschied zwischen den Regressionskonstanten als Interaktionseffekt (Differenz zur Referenzgruppe) und der Einfluß der Stimme als konditionaler Haupteffekt spezifiziert werden. Wenn alle ande-

ren Effekte zwischen den beiden Gruppen nicht variieren sollen, würde man in SPSS folgendes Regressionsmodell berechnen:

```
REGRESSION /DEPENDENT=VOTE_SIX
/ENTER KAND, EINFL, UMFELD_N UMFELD_P, NORM, KOSTEN, PANEL.
```

Das Ergebnis der Schätzung ist in Tabelle 5 festgehalten. Der Effekt der Dummyvariablen PANEL gibt die Differenz der Regressionskonstante der Gruppe der Panelteilnehmer zur Gruppe der Panelausfälle wieder. Der negative Effekt von -0.383 weist wiederum darauf hin, daß die Panelteilnehmer eine höhere Beteiligungsneigung aufweisen als die Panelausfälle. Der Einfluß der eigenen Stimme ist über konditionale Haupteffekte spezifiziert. In der Gruppe der Panelausfälle hat der Regressionskoeffizient den Wert -0.997. Bei den Panelteilnehmern beträgt der Effekt nur -0.593.

Obwohl zumindest auf den 10%-Niveau alle Regressionskoeffizienten signifikant von null verschieden sind, ist das Modell aus Tabelle 2, bei dem alle Effekte zwischen den Gruppen gleich sind, nicht signifikant schlechter. Der F-Test ergibt als Teststatistik:

$$F = \frac{(847.10275 - 843.92392) / (964 - 962)}{843.92392 / 962} = 1.812$$

Bei 2 und 962 Freiheitsgraden ergibt sich ein empirisches Signifikanzniveau von 0.164.

3. Die Logik des Gruppenvergleich im logistischen Regressionsmodell

Die vorgestellten Möglichkeiten des Gruppenvergleichs im linearen Regressionsmodell können direkt auf den Gruppenvergleich im nichtlinearen Modell der logistischen Regression übertragen werden. Zu beachten ist allein die jeweilige Interpretation der Regressionskoeffizienten, die sich bei der logistischen Regression auf die logarithmierten Wahrscheinlichkeitsverhältnisse (Logits) der Ausprägungen der abhängigen Variable beziehen. Außerdem wird beim Testen anstelle der F-Statistik der linearen Regression eine Likelihood/Ratio-Teststatistik verwendet, die bei korrekter Nullhypothese (asymptotisch) chi-quadratverteilt ist. Zur Verdeutlichung der Gleichheit der Vorgehensweise werden im folgenden die Analysen aus dem letzten Abschnitt mit einer logistischen Regression der dichotomen abhängigen Variablen VOTE_TWO wiederholt.

Zunächst wird der Gruppenvergleich wieder über eine schrittweise (logistische) Regression durchgeführt. Im ersten Schritt wird auch hier das Ausgangsmodell gleicher Koeffizienten in allen Gruppen geschätzt. Anschließend wird der Effekt der Gruppenzugehörigkeit auf die Regressionskonstante durch die zusätzliche Berücksichtigung der 0/1-kodierten Dummyvariablen PANEL getestet. Im dritten Schritt werden schließlich die Auswirkungen der Gruppenzugehörigkeit auf die Regressionsgewichte durch die Berücksichtigung von Interakti-

onseffekten untersucht. In SPSS kann die Analyse durch folgende Anweisung angefordert werden:

```
LOGISTIC REGRESSION VOTE_TWO WITH KAND, EINFL, UMFELD, NORM,
    KOSTEN, PANEL, KAND_P, EINFL_P, UMFELD_P, NORM_P, KOSTEN_P
/ENTER KAND, EINFL, UMFELD, NORM, KOSTEN
/ENTER PANEL
/ENTER KAND_P, EINFL_P, UMFELD_P, NORM_P, KOSTEN_P.
```

Tabelle 6: Gruppenvergleich im logistischen Regressionsmodell durch schrittweise Regression (n=973)

	Schritt 1		Schritt 2		Schritt 3	
-2 Log-Likelihood: (Konstantenmodell	349.435 k=6	617.984 k=1)	349.435 k=7		346.914 k=12	
LR-Index P ² :	0.435		0.435		0.439	
LR-Test:	268.549 p<0.001		268.549 p<0.001		271.071 p<0.001	
P ² -Zuwachs:			< 0.001		0.004	
LR-Test:			< 0.001 p=1.000		2.522 p=0.773	
	b	t	b	t	b	t
KAND	-1.149	-3.8	-1.149	-3.8	-0.614	-1.0
EINFL	-1.076	-3.1	-1.076	-3.1	-1.656	-2.8
UMFELD	-0.407	-4.4	-0.408	-4.4	-0.517	-3.2
NORM	-1.203	-8.9	-1.203	-8.8	-1.219	-4.5
KOSTEN	0.550	3.1	0.550	3.1	0.698	2.0
(Konstante)	0.362	1.0	0.359	0.8	0.454	0.7
PANEL	-		0.005	< 0.1	-0.264	-0.3
KAND_P	-		-		-0.692	-1.0
EINFL_P	-		-		0.839	1.1
UMFELD_P	-		-		0.150	0.8
NORM_P	-		-		0.001	0.0
KOSTEN_P	-		-		-0.142	-0.3

Die Ergebnisse der schrittweisen Regression sind in Tabelle 6 wiedergegeben. Der im zweiten Schritt berücksichtigte Effekt der Gruppenzugehörigkeit auf die Regressionskonstante hat bis auf die dritte Nachkommastelle keinen Einfluß auf die Vorhersagekraft des Modells. Der Regressionskoeffizient der Dummyvariable PANEL, die die Gruppenzugehörigkeit angibt, ist mit einem Wert von 0.005 praktisch Null. Der zugehörige t-Wert ist kleiner 0.1. Unter der Nullhypothese, daß ein Regressionskoeffizient in der Grundgesamtheit

null ist, ist der zugehörige t-Wert annähernd standardnormalverteilt. Ein auf dem 5%-Niveau signifikanter t-Wert sollte also größer als +2 bzw. -2 sein. Da diese Bedingung nicht erfüllt ist, gibt es bei einer Irrtumswahrscheinlichkeit von 5% keinen signifikanten Effekt. Das gleiche Ergebnis zeigt sich auch beim LR-Test auf eine signifikante Differenz des LR-Index P^2 . Ein bemerkbarer Zuwachs von P^2 tritt erst im dritten Schritt auf, wenn zusätzlich die Interaktionseffekte berücksichtigt werden. Allerdings ist auch dieser Zuwachs mit 0.4% sehr klein. Der LR-Test ist mit einem Wert von 2.522 nicht signifikant: das empirische Signifikanzniveau beträgt 0.774.

