



Diskussionspapiere
Discussion Papers

Diskussionspapier Nr. 54

**Berufseinmündung von Jugendlichen
als mehrstufiges Entscheidungsproblem**

von
Christof/Heldberger, Ulrich Rendtel und Johannes Schwarze*

Die in diesem Papier vertretenen Auffassungen liegen ausschließlich in der Verantwortung des Verfassers und nicht in der des Instituts.

Opinions expressed in this paper are those of the author and do not necessarily reflect views of the Institute.

Deutsches Institut für Wirtschaftsforschung

Diskussionspapier Nr. 54

Berufseinmündung von Jugendlichen als mehrstufiges Entscheidungsproblem

von
Christof/Heldberger, Ulrich Rendtel und Johannes Schwarze*

* Helberger: Technische Universität Berlin

Rendtel und Schwarze: Deutsches Institut für Wirtschaftsforschung

Berlin, im Juli 1992

Deutsches Institut für Wirtschaftsforschung, Berlin
Königin-Luise-Str. 5, 1000 Berlin 33
Telefon: 49-30 - 82 991-0
Telefax: 49-30 - 82 991-200

Die Berufseinmündung von Jugendlichen als mehrstufiges Entscheidungsproblem

Christof Helberger

Technische Universität Berlin

Uhlandstr. 4-5

W-1000 Berlin 12

Tel. 030/314-23235 (Fax -21109)

Ulrich Rendtel, Johannes Schwarze

Deutsches Institut für Wirtschaftsforschung

Königin-Luise-Str. 5

W-1000 Berlin 33

Tel. 030/ 829 91-1 (Fax -200)

Beitrag zum Workshop

“Arbeitslosigkeits-, Beschäftigungs- und Lohndynamik
auf der Basis von Mikrodaten”

am 27./28. Februar 1992 in Mannheim

Inhaltsverzeichnis

1	Einleitung	2
2	Was tun nach der Schule? - Der Weg in die Lehre	3
3	Der Übergang von der Berufsausbildung in die Erwerbstätigkeit	5
4	Ein Modell zur Analyse des Übergangs in betriebliche Berufsausbildung und in Erwerbstätigkeit	7
5	Sequentielle Probit-Modelle mit unbeobachteter Heterogenität	9
5.1	Die Modellierung der Schwellen Berufsausbildung und Erwerbstätigkeit	9
5.2	Die Schätzung des Erwerbseinkommens unter einem bivariaten Selektionsmodell	13
6	Datenbasis und Operationalisierung des Modells	16
6.1	Das Sozio-ökonomische Panel	16
6.2	Der Übergang an der ersten und an der zweiten Schwelle . . .	17
6.3	Erklärende Variable	19
7	Ergebnisse	20
7.1	Die Übergänge an der ersten und der zweiten Schwelle	20
7.2	Das Arbeitseinkommen bei Berufsbeginn	24
8	Anhang, Variablenübersicht	34

1 Einleitung

In diesem Beitrag soll die Berufseinmündung von Jugendlichen analysiert werden, d.h. der Weg von der Schule über die Berufsausbildung zur anschließenden Erwerbstätigkeit und dem erzielten Einkommen. Die folgenden Ausführungen konzentrieren sich dabei ganz auf eine Ausbildungsart: die betriebliche Berufsausbildung (Lehre).

Das Interesse für diese Fragestellung ist begründet in

- dem Umstand, daß dieser Bildungstyp nach wie vor die bei weitem häufigste Form der Berufsausbildung ist;
- gleichzeitig hat sie herausragende arbeitsmarktpolitische Eigenschaften hinsichtlich Qualität, Flexibilität und Kosten der beruflichen Bildung, des Niveaus der Arbeitslosigkeit, insbesondere der Jugendarbeitslosigkeit, und
- sie ist eine Besonderheit der deutschsprachigen Länder und daher in erster Linie hier zu analysieren, wobei ein stark zunehmendes Interesse für die Ausbildungsform "Lehre" im Ausland festgestellt werden kann.

Diese Gesichtspunkte in Verbindung mit manifesten Problemen auf dem Arbeitsmarkt - demographische Entwicklung, Lehrstellenknappheit, hohe Arbeitslosigkeit, Wandel der Bildungswege von Jugendlichen u. a. - haben dazu geführt, daß die Form der Einmündung von Jugendlichen in die betriebliche Berufsausbildung und in die Erwerbstätigkeit in der letzten Zeit wiederholt zum Gegenstand von Untersuchungen gemacht worden ist.

Im folgenden soll zuerst ein Überblick über die Struktur der Berufseinmündung von Jugendlichen in Westdeutschland gegeben werden, und es sollen die Untersuchungsansätze und -ergebnisse der jüngsten vorliegenden Publikationen erläutert werden. Darauf soll ein weiterentwickeltes Modell zur Analyse der Berufseinmündung vorgestellt werden, und es sollen seine empirische Ergebnisse aufgezeigt werden.

2 Was tun nach der Schule? - Der Weg in die Lehre

Etwa 2 von 3 Jugendlichen eines Altersjahrgangs besuchen die Haupt- oder Realschule. In welche Wege diese Jugendlichen nach Verlassen der Schule einmünden, läßt sich gut aus der Bildungsgesamtrechnung (BGR) des IAB ersehen. Bei diesem Rechenwerk handelt es sich um eine synthetische Statistik, die eine bestmögliche Ausnutzung der verschiedenen, z. T. unvollständigen und inkonsistenten, Spezialstatistiken anstrebt (vgl. Tessaring u. a. 1990, Blien, Reinberg, Tessaring 1990). Die BGR liegt derzeit für die Jahre 1975, 1980 und 1985 vor.

Tabelle 1: Übergänge der Jugendlichen nach Verlassen der Haupt- und Realschule - in % der Gesamtabgänge

in:	aus Hauptschulen			aus Realschulen		
	1975	1980	1985	1975	1980	1985
Berufsgrund-/vorber.- jahr, ber. vorber. Maßn.	7.8	17.0	18.1	2.8	3.9	5.7
betriebl. Berufsausbildung	59.8	52.0	47.1	47.8	51.1	53.0
Berufsfach u. andere Fachschulen	19.3	18.3	18.2	18.8	21.8	25.5
Fachoberschulen und Gymnasien	0.0	0.0	0.0	13.2	12.8	7.1
Fachhoch- und Hochschulen	0.0	0.0	0.0	0.0	0.0	0.0
Arbeitslosigkeit	2.2	2.7	11.1	1.4	1.2	3.3
Wehrpflicht/Zivildienst	0.0	0.1	0.1	0.4	0.3	0.3
Erwerbstätigkeit	5.9	4.4	1.8	11.2	7.3	3.7
Nichterwerbstätigkeit/ Abwanderung, Tod	5.0	5.5	3.6	4.4	1.7	1.3
Insgesamt	100.0	100.0	100.0	100.0	100.0	100.0

Quelle: Blien/Reinberg/Tessaring 1990, S.195

Tabelle 1 zeigt, daß etwa die Hälfte der Schulabgänger der Haupt- und Realschule eine Lehre aufnehmen, Hauptschüler mit sinkender, Realschüler

mit steigender Tendenz. Die Schüler, die nicht in die Lehre wechseln, verteilen sich auf zahlreiche Wege. Der direkte Übergang in Erwerbstätigkeit ist dabei - in Übereinstimmung mit bildungspolitischen Zielen und im Interesse der langfristigen Berufschancen der Jugendlichen - selten und stark sinkend.

Die Wege der Schulabgänger - von Lehre abgesehen - lassen sich zwei Hauptgruppen zuordnen: der Besuch einer weiterführenden schulischen Ausbildung und die Gesamtheit aller anderen Alternativen, die ganz überwiegend ergriffen werden, weil ein passender betrieblicher Ausbildungsplatz nicht gefunden wurde oder weil sich die Jugendlichen noch nicht für die endgültige Ausbildungsform entscheiden konnten. Hierfür hat sich die Bezeichnung "Warteschleife" durchgesetzt. Als Warteschleifen kann man vorübergehende Nichterwerbstätigkeit, Arbeitslosigkeit, das Berufsgrundbildungsjahr und berufsvorbereitende Maßnahmen sowie z. T. auch den Besuch einer Berufsfachschule bezeichnen.

