

Wissenschaftliche Diskussionspapiere

Heft 38

**Göran Kauermann**

unter Mitarbeit von

Jutta Gampe

Oliver Jost

Lisa Pritscher

Norbert Behrens

Statistische Auswertung  
des BIBB-Projekts

**„Jugend- und Berufsausbildung in Deutschland“  
unter Berücksichtigung der Panelstruktur**

Herausgeber: Bundesinstitut für Berufsbildung - Der Generalsekretär

Die WISSENSCHAFTLICHEN DISKUSSIONSPAPIERE DES BIBB werden durch den Generalsekretär herausgegeben. Sie erscheinen als Namensbeiträge ihrer Verfasser und geben deren Meinung und nicht die des Herausgebers wieder. Sie sind urheberrechtlich geschützt. Ihre Veröffentlichung dient der Diskussion mit der Fachöffentlichkeit.



Der Inhalt dieses Werkes steht unter einer Creative Commons Lizenz (Lizenztyp: Namensnennung – Keine kommerzielle Nutzung – Keine Bearbeitung – 3.0 Deutschland).

Das Werk wird durch das Urheberrecht und/oder einschlägige Gesetze geschützt. Jede Nutzung, die durch diese Lizenz oder Urheberrecht nicht ausdrücklich gestattet ist, ist untersagt.

Weitere Informationen finden Sie im Internet auf unserer Creative Commons-Infoseite <http://www.bibb.de/cc-lizenz>

Vertriebsadresse:  
Bundesinstitut für Berufsbildung  
10702 Berlin

Copyright 1998 by Bundesinstitut für Berufsbildung, Berlin und Bonn  
Herstellung: Bundesinstitut für Berufsbildung, Berlin  
Umschlag: Hoch Drei, Berlin  
Druck: Bundesinstitut für Berufsbildung

Printed in Germany  
ISBN 3-88555-650-2



Gedruckt auf Recyclingpapier, hergestellt aus 100% Altpapier

Diese Netzpublikation wurde bei Der Deutschen Bibliothek angemeldet und archiviert.  
URN: [urn:nbn:de:0035-0111-8](http://nbn-resolving.org/urn:nbn:de:0035-0111-8)

## VORWORT

Das Forschungsprojekt „Jugend und Berufsausbildung in Deutschland“ zielte auf die Analyse der Lebens- und Bildungssituation sowie der Persönlichkeit von Jugendlichen in Ostdeutschland und deren Entwicklung im Verlauf der Berufsausbildung in einer Phase des gesellschaftlichen und persönlichen Umbruchs.

Im Vordergrund stand die Frage, wie Jugendliche ihre Lebenslage bewältigen und mit welchen Strategien sie ihrer Lebenssituation, gerade in den neuen Bundesländern, begegnen. Untersucht wurden daher Einstellungen, Haltungen, Wertungen und normative Strukturen von Jugendlichen zum Zeitpunkt der Berufswahl, im Ausbildungsverlauf, vor der Berufseinstellung und ein Jahr nach Ausbildungsabschluß. Ausgehend von unterschiedlichen Bedingungen auf dem Ausbildungsstellen- und Arbeitsmarkt galt es, die Erfahrungen von Jugendlichen in Ost- und Westdeutschland während des Qualifizierungsverlaufs zu untersuchen, Unterschiede und Veränderungen herauszuarbeiten und Auswirkungen des Qualifizierungsprozesses auf das Wertgefüge ost- wie westdeutscher Jugendlicher zu analysieren.

Diese Fragestellungen ließen sich verlässlich nur auf der Basis repräsentativer Wiederholungsbefragungen bei denselben Probanden beantworten. Für Themen, die die Entwicklung im Zeitablauf betreffen, wie beispielsweise die Frage nach der Entwicklung im Ausbildungsverlauf, war daher ein Vergleich der einzelnen Erhebungszeitpunkte erforderlich. Für die Auswertung von Panelstrukturen sind die Verfahren zur Analyse von Querschnittsdaten nicht zulässig, weil es sich dabei um sog. verbundene Stichproben handelt. Hier schied in der Regel auch die einfacheren Prozeduren für die Auswertung verbundener Stichproben aus, wie etwa der Test von Mc Nemar, weil die Antwortstrukturen einen hohen Grad an Komplexität aufwiesen. Zur Untersuchung der Panelstrukturen wurden vorzugsweise angewandt:

- Tests auf Quasi-Unabhängigkeit - zur Prüfung der Strukturen in Veränderungen zwischen zwei Zeitpunkten: Analyse der individuellen Veränderungen bei „Wechslern“ - also Jugendlichen mit unterschiedlichen Antworten - zu zwei Zeitpunkten.
- Tests auf marginale Homogenität - zur Untersuchung der Veränderungen zwischen zwei Zeitpunkten in Randverteilungen.
- Analyse von Markovketten - Erklärung der Antworten durch Antworten aus (den) Erhebungen zuvor.

Diese speziellen Untersuchungen verlangten einen spezifischen methodischen Hintergrund und konnten nicht mit der im BIBB verfügbaren Standardsoftware durchgeführt werden. Diese Analysen wurden daher im Auftragswege von einem Wissenschaftlerteam an der TU Berlin durchgeführt. Diese Arbeiten sind in dem vorliegenden Diskussionspapier dokumentiert. Mit diesem Diskussionspapier soll die Methodendiskussion angeregt werden. Auch kann das Diskussionspapier die noch folgenden Auswertungsberichte von Methodenfragen entlasten.

Berlin, im Oktober 1998

Dr. Klaus Schweikert

*Leiter der Abteilung 1.1*

*Sozialwissenschaftliche Grundlagen der Berufsbildung  
im Bundesinstitut für Berufsbildung*

# Inhaltsverzeichnis

<b>1</b>	<b>Statistische Grundüberlegungen</b>	<b>1</b>
1.1	Regressionsmodelle . . . . .	1
1.2	Modellselektion . . . . .	4
1.3	Modelle mit mehreren Regressoren . . . . .	6
1.4	Software . . . . .	7
<b>2</b>	<b>Longitudinale Daten</b>	<b>8</b>
2.1	Einfache Markov-Ketten . . . . .	8
2.2	Komplexe Markov-Ketten . . . . .	10
2.3	Software . . . . .	11
<b>3</b>	<b>Kontingenztafeln mit sequentieller Struktur</b>	<b>12</b>
3.1	Unabhängigkeit und Quasiunabhängigkeit . . . . .	12
3.2	Marginale Homogenität . . . . .	15
3.3	Software . . . . .	16
<b>4</b>	<b>Marginale Modelle</b>	<b>17</b>
4.1	Modellierung . . . . .	17
4.2	Software . . . . .	20
<b>5</b>	<b>Analyse des Drop-out Prozesses</b>	<b>22</b>
<b>6</b>	<b>Allgemeiner Einfluß der beruflichen Situation</b>	<b>37</b>
6.1	Gewaltbereitschaft der Jugendlichen . . . . .	37
6.2	Allgemeinempfinden . . . . .	39
<b>7</b>	<b>Beurteilung der beruflichen Situation</b>	<b>41</b>
7.1	Allgemeine Selbsteinschätzung . . . . .	41
7.2	Zufriedenheit mit Ausbildung . . . . .	47
7.3	Zufriedenheit mit Berufsschule . . . . .	54
7.4	Zufriedenheit mit Berufswahl . . . . .	59
7.5	Leben ohne Arbeit? . . . . .	64
7.6	Selbstzufriedenheit . . . . .	66

<b>8</b>	<b>Analyse der politischen Einstellung</b>	<b>69</b>
<b>9</b>	<b>Analyse der Wertvorstellungen</b>	<b>74</b>
9.1	Querschnittsmodellierung mit Einfluß von Geschlecht . . . . .	74
9.1.1	Querschnittsmodellierung mit Einstellung zu Türken. . . . .	82
9.1.2	Querschnittsmodellierung mit Einstellung zur Politik . . . . .	90
9.2	Markovanalyse. . . . .	100
<b>10</b>	<b>Mobilitätsbereitschaft</b>	<b>107</b>
10.1	Marginale Modelle. . . . .	111
10.2	Markov Analyse. . . . .	114

1

Im Rahmen des BIBB-Projekts 1.5009 wurden im Jahr 1992 bei Jugendlichen in den neuen Bundesländern zum Abschluß ihrer allgemeinen Schulausbildung eine Befragung durchgeführt. Ziel der Untersuchung war es unter anderem, Informationen über Stellenwert von Beruf und Berufsausbildung, Mobilitätsbereitschaft, Wertgefühl und allgemeine Lebensvorstellungen der Jugendlichen zu bekommen. Um zeitliche Veränderungen aufzuspüren, wurden die Jugendlichen erneut in den Jahren 1993, 1994 und 1995/96 befragt. Somit ergibt sich ein Datenpanel, was der vorliegenden statistischen Analyse als Grundlage diene.

In den Auswertungen wurde die Panelstruktur stets explizit mit berücksichtigt. Dies verlangte komplexe und teilweise sehr neue statistische Methoden. Um die Verständlichkeit der Arbeit zu erhöhen, sind die benutzten Methoden in Teil A dieser Auswertung kurz beschrieben (Kapitel 1-4). Dabei wird weitestgehend auf komplizierte Notation verzichtet und stattdessen eine einfache Motivation versucht. Hauptaugenmerk liegt auf der Beschreibung der resultierenden Interpretationen, wobei Beispiele hier das Verständnis erleichtern sollen. Der zweite Teil dieser Auswertung (Kapitel 5-10) beschäftigt sich dann mit der Analyse des zugrunde liegenden Datensatzes. Ergebnisse sind, wenn möglich, durch Graphiken begleitet. Die Interpretation der Daten wird primär unter statistischen Gesichtspunkten behandelt. Substanz und/oder sozialwissenschaftliche Ansätze werden nur am Rande erwähnt, weil diese nicht als Hauptaugenmerk einer statistischen Auswertung erscheinen sollen.

Schwerpunktmäßig wird den folgenden Fragen nachgegangen:

- Allgemeine Zufriedenheit und Gewaltbereitschaft der Jugendlichen (Kapitel 6)
- Beurteilung und Erfahrungen mit der beruflichen Situation (Kapitel 7)
- Beschreibung der politischen Einstellung der Jugendlichen (Kapitel 8)
- Analyse der Wertvorstellungen der Jugendlichen im gesellschaftlichen und politischen Kontext (Kapitel 9)
- Mobilitätsbereitschaft der Jugendlichen (Kapitel 10)

In allen Punkten wird versucht, adäquate Modelle zu finden, um die in den Daten gefundenen Strukturen anschaulich zu beschreiben. Wegen der zeitlichen Struktur der Daten steht natürlich die Analyse von zeitlicher Variation im Vordergrund.

Um die Gültigkeit der Auswertungen zu rechtfertigen, wird ferner eine Analyse des "Drop out" Prozeß betrachtet (Kapitel 5). Es zeigt sich, daß Jugendliche rein zufällig (missing completely at random) aus dem Panel scheiden, d.h. ab einem gewissen Zeitpunkt nicht mehr zur Verfügung stehen. Dieser "Drop Out" Prozeß rechtfertigt, aus statistischer Sicht, die benutzten Methoden.

# **TeilA**

Beschreibung der Methoden



# 1 Statistische Grundüberlegungen

In den nachfolgenden Abschnitten sollen die in der Analyse des Projekts benötigten statistischen Grundüberlegungen vorgestellt und durch Beispiele illustriert werden. Der Schwerpunkt der Betrachtung liegt dabei auf statistischen Verfahren, die die Panelstruktur der Daten des Projekts berücksichtigen. Zuvor sollen jedoch einfache statistische Methoden vorgestellt werden.

## 1.1 Regressionsmodelle

Zur allgemeingültigen Darstellung der verwendeten statistischen Methoden ist es notwendig, eine mathematische Notation einzuführen. Daher sei im folgenden mit  $Y$  eine interessierende Fragestellung bezeichnet, mit  $Y_i$  notieren wir die Antwort des  $i$ -ten befragten Individuums. Neben der interessierenden Fragestellung werden erklärende Größen  $X$ , sogenannte Regressoren betrachtet, wobei mit  $X_i$  die zugehörigen Größen des  $i$ -ten Individuums gemeint sind. Ziel von Regressionsmodellen ist es, die Variable  $Y_i$  aus den Regressoren  $X_i$  zu erklären, wobei der nicht erklärte Fehler als Residuum bezeichnet wird. Ist  $Y_i$  metrisch skaliert (d.h. nimmt  $Y_i$  stetige Werte an), so lautet die klassische Regressionsgleichung (siehe Schlittgen, 1996, Seite 413)

$$Y_i = \beta_0 + X_i\beta_X + \varepsilon_i.$$

Dabei wird  $\beta_0$  als Achsenabschnitt,  $\beta_X$  als Regressionskoeffizient zu  $X_i$  und  $\varepsilon_i$  als Residuum bezeichnet. Die Grundidee der Regression ist, die Werte  $Y_i$  in eine systematische (oder erklärbare) Komponente, nämlich  $\beta_0 + X_i\beta_X$ , und eine zufällige (oder nicht erklärbare) Komponente, das Residuum  $\varepsilon_i$  zu zerlegen. Für das Residuum  $\varepsilon_i$  wird gefordert, daß es im Mittel den Wert 0 annimmt. Das bedeutet inhaltlich, daß das Residuum keine systematischen Anteile enthält. Diese Eigenschaft soll im folgenden besonders beschrieben werden. Hierzu wird der (im Mittel) zu erwartende Wert, der sogenannte Erwartungswert definiert. Er wird mit dem Symbol  $E(\cdot)$  bezeichnet, das heißt  $E(\varepsilon_i|X_i)$  ist der (mittlere) zu erwartende Wert des Residuums  $\varepsilon_i$  bei gegebenen Regressoren  $X_i$ . Die explizite Auflistung von  $X_i$  hinter dem Senkrechtrich bedeutet, daß  $X_i$  selber als gegeben bzw. bekannt betrachtet wird. Insbesondere ist somit **nicht** der mittlere Wert von  $X_i$  modelliert, sondern der mittlere Wert des Residuums  $\varepsilon_i$ . Wird nun gefordert, daß das Residuum, also die nicht erklärte Komponente, im Mittel den Wert 0 hat, so läßt sich dies mit  $E(\varepsilon_i|X_i) = 0$  ausdrücken. Damit ergibt sich

$$\begin{aligned} E(Y_i|X_i) &= E(\beta_0 + X_i\beta_X + \varepsilon_i|X_i) \\ &= \beta_0 + X_i\beta_X + E(\varepsilon_i|X_i) \\ &= \beta_0 + X_i\beta_X, \end{aligned}$$

d.h. der mittlere Wert von  $Y_i$  ist die systematischen Komponente. Somit kann ein Regressionsmodell generell als die folgende Gleichung verstanden werden:

$$E(Y_i|X_i) = \beta_0 + X_i\beta_X. \quad (1)$$

Es zeigt sich jedoch, daß diese Modellierung Schwächen aufweist, insbesondere dann, wenn  $Y_i$  nicht metrisch, sondern beispielsweise binär ist. In diesem Falle ist  $Y_i$  entweder 0 oder 1

und der mittlere Wert von  $Y_i$  liegt somit zwischen 0 und 1, d.h.  $0 \leq E(Y_i|X_i) \leq 1$ . Damit ist aber der Wertebereich der linken Seite in (1) nicht zwangsläufig gleich dem Wertebereich der rechten Seite in (1), der nicht beschränkt ist. Dieser Nachteil läßt sich durch die Einführung von Transformationsfunktionen beseitigen, die linken und rechten Wertebereich in (1) angleichen. Ist  $Y_i$  binär, so wählt man gewöhnlicherweise die logit-Funktion mit

$$\text{logit}(p) = \log(p/(1-p)) \quad \text{für } 0 \leq p \leq 1.$$

Der Quotient  $p/(1-p)$  wird auch als "odds" (Wettchance) bezeichnet. Der Wertebereich der logit-Funktion ist unbeschränkt, so daß wir (1) überführen zu

$$\text{logit}(E(Y_i|X_i)) = \beta_0 + X\beta_X. \quad (2)$$

Die resultierende Regression wird auch als **logistische Regression** bezeichnet. Die Transformationsfunktion ist auch als Link-Funktion bekannt. Invertiert man beide Seiten der Gleichung (2) so ergibt sich

$$E(Y_i|X_i) = \frac{\exp(\beta_0 + X\beta_X)}{1 + \exp(\beta_0 + \beta_X)}$$

Man beachte, daß  $E(Y_i|X_i)$ , also der mittlere Wert von  $Y_i$ , hier auch als Wahrscheinlichkeit bzw. (modellerte) relative Häufigkeit von  $Y_i = 1$  angesehen werden kann.

*Beispiel:*

Wir betrachten die Fragestellung  $Y$ : „Politik kann jemand wie ich überhaupt nicht beeinflussen“ zum Zeitpunkt 1 mit den möglichen Ausprägungen  $Y_i = 0$  für „stimme zu“ und  $Y_i = 1$  für „stimme nicht zu“. Als erklärende Größe betrachten wir das Geschlecht, mit  $X_i = 1$  für „männlich“ und  $X_i = 2$  für „weiblich“. Die zugrunde liegenden Daten sind gegeben in Tabelle 1. Wir wählen das Logitmodell (2) und erhalten die Parameterschätzung  $\beta_0 = -0,02$  und  $\beta_1 = -0,12$ . Es ergibt sich somit für  $X_i = 1$ , d.h. für Männer,  $\exp(-0,02 - 0,12)/\{1 + \exp(-0,02 - 0,12)\} = 0,466$ . Das bedeutet, 46,6% der Männer stimmen der Aussage nicht zu, wohingegen nur 43,6% der Frauen der Aussage nicht zustimmen.

Geschlecht	Politik kann jemand wie ich nicht beeinflussen!		
	Stimme zu	Stimme nicht zu	
männlich	609 53,4%	531 46,6%	100%
weiblich	478 56,4%	369 43,6%	100%

Tabelle 1: Zustimmung zur Aussage "Politik kann jemand wie ich nicht beeinflussen"

Bei Regressionsmodellen dieser Art ist insbesondere der Fall interessant, wenn  $Y_i$  ordinal skaliert ist, d.h.  $Y_i$  habe beispielsweise die Ausprägungen  $Y_i = 1$  für schlecht,  $Y_i = 2$  für mittel und  $Y_i = 3$  für gut. Nun modelliert man

$$\log \frac{P(Y_i \leq l)}{1 - P(Y_i \leq l)} = \beta_{0l} + x_i \beta_x$$

wobei  $j = 1, 2, \dots$  die möglichen Ausprägungen von  $Y$  wiedergibt. Somit haben die Regressoren  $X_j$  für alle Kategorien von  $Y_j$  den gleichen Einfluß. Je Kategorie gibt es jedoch einen unterschiedlichen Achsenabschnitt  $\beta_{0j}$ . Modelle dieser Arte werden auch als (*proportionale*) *kumulative log odds Modelle* bezeichnet.

Für kumulative logit Modelle läßt sich eine weitere Interpretation heranziehen. Ist  $Y_i$  eine ordinal gemessene Größe, so hat diese beispielsweise die Ausprägungen "schlecht"=1, "mittel"=2 und "gut"=3. Damit bedeutet  $Y_i \leq 1$  eine schlechte Ausprägung, wohingegen  $Y_i \leq 2$  eine mittlere bis schlechte Ausprägung bedeutet. Man kann nun annehmen, daß  $Y_i$  die Zusammenfassung einer zugrundeliegenden latenten stetigen Größe  $Z_i$  ist, derart daß

$$\begin{aligned} Z_i \leq a_1 &\Leftrightarrow Y_i \leq 1 \\ Z_i \leq a_2 &\Leftrightarrow Y_i \leq 2 \end{aligned}$$

mit  $a_1 < a_2$ . Somit kann  $Z_{ij}$  als latenter Score angesehen werden und die Schwellenwerte  $a_1$  und  $a_2$  zur Kategorisierung dienen.

*Beispiel:*

Die Jugendlichen sind in der ersten Periode gefragt worden, wie sie sich zur Zeit fühlen, wobei ihnen die Antwortmöglichkeiten "gut", "brauchbar", "so gut wie andere" und "schlechter" zur Verfügung standen. Wir wollen die Unzufriedenheit der Jugendlichen messen und nehmen an, daß jeder Jugendliche einen "Unzufriedenheitsscore"  $Z_i$  besitzt. Übersteigt dieser Score gewisse Werte, so fühlt sich der Jugendliche eher unzufrieden. Man modelliert den Score als "logit verteilt" und erhält aus den Daten von Tabelle 2, die Schwellenwerte

$$\begin{aligned} Z_i \leq -2,44 &\Leftrightarrow Y_i = \text{"gut"} \\ -2,44 < Z_i \leq 0,99 &\Leftrightarrow Y_i = \text{"so gut wie andere"} \\ 0,99 < Z_i \leq 2,75 &\Leftrightarrow Y_i = \text{"brauchbar"} \\ 2,75 < Z_i &\Leftrightarrow Y_i = \text{"schlechter"} \end{aligned}$$

Abbildung 1 visualisiert den Score, die Fläche der jeweils schraffierten Abschnitte entspricht den relativen Häufigkeiten aus Tabelle 2.

Kategorie	relative Häufigkeit
"schlechter"	0,06
"brauchbar"	0,21
"so gut wie andere"	0,65
"gut"	0,08

Tabelle 2: Relative Häufigkeiten für "Wie fühlen sie sich"

In Fortführung von (2) werden im folgenden unter Regressionsmodellen allgemeine Modellierungen gemäß

$$h(E(Y_i|X_i)) = \beta_0 + X_i\beta_X \tag{3}$$

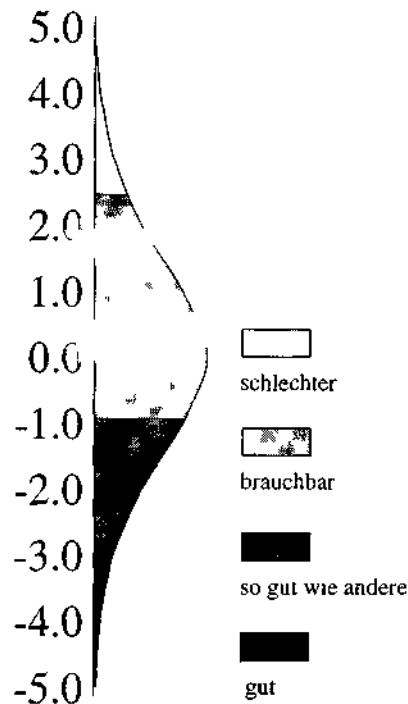


Abbildung 1: Latenter Score für "Wie fühlen sie sich?"

verstanden, wobei  $h$  als eine geeignete Transformationsfunktion (bzw. Link-Funktion) gewählt ist. Modelle dieser Art werden auch als **verallgemeinerte lineare Regressionsmodelle** (generalized linear models) bezeichnet (siehe McCullagh & Neider, 1989).

Sind mehr als eine erklärende Größe gegeben, d.h. werden beispielsweise zwei Regressoren  $X_{i1}$  und  $X_{i2}$  betrachtet, so ergeben sich Modelle gemäß

$$h(E(Y_i | X_{i1}, X_{i2})) = \beta_0 + X_{i1}\beta_1 + X_{i2}\beta_2 + X_{i1}X_{i2}\beta_{12} \quad (4)$$

wobei  $\beta_1$  den Einfluß von Regressor  $X_{i1}$  widerspiegelt,  $\beta_2$  entsprechend den Einfluß von  $X_{i2}$  und  $\beta_{12}$  den sogenannten Interaktionseffekt wiedergibt. Um die Darstellung übersichtlich zu halten, wird jedoch auf explizite Aufschlüsselung der einzelnen Regressoren im folgenden verzichtet, d.h. Modelle (4) und (3) werden notationmäßig als äquivalent angesehen.

## 1.2 Modellselektion

Im obigen Abschnitt wurden verschiedene Regressionsmodelle vorgestellt. Jetzt gilt es ein adäquates Modell auszuwählen. Adäquat heißt hier, sich auf die signifikanten Parameter zu beschränken. Exemplarisch stellt sich in einem Modell mit einem Regressor  $X_{i1}$  die Frage, ob der Einfluß von  $X_{i1}$  relevant ist, d.h. ob im Modell  $h(E(Y_i | X_{i1})) = \beta_0 + X_{i1}\beta_1$  der Parameter  $\beta_1$  signifikant ist. Ist mehr als ein Regressor gegeben, so ist zu untersuchen, welche Regressoren relevant sind und welche Interaktionen zwischen den Regressoren bestehen. Das heißt im Modell (4) ist zu prüfen, ob der Parameter  $\beta_{12}$  relevant ist (Interaktion) und ob  $\beta_1$  oder  $\beta_2$  relevant sind.

Zur Beantwortung dieser Frage steht aus statistischer Sicht mehr als nur eine Möglichkeit zur Verfügung. Der Regressionskoeffizient  $\beta x$  im Modell (3) wird aus den vorhandenen Daten geschätzt und der geschätzte Wert wird unter Berücksichtigung seiner Varianz als Konfidenzintervall angegeben. Überdeckt dieses Konfidenzintervall nicht den Wert 0, so gilt der Parameter als signifikant von 0 verschieden. Er darf also im Modell nicht vernachlässigt werden. Überdeckt das Konfidenzintervall dagegen den Wert 0, so kann die Hypothese  $H_0 : \beta x = 0$  nicht verworfen werden und das Modell  $h(E(Y_i|X_i)) = \beta_0 + X_i\beta x$  kann vereinfacht werden zu  $h(E(Y_i|X_i)) = \beta_0$ , d.h.  $Y_i$  ist nicht signifikant vom Regressor  $X_i$  abhängig.

Eine zweite Möglichkeit zur Beurteilung der Signifikanz eines Parameters ergibt sich mit dem Prinzip des Likelihood-Quotienten Tests. Sind an  $n$  Individuen die Beobachtungen  $Y_1, \dots, Y_n$  erhoben worden, so können diese retrospektiv mit dem Modell 1:  $h(E(Y_i|X_i)) = \beta_0$  und dem Modell 2:  $h(E(Y_i|X_i)) = \beta_0 + X_i\beta x$  vorhergesagt werden. Man berechnet nun den Quotienten:

$$\lambda = \frac{Ws(Y_1, \dots, Y_n | \text{Modell 1})}{Ws(Y_1, \dots, Y_n | \text{Modell 2})} \quad (5)$$

Dabei wird die Wahrscheinlichkeit  $Ws(\cdot)$  jeweils unter Modell 1 und Modell 2 berechnet. Offensichtlich ist Modell 2 komplexer als Modell 1 und es zeigt sich, daß

$$Ws(Y_1, \dots, Y_n | \text{Modell 1}) < Ws(Y_1, \dots, Y_n | \text{Modell 2}),$$

d.h. Modell 2 beschreibt die Daten besser. Ist Modell 1 nun nur unwesentlich schlechter als Modell 2, so ist der Quotient (5) nahe 1; ist dagegen Modell 1 nicht passend, so tendiert der Quotient nach 0. Die exakte Verteilung von  $\lambda$  ist kompliziert, klassische Theorie der Statistik erlaubt jedoch die Herleitung der approximativen Verteilung von  $LR = -2 \ln \lambda$  (siehe Schlittgen, 1996). Die Größe  $LR$  wird auch als log-Likelihood-Ratio bzw. log-Likelihood-Quotient bezeichnet. Ist  $Y$  eine kategoriale Größe und  $X$  ebenfalls kategorial, so ist der Likelihood Quotienten Test (asymptotisch) äquivalent zum klassischen  $\chi^2$ -Unabhängigkeitstests (vgl. Schlittgen, 1996, S. 400). Klassische Verteilungstheorie der Statistik erlaubt es, aus dieser Idee einen Test zu entwickeln, den sogenannten Likelihood-Quotienten Test (siehe Schlittgen, 1996).

*Beispiel: Fortsetzung*

*Es stellt sich in obigem Beispiel die Frage, ob sich die Einstellung zur Aussage: „Politik kann jemand wie ich überhaupt nicht beeinflussen“ signifikant bei Frauen und Männern unterscheidet. Dazu betrachten wir das Konfidenzintervall der Schätzung von  $\beta x$ . Es ergeben sich die Werte in Tabelle 3. Das Konfidenzintervall lautet somit  $-0,122 \pm 1,96 \cdot 0,091$ , d.h.  $KI(\beta x) = [-0,29; 0,06]$ . Das bedeutet aber, daß wir die Hypothese  $H_0 : \beta x = 0$  nicht verwerfen können, und somit der Geschlechtseffekt als nicht signifikant angesehen werden muß. Betrachtet man den Likelihood Quotienten (5) so ergibt sich ein Wert von 0,41 d.h. wir erhalten  $LR = 1,782$ . Es ist nun  $LR$  nach klassischer Verteilungstheorie  $\chi^2$ -verteilt mit einem Freiheitsgrad, so daß wir einen  $p$ -Wert (empirisches Signifikanzniveau) errechnen können. Es ergibt sich der  $p$ -Wert 0,18, der ebenfalls nicht signifikant ist. Dies deutet daraufhin, daß kein signifikanter Unterschied zwischen den Geschlechtern besteht. Berechnet man die Schätzer im einfachen Modell  $h(E(Y_i|X_i)) = \beta_0$ , so ergibt sich als geschätzter Wert für  $\beta_0 = -0,189$ . Die zugehörigen Wahrscheinlichkeiten findet man in Tabelle 4. Berechnet man weiter aus Tabelle 1 den  $\chi^2$ -Wert (vgl. Schlittgen, 1996, Seite 400), so ergibt sich man  $\chi^2 = 1,781$ . Beim Vergleich mit der zugehörigen  $\chi^2$ -Verteilung erhält man den  $p$ -Wert 0,18. Offensichtlich stimmen Likelihood Quotienten Test und  $\chi^2$ -Unabhängigkeitstest hier annähernd überein.*

Modell	Parameter	Schätzung	Standardabweichung
$h(E(Y_i X_i)) = \beta_0 + X_i\beta_x$	$\beta_0$	-0,015	0,138
	$\beta_x$	-0,122	0,091
$h(E(Y_i X_i)) = \beta_0$	$\beta_0$	-0,189	

Tabelle 3: Parameterschätzungen

Politik kann jemand wie ich nicht beeinflussen	
stimme zu	stimme nicht zu
1087	900
54,7%	45,3%

Tabelle 4: Relative Häufigkeiten

Grundsätzlich gibt es zwei Vorgehensweisen bei der Selektion von Modellen. Man kann einerseits mit einem komplexen Modell starten und das Modell soweit vereinfachen, bis alle verbleibenden Parameter signifikant sind. Ein derartiges Vorgehen wird als Rückwärtsselektion bezeichnet. Eine derartige Strategie wurde im obigen Beispiel durchgeführt. Alternativ kann man mit einem einfachen Modell beginnen und dieses solange erweitern, bis die zusätzlichen Parameter nicht mehr signifikant erscheinen. Diese Herangehensweise wird als Vorwärtsselektion bezeichnet.

### 1.3 Modelle mit mehreren Regressoren

Stehen bei Modellen mehr als eine Einflußgröße zur Verfügung, so stellt sich die Frage, inwieweit die Regressoren signifikanten Einfluß haben und ob signifikante Interaktionen zwischen den Einflußgrößen vorliegen. Um auch hier eine allgemeingültige Darstellung der Modelle zu bekommen, empfiehlt es sich auf die Modellnotation zurückzugreifen, wie sie von McCullagh & Neider (1989) vorgeschlagen wurde. Liegen beispielsweise zwei Regressoren  $X_1$  und  $X_2$  vor, so können diese jeweils additiv oder interaktiv wirken. Das heißt, es stehen die möglichen Modelle wie in Tabelle 5 gegeben zur Verfügung. Dabei notieren wir das interaktive Modell, d.h. das Modell in dem sowohl Regressor  $X_1$  als auch Regressor  $X_2$  samt einem interaktiven Einfluß wirken, mit  $1 \odot 2$ . Sind  $X_1$  und  $X_2$  binär (d.h. 0,1 wertig) so läßt sich der interaktive Effekt anschaulich mit Abbildung 2 beschreiben. Wirken die Regressoren dagegen nur additiv, so notieren wir dies mit  $1 \oplus 2$ . In diesem Fall ist der Regressionskoeffizient  $\beta_{12}$ , welcher die Interaktion beschreibt, gleich 0 modelliert. Die Modelle können nun weiter vereinfacht werden, indem der Einfluß von  $X_1$  oder  $X_2$  vernachlässigt wird.

	Modell	Notation
interaktiv	$E(Y X_1, X_2) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{12} X_1 X_2$	1 @ 1
additiv	$E(Y X_1, X_2) = \beta_0 + \beta_1 X_1 + \beta_2 X_2$	1 0 2
	$E(Y X_1) = \beta_0 + \beta_1 X_1$	1
	$E(Y X_2) = \beta_0 + \beta_2 X_2$	2

Tabelle 5: Modellnotation

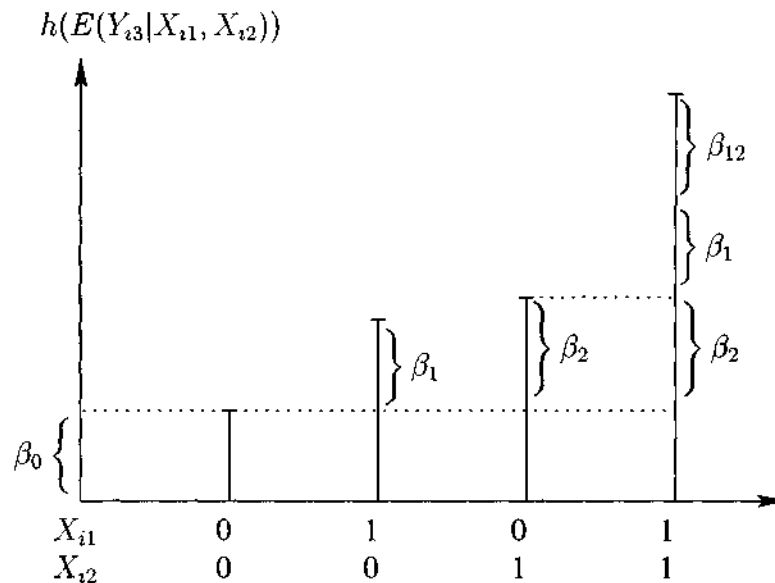


Abbildung 2: Skizze zur Beschreibung eines interaktiven Effekts

## 1.4 Software

Regressionsmodelle sind in allen gängigen Statistik Softwarepaketen enthalten, so auch in SPSS®, SPLUS©und SAS®. Alle drei Pakete wurden zur Analyse angewandt. Eine automatische Modellselektion ist in keinem Paket implementiert. Diese muß somit „per Hand“ durchgeführt werden, wobei je nach Fragestellung auf Rückwärts- oder Vorwärtsselektion zurückgegriffen wird.

## 2 Longitudinale Daten

Werden an einem Individuum mehrere Messungen zu unterschiedlichen Zeitpunkten vorgenommen, so werden die resultierenden Daten als *longitudinal* bezeichnet. Synonyme Begriffe hierfür sind *Panel Daten* oder *Daten mit Meßwiederholung*. Eine derartige Datenstruktur erfordert besondere Auswertungstechniken, weil die Daten, die an einem Individuum erhoben werden, als abhängige Größen angesehen werden müssen. Eine longitudinale Datenstruktur verbietet es insbesondere, klassische statistische Methoden anzuwenden, da Inferenzargumente ohne Berücksichtigung der Abhängigkeitsstruktur nicht greifen.

Wir bezeichnen mit  $Y_{it}$  die Messung an einem Individuum  $i$  zum Zeitpunkt  $t = 1, 2, 3$ , beispielsweise ist  $Y_{it}$  die Antwort auf eine Frage im  $t$ -ten Fragebogen, erhoben am  $i$ -ten Individuum. Wir betrachten  $Y_{it}$  als abhängige Variable oder auch Responsevariable, welche aus den Einflußgrößen oder auch Regressoren  $X_{it}$  erklärt wird. Dabei kann es sich um zeitkonstante oder zeitvariable Regressoren handeln. Als Beispiel nehme man das Geschlecht eines Individuums als zeitkonstanten Regressor, die Gewaltbereitschaft eines Individuums zum Zeitpunkt  $t$  als zeitvariable erklärende Größe. Im folgenden werden einige statistischen Methoden und Modelle zur Analyse von longitudinalen Datenstrukturen vorgestellt.

Dabei unterscheidet man generell zwei verschiedene Modellierungsansätze. Der erste Ansatz beinhaltet individuen-spezifische Methoden, bei denen die zeitliche Veränderung der Messungen eines jeden Individuum betrachtet und entsprechend modelliert wird. Der zweite Ansatz ist eine Querschnittsanalyse, bei denen die Veränderung der durchschnittlichen Ausprägung einer Messung in ihrer zeitlichen Struktur betrachtet wird.

### 2.1 Einfache Markov-Ketten

Markov-Ketten (ebenfalls bekannt als graphische Modelle, stochastische Prozesse oder sequentielle Regressionsmodelle) sind die Grundlage von Modellierungen, wie sie in der klassischen Zeitreihenanalyse vorkommen. Neuere Techniken (Lauritzen & Wermuth, 1989) beschäftigen sich dabei insbesondere mit der Darstellung von Markov-Ketten in Form eines gerichteten Graphen. Dieser Ansatz wird im folgenden vorgestellt. Er erlaubt die einfache Motivation der Strukturen und der Eigenschaften einer Markov-Kette.

Zunächst sollen der Übersichtlichkeit halber die erklärenden Größen  $X_{it}$  vernachlässigt werden. Betrachtet werden nur die Beobachtungen  $Y_{it}$  für das  $i$ -te Individuum zu den Zeitpunkten  $t = 1, 2, 3$ . Man nimmt an, daß die Messung  $Y_{i2}$  zum Zeitpunkt 2 abhängt von  $Y_{i1}$ , der zugehörigen Messung zum Zeitpunkt 1. Zum Zeitpunkt 3 erscheint  $Y_{i3}$  abhängig von den Messungen  $Y_{i1}$  und  $Y_{i2}$  in den Vorperioden. Es ergibt sich eine klare sequentielle Struktur, die durch den Graphen in Abbildung 3 visualisiert wird. Eine interessierende Fragestellung ist nun, ob die Abhängigkeit der Messung zum Zeitpunkt 3 von der Beobachtung zum Zeitpunkt 1 signifikant ist oder ob diese Abhängigkeit vernachlässigt werden kann. In diesem Fall hängt das Ergebnis  $Y_{i3}$  nur noch von der Beobachtung  $Y_{i1}$ , jedoch nicht von  $Y_{i2}$  ab. Man sagt in diesem Fall die Markov-Kette hat die Ordnung 1. Der Begriff Ordnung gibt also an, wie weit zur Regression des Wertes im Zeitpunkt  $t$  auf davor liegende Vergangenheitswerte zurückgegriffen werden



muß. Eine Markov-Kette der Ordnung 1 wird als Graph gemäß Abbildung 4 dargestellt. Wohingegen die Markov-Kette in Abbildung 3 die Ordnung 2 aufweist. Ziel der Markov-Analyse ist es nun, die Abhängigkeitsstruktur auf die signifikante Struktur zu reduzieren, d.h. insbesondere zu prüfen, ob die Struktur in Abbildung 3 im Vergleich zu Abbildung 4 aufrechterhalten werden muß.