Ich habe bereits erwähnt, daß der LR-Test dem F-Test der linearen Regression entspricht. Wie dort setzt der Test voraus, daß die beiden Modelle hierarchisch geschachtelt sind, das eine der beiden zu vergleichenden Modelle also ein Spezialfall des anderen ist. In Tabelle 6 ist wie in Tabelle 3 das Ausgangsmodell (Schritt 1) ein Spezialfall des nächsten Modells (Schritt 2) und dieses ein Spezialfall des letzten Modells (Schritt 3). Die LR-Teststatistik L^2 ist jeweils die doppelte Differenz der negativen Log-Likelihoodfunktion zweier Modelle. Wenn $-2\ln L_0$ die doppelte negative Log-Likelihoodfunktion des Spezialfalls ist und $-2\ln L_1$ der entsprechende Wert des allgemeinen Modells, berechnet sich die Teststatistik also nach:

$$L^2 = (-2\ln L_0) - (-2\ln L_1)$$

Unter der Nullhypothese, daß der Spezialfall zutrifft, ist die Teststatistik asymptotisch chi-quadratverteilt. Die Anzahl der Freiheitsgrade ergibt sich aus der Differenz der jeweils geschätzten (freien) Regressionskoeffizienten. Mit diesem läßt sich auch der Unterschied zwischen dem ersten und letzten Modell in Tabelle 6 prüfen. Das erste Modell ist ein Spezialfall, da hier alle Koeffizienten zwischen den Gruppen gleichgesetzt sind. Die Teststatistik ergibt:

$$L^2 = 349.435 - 346.914 = 2.522.$$

Da im allgemeinen Modell zwölf Regressionskoeffizienten geschätzt werden und im Spezialfall nur sechs, ist die Teststatistik mit einer Chi-quadratverteilung mit sechs (=12-6) Freiheitsgraden zu vergleichen. Das empirische Signifikanzniveau beträgt 0.866. Mit einer Wahrscheinlichkeit von 86.6% ist also eine gleiche oder höhere Differenz der doppelten negativen Log-Likelihoodfunktionen zu erwarten, wenn in der Grundgesamtheit der Spezialfall zutrifft. Die Differenzen zwischen den Modellen sind also nicht signifikant.

Betrachtet man die Werte der Regressionskoeffizienten, so sind trotz der nicht signifikanten Unterschiede betragsmäßig Differenzen festzustellen. Wie im linearen Modell können die Regressionskoeffizienten auch hier zur Berechnung von Vorhersagegleichungen verwendet werden. Vorhergesagt werden im binären Logitmodell die Logits der Kategorien der abhängigen Variable, hier also das logarithmierte Verhältnis der Nichtwähler zu den Wählern.

Da die Variable PANEL und alle Produktvariablen in der Gruppe der Panelausfälle den Wert null haben, ergibt sich aus dem letzten Modell (Schritt 3) aus Tabelle 6 für die Logits in der Gruppe der Panelausfälle folgende Vorhersagegleichung:

$$\begin{aligned} \text{Logit} &= 0.454 - 0.614 \times \text{KAND} - 1.656 \times \text{EINFL} - 0.517 \times \text{UMFELD} \\ &\quad - 1.219 \times \text{NORM} + 0.698 \times \text{KOSTEN} - 0.264 \times 0 - 0.692 \times 0 + 0.839 \times 0 \\ &\quad + 0.150 \times 0 + 0.001 \times 0 - 0.142 \times 0 \\ &= 0.454 - 0.614 \times \text{KAND} - 1.656 \times \text{EINFL} - 0.517 \times \text{UMFELD} \\ &\quad - 1.219 \times \text{NORM} + 0.698 \times \text{KOSTEN}. \end{aligned}$$

In der Gruppe der Panelteilnehmer gilt dagegen:

$$\begin{aligned} \text{Logit} &= 0.454 - 0.614 \times \text{KAND} - 1.656 \times \text{EINFL} - 0.517 \times \text{UMFELD} \\ &\quad - 1.219 \times \text{NORM} + 0.698 \times \text{KOSTEN} - 0.264 \times 1 - 0.692 \times \text{KAND} \\ &\quad + 0.839 \times \text{EINFL} + 0.150 \times \text{UMFELD} + 0.001 \times \text{NORM} - \\ &\quad 0.142 \times \text{KOSTEN} \\ &= 0.190 - 1.306 \times \text{KAND} - 0.817 \times \text{EINFL} - 0.367 \times \text{UMFELD} \\ &\quad - 1.218 \times \text{NORM} + 0.556 \times \text{KOSTEN} \end{aligned}$$

Wie im linearen Modellen messen die Produktvariablen also die Differenz der Effekte in der Gruppe der Panelteilnehmer relativ zur Gruppe der Referenzkategorie der Panelausfälle. Die Regressionskoeffizienten in den einzelnen Gruppen können auch direkt geschätzt werden. Dazu wird analog zur linearen Regression ein Modell mit konditionalen Haupteffekten spezifiziert. Wieder ist dabei zu beachten, daß dann keine gemeinsame Regressionskonstante geschätzt wird. In SPSS wird dazu wieder die Option ORIGIN verwendet. Das Logitmodell mit konditionalen Haupteffekten kann durch folgende Anweisung berechnet werden:

```
LOGISTIC REGRESSION VOTE_TWO WITH PANEL KAND_P EINFL_P
UMFELD_P NORM_P KOSTEN_P NOPANEL KAND_N EINFL_N
UMFELD_N NORM_N KOSTEN_N
/ORIGIN.
```

Die Ergebnisse der Schätzung sind in Tabelle 7 wiedergegeben, die ganz analog zu Tabelle 4 aufgebaut ist. Da das Modell empirisch äquivalent zu dem letzten Modell (Schritt 3) aus Tabelle 6 ist, ist die doppelte negative Log-Likelihoodfunktion identisch (346.914). Der Vergleich zeigt zudem, daß die Koeffizienten der Panelausfälle mit den Schätzungen der Effekte der unabhängigen Variablen in Tabelle 6 übereinstimmt und daß die Effekte für die Panelteilnehmer genau die Summen der entsprechenden Haupt- und Interaktionseffekte aus Tabelle 6 sind. Wieder zeigt sich eine Analogie zu dem entsprechenden Vergleich im linearen Modell aus Tabelle 3 und 4.