Der Anteil der Jugendlichen, die nach Abschluß aller Warteschleifen die Lehre aufnehmen, ist daher deutlich höher als jene, die gleich nach der Schule dorthin gehen. Brandes, Brosi und Menk (1986) haben in Wiederholungsbefragungen des Schulabgängerjahrgangs 1977/78 ermittelt, daß die Gesamtzahl der Jugendlichen (dieses Jahrgangs), die letztendlich eine Lehre aufnehmen, etwa doppelt so hoch war wie die Zahl derer, die unmittelbar nach Schulabschluß in die Lehre übergehen. Aus der BIBB-Untersuchung ergab sich: "Insgesamt haben ca. 70 % des Schulabgängerjahrgangs 1983 eine Ausbildung im dualen System aufgenommen. 36 % sind als 'Direkteinsteiger' anzusehen, 23 % hatten eine oder mehrere schulische 'Schleifen' absolviert, 11 % waren nach außerschulischen Schleifen in eine betriebliche Berufsausbildung eingemündet." (Palamidis, Schwarze, S. 116)

Man kann nun fragen, welche Jugendlichen es sind, die einen unverzögerten und erfolgreichen Übergang in die betriebliche Berufsausbildung vollziehen. Wovon hängt dieser Übergang ab? Wie stark wirken sich Geschlecht und Nationalität aus? Liegt Diskriminierung vor? Welchen Einfluß haben Faktoren wie schulische Leistungen, soziale Herkunft sowie die aktuelle Arbeitsmarktsituation?

Zu diesen Fragen liegt inzwischen eine Anzahl von Untersuchungen vor. Besonders umfassend und differenziert sind die Auswertungen der Lehrlingsbefragung des Bundesinstituts für Berufsbildung (BIBB) von 1983 und 1984 (Brandes, Brosi und Menk 1986) und die Verbleibsuntersuchung des IAB (vgl. Stegmann und Kraft 1987 sowie die dort zitierten Arbeiten). Die Ergebnisse

dieser Untersuchungen werden bei Palamidis und Schwarze (1989) übersichtlich dargestellt. Die Untersuchungen zeigen, daß von den Jugendlichen aus Haupt- und Realschulen Ausländer und Mädchen geringere Chancen haben, eine Ausbildungsstelle zu erhalten (im Gegensatz zu den höhergesteckten Ausbildungswünschen). Nachteiligt wirkt sich auch eine niedrigere berufliche Stellung der Eltern aus. Jugendliche mit guten Schulnoten tendieren stärker zu schulischen Formen der Berufsausbildung.

Von Interesse wäre auch eine Unterscheidung nach der Qualität der beruflichen Ausbildungsstellen, da das Niveau der Vermittlung von Fähigkeiten sehr verschieden ist und auch die beruflichen Verwertungschancen im Anschluß an die Ausbildung stark differieren. Diese Effekte lassen sich allerdings bei der "zweiten Schwelle", d. h. beim Übergang von der Lehre in die Erwerbstätigkeit, direkt beobachten.

3 Der Übergang von der Berufsausbildung in die Erwerbstätigkeit

Für den Übergang von der betrieblichen Berufsausbildung in die Erwerbstätigkeit ergeben sich aus der Bildungsgesamtrechnung die in Tabelle 2 wiedergegebenen Strukturen.

Ca. 70 % gehen direkt in Erwerbstätigkeit über. Etwa 10 % nehmen eine der verschiedenen Formen weiterführender schulischer Ausbildung auf. Die übrigen verteilen sich auf Arbeitslosigkeit, Wehr-/Zivildienst, Nichterwerbstätigkeit und sonstige Abwanderung. Arbeitslosigkeit stieg bis 1985 an. Der relativ starke Anstieg des Übergangs in Fachhoch- und Hochschulen hängt damit zusammen, daß in diesen Zahlen auch die Abiturienten enthalten sind, die anschließend eine Lehre aufgenommen haben - ein Weg mit steil ansteigenden Teilnehmerzahlen.

Als Problem für die Jugendlichen wie auch für die Arbeitsmarktpolitik stellt sich insbesondere der Übergang in Arbeitslosigkeit dar, während der Übergang in weiterführende schulische Bildungswege meist den Präferenzen der Jugendlichen entspricht. Andere Untersuchungen ergeben eine deutlich höhere Betroffenheit von Arbeitslosigkeit nach Abschluß der Lehre. Die Differenz zwischen solchen Befunden ist aus den verwendeten Befragungskonzepten erklärbar. Die Ergebnisse hängen stark davon ab, in welchem zeitlichen

Tabelle 2 Übergänge aus beruflicher Ausbildung (Lehre)
- in % der Gesamtabgänge -

in:	1975	1980	1985
Berufsgrund-/vorber.- jahr, ber.vorber.Maßn.	0,0	0,0	0,0
Fachschulen	5,0	3,7	4,1
Fachoberschulen und Gymnasien	5,7	4,4	3,7
Fachhoch- und Hochschulen	0,5	1,2	2,4
Arbeitslosigkeit	3,3	2,9	7,9
Wehrpflicht/ Zivildienst	4,7	4,2	4,0
Erwerbstätigkeit	73,6	76,4	69,6
Nichterwerbstätigkeit/ Abwanderung, Tod	7,2	7,2	8,3
Insgesamt	100,0	100,0	100,0

Quelle: Blien/Reinberg/Tessaring 1990, S.195

Abstand nach dem Lehrabschluß die Arbeitslosigkeit erfragt wird oder ob gar danach gefragt wird, ob *irgendwann* zwischen Abschluß und Befragung eine Arbeitslosigkeit vorlag (Weißhuhn und Rohwer 1991).

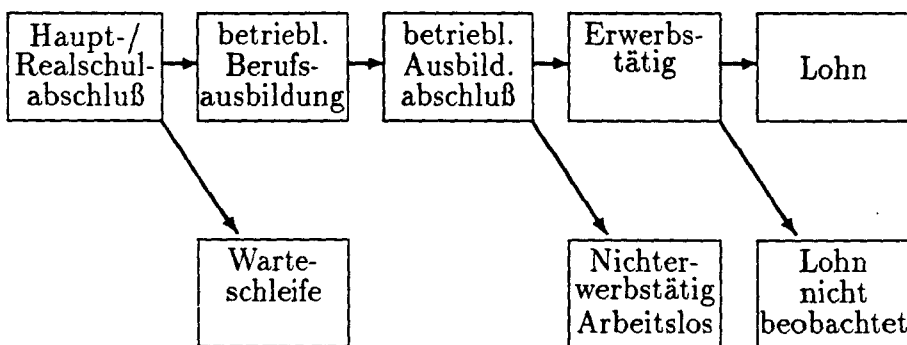
Das Arbeitslosigkeitsrisiko ist nach den bisherigen Ergebnissen kaum geschlechtsspezifisch, aber Ausländer, Hauptschüler und Personen, die bereits beim Übergang in die Lehre Schwierigkeiten hatten, sind stärker betroffen (Palamidis und Schwarze 1989, S. 117).

Von Interesse ist allerdings nicht nur, ob eine Einmündung in die Erwerbstätigkeit erfolgte, sondern auch, wie gut sie gelang. Einer der zentralen Indikatoren ist hierbei die Höhe des erzielten Einkommens.

4 Ein Modell zur Analyse des Übergangs in betriebliche Berufsausbildung und in Erwerbstätigkeit

Wie gezeigt wurde, ist der am häufigsten beschrittene und auch am meisten präferierte Weg der Jugendlichen, welche die Haupt- und Realschulen verlassen, eine betriebliche Lehrausbildung und daran anschließend ein Arbeitsplatz als Erwerbstätiger. Dieser geplante Ablauf gelingt nur einem Teil der Schulentlassenen. Ein beträchtlicher Teil der Jugendlichen hat Schwierigkeiten an einer dieser beiden "Schwellen" zum Berufsleben. Bei der ersten Schwelle kommt dies in eingelegten Warteschleifen zum Ausdruck, bei der zweiten in dem Risiko, keinen Arbeitsplatz zu erhalten.

Diese Abläufe können wie folgt dargestellt werden:



Zur Analyse dieser Verläufe kann ein sequentielles Entscheidungsmodell mit jeweils zwei wählbaren Ausprägungen auf jeder Entscheidungsstufe herangezogen werden. Die Bildungsverläufe können - anhand von Zeitreihen - auf Makroebene oder - anhand von Individualdaten - auf Mikroebene betrachtet werden.