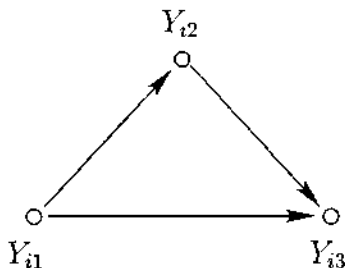


Abbildung 3: Markov-Kette mit Ordnung 2

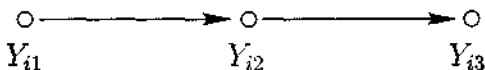


Abbildung 4: Markov-Kette mit Ordnung 1

Betrachtet man das anzuwendende statistische Repertoire, so zeigt sich, daß eine Markov-Kette als sequentielle Regression begriffen werden kann. Das heißt man modelliert

$$\begin{aligned} h(E(Y_{i2}|Y_{i1})) &\approx \beta_{02} + Y_{i1}\beta_{12} \\ h(E(Y_{i3}|Y_{i1}, Y_{i2})) &= \beta_{03} + Y_{i1}\beta_{13} + Y_{i2}\beta_{23} \end{aligned} \quad (6)$$

wobei in der zweiten Zeile keine Interaktion zwischen  $Y_n$  und  $Y_{l2}$  berücksichtigt wurde. Modelliert man diese mit, so erhält man

$$h(E(Y_{i3}|Y_{i1}, Y_{i2})) = \beta_{03} + Y_{i1}\beta_{13} + Y_{i2}\beta_{23} + Y_{i1}Y_{i2}\beta_{123}. \quad (7)$$

Die Schätzungen der Parameter werden nun bzgl. ihrer Signifikanz betrachtet, ähnlich wie dies in Kapitel 1.2 geschehen ist. Dabei werden grundsätzlich zwei Fragestellungen verfolgt. Einerseits wird untersucht, ob eine Interaktion signifikant ist und andererseits wird gefragt inwieweit die Vergangenheitswerte benötigt werden, um die Gegenwart zu erklären. Sukzessive wird somit getestet, ob  $\beta_{23}$  bzw.  $\beta_{13}$  und  $\beta_{12}$  oder  $\beta_{23}$  von 0 verschieden sind. Ein derartiges Vorgehen kommt einer Rückwärtsselektion gleich, wie sie in Kapitel 1.2 skizziert worden ist. Alternativ kann eine Vorwärtsselektion angewendet werden, indem  $Y_{it}$  sukzessive aus immer mehr Vergangenheitswerten erklärt wird, solange bis diese keinen signifikanten Erklärungsbeitrag mehr leisten.

*Beispiel:*

*Betrachten wir folgendes Beispiel. Die Jugendlichen sollten ihr Urteil über die Fragestellung „Gesetz*

den Fall, Sie würden über ausreichend Geld verfügen: Könnten Sie sich dann ein Leben ohne Arbeit vorstellen?' zu den Zeitpunkten  $t = 1, 2, 3$  abgeben. Hat ein Jugendlicher zum Zeitpunkt 1 sich mit dem Slogan solidarisch erklärt, so wird er zum Zeitpunkt 2 ebenfalls dazu neigen. Einen ähnlichen Einfluß gibt es zwischen der zweiten und dritten Welle. Der interessierende Aspekt ist nun, a) ob eine Interaktion zwischen dem Einfluß der Messung zum Zeitpunkt 1 und 2 auf die Messung  $Y^{\wedge}$  besteht und wenn dies nicht der Fall ist, ob b)  $Y_2$  zur Erklärung von  $Y_3$  bereits ausreicht. Wir betrachten Modell (7) und erhalten die Parameterschätzer in Tabelle 6. Es zeigt sich, daß der interaktive Effekt nicht signifikant ist. Somit reduzieren wir das Modell um den Effekt und schätzen das Modell (6). In diesem Haupteffektmodell sind alle Größen signifikant, d.h. insbesondere, daß die Markov-Kette die Ordnung 2 besitzt. Die zugehörige Wahrscheinlichkeiten sind in Tabelle 7 aufgeführt. Es ergibt sich bei dem Datenbeispiel keine weitere Vereinfachung der Markov-Struktur, wohl jedoch ist die Abhängigkeitsstruktur auf ihre signifikante, hier additive Struktur reduziert.

Der signifikante additive Einfluß von der ersten Periode kann in der Tabelle 7 gesehen werden. Der Anteil der Jugendlichen die sich in Periode 3 ein Leben ohne Arbeit vorstellen können ist um ca. 20 Prozentpunkte höher für Jugendliche, die bereits in der Periode 2 sich ein Leben ohne Arbeit vorstellen können. Konnten sich die Jugendlichen bereits in der Periode 1 ein Leben ohne Arbeit vorstellen, so steigt der Anteil nochmals um ca. 10%. Diese Effekte reagieren additiv.

Modell	Parameter	Schätzer	Standardabweichung
$h\{E(Yx Y_{1u}Y_{12})\}$ $= \beta_0 + Y_{1j}\beta_{1z} + Y_{12}\beta_{13} + Y_{1X}Y_{12}\beta_{12Z}$	$\beta_0$	-2,617	0,997
	$\beta_{1n}$	0,956	0,691
	A3	1,328	0,610
	$\beta_{123}$	<b>-0,302</b>	<b>0,361</b>
$h(E\{Y_{13} Y_{1U}Y_{12}\})$ $= \beta_0 + Y_{1n}\beta_{13} + Y_{12}\beta_{2Z}$	$\beta_0$	-1,850	0,374
	$\beta_{1n}$	0,470	0,175
	A3	0,841	0,175

Tabelle 6: Parameterschätzer

Zeitpunkt 1	Zeitpunkt 2	ja (beobachtet)	ja (ohne Interaktion)
ja	ja	60,4%	63,1%
	nein	40,4%	42,5%
nein	ja	49,9%	51,7%
	nein	32,2%	31,6%

Tabelle 7: Vorstellung Leben ohne Arbeit, obere Zeile beobachtet, untere Zeile Modell ohne Interaktion

## 2.2 Komplexe Markov-Ketten

Bisher wurden zusätzliche Regressoren  $X_{it}$  in der Analyse vernachlässigt. Ihre Berücksichtigung ist jedoch einfach zu handhaben, indem das Modell (6) um die entsprechenden Größen

erweitert wird, d.h. wir setzen

$$\begin{aligned} E\{Y_{12}|Y_{11}, X_{12}\} &= h(\beta_{02} + Y_{11}\beta_{12} + X_{12}\beta_{X2}) \\ E(Y_a|Y_{11}, Y_{12}, X_{13}) &= h(\beta_m + Y_{11}\beta_{13} + Y_{12}\beta_{23} + X_{13}\beta_{X3}). \end{aligned} \quad (8)$$

Das bedeutet, zur Erklärung von  $Y_{1t}$  wird auf die Vergangenheitswerte und auf die Regressoren  $X_{1t}$  der entsprechenden Zeitperiode zurückgegriffen. Mit analogen Methoden wie oben kann nun die Modellstruktur (8) vereinfacht werden. Dabei werden Tests auf die Ordnung der Markov-Kette getrennt von Tests auf die Signifikanz der Regressoren durchgeführt.

## 2.3 Software

In Analogie zu Regressionsmodellen kann nach entsprechender Aufarbeitung der Daten auf klassische Software zurückgegriffen werden, d.h. Produkte wie SPSS®, SPLUS® oder SAS®. Man muß jedoch beachten, daß die Zerlegung der Modellierung in sequentielle Regressionsmodelle und die daraus resultierende sequentielle numerische Berechnung weitere Probleme aufwirft. Insbesondere entsteht die Problematik des sequentiellen Testens. Es werden Signifikanzen zwischen den Zeitpunkten 1 und 2 und zwischen den Zeitpunkten (1,2) und 3 getrennt berechnet. Das führt dazu, daß ein globales Signifikanzniveau nicht notwendigerweise eingehalten wird. Auswege aus diesem Dilemma ergeben sich mit Zeitreihenmodellen und entsprechender Software. Alternativ dazu können Anpassungen des Signifikanzniveaus nach Bonferroni Fahrmeir, Hamerle & Tutz (1996, Seite 92) vorgenommen werden. Der Punkt soll jedoch hier nicht problematisiert werden.

### 3 Kontingenztafeln mit sequentieller Struktur

Im folgenden Abschnitt werden Kontingenztafeln betrachtet, denen eine sequentielle Struktur zugrunde liegt, d.h. die Messungen  $Y_{it}$  und  $Y_{i,t-1}$  werden gegeneinander abgetragen, wobei  $Y_{it}$  und  $Y_{i,t-1}$  hier kategoriale Variablen sind. Die Tafel zeichnet sich insbesondere dadurch aus, daß Ausprägungen der Zeilen und Spalten gleich sind.

Zur Analyse derartiger Tafeln werden andere Modelle herangezogen als sie standardmäßig zur Kontingenztafelanalyse angewandt werden. Nachfolgend werden die zur Projektauswertung verwendeten Methoden vorgestellt.

#### 3.1 Unabhängigkeit und Quasiunabhängigkeit

Es soll zunächst die Abhängigkeitsstruktur zwischen dem kategorialen Merkmalen  $Y_{it}$  und  $Y_{i,t-1}$  betrachtet werden. Die einfachste Form der Analyse ergibt sich aus der Fragestellung, ob  $Y_{it}$  und  $Y_{i,t-1}$  voneinander unabhängig sind. Ausgedrückt im Kontext der Markov Ketten des vergangenen Kapitels bedeutet dies, daß im Modell

$$E(Y_{it}|Y_{i,t-1}) = h(\beta_0 + Y_{i,t-1}\beta_1)$$

die Signifikanz des Parameters  $\beta_1$  untersucht wird, d.h. untersucht wird, ob die Markov Kette die Ordnung 0 hat. Die zugehörige statistische Methodik ergibt sich mit dem klassischen  $\chi^2$ -Unabhängigkeitstest. Bezeichnen wir mit  $h_{rs}$  die Anzahl der Beobachtungen mit Ausprägungen  $Y_{it} = r$  und  $Y_{i,t-1} = s$ , so erhalten wir die Teststatistik

$$\chi^2 = \sum_{r,s} \frac{(h_{rs} - \frac{h_{r.}h_{.s}}{n})^2}{\frac{h_{r.}h_{.s}}{n}}$$

wobei  $h_{r.} = \sum_s h_{rs}$  und  $h_{.s} = \sum_r h_{rs}$  die Randhäufigkeiten der Tabelle angeben. Betrachtet man die in Tabelle 8 kreuzklassifizierten Beobachtungen, so ist offensichtlich, daß eine Unabhängigkeit von  $Y_{it}$  und  $Y_{i,t-1}$  nicht vorliegt. Die Diagonale der Kontingenztafel ist unverhältnismäßig stark besetzt. Das bedeutet inhaltlich, daß die Individuen in Periode  $t$  tendenziell zur gleichen Antwort neigen wie in der Vorperiode. Das primäre Interesse liegt somit nicht auf den Individuen, die ihre Antwort in Perioden  $t$  und  $t - 1$  gleich wählen, sondern auf denjenigen, deren Meinungsbild sich ändert, d.h. die in Periode  $t$  eine andere Antwort geben als in Periode  $t - 1$ . Wir bezeichnen diese Individuen als **Wechsler**, die Anderen als Nichtwechsler. Die interessierende Fragestellung ist nun, ob unter den Wechslern Struktur vorliegt, oder ob im Gegensatz dazu, das Wechseln eher zufällig geschieht. Struktur bedeutet hier, daß der Quotient aus  $h_{rs}$  (Anzahl der Wechsler von Kategorie  $r$  nach Kategorie  $s$ ) und  $h_{rt}$  (Anzahl Wechsler von  $r$  nach  $t$ ) nicht von der Herkunftskategorie  $r$  abhängt. Das heißt, daß das Verhältnis unter den Wechslern nicht von der Antwort in der Vorperiode abhängt. Formelmäßig ausgedrückt bedeutet das:  $h_{rs} = a_r \beta_s$  für  $r \neq s$  und geeignetem  $a_r$  und  $\beta_s$ . Insbesondere werden nur die Nichtdiagonalfelder der Tafel analysiert, d.h. nur die Wechsler werden betrachtet. Die Überprüfung der Hypothese, daß unter den Wechslern keine Struktur vorliegt, ist nun mit

einem leicht modifizierten  $\chi^2$  Test möglich, d.h. wir betrachten

$$T^2 = \sum_{r \neq s} \frac{(h_{rs} - \hat{a}_r \hat{\beta}_s)^2}{\hat{a}_r \hat{\beta}_s} \quad (9)$$

mit  $\hat{a}_r$  und  $\hat{\beta}_s$ , gemäß dem Likelihood-Ansatz ermittelt (siehe Agresti, 1990, Seite 356). Hypothesen dieser Form werden als Quasi-Unabhängigkeit bezeichnet. Die Hypothesen der Quasi-unabhängigkeit ist hauptsächlich auf die Analyse von Daten, die zu zwei Zeitpunkten erhoben werden, beschränkt. Eine Erweiterung auf mehr als zwei Zeitpunkte ist wissenschaftlich nicht geläufig. Die Quasiunabhängigkeit wird somit nur jeweils paarweise zwischen Zeitpunkten 1 und 2 sowie zwischen Zeitpunkten 2 und 3 überprüft.

Wie empfinden Sie ihre jetzige Ausbildung?

	Zeitpunkt 2				
Zeitpunkt 1	1	2	3	4	5
1	0	1	4	9	4
		0.5	1.1	13.1	2.1
2	2	0	2	12	1
	0.4		1.6	12.5	2.5
3	1	0	9	21	5
	0.6	0.9		21.2	4.3
4	9	13	40	212	62
	9.3	12.5	39.8		62.5
5	5	9	26	227	151
	6.1	9.1	28.9	222.3	

Tabelle 8: Beurteilung der Ausbildung: (1="Möchte lieber heute als morgen aufhören", 2="Fällt recht schwer", 3="Es geht", 4="Ist ganz gut" und 5="Macht richtig Spaß") zu den Zeitpunkten 1 und 2. Beobachtete Fälle (obere Zeile) und geschätzte Fälle (untere Zeile) bei Quasi-Unabhängigkeit.

*Beispiel:*

*Wir betrachten als Beispiel die Zufriedenheit der Jugendlichen mit ihrer Ausbildung zu den Zeitpunkten 1 und 2. Die Ausprägungen von  $Y_u$  sind in den folgenden Kategorien erhoben:  $Y_{1t} = 1$  für "Möchte lieber heute als morgen aufhören",  $Y^{\wedge} = 2$  für "Es fällt mir schwer",  $Y_u = 3$  für "Es geht, ist für mich ein notwendiges Übel",  $Y_u = 4$  für "Ist ganz gut" und  $Y_{1t} = 5$  für "Macht richtig Spaß".*

*Wir erhalten die Daten, wie sie in Tabelle 8 und 9 gegeben sind. Offensichtlich sind die Diagonalen in beiden Tabellen stark besetzt, was natürlich eine Abhängigkeit der Merkmale hervorruft. Wir schätzen unter der Annahme von Quasi-Unabhängigkeit, d.h. wir analysieren, ob in den Nichtdiagonalfeldern Unabhängigkeit vorliegt. Die entsprechenden erwarteten Größen sind in den Tabellen als zweite Zeilen gegeben. In beiden Tabellen wird nun ein  $\chi^2$ -Quasi-Unabhängigkeitstest durchgeführt. Wir erhalten für die erste Tabelle  $T^2_2 = 12,24$  bei 11 Freiheitsgraden. Das entspricht einem p-Wert von 0,346. Das bedeutet, daß die Hypothese der Quasi-Unabhängigkeit nicht verworfen werden kann. Antwortet somit ein Jugendlicher in Periode 2 anders als in Periode 1, d.h. ist er ein sogenannter Wechsler, so ist seine Antwort in Periode 2 nicht abhängig von seiner Antwort in Periode 1. Anders sieht es aus*

Wie empfinden Sie ihre jetzige Ausbildung?

	Zeitpunkt 3				
Zeitpunkt 2	1	2	3	4	5
1	0	0	3	6	1
		0.1	1.0	6.1	2.8
2	0	2	3	5	4
	0.1		1.2	1.4	3.3
3	2	3	11	29	9
	0.3	0.1		29.0	13.0
4	1	4	35	173	112
	2.1	5.7	31.3		106.9
5	0	1	8	64	83
	0.5	1.5	9.6	61.4	

Tabelle 9: Beurteilung der Ausbildung: (1="Möchte lieber heute als morgen aufhören", 2="Fällt recht schwer", 3="Es geht", 4="Ist ganz gut" und 5="Macht richtig Spaß") zu den Zeitpunkten 2 und 3. Beobachtete Fälle und geschätzte Fälle bei Quasi-Unabhängigkeit.

zwischen Perioden 2 und 3. Dort erhält man  $T_{23}^2 = 21,31$  mit dem  $p$ -Wert 0,030. Somit liegt keine Quasiunabhängigkeit vor. Ändert ein Jugendlicher seine Meinung bzgl. seiner Ausbildungssituation, so ist seine neue Meinung signifikant abhängig von seiner bisherigen Meinung. Eine Untersuchung der Residuen ergibt, daß die Verletzung der Quasi-Unabhängigkeit insbesondere durch 3 Zellen verursacht ist. Tabelle 10 zeigt die Pearson-Residuen  $(h_{rs} - \hat{h}_{rs})/\sqrt{\hat{h}_{rs}}$ , wobei  $\hat{h}_{rs} = \hat{\alpha}_r \hat{\beta}_s$  die unter Quasi-Unabhängigkeit erwarteten Häufigkeiten sind. Insbesondere Zellen  $Y^{\wedge} = 3$  (Periode 2: "Es geht") und  $F_{13} = 1$  oder 2 (Periode 3: "Möchte aufhören" oder "Fällt schwer") als auch die Zellen  $Y^{\wedge} = 1$  oder 2 (Periode 2: "Möchte aufhören" oder "Fällt schwer") und  $Y_{13} = 3$  (Periode 3: "Es geht") fallen durch signifikant hohe Residuen auf, d.h. es wurden in diesen Zellen mehr Fälle beobachtet als bei Quasi-Unabhängigkeit zu erwarten waren.

Wie empfinden Sie ihre jetzige Ausbildung?

	Zeitpunkt 3				
Zeitpunkt 2	1	2	3	4	5
1		-0.38	<b>2.08</b>	-0.06	-1.05
2	-0,26		<b>1.10</b>	-0.89	0.36
3	<b>3.42</b>	<b>2.11</b>		-0.06	-1.11
4	-0.11	-0.11	-0.31		0.49
5	-0.14	-0.38	-0.51	0.33	

Tabelle 10: Residuen (signifikante Werte [bei  $\alpha = 10\%$ ] sind fettgedruckt)

### 3.2 Marginale Homogenität

Steht bei einem Quasi-Unabhängigkeits-Modell der individuenspezifische Gesichtspunkt im Vordergrund, so liegt der Schwerpunkt bei marginaler Homogenität auf einer Querschnittsbeobachtung. Das bedeutet, daß bei Quasi Unabhängigkeit das Innere der Tafel betrachtet wird, bei marginaler Homogenität werden dagegen die Ränder der Tafel untersucht. Ziel der Analyse ist es, zu beurteilen, ob die Randwahrscheinlichkeiten als gleich, d.h. als homogen angesehen werden können. Dabei wird die Struktur im Inneren der Tafel nicht berücksichtigt.

Modelle, in denen die Ränder, d.h. die Marginalien von Tafeln untersucht werden, behandeln wir ausführlicher im nächsten Abschnitt. Ein klassischer Test auf marginale Homogenität in einer 2 x 2 Tafel ist der McNemar Test:

$$T = \frac{(h_{12} - h_{21})^2}{(h_{12} + h_{21})}$$

Haben die Messungen  $Y_u$  und  $Y_{it}$  mehr als zwei Ausprägungen, so stehen Likelihood-Quotienten Test zur Überprüfung von marginaler Homogenität zur Verfügung (siehe Agresti, 1990, Seite 358).

*Beispiel:*

*Wir betrachten die in Tabelle 11 gegebenen Daten. Sie geben die Klassifizierung der Empfindung beim Besuch der Berufsschule in den Perioden 1 und 2 wieder. Die möglichen Ausprägungen sind  $Y_u = 1$  für "Möchte lieber heute als morgen aufhören",  $Y_u = 2$  für "Fällt mir schwer",  $Y_u = 3$  für "Es geht, ist notwendiges Übel",  $Y_{it} = 4$  für "Ist ganz gut" und  $Y_{it} = 5$  für "Macht richtig Spaß". Betrachtet man die Ränder der Tafel so zeigt sich, daß diese eine ähnliche Verteilung aufweisen. Es wird somit auf marginale Homogenität getestet. Man erhält den  $\chi^2$ -Wert 6.39 bei 4 Freiheitsgraden. Das entspricht einem p-Wert von 0,112. Die Hypothese wird somit angenommen, d.h. die durchschnittliche Empfindung beim Besuch der Berufsschule weist zwischen der ersten und zweiten Periode keinen signifikanten Unterschied auf. (In Tabelle 12 ist die entsprechende homogene marginale Randtafel gegeben)*

Empfindung beim Besuch der Berufsschule						
	Zeitpunkt 2					
Zeitpunkt 1	1	2	3	4	5	
1	1	0	4	10	3	18
2	3	4	17	15	4	43
3	9	9	102	92	15	227
4	8	15	122	241	48	434
5	2	3	16	56	24	101
	23	31	261	414	94	823

Tabelle 11: Beurteilung der Berufsschule: (1="Möchte lieber heute als morgen aufhören", 2="Fällt recht schwer", 3="Es geht", 4="Ist ganz gut" und 5="Macht richtig Spaß") zu den Zeitpunkten 1 und 2. Beobachtete Fälle und geschätzte Fälle bei Quasi-Unabhängigkeit.

Abschließend sei noch auf die Zusammenhänge zwischen Symmetrie und marginaler Homogenität hingewiesen. Marginale Homogenität folgt aus Symmetrie, jedoch nicht umgekehrt.

### Empfindung beim Besuch der Berufsschule

1	2	3	4	5
20,4	35,8	243,0	425,7	98,1

Tabelle 12: Marginale Häufigkeit bei Homogenitätsannahme bzgl. der Beurteilung der Berufsschule: (1="Möchte lieber heute als morgen aufhören", 2="Fällt recht schwer", 3="Es geht", 4="Ist ganz gut" und 5="Macht richtig Spaß") zu den Zeitpunkten 2 und 3. Beobachtete Fälle und geschätzte Fälle bei Quasi-Unabhängigkeit.

Unter Symmetrie versteht man hier, daß die Kontingenztafel eine symmetrische Struktur um die Diagonale aufweist. Es zeigte sich jedoch, daß die Interpretation von Symmetrie schwierig ist, weshalb hier auf eine weitere Diskussion verzichtet wird (vergleiche Agresti, 1990, Kapitel 10).

### 3.3 Software

Zur numerischen Umsetzung von Modellen für Kontingenztafeln mit sequentieller Struktur stehen die Produkte SPSS® und SAS® bereit. SPSS® weist jedoch erhebliche Einschränkungen auf. So ist es in SPSS® beispielsweise nicht möglich auf marginale Homogenität zu testen.



## 4 Marginale Modelle

### 4.1 Modellierung

Der folgende Abschnitt beschreibt eine junge Analyserichtung der Statistik. Wie bereits im Abschnitt über marginale Homogenität motiviert, soll der Schwerpunkt nun auf Querschnittsanalysen gelegt werden. Dies steht im Gegensatz zu Markov Modellen, in denen individuen-spezifische Gesichtspunkte im Vordergrund standen. Bei Markovketten gilt es, das Ergebnis  $Y_{it}$  des  $i$ -ten Individuum zum Zeitpunkt  $t$  aus den Ergebnissen  $Y_{i,t-1}, Y_{i,t-2}$  der Vorperioden vom  $(t-1)$ -ten Individuum zu erklären. Der Schwerpunkt liegt also auf der Modellierung und Analyse der zeitlichen Abhängigkeitsstruktur von an einem Individuum erhobene Messungen. Betrachtet man die Messungen nur zu einem festen Zeitpunkt  $t$ , so spricht man von einer Querschnittsanalyse. Hier betrachtet man also nur die Verteilung von  $Y_{it}$  unabhängig von den Ergebnissen der Vor- und Nachperioden. Diese Querschnittsanalyse hängt natürlich von dem gewählten Zeitpunkt  $t$  ab. Wird nun diese Abhängigkeit der Querschnittsanalysen vom jeweils betrachteten Zeitpunkt  $t$  modelliert, so spricht man von marginalen Modellen. Bei der Betrachtung von kategorialen Messungen  $Y_{it}$  und  $Y_{it+i}$  zu zwei Zeitperioden, ergeben sich Fragestellungen, wie sie bereits als marginale Homogenität im letzten Kapitel vorgestellt wurden.

Folgend dem Regressionsansatz (1) lassen sich marginale Modelle schreiben als

$$h(E(Y_{it}|t)) = \beta_t \quad (10)$$

Es wird also angenommen, daß sich der mittlere Wert von  $Y_{it}$  (d.h. der Durchschnitt der Querschnittsanalyse zum Zeitpunkt  $t$ ) mit  $t$  ändert. Die interessierende Fragestellung ist nun, ob  $\beta_t$  konstant ist für alle  $t$ , d.h.  $\beta_1 = \beta_2 = \beta_3$ . Diese Fragestellung beinhaltet auch die Hypothese der marginalen Homogenität (siehe vorheriges Kapitel).

*Beispiel:*

*Wir betrachten die Meinung der Jugendlichen zur Aussage „Die Politik kann jemand wie ich überhaupt nicht beeinflussen“, erhoben zu den Zeitpunkten 1, 2 und 3. Die möglichen Ausprägungen sind  $Y_{it} = 1$  für „richtig“ und  $Y_{it} = 0$  für „falsch“. Wir modellieren  $h(E(Y_{it}|t)) = \beta_t$  mit  $h(\cdot)$  als logit Funktion. Die marginalen Häufigkeiten für die Zustimmung zu der Aussage sind in Tabelle 13 bzw. Abbildung 5 gegeben. Es zeigt sich, daß der Test auf die Hypothese  $H_0: \beta_1 = \beta_2 = \beta_3$  die Hypothese nicht verwirft bei einem  $p$ -Wert von 0,647. Somit zeigt sich, daß die Einstellung bzgl. der Beeinflussung von Politikern im Mittel keiner zeitlichen Veränderung unterliegt, ebenfalls dargestellt in Abbildung 5.*

*Das Bild ändert sich, wenn wir die Einstellung zur Aussage „Politiker bei uns machen was sie wollen“ betrachten. Hier zeigt sich, daß der Zeiteffekt signifikant ist ( $p$ -Wert  $< 0.001$ ). Die zugehörigen Wahrscheinlichkeiten sind in Tabelle 14 bzw. Graphik 5 gegeben. Offensichtlich stimmen in Periode zwei signifikant mehr Jugendliche der Aussage zu, als in Perioden eins und drei.*

Marginale Modelle der Form (10) können nun um zusätzliche Regressoren  $X_{it}$  erweitert werden. Hierzu geht man über zum Modell

$$h(E(Y_{it}|X_{it})) = \beta_t + X_{it}\beta_x \quad (11)$$

wobei  $\beta_x$  den Einfluß der Regressoren  $X_{it}$  widerspiegelt. Dieser Einfluß ist zunächst als zeitkonstant modelliert, d.h. die Regressoren  $X_{it}$  haben zu jedem Zeitpunkt den gleichen Einfluß.

Politik kann ich nicht beeinflussen

	Zeitpunkt		
	1	2	3
richtig	367 53,7%	382 55,8%	377 55,1%
falsch	317 46,3%	302 44,2%	307 44,9%

Tabelle 13: Randhäufigkeit zur Fragestellung: *Politik kann jemand wie ich nicht beeinflussen*

Politiker machen was sie wollen

	Zeitpunkt		
	1	2	3
richtig	518 76,2%	546 80,3%	482 70,9%
falsch	162 23,8%	134 19,7%	189 29,1%

Tabelle 14: Randhäufigkeiten zu der Fragestellung: *Politiker bei uns machen was sie wollen*

Gibt man diese Annahme auf und erlaubt, daß der Einfluß der Regressoren über die Zeit variiert, so erweitert sich das Modell (11) zu

$$h\{E\{Y_{xt}|X_{it}\}) = \beta_t + X_{it}\beta_{tX}. \quad (12)$$

In (12) wird für jede Periode  $t$  ein eigener Effekt der Regressoren  $X_{it}$  modelliert. Dies bedeutet, daß in (12) eine Interaktion zwischen Regressoren und Zeit modelliert ist, in (11) dagegen agieren Zeit und Regressoren additiv. Die üblichen Überlegungen erlauben nun die Überprüfung der Signifikanzen der Parameter. Wie bisher können nun entweder vor- oder rückwärtiges Selektionverfahren angewandt werden (siehe Kapitel 1.2), um ein adäquates Modell zu wählen.

*Beispiel:*

*Wir setzen die Analyse des vorherigen Beispiels fort. Insbesondere untersuchen wir, ob das Geschlecht einen relevanten Einfluß auf die Meinung zur Frage „Politiker machen was sie wollen“ hat. Wir modellieren zunächst einen über die Zeit variierenden Geschlechtseffekt, d.h. wir benutzen Modell (12). Die zugehörigen relativen Häufigkeiten entsprechen den beobachteten Häufigkeiten und sind in Tabelle 15 gegeben.*

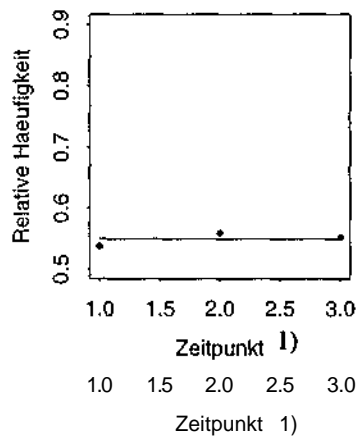
*Zunächst soll untersucht werden, ob der Einfluß des Geschlechts über die Zeit variiert, d.h. ob die Hypothese  $H_0 : \beta_{x1} = \beta_{x2} = \beta_{x3}$  gültig ist wobei  $\beta_{xt}$  der Regressionskoeffizient des Geschlechts zum Zeitpunkt  $t$  ist. Der Test liefert einen  $p$ -Wert von  $p = 0,075$ , basierend auf einem 5% Niveau ist somit keine zeitliche Variation nachweisbar. Somit gehen wir über zum Haupteffektmodell (11). Die zugehörigen geschätzten Wahrscheinlichkeiten sind in der Abbildung 6 dargestellt. Ein weiterer Test auf  $H_0 : \beta_x = 0$  wird mit einem  $p$ -Wert von  $p = 0,040$  und auf  $H_0 : \beta_t = 0$  mit einem  $p$ -Wert von  $p < 0,001$  abgelehnt. Somit agieren Zeit und Geschlecht additiv, jedoch nicht interaktiv, was in Abbildung 6 anschaulich wird.*

Politiker machen was sie wollen

Geschlecht	Zeitpunkt		
	1	2	3
männlich	308 76,6%	314 78,1%	267 67,2%
weiblich	210 75,5%	232 83,5%	212 76,3%

Tabelle 15: Zustimmung zur Aussage: Politiker machen was sie wollen

Politik kann ich nicht beeinflussen



Politiker machen was sie wollen

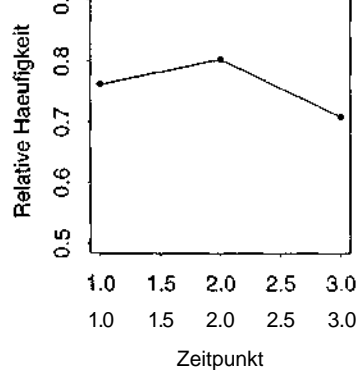


Abbildung 5: Zustimmung zu den Aussagen "Politik kann ich nicht beeinflussen" und "Politiker machen was sie wollen"

1) Das verwendete Grafikprogramm ließ keine Ausblendung der Zwischenzeitpunkte  $t = 1,5$  und  $2,5$  zu.

## Politiker machen was sie wollen

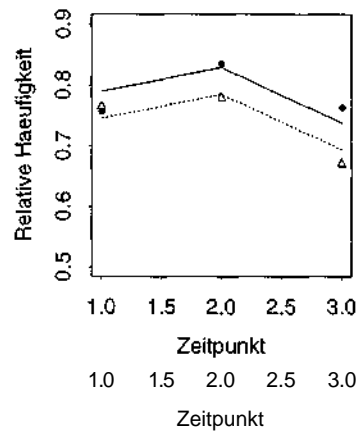


Abbildung 6: Zustimmung zur Aussage "Politiker machen was sie wollen" in Abhängigkeit vom Geschlecht (Männer—und Frauen —, • = Männer beobachtet, A = Frauen beobachtet)

## 4.2 Software

Im Falle von stetigen Messungen  $Y_{it}$  sind Modelle der Form (11) auch bekannt als Modelle mit Meßwiederholungen und numerisch im Rahmen der Varianzanalyse umsetzbar (vgl. Bortz, 1989, Kapitel 9). Derartige Modelle sind mit allen gängigen Softwareprodukten (SAS®, SPSS® oder SPLUS®) umsetzbar. Liegen jedoch zeitvariierende Regressoren  $X_{it}$  vor, so ist eine Berechnung mit keinem der erwähnten Produkte möglich. In diesem Fall sind die Berechnungen jedoch von einfacher Struktur und leicht implementierbar, so geschehen zur Auswertung. Ist  $Y_{it}$  dagegen von kategorialer Natur, so stehen in SPSS® keine Auswertungsroutinen zur Verfügung. Unter SAS® lassen sich dagegen Modelle der Form (10) schätzen. Neben nominal skalierten Variablen  $Y_{it}$  werden auch ordinal skalierte Größen (kumulatives Logitmodell) unterstützt. Ebenfalls zulässig unter SAS® sind Modelle der Form (11) bzw. (12), sofern die Regressoren  $X_{it}$  zeitkonstante Faktoren darstellen. Dies ist beispielsweise im obigen Beispiel das Geschlecht eines Individuums. Ist  $X_{it}$  dagegen von stetiger Natur oder auch zeitvariabel, so sind die Modelle mit SAS® nicht mehr berechenbar. Hier muß auf (bisher) nicht gewerbliche Software zurückgegriffen werden, wie sie von Glonek & McCullagh (1995) zur Verfügung gestellt wird.

# **TeilB**

Statistische Analyse

## 5 Analyse des Drop-out Prozesses

Mit Drop-out wird der Vorgang bezeichnet, daß Teilnehmer an einer Panelstudie ab einer bestimmten Welle für alle folgenden Wellen nicht mehr an der Befragung teilnehmen. Im Gegensatz dazu führen Personen, welche an einer oder mehreren Wellen nicht, aber zu einem späteren Zeitpunkt wieder an der Studie teilnehmen, zu sog. intermittierenden fehlenden Werten.

An der ersten Welle der vorliegenden Studie nahmen 2007 Personen teil. Intermittierende fehlende Werte treten bei insgesamt 11 Fällen<sup>1</sup> auf. Von den verbleibenden 1996 Fällen beenden insgesamt 351 Personen alle vier Wellen des Panels, die restlichen 1645 sind Beobachtungen mit Drop-out. Dabei verteilt sich der Drop-out Vorgang auf die einzelnen Wellen wie folgt:

	Anzahl	Prozent	kum.Anzahl	kum.Prozent
Drop-out nach 1.Welle	953	47.7	953	47.7
Drop-out nach 2.Welle	355	17.8	1308	65.5
Drop-out nach 3.Welle	337	16.9	1645	82.4
kein Drop-out	351	17.6	1996	100.0

Tabelle 16: Verteilung des Dropout

Nur 17.6 % der Befragten nahmen also an allen vier Wellen teil. Die Verteilung nach Geschlecht und Drop-out zeigt die folgende Tabelle. Vergleicht man das Drop-out Geschehen

	Drop-out nach 1.Welle	Drop-out nach 2.Welle	Drop-out nach 3.Welle	kein Drop-out	Total	
männlich	546	193	219	187	1145	Frequency
	27.38	9.68	10.98	9.38	57.42	Percent
	47.69	16.86	19.13	16.33		Row Pct
	57.41	54.37	64.99	53.28		Col Pct
weiblich	405	162	118	164	849	Frequency
	20.31	8.12	5.92	8.22	42.58	Percent
	47.70	19.08	13.90	19.32		Row Pct
	42.59	45.63	35.01	46.72		Col Pct
Total	951	355	337	351	1994	
	47.69	17.80	16.90	17.60	100.00	

Tabelle 17: Verteilung des Dropout nach Geschlecht

<sup>1</sup> Da diese intermittierenden fehlenden Werte auch innerhalb der erhobenen Wellen teilweise lückenhaft sind oder die Beobachtungen nur aus der ersten und letzten Welle bestehen, werden sie im weiteren aus der Betrachtung ausgeschlossen.

nach dem Anteil der im Panel verbleibenden Personen bezogen auf die Teilnehmerzahl an der vorausgehenden Welle, so ergibt sich folgendes Bild:

	Gesamt	männlich	weiblich
1.Welle-»2.Welle	52.3%	52.3%	52.3%
2.Welle -> 3.Welle	66.0%	67.8%	63.5%
3.Welle -> 4.Welle	51.0%	46.1%	58.2%

Tabelle 18: Prozent verbleibender Teilnehmer

Eine einfache Möglichkeit zur Untersuchung der Fragestellung, ob beobachtete Größen den Drop-out-Mechanismus beeinflussen, besteht darin, die Verteilung der entsprechenden Variablen in den zwei Teilgruppen zu betrachten, die zum einen durch jene Individuen gebildet werden, welche zum nächsten Erhebungszeitpunkt aus dem Panel ausscheiden, zum anderen durch jene Personen, die weiter beobachtet werden. Ist die Annahme zutreffend, daß die nach dem betrachteten Zeitpunkt ausfallenden Personen eine Zufallsstichprobe aus all jenen zu diesem Zeitpunkt noch im Panel befindlichen Personen sind, so sollten - bis auf zufällige Schwankungen - die Verteilungen zwischen Drop-outs und non-Drop-outs übereinstimmen. Zeigt die Verteilung einer Variablen zwischen diesen beiden Teilgruppen merkliche Unterschiede, so spricht dies gegen die Annahme eines völlig zufälligen Drop-out Mechanismus.

Im folgenden wird getrennt für die drei Panel-Wellen für einige Variablen dieser Vergleich angestellt. Die Darstellung erfolgt dabei stets in der gleichen Form. So zeigt beispielsweise Abbildung 7 oben auf Seite 25 die Verteilung der Variablen 'Geschlecht' für alle Personen, die nach der ersten Panel-Welle noch weiter an der Studie teilnehmen (non-Drop-out 1.Welle, links), im Vergleich zur Gruppe derjenigen, die ihre Teilnahme nach der 1.Welle abbrechen (Drop-out 1.Welle, rechts). In beiden Teilgruppen ist der Frauenanteil annähernd gleich, nämlich 42.57% ( $\frac{442}{1043}$ ) für die Gruppe der non-Drop-outs im Vergleich zu 42.59% ( $\frac{40}{94}$ ) für die Drop-out-Gruppe.

Die zum im Vergleich betrachteten Variablen sind neben dem Geschlecht zunächst Alter, Anzahl der Schuljahre (Abbildung 7), Bundesland, Wohnort (nach Siedlungsform) sowie Haushaltseinkommen (Abbildung 8). Die Betrachtung erfolgt dabei getrennt nach Geschlecht. Abbildung 9 auf Seite 27 zeigt Variablen, die im Zusammenhang mit der aktuellen Tätigkeit bzw. Lehrstellensuche stehen. In Abbildung 10 auf Seite 28 werden Variablen betrachtet, die Einschätzungen über die konkrete Ausbildungsstelle erfassen, Abbildungen 11 und 12 auf Seite 29 - 30 widmen sich der allgemeinen Stimmungslage, der Selbsteinschätzung sowie ausgewählten markanten Einstellungsvariablen.

Für den Drop-out Vorgang nach der 2. bzw. 3.Panel-Welle zeigen Abbildungen 13 bis 15 sowie Abbildungen 16 bis 18 die entsprechenden Vergleiche.

Mit Ausnahme der Variablen 'Bundesland' besitzen die betrachteten Variablen in allen Panel-Wellen homogene Verteilungen. Die besondere Rolle des Wohnortes nach Bundesland wird deutlich, wenn man die Drop-out-Quote für die einzelnen Wellen getrennt nach Bundesland betrachtet. Die folgende Tabelle gibt den Anteil der Personen an, die nach der betrachteten

Welle aus dem Panel ausscheiden. Zusätzlich ist angegeben, wieviele Personen an der Welle in den einzelnen Ländern teilgenommen haben.