Tabelle 7: Getrennte Koeffizientenschätzung der Gruppen mit einer Regressionsgleichung

	Panelteilnehmer		Panelausfälle	
	b	t	b	t
-2 Log-Likelihood	346.914		df=961	
Kandidatenpräferenz	-1.306	-4.3	-0.614	-1.0
Einfluß der Stimme	-0.817	-1.9	-1.656	-2.8
Soziales Umfeld	-0.367	-3.3	-0.517	-3.2
Wahlbeteiligungsnorm	-1.218	-7.4	-1.219	-4.5
Beteiligungskosten	0.556	2.5	0.698	2.0
(Konstante)	0.190	0.4	0.454	0.7

4. Gruppenvergleich über die Schätzung gruppenspezifischer Regressionsgleichungen

Bei der bisherigen Betrachtung wurde über alle Gruppen hinweg eine einzige lineare oder logistische Regressionsgleichung geschätzt. Durch die Spezifikation konditionaler Haupteffekte ergeben sich zwar gruppenspezifische Koeffizienten, diese werden aber gleichwohl im Rahmen einer einzigen Modellgleichung geschätzt. Es ist jedoch auch möglich, für jede Gruppe eine eigene Regressionsgleichung zu spezifizieren und zu schätzen. In SPSS kann dies beispielsweise über die Anweisung SPLIT FILE realisiert werden. Soll das lineare Regressionsmodell aus Tabelle 2 für jede Gruppe getrennt berechnet werden, werden folgende Anweisungen formuliert.

```

SORT CASES BY PANEL.
SPLIT FILE BY PANEL.
REGRESSION /STATISTICS DEFAULT BCOV
           /DEPENDENT VOTE_SIX
           /ENTER KAND, EINFL, UMFELD, NORM, KOSTEN.
SPLIT FILE OFF.

```

Durch das in der ersten Anweisung angeforderte Sortieren der Fälle nach der Gruppenzugehörigkeit wird sichergestellt, daß die nachfolgende Anweisung SPLIT FILE richtig funktioniert. Aufgrund dieser Anweisung wird die nachfolgende Regression für jede Gruppe getrennt berechnet. Anschließend wird die Split-Funktion mit der Anweisung SPLIT FILE OFF wieder ausgeschaltet.

Das Ergebnis der Schätzung ist in Tabelle 8 wiedergegeben. Vergleicht man die Regressionskoeffizienten dieser Tabelle mit den entsprechenden Koeffizienten aus Tabelle 4, so zeigt sich, daß die getrennte Schätzung in zwei Gleichungen zu demselben Ergebnis führt wie die gemeinsame Schätzung über konditionale Haupteffekte. Summiert man die Quadratsummen der Residuen in den beiden Gruppen (545.66712+297.53922), so ist auch diese Summe mit der Summe der Residuen aus Tabelle 4 (843.20634) identisch. Die getrennte Schätzung der

Modelle scheint also zu keinem neuen Ergebnis zu führen. Anders sieht es jedoch aus, wenn man die korrespondierenden t-Werte der beiden Tabellen vergleicht. Diese unterscheiden sich: in Tabelle 8 sind die t-Werte für die Gruppe der Panelteilnehmer größer und für die Gruppe der Panelausfälle kleiner.

Tabelle 8: Getrennte Schätzung von linearen Regressionsgleichungen in den Gruppen der Panelausfälle und der Panelteilnehmer

	Panelteilnehmer		Panelausfälle	
R ²	0.338		0.346	
F-Test	71.857	< 0.001	26.883	< 0.001
Quadratsummen:				
Regression	278.47936	df=5	157.45693	df=5
Residuen	545.66712	df=704	297.53922	df=254
	b	t	b	t
Kandidatenpräferenz	-0.396	-4.8	-0.492	-2.6
Einfluß der Stimme	-0.596	-5.0	-0.980	-4.4
Soziales Umfeld	-0.154	-5.7	-0.182	-3.6
Wahlbeteiligungsnorm	-0.379	-10.8	-0.342	-5.3
Beteiligungskosten	0.109	2.1	0.085	0.9
(Konstante)	2.960	22.6	3.440	14.0

Ursache dieser Unterschiede ist die getrennte Schätzung der Residualvarianzen in den Gruppen. Bei der gemeinsamen Schätzung aller Koeffizienten in einer Regressionsgleichung wird Homoskedastizität zwischen den Gruppen angenommen und entsprechend eine gemeinsame Populationsfehlervarianz geschätzt. Diese ist gerade die Summe der quadrierten Residuen geteilt durch die Freiheitsgrade, in Tabelle 4 also 0.880 (=843.20634/958). bei der getrennten Schätzung der Regressionsgleichungen in den beiden Gruppen wird auch für jede Gruppe getrennt eine eigene Schätzung der (Sub-) Populationsfehlervarianzen vorgenommen. Für die Gruppe der Panelteilnehmer ergibt sich der Wert 0.775 (=545.66712/704); in der Gruppe der Panelausfälle beträgt die Schätzung dagegen 1.171 (=297.53922/254). Da die Standardschätzfehler der Regressionskoeffizienten eine Funktion der geschätzten Fehlvvarianzen sind, sind diese in der Gruppe der Panelteilnehmer kleiner als in der Gruppe der Panelausfälle. Bei den t-Werten gilt dann das Umgekehrte: je größer die Fehlvvarianz, desto kleiner sind die t-Werte.