Für die folgenden Untersuchungen wird letzteres gewählt. Gegenüber den

vielfältigen realen Abläufen ist die obige Darstellung eine deutliche Vereinfachung:

- Es gibt Jugendliche, welche die Schulausbildung auf einer anderen allgemeinbildenden oder berufsbildenden Schule fortsetzen. Diesen Weg kann man als eine vorgelagerte (weitere) Entscheidungsstufe interpretieren (Fortsetzung der Schulausbildung versus Schulabgang).
- Nicht alle Jugendlichen, welche nach Schulabgang erwerbstätig oder nichterwerbstätig sind, warten auf einen betrieblichen Ausbildungsplatz (d. h. befinden sich in der Warteschleife).
- Ein Teil der Lehrlinge bricht die Ausbildung ab.
- Aus der Warteschleife erfolgen zeitverzögerte Übergänge in die betriebliche Berufsausbildung.
- Nach Abschluß der betrieblichen Berufsausbildung gibt es als weitere Alternativen Wehr- oder Zivildienst sowie den Besuch einer weiterführenden Schule oder Hochschule.

Die binäre Struktur des sequentiellen Entscheidungsmodells macht es erforderlich, die zusätzlichen Alternativen aus der Analyse auszuschneiden. Es werden daher nur die in Abbildung 1 dargestellten Übergänge in das Modell aufgenommen.

Auf allen Entscheidungsstufen sollen individuelle und überindividuelle - betriebliche, regionale, gesamtwirtschaftliche - Merkmale zur Erklärung herangezogen werden. Da ein Teil der relevanten Merkmale nicht beobachtet werden kann, liegt unbeobachtete Heterogenität vor. Die gemeinsame Analyse beider Stufen bietet die Möglichkeit, unbeobachtete Heterogenität - soweit sie individuell, jedoch nicht zeitlich variabel ist - in der empirischen Schätzung des Modells zu berücksichtigen. Als zusätzliches Problem ergibt sich dabei, daß für einen Teil der Jugendlichen nicht für beide Entscheidungen Beobachtungen vorliegen. Dies macht es erforderlich, unvollständige Beobachtungen in der Schätzung zu berücksichtigen. Einerseits sind dies Personen, bei denen nur die erste Stufe beobachtet wurde, meist, weil ihre Biographie in dem verwendeten Datensatz noch nicht lange genug beobachtet worden ist -, andererseits Personen, bei denen nur die Entscheidung auf der zweiten Stufe bekannt ist. Unvollständige Beobachtungen treten auch auf der letzten Stufe, der Einkommenserzielung, auf.

5 Sequentielle Probit-Modelle mit unbeobachteter Heterogenität

Ein sequentielles Probit-Modell beschreibt eine Situation, in der unterschiedlich lange Folgen von 0-1 Zufallsgrößen auftreten. Eine Folge endet, sobald eine Beobachtung mit 0 auftritt. Wird hingegen eine 1 beobachtet, so wird eine weitere 0-1 Zufallsgröße realisiert. Typische Situationen, die durch sequentielle Probit-Modelle beschrieben werden können, haben folgende Struktur. Es wird die Zeit betrachtet, während der ein interessierender Zustand anhält. An diskreten Punkten wird festgestellt, ob dieser Zielzustand noch besteht oder beendet worden ist. Damit erhält man für jeden Zeitpunkt t eine 0-1 Variable Y_t , die Auskunft über das Andauern des Zielzustandes gibt.¹ Derartige Modelle wurden von Allison (1987) vorgestellt.

5.1 Die Modellierung der Schwellen Berufsausbildung und Erwerbstätigkeit

Im vorliegenden Fall ist das interessierende Ereignis das Verbleiben auf einem definierten Ausbildungsweg. Dieser Bildungsweg ist durch einen direkten Übergang von der Haupt- bzw. der Realschule in die berufliche Bildung und nach der beruflichen Bildung in die Erwerbstätigkeit gekennzeichnet. Hierbei bezieht sich die Definition vom "direktem Übergang" auf einen Zeitraum von ca. einem Jahr nach dem jeweiligen Abschluß. Für jeden Bildungsweg existieren daher maximal zwei 0-1 Zufallsgrößen. Für $Y_1 = 0$ wird der vergebene Bildungsweg über die Warteschleife und für $Y_2 = 0$ über die Nichterwerbstätigkeit nach der Berufsausbildung verlassen. Es sei an dieser Stelle vermerkt, daß über diese zwei 0-1 Variablen nicht alle möglichen Bildungsverläufe abgebildet werden. Beispielsweise besteht die Möglichkeit, auch nach einem längeren Zeitintervall von einem Jahr, von der Haupt- Realschule in eine Berufsausbildung zu wechseln und einen Lehrabschluß zu erwerben. Dieses Ereignis ist in der Stichprobe doch relativ selten. (12 von 158 Personen = 7,5%).

¹Auf den Personenindex i wird im folgenden verzichtet.

Je nach Beobachtung liefern Personen unterschiedliche Beiträge zur Likelihoodfunktion:

<i>Beobachtung</i>	<i>Likelihoodbeitrag</i>	
$Y_1 = 0$	$P(Y_1 = 0)$	(1)
$Y_1 = 1, Y_2 = y_2$	$P(Y_1 = 1)P(Y_2 = y_2 Y_1 = 1)$	(2)

Für $P(Y_1 = y_1)$ bzw. $P(Y_2 = y_2 | Y_1 = 1)$ wird ein Schwellenwertmodell benutzt. In diesem Modell bezeichnet Y_1^* die Neigung einer Person, die Entscheidung $Y_1 = 1$ zu treffen und Y_2^* die Neigung, die Entscheidung $Y_2 = 1$ zu treffen, wenn die Entscheidung $Y_1 = 1$ getroffen worden ist. Skaliert man die Entscheidungsschwelle auf den Wert 0, so gilt:

$$P(Y_1 = 1) = P(Y_1^* > 0) \quad (3)$$

$$P(Y_2 = 1 | Y_1 = 1) = P(Y_2^* > 0) \quad (4)$$

Die Neigungen Y_1^* und Y_2^* lassen sich über einen linearen Ansatz parametrisieren. Weiterhin bietet sich die Möglichkeit aufgrund von Personen mit 2 Beobachtungen eine Aufspaltung in eine zeitunabhängige Komponente α - die unbeobachtete Heterogenität - und in einen hiervon unabhängigen Rest u_t zu identifizieren. Hieraus resultiert:

$$Y_t^* = X_t \beta_t + \alpha + u_t \quad (5)$$

mit $t = 1, 2$

Hierbei sind die Störterme α und u_t unabhängig von den Kovariaten X_t . Die Varianzkomponenten α , u_1 und u_2 sind wiederum paarweise unabhängig. Dieses Modell ist ein Analogon zum Varianzkomponenten-Modell im metrischen linearen Fall, vgl. Hsiao (1986, S. 32 ff). Unter der Annahme der Normalverteilung für die $u_t (t = 1, 2)$ und α , erhält man aus (3) und (4) :

$$P(Y_1 = 0 | X_1) = \int \Phi \left[- \left(X_1' \frac{\beta_1}{\sigma_1} + \frac{\alpha}{\sigma_1} \right) \right] f(\alpha) d\alpha \quad (6)$$

$$P(Y_1 = 1 | X_1) P(Y_2 = y_2 | X_2, Y_1 = 1) = \int \Phi \left[X_1' \frac{\beta_1}{\sigma_1} + \frac{\alpha}{\sigma_1} \right] \Phi \left[(2y_2 - 1) \left(X_2' \frac{\beta_2}{\sigma_2} + \frac{\alpha}{\sigma_2} \right) \right] f(\alpha) d\alpha \quad (7)$$

Hierbei ist $\sigma_t(t = 1, 2)$ die Standardabweichung von $u_t(t = 1, 2)$, $f(\alpha, \cdot)$ die Dichte von α , und Φ die Verteilungsfunktion der Standard-Normalverteilung.

Wie bei allen Schwellenwert-Modellen sind die latenten Größen Y_1^* und Y_2^* nur bis auf einen Skalenfaktor identifiziert. Die Skala kann daher für jeden Zeitpunkt so gewählt werden, daß $\sigma_t = 1(t = 1, 2)$ gilt.