	1. Welle	2. Welle	3. Welle
M-Vp	39.2% (250)	13.8% (152)	52.7% (129)
Brb	54.9% (266)	18.3% (120)	49.5% (99)
S-A	47.8% (479)	54.8% (248)	42.2% (109)
S	51.7% (575)	23.1% (282)	50.9% (220)
Th	40.4% (277)	42.1% (164)	45.8% (96)
B(O)	47.7% (132)	52.9% (70)	52.9% (34)
B(W)	37.5% (8)	60.0% (5)	. (0)
Sonst.	55.6% (9)	100% (2)	0.0% (1)

Tabelle 19: Drop-out Quote nach Bundesländern, Anzahl der im Bundesland wohnenden Befragten in Klammern

Besonders bei der zweiten Panel-Welle sind die Drop-out Quoten extrem unterschiedlich: Sie reichen von 13.8% in Mecklenburg-Vorpommern bis zu 54.8% in Sachsen-Anhalt (läßt man die 2 Personen in den alten Bundesländern außer acht). Die Gründe für diese Unterschiede sind vermutlich weniger bei den Befragten, sondern eher bei der Erhebung zu suchen.



## Drop-out 1. Welle

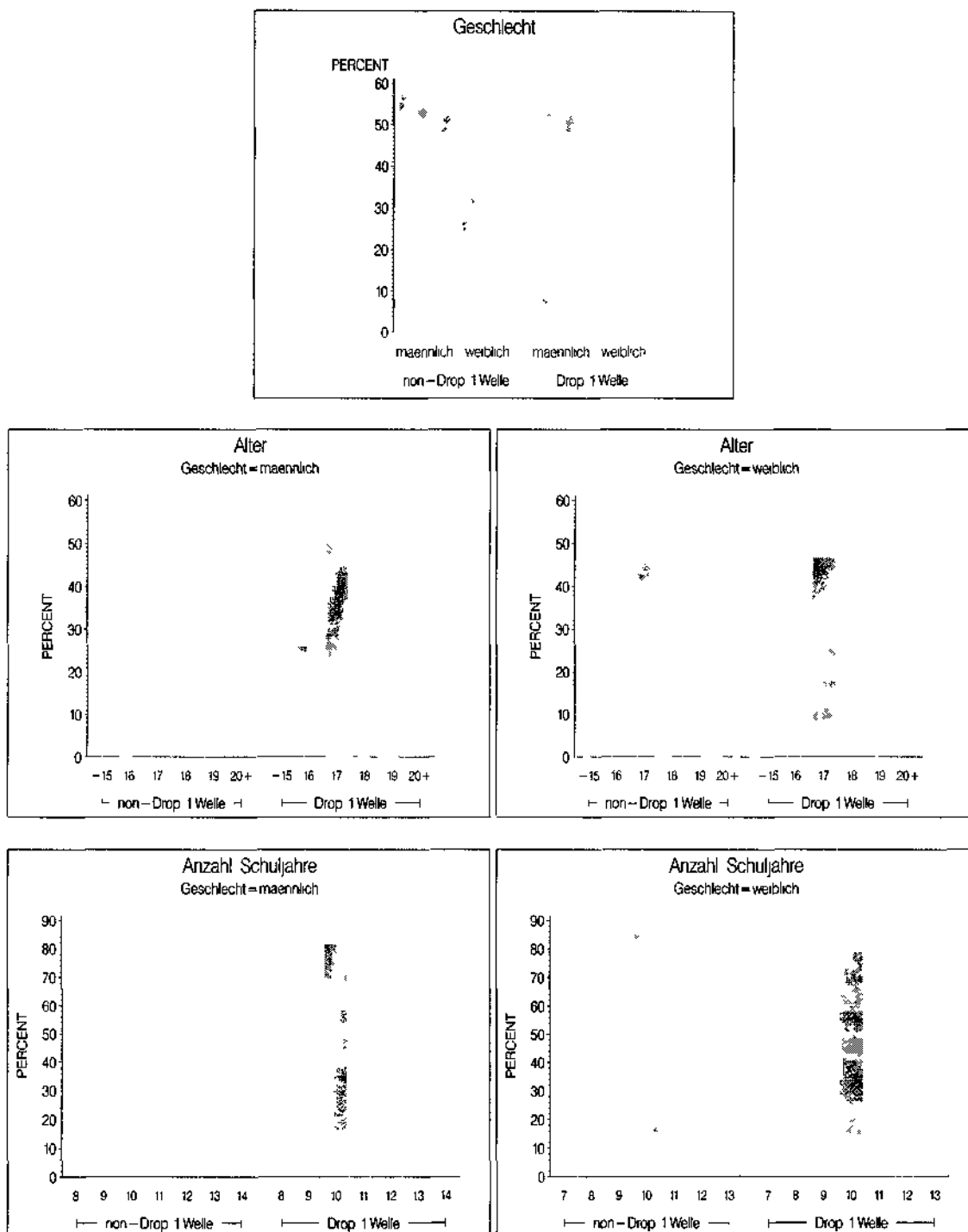


Abbildung 7: Drop-out nach Welle I: Frage Nr. I/53, Nr. I/54, Nr. I/57

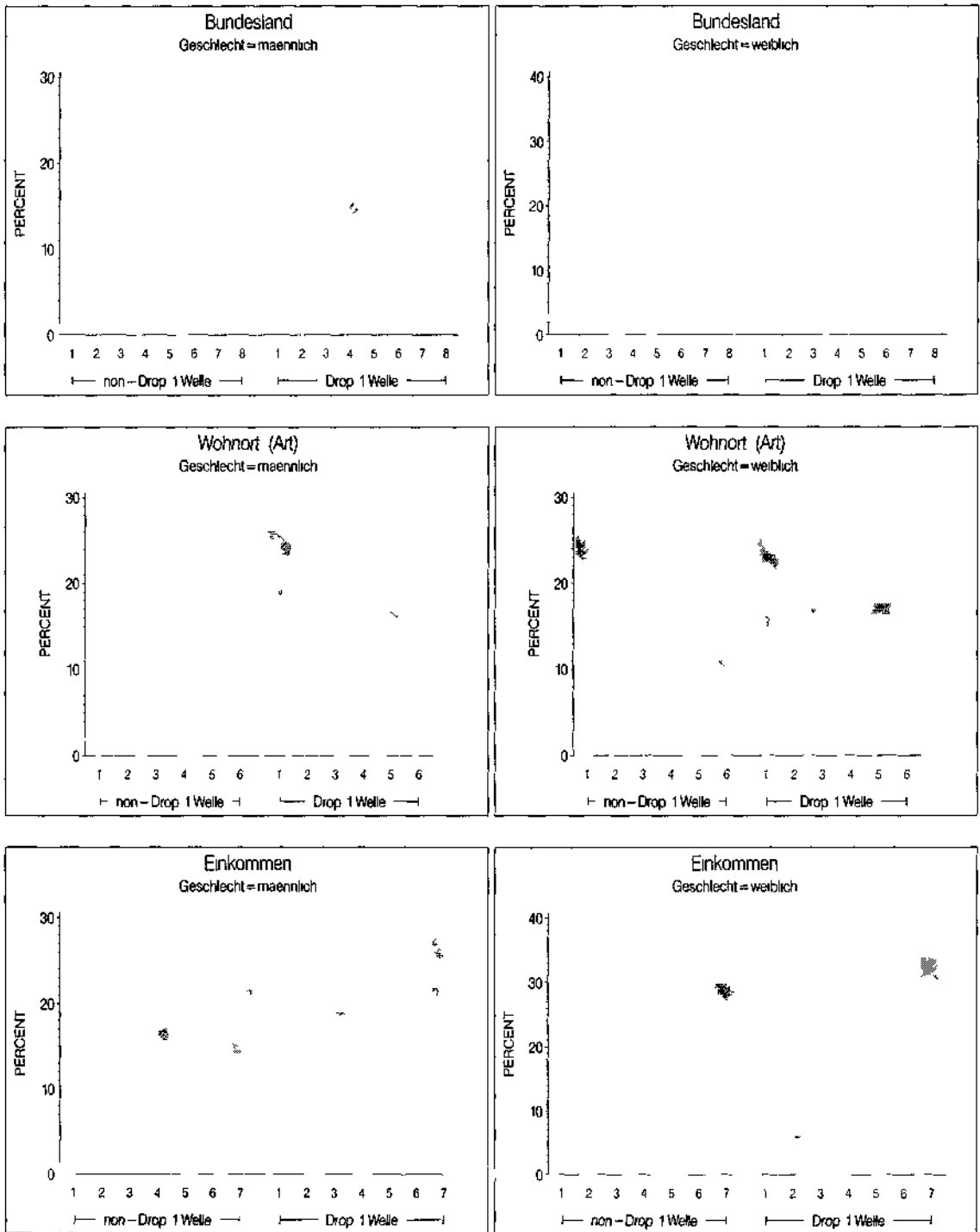


Abbildung 8 Diop-out nach Welle 1 Frage Nr 1/71, Nr 1/68, Nr 1/67

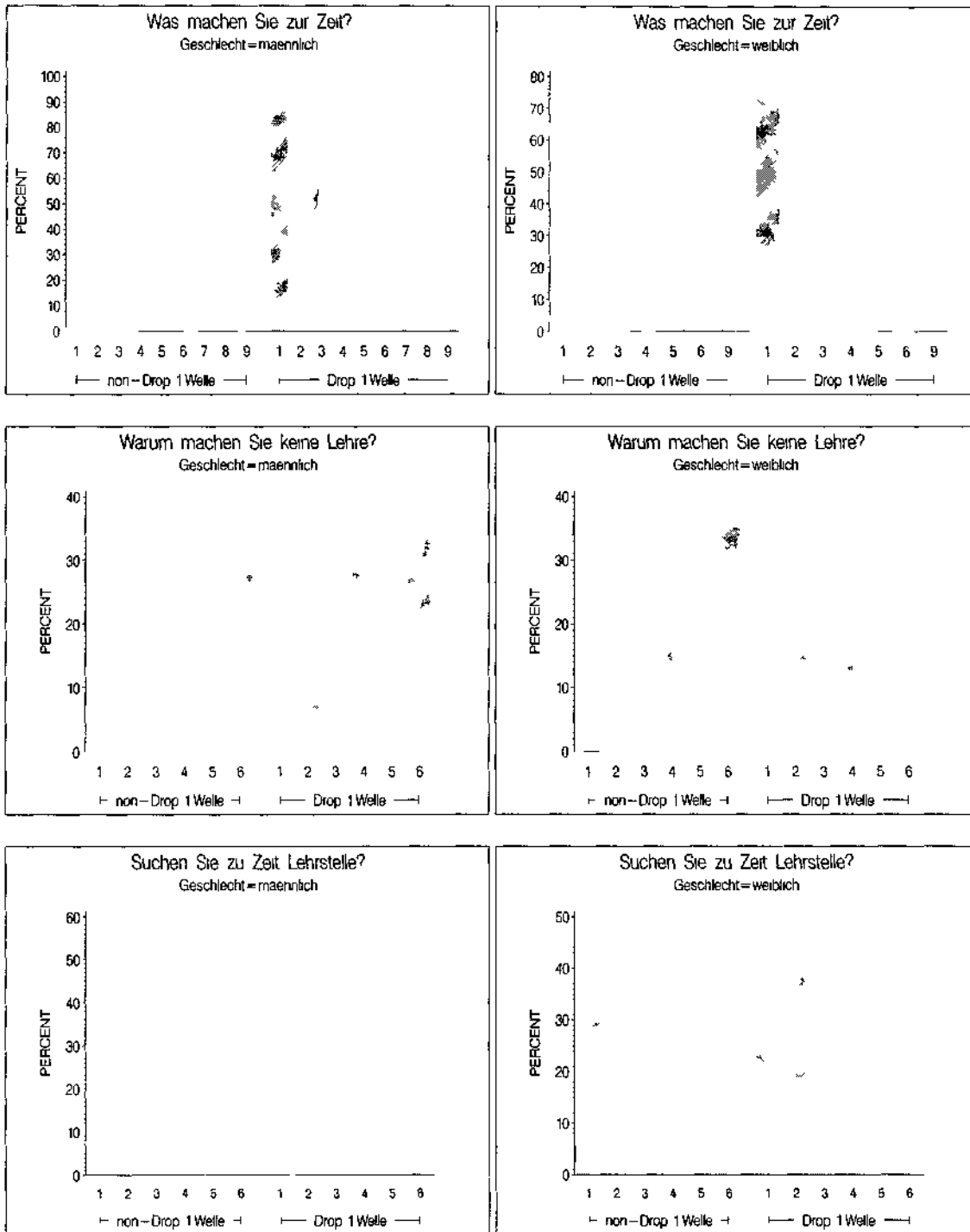


Abbildung 9- Drop-out nach Welle 1 Frage Nr 1/5, Nr 1/6, Nr 1/7

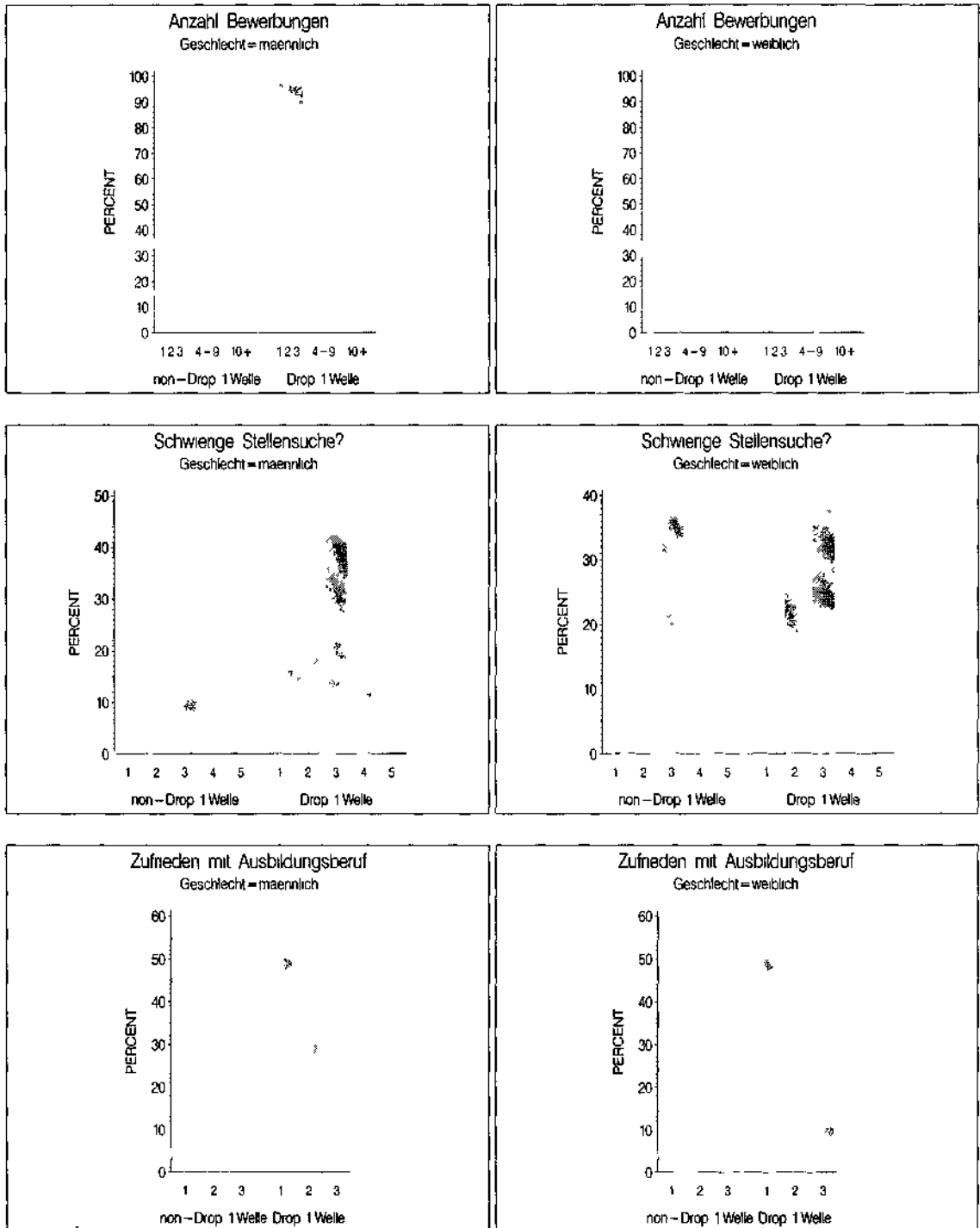


Abbildung 10: Drop-out nach Welle 1: Frage Nr. I/8a , Nr. 1/22, Nr. I/1 1

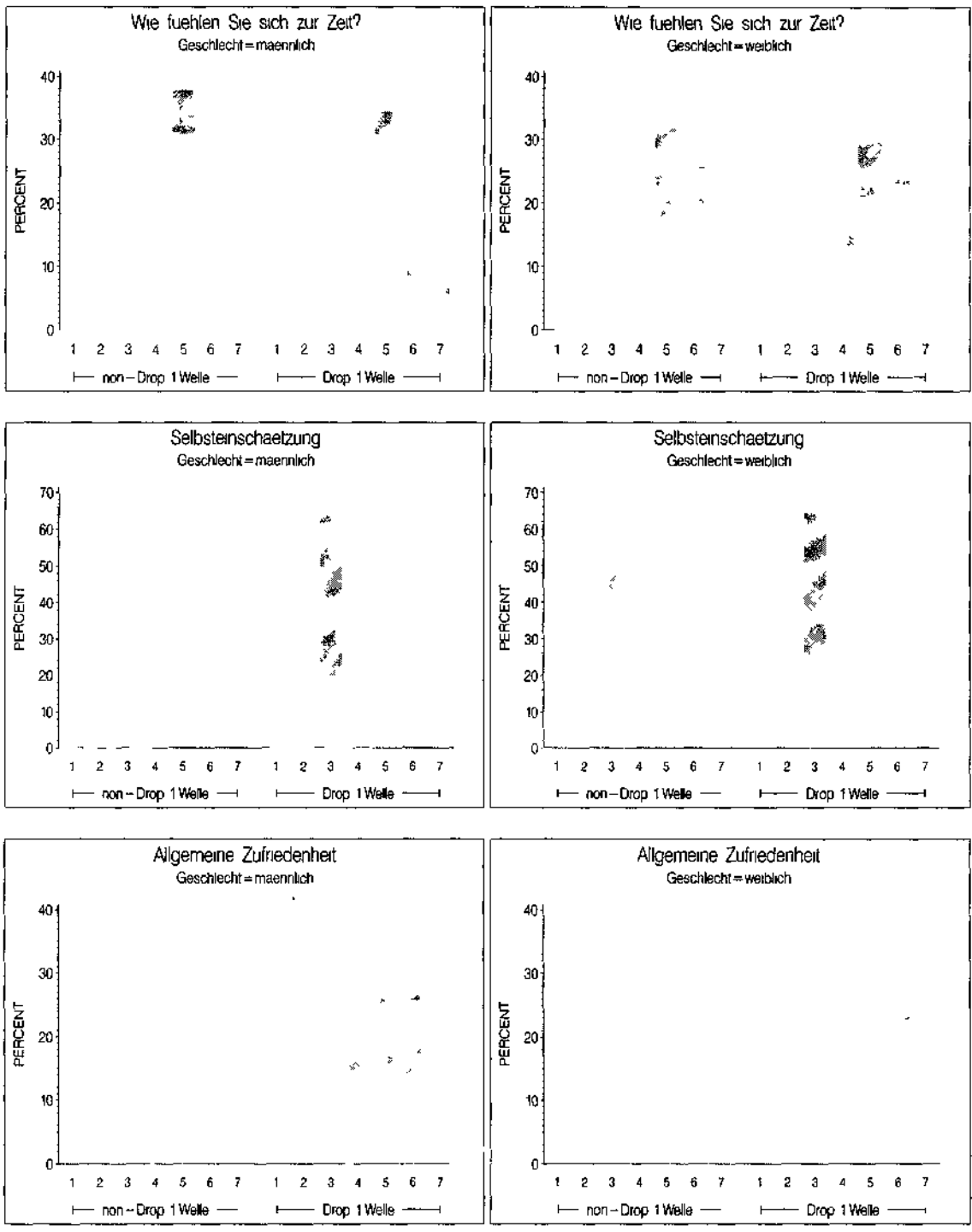


Abbildung 11: Drop-out nach Welle 1: Frage Nr. I/1, Nr. I/38, Nr. I/72

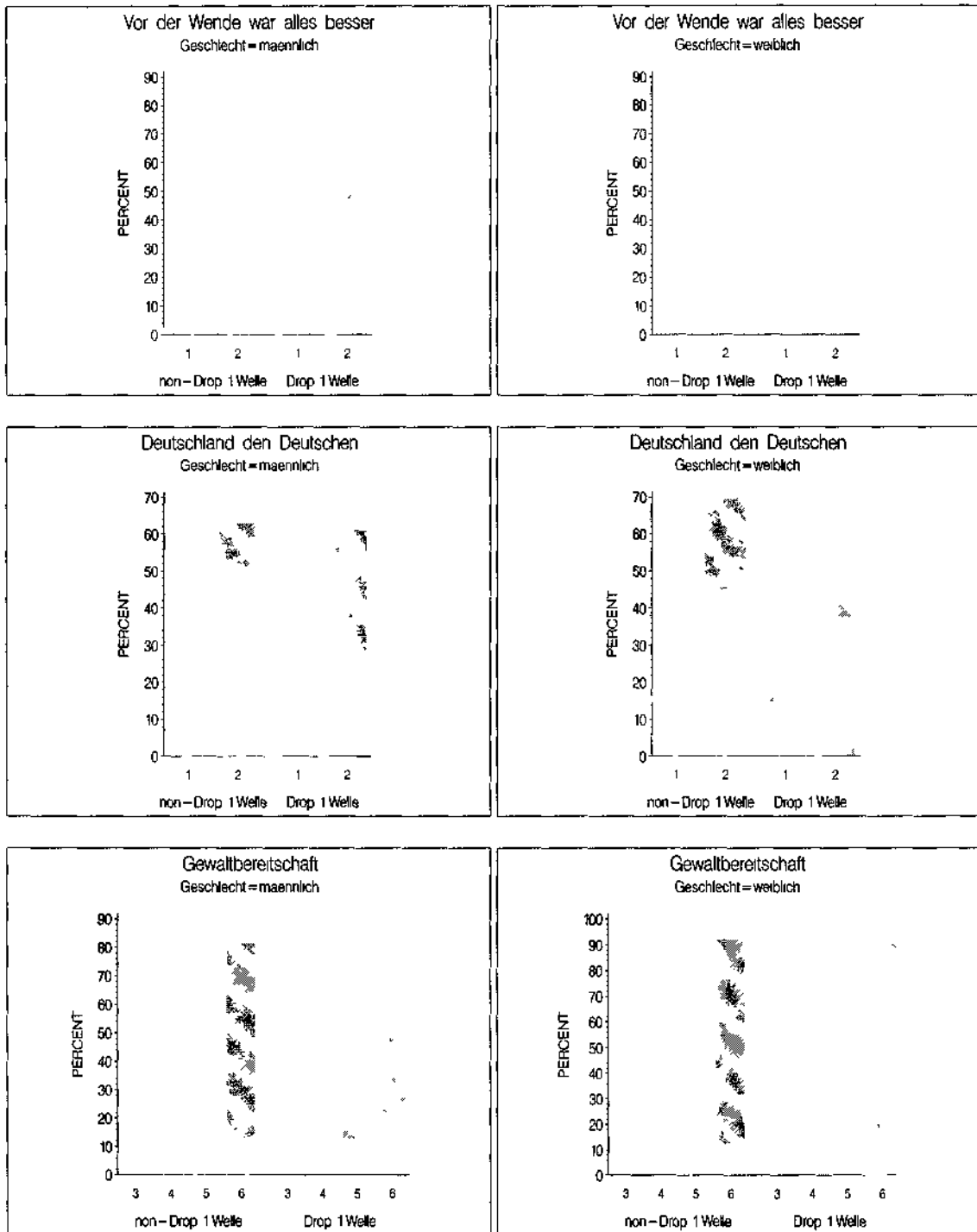


Abbildung 12: Drop-out nach Welle 1: Frage Nr. I/50 Item 1, Nr. I/50 Item 9, Summe aus Nr. I/51