Die getrennte Schätzung der Residualvarianzen erklärt die Unterschiede bei den t-Werten. Offen bleibt aber zunächst die Frage, welchen Werten eher zu trauen ist. Die Antwort hängt davon ab, ob die Annahme der gleichen Residualvarianzen zwischen den Gruppen erfüllt ist

oder nicht. Wenn in der Population die Residualvarianzen in den Gruppen tatsächlich gleich sind, ist die gemeinsame Schätzung aus Tabelle 4 vorzuziehen, anderenfalls die getrennte Schätzung aus Tabelle 8. Hilfreich kann hier ein Test sein, der die Gleichheit oder Verschiedenheit der Residualvarianzen prüft. Bei nur zwei Gruppen kann ein F-Test auf Gleichheit der Varianzen angewendet werden. Bei dieser in der ökonometrischen Literatur als **Goldfeld-Quandt** Test bezeichneter Methode wird als Teststatistik der Quotient der geschätzten Residualvarianzen in den Gruppen als Teststatistik verwendet. Da die zu vergleichenden Gruppen unabhängig voneinander sind, ist dieser Quotient bei den üblichen Annahmen des Regressionsmodells F-verteilt, wenn die Residualvarianzen in der Grundgesamtheit gleich sind. Die Freiheitsgrade sind die Freiheitsgrade der Residual-varianzen. Angewendet auf das Anwendungsbeispiel ergibt sich eine F-Statistik von¹¹

$$F = (297.53922/254) / (545.66712/704) = 1.511.$$

Bei 254 und 704 Freiheitsgraden ergibt sich ein empirisches Signifikanzniveau kleiner 0.001. Die Nullhypothese gleicher Varianzen ist also abzulehnen.

Beim Vergleich der Residualvarianzen von mehr als zwei Gruppen ist der Test nicht anwendbar. Hier kann bei normalverteilten Residuen ein LR-Test angewendet werden, bei dem folgende Statistik berechnet wird:

$$L^2 = n \times \ln\left(\frac{SS(E)}{n}\right) - \sum_{g} n_g \times \ln\left(\frac{SS(E_g)}{n_g}\right)$$

In der Formel bezeichnet g die Gruppenzugehörigkeit. $SS(E_g)$ ist die Summe der quadrierten Residuen in der g -ten Gruppe, $SS(E)$ die entsprechende Summe über alle Gruppen, n_g die Fallzahl in der g -ten Gruppe und n die Gesamtfallzahl. Für das Beispiel aus Tabelle 8 ergibt sich folgender Wert:

$$\begin{aligned} L^2 &= 970 \times \ln\left(\frac{843.20634}{970}\right) - 710 \times \ln\left(\frac{545.66712}{710}\right) - 260 \times \ln\left(\frac{297.53922}{260}\right) \\ &= 9.89 \end{aligned}$$

Unter der Nullhypothese gleicher Varianzen in den Gruppen ist die Teststatistik asymptotisch chi-quadratverteilt. Die Zahl der Freiheitsgrade ergibt sich aus der Zahl der zu vergleichenden Gruppen minus 1; im Beispiel somit $df = 1$. Das empirische Signifikanzniveau ist hier < 0.01 . Die Varianzen sind also auf dem 1%-Niveau signifikant verschieden.

¹¹ Bei der Berechnung bildet die größere Varianz den Zähler und die kleinere den Nenner der Teststatistik.

Die Tests sprechen also dafür, daß die Annahme gleicher Residualvarianzen zwischen den Gruppen nicht erfüllt ist. Zumindest einen Unterschied gibt es, der auf die Gruppenzugehörigkeit zurückzuführen ist: in der Gruppe der Panelausfälle ist die Residualvarianz signifikant größer als in der Gruppe der Panelteilnehmer. Bedeutet dies, daß die bisherigen Ergebnisse des Gruppenvergleichs im linearen Modell haltlos sind, weil sie auf falschen Annahmen beruhen? Glücklicherweise muß dies nicht der Fall sein. Die Verletzung der Homoskedastizitätsannahme ist nämlich bei sehr großen Stichproben weitgehend unproblematisch, solange die Residualvarianzen nicht mit den unabhängigen Variablen des Modells variieren. Weil das Ignorieren von ungleichen Residualvarianzen aber nicht optimal ist und zumindest bei kleineren Stichproben zu zweifelhaften Ergebnisse beim statistischen Testen führt, kann es durchaus sinnvoll sein, dem Beispiel aus Tabelle 8 zu folgen und die Regressionsgleichungen für jede Gruppe getrennt zu schätzen.

Abbildung 1: Varianzen, Kovarianzen und Korrelationen der Regressionskoeffizienten in den Gruppen (SPSS-Ausgabe)

Varianzen, Kovarianzen und Korrelationen der Regressionsgewichte

a. für Gruppe der Panelausfälle:

Var-Covar Matrix of Regression Coefficients (B)
Below Diagonal: Covariance Above: Correlation

	KAND	INFL	UMFELD	DUTY	COST
KAND	.03469	-.20758	-.13259	-.02431	.27485
INFL	-.00855	.04887	-.11985	-.14118	.05493
UMFELD	-.00125	-.00134	.00258	-.29088	-.03127
DUTY	-2.895E-04	-.00199	-9.438E-04	.00409	.05189
COST	.00506	.00120	-1.570E-04	3.281E-04	.00978

b. für Gruppe der Panelteilnehmer:

Var-Covar Matrix of Regression Coefficients (B)
Below Diagonal: Covariance Above: Correlation

	KAND	INFL	UMFELD	DUTY	COST
KAND	.00694	-.17475	-.06003	-.06795	.01808
INFL	-.00174	.01432	-.11749	-.15726	.08913
UMFELD	-1.340E-04	-3.770E-04	7.187E-04	-.19862	.03648
DUTY	-1.981E-04	-6.589E-04	-1.864E-04	.00123	.21070
COST	7.710E-05	5.462E-04	5.007E-05	3.777E-04	.00262