Das hier vorgestellte Modell verbindet Allison's sequentielles Probit-Modell mit der Möglichkeit, unbeobachtete Heterogenitäten zu identifizieren. Im Vergleich zu einer unverbundenen Schätzung durch 2 getrennte Probit-Modelle (Schätzung von $P(Y_1 = 1)$ auf Basis aller Personen, Schätzung von $P(Y_2 = y_2 \mid Y_1 = 1)$ auf Basis aller Personen mit $Y_1 = 1$) wird die Stichprobenninformation genauer ausgewertet. Die unverbundene Schätzung ist unter den Modellannahmen nicht effizient. Obwohl das hier beschriebene Modell große Ähnlichkeit mit dem multivariaten Probit-Modell von Heckman/Willis (1976) hat (vgl. z.B. Heckman 1981), sei hier auf einen wesentlichen Unterschied hingewiesen. Im Heckman/Willis-Modell bestehen keine Restriktionen hinsichtlich der Beobachtbarkeit der 0-1-Variablen. In diesem Modell kann zu jedem Zeitpunkt die abhängige 0-1 Variable beobachtet werden, unabhängig vom Ausgang der vorherigen Beobachtungen. Unter diesem Modell bedeutet die Einschränkung auf vollständig beobachtete Verläufe nicht notwendig eine Einschränkung für die Schätzbarkeit der Parameter.

Im sequentiellen Probit-Modell bedeutet jedoch die Einschränkung auf die vollständigen Verläufe den Ausschluß aller Beobachtungen mit $Y_1 = 0$. Auf dieser Datenbasis ist β_1 nicht mehr schätzbar. Im sequentiellen Probit-Modell müssen daher auch die unvollständigen Verläufe in die Schätzung eingehen. Aus diesem Grunde ist auch der Schätzansatz von Chamberlain (1980, 1985) auf dieses Modell nicht anwendbar. Bei diesem Ansatz werden die α_i als fixe Größen betrachtet und es wird die bedingte Verteilung von Y_1, Y_2 gegeben $Y_1 + Y_2$ untersucht. Unter einem Logit-Ansatz für die $Y_i(t = 1, 2)$ ist nämlich $Y_1 + Y_2$ suffizient für α und die bedingte Verteilung von Y_1, Y_2 gegeben $Y_1 + Y_2 = s$ ist unabhängig von α . Jedoch ist diese Verteilung nur für $s=1$ nicht entartet, d.h. Y_1 und Y_2 können verschiedene Werte annehmen - und nur diese Beobachtungen gehen in die Schätzung nach dem Chamberlain-Ansatz ein. Im sequentiellen Probit-Modell wird für $s=1$ jedoch nur $Y_1 = 1, Y_2 = 0$ beobachtet, während der symmetrische Ausdruck $Y_1 = 0, Y_2 = 1$ nicht beobachtet wird. D.h. auch für $s=1$ ist die Verteilung von Y_1, Y_2 entartet.

Wie Allison (1987) darstellt, steht das sequentielle Probit-Modell in enger Analogie zur Ereignisanalyse. Auch in dieser konkreten Problemstellung

treten massiv Links- und Rechtszensierungen auf. Rechtszensierungen treten auf, wenn Personen im Verlaufe des Panels die Schule verlassen haben, aber noch nicht ihre Lehre abgeschlossen haben. In diesem Fall ist Y_2 noch nicht bekannt. Der Likelihood-Beitrag dieser Personen ist daher $P(Y_1 = y_1)$.

Wesentlich schwieriger zu behandeln sind im vorliegenden Fall die linkszensierten Beobachtungen. Bei diesen Personen (es handelt sich um 37% der insgesamt 808 Personen) liegt keine Information über das Überschreiten der ersten Bildungsschwelle vor. Zwar kann aufgrund der erwähnten seltenen Kombination von Warteschleife und Lehrabschluß in der Regel $Y_1 = 1$ angenommen werden, allerdings sind viele Kovariaten zur Erklärung des Übergangs Schule/Berufsausbildung unbekannt. Hierzu gehören beispielsweise das Alter bei Schulabschluß, die Arbeitslosenquote am Wohnort sowie das Jahr des Abschlusses. Der Likelihood-Beitrag von Personen mit Linkszensuren reduziert sich daher auf $P(Y_2 = y_2 \mid Y_1 = 1)$.

Die Berechnung der ML-Lösung läßt sich dadurch vereinfachen, daß immer nur ein funktionaler Ausdruck für die Likelihood verwendet wird. Dieser Ausdruck ist durch

$$\tilde{L} = \int \prod_{t=1,2} \Phi[(2Y_t - 1)(X'_t \beta_t + \alpha)] f(\alpha) d\alpha \quad (8)$$

gegeben, wenn für nicht beobachtete Werte von Y_t der Wert $Y_t = 0.5$ gesetzt wird. Beispielsweise erhält man für $Y_2 = 0.5$

$$\tilde{L} = \int \Phi[(2Y_1 - 1)(X'_1 \beta_1 + \alpha)] \Phi(0) f(\alpha) d\alpha$$

Da der Faktor $\Phi(0) = 0.5$ nicht von Modellparametern abhängt, liefert die Maximierung von $\tilde{L}$ dieselben Resultate wie die Maximierung der ursprünglichen Likelihood. Die Auswertung von $\tilde{L}$ verlangt eine numerische Integration für jede Person. Die Integration bezüglich der Dichte der Normalverteilung kann über die Gauss-Hermite-Approximation durchgeführt werden. Wie Butler/Moffitt (1982) gezeigt haben, liefert die Gauss-Hermite Approximation von (8) schon mit wenigen Funktionsauswertungen hinreichend akkurate ML-Schätzungen.

5.2 Die Schätzung des Erwerbseinkommens unter einem bivariaten Selektionsmodell

Möchte man das Erwerbseinkommen E (genauer: den Logarithmus des Bruttoerwerbseinkommens), das direkt am Anschluß an die Lehre erzielt wird, analysieren, so muß man zunächst berücksichtigen, daß bei einem nennenswerten Anteil ² der Personen kein Einkommen beobachtet wird, weil

1. die betreffende Person im Jahr nach dem Lehrabschluß nicht erwerbstätig ist
oder
2. im Falle der Erwerbstätigkeit keine Angaben zum Einkommen (bzw. zu interessierenden Kovariablen) vorliegen.

Wiederum läßt sich der Prozeß der Beobachtung des Erwerbseinkommens als sequentielles Probit-Modell beschreiben. Die Stufe (1) wird hierbei durch die 0-1 Variable $R_1 = Y_2$ von Abschnitt 5.1 erfaßt. Die Stufe (2) sei durch die Variable R_2 beschrieben, die anzeigt, ob bei Erwerbstätigkeit Angaben zum Einkommen vorliegen.

Je nach Beobachtungsmuster ergeben sich die folgenden Beiträge zur Likelihoodfunktion:

<i>Beobachtungsmuster</i>	<i>Likelihoodbeitrag</i>	
$R_1 = 0$	$P(R_1 = 0)$	(9)

$R_1 = 1, R_2 = 0$	$P(R_1 = 1, R_2 = 0)$	(10)
--------------------	-----------------------	------

$R_1 = 1, R_2 = 1, E = e$	$P(R_1 = 1, R_2 = 1 \mid E = e)f_E(e)$	(11)
---------------------------	--	------

Hierbei ist $f_E(e)$ die Dichte für die Einkommensverteilung an der Stelle e . Wie im vorherigen Abschnitt werden die Wahrscheinlichkeiten $P(R_1 = r_1)$ und $P(R_2 = r_2 \mid R_1 = 1)$ durch

$$P(R_1 = 1) = P(Z'_1\gamma_1 + \delta_1 > 0) \quad (12)$$

$$P(R_2 = 1 \mid R_1 = 1) = P(Z'_2\gamma_2 + \delta_2 > 0) \quad (13)$$

²Im vorliegenden Fall 105 von 386 Personen ($\approx 27\%$.)

modelliert. Die Gleichungen (12) und (13) werden mit der Einkommensgleichung

$$E = X'_E \beta_E + \delta_E \quad (14)$$

über die Verteilungsannahme

$$\begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_E \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 + \sigma_u^2 & & \\ \sigma_u^2 & 1 + \sigma_u^2 & \\ \sigma_{1E} & \sigma_{2E} & \sigma_E^2 \end{pmatrix} \right] \quad (15)$$

gekoppelt.