## Drop-out 2. Welle

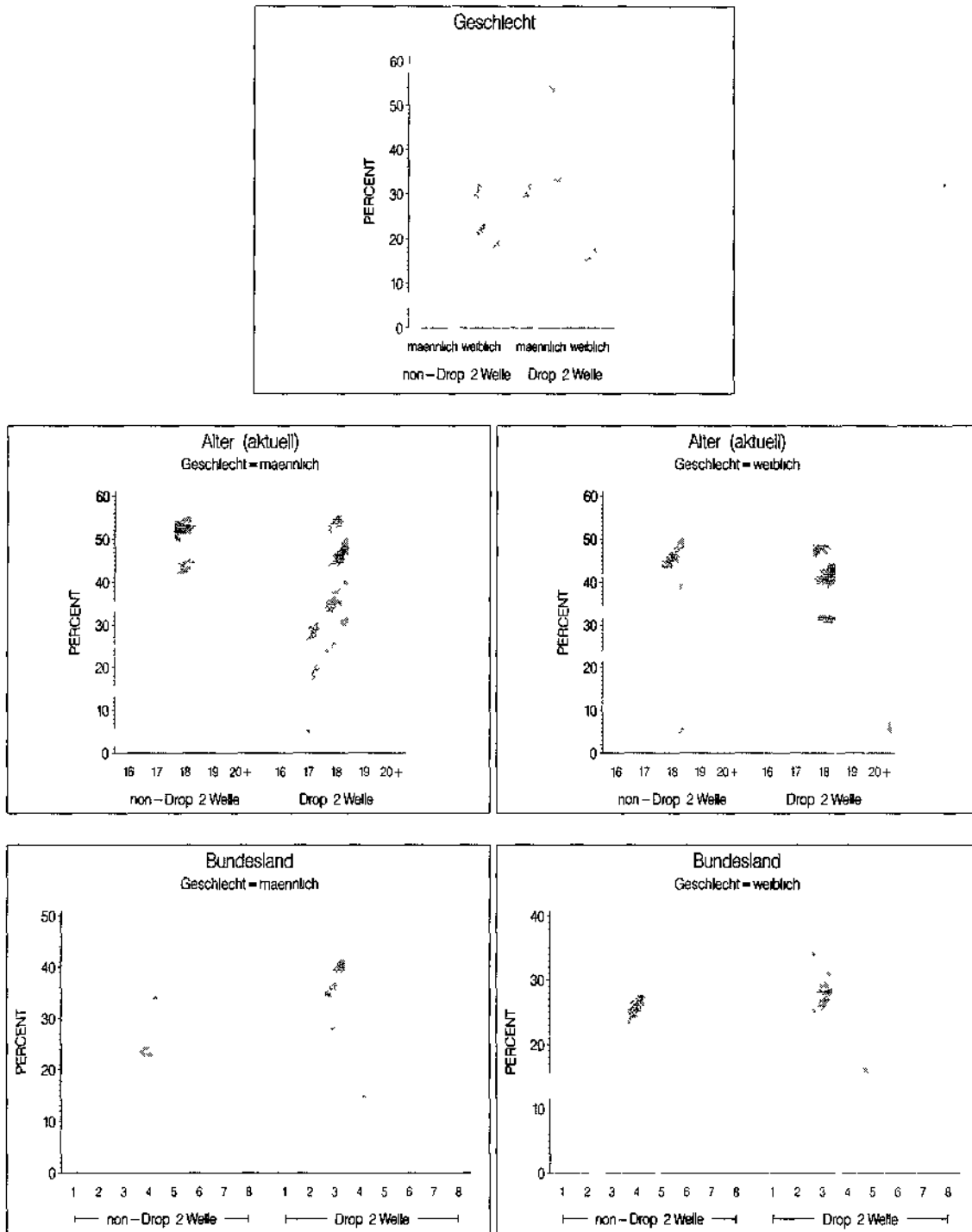


Abbildung 13. Drop-out nach Welle 2: Frage Nr. I/53 , Nr. II/66, Nr. II/70

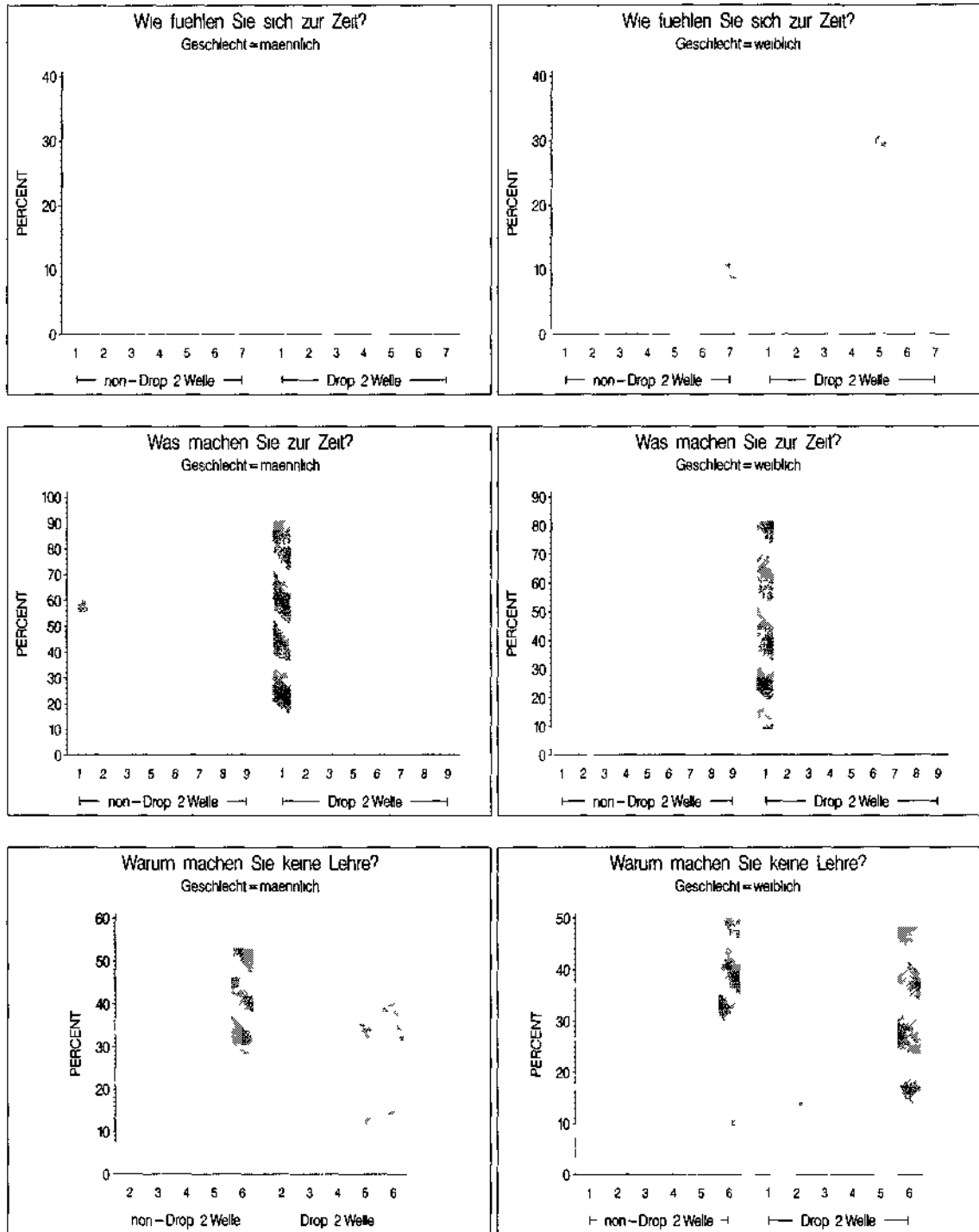


Abbildung 14 Drop-out nach Welle 2 FiageNi II/1, Ni II/2, Nr II/4



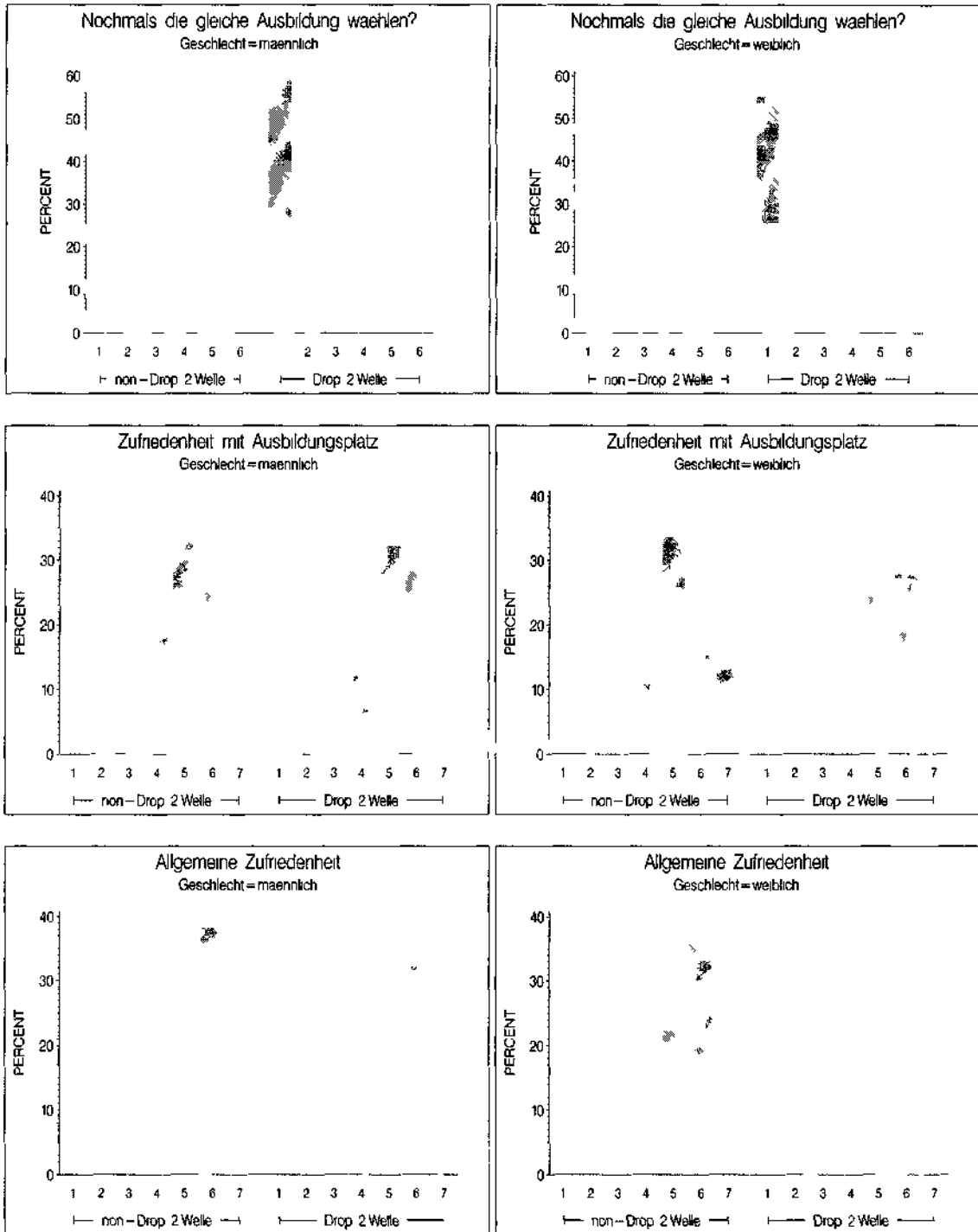


Abbildung 15 Drop-out nach Welle 2 Frage Nr 11/13, Nr 11/16, Nr 11/56

## Drop-out 3. Welle

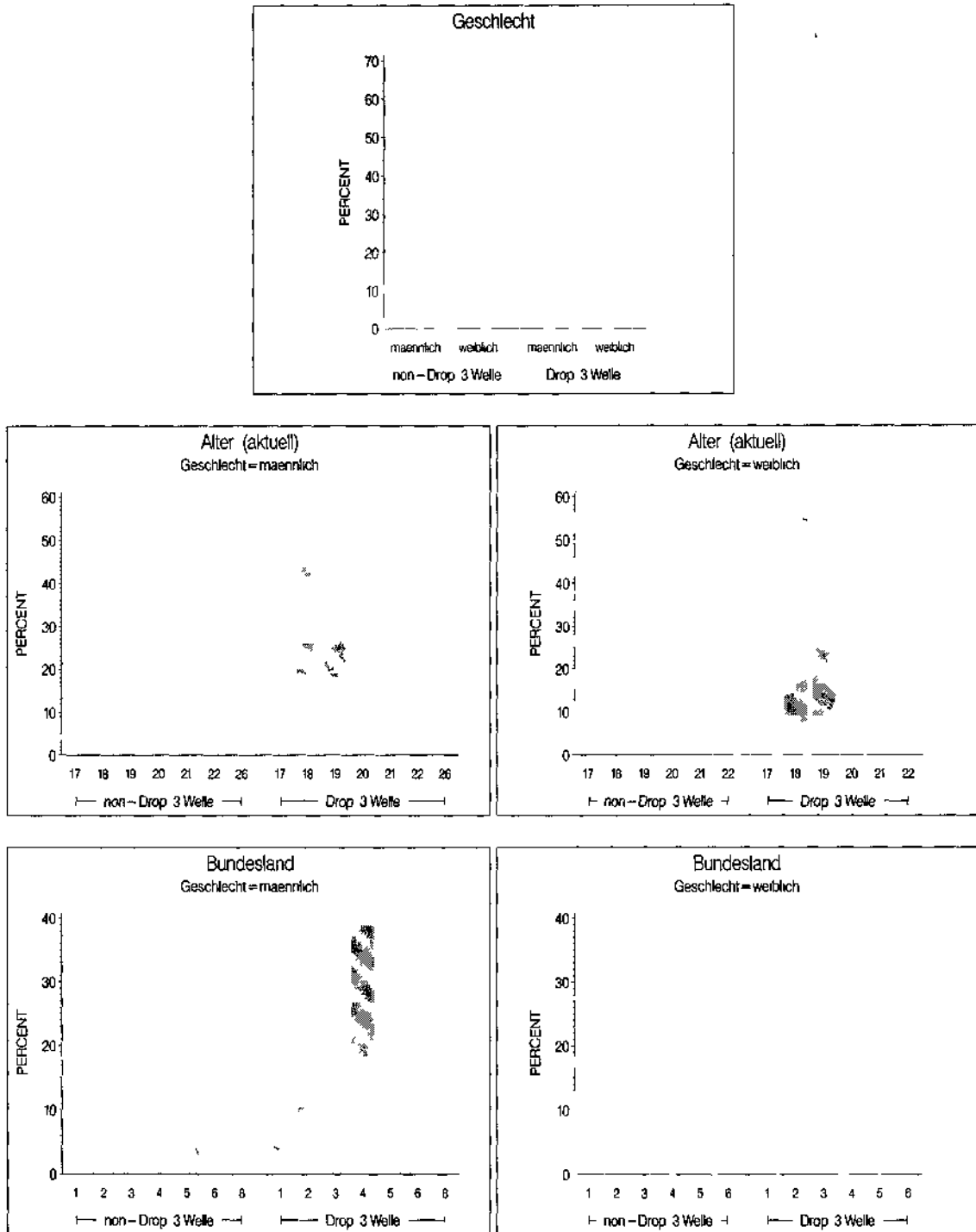


Abbildung 16 Drop-out nach Welle 3 Frage Nr U53, Nr III/72, Nr III/79

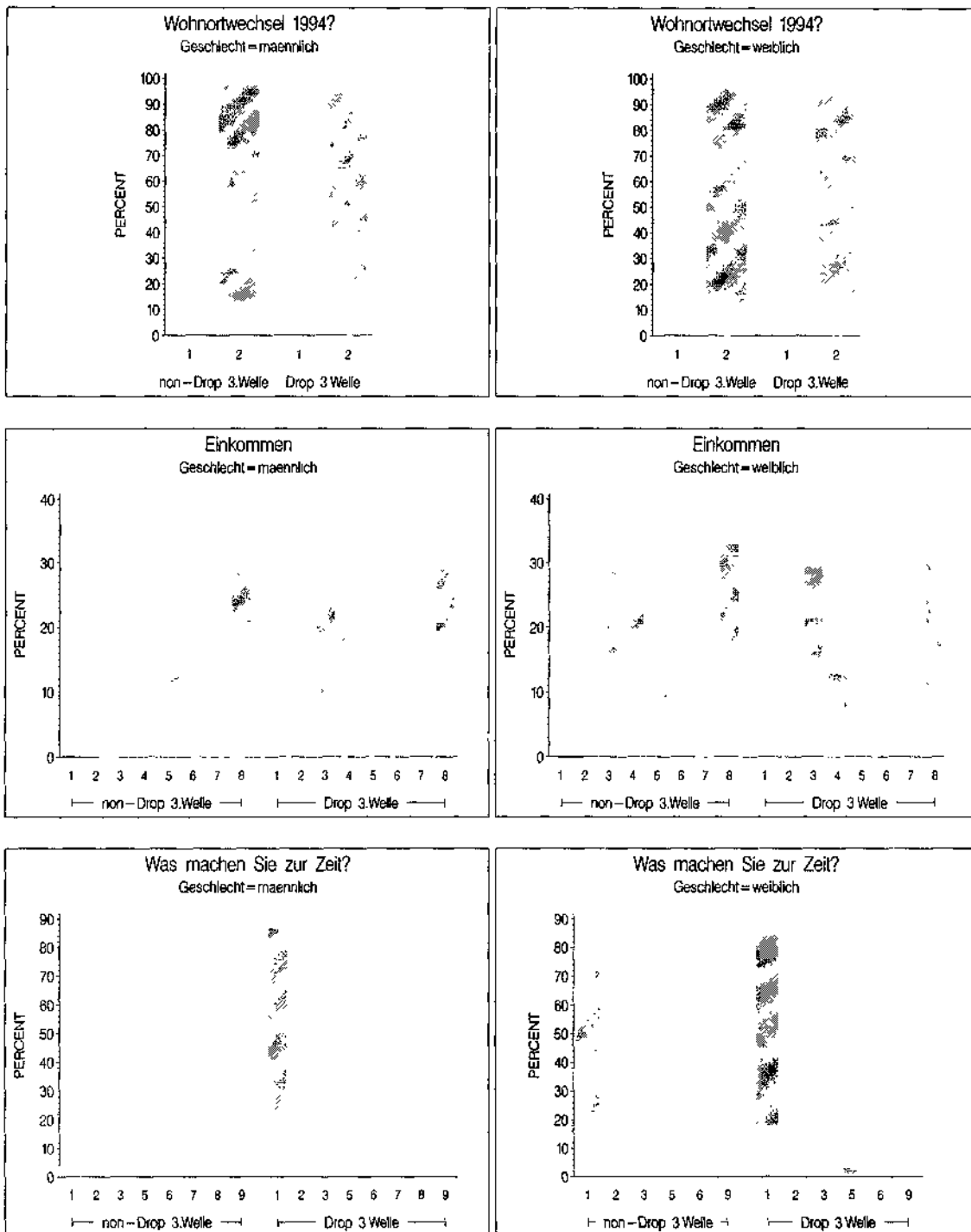


Abbildung 17: Drop-out nach Welle 3: Frage Nr. III/76, Nr. III/75, Nr. III/2

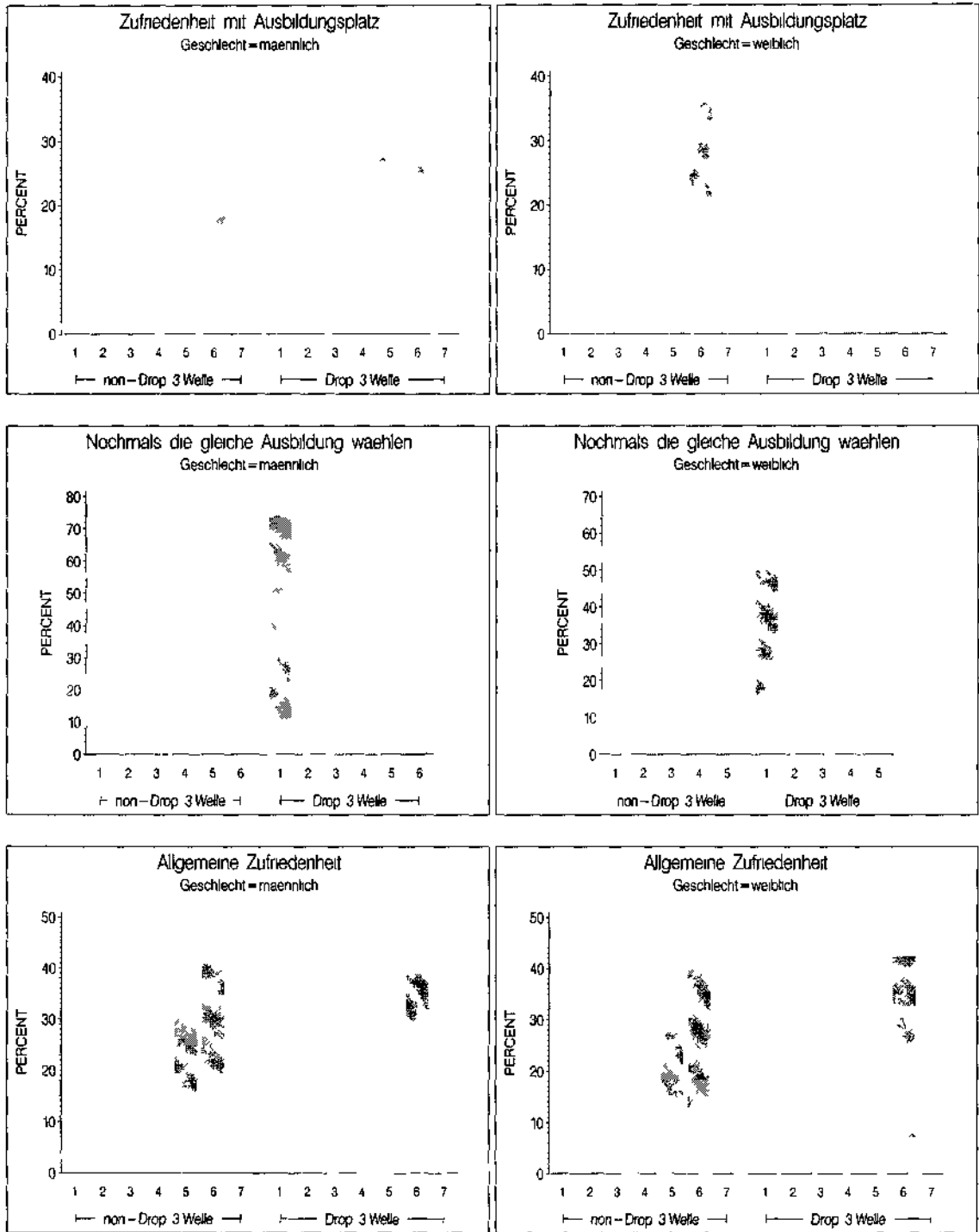


Abbildung 18 Drop-out nach Welle 3 Frage Nr III/20,Nr 111/12, Nr 111/59

## 6 Allgemeiner Einfluß der beruflichen Situation

### 6.1 Gewaltbereitschaft der Jugendlichen

Es soll nun untersucht werden, welchen Einfluß das Innehaben einer Lehrstelle auf die Gewaltbereitschaft von Jugendlichen hat. Die Jugendlichen wurden zu allen drei Zeitpunkten befragt, in welchen von 10 geschilderten Situationen sie bereit wären Gewalt anzuwenden. Die Situationen waren:

- *Wenn Sie Ihre Wohnung verlieren würden*
- *Wenn noch mehr Asylanten kommen*
- *Wenn die Umweltverschmutzung weiter zunimmt*
- *Wenn Sie jemanden beim Diebstahl Ihres Autos oder beim Einbruch in Ihre Wohnung erwischen würden*
- *Wenn Sie sehen würden, daß ein Deutscher von Ausländern bedroht wird*
- *Wenn Sie Ihre Arbeit oder Ihre Lehrstelle verlieren würden*
- *Wenn Sie sehen würden, daß ein Ausländer von Deutschen bedroht wird*
- *Wenn sich die Parteien und Politiker als unfähig erweisen, die Probleme zu lösen*
- *Wenn die Kriminalität weiter zunimmt*
- *Wenn immer mehr Ausländer ins Land kommen*

Die möglichen Antworten waren "ja, auf jeden Fall" (= 1), "vielleicht" (= 0) und "nein, auf keinen Fall" (= -1). Als Gewaltscore eines Jugendlichen definieren wir die Summe der Ausprägungen zu obigen Fragen. Eine Hauptkomponentenanalyse (Abbildung 19) zeigt, daß die einfache Summenbildung als Score durchaus gerechtfertigt erscheint. Die Variablen weisen recht ähnliche Ladungsbeiträge in ihren Hauptkomponenten auf. Nur die Aussagen "Wenn noch mehr Asylanten kommen" und "Wenn immer mehr Ausländer ins Land kommen" auf der einen Seite und "Wenn Sie sehen würden, daß Ausländer von Deutschen bedroht wird" auf der anderen Seite weisen Unterschiede auf. Die erste Hauptkomponente gibt somit die generelle Gewaltbereitschaft wieder, die zweite Hauptkomponente scheint die durch Fremdenfeindlichkeit hervorgerufene Gewaltbereitschaft widerzuspiegeln.

Als Einflußgröße dient die momentane Berufssituation der Jugendlichen, wobei diese in die Ausprägungen "Mache eine Lehre" (= 1), "gehe zur Schule" (= 2), "studiere" (= 3) und "bin unversorgt" (= 4) kategorisiert wurde. Dabei wurden in der Kategorie "bin unversorgt" alle Jugendlichen zusammengefaßt, die einem "BVJ/BVG" nachgehen, "arbeiten/jobben" oder "arbeitslos" sind. Bundeswehr und Zivildienstleistende wurden ebenso vernachlässigt wie Jugendliche, die keine der genannten Kategorien erfüllen. Wir erstellen das marginale Modell

$$E(y_{it}|x_{it}) = \beta t + x_{it}\beta x_i$$

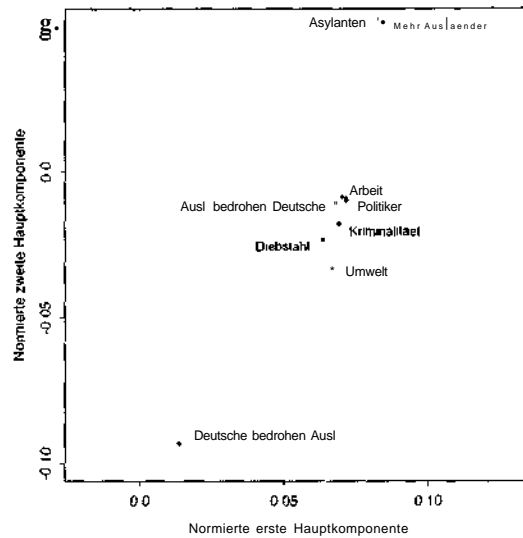


Abbildung 19: Normierte Hauptkomponenten der Fragen zur Gewaltbereitschaft von Jugendlichen

wobei  $y_{it}$  der Gewaltscore des  $i$ ten Individuums zum Zeitpunkt  $t$  und  $x_{it}$  die Berufssituation des Jugendlichen zum Zeitpunkt  $t$  wiedergibt (Ein zeitvariierender Effekt von  $x_{it}$  wurde getestet, erwies sich jedoch als nicht signifikant). Somit ist also der Gewaltscore eines Jugendlichen als abhängig von Zeitpunkt der Befragung und der Berufssituation modelliert. Derartige Modelle gehören in den Kontext von Varianzanalysemodellen. Wir testen, ob die Berufssituation Einfluß auf die Gewaltbereitschaft des Jugendlichen hat, d.h. ob  $\beta_x = 0$  angenommen werden kann. Die Hypothese liefert einen p-Wert von 0,05007, d.h. der Test liefert keine Signifikanz für den Einfluß der beruflichen Situation.

Dies liegt insbesondere an den schwachen Besetzungszahlen der Studenten wie aus Tabelle 20 ersichtlich wird.<sup>2</sup> Faßt man Schüler und Studenten zu einer Gruppe zusammen, so wird der Einfluß der beruflichen Situation mit einem p-Wert von 0,023 signifikant. Parameterschätzungen sind in Tabelle 21 gegeben. Es zeigt sich, daß Auszubildende eine höhere Gewaltbereitschaft aufweisen als Unversorgte oder Schüler und Studenten. Die Gewaltbereitschaft nimmt jedoch generell signifikant mit der Zeit ab. Die Graphiken 20 und 21 verdeutlichen das Ergebnis.

Zeitpunkt	Auszubildende	Schule	Studenten	Unversorgte
1	574	62	5	14
2	587	46	4	18
3	573	40	7	35

Tabelle 20: Verteilung der beruflichen Situation

<sup>2</sup> Aufgeschlüsselt in Tabelle 20 sind nur Jugendliche bei denen Messungen ihrer beruflichen Situation in allen drei Zeitpunkten vorlagen und die zu keinem der Zeitpunkte bei der Bundeswehr oder im Zivildienst waren oder "Sonstiges" als Tätigkeit angaben.

	Schätzung	Std.abweichung	Schätzung	Std.abweichung
$\beta_1$	-1,52	0,16	-1,52	0,16
$\beta_2$	-0,68	0,18	-0,68	0,18
$\beta_3$	-2,73	0,22	-2,74	0,22
$\hat{\rho}^{\wedge}$ Schüler	-0,91	0,38	-0,86	0,36
$\hat{\rho}$ Student	-0,39	1,08		
$\hat{\rho}$ Unversorgte	-0,81	0,51	-0,81	0,51

Tabelle 21: Parameterschätzungen zur Gewaltbereitschaft ohne und mit Zusammenfassung von Schülern und Studenten

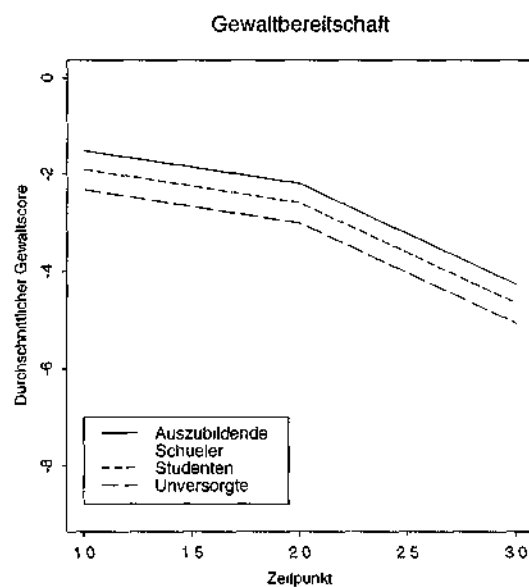


Abbildung 20: Gewaltbereitschaft der Jugendlichen

## 6.2 Allgemeinempfinden

Wir betrachten weiter, welchen Einfluß die berufliche Situation auf das Allgemeinbefinden der Jugendlichen hat. Dazu wird obige Einflußgröße "*Berufssituation*" als Regressor angesehen für die Zielvariable "*Wie fühlen Sie sich zur Zeit*" (mit den Ausprägungen 1 für "*eher schlecht*" bis 7 für "*eher gut*"). Wieder benutzen wir ein marginales Regressionsmodell, es zeigt sich jedoch, daß die berufliche Situation keinen Einfluß auf das persönliche Empfinden hat (p-Wert 0,65).

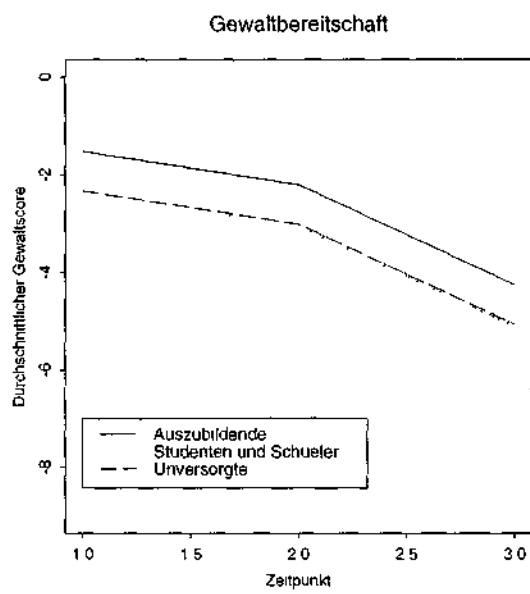


Abbildung 21: Gewaltbereitschaft der Jugendlichen



## 7 Beurteilung der beruflichen Situation

### 7.1 Allgemeine Selbsteinschätzung

Im folgenden Abschnitt soll untersucht werden, welche Faktoren Einfluß auf die Zufriedenheit eines Jugendlichen mit seiner jeweiligen Ausbildung, Arbeits- und Lebenssituation haben. Dabei werden Modelle der Quasiunabhängigkeit (siehe Abschnitt 3) sowie marginale Modellierungen (Abschnitt 4) herangezogen.

Als erstes betrachten wir, wie sich die Jugendlichen derzeit generell einschätzen. Dabei sollten sie eine Selbsteinschätzung mit den möglichen Ausprägungen:

- 1 *"Alles, was ich mache, mache ich gut"*
- 2 *"Ich kann alles so gut wie andere Leute "*
- 3 *"Ich bin ein brauchbarer Mensch"*
- 4 *"Machmal komme ich mir nutzlos vor"*
- 5 *"Mir gelingt nichts richtig "*
- 6 *"Ich kann aufmich nicht stolz sein "*
- 7 *"Ich werde im Leben nie etwas erreichen "*

Die Antwortmöglichkeiten werden als ordinal skaliert betrachtet.

Zunächst untersuchen wir, wie die Jugendlichen ihre Meinung über die ersten drei Jahre hinweg ändern. Dazu betrachten wir, ob in den Tabellen 22 und 23 Quasi-Unabhängigkeit vorliegt. Hier zeigt sich, daß zwischen der ersten und zweiten Periode Quasi-Unabhängigkeit angenommen werden kann (p-Wert 0.161), d.h. daß unter den Jugendlichen, die ihre Meinung ändern, keine Struktur nachweisbar ist. Anders sieht es zwischen zweiter und dritter Periode aus (p-Wert 0.01). In Tabelle 24 sind die Residuen  $(/i_i; - \hat{h}^x/\hat{h}^{ij}$  gegeben, wobei  $h_{ij}$  die beobachteten Häufigkeiten wie sie in Tabelle 23 gegeben sind und  $\hat{h}^x$  die unter Quasi-Unabhängigkeit erwarteten. Eine klare Struktur der Residuen ist nicht erkennbar, tendenziell läßt sich jedoch sagen, daß Jugendliche die sich in Perioden 2 schlecht eingeschätzt haben (Kategorie 4-6) auch in Periode 3 zu diesem Urteil neigen (hauptsächlich negative Residuen in unterer linken Hälfte der Tafel) was in umgekehrter Form auch für Jugendliche gilt, die sich in Periode 2 gut gefühlt haben (hauptsächlich negative Residuen im rechten oberen Teil der Tabelle). Somit verstärkt sich im Laufe der Zeit der Effekt, daß die Jugendlichen sich in zwei Gruppen teilen lassen, diejenigen, die sich gut fühlen (Kategorie 1-3) und diejenigen, die sich unzufrieden fühlen.

Zeitpunkt1	Zeitpunkt 2							Gesamt
	Was ich mache ist gut	Bin so gut wie andere	Bin ein brauchbarer Mensch	Komme mir manchmal nutzlos vor	Mir gelingt nichts richtig	Kann nicht stolz sein	Ich werde nie etwas erreichen	
Was ich mache ist gut	16 1.54	15 1.44	28 2.69	2 0.19	0 0.00	0 0.00	0 0.00	61 5.87
Bin so gut wie andere	26 2.50	93 8.94	99 9.52	7 0.67	0 0.00	1 0.10	0 0.00	226 21.73
Bin ein brauchbarer Mensch	48 4.62	161 15.48	410 39.42	30 2.88	2 0.19	3 0.29	0 0.00	654 62.88
Komme mir manchmal nutzlos vor	9 0.87	14 1.35	42 4.04	9 0.87	0 0.00	1 0.10	1 0.10	76 7.31
Mir gelingt nichts richtig	0 0.00	2 0.19	1 0.10	3 0.29	0 0.00	0 0.00	0 0.00	6 0.58
Kann nicht stolz sein	0 0.00	8 0.77	8 0.77	0 0.00	0 0.00	0 0.00	0 0.00	16 1.54
Ich werde nie etwas erreichen	0 0.00	1 0.10	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	1 0.10
Gesamt	99 9.52	294 28.27	588 56.54	51 4.90	2 0.19	5 0.48	1 0.10	1040 100.00

Tabelle 22: Selbsteinschätzung der Jugendlichen in Periode 1 und 2. Absolute Häufigkeit und Prozent

Zeitpunkt 2	Zeitpunkt 3							Gesamt
	Was ich mache ist gut	Bin so gut wie andere	Bin ein brauchbarer Mensch	Komme mir manchmal nutzlos vor	Mir gelingt nichts richtig	Kann nicht stolz sein	Ich werde nie etwas erreichen	
Was ich mache ist gut	21 3.05	22 3.20	19 2.76	1 0.15	0 0.00	1 0.15	0 0.00	64 9.30
Bin so gut wie andere	47 6.83	74 10.76	74 10.76	2 0.29	0 0.00	0 0.00	0 0.00	197 28.63
Bin ein brauchbarer Mensch	54 7.85	122 17.73	196 28.49	16 2.33	1 0.15	5 0.73	1 0.15	395 57.41
Komme mir manchmal nutzlos vor	3 0.44	5 0.73	14 2.03	5 0.73	0 0.00	0 0.00	0 0.00	27 3.92
Mir gelingt nichts richtig	0 0.00	0 0.00	0 0.00	2 0.29	0 0.00	0 0.00	0 0.00	2 0.29
Kann nicht stolz sein	0 0.00	3 0.44	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	3 0.44
Ich werde nie etwas erreichen	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00
Gesamt	125 18.17	226 32.85	301 44.04	26 3.78	1 0.15	6 0.87	1 0.15	688 100.00

Tabelle 23: Selbsteinschätzung der Jugendlichen in Periode 2 und 3. Absolute Häufigkeit und Prozent

Zeitpunkt 2	Zeitpunkt 3					
	Was ich mache ist gut	Bin so gut wie andere	Bin ein brauchbarer Mensch	Komme mir manchmal nutzlos vor	Mir gelingt nichts richtig	Negativ
Was ich mache ist gut		0.46	-0.32	-0.67	-0.30	0.48
Bin so gut wie andere	1.58		-0.21	-1.93	-0.57	-1.52
Bin ein brauchbarer Mensch	-0.94	0.04		1.24	0.63	1.14
Komme mir manchmal nutzlos vor	-0.64	-1.22	<b>1.76</b>		-0.20	-0.52
Mir gelingt nichts richtig	-0.62	-0.87	-0.88	<b>7.15</b>		-0.15
Negativ	-0.76	<b>1.74</b>	-1.08	-0.33	-0.07	

Tabelle 24: Residuen (signifikante Werte [bei  $\alpha = 10\%$ ] sind fettgedruckt). Ausprägung "Negativ" umfaßt "kann nicht stolz sein" und "Ich werde nie etwas erreichen"

Nun soll untersucht werden, wie sich die durchschnittliche Zufriedenheit der Jugendlichen über die Zeit hinweg verändert. Hierzu fassen wir Kategorien 4, 5, 6 und 7 wegen schwacher Besetzungszahlen zusammen und kategorisieren diese Jugendlichen als sich schlecht Fühlende. Wir betrachten ein marginales Modell

$$E(y_{it}|x_i, t) = h(\beta_t + x_i\beta_x)$$

wobei  $y_{it}$  hier die angegebene Zufriedenheit zum Zeitpunkt  $t$  ist und das Geschlecht als möglichen Regressor  $x_i$  dient. Die Zufriedenheit ist somit in Abhängigkeit von Zeit und Geschlecht modelliert. Es zeigt sich jedoch, daß der Einfluß des Geschlechts nicht signifikant ist. Durch die ordinale Skalierung der Selbsteinschätzung bedingt verwenden wir ein proportionales kumulatives Logit Modell. Es ergibt sich jedoch, daß dieses Modell die Daten nicht ausreichend beschreibt. Das bedeutet, daß sich die Zuordnung zu den einzelnen Kategorien 1-4 nicht proportional über die Zeit verändert. Wie aus Tabelle 25 als auch aus Abbildung 22 ersichtlich ist, nimmt der Anteil der Jugendlichen, die sich gut (Kategorie 1) oder so gut wie andere (Kategorie 2) fühlen, klar zu und der Anteil der Jugendlichen, die sich brauchbar fühlen (Kategorie 3), verringert sich. Dennoch fühlen sich sowohl in der zweiten auch in der dritten Periode ca. 5% der Jugendlichen schlecht (Kategorie 4) (Verletzung der kumulierten Logit Annahme). Somit ist kein genereller positiver Trend in der Selbsteinschätzung nachweisbar, sofern man die sich schlecht fühlenden Jugendlichen mit berücksichtigt. Betrachtet man dagegen nur die Jugendlichen, die sich zumindest brauchbar fühlen, so ergibt sich eine klare positive Tendenz in der Zufriedenheit.

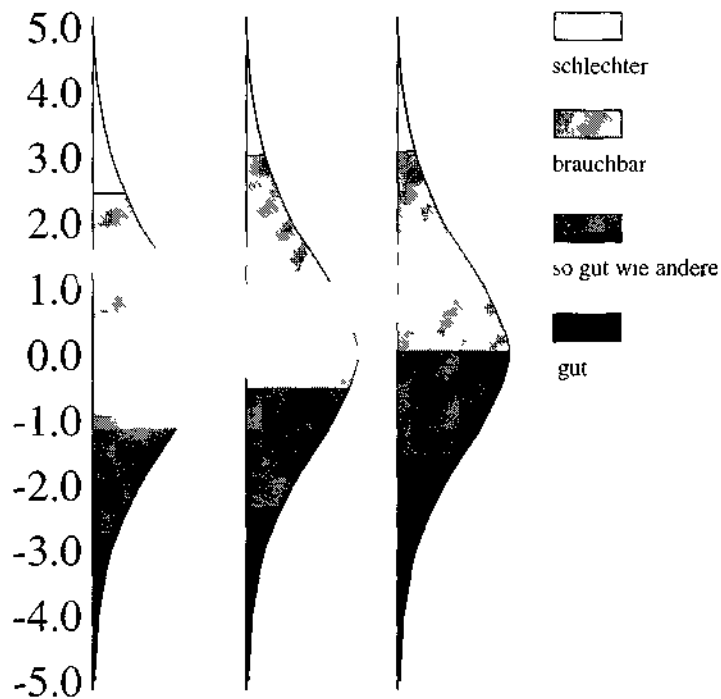


Abbildung 22. Selbsteinschätzung in Zeitpunkten 1, 2 und 3

	Geschlecht	Kategorie	Zeitpunkt		
			1	2	3
beobachtet	männlich	1	0,057	0,089	0,165
		2	0,197	0,281	0,337
		3	0,677	0,591	0,463
		≥4	0,069	0,039	0,034
	weiblich	1	0,064	0,100	0,207
		2	0,236	0,293	0,314
		3	0,607	0,550	0,411
		≥4	0,093	0,057	0,068
modelliert	männlich & weiblich	1	0,061	0,092	0,183
		2	0,211	0,287	0,331
		3	0,650	0,576	0,443
		≥4	0,078	0,045	0,043

Tabelle 25: Randhäufigkeit für Selbsteinschätzung

## 7.2 Zufriedenheit mit Ausbildung

Als nächstes betrachten wir, ob Jugendliche mit der jeweiligen Ausbildung im Beruf zufrieden sind. Als mögliche Antworten auf die Frage "Wie empfinden Sie Ihre jetzige Ausbildung" stehen den Jugendlichen die Antworten:

- 1 "Ich möchte lieber heute als morgen aufhören"
- 2 "Es fällt mir schwer"
- 3 "Es geht so"
- 4 "Es gefällt mir gut"
- 5 "Es macht richtig Spaß"

zur Auswahl. In Tabelle 26 und 27 geben wir die entsprechenden Daten wieder. Wir untersuchen, ob zwischen den einzelnen Perioden Quasi-Unabhängigkeit vorliegt. Das bedeutet, wir untersuchen, ob unter den Jugendlichen, die ihre Meinung zu obiger Frage von einer Periode zur nächsten ändern, Struktur vorhanden ist. Wir betrachten somit nur die Jugendlichen, die ihre Meinung wechseln (Wechsler). Es zeigt sich, daß zwischen Perioden 1 und 2 der Meinungswechsel strukturlos ist (p-Wert 0.35), zwischen Periode 2 und 3 hängt ein Wechsel der Einstellung jedoch signifikant von der Einstellung in der Vorperiode ab (p-Wert 0.03).

Zeitpunkt 1	Zeitpunkt 2					Gesamt
	Lieber heute als morgen aufhören	Fällt recht schwer	Es geht, ist notwendiges Übel	Es ist ganz gut	Es macht richtig Spaß	
Lieber heute als morgen aufhören	0 0.00	1 0.12	4 0.48	9 1.09	4 0.48	18 2.18
Fällt recht schwer	2 0.24	0 0.00	2 0.24	12 1.45	1 0.12	17 2.06
Es geht, ist notwendiges Übel	1 0.12	0 0.00	9 1.09	21 2.55	5 0.61	36 4.36
Es ist ganz gut	9 1.09	13 1.58	40 4.85	212 25.70	62 7.52	336 40.73
Es macht richtig Spaß	5 0.61	9 1.09	26 3.15	227 27.52	151 18.30	418 50.67
Gesamt	17 2.06	23 2.79	81 9.82	481 58.30	223 27.03	825 100.00

Tabelle 26: "Wie empfinden Sie jetzige Ausbildung im Betrieb", Zeitpunkte 1 und 2. Absolute Häufigkeit und Prozent

Zeitpunkt 2	Zeitpunkt 3					Gesamt
	Lieber heute als morgen aufhören	Fällt recht schwer	Es geht, ist notwendiges Übel	Es ist ganz gut	Es macht richtig Spaß	
Lieber heute als morgen aufhören	0 0.00	0 0.00	3 0.54	6 1.07	1 0.18	10 1.79
Fällt recht schwer	0 0.00	2 0.36	3 0.54	5 0.89	4 0.72	14 2.50
Es geht, ist notwendiges Übel	2 0.36	3 0.54	11 1.97	29 5.19	9 1.61	54 9.66
Es ist ganz gut	1 0.18	4 0.72	35 6.26	173 30.95	112 20.04	325 58.14
Es macht richtig Spaß	0 0.00	1 0.18	8 1.43	64 11.45	83 14.85	156 27.91
Gesamt	3 0.54	10 1.79	60 10.73	277 49.55	209 37.39	559 100.00

Tabelle 27: "Wie empfinden Sie jetzige Ausbildung im Betrieb", Zeitpunkte 2 und 3

Wir betrachten nun die Verteilung der Zufriedenheit im Querschnitt über die Beobachtungspunkte (marginale Modellierung). Wegen der schwachen Besetzungszahlen fassen wir die ersten beiden Kategorien zu einer Kategorie zusammen. Als zusätzliche Einflußgröße ziehen wir das Geschlecht einer Person heran. Das bedeutet wir modellieren

$$E(y_{it}|x_i, t) = h(\beta_0 + X_i\beta_x),$$

wobei  $y_{it}$  hier die Zufriedenheit mit der Ausbildung zum Zeitpunkt  $t$  angibt und  $X_i$  das Geschlecht ist. Die Zufriedenheit mit der Ausbildung wird also als abhängig von Geschlecht und Zeit angenommen. Die ordinale Struktur von  $y_{it}$  verlangt die Anwendung eines kumulativen Logit Modells. Es zeigt sich, daß sowohl die Zeit, als auch das Geschlecht signifikanten Einfluß auf die Zufriedenheit haben. Es zeigt sich weiter, daß das Modell mit proportionalen kumulativen Logits nicht aufrecht erhalten werden kann, denn die Residuen weisen (signifikante) Struktur auf. Das bedeutet inhaltlich, daß sich die Klasseneinteilungen in den vier ordinalen Kategorien "Ich möchtelieber heute als morgen aufhören", "Es geht so", "Es gefällt mir gut" und "Es macht richtig Spaß" unproportional über die Zeit verändert. Aus diesem Grund verwenden wir ein Modell, in dem die Kategorien mit unterschiedlichen Parametern modelliert werden. Die marginalen Wahrscheinlichkeiten der einzelnen Gruppen sind in Abbildung 23 dargestellt, die entsprechenden relativen Häufigkeiten sind in Tabelle 28 gegeben. In Tabelle 29 geben wir die empirischen Signifikanzen der entsprechenden Tests an. Es zeigt sich, daß Männer deutlich öfter die Kategorie "Es gefällt mir gut" wählen als Frauen (Geschlechtseffekt). Ferner variiert die positive Einstellung "Es macht Spaß" deutlich über die Zeit (Zeiteffekt), wobei sie in der zweiten Periode ihr Minimum annimmt. Dabei ist die Veränderung in den anderen Kategorien insbesondere zwischen der zweiten und dritten Periode weitaus weniger variabel (kein proportionales kumulatives Logit Modell). In der zweiten Periode sind die meisten Jugendlichen der



Meinung, daß die Ausbildung "ganz gut" ist, wobei die positive Einstellung (Kategorien 3 und 4) von der zweiten zur dritten Periode kaum Veränderung erfährt. Es unterscheiden sich aber die positiv als auch die negativ urteilenden Jugendlichen in sich. Die positiv Urteilenden neigen zur besten Kategorie ("Spaß") von der zweiten zur dritten Periode. Außerdem nimmt der Anteil der sehr negativ urteilenden Jugendlichen (Kategorie 0 und 1) von der zweiten zur dritten Periode ab.

	Geschlecht	Kategorie	Zeitpunkt		
			1	2	3
beobachtet	männlich	≤ 1	0,031	0,043	0,019
		2	0,037	0,102	0,115
		3	0,433	0,591	0,523
		4	0,498	0,263	0,344
	weiblich	≤ 1	0,051	0,046	0,036
		2	0,046	0,097	0,108
		3	0,333	0,549	0,451
		4	0,569	0,308	0,405
modelliert	männlich	≤ 1	0,032	0,039	0,019
		2	0,039	0,099	0,111
		3	0,424	0,603	0,524
		4	0,505	0,260	0,346
	weiblich	≤ 1	0,048	0,054	0,035
		2	0,042	0,102	0,114
		3	0,350	0,529	0,450
		4	0,560	0,315	0,401

Tabelle 28: Randhäufigkeiten für Zufriedenheit mit Ausbildung

$H_0$ versus $H_1$			
$H_0$	# i	p-Wert	Testergebnis
Zeit 0 Geschlecht	Zeit	0.04	Zeit © Geschlecht
Zeit © Geschlecht	Geschlecht	<0.01	Zeit © Geschlecht

Tabelle 29: Signifikanzen für Zufriedenheit mit der Ausbildung

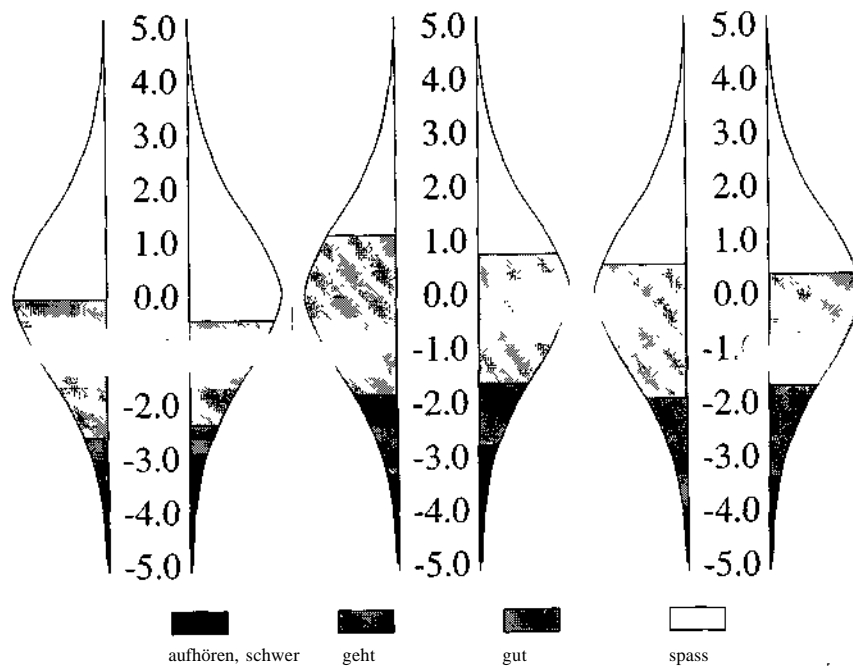


Abbildung 23: Zufriedenheit mit der Ausbildung (links Männer, rechts Frauen), zu Zeitpunkten 1, 2 und 3

Wir wollen nun untersuchen, welchen Einfluß die Betriebsgröße auf die Zufriedenheit der Jugendlichen hat. Dazu teilen wir die Größe eines Betriebs in die Kategorien Kleinbetrieb (bis unter 10 Beschäftigte), mittelgroßer Betrieb (10 bis unter 100 Beschäftigte) und über 100 Beschäftigte ein. Auch hier zeigt sich, daß ein proportionales kumulatives Logit Modell nicht adäquat ist, so daß jede Kategorie der Zufriedenheit getrennt modelliert werden muß. Die entsprechenden Daten finden sich in Tabelle 30, die zugehörigen Tests in Tabelle 31, eine graphische Repräsentation liefern die Abbildungen 24-26. Es zeigt sich, daß Jugendliche in kleinen und mittelgroßen Betrieben weitaus eher zur vollständigen Zufriedenheit {"Macht mir Spaß"} neigen als in großen Betrieben, wo man eher zur Antwort {"Gefällt mir gut"} neigt (Betriebsgrößeneffekt). Die zeitliche Veränderung vollzieht sich dabei ähnlich wie oben bereits geschildert, d.h. die Einstellung "Macht mir Spaß" variiert über die Zeit und weicht insbesondere in der zweiten Periode der Einstellung "Gefällt mir gut" (Zeiteffekt).