Bei der getrennten Schätzung der Regressionskoeffizienten stellt sich die Frage, ob auch hier ein statistischer Test auf Gleichheit der Koeffizienten zwischen den Gruppen durchgeführt werden können. Solche Tests sind möglich, wenn man die Korrelationen bzw. Kovarianzen zwischen den geschätzten Regressionskoeffizienten in jeder Gruppe berücksichtigt. Aus diesem Grunde ist oben bei der SPSS-Anweisung der getrennten Schätzung in jeder Gruppe das Schlüsselwort BCOV in der Option STATISTICS der REGRESSION-Anweisung auf-

geführt. Mit diesem Schlüsselwort wird für jede Regressionsgleichung die Varianz/Kovarianzmatrix der geschätzten Koeffizienten ausgedrückt. Abbildung 1 zeigt die SPSS-Ausgabe dieser Matrizen für das Anwendungsbeispiel. Unter Benutzung dieser Statistiken kann die Methode des Wald-Tests angewendet werden, um die Gleichheit aller oder einiger Regressionskoeffizienten zu prüfen. Die Teststatistik W^2 berechnet sich nach:

$$W^2 = [\mathbf{b}_1 - \mathbf{b}_2]' [\mathbf{S}_{b_1} + \mathbf{S}_{b_2}]^{-1} [\mathbf{b}_1 - \mathbf{b}_2]$$

In der Gleichung steht \mathbf{b}_1 für den Vektor der zu vergleichenden Regressionskoeffizienten in der ersten Gruppe und \mathbf{b}_2 für den Vektor der entsprechenden Koeffizienten in der zweiten Gruppe. \mathbf{S}_{b_1} und \mathbf{S}_{b_2} stehen für zwei Matrizen, die die Varianzen und Kovarianzen der Regressionskoeffizienten der ersten bzw. zweiten Gruppe enthalten. Die Werte der Regressionskoeffizienten sind in Tabelle 8 angegeben, die Varianzen und Kovarianzen in Abbildung 1. Um den Test auszuführen, benötigt man ein Programm zur Matrixrechnung. In SPSS steht hierfür die Prozedur MATRIX zur Verfügung. Die Anweisungen zur Durchführung des Tests und das Ergebnis sind in Abbildung 2 wiedergegeben.

Mit der ersten Anweisung wird die Prozedur MATRIX aufgerufen. Danach werden mit vier COMPUTE-Anweisungen die Werte der Regressionskoeffizienten in die Vektoren bNoPanel für die Panelausfälle und bPanel für die Panelteilnehmer sowie die Varianzen und Kovarianzen in die Matrizen SNoPanel und SPanel eingetragen. Die Werte sind jeweils aus den SPSS-Ausdruck abgeschrieben bzw. in die Anweisungen kopiert. Bei den Varianz/Kovarianzmatrizen ist zu beachten, daß die SPSS-Ausgabe der Prozedur Regression- (Abbildung 1) oberhalb der Diagonale Korrelationen enthält. Die entsprechenden Kovarianzen müssen aus den entsprechenden Zellen unterhalb der Diagonale entnommen werden.

Bei der Eingabe von Elementen von Vektoren und Matrizen mit COMPUTE-Anweisungen ist zu beachten, daß die Elemente innerhalb einer Zeile durch jeweils ein Komma zu trennen sind. Ein Semikolon weist darauf hin, daß eine neue Zeile beginnt. Die Elemente eines Vektors oder einer Matrix müssen durch geschweifte Klammern ($\{, \}$) eingeschlossen werden.

Nach der Eingabe erfolgt die Berechnung der Teststatistik, der Freiheitsgrade und des empirischen Signifikanzniveaus. Mit der vorletzten Anweisung werden die Ergebnisse ausgedrückt und anschließend die Prozedur mit der Anweisung END MATRIX verlassen.

Abbildung 2: Berechnung des Wald-Tests für die lineare Regression mit der SPSS-Prozedur MATRIX.

```
matrix.
* Eingabe der Regressionskoeffizienten der ersten Gruppe in Vektor bNoPanel.
compute bNoPanel={-.491522; -.980174; -.182192; -.341713; .085125}.

* Eingabe der Varianz/Kovarianzmatrix der Koeffizienten in Matrix SNoPanel.
compute SNoPanel={
  .03469,    -.00855,    -.00125,  -2.895E-04,    .00506;
  -.00855,    .04887,    -.00134,   -.00199,    .00120;
  -.00125,   -.00134,    .00258,  -9.438E-04,  -1.570E-04;
  -2.895E-04, -.00199,  -9.438E-04,    .00409,    3.281E-04;
  .00506,    .00120,  -1.570E-04,    3.281E-04,    .00978}.

* Eingabe der Regressionskoeffizienten der zweiten Gruppe in Vektor bPanel.
compute bPanel={-.395702; -.596240; -.153959; -.379021; .109118}.

* Eingabe der Varianz/Kovarianzmatrix der Koeffizienten in Matrix SPanel.
compute SPanel={
  .00694,    -.00174,  -1.340E-04,  -1.981E-04,    7.710E-05;
  -.00174,    .01432,  -3.770E-04,  -6.589E-04,    5.462E-04;
  -1.340E-04, -3.770E-04,    7.187E-04,  -1.864E-04,    5.007E-05;
  -1.981E-04, -6.589E-04,  -1.864E-04,    .00123,    3.777E-04;
  7.710E-05,    5.462E-04,    5.007E-05,    3.777E-04,    .00262}.

* Berechnung von Teststatistik, Freiheitsgraden und Signifikanz.
compute chisqr=t(bNoPanel-bPanel)*INV(SNoPanel+SPanel)*(bNoPanel-bPanel).
compute df=nrow(bNoPanel).
compute prob=1-chicdf(chisqr,df).

* Ausgabe der Ergebnisse.
print {chisqr,df,prob} /Format F8.3 /clab="Chisqr" "DF" "Prob" .
end matrix.

{CHISQR,DF,PROB}
  Chisqr    DF    Prob
    3.613    5.000    .606
```

Wenn die Nullhypothese zutrifft, daß sich die Regressionskoeffizienten nicht unterscheiden, ist die Teststatistik asymptotisch chiquadratverteilt. Die Zahl der Freiheitsgrade ergibt sich aus der Anzahl der zu vergleichenden Koeffizienten. In Abbildung 2 werden die fünf Regressionsgewichte der beiden Gruppen aus Tabelle 8 verglichen. Die berechnete Waldstatistik von 3.613 ergibt bei fünf Freiheitsgraden ein empirisches Signifikanzniveau von 0.606: die Nullhypothese kann also nicht abgelehnt werden. Im Beispiel kommt der Wald-Test bei getrennter Schätzung der Regressionsgleichungen also zu dem gleichen Resultat wie der vorher durchgeführte F-Test. Zwischen den Panelteilnehmern und Panelausfällen scheint es keine signifikanten Unterschiede in der Beziehung der unabhängigen Variablen auf die sechsstufige Wahlbeteiligungsskala zu geben.