Die Verteilungsannahme (15) besagt, daß die marginale Verteilung von (δ_1, δ_2) eine Varianzkomponenten-Struktur hat, d.h.

$$\sigma_1 = u + v$$

$$\sigma_2 = u + w$$

mit $\sigma_v^2 = \sigma_w^2 = 1$. Dies entspricht der Modell- Annahme des sequentiellen Probit-Modells mit unbeobachteter Heterogenität. Die Kovarianzen σ_{1E} und σ_{2E} modellieren eine selektive Kopplung der Einkommensgleichung mit den beiden Beobachtungsgleichungen.

Zur Berechnung von $P(R_1 = 1, R_2 = 1 \mid E = e)$ muß die bedingte Verteilung von δ_1, δ_2 gegeben δ_E bestimmt werden. Partitioniert man die Kovarianzmatrix der Verteilung von $\delta_1, \delta_2, \delta_E$ in

$$\Sigma = \begin{pmatrix} \Sigma_{12} & \Sigma_{12,E} \\ \Sigma_{E,12} & \Sigma_E \end{pmatrix}$$

so ist die bedingte Verteilung von δ_1, δ_2 , gegeben δ_E wiederum eine Normalverteilung.

Deren Erwartungswert $\tilde{\mu}$ und Kovarianzmatrix $\tilde{\Sigma}$ sind bestimmt durch (vgl. Johnson/Kotz (1972), S.41):

$$\tilde{\mu}' = \delta_E \Sigma_E^{-1} \Sigma_{E,12} \quad (16)$$

$$\tilde{\Sigma} = \Sigma_{12} - \Sigma_{12,E} \Sigma_E^{-1} \Sigma_{E,12} \quad (17)$$

Mit $\sigma_{tE} = \rho_{tE}\sigma_E(1 + \sigma_u^2)^{\frac{1}{2}} (t = 1, 2)$ erhält man aus (16)

$$\tilde{\mu} = \delta_E(1 + \sigma_u^2)^{\frac{1}{2}}/\sigma_E \begin{pmatrix} \rho_{1E} \\ \rho_{2E} \end{pmatrix} \quad (18)$$

Für den zweiten Term in (17) erhält man:

$$\begin{aligned} \Sigma_{12,E}\Sigma_E^{-1}\Sigma_{E,12} &= \frac{1}{\sigma_E^2} \begin{pmatrix} \sigma_{1E}^2 & \sigma_{1E}\sigma_{2E} \\ \sigma_{1E}\sigma_{2E} & \sigma_{2E}^2 \end{pmatrix} \\ &= (1 + \sigma_u^2) \begin{pmatrix} \rho_{1E}^2 & \rho_{1E}\rho_{2E} \\ \rho_{1E}\rho_{2E} & \rho_{2E}^2 \end{pmatrix} \end{aligned}$$

Damit hat auch $\tilde{\Sigma}$ wieder eine Varianzkomponentenstruktur mit

$$\tilde{\Sigma} = \begin{pmatrix} \sigma_v^2 & 0 \\ 0 & \sigma_w^2 \end{pmatrix} + \sigma_u^2 \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad (19)$$

wobei die Größe der Varianzkomponenten gegeben ist durch

$$\sigma_v^2 = 1 - (1 + \sigma_u^2)(\rho_{1E}^2 - \rho_{1E}\rho_{2E}) \quad (20)$$

$$\sigma_w^2 = 1 - (1 + \sigma_u^2)(\rho_{2E}^2 - \rho_{1E}\rho_{2E}) \quad (21)$$

$$\sigma_u^2 = \sigma_u^2 - (1 + \sigma_u^2)\rho_{1E}\rho_{2E} \quad (22)$$

Damit σ_v^2, σ_w^2 und σ_u^2 positiv sind, müssen einige Restriktionen an die Relation von ρ_{1E}, ρ_{2E} und σ_u^2 erfüllt sein. Diese Restriktionen ergeben sich aus der Tatsache, daß im Falle eines Varianzkomponentenmodells für δ_1, δ_2 die Korrelationen von δ_E mit δ_1 und δ_2 nicht beliebig sein können. So beschreibt die Positivität von σ_u^2 in (22) den Sachverhalt, daß die über δ_E geschätzte gemeinsame Varianzkomponente von δ_1 und δ_2 nicht größer als σ_u^2 sein kann. Die Varianz der Schocks ist in (20) durch $(1 + \sigma_u^2)(\rho_{1E}^2 - \rho_{1E}\rho_{2E})$ gegeben. Diese Varianz wird durch die Varianz der Schocks in der unbedingten Verteilung von δ_1 begrenzt. Analog ist die Positivität von σ_w^2 in Gleichung (22) zu interpretieren.

Insgesamt erhält man:

$$P[R_1 = 1, R_2 = 1 \mid \delta_E] = \int \left(\prod_{t=1,2} \Phi \left(\frac{Z'_t \beta_t + \delta_E \frac{\rho_{1E}}{\sigma_E} \sqrt{1 + \sigma_u^2} + \tilde{u}}{s_t} \right) \right) \tilde{f}(\tilde{u}) d\tilde{u} \quad (23)$$

mit $s_t^2 = 1 - (1 + \sigma_u^2)(\rho_{1E}^2 - \rho_{1E}\rho_{2E})$ $t = 1, 2$

und $\tilde{f}$ als Dichte einer Normalverteilung mit Varianz $\sigma_u^2 - (1 + \sigma_u^2)\rho_{1E}\rho_{2E}$.

Für eine Beobachtung mit $R_1 = 1, R_2 = 1$ und $E = e$ erhält man damit den Likelihoodbeitrag

$$L = P[R_1 = 1, R_2 = 1 \mid \delta_E = e - X'_E \beta_E] f_E(e) \quad (24)$$

Damit ist die Likelihoodfunktion für alle möglichen Beobachtungen bestimmt und die Modell-Parameter $\beta_E, \gamma_1, \gamma_2, \sigma_u, \sigma_E, \rho_{1E}$ und ρ_{2E} können durch rekursive Maximierung der Likelihood-Funktion geschätzt werden.

Dieses bivariate Selektionsmodell enthält die folgenden Modelle als Spezialfälle:

- (i) $\sigma_u = 0, \rho_{2E} = 0$ Univariates Selektionsmodell bzgl. R1
- (ii) $\sigma_u = 0, \rho_{1E} = 0$ Univariates Selektionsmodell bzgl. R2
- (iii) $\rho_{1E} = \rho_{2E} = 0$ Sequentielles Probit-Modell mit Heterogenität.

6 Datenbasis und Operationalisierung des Modells

6.1 Das Sozio-ökonomische Panel

Datenbasis für das zweistufige Entscheidungsmodell ist das Sozio-ökonomische Panel (SOEP) für die alten Bundesländer. Das SOEP ist eine Stichprobenerhebung im Rahmen der Bevölkerungs- und Erwerbsstatistik, mit der seit 1984 durch wiederholte Befragungen jährlich repräsentative Mikro-Längsschnittdaten über Personen, Familien und Haushalte in der Bundesrepublik bereitgestellt werden. Inzwischen stehen Daten aus sieben Erhebungswellen (1984 - 1990) für Auswertungen zur Verfügung (vgl. Projektgruppe 'Das Sozio-ökonomische Panel' 1990). Die erste Befragungswelle wurde 1984 in

rund 6000 Haushalten mit 12000 Personen durchgeführt. In jedem Haushalt werden alle Personen befragt, die 16 Jahre und älter sind. Das breitgefächerte Fragenprogramm des SOEP erstreckt sich auf die folgenden Themenfelder: Demographie, Arbeitsmarkt und Beschäftigung, Einkommen, Steuern und soziale Sicherung, Wohnen, Gesundheit, Bildung und Qualifikation, Leistungen privater Haushalte und schließlich Grundorientierungen und Integration.

6.2 Der Übergang an der ersten und an der zweiten Schwelle

Im Standardprogramm jeder Erhebungswelle wird danach gefragt, ob seit dem letzten Befragungstermin (vor etwa einem Jahr) eine Schule, eine Berufsausbildung oder ein Hochschulstudium abgeschlossen wurde. Diejenigen Personen, die angeben, den Haupt- oder Realschulabschluß erworben zu haben, bilden die Grundgesamtheit für die erste Stufe des Entscheidungsmodells: den Übergang von der Schulausbildung in die betriebliche Berufsausbildung (Y_1). Für die zweite Stufe - das ist der Übergang von der Berufsausbildung in die Erwerbstätigkeit (Y_2) - werden die Personen ausgewählt, die eine Lehre, bzw. eine Beamtenausbildung abgeschlossen haben.