	Betriebsgröße	Kategorie	Zeitpunkt		
			1	2	3
beobachtet	klein	≤ 1	0,050	0,040	0,040
		2	0,020	0,080	0,110
		3	0,300	0,550	0,430
		4	0,630	0,330	0,420
	mittel	≤ 1	0,033	0,047	0,024
		2	0,038	0,128	0,104
		3	0,351	0,536	0,498
		4	0,578	0,289	0,374
	groß	≤ 1	0,029	0,046	0,023
		2	0,640	0,087	0,127
		3	0,486	0,613	0,532
		4	0,422	0,254	0,318
modelliert	klein	≤ 1	0,045	0,050	0,035
		2	0,024	0,086	0,094
		3	0,325	0,505	0,436
		4	0,607	0,359	0,434
	mittel	≤ 1	0,034	0,039	0,025
		2	0,043	0,105	0,113
		3	0,366	0,547	0,478
		4	0,556	0,309	0,384
	groß	≤ 1	0,034	0,039	0,024
		2	0,048	0,110	0,119
		3	0,443	0,624	0,555
		4	0,475	0,227	0,302

Tabelle 30: Randhäufigkeiten für Zufriedenheit mit Ausbildung

<i>H<sub>0</sub></i> versus <i>H<sub>1</sub></i>			
<i>H<sub>0</sub></i>	# <i>i</i>	p-Wert	Testergebnis
Zeit © Betriebsgröße	Zeit	< 0.001	Zeit © Betriebsgröße
Zeit © Betriebsgröße	Betriebsgröße	0.017	Zeit © Betriebsgröße

Tabelle 31: Signifikanzen für Zufriedenheit mit Ausbildung

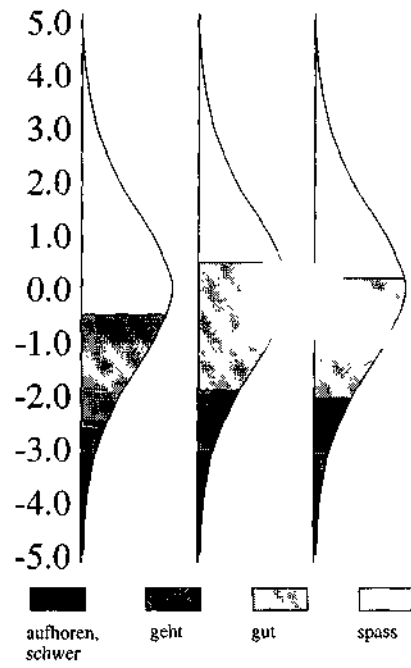


Abbildung 24: Kleinbetriebe

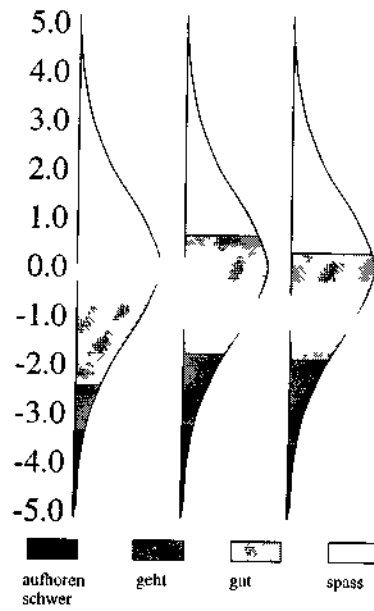


Abbildung 25. Mittlere Betriebe

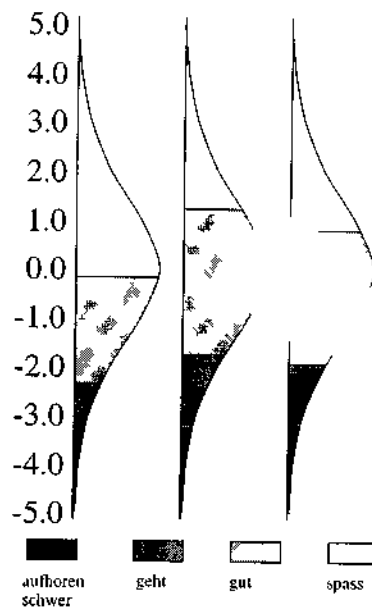


Abbildung 26 Großbetriebe

### 7.3 Zufriedenheit mit Berufsschule

Wir wollen nun untersuchen, welche Faktoren die Zufriedenheit mit der Berufsschule beeinflussen. Dabei konnten die Jugendlichen ihre Zufriedenheit ausdrücken durch die Antworten:

1: *"Ich möchte lieber heute als morgen aufhören "*

2: *"Es fällt schwer"*

2: *"Es geht so"*

3: *"Es gefällt mir gut"*

4: *"Es macht richtig Spaß "*

Zunächst untersuchen wir, inwieweit die Zufriedenheit in einer Periode von der Einstellung der Vorperiode beeinflusst wird. Dazu betrachten wir die jeweils kreuzklassifizierte Ausprägungen in Tabelle 32 und 33 und testen auf Quasi-Unabhängigkeit. Es zeigt sich zwischen Perioden 1 und 2 eine signifikante Struktur unter den Wechslern (p-Wert 0.044) zwischen Perioden 2 und 3 liegt dagegen Quasi-Unabhängigkeit vor (p-Wert 0.52). Eine Residuenanalyse ergibt, daß Jugendliche zwischen der ersten und zweiten Periode ihrer generellen Meinung eher treu bleiben. In Tabelle 34 sind die Residuen  $(h_{ij} - \hat{K}^i \hat{L}_j \hat{K}_j)$  gegeben, wobei wie gehabt  $h_{ij}$  die Besetzungszahl der Tafel und  $\hat{h}_3$  die unter Quasi-Unabhängigkeit erwartete Anzahl von Jugendlichen ist. Man erkennt, daß Jugendliche mit einer guten Meinung *"Es ist ganz gut"* bzw. *"Es macht richtig Spaß "* in der nächsten Periode wieder zu einer positiven Antwort neigen (negative Residuen linke untere Tafel) wohingegen eher unzufriedene Jugendliche ihre negative Meinung beibehalten, ja sogar verstärken (positive Residuen in Zellen (2, 1) und (3, 1) sowie negative Residuen in rechter oberer Tafel).

Zeitpunkt 1	Zeitpunkt 2					Gesamt
	Lieber heute als morgen aufhören	Fällt recht schwer	Es geht, ist notwendiges Übel	Es ist ganz gut	Es macht richtig Spaß	
Lieber heute als morgen aufhören	1 0.12	0 0.00	4 0.49	10 1.22	3 0.36	18 2.19
Fällt recht schwer	3 0.36	4 0.49	17 2.07	15 1.82	4 0.49	43 5.22
Es geht, ist notwendiges Übel	9 1.09	9 1.09	102 12.39	92 11.18	15 1.82	227 27.58
Es ist ganz gut	8 0.97	15 1.82	122 14.82	241 29.28	48 5.83	434 52.73
Es macht richtig Spaß	2 0.24	3 0.36	16 1.94	56 6.80	24 2.92	101 12.27
Gesamt	23 2.79	31 3.77	261 31.71	414 50.30	94 11.42	823 100.00

Tabelle 32: "Wie empfinden Sie den Besuch der Berufsschule?", Zeitpunkte 1 und 2. Absolute Häufigkeiten und Prozent

Zeitpunkt 2	Zeitpunkt 3					Gesamt
	Lieber heute als morgen aufhören	Fällt recht schwer	Es geht, ist notwendiges Übel	Es ist ganz gut	Es macht richtig Spaß	
Lieber heute als morgen aufhören	0 0.00	1 0.18	6 1.07	7 1.25	1 0.18	15 2.67
Fällt recht schwer	0 0.00	1 0.18	6 1.07	11 1.96	3 0.53	21 3.74
Es geht, ist notwendiges Übel	4 0.71	5 0.89	59 10.52	97 17.29	10 1.78	175 31.19
Es ist ganz gut	3 0.53	7 1.25	60 10.70	188 33.51	30 5.35	288 51.34
Es macht richtig Spaß	0 0.00	1 0.18	14 2.50	32 5.70	15 2.67	62 11.05
Gesamt	7 1.25	15 2.67	145 25.85	335 59.71	59 10.52	561 100.00

Tabelle 33: "Wie empfinden Sie den Besuch der Berufsschule?", Zeitpunkte 2 und 3. Absolute Häufigkeiten und Prozent

Zeitpunkt 1	Zeitpunkt 2				
	Lieber heute als morgen aufhören	Fällt recht schwer	Es geht, ist notwendiges Übel	Es ist ganz gut	Es macht richtig Spaß
Lieber heute als morgen aufhören		-0.82	-0.41	0.15	0.82
Fällt recht schwer	1.59		1.67	-1.51	-0.16
Es geht, ist notwendiges Übel	1.62	0.89		-0.25	-0.83
Es ist ganz gut	-1.36	-0.32	0.31		0.41
Es macht richtig Spaß	-0.38	-0.18	-1.66	1.33	

Tabelle 34: Wie empfinden Sie ihre jetzige Ausbildung? Residuen



Als nächstes betrachten wir, wie sich die durchschnittliche Zufriedenheit über die Zeit hinweg ändert. Man stellt fest, daß kein Geschlechtseffekt signifikant nachweisbar ist (p-Wert 0,60), wohl aber ein Zeiteffekt. Die entsprechenden modellierten Häufigkeiten finden sich in Tabelle 35, die graphische Darstellung ist in Abbildung 27 gegeben. Es zeigt sich, daß die Jugendlichen zum zweiten Zeitpunkt die Berufsschule tendenziell schlechter beurteilen als zum ersten und dritten Zeitpunkt, zu denen eine annähernd gleiche Meinung der Jugendlichen vorliegt (Zeiteffekt).

Ebenso wie kein Geschlechtseffekt nachweisbar ist, ergibt sich, daß auch kein Betriebsgrößeneffekt nachweisbar ist (p-Wert 0,26). Das heißt, Jugendliche in kleinen wie großen Betrieben weisen die gleiche Einstellung gegenüber der Berufsschule auf (siehe Tabelle 36).

	Geschlecht	Kategorie	Zeitpunkt		
			1	2	3
beobachtet	männlich	≤ 1	0,065	0,062	0,040
		2	0,252	0,329	0,267
		3	0,571	0,494	0,606
		4	0,112	0,115	0,087
	weiblich	≤ 1	0,042	0,078	0,042
		2	0,276	0,302	0,229
		3	0,526	0,521	0,594
		4	0,156	0,099	0,135
modelliert	männlich & weiblich	≤ 1	0,051	0,067	0,050
		2	0,264	0,311	0,259
		3	0,558	0,523	0,561
		4	0,126	0,099	0,130

Tabelle 35: Randhäufigkeit für Zufriedenheit mit Berufsschule

<i>H<sub>0</sub> versus H<sub>x</sub></i>			
<i>H<sub>0</sub></i>	# i	p-Wert	Testergebnis
Zeit © Geschlecht	Zeit	0,60	Zeit
Zeit © Betriebsgröße	Zeit	0,26	Zeit

Tabelle 36: Empirische Signifikanzen für Zufriedenheit mit Berufsschule

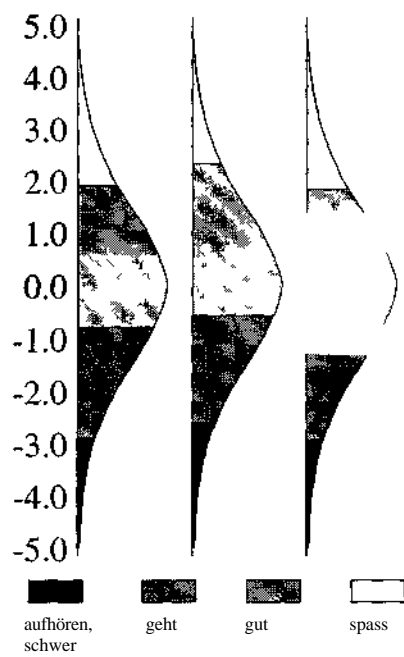


Abbildung 27: Zufriedenheit mit der Berufsschule zu Zeitpunkten 1, 2 und 3

## 7.4 Zufriedenheit mit Berufswahl

Die Jugendlichen wurden weiter gefragt, ob sie mit ihrer Berufswahl zufrieden sind. Dabei standen ihnen jeweils in der zweiten, dritten und vierten Periode die möglichen Antworten:

- 1: *"Ich würde die gleiche Ausbildung im selben Betrieb noch einmal machen "*
- 2: *"Ich würde die gleiche Ausbildung, jedoch in einem anderen Betrieb noch einmal machen"*
- 3: *"Ich würde einen anderen Beruf erlernen"*
- 4: *"Ich würde zur Schule oder Hochschule gehen "*
- 5: *"Ich würde ohne Ausbildung gleich arbeiten gehen "*
- 6: *"Ich würde etwas ganz anderes machen "*

zur Auswahl. Die möglichen Antworten *"Ich würde ohne Ausbildung gleich arbeiten gehen"* und *"Ich würde etwas ganz anderes machen "* werden nachfolgend aufgrund der schwachen Besetzungszahlen bei der marginalen Analyse vernachlässigt. Zunächst wird untersucht, ob zwischen zweiter und dritter Periode, sowie zwischen dritter und vierter Periode Struktur vorliegt oder ob Quasi-Unabhängigkeit angenommen werden kann. Die entsprechenden Kreuztafeln sind in den Tabellen 37 und 38 gegeben. Ein Test auf Quasi-Unabhängigkeit wird jedoch mit  $p$ -Werten von 0.13 (zweite-dritte Periode) und 0.17 (dritte-vierte Periode) abgelehnt. Das bedeutet, daß unter den Wechslern keine Struktur vorliegt, d.h. entscheidet sich ein Jugendlicher in der dritten bzw. vierten Periode anders als in der Vorperiode, so ist sein Meinungswechsel nicht von seiner Meinung in der Vorperiode abhängig.

Zeitpunkt 1	Zeitpunkt 2						Gesamt
	Gleiche Ausbildung im selben Betrieb	Gleiche Ausbildung im anderen Betrieb	Anderen Beruf	Keine Ausbildung - gleich Arbeiten	Lieber weiter zur Schule	Etwas ganz anderes machen	
Gleiche Ausbildung im selben Betrieb	250 44.40	40 7.10	18 3.20	1 0.18	9 1.60	1 0.18	319 56.66
Gleiche Ausbildung im anderen Betrieb	48 8.53	27 4.80	10 1.78	1 0.18	5 0.89	0 0.00	91 16.16
Anderen Beruf	48 8.53	7 1.24	22 3.91	2 0.36	3 0.53	2 0.36	84 14.92
Keine Ausbildung - gleich Arbeiten	6 1.07	1 0.18	1 0.18	0 0.00	0 0.00	0 0.00	8 1.42
Lieber weiter zur Schule	24 4.26	3 0.53	11 1.95	0 0.00	10 1.78	1 0.18	49 8.70
Etwas ganz anderes machen	8 1.42	1 0.18	3 0.53	0 0.00	0 0.00	0 0.00	12 2.13
<b>Gesamt</b>	<b>384</b> <b>68.21</b>	<b>79</b> <b>14.03</b>	<b>65</b> <b>11.55</b>	<b>4</b> <b>0.71</b>	<b>27</b> <b>4.80</b>	<b>4</b> <b>0.71</b>	<b>563</b> <b>100.00</b>

Tabelle 37: "Noch einmal Berufswahl - Welche?", Zeitpunkte 1 und 2. Absolute Häufigkeiten und Prozent

Zeitpunkt 2	Zeitpunkt 3						Gesamt
	Gleiche Ausbildung im selben Betrieb	Gleiche Ausbildung im anderen Betrieb	Anderen Beruf	Keine Ausbildung - gleich Arbeiten	Lieber weiter zur Schule	Etwas ganz anderes machen	
Gleiche Ausbildung im selben Betrieb	250 44.40	40 7.10	18 3.20	1 0.18	9 1.60	1 0.18	319 56.66
Gleiche Ausbildung im anderen Betrieb	48 8.53	27 4.80	10 1.78	1 0.18	5 0.89	0 0.00	91 16.16
Anderen Beruf	48 8.53	7 1.24	22 3.91	2 0.36	3 0.53	2 0.36	84 14.92
Keine Ausbildung - gleich Arbeiten	6 1.07	1 0.18	1 0.18	0 0.00	0 0.00	0 0.00	8 1.42
Lieber weiter zur Schule	24 4.26	3 0.53	11 1.95	0 0.00	10 1.78	1 0.18	49 8.70
Etwas ganz anderes machen	8 1.42	1 0.18	3 0.53	0 0.00	0 0.00	0 0.00	12 2.13
<b>Gesamt</b>	<b>384</b> <b>68.21</b>	<b>79</b> <b>14.03</b>	<b>65</b> <b>11.55</b>	<b>4</b> <b>0.71</b>	<b>27</b> <b>4.80</b>	<b>4</b> <b>0.71</b>	<b>563</b> <b>100.00</b>

Tabelle 38: "Noch einmal Berufswahl - Welche?", Zeitpunkte 2 und 3. Absolute Häufigkeiten und Prozent

Wir wollen nun untersuchen, inwieweit sich das Meinungsbild der Jugendlichen im Durchschnitt über die Zeit verändert. Wie in den vorherigen Kapiteln betrachten wir ein marginales Modell. Dabei zeigt sich, daß der Zeiteffekt klar signifikant ist (p-Wert < 0.001), ein Geschlechtsspezifischer Unterschied ist jedoch nicht nachweisbar (p-Wert 0,26). Tabelle 39 gibt die beobachteten und modellierten relativen Häufigkeiten wieder, die in Abbildung 28 weiter verdeutlicht sind. Tabelle 40 zeigt die Ergebnisse der Signifikanztests. Man sieht, daß zwischen zweiter und dritter Periode kaum Veränderung vorliegt, nach Abschluß der Ausbildung ist jedoch zwischen dritter und vierter Periode eine deutliche Veränderung der Zufriedenheit der Jugendlichen mit der gewählten Ausbildung zu beobachten. Dabei sind nur noch halb so viele Jugendliche mit der gewählten Ausbildung und dem gewählten Betrieb zufrieden wie in der dritten Periode zufrieden waren. Entsprechend steigt der Anteil derer, die einen anderen Beruf ergreifen würden, auf über 36%. Die anderen beiden möglichen Kategorien weisen weitaus weniger zeitliche Veränderung auf.

	Geschlecht	Kategorie	zeitpunkt		
			1	2	3
beobachtet	männlich	1	0,631	0,688	0,340
		2	0,163	0,128	0,163
		3	0,092	0,121	0,333
		4	0,113	0,064	0,163
	weiblich	1	0,580	0,652	0,295
		2	0,152	0,161	0,188
		3	0,161	0,152	0,402
		4	0,107	0,036	0,116
modelliert	männlich	1	0,616	0,679	0,320
		2	0,159	0,141	0,178
		3	0,116	0,132	0,364
		4	0,109	0,048	0,138
	weiblich	1	0,616	0,679	0,320
		2	0,159	0,141	0,178
		3	0,116	0,132	0,364
		4	0,109	0,048	0,138

Tabelle 39: Randhäufigkeiten für Zufriedenheit mit Berufswahl

$H_0$ versus $H_1$			
$H_0$	$H_1$	p-Wert	Testergebnis
Zeit © Betrieb	Zeit	0.618	Zeit
Zeit © Geschlecht	Zeit	0.264	Zeit
Zeit	0	< 0.001	Zeit

Tabelle 40: Empirische Signifikanzen für Zufriedenheit mit Berufswahl

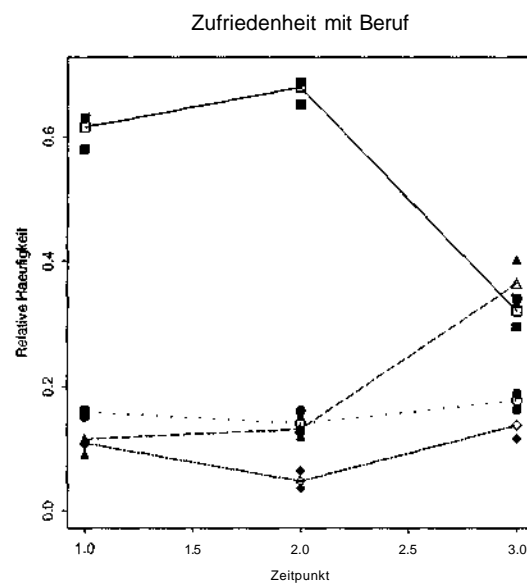


Abbildung 28: Beobachtete und modellierte relative Häufigkeiten für Beurteilung der Zufriedenheit mit der Berufswahl (— gleiche Ausbildung, gleicher Betrieb, ••• gleiche Ausbildung, anderer Betrieb, — anderer Beruf, — Schule)

## 7.5 Leben ohne Arbeit?

Abschließend wurden die Jugendlichen in der ersten, zweiten und dritten Periode gefragt, ob sie sich ein Leben ohne Arbeit vorstellen könnten (Tabellen 41 und 42). Es zeigt sich, daß zwischen erster und zweiter Periode Quasi-Unabhängigkeit vorliegt (p-Wert 0.548). Der entsprechende Test ist in diesem Fall der McNemar Test. Dies bedeutet, daß annähernd so viele Jugendliche zu beobachten sind, die sich am Anfang der Lehre ein Leben ohne Arbeit vorstellen konnten, dies jedoch in der zweiten Periode nicht mehr können, wie Jugendliche, die sich nach einem Jahr Lehre ein Leben ohne Arbeit vorstellen können, obgleich sie dies zu Beginn der Lehre nicht konnten. Zwischen der zweiten und dritten Periode ändert sich das Bild jedoch (p-Wert < 0.001), weil weitaus mehr Jugendliche, die sich in der zweiten Periode ein Leben ohne Arbeit nicht vorstellen konnten, dies in der dritten Welle durchaus können als umgekehrt.

		Zeitpunkt 2		
Zeitpunkt 1	Ja	Nein	Gesamt	
Ja	153 14.78	173 16.71	326 31.50	
Nein	162 15.65	547 52.85	709 68.50	
Gesamt	315 30.43	720 69.57	1035 100.00	

Tabelle41: "Bei ausreichend Geld: Leben ohne Arbeit", Zeitpunkte 1 und 2

		Zeitpunkt 3		
Zeitpunkt 2	Ja	Nein	Gesamt	
Ja	122 17.73	90 13.08	212 30.81	
Nein	161 23.40	315 45.78	476 69.19	
Gesamt	283 41.13	405 58.87	688 100.00	

Tabelle 42: "Bei ausreichend Geld: Leben ohne Arbeit", Zeitpunkte 2 und 3



Betrachten wir nun, welche Faktoren die Bereitschaft zu einem Leben ohne Arbeit beeinflussen, so zeigt sich, daß sowohl ein Zeit- als auch ein Geschlechtseffekt vorliegt. Aus Tabelle 43 sowie auch aus Abbildung 29 erkennt man, daß die Bereitschaft zu einem Leben ohne Arbeit über die Zeit zunimmt, wobei hier insbesondere nach der zweiten Periode eine deutliche Veränderung eintritt. Ferner können sich weitaus mehr Männer ein Leben ohne Arbeit vorstellen als dies Frauen können.

	Geschlecht	Kategorie	Zeitpunkt		
			1	2	3
beobachtet	männlich	ja	0,383	0,341	0,445
		nein	0,617	0,659	0,555
	weiblich	ja	0,210	0,263	0,367
		nein	0,790	0,737	0,633
modelliert	männlich	ja	0,355	0,356	0,459
		nein	0,645	0,644	0,541
	weiblich	ja	0,241	0,242	0,345
		nein	0,759	0,758	0,655

Tabelle 43: Randhäufigkeit für die Vorstellung nach einem Leben ohne Arbeit

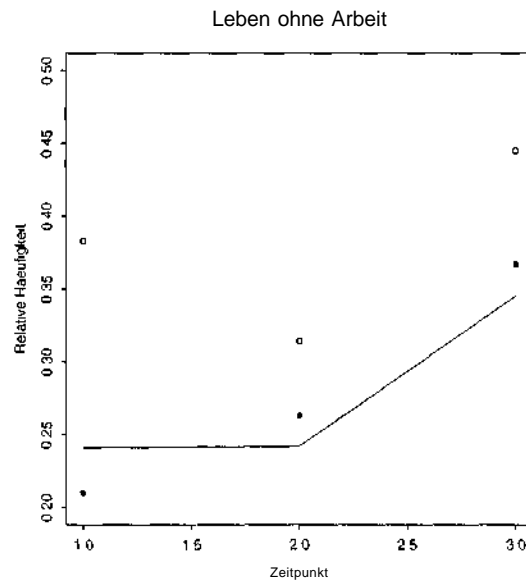


Abbildung 29: Relative Häufigkeiten der Jugendlichen, die sich ein Leben ohne Arbeit vorstellen können (— Frauen, ••• Männer)

## 7.6 Selbstzufriedenheit

Um abschließend einen Eindruck der generellen Zufriedenheit der Jugendlichen zu erhalten, wurden die Jugendlichen gefragt, wie sie sich zur Zeit fühlen. Dabei konnten sie sich sieben möglichen Stimmungslagen zuordnen, die wir zur einfacheren Auswertung zu drei Gruppen zusammengefaßt haben. Wir beurteilen die mögliche Stimmung eines Jugendlichen als *negativ* (Kategorie 1-3), *mittel* (Kategorie 4-5) und *positiv* (Kategorie 6-7). Wir betrachten zunächst, ob ein Wandel in der Zufriedenheit von der bisherigen Zufriedenheit abhängig ist. Es zeigt sich jedoch, daß in den einzelnen Tabellen (Tabelle 44 und 45) Quasi-Unabhängigkeit vorliegt (p-Wert 0.16 für erste-zweite Periode, p-Wert 0.20 für zweite-dritte Periode).

	Zeitpunkt 2			
Zeitpunkt 1	Negativ	Mittel	Positiv	Gesamt
Negativ	20 1.94	40 3.88	44 4.27	104 10.10
Mittle	48 4.66	272 26.41	197 19.13	517 50.19
Positiv	32 3.11	173 16.80	204 19.81	409 39.71
Gesamt	100 9.71	485 47.09	445 43.20	1030 100.00

Tabelle 44: "Wie fühlen Sie sich zur Zeit?"

	Zeitpunkt 3			
Zeitpunkt 2	Negativ	Mittel	Positiv	Gesamt
Negativ	4 1.23	11 3.38	17 5.23	32 9.85
Mittle	17 5.23	81 24.92	48 14.77	146 44.92
Positiv	14 4.31	64 19.69	69 21.23	147 45.23
Gesamt	35 10.77	156 48.00	134 41.23	325 100.00

Tabelle 45: "Wie fühlen Sie sich zur Zeit?"

Ferner soll untersucht werden, ob sich die allgemeine Zufriedenheit der Jugendlichen über die Zeit ändert. Es lässt sich jedoch weder eine zeitliche Veränderung noch ein geschlechtsspezifischer Unterschied nachweisen (siehe Tabelle 46), was auch aus Tabelle 47 deutlich wird. Somit ergibt sich ein zeitkonstantes und geschlechts-unabhängiges Abbild der Zufriedenheit der Jugendlichen, wie es mit Abbildung 30 verdeutlicht wird.

$H_o$ versus $H_1$			
$H_o$	# i	p-Wert	Testergebnis
Zeit & Geschlecht	Zeit	0.654	Zeit
Zeit	0	0.505	0

Tabelle 46: Empirische Signikanzen für allgemeine Zufriedenheit

	Geschlecht	Kategorie	zeitpunkt		
			1	2	3
beobachtet	männlich	1	0,088	0,088	0,088
		2	0,538	0,450	0,485
		3	0,374	0,462	0,427
	weiblich	1	0,098	0,111	0,131
		2	0,458	0,451	0,471
		3	0,444	0,438	0,399
modelliert	männlich	1	0.100		
	weiblich	2	0.474		
		3	0.426		

Tabelle 47: Randhäufigkeit für allgemeine Zufriedenheit

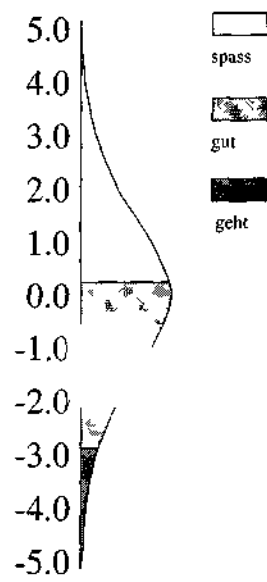


Abbildung 30: Allgemeine Zufriedenheit zu allen drei Zeitpunkten

## 8 Analyse der politischen Einstellung

### Parteipräferenzen

Die Jugendlichen wurden in der ersten, zweiten und dritten Periode gefragt welcher Partei gegenüber sie die meiste Sympathie aufbringen. Dabei wurden die Parteien zu "CDU/CSU", "SPD", "FDP", "Grüne", "PDS" und "rechts" d.h. "NPD/DVU" und "Republikaner" zusammengefaßt. Als weitere Kategorie stand "keine Präferenz" zur Verfügung (siehe Tabelle). Eine Untersuchung auf Quasi-Unabhängigkeit lieferte keine signifikanten Ergebnisse (p-Wert 0.29 erste-zweite Periode, p-Wert 0.567 zweite-dritte Periode). Das bedeutet inhaltlich, daß unter den Jugendlichen, die ihre Parteipräferenz über die Zeit wechseln, keine Struktur vorliegt, also keine Abhängigkeit zur vorherigen Parteipräferenz nachweisbar ist. Es zeigt sich jedoch sowohl ein Zeiteffekt als auch ein Geschlechtseffekt für die durchschnittlichen Parteipräferenzen (p-Wert < 0.001). Auf die Darstellung und Interpretation eines Zeit-Geschlechts-Interaktionseffekts wurde aus Gründen der Übersichtlichkeit verzichtet (p-Wert 0.026). Die entsprechenden Parteipräferenzen sind in Abbildung 31 gegeben, die entsprechenden Häufigkeiten finden sich in Tabelle 50.

Zeitpunkt 1	Zeitpunkt 2							Gesamt
	CDU/ CSU	SPD	FDP	Grüne/ Bündnis 90	PDS	Rechts	Sonstige	
CDU/CSU	17 1.64	6 0.58	5 0.48	10 0.97	8 0.77	2 0.19	43 4.15	91 8.79
SPD	5 0.48	38 3.67	2 0.19	5 0.48	8 0.77	1 0.10	43 4.15	102 9.86
FDP	3 0.29	4 0.39	6 0.58	6 0.58	2 0.19	0 0.00	15 1.45	36 3.48
Grüne/ Bündnis 90	6 0.58	5 0.48	1 0.10	29 2.80	6 0.58	0 0.00	37 3.57	84 8.12
PDS	0 0.00	2 0.19	1 0.10	5 0.48	12 1.16	2 0.19	5 0.48	27 2.61
Rechts	5 0.48	5 0.48	1 0.10	3 0.29	5 0.48	20 1.93	31 3.00	70 6.76
Sonstige	42 4.06	73 7.05	18 1.74	55 5.31	31 3.00	25 2.42	381 36.81	625 60.39
Gesamt	78 7.54	133 12.85	34 3.29	113 10.92	72 6.96	50 4.83	555 53.62	1035 100.00

Tabelle 48: "Welche politische Partei gefällt Ihnen zur Zeit am besten?", Zeitpunkt 1, 2

Zeitpunkt 2	Zeitpunkt 3							Gesamt
	CDU/ CSU	SPD	FDP	Grüne/ Bündnis 90	PDS	Rechts	Sonstige	
CDU/CSU	23 3.34	10 1.45	0 0.00	3 0.44	1 0.15	1 0.15	13 1.89	51 7.41
SPD	16 2.33	36 5.23	1 0.15	6 0.87	12 1.74	1 0.15	24 3.49	96 13.95
FDP	7 1.02	5 0.73	2 0.29	1 0.15	2 0.29	1 0.15	5 0.73	23 3.34
Grüne/ Bündnis 90	10 1.45	13 1.89	2 0.29	20 2.91	8 1.16	1 0.15	24 3.49	78 11.34
PDS	5 0.73	9 1.31	1 0.15	4 0.58	15 2.18	2 0.29	6 0.87	42 6.10
Rechts	6 0.87	7 1.02	1 0.15	1 0.15	1 0.15	6 0.87	7 1.02	29 4.22
Sonstige	94 13.66	70 10.17	2 0.29	22 3.20	40 5.81	11 1.60	130 18.90	369 53.63
Gesamt	161 23.40	150 21.80	9 1.31	57 8.28	79 11.48	23 3.34	209 30.38	688 100.00

Tabelle 49: "Welche politische Partei gefällt Ihnen zur Zeit am besten?", Zeitpunkte 2, 3

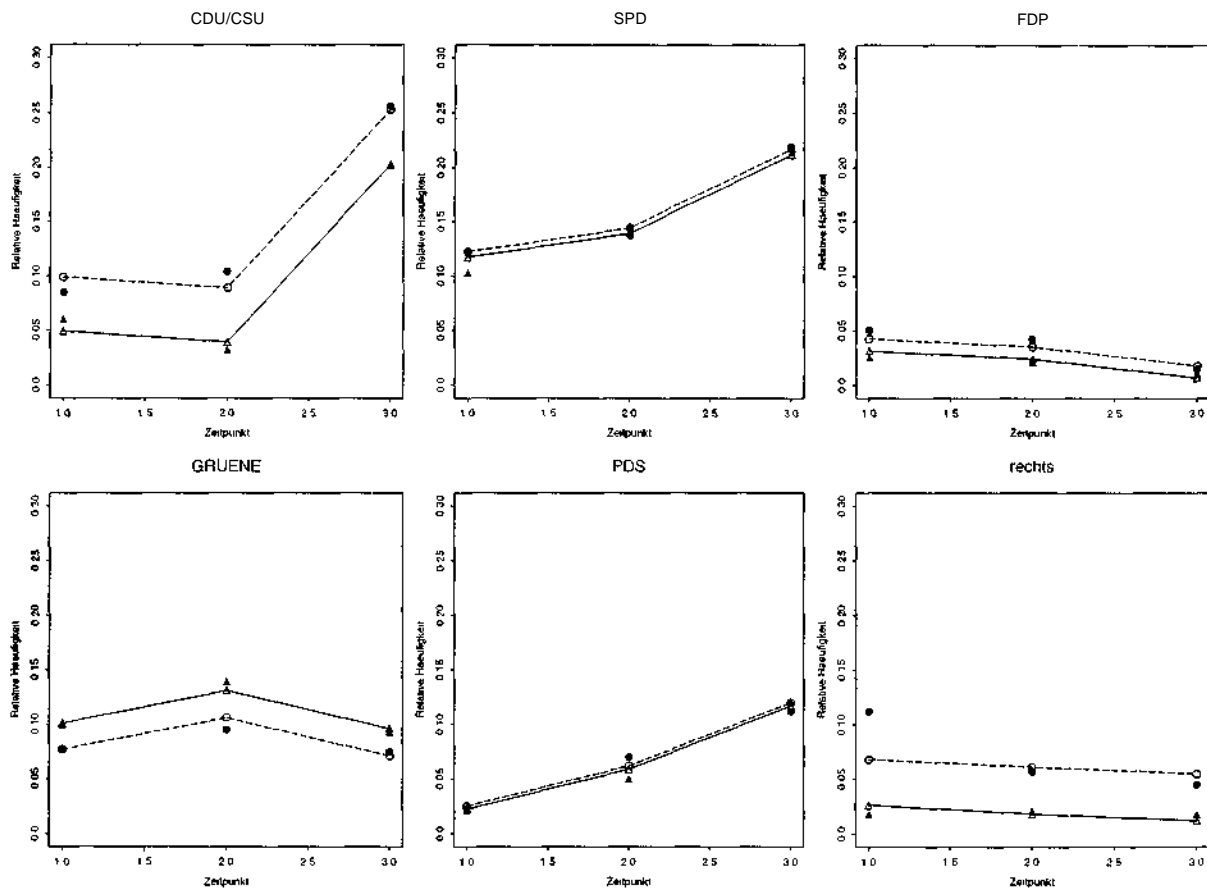


Abbildung 31: Relative Häufigkeit der Jugendlichen die der entsprechenden Partei gegenüber Sympathie bekunden. (— Frauen, ——— Männer, ○ = modelliert, • = beobachtet)

	Geschlecht	Kategorie	Zeitpunkt		
			1	2	3
beobachtet	männlich	CDU	0,085	0,104	0,256
		SPD	0,122	0,137	0,219
		FDP	0,050	0,042	0,015
		Grüne	0,077	0,095	0,075
		PDS	0,022	0,070	0,112
		rechts	0,112	0,057	0,045
		sonst	0,532	0,495	0,279
	weiblich	CDU	0,060	0,032	0,203
		SPD	0,103	0,146	0,214
		FDP	0,025	0,021	0,011
		Grüne	0,100	0,139	0,093
		PDS	0,025	0,050	0,121
		rechts	0,018	0,021	0,018
		sonst	0,669	0,591	0,342
modelliert	männlich	CDU	0,099	0,089	0,253
		SPD	0,122	0,144	0,217
		FDP	0,042	0,035	0,018
		Grüne	0,077	0,106	0,071
		PDS	0,025	0,062	0,120
		rechts	0,068	0,061	0,055
		sonst	0,567	0,503	0,265
	weiblich	CDU	0,049	0,039	0,203
		SPD	0,117	0,139	0,212
		FDP	0,031	0,024	0,007
		Grüne	0,101	0,131	0,096
		PDS	0,022	0,059	0,117
		rechts	0,026	0,018	0,012
		sonst	0,655	0,591	0,353

Tabelle 50: Randhäufigkeiten für Parteipräferenzen



Der Zeiteffekt verdeutlicht sich zwischen der zweiten und dritten Periode wo eine klare Zunahme der Präferenz zu den beiden großen Parteien zu beobachten ist. Ebenso unterliegt der Zuspruch zur PDS einer zeitlichen Veränderung, wobei hier eine eher gleichmäßige Zunahme an Zuspruch über die Zeit hinweg erfolgt. Auffällig ist ferner, daß der Zuspruch zu rechten Parteien über die Zeit eher konstant ist.

Der Geschlechtseffekt wird bei der CDU als auch bei den rechten Parteien deutlich, wobei Männer in beiden Fällen eine erhöhte Präferenz aufweisen. Beim Zuspruch zur PDS gibt es dagegen kaum geschlechtsspezifische Unterschiede. Betrachtet man nun nur die Jugendlichen, die keine Präferenz zu den genannten Parteien aufweisen, so zeigt sich hier der Zeiteffekt deutlich durch eine Abnahme des Anteils der Jugendlichen, die keinerlei Präferenz aufweisen. Der Geschlechtseffekt ergibt hier, daß Frauen häufiger keine Parteisympathie aufweisen als dies Männer tun (siehe Abbildung 32).

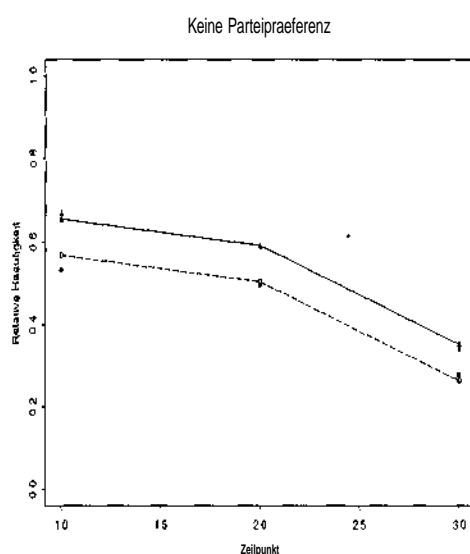


Abbildung 32: Relative Häufigkeit von Jugendlichen, die keine Parteipräferenz aufweisen. (— Frauen, ———Männer, o = modelliert, • = beobachtet)

## 9 Analyse der Wertvorstellungen

Wir betrachten im folgenden die Einstellung der Jugendlichen zu nationalen wie gesellschaftlichen Fragestellungen, um einen Eindruck über die Wertvorstellungen der Jugendlichen zu bekommen. Dabei wurden die Jugendlichen zu den ersten drei Zeitpunkten nach Zustimmung zu den Aussagen:

- *"Ich bin stolz ein Deutscher zu sein"*
- *"Ich möchte etwas tun, womit ich meinem Vaterland dienen kann"*
- *"Als Mensch und Staatsbürger fühle ich mich zuerst meinen Landsleuten hier in der ehemaligen DDR zugehörig"*
- *"Ich würde auch meine eigenen Interessen zurückstellen, wenn ich damit dem Aufbau und der Entwicklung des vereinten Deutschlands dienen könnte"*
- *"Für mich ist es jetzt am wichtigsten, daß ich privat klar komme - in meiner Ausbildung, meinem Beruf und meinen persönlichen Beziehungen. Von gesellschaftlichen Verpflichtungen habe ich vorläufig die Nase voll."*

gefragt. Die Daten wurden dabei individualspezifisch mit Hilfe von Markov-Ketten, als auch querschnittsmäßig in Form von marginalen Modellen untersucht. Ferner wurden Quasiunabhängigkeitsanalysen durchgeführt.

### 9.1 Querschnittsmodellierung mit Einfluß von Geschlecht

Wir wollen nun die Modellierung der Randverteilung verfolgen. Das heißt wir betrachten marginale Modelle wie sie in Kapitel 4 vorgestellt worden sind. Es soll insbesondere untersucht werden, ob zeitliche Veränderungen bei der Zustimmung auf obige Aussagen bestehen. Es ist weiter von Interesse zu beurteilen, ob geschlechtsspezifische Unterschiede bestehen und ob diese zeitlich variieren oder konstant sind, d.h. ob Geschlechtseffekt und Zeit interagieren.