Mit den in Abbildung 2 abgedruckten SPSS-Anwendungen können beliebige Anwendungen des Tests durchgeführt werden. Anpassen sind nur jeweils die vier COMPUTE-Anweisungen, bei denen die Werte für die Regressionskoeffizienten und deren Varianzen und Kovarianzen einzutragen sind. Sollen mehr als fünf Koeffizienten verglichen werden, müssen entsprechend mehr Werte eingetragen werden. Bei einer geringeren Zahl von Koeffizienten reduziert sich die Zahl. Bei der Benutzung von SPSS ist zu beachten, daß die Reihenfolge der Elemente in der Varianz/Kovarianzmatrix aus Abbildung 1 nicht mit der Reihenfolge der Elemente bei der Ausgabe der Regressionskoeffizienten übereinstimmen muß.

Soll ein solcher Fall vermieden werden, muß in schrittweiser Regression die Reihenfolge jedes Koeffizienten explizit festgelegt werden.¹²

Mit dem Wald-Test können auch die Koeffizienten aus mehr als zwei Gruppen verglichen werden. Bei drei Gruppen ergibt sich folgende Teststatistik:

$$W^2 = \begin{bmatrix} \mathbf{b}_1 - \mathbf{b}_2 \\ \mathbf{b}_1 - \mathbf{b}_3 \end{bmatrix} / \begin{bmatrix} \mathbf{S}_{b_1} + \mathbf{S}_{b_2} & \mathbf{S}_{b_1} \\ \mathbf{S}_{b_1} & \mathbf{S}_{b_1} + \mathbf{S}_{b_3} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{b}_1 - \mathbf{b}_2 \\ \mathbf{b}_1 - \mathbf{b}_3 \end{bmatrix}$$

In analoger Weise kann die Formel für beliebig viele Gruppen erweitert werden. Der Wald-Test kann auch bei der logistischen Regression angewendet werden. In Abbildung 3 sind die entsprechenden SPSS-Anweisungen und das Ergebnis festgehalten. Im ersten Schritt wird wieder eine getrennte Schätzung der Regressionsmodelle in beiden Gruppen vorgenommen. Mit dem Schlüsselwort COR der Option PRINT wird die zusätzliche Ausgabe von Korrelationen der geschätzten Regressionskoeffizienten angefordert.¹³ In der ProzedurMATRIX müssen dann zusätzlich aus diesen Korrelationen und Standardfehlern die Kovarianzen berechnet werden. Bei dem Beispiel des Wald-Tests bei der logistischen Regression aus Abbildung 3 werden daher zwei zusätzliche COMPUTE-Anweisungen zur Eingabe der Standardfehler der Regressionskoeffizienten in die Zeilenvektoren SENop für die Panelausfälle und SEPan für die Panelteilnehmer benötigt. Die Teststatistik beträgt bei der logistischen Regression 1.101. bei fünf Freiheitsgraden kann die Nullhypothese, daß die Koeffizienten in den beiden Gruppen gleich sind, wiederum nicht abgelehnt werden (Prob.0.954).

Obwohl also wiederum die gleiche Vorgehensweise wie im linearen Fall möglich ist, ist bei der logistischen Regression eine getrennte Schätzung der Modellgleichungen in den beiden Gruppen eigentlich überflüssig. Da die logistische Regression grundsätzlich von heteroskedastischen Residuen ausgeht, sind hier nicht nur wie beim Vergleich von Tabelle 5 und 8 die

¹² SPSS gibt bei Verwendung der Prozedur REGRESSION leider nicht die Varianz und die Kovarianzen der Regressionskonstante aus. Mit einem kleinen Trick kann diese Einschränkung umgangen werden. Dazu muß mit der Option ORIGIN eine Regression durch den Ursprung angefordert werden. Die Regressionskonstante kann dann als zusätzlicher Prädiktor mit dem Wert "1" für alle Fälle in das Modell aufgenommen werden.

¹³ Im Unterschied zur Prozedur REGRESSION gibt LOGISTIC REGRESSION nur Korrelationen, aber keine Varianzen und Kovarianzen der Regressionskoeffizienten aus. Dafür werden aber auch die Korrelationen der Regressionskonstante mit den Regressionsgewichten ausgegeben. Um das Beispiel mit dem aus Abbildung 1 und 2 gleich zu halten, wird hier auf die Berücksichtigung der Konstanten im Wald-Test verzichtet.

geschätzten Regressionskoeffizienten gleich, sondern auch die Standardfehler.¹⁴ Im linearen Modell ist die gruppenspezifische Schätzung der Regressionskoeffizienten über mehrere Regressionsgleichungen dagegen sinnvoll, wenn es gruppenspezifische Unterschiede in den Residualvarianzen gibt. Wenn dann jedoch - wie in dem Beispiel aus Abbildung 2 - der Wald-Test den Schluß nahelegt, daß sich alle oder zumindest einige Regressionskoeffizienten zwischen den Gruppen nicht unterscheiden, ist es sinnvoll, für diese Koeffizienten eine gemeinsame Schätzung zu erhalten, die gleichwohl die Heteroskedastizität zwischen den Gruppen berücksichtigt.