Die Entscheidung an der ersten Schwelle (Zustand im Jahr nach Ausbildungsabschluß) wird wie folgt formuliert:

$$Y_{1it} = \begin{cases} 1 & \text{"betriebliche Berufsausbildung"} \\ 0 & \text{"sog. Warteschleife"} \end{cases}$$

Die Entscheidung an der zweiten Schwelle:

$$Y_{2it} = \begin{cases} 1 & \text{"Erwerbstätigkeit"} \\ 0 & \text{"arbeitslos, nicht erwerbstätig"} \end{cases}$$

Die betriebliche Berufsausbildung umfaßt auch die Ausbildung im Beamtenverhältnis. In der sogenannten Warteschleife (1. Schwelle) sind Personen zusammengefaßt, die ein Berufsgrundbildungs- bzw. -vorbereitungsjahr absolvieren oder eine Berufsschule ohne Lehre besuchen. Hinzu kommen diejenigen Personen, die sich weder in Ausbildung befinden noch einer Erwerbstätigkeit nachgehen. Ausbildungsabsolventen, die einen weiterführenden schulischen Ausbildungsabschluß anstreben (z.B. Haupt- oder

Realschüler, die ein Gymnasium besuchen oder Personen, die nach Abschluß der betrieblichen Ausbildung eine Hochschule besuchen), können den oben genannten Zuständen nicht zugeordnet werden und gehen somit nicht in die Analysepopulation ein.

Für die Schätzung des oben formulierten Entscheidungsmodells können in jeder Erhebungswelle des SOEP durchschnittlich etwa 150 Personen beobachtet werden. Insgesamt können 6 Wellen ausgewertet werden. Die Ereignisse über die Erhebungswellen 1985 bis 1990 werden "gepoolt" in das Modell aufgenommen. Neben der Vergrößerung der Fallzahl hat dies vor allem den Vorteil, daß für einen Teil der Personen die Entscheidung an der ersten und der zweiten Schwelle zu beobachten ist. Die zeitlichen Effekte gehen durch die "Poolung" nicht verloren, sie können als erklärende Variable kontrolliert werden.

Insgesamt basieren die folgenden Schätzungen auf den Entscheidungen von 808 Personen (vgl. Tabelle 3).

Tabelle 3: Der Übergang von Haupt- und Realschulabgängern in die Berufsausbildung und in Erwerbstätigkeit, 1984 bis 1990

Verbleib im Zeitraum 1985 bis 1990			
	0 ¹ in %	1 ² in %	Fälle
Erste Schwelle	31,0	69,0	509
Zweite Schwelle	17,9	82,1	386
Zusammen			895
darunter:			
Personen, für die beide Abschlüsse beobachtet werden konnten			87

- 1) 1. Schwelle: "Warteschleife"
2. Schwelle: "arbeitslos, nicht erwerbstätig"
- 2) 1. Schwelle: "betriebliche Berufsausbildung"
2. Schwelle: "erwerbstätig"

Quelle: Das Sozio-ökonomische Panel 1984 - 1990.

6.3 Erklärende Variable

In die zu schätzenden Modelle werden die folgenden Gruppen erklärender Variable einbezogen:

- sozio-demographische Merkmale der Ausbildungsabsolventen,
- Merkmale des Ausbildungsbetriebes (nur für die 2. Schwelle),
- allgemeine Merkmale des Arbeitsmarktes,
- Hintergrundmerkmale des Elternhauses sowie
- zeitbezogene fixe Effekte.

Zunächst einige allgemeine Anmerkungen zur Generierung der erklärenden Merkmale. Je nach Relevanz erfolgte die Zuweisung für das Jahr des Ausbildungsabschlusses (z.B. Merkmale des Ausbildungsbetriebes) oder für das Jahr der Ausbildungs- bzw. Erwerbstätigkeitsentscheidung (z.B. Charakteristika des Arbeitsmarktes). Bei dieser Vorgehensweise sind gewisse Unschärfen nicht auszuschließen, da nur der Zeitraum (maximal ein Jahr) und nicht der genaue Zeitpunkt des Ausbildungsbeginns bzw. des Beginns der Erwerbstätigkeit, bekannt ist.

Der zeitliche Einfluß auf die jeweiligen Entscheidungen wird zum einen durch das Jahr des Bildungsabschlusses in der Schätzung kontrolliert. Die zeitliche Entwicklung der Konjunktur- und Arbeitsmarktsituation wird zum anderen auch durch verschiedene Arbeitsmarkt- und Ausbildungsplatzindikatoren berücksichtigt, die auf der Ebene von Bundesländern berechnet wurden.

Die soziale Herkunft der Ausbildungsabsolventen wurde approximativ durch die berufliche Stellung des Vaters und das verfügbare Haushalts-Pro-Kopf-Einkommen abgebildet. Diese Informationen sind jedoch nur für die Ausbildungsabsolventen verfügbar, die noch im elterlichen Haushalt leben, oder für die der Vater in einem anderen Panelhaushalt identifiziert werden kann. Dies ist immerhin für 80 % der Ausgangspopulation der Fall. Diese Gruppe dürfte dennoch eine systematische Auswahl darstellen, da vor allem für jüngere Personen die Wahrscheinlichkeit groß ist, den Vater im selben Haushalt bzw. in einem anderen Panelhaushalt ausfindig zu machen. Aus

diesem Grund wurden auch zwei Modellvarianten geschätzt: Eine auf Basis der gesamten Population ohne die Hintergrundsmerkmale des Haushaltes und eine auf Basis der 80 %- Population unter Berücksichtigung der Hintergrundsmerkmale. Die Variablen werden im Anhang erläutert.

7 Ergebnisse

7.1 Die Übergänge an der ersten und der zweiten Schwelle

Tabelle 4 enthält die Koeffizienten für die Variablen der ersten Schwelle. Die 1. Spalte in den Tabellen ist eine Schätzung, welche ausschließlich auf der Basis der Beobachtungen der betreffenden Schwelle durchgeführt wurde. Die 2. Spalte enthält die Schätzung des sequentiellen Modells, wie in Abschnitt 5 beschrieben. Die Schätzungen in den Spalten 3 und 4 (Modelle 1a und 2a) sind jeweils analog zu den Spalten 1 und 2, aber um Merkmale der sozialen Herkunft erweitert.

Positive Koeffizienten bedeuten eine höhere Neigung zum direkten Übergang in die betriebliche Berufsausbildung (Tab. 4) bzw. zum direkten Übergang in die Erwerbstätigkeit (Tab. 5).

Beim Übergang an der ersten Schwelle (vgl. Tab. 4) hat das Geschlecht der Haupt- und Realschulabgänger einen signifikanten Einfluß auf den Erfolg auf dem Ausbildungsmarkt. Weibliche Bewerber haben deutlich geringere Chancen einen Ausbildungsplatz zu erhalten. Ausländische Schulabgänger erhalten seltener einen Ausbildungsplatz als Deutsche. Auch die Art des Schulabschlusses der Lehrstellenbewerber ist von Bedeutung. Hier schneiden Realschulabgänger signifikant besser ab, als Personen, die nur einen Hauptschulabschluß nachweisen können. Wie erwartet, erweist sich auch die regionale Verfassung des Arbeitsmarktes als einflußreich für den Einstieg in die Berufsausbildung. Gute Chancen, einen Ausbildungsplatz zu erhalten, bestehen in den Ballungsräumen (Städten mit mehr als 500 000 Einwohnern) und in Bundesländern mit einer geringen Arbeitslosenquote. Hingegen sind zeitbezogene Effekte beim Übergang in die Berufsausbildung, wie auch in die Erwerbstätigkeit, nicht zu beobachten. Dies deutet zugleich darauf hin, daß sich die Entscheidungsstruktur im betrachteten Zeitraum nicht verändert hat und sich damit die Entscheidung, das Modell mit zeitlich konstanten Struk-

Tabelle 4: Der Übergang von Haupt- und Realschülern in betriebliche Berufsausbildung oder Warteschleife (1. Schwelle)