- *"Ich bin stolz ein Deutscher zu sein"*  
In Tabelle 51 geben wir die Zustimmung auf die Fragestellung aufgeschlüsselt nach Geschlecht für die drei Zeitpunkte an. Es zeigt sich deutlich, daß ein geschlechtsspezifischer Effekt vorliegt. Wir modellieren die Daten so, daß die Zustimmung zur obigen Aussage sowohl vom Zeitpunkt der Befragung als auch vom Geschlecht abhängt, jedoch keine Interaktion zwischen Geschlecht und Zeit existiert. In der im Kapitel 1.2 vorgeschlagenen Modellnotation ergibt sich somit  $\text{Zeit} \otimes \text{Geschlecht}$ . Tests auf komplexere oder einfachere Modelle werden basierend auf einem Signifikanzniveau von  $\alpha = 0.05$  abgelehnt. Tabelle 52 zeigt die entsprechenden p-Werte. Die Einstellung zur Aussage *"Ich bin stolz ein Deutscher zu sein"* variiert somit signifikant über die Zeit und ist bei Männern und Frauen unterschiedlich. Der Geschlechtseffekt ist jedoch konstant über die Zeit. Die entsprechend modellierten Häufigkeiten sind ebenfalls in Tabelle 51 gegeben. Abbildung 33

verdeutlicht die Daten. Es zeigt sich, daß Männer häufiger der Aussage zustimmen als Frauen. Dabei zeigt sich in der zweiten Welle eine deutliche Zurückhaltung bei der Zustimmung zu der Aussage im Vergleich zur ersten und dritten Periode.

	Geschlecht	Zeitpunkt		
		1	2	3
beobachtet	männlich	63,4	55,7	72,3
	weiblich	57,7	47,0	55,9
modelliert	männlich	65,4	56,6	70,3
	weiblich	54,7	46,0	59,6

Tabelle 51: Anteil der Jugendlichen, die der Aussage "Ich bin stolz ein Deutscher zu sein" zustimmen

$H_0$ versus $H_1$			
$H_0$	$H_1$	p-Wert	Testergebnis
Zeit © Geschlecht	Zeit © Geschlecht	0.056	Zeit © Geschlecht
Zeit	Zeit © Geschlecht	< 0.001	Zeit © Geschlecht
Geschlecht	Zeit © Geschlecht	< 0.001	Zeit © Geschlecht

Tabelle 52: Testergebnisse zur Aussage: "Ich bin stolz ein Deutscher zu sein "

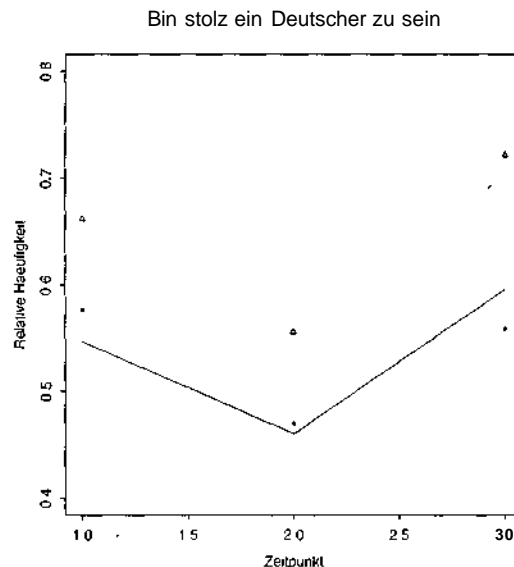


Abbildung 33: Anteil der Jugendlichen, die der Aussage "Ich bin stolz ein Deutscher zu sein" zustimmen in Abhängigkeit vom Geschlecht (Männer—und Frauen ---)

- "Ich möchte etwas tun, womit ich meinem Vaterland dienen kann"

In analoger Weise betrachten wir nun die Zustimmung zu obiger Aussage. Die zugehörigen Daten sind in Tabelle 53 gegeben. Es zeigt sich, daß Zeit und Geschlecht zu dieser Fragestellung interagieren, was in Tabelle 54 mit statistischen Tests verdeutlicht ist. Das bedeutet, daß sich der Geschlechtseffekt mit der Zeit verändert, in diesem Falle verstärkt, wie aus Abbildung 34 zu sehen ist. Wieder neigen Männer eher zur Zustimmung als Frauen. Über die Zeit hinweg nimmt die Bereitschaft, "etwas für sein Vaterland tun" zu wollen jedoch generell ab.

	Geschlecht	Zeitpunkt		
		1	2	3
beobachtet	männlich	66,6	59,0	56,5
und modelliert	weiblich	55,3	40,3	31,9

Tabelle 53: Anteil der Jugendlichen, die der Aussage "Ich möchte etwas tun, womit ich meinem Vaterland dienen kann" zustimmen

$H_0$ versus $H_1$		p-Wert	Testergebnis
$H_0$	# i		
Zeit $\times$ Geschlecht	Zeit $\times$ Geschlecht	0.023	Zeit $\otimes$ Geschlecht

Tabelle 54: Testergebnisse zur Aussage: "Ich möchte etwas tun, womit ich meinem Vaterland dienen kann"

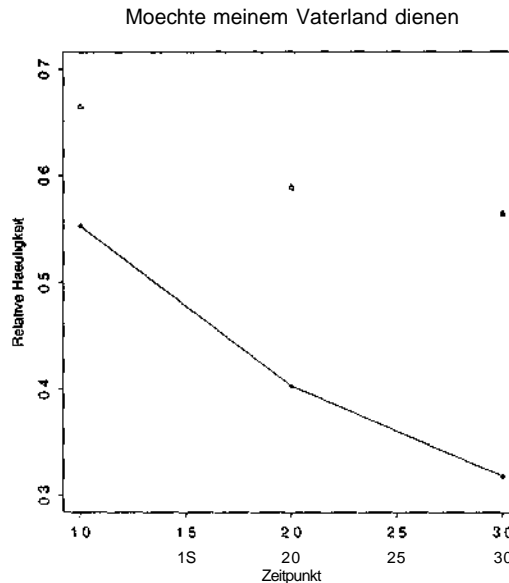


Abbildung 34: Anteil der Jugendlichen, die der Aussage "Ich möchte etwas tun um meinem Vaterland zu dienen" zustimmen, in Abhängigkeit vom Geschlecht (Männer—und Frauen —)

- *"Als Mensch und Staatsbürger fühle ich mich zuerst meinen Landsleuten hier in der ehemaligen DDR zugehörig"*

Für diese Aussage zeigt sich, daß geschlechtsspezifische Unterschiede nicht signifikant sind, Frauen und Männer somit eine gleiche Einstellung haben. Deutlich ist jedoch eine zeitliche Variation nachweisbar. Die entsprechenden statistischen Tests sind in Tabelle 55 enthalten, die zugehörigen Daten sind in Tabelle 56 gegeben. Modell und Daten sind in Abbildung 35 dargestellt. Hier zeigt sich, daß die Jugendlichen sich in der zweiten Periode am deutlichsten *"den Landsleuten der DDR zugehörig fühlen"*.

<i>H<sub>0</sub> versus H<sub>1</sub></i>			
<i>H<sub>0</sub></i>	# i	p-Wert	Testergebnis
Zeit	Zeit © Geschlecht	0.246	Zeit
1	Zeit	0.001	Zeit

Tabelle 55: Testergebnisse zur Aussage: *"Als Mensch und Staatsbürger fühle ich mich zuerst meinen Landsleuten hier in der ehemaligen DDR zugehörig"*

	Geschlecht	Zeitpunkt		
		1	2	3
beobachtet	männlich	75,2	77,4	70,5
	weiblich	76,1	81,8	72,5
modelliert	männlich	75,6	79,4	71,5
	weiblich	75,6	79,4	71,5

Tabelle 56: Anteil der Jugendlichen, die der Aussage *"Als Mensch und Staatsbürger fühle ich mich zuerst meinen Landsleuten hier in der ehemaligen DDR zugehörig"* zustimmen

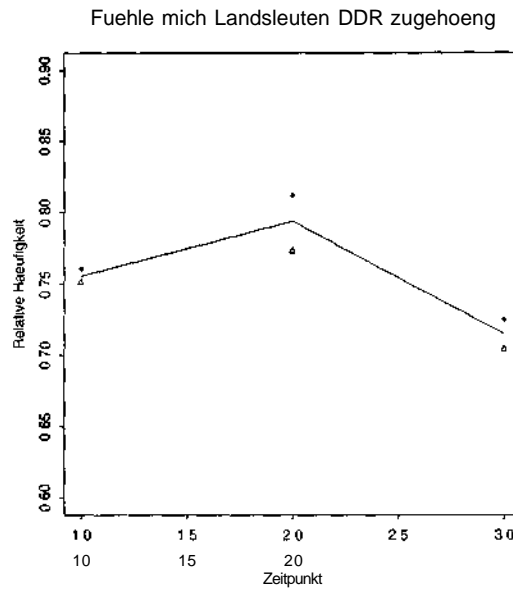


Abbildung 35: Anteil der Jugendlichen, die der Aussage "Als Mensch und Staatsbürger fühle ich mich zuerst meinen Landsleuten hier in der ehemaligen DDR zugehörig" zustimmen

- "Ich würde auch meine eigenen Interessen zurückstellen, wenn ich damit dem Aufbau und der Entwicklung des vereinten Deutschlands dienen könnte"

Die Einstellung zu obiger Aussage unterscheidet sich signifikant bei Frauen und Männern, dieser Unterschied ist jedoch zeitkonstant. Die Einstellung variiert jedoch mit der Zeit, wobei, wie in Abbildung 36 zu sehen ist, deutlich weniger Jugendliche im Laufe der Zeit der Aussage zusprechen. Die zugehörigen Daten sind in Tabelle 57 aufgelistet, die Ergebnisse der Modelltests finden sich in Tabelle 58.

	Geschlecht	Zeitpunkt		
		1	2	3
beobachtet	männlich	46,3	39,8	30,5
	weiblich	41,3	32,8	25,5
modelliert	männlich	46,6	39,3	30,7
	weiblich	40,9	33,6	25,1

Tabelle 57: Anteil der Jugendlichen, die der Aussage "Ich würde auch meine eigenen Interessen zurückstellen, wenn ich damit dem Aufbau und der Entwicklung des vereinten Deutschlands dienen könnte" zustimmen

$H_0$ versus $H_1$		p-Wert	Testergebnis
$H_0$	$H_1$		
Zeit © Geschlecht	Zeit © Geschlecht	0.897	Zeit © Geschlecht
Zeit	Zeit © Geschlecht	< 0.001	Zeit © Geschlecht
Geschlecht	Zeit © Geschlecht	< 0.021	Zeit © Geschlecht

Tabelle 58: Testergebnisse zur Aussage: "Ich würde auch meine eigenen Interessen zurückstellen, wenn ich damit dem Aufbau und der Entwicklung des vereinten Deutschlands dienen könnte "

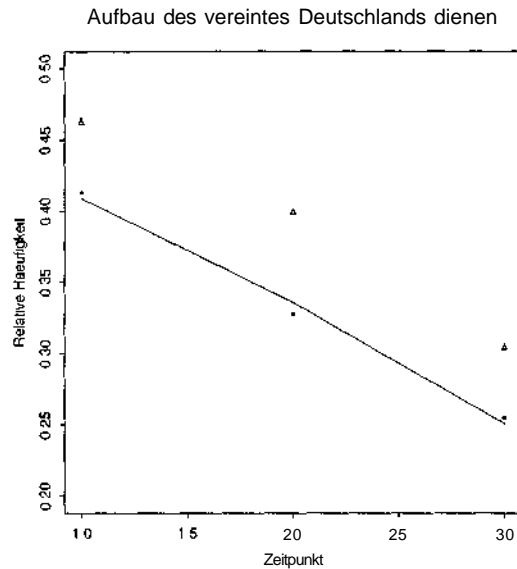


Abbildung 36: Anteil der Jugendlichen, die der Aussage "Ich würde auch meine eigenen Interessen zurückstellen, wenn ich damit dem Aufbau und der Entwicklung des vereinten Deutschlands dienen könnte " zustimmen, Abhängigkeit vom Geschlecht (Männer——und Frauen —)

- *"Für mich ist es jetzt am wichtigsten, daß ich privat klar komme - in meiner Ausbildung, meinem Beruf und meinen persönlichen Beziehungen. Von gesellschaftlichen Verpflichtungen habe ich vorläufig die Nase voll."*

Die Einstellung zum gesellschaftlichen Engagement unterscheidet sich signifikant über die Zeit, ein geschlechtsspezifischer Unterschied ist jedoch nicht nachweisbar, wie aus Tabelle 59 ersichtlich. Die Daten sind in Tabelle 60 gegeben, Abbildung 37 veranschaulicht das Ergebnis. Insbesondere in der zweiten Periode ist das private Auskommen weniger wichtig als in Periode eins und drei.

<i>Ho</i> versus <i>H</i>			
<i>Ho</i>	<i>Hi</i>	p-Wert	Testergebnis
Zeit © Geschlecht	Zeit	0.203	Zeit
1	Zeit	< 0.002	Zeit

Tabelle 59: Testergebnisse zur Aussage: *"Für mich ist es jetzt am wichtigsten, daß ich privat klar komme - in meiner Ausbildung, meinem Beruf und meinen persönlichen Beziehungen. Von gesellschaftlichen Verpflichtungen habe ich vorläufig die Nase voll"*

	Geschlecht	Zeitpunkt		
		1	2	3
beobachtet	männlich	77,4	71,1	78,9
	weiblich	78,0	74,7	82,3
modelliert	männlich	77,6	72,7	80,4
	weiblich	77,6	72,7	80,4

Tabelle 60: Anteil der Jugendlichen, die der Aussage *"Für mich ist es jetzt am wichtigsten, daß ich privat klar komme - in meiner Ausbildung, meinem Beruf und meinen persönlichen Beziehungen. Von gesellschaftlichen Verpflichtungen habe ich vorläufig die Nase voll"* zustimmen



Keine gesellschaftlichen Verpflichtungen

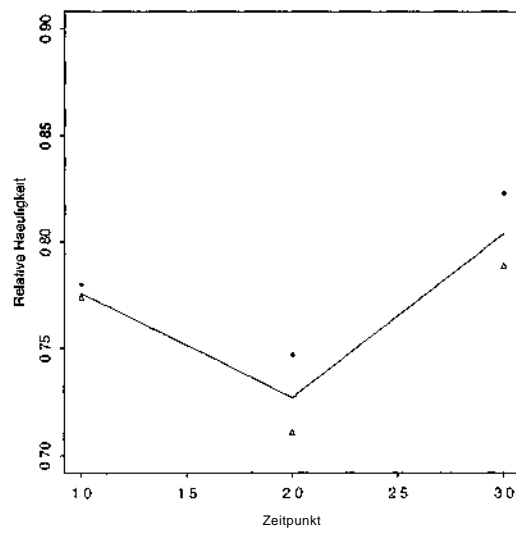


Abbildung 37: Anteil der Jugendlichen, die der Aussage "Für mich ist es jetzt am wichtigsten, daß ich privat klar komme - in meiner Ausbildung, meinem Beruf und meinen persönlichen Beziehungen. Von gesellschaftlichen Verpflichtungen habe ich vorläufig die Nase voll" zustimmen

### 9.1.1 Querschnittsmodellierung mit Einstellung zu Türken

Im folgenden soll der Einfluß einer weiteren Variablen berücksichtigt werden, die die Einstellung zu türkischen Mitbürgern in Form eines Score mißt. Dieser Score wurde gebildet aus den Antworten auf die Frage ob man einen Türken *"als Wohnungsnachbar"*, *"im Freundeskreis"*, *"als Kollegen"*, *"als Vorgesetzten"*, *"in der Familie"* oder *"im Gemeinderat"* eher positiv oder negativ gegenübersteht. Die möglichen Antworten waren von 1 für *"würde ich begrüßen"* bis 5 *"würde ich ablehnen"*. Die entsprechenden Ausprägungen wurden summiert und ergaben einen Score, große Werte des Scores zeugen also von einer Ablehnung, kleine Werte von einer Akzeptanz türkischer Mitbürger. Der Score selber kann als normalverteilt angenommen werden.

Die ursprünglich erhobene Fragestellung, ob man sich einen Türken *"als Freund"* vorstellen kann, wurde zur Bildung des Scores nicht berücksichtigt. Grund hierfür war, daß Frauen und Männer auf diese Frage völlig unterschiedliche Antworten gaben, was wohl an einer unterschiedlichen Interpretation der Frage gelegen haben kann. Um dies zu verdeutlichen sind in Abbildung 38 die normierten ersten beiden Hauptkomponenten einer Hauptkomponentenanalyse dargestellt (vgl. Fahrmeir, Hamerle & Tutz, 1996). Man erkennt, daß alle Fragen bis auf *"Türke als Freund"* in den beiden Geschlechtergruppen die gleiche Stellung zu den ersten beiden Hauptkomponenten haben, die Antwort auf *"Türke als Freund"* jedoch gänzlich unterschiedlich bemessen wird.

Kehren wir nun zurück zur Fragestellung inwieweit die Einstellung zu Türken Einfluß auf die Einstellung eingangs gestellter Aussagen hat.

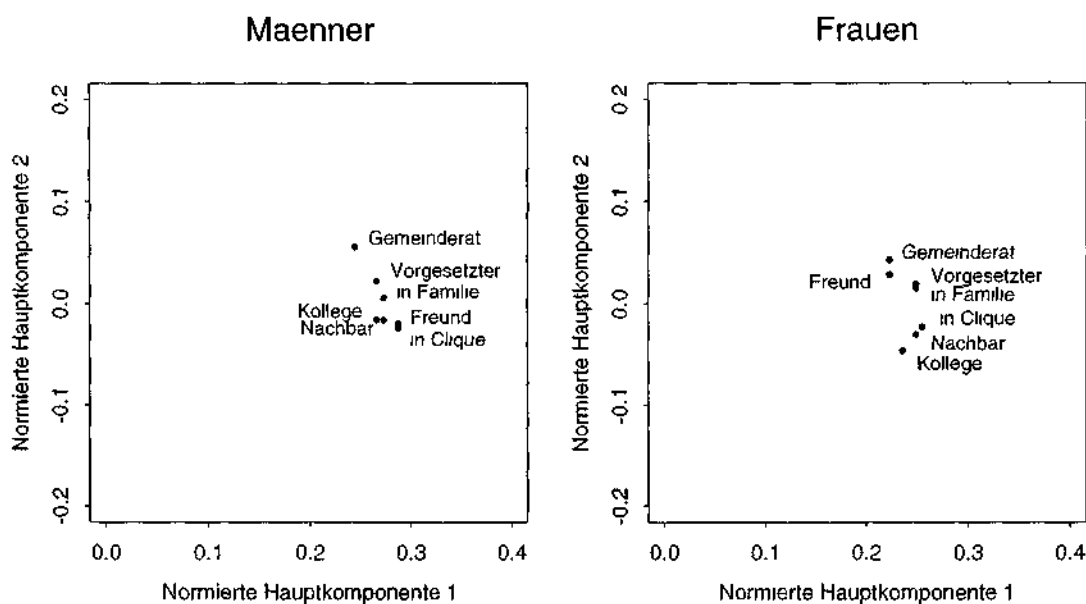


Abbildung 38: Hauptkomponentenanalyse zur Bildung eines Scores zur Beschreibung der Einstellung zu Türken

- "Ich bin stolz ein Deutscher zu sein"

Wir modellieren die Daten gemäß dem marginalen Modell (12), wobei der Score als linearer Einfluß modelliert wird. Ferner erlauben wir, daß Geschlecht und Score interagieren. Um die Modellkomplexität nicht zu überfrachten, wird diese Interaktion jedoch als zeitkonstant angesehen. Das heißt, wir betrachten das Modell

$$h(E(Y_{it}|G, S, t)) = \beta t + G\beta_{Gt} + S\beta_{St} + G \cdot S\beta_{GS}$$

wobei  $G$  für Geschlecht,  $S$  für Score und  $t$  für Zeit steht. In Modellschreibweise erhält man

$$(Zeit \odot Geschlecht) \odot (Zeit \odot Score) \odot (Geschlecht \odot Score)$$

Derartige Modelle lassen sich wegen des stetigen Scores nicht mehr mit SAS® berechnen. Wir greifen daher auf die von Glonek & McCullagh (1995) bereitgestellte Software zurück. Es zeigt sich, daß alle Größen im Modell signifikanten Einfluß haben. Das wird insbesondere auch aus Tabelle 61 ersichtlich.<sup>3</sup> Die zugehörigen standardisierten Schätzer der Interaktionen (Schätzer geteilt durch Standardabweichung) sind von signifikanter Größe, d.h. betragsmäßig größer als 1.96.

Parameter	Modellkomponente	Schätzung	Standardabweichung
$\beta_1$	} t	-4,19	0,52
$\beta_2$		-1,52	0,53
$\beta_3$		1,39	0,19
$\beta_{G1}$	G	1,07	0,55
$\beta_{S1}$	S	0,26	0,03
$\beta_{GS}$	G⊙S	-0,06	0,03
A?2 — $\beta_d$	} G ⊙ t	-0,15	0,21
\$33 — $\beta_{ü1}$		-0,58	0,22
$\beta_{S2}$ — $\beta_{S\setminus}$	} S ⊙ t	-0,11	0,03
$\beta_{S3}$ — $\beta_{Si}$		-0,16	0,03

Tabelle 61: Parameterschätzungen zur Aussage "Ich bin stolz ein Deutscher zu sein"

<sup>3</sup>Die Modelle werden mit einer sogenannten Dummykodierung geschätzt, d.h. z.B.  $\beta_{1i}$  dient als sogenannte baseline während  $\beta_{0-i}$  als Differenz zwischen ersten und zweiten Zeitpunkt bestimmt worden ist.

Zur besseren Übersichtlichkeit stellen wir die modellierten Daten in Abbildung 39 dar. Es zeigt sich, daß je höher der Score, d.h. je distanzierter die Jugendlichen türkischen Mitbürgern gegenüberstehen, desto höher die Zustimmung zur Aussage "Ich bin stolz ein Deutscher zu sein". Dieser Score Effekt variiert jedoch über die Zeit und insbesondere in der dritten Periode zeigt sich ein recht homogenes Bild unter den Jugendlichen, d.h. die Zustimmung ist weitaus weniger abhängig vom Score (Interaktion Score © Zeit). In den ersten zwei Perioden neigen die männlichen Jugendlichen mit hohem Score eher zu einer Zustimmung als Frauen, für geringe Scores dreht sich das Bild jedoch um und Frauen tendieren eher zu einer Zustimmung (Interaktion Geschlecht © Score). Außerdem zeigt sich deutlich, daß der Geschlechtseffekt über die Zeit variiert (Interaktion Geschlecht © Zeit)

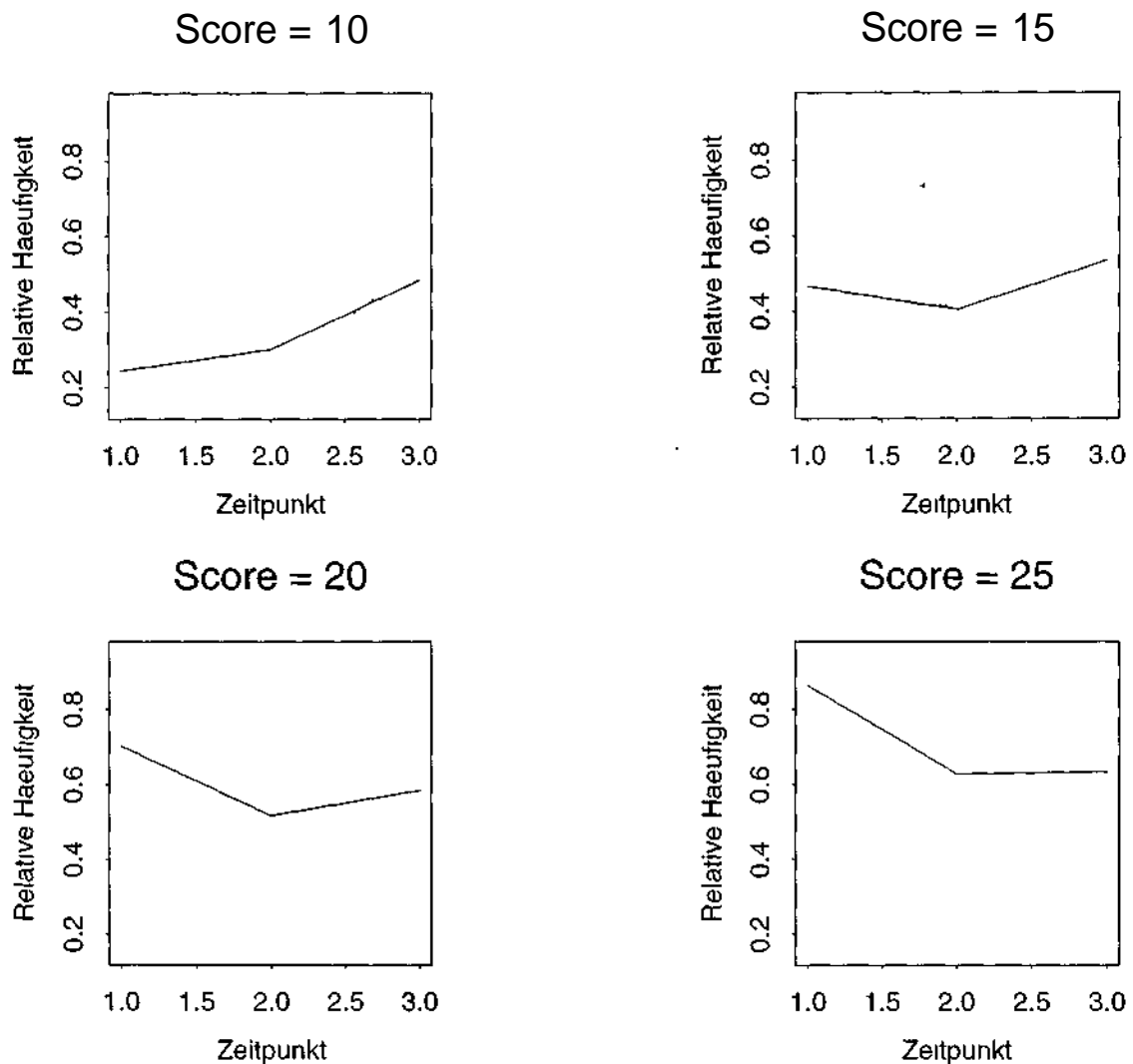


Abbildung 39 Zustimmung zur Aussage: "Ich bin stolz ein Deutscher zu sein" in Abhängigkeit zur Einstellung zu türkischen Mitbürgern (gemessen als Score, 10 = eher positiv bis 25 = eher negativ, — = Männer, - - = Frauen)

- "Ich möchte etwas tun, womit ich meinem Vaterland dienen kann "
- In analoger Form wie oben modellieren wir die Zustimmung zu obiger Aussage in Abhängigkeit von Geschlecht, Score und Zeit gemäß:

$$(Zeit \odot Geschlecht) \odot (Zeit \odot Score) \odot (Geschlecht \odot Score).$$

Es zeigt sich jedoch, daß der Score nicht signifikant mit der Zeit variiert (p-Wert 0.16 basierend auf einen Likelihood Quotienten Test). Somit gehen wir über zum Modell

$$(Zeit \odot Geschlecht) \odot (Geschlecht \odot Score)$$

wobei die Parameterschätzungen in Tabelle 62 gegeben sind. Die zugehörigen modellierten relativen Häufigkeiten sind in Abbildung 40 gegeben.

Parameter	Modellkomponente	Schätzung	Std.abw.
$\beta_i$	I	-0,97	0,31
$\beta_2$	} t	0,35	0,13
$\beta_3$		1,07	0,18
$\beta_{ö1}$	G	0,94	0,50
$\beta_{si}$	S	0,09	0,02
$\beta_{os}$	G⊙S	-0,08	0,03
$\beta_{G2} - \beta_{G1}$	} G ⊙ t	-0,32	0,21
$\beta_{G3} - \beta_{G1}$		-0,53	0,22

Tabelle 62: Parameterschätzung zur Aussage "Ich möchte etwas tun, womit ich meinem Vaterland dienen kann "

Man sieht, daß die Zustimmung zu der gestellten Aussage im Laufe der Zeit klar abnimmt. Ferner unterscheidet sich der Anteil der Zustimmer deutlich zwischen den Geschlechtern. Insbesondere für hohen Score, d.h. negative Einstellung zu Türken, ist der Geschlechtseffekt deutlich, wobei Männer deutlich häufiger der Aussage "Ich möchte etwas tun, womit ich meinem Vaterland dienen kann" zustimmen. Für geringen Score, d.h. positive Einstellung zu Türken, verschwindet der geschlechtsspezifische Unterschied nahezu (Interaktion Geschlecht ⊙ Score). Es zeigt sich ferner, daß sich der Geschlechtseffekt (insbesondere für hohe Scores) im Laufe der Zeit deutlich verstärkt (Interaktion Zeit ⊙ Geschlecht), was mit der Analyse unter 9.1 einhergeht.

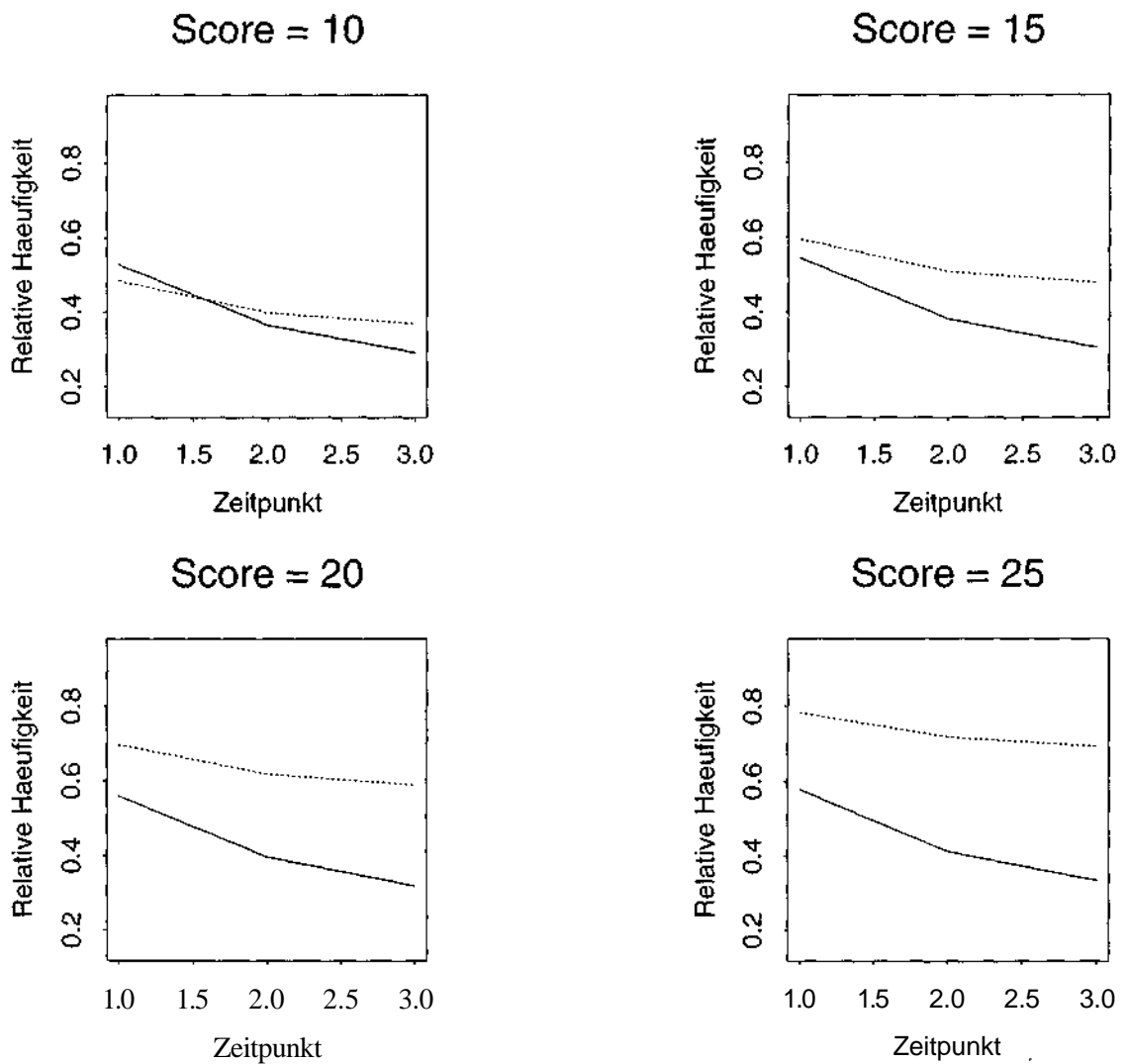


Abbildung 40: Zustimmung zur Aussage: "Ich möchte etwas tun, womit ich meinem Vaterland dienen kann" in Abhängigkeit zur Einstellung zu türkischen Mitbürgern (gemessen als Score, 10 = eher positiv bis 25 = eher negativ, — = Männer, - - = Frauen)

- *"Als Mensch und Staatsbürger fühle ich mich zuerst meinen Landsleuten hier in der ehemaligen DDR zugehörig"*

Es zeigt sich, daß weder die Einstellung zu Türken noch das Geschlecht einen signifikanten Einfluß auf die Zustimmung zu obiger Aussage hat. Die entsprechenden Parameter im Modell

Zeit © Geschlecht © Score

wie sie in Tabelle 63 gegeben sind, sind nicht signifikant. Somit ergibt sich im Vergleich zur Analyse in Kapitel 9.1 kein Unterschied.

Parameter	Modellkomponente	Schätzung	Standardabweichung
$\beta_i$	} t	0,88	0,27
$\beta_2$		-0,25	0,12
$\beta_s$		0,19	0,28
$\beta_c$	G	0,03	0,13
$\beta_s$	S	0,01	0,01

Tabelle 63: Parameterschätzung zur Aussage *"Als Mensch und Staatsbürger fühle ich mich zuerst meinen Landsleuten hier in der ehemaligen DDR zugehörig"*

- *"Ich würde auch meine eigenen Interessen zurückstellen, wenn ich damit dem Aufbau und der Entwicklung des vereinten Deutschlands dienen könnte"*

Auch bezüglich obiger Aussage ergibt sich kein Unterschied zur Analyse wie sie im Kapitel 9.1 zu finden ist. Dies bedeutet, daß der Einfluß der Einstellung zu Türken nicht signifikant ist, wohl aber, wie im Kapitel 9.1 bereits erhalten, der Geschlechtseffekt. Die Parameter finden sich in Tabelle 64.

Parameter	Modellkomponente	Schätzung	Standardabweichung
$\beta_i$	} t	-0,21	0,24
$\beta_2$		0,33	0,10
$\beta_z$		0,74	0,17
$\beta_a$	G	-0,28	0,11
$\beta_{si}$	S	0,01	0,01

Tabelle 64: Parameterschätzung zur Aussage *"Ich würde auch meine eigenen Interessen zurückstellen, wenn ich damit dem Aufbau und der Entwicklung des vereinten Deutschlands dienen könnte"*

- "Für mich ist es jetzt am wichtigsten, daß ich privat klar komme - in meiner Ausbildung, meinem Beruf und meinen persönlichen Beziehungen. Von gesellschaftlichen Verpflichtungen habe ich vorläufig die Nase voll."

Wir wollen nun den Einfluß der Einstellung zu Türken auf die Einstellung zu obiger Aussage analysieren. Es hatte sich bereits in Kapitel 9.1 gezeigt, daß kein signifikanter Geschlechtseffekt vorliegt. Wir können jedoch einen signifikanten Effekt des Scores nachweisen, der ebenfalls signifikant über die Zeit variiert. Die entsprechenden Parameter sind in Tabelle 65 gegeben.

Parameter	Modellkomponente	Schätzung	Standardabweichung
$\beta_i$	} t	0,59	0,38
$\beta_i$		-0,15	0,50
$\beta_z$		0,35	0,21
$\beta_{si}$	S	0,04	0,02
$\beta_{s2} - \beta_{si}$	} S ⊗ t	-0,02	0,03
$\beta_{s3} - \beta_{si}$		-0,06	0,03

Tabelle 65: parameterschätzer zur Aussage "Für mich ist es jetzt am wichtigsten, daß ich privat klar komme - in meiner Ausbildung, meinem Beruf und meinen persönlichen Beziehungen. Von gesellschaftlichen Verpflichtungen habe ich vorläufig die Nase voll"

Abbildung 41 verdeutlicht die gefundenen Aussagen. Man sieht, daß ein hoher Score, d.h. eine negative Einstellung zu Türken, im ersten Zeitpunkt zu einer deutlich höheren Zustimmung zu obiger Aussage führt d.h. zu einer höheren Gewichtung von privaten Interessen. Dieser Effekt verschwindet jedoch im Laufe der Zeit und kehrt sich zum dritten Zeitpunkt sogar um. Das bedeutet daß zum Zeitpunkt 3 eine ursprünglich negative Einstellung zu Türken im Vergleich zu einer positiven eher zur Ablehnung der gestellten Aussage führt.



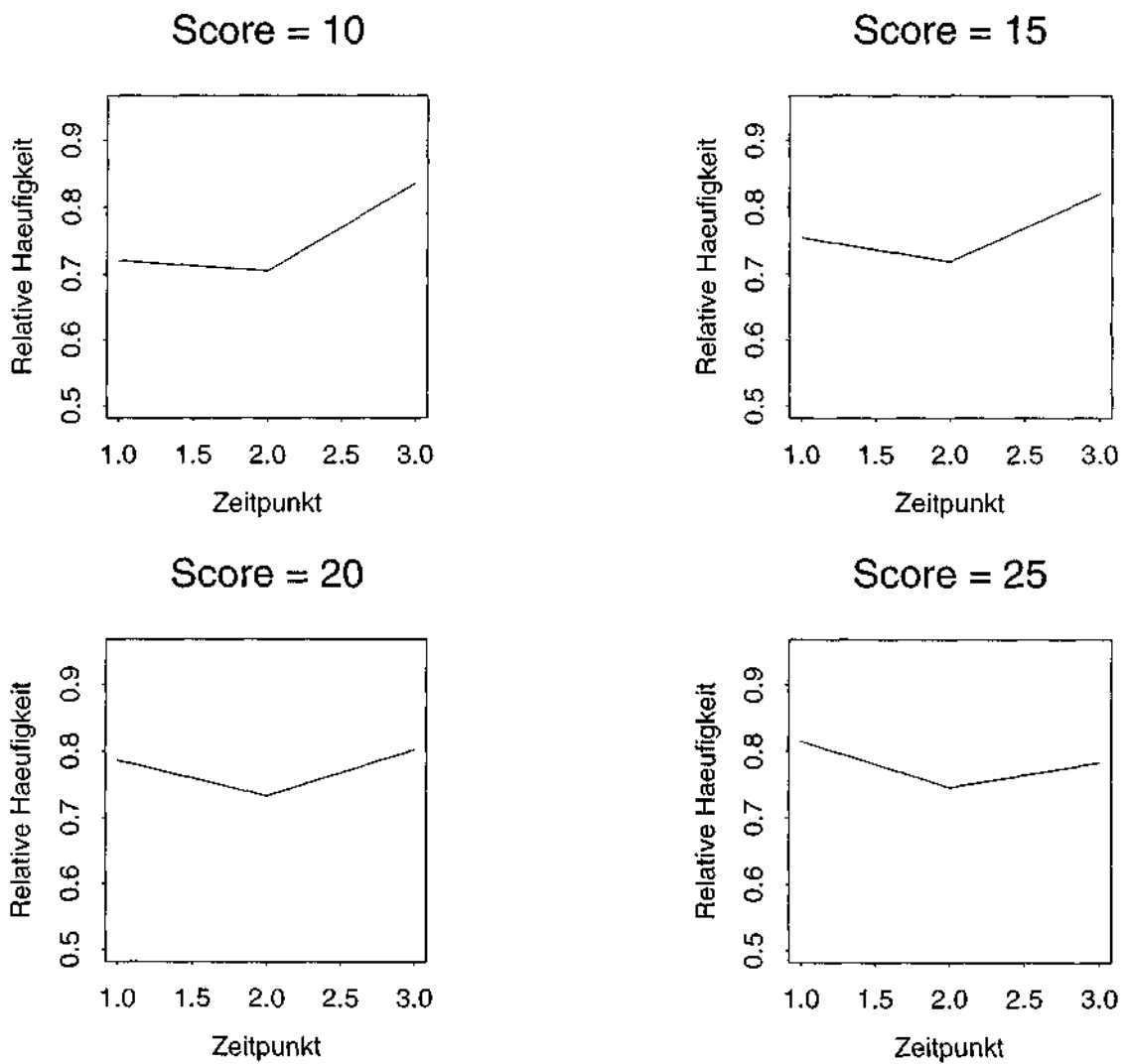


Abbildung 41: Zustimmung zur Aussage: "Für mich ist es jetzt am wichtigsten, daß ich privat klar komme, ( ... )" in Abhängigkeit zur Einstellung zu türkischen Mitbürgern (gemessen als Score, 10 = eher positiv bis 25 = eher negativ). (Kein signifikanter Geschlechtseffekt)

### 9.1.2 Querschnittsmodellierung mit Einstellung zur Politik

Es soll nun untersucht werden, inwieweit die Einstellung zur Politik und zu Politikern Einfluß auf die obigen Aussagen hat. Dabei wurden die Jugendlichen befragt, ob sie den Aussagen:

- *"Die Politik kann jemand wie ich nicht beeinflussen "*
- *"Die Politiker bei uns machen was sie wollen"*
- *"Die Politiker hören erst zu, wenn Randalen gemacht wird "*

zustimmen. Wir untersuchen den Einfluß der obigen Politikeinstellung getrennt und vernachlässigen somit eventuelle interaktive Effekte. Ferner berücksichtigen wir das Geschlecht, vernachlässigen jedoch eventuelle interaktive Effekte zwischen Geschlecht und Politikeinstellung. Wir modellieren somit gemäß dem marginalen Modell

$$h\{E\{Y_t|G, P_t\}\} = \beta_t + G\beta_{Gt} + P_t\beta_{P_t}$$

wobei  $P_t$  die Einstellung zur Politik zum Zeitpunkt  $t$  ist und  $G$ , wie oben, das Geschlecht. In Modellschreibweise ergibt sich somit

$$(\text{Zeit} \otimes \text{Geschlecht}) \otimes (\text{Zeit} \otimes \text{Politikeinstellung})$$

- *"Die Politik kann jemand wie ich nicht beeinflussen"*

Es zeigt sich, daß der Glaube an Beeinflussungsmöglichkeit in der Politik keinen signifikanten Einfluß auf die Einstellung zur Aussage *"Ich bin stolz ein Deutscher zu sein"* hat. Somit ergibt sich das Modell wie in Kapitel 9.1, hergeleitet. Anders sieht es jedoch bei den verbleibenden vier Aussagen aus.