Abbildung 3: Berechnung des Wald-Tests für die logistische Regression mit SPSS

```

sort cases by PANEL.
split file by PANEL.
logistic regression VOTE_TWO with KAND EINFL UMFELD NORM KOSTEN
  /print def cor.
split file off.
matrix.
compute bNoPanel={-.6138; -1.6562; -.5168; -1.2187; .6983}.
compute SENoP={.6032; .5980; .1637; .2682; .3467}.
compute RNoPanel={ 1.00000, -.25204, -.16858, -.01354, .32814;
                  -.25204, 1.00000, -.02903, .10019, -.14448;
                  -.16858, -.02903, 1.00000, .03557, -.15429;
                  -.01354, .10019, .03557, 1.00000, -.30811;
                  .32814, -.14448, -.15429, -.30811, 1.00000}.
compute SNoPanel=t(SENoP)*RNoPanel*SENoP.

compute bPanel={ -1.3062; -.8169; -.3671; -1.2176; .5563}.
compute SEPan={ .3600; .4267; .1129; .1642; .2247}.
compute RPanel={ 1.00000, -.16627, .03727, .13849, -.06786;
                 -.16627, 1.00000, -.16564, -.05882, .09389;
                 .03727, -.16564, 1.00000, -.03653, .01047;
                 .13849, -.05882, -.03653, 1.00000, .12085;
                 -.06786, .09389, .01047, .12085, 1.00000}.
compute SPanel=t(SEPan)*RPanel*SEPan.

compute chisqr=t(bNoPanel-bPanel)*INV(SNoPanel+SPanel)*(bNoPanel-bPanel).
compute df=nrow(bNoPanel).
compute prob=1-chicdf(chisqr,df).
print {chisqr,df,prob} /Format F8.3 /clab="Chisqr" "DF" "Prob" .
end matrix.

{CHISQR,DF,PROB}
  Chisqr      DF      Prob
    1.101     5.000     .954

```

¹⁴ Die Möglichkeit unterschiedlicher Residualvarianzen zwischen den Gruppen macht im logistischen Modell nur Sinn, wenn man das Modell als Schätzung der Koeffizienten einer linearen Gleichung einer unbeobachtbaren latenten abhängigen Variable interpretiert, die über ein Schwellenwertmodell durch einen dichotomen Indikator gemessen wird. Wenn sich bei dieser Sichtweise die Residualvarianzen der latenten abhängigen Variablen in den Gruppen unterscheiden, benötigt man komplexere statistische Modelle, die noch nicht standardmäßig verfügbar sind.

Im Prinzip gibt es hierzu mehrere Lösungen. Eine Lösung besteht darin, die getrennte Formulierung der Regressionsgleichungen in den Gruppen aufzugeben und doch eine gemeinsame Gleichung zu schätzen, bei der jedoch gruppenspezifische Residualvarianzen zugelassen werden. Möglich ist dies mit iterativen Algorithmen zur GLS- oder ML-Schätzung von Regressionsmodellen mit heteroskedastischen Fehlern. Falls die Heteroskedastizität der Residuen bei der Kleinstquadratschätzung der Regressionskoeffizienten ignoriert wird, können mit Hilfe der sogenannten White-Korrektur zumindest statistisch akzeptable Schätzungen der Standardfehler berechnet werden. Schließlich besteht auch die Möglichkeit, die Regressionsgleichungen für die Gruppen getrennt zu formulieren, aber eine gemeinsame Schätzung vorzunehmen, bei der alle oder einige Regressionskoeffizienten zwischen den Gruppen gleichgesetzt werden. Notwendig ist hierfür ein Algorithmus, der bei getrennter Formulierung von Regressionsgleichungen in den Gruppen eine gemeinsame Schätzung mit Restriktionen über die Gruppen hinweg erlaubt. Ein solcher Algorithmus wird als sogenannter simultaner Gruppenvergleich in Programmen zur Schätzung von Strukturgleichungsmodellen zur Verfügung gestellt.

Tabelle 9: Getrennte Schätzung von linearen Regressionsgleichungen im simultanen Gruppenvergleich
(AGLS-Schätzung mit LISREL 8, $\chi^2_{df=6} = 2.290$, Prob.=0.891)

	Panelteilnehmer		Panelausfälle	
	b	t	b	t
R ²	0.354		0.305	
Residualvarianz	0.767		1.186	
Kandidatenpräferenz	-0.407	-4.6	-0.407	-4.6
Einfluß der Stimme	-0.714	-5.2	-0.714	-5.2
Soziales Umfeld	-0.163	-5.4	-0.163	-5.4
Wahlbeteiligungsnorm	-0.371	-11.0	-0.371	-11.0
Beteiligungskosten	0.108	2.4	0.108	2.4
(Konstante)	3.094	16.8	3.094	16.8

In Tabelle 9 sind die Ergebnisse einer solchen Schätzung aufgeführt. Die Schätzung erfolgte mit dem Programm LISREL 8.¹⁵ Die verwendete AGLS-Schätzung nimmt im Unterschied zu den klassischen Annahmen der Inferenzstatistik in Regressionsmodellen keine Normal-

¹⁵ Aus Platzgründen kann an dieser Stelle nicht auf die Einzelheiten der Programm Benutzung eingegangen werden. Vergl. hierzu *Jöreskog/Sörbom* 1993.

verteilung der Residuen an. In der Schätzung sind alle Regressionsgewichte sowie die Regressionskonstanten zwischen den Gruppen gleichgesetzt. Für die Gruppen getrennt geschätzt sind dagegen die Residualvarianzen. Die Anpassung des Modells an die Daten kann durch einen Chiquadratstest beurteilt werden, der hier zu einem Wert von 2.290 bei 6 Freiheitsgraden führt. Das empirische Signifikanzniveau von 0.891 weist darauf hin, daß das Modell sehr gut auf die Daten paßt. Da die einzigen Restriktionen die gleichen Regressionskoeffizienten sind, ist dieser Anpassungstest gleichzeitig ein Test der Nullhypothese, daß die Koeffizienten tatsächlich nicht zwischen den Gruppen variieren. Diese Nullhypothese kann nicht abgelehnt werden. Vergleicht man die Koeffizienten aus Tabelle 9 mit denen aus Tabelle 2, zeigen sich geringfügige Unterschiede. Diese sind auf die unterschiedlichen Modellannahmen zurückzuführen. In Tabelle 2 wurde zusätzlich zu den Annahmen gleicher Regressionskoeffizienten in den Gruppen sowohl gleiche Residualvarianz zwischen den Gruppen wie die Normalverteilung der Residuen unterstellt. Die Unterschiede zwischen den Schätzungen sind allerdings nicht sehr groß, so daß die inhaltliche Interpretation sich nicht substantiell unterscheidet.