	Modell			
	1	2	1a	2a
Konstante	-0.315	-0.368	-2.559	-2.641
Alter	-0.044	-0.049	-0.008	-0.008
Nation (Ausl=1)	-0.546**	-0.599**	-0.518**	-0.533**
Geschlecht (M=1)	0.428**	0.473**	0.499**	0.515**
Realabschluß	0.482**	0.535**	0.508**	0.526**
Ortsgr 1	-0.293*	-0.315*	-0.273	-0.281
Ortsgr 2	-0.189	-0.205	-0.244	-0.252
Ortsgr 3	-0.404**	-0.439**	-0.423**	-0.435*
Azu-Quote	0.024*	0.027*	0.028*	0.029*
Al-Quote	-0.053**	-0.057**	-0.027	-0.027
Jahr86	-0.152	-0.170	-0.241	-0.251
Jahr87	0.202	0.213	0.001	0.000
Jahr88	0.313	0.354	0.282	0.294
Jahr89	0.283	0.317	0.237	0.247
Jahr90	0.029	0.037	-0.153	-0.155
VaterBni			-0.162	-0.169
VaterBho			-0.811**	-0.835**
Vater-Al			-0.114	-0.110
HH-Einkom			1.177**	0.183**
$HH - Einkom^2$			-0.005**	-0.006**
σ_α	-	0.475	-	0.257
Fälle	509	808	448	683
Likelihood	-123.4	-150.5	-100.2	-118.2

Modell 1 und 1a: gesonderte Schätzung für die 1. Schwelle

Modell 2 und 2a: Schätzung des sequentiellen Modells

* signifikant zum Niveau 0.1

** signifikant zum Niveau 0.05

Tabelle 5: Übergang vom Abschluß der beruflichen Berufsausbildung in Erwerbstätigkeit oder Arbeitslosigkeit/Nichterwerbstätigkeit (2. Schwelle)

	Modell			
	1	2	1a	2a
Konstante	1.748**	1.874**	4.221**	4.315**
Alter	0.010	0.015	-0.081	-0.081
Nation (Ausl=1)	0.234	0.249	0.327	0.331
Geschlecht (M=1)	-0.518**	-0.581**	-0.607**	-0.630**
Fam-stand(verh=1)	-0.076	-0.102	5.451	5.451
Realabschluß	-0.145	-0.169	-0.085	-0.090
Ortsgr 1	-0.142	-0.156	-0.051	-0.049
Ortsgr 2	-0.138	-0.148	-0.278	-0.282
Ortsgr 3	-0.005	-0.000	-0.043	-0.046
Al-Quote	-0.086**	-0.098**	-0.093**	-0.097**
Br-Grundstoff	-0.160	-0.189	-0.310	-0.324
Br-Verbrauch	0.298	0.341	-0.139	-0.146
Br-Bau	-0.260	-0.286	-0.535	-0.550
Br-HandelGast	-0.130	-0.142	-0.460	-0.477
Br-Dienst	-0.047	-0.058	-0.343	-0.358
Br-Übrige	0.053	0.062	0.037	0.041
Öff-Dienst	0.526*	0.603*	0.660**	0.690*
Betr-Größe	0.063	0.073	-0.008	-0.008
Jahr86	0.017	0.026	-0.021	-0.022
Jahr87	0.032	0.033	-0.179	-0.185
Jahr88	-0.214	-0.280	-0.525	-0.555
Jahr89	0.289	0.297	-0.033	-0.027
Jahr90	-0.243	-0.291	-0.512	-0.535
VaterBni			-0.259	-0.265
VaterBho			-0.059	-0.006
Vater-Al			0.592	-0.605
σ_α	-	0.475	-	0.257
Fälle	386	808	314	683
Likelihood	-1.92	-150.5	-54.1	-118.2

Anmerkungen zu Tabelle 5:

Modell 1 und 1a: gesonderte Schätzung für die 2. Schwelle

Modell 2 und 2a: Schätzung des sequentiellen Modells

* signifikant zum Niveau 0.1

** signifikant zum Niveau 0.05

turparametern zu schätzen, als richtig erweist.

Beim Übergang von der Berufsausbildung in die Erwerbstätigkeit (Tabelle 5) dreht sich der Einfluß des Geschlechtes um. Daraus sollte jedoch nicht auf eine Diskriminierung der männlichen Absolventen geschlossen werden. Vielmehr dürfte in dem negativen Koeffizienten zum Ausdruck kommen, daß viele Männer nach Abschluß ihrer Berufsausbildung den Wehr- oder Zivildienst absolvieren. Insgesamt läßt sich feststellen, daß die in vielen Untersuchungen angenommenen Nachteile von Frauen nicht bestätigt werden können. Das gleiche gilt für Ausländer und Personen mit niedrigem Schulabschluß, wo Nachteile nicht bestätigt werden können. Personenbezogene Merkmale spielen beim Übergang in die Erwerbstätigkeit offenbar keine ausgeprägte Rolle. Wie bei Schwelle 1 ist aber auch hier die Situation am Arbeitsmarkt von Einfluß. Von den Merkmalen des Ausbildungsbetriebes ist einzig die Variable öffentlicher Dienst positiv signifikant. In den Modellen 1a und 2a wurden zusätzlich noch Merkmale des Haushaltes, in dem die Personen aufgewachsen sind, berücksichtigt. Signifikante Effekte sind aber nur für den Übergang an der ersten Schwelle zu beobachten. Dabei ist allerdings ein auffälliger Befund zu konstatieren. Absolventen von Haupt- und Realschulen, deren Väter in hohen beruflichen Positionen tätig sind, erhalten mit geringerer Wahrscheinlichkeit einen Ausbildungsplatz³. Dies ist vermutlich ein Ausdruck negativer Selektion, die ihre eigentliche Ursache schon in der Wahl der Schulausbildung hat, wenn man davon ausgeht, daß Kinder aus Haushalten mit hoher Sozialstellung i.d.R. die Hochschulreife anstreben. Bei Haupt- und Realschülern aus solchen Haushalten handelt es sich dann häufig um lernschwache oder wenig motivierte Jugendliche, die auf dem Ausbildungsmarkt c.p. geringere Chancen haben.

σ_a ist der geschätzte Parameter für die Streuung der nicht direkt beobachteten Heterogenität, d. h. für die über die Zeit konstante, aber zwischen

³Dieses Ergebnis wurde zunächst in einer Schätzung beobachtet, in die das Pro-Kopf-Einkommen des Haushalts lediglich linear einging. Dabei lag der Verdacht nahe, daß es sich bei dem negativen Einfluß der hohen beruflichen Position des Vaters um eine "Korrektur" dieser Linearitätsannahme handelt. Darum wurde zusätzlich das quadrierte Einkommen in die Gleichung aufgenommen. Dieses zeigt erwartungsgemäß das negative Vorzeichen (beide Einkommensvariablen sind auch signifikant). Der Einfluß der beruflichen Position des Vaters ist allerdings weiter zu beobachten. Zur Kontrolle wurde auch noch eine Schätzung ohne dieses Merkmal durchgeführt (Tabelle 4). Dabei verändert sich der Einfluß des Haushaltseinkommens kaum.

den Personen variierenden Varianzkomponente α_i . Der geschätzte Parameter beträgt im Modell 2 lediglich das 0,48-fache der zeitpunktbezogenen Streuung σ_i und ist zudem nicht signifikant (t-Wert ist 0,9). Für Modell 2a beträgt der geschätzte Parameter jedoch nur noch 0,26. Dies deutet darauf hin, daß die Hintergrundmerkmale des Haushaltes einen Beitrag zur Aufdeckung nicht beobachteter individueller Heterogenität leisten.

Insgesamt läßt sich feststellen, daß die individuelle Heterogenität in beiden Modellen keine große Rolle spielt. D.h., bei getrennten Schätzungen für beide Schwellen mit einem univariaten Probit-Modell begeht man keinen gravierenden Fehler. Im Vergleich zum univariaten Probit-Modell (Modelle 1 und 1a) erhält man mit dem sequentiellen Modell (Modelle 2 und 2a) größere Werte für die geschätzten Koeffizienten. Das liegt daran, daß man eigentlich keine Schätzwerte für β , sondern nur für β/σ erhält. Im sequentiellen Modell ist σ jedoch um die geschätzte Komponente für α_i geringer. Dies muß allerdings nicht bedeuten, daß man auch vermehrt signifikante Koeffizienten erhält, da auch die Varianzen der Parameterschätzungen größer werden.