— *"Ich möchte etwas tun, womit ich meinem Vaterland dienen kann "*

Hält ein Jugendlicher Politik für beeinflussbar, so neigt er eher zur Bereitschaft etwas zu tun, womit er dem Vaterland dienen kann. Dieser Effekt verschwindet jedoch über die Zeit und kehrt sich zum Zeitpunkt 3 sogar um (Interaktion  $P \otimes t$ ). Ebenfalls ist, wie in Kapitel 9.1, der Geschlechtseffekt zeitvariabel (Interaktion  $G \otimes t$ ). Die zugehörigen Schätzer sind in Tabelle 66 gegeben, die modellierten relativen Häufigkeiten sind in Abbildung 42 aufgelistet. Als Referenzlinien sind die Häufigkeiten ohne Berücksichtigung der Beeinflussungsmöglichkeit von Politik gegeben.

Parameter	Modellkomponente	Schätzung	Standardabweichung
$\beta_1$	1 t	0,83	0,29
$\beta_2$		0,01	0,30
$\beta_3$		1,19	0,18
$\beta_{ai}$	G	-0,47	0,17
$\beta_p$	P*	0,22	0,09
$\beta_{\ddot{o}2} \sim \beta_d$	} $P_t \text{ (3)t}$	-0,31	0,20
$\beta_{\ddot{u}z} - \beta_{\ddot{o}i}$		-0,53	0,21

Tabelle 66: Parameterschätzer

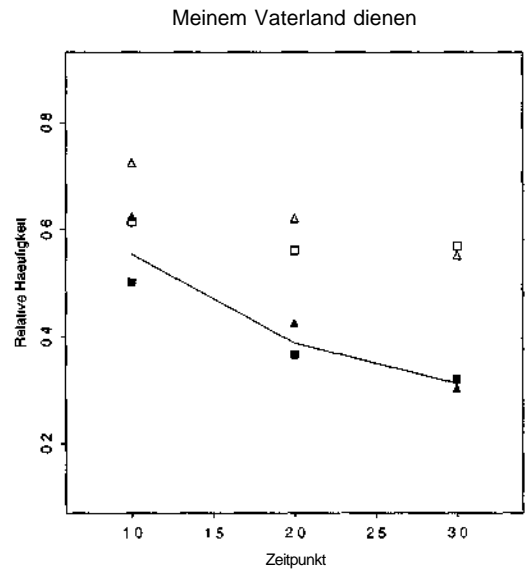


Abbildung 42: Zustimmung zur Aussage "Ich möchte etwas tun, womit ich meinem Vaterland dienen kann" in Abhängigkeit vom Geschlecht und der Meinung ob Politik beeinflussbar ist. Männer (—) die Politik für beeinflussbar (A) bzw. nicht beeinflussbar (D) halten. Frauen (---) die Politik für beeinflussbar (A) bzw. nicht beeinflussbar (D) halten

- "Als Mensch und Staatsbürger fühle ich mich zuerst meinen Landsleuten hier in der ehemaligen DDR zugehörig"

Auf die Frage nach dem Zugehörigkeitsgefühl mit der ehemaligen DDR hat der Glaube an Beeinflussungsmöglichkeit der Politik einen zeitkonstanten Effekt ( $p$ -Wert 0,16 bei einem Test auf Zeitvariation), wie Abbildung 43 verdeutlicht. Dabei fühlen sich Jugendliche, die nicht an die Möglichkeit der Beeinflussung der Politik glauben, eher der DDR zugehörig als Jugendliche, die Politik für beeinflussbar halten. Ein geschlechtsspezifischer Unterschied ist nicht nachweisbar. Die Schätzungen der Parameter finden sich in Tabelle 67.

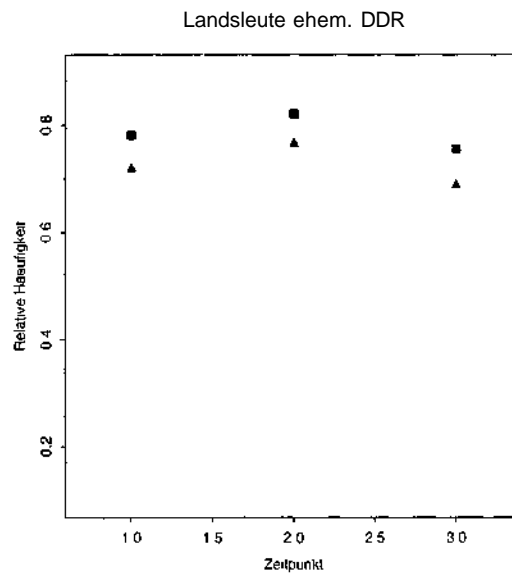


Abbildung 43: Zustimmung zur Aussage "Als Mensch und Staatsbürger fühle ich mich zuerst meinen Landsleuten hier in der ehemaligen DDR zugehörig" in Abhängigkeit vom Geschlecht und der Meinung ob Politik beeinflussbar ist. Jugendliche (—) die Politik für beeinflussbar ( $k$ ) bzw. nicht beeinflussbar ( $\bullet$ ) halten

Parameter	Modellkomponente	Schätzung	Standardabweichung
$\beta_i$	} t	1,60	0,19
$\beta_i$		-0,25	0,12
$\beta_s$		1,03	0,21
$\beta_p$	$P_t$	-0,31	0,11

Tabelle 67: Parameterschätzer

- "Ich würde auch meine eigenen Interessen zurückstellen, wenn ich damit dem Aufbau und der Entwicklung des vereinten Deutschlands dienen könnte "

In ähnlicher Form zeigt der Glaube an Politikbeeinflussung einen additiven Effekt auf die Bereitschaft für ein vereintes Deutschland zu arbeiten, wie Abbildung 44 zeigt. Halten Jugendliche Politik für beeinflussbar, so streben sie eher nach der Entwicklung eines vereinten Deutschlands, was auch an den Parameterschätzungen in Tabelle 68 ersehen werden kann. Wie oben bereits gezeigt, ergibt sich außerdem ein geschlechtsspezifischer Unterschied.

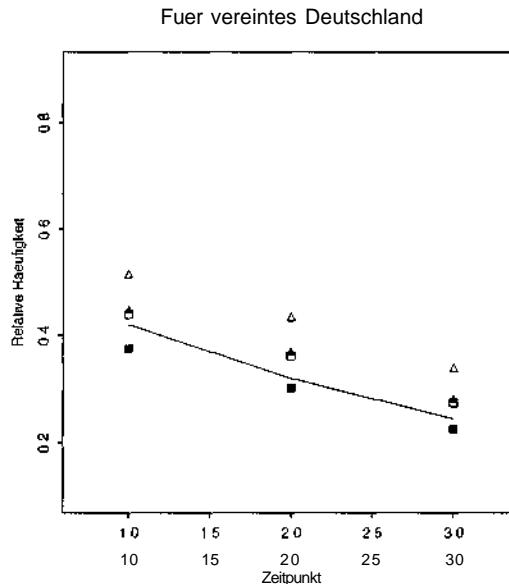


Abbildung 44: Zustimmung zur Aussage "Ich würde auch meine eigenen Interessen zurückstellen, wenn ich damit dem Aufbau und der Entwicklung des vereinten Deutschlands dienen könnte " in Abhängigkeit vom Geschlecht und der Meinung ob Politik beeinflussbar ist. Männer (—) die Politik für beeinflussbar (A) bzw. nicht beeinflussbar (D) halten. Frauen (---) die Politik für beeinflussbar (A) bzw. nicht beeinflussbar (D) halten

Parameter	Modellkomponente	Schätzung	Standardabweichung
$\beta_1$	t	-0,26	0,23
$\beta_2$		0,32	0,10
$\beta_3$		0,75	0,17
$\beta_a$	G	-0,27	0,11
$\beta_p$	P*	0,31	0,10

Tabelle 68: Parameterschätzer

- "Für mich ist es jetzt am wichtigsten, daß ich privat klar komme - in meiner Ausbildung, meinem Beruf und meinen persönlichen Beziehungen. Von gesellschaftlichen Verpflichtungen habe ich vorläufig die Nase voll."

Hier zeigt sich ein additiver jedoch zeitkonstanter Effekt der Politikerbeeinflussung (Abbildung 45) auf die Bereitschaft zu gesellschaftlichen Verpflichtungen, wobei Jugendliche, die Politik für nicht beeinflussbar halten, ihr Engagement eher im privaten Bereich suchen als diejenigen, die Beeinflussung von Politik für möglich halten.

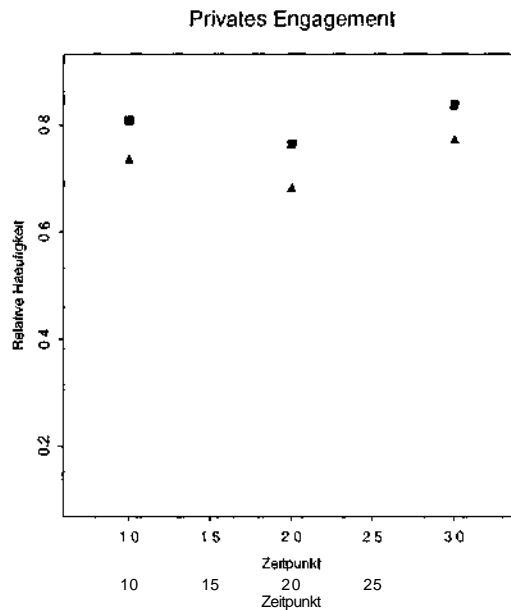


Abbildung 45: Zustimmung zur Aussage "Für mich ist es jetzt am wichtigsten, daß ich privat klar komme - in meiner Ausbildung, meinem Beruf und meinen persönlichen Beziehungen. Von gesellschaftlichen Verpflichtungen habe ich vorläufig die Nase voll." in Abhängigkeit vom Geschlecht und der Meinung ob Politik beeinflussbar ist. Jugendliche (—) die Politik für beeinflussbar (A) bzw. nicht beeinflussbar (•) halten

- "Die Politiker bei uns machen was sie wollen "

Wir betrachten nun, welchen Einfluß die Meinung über die Moral von Politikern auf die Einstellung zu den fünf gestellten Aussagen hat. Dabei läßt sich kein signifikanter Einfluß nachweisen auf die Aussagen "Ich bin stolz ein Deutscher zu sein " und "Ich möchte etwas tun, womit ich meinem Vaterland dienen kann".

- "Als Mensch und Staatsbürger fühle ich mich zuerst meinen Landsleuten hier in der ehemaligen DDR zugehörig "

Es zeigt sich, daß Jugendliche mit geringschätziger Meinung über die Moral von Politikern, d.h. Zustimmung zur Aussage "Die Politiker bei uns machen was sie wollen ", sich eher den Landsleuten der ehemaligen DDR zugehörig fühlen als Jugendliche mit guter Meinung über Politiker. Dieser Effekt verschwindet jedoch über die Zeit fast vollständig (Interaktion  $P \otimes t$ ), wie aus Abbildung 46 zu sehen. Die zugehörigen Schätzer sind in Tabelle 69 gegeben.

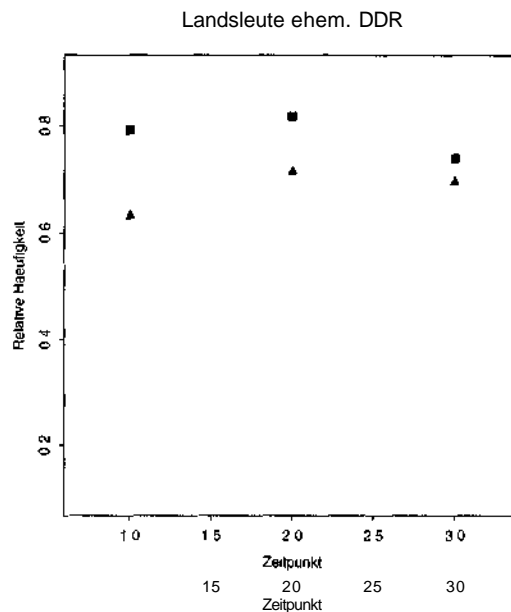


Abbildung 46: Zustimmung zur Aussage "Als Mensch und Staatsbürger fühle ich mich zuerst meinen Landsleuten hier in der ehemaligen DDR zugehörig " in Abhängigkeit von der Meinung über die Moral von Politikern (gute  $k$  bzw. geringschätzige  $\bullet$ )

Parameter	Modellkomponente	Schätzung	Standardabweichung
$\beta_i$	1 t	2,12	0,28
$\beta_2$		0,06	0,39
$\beta_s$		0,94	0,22
$\beta_P$	Pt	-0,79	0,20
$\beta_{pi}$	P_t ⊗ t	0,22	0,29
$\beta_{p2}$		0,58	0,27

Tabelle 69: Parameterschätzer

- "Ich würde auch meine eigenen Interessen zurückstellen, wenn ich damit dem Aufbau und der Entwicklung des vereinten Deutschlands dienen könnte "

Eine gute Meinung über die Moral von Politikern führt zu einer höheren Bereitschaft für ein vereintes Deutschland zu arbeiten (siehe Abbildung 47 und Tabelle 70). Interaktive Effekte sind jedoch nicht nachweisbar wohl aber ein Geschlechtseffekt.

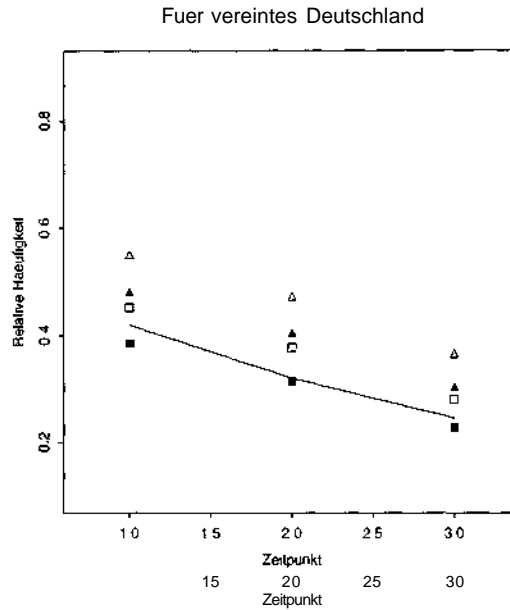


Abbildung 47: Zustimmung zur Aussage "Ich würde auch meine eigenen Interessen zurückstellen, wenn ich damit dem Aufbau und der Entwicklung des vereinten Deutschlands dienen könnte" in Abhängigkeit von der Meinung über die Moral von Politikern. Männer (—) die eine gute (A) bzw. schlechte (D) Meinung haben. Frauen (—) die eine gute (A) bzw. schlechte (•) Meinung haben

Parameter	Modellkomponente	Schätzung	Standardabweichung
$\beta_1$	} t	-0,31	0,23
$\beta_2$		0,31	0,10
$\beta_z$		0,77	0,17
$\beta_0$	G	-0,28	0,11
$\beta_p$	Pt	0,39	0,11

Tabelle 70: Parameterschätzer



- "Für mich ist es jetzt am wichtigsten, daß ich privat klar komme - in meiner Ausbildung, meinem Beruf und meinen persönlichen Beziehungen. Von gesellschaftlichen Verpflichtungen habe ich vorläufig die Nase voll."

Bei dieser Fragestellung zeigt sich, daß eine geringschätzige Meinung über die Moral von Politikern zu einem Zuspruch für privates Engagement anstelle von gesellschaftlichen Verpflichtungen führt (Abbildung 48, Tabelle 71).

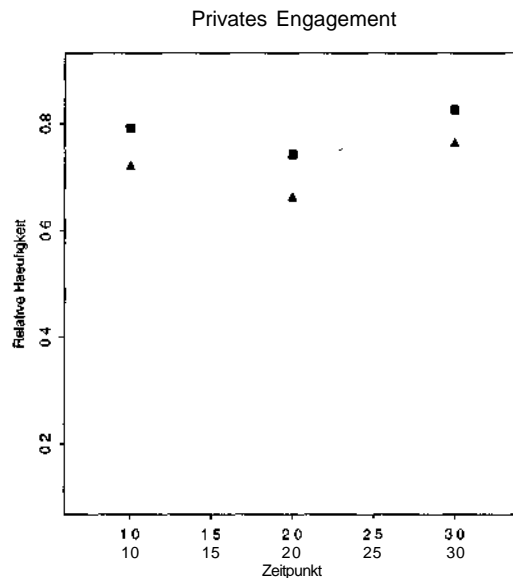


Abbildung 48: Zustimmung zur Aussage "Für mich ist es jetzt am wichtigsten, daß ich privat klar komme - in meiner Ausbildung, meinem Beruf und meinen persönlichen Beziehungen. Von gesellschaftlichen Verpflichtungen habe ich vorläufig die Nase voll." in Abhängigkeit von der Meinung über die Moral von Politikern (gute  $k$  bzw. geringschätzige  $\bullet$ )

Parameter	Modellkomponente	Schätzung	Standardabweichung
$\beta_i$	$\left. \begin{matrix} 1 \\ t \\ s \end{matrix} \right\}$	1,72	0,18
$\beta_i$		0,27	0,13
$\beta_s$		0,35	0,21
$\beta_p$	P*	-0,38	0,12

Tabelle 71: Parameterschätzer

- *"Die Politiker hören erst zu, wenn Randalen gemacht wird"*  
 Als letztes soll der Einfluß von Protest (Randalen) auf Politiker als Erklärungsgröße dienen. Dabei läßt sich kein signifikanter Einfluß auf *"Ich bin stolz ein Deutscher zu sein"*, *"Ich möchte etwas tun, womit ich meinem Vaterland dienen kann"* und *"Ich würde auch meine eigenen Interessen zurückstellen, wenn ich damit dem Aufbau und der Entwicklung des vereinten Deutschlands dienen könnte"* zeigen.
- *"Als Mensch und Staatsbürger fühle ich mich zuerst meinen Landsleuten hier in der ehemaligen DDR zugehörig"* und *"Für mich ist es jetzt am wichtigsten, daß ich privat klar komme (...)"*

In beiden Fragestellungen wirkt das Empfinden, daß nur Randalen Politiker zum Zuhören bringt, hemmend auf das Solidaritätsgefühl zu Landsleuten der ehemaligen DDR und auf die Zurückhaltung von gesellschaftlichen Verpflichtungen, wobei eine Zeitvariation nicht nachweisbar ist (siehe Graphiken 49, 50 und Tabellen 73, 73).

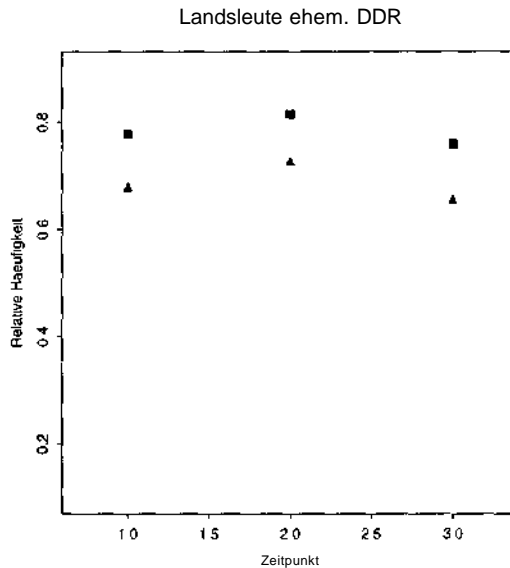


Abbildung 49: Zustimmung auf die Frage *"Als Mensch und Staatsbürger fühle ich mich zuerst meinen Landsleuten hier in der ehemaligen DDR zugehörig"* in Abhängigkeit von der Meinung über Politiker. Politiker hören erst zu, wenn man Randalen macht (A richtig, • falsch)

Parameter	Modellkomponente	Schätzung	Standardabweichung
$\beta_1$	) t	1,75	0,18
$\beta_2$		-0,22	0,12
$\beta_z$		0,96	0,21
$\beta_P$	P	-0,50	0,12

Tabelle 72: Parameterschätzer

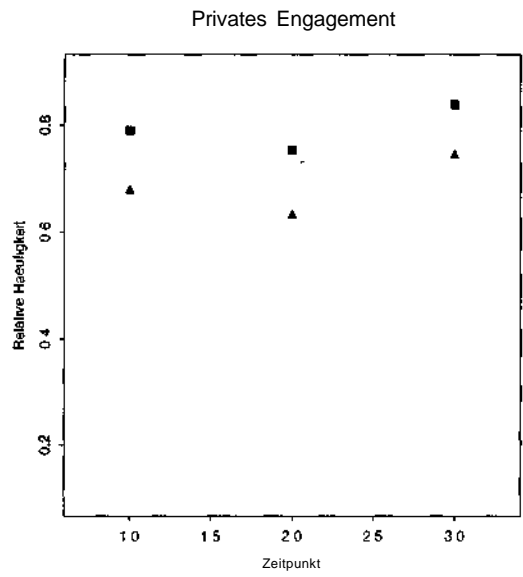


Abbildung 50: Zustimmung zur Aussage "Für mich ist es jetzt am wichtigsten, daß ich privat klar komme - in meiner Ausbildung, meinem Beruf und meinen persönlichen Beziehungen. Von gesellschaftlichen Verpflichtungen habe ich vorläufig die Nase voll" in Abhängigkeit von der Meinung über Politiker. Politiker hören erst zu, wenn man Randalen macht (A richtig, • falsch)

Parameter	Modellkomponente	Schätzung	Standardabweichung
$\beta_1$	} t	1,90	0,18
$\beta_2$		0,21	0,13
$\beta_3$		0,37	0,21
$\beta_P$	P	-0,57	0,13

Tabelle 73: Parameterschätzer

## 9.2 Markovanalyse

Abschließend und zusammenfassend sollen die obigen Fragestellungen mit Markovketten (vgl. Abschnitt 2.1) analysiert werden. Hierzu betrachten wir in allen drei Perioden, welche Faktoren Einfluß auf die Zustimmung oder Ablehnung der fünf genannten politischen Aussagen haben. Hierbei sollen individuenspezifische Gesichtspunkte im Vordergrund stehen. Wir modellieren (vgl. 8)

$$E(y_{it}|x_{it}, y_{it-1}) = h(\beta_t + x_{it}\beta_{xt} + y_{it-1}\beta_{yt})$$

wobei  $y_{it}$  die Zustimmung oder Ablehnung des  $i$ -ten Jugendlichen zu einer der gestellten Aussagen in Periode  $t$  angibt und  $x_{it}$  zusätzliche Regressoren sind. Der Koeffizient  $\beta_{xt}$  beschreibt dabei den Einfluß der Regressoren  $x_{it}$  und  $\beta_{yt}$  spiegelt den Einfluß der Meinung des Jugendlichen in der Vorperiode wider. Somit hängt also die Meinung des Jugendlichen zu einer der gegebenen Fragestellungen von erklärenden Größen  $x_{it}$  und seiner bisherigen Meinung ab. Eine Modellierung derart, daß auch Ergebnisse der Vor-Vorperiode berücksichtigt werden, erwies sich als nicht signifikant. Als Regressoren  $x_{it}$  ziehen wir

- das *Geschlecht*
- die *Einstellungen zu Türken* (im folgenden auch abgekürzt mit TE)
- Einstellung zu Politik in den drei Aussagen
  - *Politik kann man nicht beeinflussen* (Politik 1)
  - *Politiker machen was sie wollen* (Politik 2)
  - *Politiker hören erst zu, wenn man Randalen macht* (Politik 3)

heran. Der Einfachheit halber wurde dabei generell auf die Modellierung von Interaktionen verzichtet. Wir beziehen uns im folgenden nur auf signifikante Parameter. Die Ergebnisse sind, wie in Abschnitt 2.1 motiviert, mit Graphen veranschaulicht.

- *"Ich bin stolz ein Deutscher zu sein"*  
 Aus Abbildung 51 (bzw. Tabellen 74, 75 und 76) zeigt sich, daß in allen drei Wellen eine negative Einstellung zu Türken die Zustimmung zu obiger Aussage fördert. In der dritten Welle ergibt sich ferner ein klarer geschlechtsspezifischer Unterschied. In zweiter und dritter Welle neigen Jugendliche, die in Vorperiode stolz waren Deutsche zu sein, weiter klar zu dieser Aussage

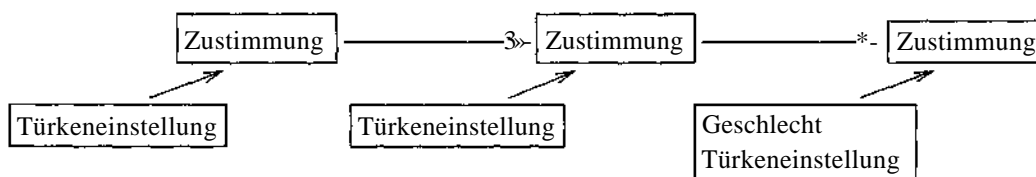


Abbildung 51: Abhängigkeit der Zustimmung zur Aussage "Ich bin stolz ein Deutscher zu sein"

Parameter	Schätzer	Std.abw.	t-Wert
$\beta_0$	-4.033	0.276	-14.609
Türkeneinst.	0.231	0.014	16.079

Tabelle 74: Parameterschätzer erste Welle

Parameter	Schätzer	Std.abw.	t-Wert
$\beta_0$	-1.088	0.323	-3.359
Vorwelle	0.645	0.074	8.705
Türkeneinst.	0.058	0.016	3.468

Tabelle 75: Parameterschätzer zweite Welle

Parameter	Schätzer	Std.abw.	t-Wert
$\beta_0$	-2.741	0.511	-5.364
Vorwelle	0.716	0.092	7.765
Geschlecht	0.314	0.091	3.448
Türkeneinst.	0.184	0.027	6.673

Tabelle 76: Parameterschätzer dritte Welle

- *"Ich möchte etwas tun, womit ich meinem Vaterland dienen kann"*

Auch bei dieser Fragestellung tritt die negative Einstellung zu Türken signifikant in allen drei Zeitpunkten als Einfluß auf. Ferner zeigen sich zu allen drei Zeitpunkten geschlechtsspezifische Unterschiede, wobei sich Männer deutlich häufiger mit der obigen Aussage identifizieren. In den ersten beiden Perioden führt weiterhin eine negative Einstellung zur Politik (*"Politik kann jemand wie ich nicht beeinflussen"*) wie auch eine schlechte Meinung über die Moral von Politikern (*"Politiker machen was sie wollen"*) zu einer Zurückhaltung, *"etwas für sein Vaterland"* tun zu wollen. Zu den letzten beiden Zeitpunkten ist außerdem die bisherige Meinung beeinflussend.

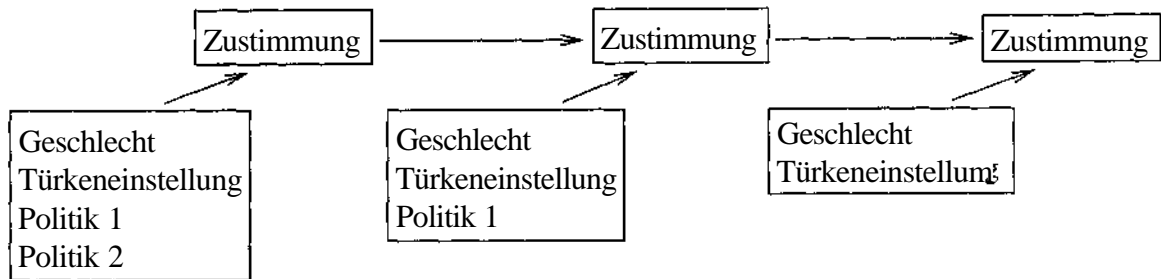


Abbildung 52: Abhängigkeit der Zustimmung zur Aussage *"Ich möchte etwas tun, womit ich meinem Vaterland dienen kann"*

Parameter	Schätzer	Std.abw.	t-Wert
$\beta_0$	-1.383	0.215	-6.412
Geschlecht	0.242	0.048	4.962
Türkeneinst.	0.096	0.011	8.637
Politik 1	-0.155	0.050	-3.087
Politik 2	-0.167	0.060	-2.791

Tabelle 77: Parameterschätzer erste Welle

Parameter	Schätzer	Std.abw.	t-Wert
$\beta_0$	-0.976	0.300	-3.253
Vorwelle	0.532	0.069	7.627
Geschlecht	0.274	0.068	4.011
Türkeneinst.	0.048	0.015	3.124
Politik 1	-0.229	0.069	-3.302

Tabelle 78: Parameterschätzer zweite Welle

Parameter	Schätzer	Std.abw.	t-Wert
$\beta_0$	-2.238	0.434	-5.149
Vorwelle	0.560	0.084	6.638
Geschlecht	0.426	0.086	4.920
Türkeneinst.	0.103	0.022	4.599

Tabelle 79: Parameterschätzer dritte Welle

- *"Als Mensch und Staatsbürger fühle ich mich zuerst meinen Landsleuten hier in der ehemaligen DDR zugehörig"*  
Das Zugehörigkeitsgefühl zur ehemaligen DDR ist deutlich beeinflusst von der Meinung über die Politik. Jugendliche, die der Meinung sind, daß Politiker nur durch "Randalen" beeinflussbar sind, fühlen sich der DDR eher zugehörig. Ebenso hat die generelle Meinung *"Politik ist von mir nicht beeinflussbar"* positiven Effekt auf das Zugehörigkeitsempfinden zur DDR. In den letzten beiden Perioden zeigt sich ferner eine Abhängigkeit zur Meinung in der Vorperiode.

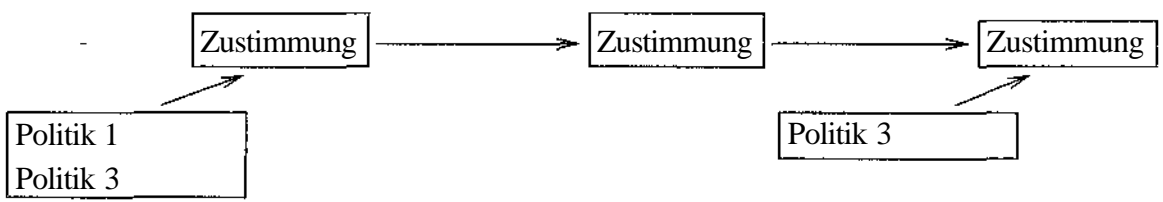


Abbildung 53: Abhängigkeit der Zustimmung zur Aussage "Als Mensch und Staatsbürger fühle ich mich zuerst meinen Landsleuten hier in der ehemaligen DDR zugehörig"

Parameter	Schätzer	Std.abw.	t-Wert
$\beta_0$	0.823	0.069	11.904
Politik 1	0.184	0.062	2.965
Politik 3	0.197	0.072	2.735

Tabelle 80: Parameterschätzer erste Welle

Parameter	Schätzer	Std.abw.	t-Wert
$\beta_0$	0.651	0.097	6.686
Vorwelle	0.484	0.097	4.971

Tabelle 82: Parameterschätzer dritte Welle

Parameter	Schätzer	Std.abw.	t-Wert
$\beta_0$	0.959	0.090	10.649
Vorwelle	0.510	0.080	6.332
Politik 3	0.230	0.087	2.633

Tabelle 81: Parameterschätzer zweite Welle

- *"Ich würde auch meine eigenen Interessen zurückstellen, wenn ich damit dem Aufbau und der Entwicklung des vereinten Deutschlands dienen könnte "*

Bei dieser Aussage ergibt sich, daß eine negative Meinung über die Moral von Politikern (Politik 2) demotivierend für den Aufbau eines vereinten Deutschland wirkt. Dieser Effekt ist in den ersten beiden Perioden signifikant, wobei in der ersten Periode generell eine negative Meinung über Politik (Politik 1) eher zur Ablehnung der obigen Aussage führt. Wie bei den bisherigen Fragestellungen zeigt sich die Abhängigkeit der Meinung zur Meinung in der Vorperiode.

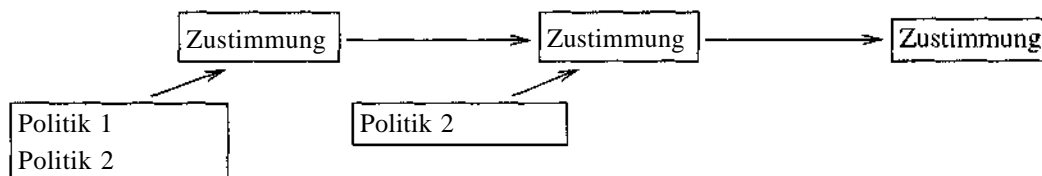


Abbildung 54: Abhängigkeit der Zustimmung zur Aussage *"Ich würde auch meine eigenen Interessen zurückstellen, wenn ich damit dem Aufbau und der Entwicklung des vereinten Deutschlands dienen könnte "*

Parameter	Schätzer	Std.abw.	t-Wert
$\beta_0$	-0.234	0.054	-4.310
Politik 1	-0.126	0.047	-2.659
Politik 2	-0.151	0.055	-2.725

Tabelle 83: Parameterschätzer erste Welle

Parameter	Schätzer	Std.abw.	t-Wert
$\beta_0$	-0.345	0.081	-4.266
Vorwelle	0.383	0.066	5.782
Politik 2	-0.215	0.080	-2.660

Tabelle 84: Parameterschätzer zweite Welle



Parameter	Schätzer	Std.abw.	t-Wert
$\beta_0$	-0.878	0.087	-10.082
Vorwelle	0.359	0.087	4.122

Tabelle 85: Parameterschätzer dritte Welle

- *"Für mich ist es jetzt am wichtigsten, daß ich privat klar komme - in meiner Ausbildung, meinem Beruf und meinen persönlichen Beziehungen. Von gesellschaftlichen Verpflichtungen habe ich vorläufig die Nase voll"*

Die Bereitschaft sich gesellschaftlichen Pflichten zu stellen ist klar von der Meinung über Politik und Politiker abhängig, ein geschlechtsspezifischer Unterschied ist nicht nachweisbar. Allgemein gesprochen ergibt sich ein ähnliches Bild wie bei der vorherigen Fragestellung.

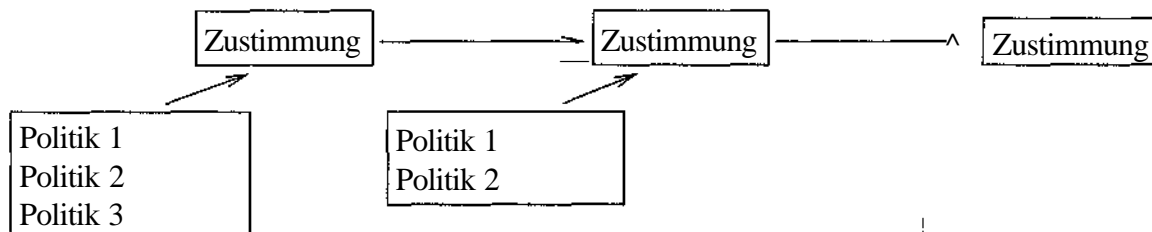


Abbildung 55: Abhängigkeit der Zustimmung zur Aussage *"Für mich ist es jetzt am wichtigsten, daß ich privat klar komme — in meiner Ausbildung, meinem Beruf und meinen persönlichen Beziehungen. Von gesellschaftlichen Verpflichtungen habe ich vorläufig die Nase voll"*

Parameter	Schätzer	Std.abw.	t-Wert
$\beta_0$	1.055	0.072	14.595
Politik 1	0.235	0.058	4.049
Politik 2	0.181	0.067	2.684
Politik 3	0.287	0.075	3.826

Tabelle 86: Parameterschätzer erste Welle

Parameter	Schätzer	Std.abw.	t-Wert
$\beta_0$	0.897	0.084	10.678
Politik 1	0.217	0.075	2.889
Politik 2	0.255	0.087	2.935

Tabelle 87: Parameterschätzer zweite Welle

Parameter	Schätzer	Std.abw.	t-Wert
$\beta_0$	1.311	0.102	12.806
Vorwelle	0.255	0.102	2.492

Tabelle 88: Parameterschätzer dritte Welle

### **Zusammenfassung**

Während die ersten beiden Aussagen eher nationalistische Gedanken der Befragten zutage fördern, ist in den letzten drei Aussagen die private Lebenssituation der Befragten im Rahmen des Vereinigungsprozesses aufgedeckt. Dies ist insbesondere auch daran zu erkennen, daß sich die relevanten Einflußgrößen relativ gut trennen lassen. Die Einstellung gegenüber Türken wirkt sich bei den letzten drei Aussagen nicht aus, dafür die Meinung über die Politik umso mehr.

## 10 Mobilitätsbereitschaft

Der folgende Abschnitt beschäftigt sich mit der Bereitschaft der Jugendlichen, ihren Lebensort zu wechseln. Dabei sollten die Jugendlichen zum ersten und dritten Zeitpunkt angeben, ob und wie weit sie bereit wären aus beruflichen Gründen umzuziehen. Als mögliche Zielorte wurden gegeben

- 1: *In einem Ort hier in der Nähe*
- 2: *In einen Ort weiter weg, aber im gleichen Bundesland*
- 3: *In einen weiter entfernten Ort hier in den neuen Bundesländern*
- 4: *Nach Westdeutschland*
- 5: *Ins osteuropäische Ausland (Polen, CSFR etc.)*
- 6: *Ins westeuropäische Ausland (Österreich, Schweiz, Dänemark, England etc.)*
- 7: *Nach Übersee (Amerika, Kanada, Australien etc.)*

Als mögliche Antworten standen den Jugendlichen "Ja, sicher", "Vielleicht" oder "Nein, nicht" zur Auswahl. Die entsprechenden Daten sind in Tabellen 89-95 gegeben.