5. Diskussion

In Gruppenvergleichen wird untersucht, ob sich die Parameter eines statistischen Modells zwischen verschiedenen Gruppen unterscheiden. Anwendungsmöglichkeiten finden sich im interkulturellen oder internationalen Vergleich, aber auch im Vergleich von Gruppen innerhalb einer Gesamtpopulation oder bei einer Trendanalyse, bei der mit zeitlichen Abständen Stichproben aus einer Grundgesamtheit gezogen werden. Voraussetzung für die Anwendung der in diesem Beitrag vorgestellten Möglichkeiten des Gruppenvergleichs ist allein, daß die einzelnen Gruppen (Stichproben) und die Untersuchungseinheiten in den Stichproben jeweils unabhängig voneinander gezogen worden sind.

Die hier vorgestellten Analysen bilden keine eigenständige Methodologie des Gruppenvergleichs. Es handelt sich vielmehr um Anwendungen von Standardverfahren der Statistik. So kann ein Gruppenvergleich dadurch realisiert werden, daß die Gruppenzugehörigkeit als eine nominalskalierte unabhängige Variable in eine Regressionsgleichung aufgenommen wird. Die Auswirkungen der Gruppenzugehörigkeit können dann als Haupteffekt dieser zusätzlichen Variable oder auch als Interaktionseffekte mit den übrigen Variablen modelliert werden. Die unterschiedlichen Parametrisierungen als klassische Interaktionseffekte oder als konditionale Haupteffekte sind aus der Varianzanalyse bekannt.

Im linearen Modell besteht ein Nachteil des Gruppenvergleichs über die Modellierung als zusätzliche unabhängige Variable darin, daß wie in der Varianzanalyse die Annahme gleicher Residualvarianzen angenommen werden muß. Genaugenommen gilt diese Einschränkung allerdings nur für die einfache Kleinstquadratschätzung. In anderen Schätzverfahren können ungleiche Residualvarianzen in den Gruppen berücksichtigt werden. Es ist aber auch möglich, Regressionsgleichungen für jede Gruppe getrennt zu rechnen. Auch in

diesem Fall lassen sich statistische Tests auf Gleichheit der Regressionskoeffizienten durchführen. Diese Tests sind Anwendungen des klassischen Wald-Tests und basieren darauf, daß die geschätzten Modellparameter (zumindest bei großen Fallzahlen) normalverteilt sind. Komplexere Schätzverfahren sind nur dann notwendig, wenn bei einer getrennten Formulierung von Regressionsgleichungen in den einzelnen Gruppen Restriktionen über die Gruppen hinweg berücksichtigt werden sollen. Als Beispiel hierfür habe ich einen simultanen Gruppenvergleich mit LISREL durchgeführt.

Da die beim Gruppenvergleich eingesetzten statistischen Methoden sehr allgemein sind, sind sie nicht auf die klassische lineare Regression beschränkt. Um dies zu verdeutlichen, wurden die Analysen nicht nur bei einer linearen Regression, sondern auch bei einer logistischen Regression einer dichotomen abhängigen Variablen durchgeführt. Die gleiche Anwendungslogik läßt sich auch bei anderen Modellen anwenden, z.B. bei multinomialen Logitmodellen, bei Probit- oder Poissonregression, in log-linearen Modellen oder bei der Analyse von Ereignisdaten.

Wie bei allen formalen Methoden ist auch die Logik des Gruppenvergleichs offen gegenüber den inhaltliche Fragestellungen. Mein Anwendungsbeispiel bezieht sich auf den Vergleich der Teilnehmer einer Wiederholungsbefragung mit denjenigen Personen, die nur in der ersten, aber nicht in der zweiten Befragung nicht erreicht worden sind. Anhand von Daten aus der Erstbefragung wurde geprüft, ob sich die Regressionskoeffizienten bei einer Regression der Wahlbeteiligungsabsicht auf verschiedene Prädiktoren signifikant unterscheiden. Die Ergebnisse des Gruppenvergleichs weisen darauf hin, daß sich die beiden Gruppen nicht signifikant unterscheiden. Dies Ergebnis gilt für das lineare wie das logistische Modell und sowohl bei einer gemeinsamen Schätzung der Koeffizienten in einer Gleichung wie bei einer getrennten Schätzung in gruppenspezifischen Gleichungen. Der einzige auffällige Unterschied zwischen den Gruppen besteht darin, daß im linearen Modell die Residualvarianz der Panelteilnehmer geringer war als die der Panelausfälle.

Zumindest für das Anwendungsbeispiel läßt sich insgesamt die Schlußfolgerung ziehen, daß der Panelausfall die Beziehung zwischen abhängiger und unabhängiger Variable nicht stört. Selbst wenn es systematische Gründe dafür gibt, ob eine Person wiederbefragt werden kann oder nicht, so verzerren diese doch nicht die Ergebnisse der Analyse. Auch wenn eine Verallgemeinerung auf die Analyse anderer Beziehungen nicht begründet werden kann, läßt dieses Ergebnis doch zumindest den Schluß zu, daß Ausfälle nicht grundsätzlich und ausnahmslos jede Verallgemeinerung über einen Datensatz hinaus hinfällig machen.

Literatur:

Greene, William H. (1993) *Econometric Analysis*. 2. Auflage, New York: Macmillan.

Jöreskog, Karl G., and Sörbom, Dag (1993) *LISREL 8 User's Reference Guide*. Chicago: Scientific Software, Inc.

Kühnel, Steffen-M., Jagodzinski, Wolfgang und Terwey, Michael (1989) *Teilnehmen oder Boykottieren*. Ein Anwendungsbeispiel der binären logistischen Regression mit SPSSx. *ZA-Information* 25, S. 44-75.