Zeitpunkt 1	Zeitpunkt 3			Gesamt
	Ja, sicher	Vielleicht	Nein, nicht	
Ja, sicher	232 33.62	154 22.32	25 3.62	411 59.57
Vielleicht	112 16.23	90 13.04	24 3.48	226 32.75
Nein, nicht	21 3.04	23 3.33	9 1.30	53 7.68
Gesamt	365 52.90	267 38.70	58 8.41	690 100.00

Tabelle 89: *Umzug in einen anderen Ort hier in der Nähe?*

	Zeitpunkt 3			
Zeitpunkt 1	Ja, sicher	Vielleicht	Nein, nicht	Gesamt
Ja, sicher	86 12.52	108 15.72	24 3.49	218 31.73
Vielleicht	111 16.16	179 26.06	44 6.40	334 48.62
Nein, nicht	40 5.82	68 9.90	27 3.93	135 19.65
Gesamt	237 34.50	355 51.67	95 13.83	687 100.00

Tabelle 90: *Umzug in einen anderen Ort weiter weg, aber im gleichen Bundesland?*

	Zeitpunkt 3			
Zeitpunkt 1	Ja, sicher	Vielleicht	Nein, nicht	Gesamt
Ja, sicher	40 5.81	87 12.65	27 3.92	154 22.38
Vielleicht	59 8.58	134 19.48	63 9.16	256 37.21
Nein, nicht	46 6.69	130 18.90	102 14.83	278 40.41
Gesamt	145 21.08	351 51.02	192 27.91	688 100.00

Tabelle 91: *Umzug in einen weiter entfernten Ort hier in den neuen Bundesländern?*

	Zeitpunkt 3			
Zeitpunkt 1	Ja, sicher	Vielleicht	Nein, nicht	Gesamt
Ja, sicher	23 3.35	61 8.89	46 6.71	130 18.95
Vielleicht	46 6.71	129 18.80	91 13.27	266 38.78
Nein, nicht	34 4.96	94 13.70	162 23.62	290 42.27
Gesamt	103 15.01	284 41.40	299 43.59	686 100.00

Tabelle 92: *Umzug nach Westdeutschland?*

	Zeitpunkt 3			
Zeitpunkt 1	Ja, sicher	Vielleicht	Nein, nicht	Gesamt
Ja, sicher	1 0.15	7 1.02	14 2.03	22 3.20
Vielleicht	1 0.15	18 2.62	58 8.43	77 11.19
Nein, nicht	9 1.31	80 11.63	500 72.67	589 85.61
Gesamt	11 1.60	105 15.26	572 83.14	688 100.00

Tabelle 93: *Umzug ins osteuropäische Ausland (Polen, CSFR etc.)?*

	Zeitpunkt 3			
Zeitpunkt 1	Ja, sicher	Vielleicht	Nein, nicht	Gesamt
Ja, sicher	21 3.04	59 8.55	58 8.41	138 20.00
Vielleicht	29 4.20	99 14.35	122 17.68	250 36.23
Nein, nicht	20 2.90	79 11.45	203 29.42	302 43.77
Gesamt	70 10.14	237 34.35	383 55.51	690 100.00

Tabelle 94: *Umzug ins westeuropäische Ausland (Österreich, Schweiz, Dänemark, England etc.)?*

Zeitpunkt 1	Zeitpunkt 3			Gesamt
	Ja, sicher	Vielleicht	Nein, nicht	
Ja, sicher	30 4.35	57 8.26	81 11.74	168 24.35
Vielleicht	19 2.75	74 10.72	118 17.10	211 30.58
Nein, nicht	28 4.06	67 9.71	216 31.30	311 45.07
Gesamt	77 11.16	198 28.70	415 60.14	690 100.00

Tabelle 95: *Umzug nach Übersee (Amerika, Kanada, Australien etc.)?*

## 10.1 Marginale Modelle

Um einen ersten Eindruck der Daten zu vermitteln, sind in Abbildung 56 die Häufigkeiten für Mobilitätsbereitschaft gegeben, d.h. die relative Häufigkeit der "Ja-Sager". Generell läßt sich sagen, daß die Bereitschaft zu migrieren mit wachsender Entfernung vom jetzigen Wohnsitz sinkt, wobei das Ausland, insbesondere in Periode 3, die wenigsten Jugendlichen reizt. Ganz besonders stößt eine Auswanderung nach Osteuropa auf Ablehnung. Es zeigt sich ferner, daß Jugendliche in der dritten Periode weniger zum Wechsel in ein europäisches Ausland oder nach Übersee bereit sind als zu Beginn ihrer Lehre. Die zeitliche Veränderung in den beiden Graphiken soll Gegenstand der folgenden Untersuchungen sein.

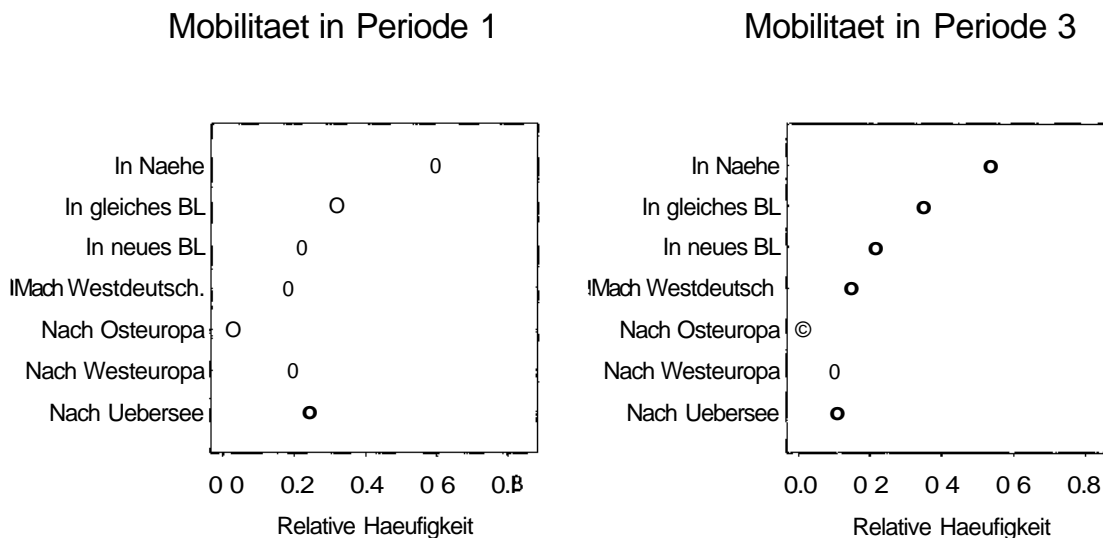


Abbildung 56. Mobilitätsbereitschaft der Jugendlichen

Als erstes untersuchen wir, wie sich die Mobilitätsbereitschaft über die Zeit verändert. Hierzu betrachten wir, ob in den einzelnen Tabellen 89-95 Quasi-Unabhängigkeit angenommen werden kann. Das bedeutet, wir untersuchen, ob unter den Jugendlichen, deren Mobilitätsbereitschaft sich ändert (Wechsler), Struktur vorliegt, oder ob die Änderung zufällig geschieht. Es zeigt sich, daß man in allen sieben Fragestellungen von Quasi-Unabhängigkeit ausgehen kann (siehe Tabelle 96). Das bedeutet, wechselt ein Jugendlicher seine Meinung, so geschieht dies unabhängig von seiner bisherigen Mobilitätsbereitschaft. Nur bei der letzten Fragestellung "Umzug nach Übersee (Amerika, Kanada, Australien etc.)" zeigt sich mit einem p-Wert von 0,091 eine leichte, wenn auch nicht signifikante Struktur.

Als nächstes soll untersucht werden, wie sich die Mobilitätsbereitschaft über die Zeit verändert. Dazu testen wir auf marginale Homogenität (vgl. 3.2), d.h. wir untersuchen, ob die Mobilitätsbereitschaft im Mittel mit der Zeit variiert. Der Größe der Stichprobe wegen, setzen wir hier ein Signifikanzniveau von  $\alpha = 0,01$  an, d.h. nur wenn die empirischen Signifikanzen kleiner als 0,01 sind, wird die marginale Homogenität signifikant verworfen. Zum Testen wurde dabei nur auf die Beobachtungen zurückgegriffen, von denen vollständige Information zu

Frage	p-Wert
In Nähe	0,667
In gleiches BL	0,691
In neues BL	0,495
Nach Westdeutsch.	0,914
Nach Osteuropa	0,229
Nach Westeuropa	0,776
Nach Übersee	0,091

Tabelle 96: Empirische Signifikanzen bei Test auf Quasi-Unabhängigkeit

beiden Zeitpunkten und zu allen Fragestellungen vorlag. Der Gesamtstichprobenumfang umfaßte damit 679 Beobachtungen. Es zeigt sich aus Tabelle 97, daß marginale Homogenität für die Bereitschaft ins gleiche oder in ein anderes jedoch neues Bundesland zu ziehen verworfen werden muß. Ebenso wird die marginale Homogenität für Umzüge nach Westeuropa oder nach Übersee verworfen.

Die nachweisbare zeitliche Struktur ist in Abbildung 57 visualisiert. Es zeigt sich, daß die Bereitschaft ins gleiche Bundesland zu ziehen steigt und die entsprechende Ablehnung sinkt. Ebenso sinkt die Ablehnung in ein anderes, neues Bundesland zu ziehen, wobei der Anteil der Befürworter eher gleich bleibt, somit also der Anteil der Unentschlossenen abnimmt.

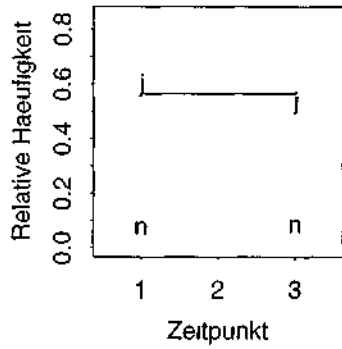
Frage	p-Wert
In Nähe	0,041
In gleiches BL	0,007
In neues BL	<0,001
Nach Westdeutsch.	0,153
Nach Osteuropa	0,034
Nach Westeuropa	<0,001
Nach Übersee	<0,001

Tabelle 97: Empirische Signifikanzen beim Test auf marginale Homogenität

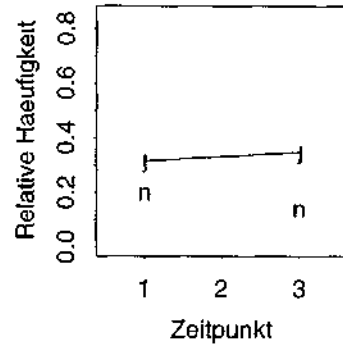
Die Bereitschaft nach Westdeutschland zu ziehen unterliegt keiner nachweisbaren Veränderung. Ebenso ist die Ablehnung eines Umzugs ins osteuropäische Ausland über beide Perioden hinweg unverändert. Die deutlichsten Veränderungen zeigen sich in der Migrationsbereitschaft nach Westeuropa und nach Übersee, wobei sich beide Fragestellungen annähernd gleich verhalten. In beiden Fällen sinkt die Bereitschaft eines Umzugs und die Ablehnung steigt.



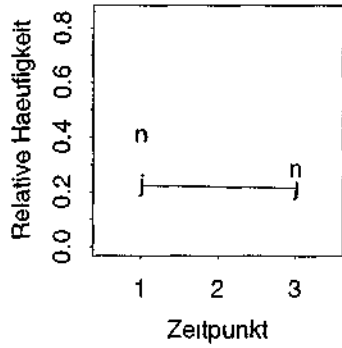
Ort in der Naeh



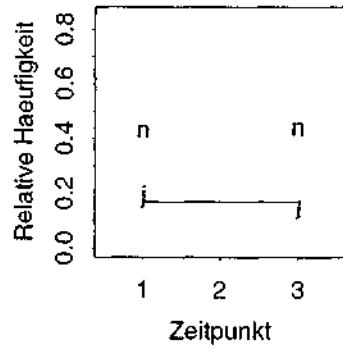
in gleiches Bundesl.



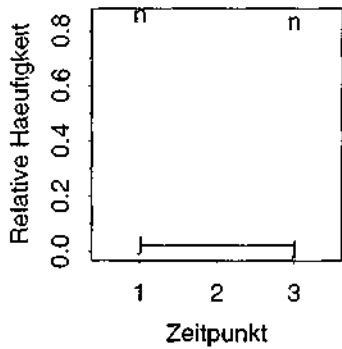
In neue Bundesl.



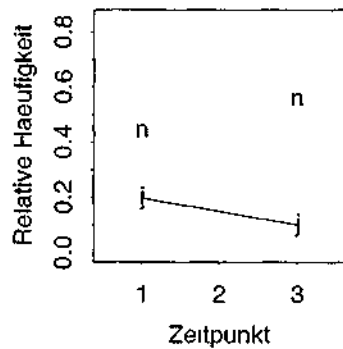
nach Westdeutchl.



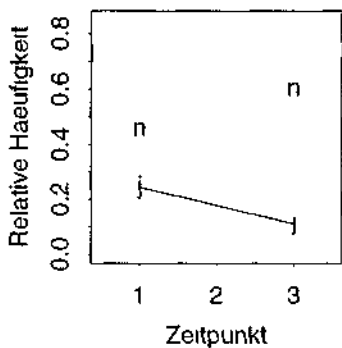
nach Osteuropa



nach Westeuropa



nach Uebersee



## 10.2 Markov Analyse

Die Fragestellung der Mobilitätsbereitschaft soll nun individuenpezifisch mit Hilfe einer Markov Analyse behandelt werden. Damit wird die Einstellung zur Mobilität der ersten Welle anhand der zu diesen Zeitpunkten zur Verfügung stehenden Einflußgrößen (Wirtschaftsbereich, Geschlecht) modelliert. In der dritten Welle fließt neben den vorhandenen Einflußgrößen auch die Aussage der ersten Welle in die Betrachtung ein.

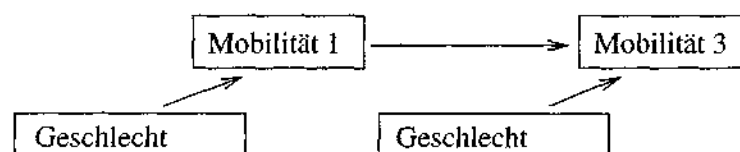
Da die Antworten zur Mobilitätsbereitschaft eine ordinale Struktur aufweisen, wird ein kumulatives logistisches Regressionsmodell verwendet (vergleiche Abschnitt 1.1). Die verwendeten Einflußgrößen sind überwiegend nominal skaliert und bedürfen einer Kodierung, um in Regressionsmodellen verwendet werden zu können. Aus Interpretationsgründen wurde die Effektkodierung verwendet. Diese Kodierungsform gestatten die Interpretation der Parameter als Abweichungen vom (hypothetischen) Durchschnitt. Aus technischen Gründen können bei einer Variablen mit  $k$  Kategorien (bspw. Geschlecht mit männlich/weiblich) nur  $k - 1$  Kategorien (hier: männlich) explizit erfaßt werden. Die  $k$ -te Kategorie ergibt sich aus der verwendeten Kodierungsform. Das verwendete Kodierschema für die nachfolgend betrachteten Variablen läßt sich auch etwas formaler darstellen:

$$\text{Sex} = \begin{cases} 1 & \text{männlich} \\ -1 & \text{weiblich} \end{cases} \quad \text{vorwelle1} = \begin{cases} 1 & \text{ja} \\ 0 & \text{vielleicht} \\ -1 & \text{nein} \end{cases} \quad \text{vorwelle2} = \begin{cases} 1 & \text{ja} \\ 0 & \text{vielleicht} \\ -1 & \text{nein} \end{cases}$$

$$\text{wiber1} = \begin{cases} 1 & \text{Landw.+Handwerk} \\ 0 & \text{sonst} \\ -1 & \text{freie, Vers., sonstiges} \end{cases} \quad \text{wiber2} = \begin{cases} 1 & \text{Industrie} \\ 0 & \text{sonst} \\ -1 & \text{freie, Vers., sonstiges} \end{cases}$$

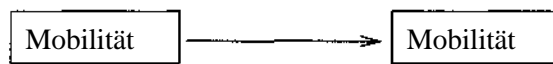
$$\text{wiber3} = \begin{cases} 1 & \text{Handel} \\ 0 & \text{sonst} \\ -1 & \text{freie, Vers., sonstiges} \end{cases} \quad \text{wiber4} = \begin{cases} 1 & \text{öffentl. Dienst} \\ 0 & \text{sonst} \\ -1 & \text{freie, Vers., sonstiges} \end{cases}$$

Zunächst untersuchen wir, welchen Einfluß das Geschlecht auf die Mobilitätsbereitschaft spielt. Hierzu betrachten wir ein Markov Modell der Form

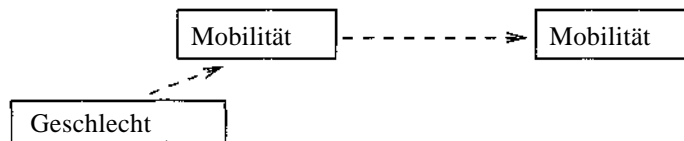


wobei Mobilität 1 bzw. Mobilität 3 die Mobilitätsbereitschaft in Perioden eins und drei angibt. Wir untersuchen nun, welche Einflüsse signifikant nachweisbar sind z.B. in welchen Perioden beim Geschlecht signifikante Unterschiede auffallen. Um einen Überblick über die Stärke des Zusammenhangs zu bekommen, stellen wir Verbindungen die zu einem Signifikanzniveau von 0,01 existieren als durchgezogene Linien dar. Abhängigkeiten, die nur bei einem Signifikanzniveau von 0,05 als signifikant auffallen, werden gestrichelt dargestellt<sup>4</sup>. Es ergeben sich folgende Bilder:

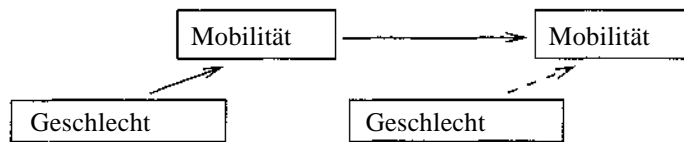
- *"In einen Ort in der Nähe "*



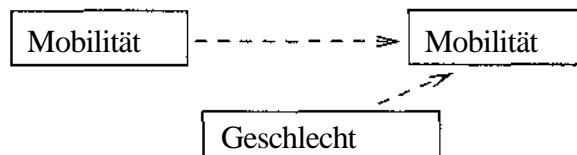
- *"In gleiches Bundesland"*



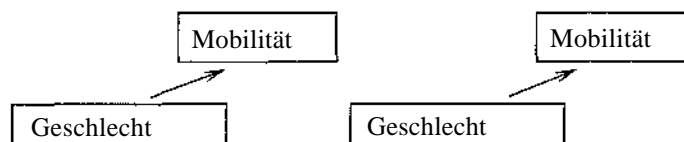
- *"In neue Bundesländer"*



- *"Nach Westdeutschland"*

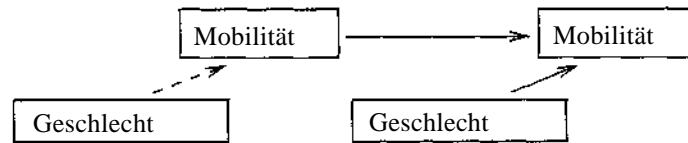


- *"Nach Osteuropa"*

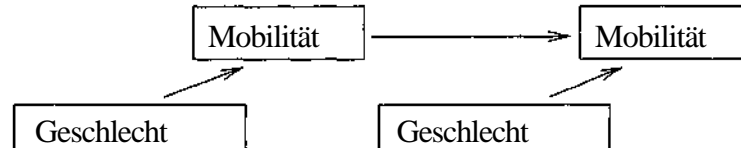


<sup>4</sup>Bei der Abhängigkeit zur Vorperiode wurde das Signifikanzniveau gemäß der Bonferoni-Methode angepaßt. Das bedeutet, weil die Mobilität in 2 Ausprägungen gemessen wurde, ist jeder der zugehörigen Parameter mit  $\alpha/2$  beurteilt worden, also 0,025 und 0,005.

- "Nach Westeuropa"



- "Nach Übersee"



Die Darstellung ergibt, daß die Bereitschaft zu Mobilität in der ersten Periode Einfluß auf die Bereitschaft in der dritten Periode hat. Ausnahmen sind hier die Motivation in das gleiche Bundesland oder nach Westdeutschland zu ziehen, was nur eine schwache Abhängigkeit aufweist, als auch die Bereitschaft in den osteuropäischen Raum auszuwandern. In diesem Fall liegt Unabhängigkeit vor, d.h. die Mobilitätsbereitschaft nach Osteuropa zu ziehen ist eine Entscheidung, die nicht auf der Meinung der Vergangenheit beruht. Das Geschlecht des Jugendlichen spiegelt sich nicht wider in der Fragestellung "in einen Ort in der Nähe" zu ziehen. Ebenso ist die Abhängigkeit bei den Fragen "in gleiches Bundesland" oder "nach Westdeutschland" eher schwach. Anders sieht es bei der Motivation in ein neues Bundesland zu ziehen aus, was Frauen in diesem Fall zum ersten Zeitpunkt klar präferieren. Männer dagegen weisen höhere Mobilitätsbereitschaft auf, wenn es darum geht nach Osteuropa, nach Westeuropa oder nach Übersee zu ziehen. In Tabelle 98 sind die signifikanten geschlechtsspezifischen Ergebnisse nochmals aufgeführt. In Tabellen 99 bis 112 finden sich die entsprechenden Parameter.

	Welle 1	Welle 3
In Nähe		
In gleiches BL	<i>F</i>	
In neue BL	F	<i>M</i>
nach Westdeutschland		<i>M</i>
nach Osteuropa	M	M
nach Westeuropa	<i>M</i>	M
nach Übersee	M	M

Tabelle 98: Signifikante geschlechtsspezifische Unterschiede: M=Männer neigen zu größerer Mobilität, F = Frauen neigen zu größerer Mobilität. Kursiv geschrieben fällt nur bei Signifikanzniveau von 0,05 auf, ansonsten ist  $\alpha = 0,01$

- "In einem Ort in der Nähe "

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
$\theta_1$	0.2258	0.0457	24.4376	0.0001
$\theta_2$	2.2605	0.0769	864.6163	0.0001
männlich	-0.0521	0.0444	1.3771	0.2406

Tabelle 99: Parameterschätzung erste Welle

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
$\theta_1$	-0.1044	0.1064	0.9627	0.3265
$\theta_2$	2.1927	0.1510	210.9091	0.0001
Vorwelle	0.3702	0.1171	9.9983	0.0016
	0.0538	0.1263	0.1816	0.6700
männlich	0.0661	0.0756	0.7639	0.3821

Tabelle 100: Parameterschätzung dritte Welle

- "Ins gleiche Bundesland"

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
$\theta_1$	-0.8432	0.0495 <sup>^</sup>	290.1446	0.0001
$\theta_2$	1.3654	0.0562	589.2963	0.0001
männlich	-0.1095	0.0430	6.4830	0.0109

Tabelle 101: Parameterschätzung erste Welle

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
$\theta_1$	-0.6838	0.0872	61.5631	0.0001
$\theta_2$	1.8053	0.1143	249.6305	0.0001
Vorwelle	0.2618	0.1094	5.7314	0.0167
	0.0135	0.0990	0.0186	0.8915
männlich	-0.0009	0.0748	0.0001	0.9903

Tabelle 102: Parameterschätzung dritte Welle

- "In neue Bundesländer"

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
$\delta_i$	-1.4760	0.0583	641.8447	0.0001-
$\theta_2$	0.4503	0.0467	93.1016	0.0001
männlich	-0.1512	0.0426	12.5863	0.0004

Tabelle 103: Parameterschätzung erste Welle

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
$\delta_i$	-1.3196	0.0966	186.5388	0.0001
$\theta_2$	1.0135	0.0900	126.8200	0.0001
Vorwelle	0.3720	0.1175	10.0306	0.0015
	0.0930	0.1017	0.8352	0.3608
männlich	0.1495	0.0745	4.0249	0.0448

Tabelle 104: Parameterschätzung dritte Welle

- "nach Westdeutschland"

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
$\delta_i$	-1.5109	0.0588	660.9407	0.0001
$\theta_2$	0.2285	0.0457	24.9686	0.0001
männlich	0.0380	0.0427	0.7954	0.3725

Tabelle 105: Parameterschätzung erste Welle

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
$\delta_i$	-1.7609	0.1122	246.3161	0.0001
$\theta_2$	0.2973	0.0840	12.5222	0.0004
Vorwelle	0.2432	0.1226	3.9389	0.0472
	0.2668	0.1020	6.8363	0.0089
männlich	0.1473	0.0744	3.9213	0.0477

Tabelle 106: Parameterschätzung dritte Welle

- "nach Osteuropa"

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
$d_i$	-3.4385	0.1284	716.6039	0.0001
$\beta_2$	-1.6853	0.0635	703.9782	0.0001
männlich	0.1790	0.0634	7.9799	0.0047

Tabelle 107: Parameterschätzung erste Welle

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
$\theta_1$	-3.8273	0.3377	128.4814	0.0001
$\theta_2$	-1.2684	0.1842	47.4426	0.0001
Vorwelle	0.5484	0.3115	3.1000	0.0783
	-0.0050	0.2344	0.0005	0.9828
männlich	0.3351	0.1125	8.8697	0.0029

Tabelle 108: Parameterschätzung dritte Welle

- "nach Westeuropa"

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
$d_i$	-1.5107	0.0588	660.6466	0.0001
$\theta_2$	0.1546	0.0457	11.4686	0.0007
männlich	0.0956	0.0428	4.9776	0.0257

Tabelle 109: Parameterschätzung erste Welle

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
$\theta_i$	-2.2313	0.1314	288.2300	0.0001
$\theta_2$	-0.1828	0.0840	4.7351	0.0296
Vorwelle	0.4119	0.1223	11.3433	0.0008
	0.1624	0.1055	2.3729	0.1235
männlich	0.2834	0.0786	12.9848	0.0003

Tabelle 110: Parameterschätzung dritte Welle

- "Nach Übersee"

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
$\beta_1$	-1.2374	0.0542	520.8085	0.0001
$\beta_2$	0.1304	0.0456	8.1639	0.0043
männlich	0.1574	0.0427	13.5903	0.0002

Tabelle 111: Parameterschätzung erste Welle

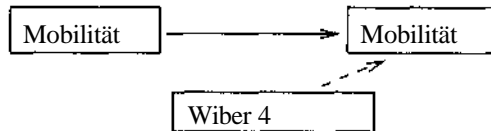
Parameter	Schätzung	Std.err.	t-Wert	p-Wert
$\theta_1$	-2.1857 ^	0.1275	293.6917	0.0001
$\theta_2$	-0.4365	0.0852	26.2368	0.0001
Vorwelle	0.3941	0.1162	11.4984	0.0007
	0.0407	0.1114	0.1338	0.7145
männlich	0.4340	0.0824	27.7563	0.0001

Tabelle 112: Parameterschätzung dritte Welle

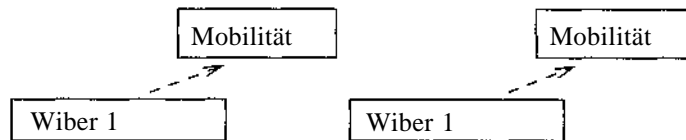


Abschließend soll nun untersucht werden, welchen Einfluß der Wirtschaftsbereich hat. Dazu betrachte man nochmals die Kodierung des Wirtschaftsbereichs und seine Einteilung in Wiber 1-4 (siehe Seite 114). In analoger Form wie oben modellieren wir nun den Wirtschaftsbereich als Einflußgröße anstelle des Geschlechts. Eine Beschränkung auf signifikante Komponenten liefern die folgenden graphischen Darstellungen.

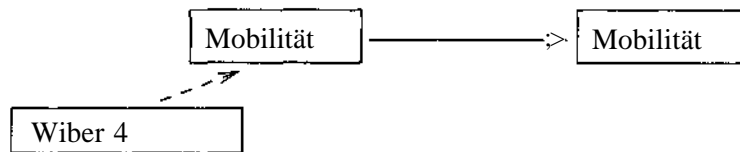
- *"In einen Ort in der Nähe "*



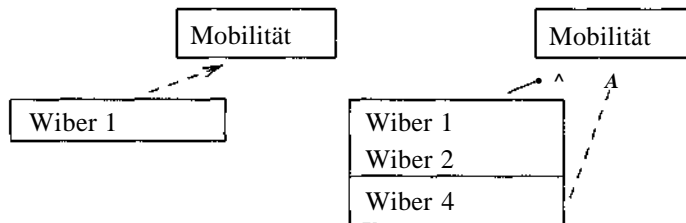
- *"Im gleichen Bundesland"*



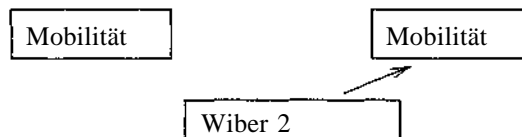
- *"In neue Bundesländer"*



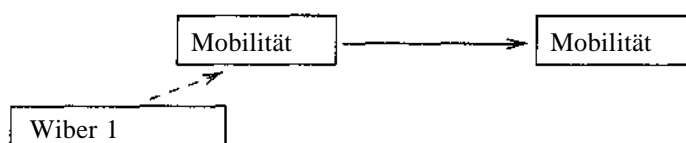
- *"Nach Westdeutschland"*



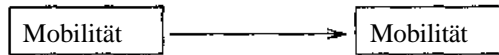
- *"Nach Osteuropa"*



- *"Nach Westeuropa"*



- "Nach Übersee"



Es zeigt sich, daß Jugendliche, die in Landwirtschaft und Handwerk eine Lehrstelle gefunden haben, sich eher immobil geben. Der Einfluß fällt generell auf, zeigt sich aber besonders in der Ablehnung einer Migration nach Westdeutschland bzw. ins Westeuropäische Ausland. Im Unterschied dazu fällt auf, daß Jugendliche mit Arbeitsplatz in der Industrie sich nach der Lehre (Welle 3) einen Umzug nach Westdeutschland oder ins Osteuropäische Ausland klar vorstellen können. Weitere signifikante Einflüsse zeigen sich für Jugendliche, die im öffentlichen Dienst ihre Lehre absolvieren. Diese Jugendlichen weisen in Periode 3 wenig Motivation auf nach Westdeutschland zu ziehen und präferieren zum Beginn ihrer Lehre einen Umzug in die neuen Bundesländer. Sie sind ferner nicht bereit in einen Ort in der Nähe zu ziehen.

Weiter zeigt sich, daß die Bereitschaft in der dritten Welle nach Westdeutschland zu ziehen nur vom Wirtschaftsbereich abhängt, nicht jedoch von der Vorperiode. Gleiches gilt für Osteuropa oder das gleiche Bundesland. Eine Zusammenfassung der Ergebnisse liefert die Tabelle 113, die entsprechenden Parameterschätzungen sind in den Tabellen 114 bis 127 gegeben.

	Welle 1	Welle 3
In Nähe		<i>öffentl. Dienstr</i>
In gleiches BL	<i>Landw. + Handw."</i>	<i>Landw. + Handw."</i>
In neue BL	<i>öffentl. Dienst<sup>+</sup></i>	
nach Westdeutschland	<i>Landw. + Handw."</i>	<i>öffentl. Dienst~, Industrie<sup>+</sup>, Landw. + Handw."</i>
nach Osteuropa		Industrie <sup>111</sup>
nach Westeuropa	<i>Landw. + Handw."</i>	
nach Übersee		

Tabelle 113: Signifikante Einflüsse des Wirtschaftsbereichs auf die Mobilitätsbereitschaft. Kursiv geschrieben fällt nur bei einem Signifikanzniveau von  $\alpha = 0,05$  auf, ansonsten ist  $\alpha = 0,01$ , Zeichen geben an, ob Wirtschaftsbereich förderlich (+) oder ablehnend (—) für die Mobilitätsbereitschaft wirkt

Abschließend sind nachfolgend die entsprechenden Schätzer aufgelistet.

- *"In eine Ort in der Nähe "*

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
$\delta_i$	0.2699	0.0573	22.1941	0.0001
$\delta_2$	2.2426	0.0881	647.8709	0.0001
WIBER1	-0.1171	0.0785	2.2291	0.1354
WIBER2	0.0701	0.1211	0.3346	0.5630
WIBER3	0.1319	0.1239	1.1332	0.2871
WIBER4	-0.0460	0.1081	0.1813	0.6703

Tabelle 114: Parameterschätzung erste Welle

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
$\delta_i$	-0.1823	0.1205	2.2864	0.1305
$\delta_2$	2.1260	0.1624	171.3130	0.0001
Vorwelle	0.4726	0.1261	14.0546	0.0002
	0.0949	0.1363	0.4849	0.4862
WIBER1	-0.0415	0.1301	0.1015	0.7501
WIBER2	0.2974	0.1999	2.2132	0.1368
WIBER3	-0.0112	0.1931	0.0034	0.9535
WIBER4	-0.4374	0.1781	6.0297	0.0141

Tabelle 115: Parameterschätzung dritte Welle

- *"In gleiches Bundesland"*

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
$\delta_i$	-0.8169	0.0602	183.9931	0.0001
$\delta_2$	1.3337	0.0668	398.5373	0.0001
WIBER1	-0.1531	0.0758	4.0818	0.0433
WIBER2	-0.1454	0.1154	1.5890	0.2075
WIBER3	0.0747	0.1179	0.4017	0.5262
WIBER4	0.1640	0.1046	2.4595	0.1168

Tabelle 116: Parameterschätzung erste Welle

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
öi	-0.6732	0.1011	44.3676	0.0001
$\theta_2$	1.8501	0.1295	203.9545	0.0001
Vorwelle	0.1959	0.1186	2.7276	0.0986
	0.0235	0.1073	0.0480	0.8265
WIBER1	-0.2723	0.1297	4.4102	0.0357
WIBER2	0.2353	0.1919	1.5029	0.2202
WIBER3	0.1494	0.1897	0.6202	0.4310
WIBER4	-0.2169	0.1789	1.4688	0.2255

Tabelle 117: Parameterschätzung dritte Welle

- "In neue Bundesländer"

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
öi	-1.5391	0.0707	473.6541	0.0001
$\theta_2$	0.3772	0.0572	43.4441	0.0001
WIBER1	-0.1436	0.0757	3.5943	0.0580
WIBER2	-0.0408	0.1150	0.1257	0.7229
WIBER3	-0.1420	0.1179	1.4504	0.2285
WIBER4	0.2086	0.1032	4.0820	0.0433

Tabelle 118: Parameterschätzung erste Welle

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
öi	-1.2932	0.1113	135.0273	0.0001
o2	1.1087	0.1071	107.2355	0.0001
Vorwelle	0.2497	0.1307	3.6505	0.0561
	0.1184	0.1124	1.1102	0.2920
WIBER1	-0.2181	0.1278	2.9124	0.0879
WIBER2	0.2875	0.1913	2.2575	0.1330
WIBER3	-0.2147	0.1899	1.2782	0.2582
WIBER4	-0.0658	0.1770	0.1384	0.7098

Tabelle 119: Parameterschätzung dritte Welle<sup>6</sup>

- "nach Westdeutschland"

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
$\theta_1$	-1.5430	0.0709	473.5409	0.0001
$\theta_2$	0.1584	0.0566	7.8426	0.0051
WIBER1	-0.1513	0.0761	3.9567	0.0467
WIBER2	-0.0764	0.1159	0.4350	0.5095
WIBER3	0.1380	0.1171	1.3885	0.2387
WIBER4	0.0174	0.1040	0.0279	0.8672

Tabelle 120: Parameterschätzung erste Welle

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
$\theta_1$	-1.6551	0.1259	172.7069	0.0001
$\theta_2$	0.4716	0.1006	21.9840	0.0001
Vorwelle	0.2611	0.1351	3.7367	0.0532
	0.2078	0.1126	3.4058	0.0650
WIBER1	-0.4931	0.1296	14.4781	0.0001
WIBER2	0.5622	0.1891	8.8357	0.0030
WIBER3	0.1813	0.1879	0.9302	0.3348
WIBER4	-0.4096	0.1794	5.2121	0.0224

Tabelle 121: Parameterschätzung dritte Welle

- "nach Osteuropa"

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
$\theta_1$	-3.4252	0.1492	526.7265	0.0001
$\theta_2$	-1.6982	0.0784	469.1879	0.0001
WIBER1	-0.0808	0.1123	0.5176	0.4719
WIBER2	0.0771	0.1657	0.2166	0.6416
WIBER3	0.0359	0.1705	0.0444	0.8331
WIBER4	-0.2339	0.1623	2.0786	0.1494

Tabelle 122: Parameterschätzung erste Welle

---

<sup>6</sup>Die Vorwelle liefert einen signifikanten Einfluß, wenn die nicht signifikanten Größen Wiber 1–4 aus dem Modell genommen werden.

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
öi	-3.7941	0.3662	107.3336	0.0001
$\theta_2$	-1.3598	0.2164	39.4805	0.0001
Vorwelle	0.7038	0.3542	3.9470	0.0470
	-0.3097	0.2905	1.1369	0.2863
WIBER1	0.0241	0.1867	0.0166	0.8975
WIBER2	0.8547	0.2366	13.0470	0.0003
WIBER3	-0.0749	0.2796	0.0717	0.7889
WIBER4	-0.5038	0.3003	2.8147	0.0934

Tabelle 123: Parameterschätzung dritte Welle

- "nach Westeuropa"

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
öi	-1.5089	0.0703	460.6004	0.0001
$\theta_2$	0.1373	0.0565	5.9046	0.0151
WIBER1	-0.1868	0.0762	6.0003	0.0143
WIBER2	0.0151	0.1153	0.0171	0.8959
WIBER3	0.0900	0.1170	0.5919	0.4417
WIBER4	-0.0948	0.1044	0.8251	0.3637

Tabelle 124: Parameterschätzung erste Welle

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
öi	-2.2147	0.1495	219.5220	0.0001
$\theta_2$	-0.1644	0.0971	2.8683	0.0903
Vorwelle	0.3526	0.1349	6.8335	0.0089
	0.1799	0.1161	2.4011	0.1213
WIBER1	-0.1176	0.1339	0.7706	0.3800
WIBER2	0.1430	0.1952	0.5369	0.4637
WIBER3	-0.0091	0.1956	0.0021	0.9631
WIBER4	-0.2392	0.1879	1.6210	0.2030

Tabelle 125: Parameterschätzung dritte Welle

- "nach Übersee"

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
$\theta_1$	-1.3368	0.0667	401.1083	0.0001
$\theta_2$	0.0607	0.0563	1.1601	0.2815
WIBER1	0.0430	0.0755	0.3240	0.5692
WIBER2	-0.0132	0.1156	0.0130	0.9093
WIBER3	0.0306	0.1171	0.0681	0.7942
WIBER4	-0.1848	0.1053	3.0820	0.0792

Tabelle 126: Parameterschätzung erste Welle

Parameter	Schätzung	Std.err.	t-Wert	p-Wert
$\theta_i$	-2.1508	0.1446	221.2188	0.0001
$\theta_2$	-0.3573	0.0980	13.2893	0.0003
Vorwelle	0.3951	0.1282	9.5016	0.0021
	0.0842	0.1198	0.4935	0.4824
WIBER1	0.0709	0.1345	0.2779	0.5981
WIBER2	0.2551	0.1972	1.6733	0.1958
WIBER3	-0.0913	0.2021	0.2041	0.6514
WIBER4	-0.2874	0.1952	2.1684	0.1409

Tabelle 127: Parameterschätzung dritte Welle

## Literatur

- Agresti, A. (1990). *Categorical Data Analysis*. New York: Wiley.
- Bortz, J. (1989). *Statistik für Sozialwissenschaftler* (3. Aufl.). Berlin: Springer Verlag.
- Fahrmeir, L., Hamerle, A., und Tutz, G. (1996). *Multivariate statistische Verfahren* (2. Aufl.). Berlin, New York: de Gruyter.
- Glonek, G. F. V. und McCullagh, P. (1995). Multivariate logistic models. *J. R. Statist. Soc.* 57, 533-546.
- Lauritzen, S. L. und Wermuth, N. (1989). Graphical models for associations between variables, some of which are qualitative and some quantitative. *Ann. Statist.* 17, 31-57. (Corr: V17p1916).
- McCullagh, P. und Neider, J. A. (1989). *Generalized Linear Models* (2. Aufl.). New York: Chapman and Hall.
- Schlittgen, R. (1996). *Statistische Inferenz*. München: R. Oldenbourg.