Die geschätzten Modelle bestätigen die Ergebnisse von Palamidis und Schwarze (1989), doch sind hier die Koeffizienten in deutlich höherem Maße signifikant. Dies dürfte an der größeren Zahl der Beobachtungen, etwas veränderten Variablenabgrenzungen und dem hier gewählten Schätzansatz liegen. Problematisch bei der gewählten Modellspezifikation ist, daß die Übergänge definiert sind als die Veränderung des 'Status' einer Person von einem Stichtag zum anderen. Kurzzeitige Episoden, z. B. Warteschleifen, oder Entscheidungen, die erst nach dem zweiten Stichtag getroffen wurden, bleiben dabei unberücksichtigt. Will man dies berücksichtigen, bieten sich Ratenmodelle an (vgl. Weißhuhn und Rohwer 1991). Inhaltlich führt aber auch dieser Ansatz zu sehr ähnlichen Ergebnissen.

7.2 Das Arbeitseinkommen bei Berufsbeginn

Nachdem im vorangehenden Abschnitt betrachtet worden ist, welche Faktoren den erfolgreichen Übergang in eine Lehre und in anschließende Erwerbstätigkeit bestimmen, soll nunmehr nach den Bestimmungsgründen des Einkommens gefragt werden, das in der Erwerbstätigkeit erzielt wird. Die erklärenden Variablen sind im wesentlichen dieselben wie in den Tabellen 5. Zusätzlich berücksichtigt wurde berufliche Mobilität in dem Jahr zwischen dem Abschluß der Lehre und der beobachteten Erwerbstätigkeit. Es wurde

ermittelt, ob

- der Betrieb gewechselt wurde und, wenn ja, ob
- die Branche gewechselt wurde.

Bei der Einkommensschätzung wurde, der Tradition der Humankapitaltheorie folgend, ein semi-logarithmischer Ansatz gewählt. Zu berücksichtigen ist, daß die beobachtete Personengruppe hinsichtlich der zentralen Variablen der Humankapitaltheorie weitgehend homogen ist: Alle Personen haben Lehrabschluß. Unterschiede gibt es nur beim Schulabschluß (Haupt- oder Realschule). Auch das Alter streut nur sehr wenig.

Wie in Abschnitt 5 ausgeführt, wurde ein Modell formuliert, welches berücksichtigt, daß es sich um ein mehrstufiges Problem handelt.

1. Stufe: Übergang in Erwerbstätigkeit
2. Stufe: Verfügbarkeit von Einkommensangaben bei den Erwerbstätigen
3. Stufe: Höhe des Einkommens

Die Ergebnisse sind in den Tabellen 6, 7 und 8 wiedergegeben. Außerdem standen mehrere Spezifikationsmöglichkeiten des Modells zur Verfügung. Die Stufen 1 und 2 wurden jeweils als

- univariates Probitmodell (Spalte 1),
- sequentielles Probitmodell (Spalte 2),
- univariates Selektionsmodell mit ML-Schätzung (Spalte 3)

sowie als – sequentielles Selektionsmodell mit ML-Schätzung (Spalte 4) geschätzt.

Tabelle 6: Übergang von Jugendlichen in Erwerbstätigkeit

1. Spalte : Univ. Probit
2. Spalte : Sequentielles Probit
3. Spalte : univariates Selektions-Modell ;ML-Lösung
4. Spalte : Sequentielles Selektions-Modell ;ML-Lösung

Variable	(1)	(2)	(3)	(4)
Konstante	1.748*	1.748*	1.581**	1.750*
Alter	0.010	0.010	0.009	0.011
Nation (Ausl=1)	0.234	0.234	0.195	0.237
Geschlecht (M=1)	-0.518*	-0.518*	-0.466*	-0.527*
Fam-stand (verh=1)	-0.076	-0.076	-0.084	-0.077
Realabschluß	-0.145	-0.145	-0.138	-0.145
Ortsgr 1	-0.142	-0.142	-0.136	-0.145
Ortsgr 2	-0.138	-0.138	-0.071	-0.139
Ortsgr 3	-0.005	-0.005	-0.014	0.000
AL-Quote	-0.086*	-0.086*	-0.085*	-0.087*
Br-Grundstoff	-0.160	-0.160	-0.057	-0.169
Br-Verbrauch	0.298	0.298	0.397	0.306
Br-Bau	-0.260	-0.260	-0.237	-0.265
Br-HandelGast	-0.130	-0.130	-0.041	-0.122
Br-Dienst	-0.047	-0.047	0.073	-0.038
Br-Übrige	0.053	0.053	0.047	0.049
Betr-Größe	0.063	0.063	0.062	0.067
Öff-Dienst	0.526**	0.526**	0.558**	0.535***
Jahr 86	0.017	0.017	-0.022	0.020
Jahr 87	-0.032	-0.032	-0.026	-0.030
Jahr 88	-0.214	-0.214	-0.172	-0.216
Jahr 89	0.289	0.289	0.356	0.293
Jahr 90	-0.243	-0.243	-0.252	-0.245

* signifikant zum 0.05-Niveau

** signifikant zum 0.10-Niveau

*** t-Wert nahe dem 0.10-Niveau

Das Einkommen wurde zuerst einmal mit einer einfachen OLS- Schätzung erklärt (Spalte 1). Die Interaktion mit Stufe 1 bzw. Stufe 2 wurde in einer Heckman Korrektur der OLS-Schätzung berücksichtigt (Spalte 2 bzw. Spalte 3). Statt der zweistufigen Heckman-Routine wurde auch eine ML-Schätzung für beide Alternativen getestet (Spalten 4 und 5). Schließlich wurde, als umfassendstes Modell, auch eine Maximum-Likelihood-Schätzung in Verbindung mit dem sequentiellen Probitmodell gerechnet (Spalte 6).

Die Ergebnisse von Tabelle 6 entsprechen inhaltlich denen von Tabelle 5 und brauchen hier nicht noch einmal kommentiert zu werden. Hingewiesen sei allerdings auf das Ergebnis, daß die Koeffizientenwerte in den 4 geschätzten Modellen bemerkenswert stabil sind.

Fehlende Angaben beim Einkommen (Tabelle 7) treten gehäuft bei Ausländern auf und bei Frauen. Die übrigen in der Schätzung berücksichtigten Variablen sind entweder nicht signifikant oder inhaltlich wenig aussagefähig.

Bei der Erklärung des Einkommens (Tabelle 8) bestätigt sich auch hier wieder eine positive Korrelation mit dem Lebensalter, das hier aber eher als persönliche Reife zu interpretieren sein dürfte, denn als Berufserfahrung. Ausländerstatus ist nicht mit einer (signifikanten) Einkommensreduktion verbunden. Männliche Jugendliche verdienen jedoch deutlich mehr als weibliche. Der Unterschied zwischen Haupt- und Realschulabschluß erweist sich als wenig einflußreich, aber eine hohe regionale Arbeitslosenquote senkt das Einkommen. Größere Betriebe zahlen höhere Einkommen. Einige Branchen zahlen signifikant niedrigere Löhne (die Verbrauchsgüterindustrie, Handel und Gaststätten, die übrigen Dienstleistungszweige; Bezugsgruppe ist die Metallindustrie). Betriebswechsel steigern das Einkommen, vor allem dann, wenn sie innerhalb der Branche erfolgen (wodurch Humankapitalentwertung vermieden wird).

Tabelle 7: Fehlende Einkommensangabe

1. Spalte : Univ. Probit
2. Spalte : Sequentielles Probit
3. Spalte : univariates Selektions-Modell ;ML-Lösung
4. Spalte : Sequentielles Selektions-Modell ;ML-Lösung

Variable	(1)	(2)	(3)	(4)
Konstante	0.631	0.631	0.277	0.630
Alter	0.007	0.007	0.007	0.004
Nation (Ausl=1)	-0.502*	-0.502*	-0.459**	-0.492**
Geschlecht (M=1)	0.647*	0.647*	0.782*	0.682*
Fam-stand (verh=1)	0.479	0.479	0.340	0.466
Realabschluß	0.127	0.127	0.057	0.112
Ortsgr 1	-0.150	-0.150	-0.108	-0.149
Ortsgr 2	0.458	0.458	0.683	0.474
Ortsgr 3	-0.391	-0.391	-0.339	-0.384
AL-Quote	0.000	0.000	0.007	0.004
Br-Verbrauch	0.562	0.562***	0.681**	0.558
Br-Bau	-0.147	-0.147	-0.057	-0.141
Br-HandelGast	0.405	0.405	0.553	0.417
Br-Dienst	0.663**	0.663**	0.877*	0.717**
Br-Übrige	0.032	0.032	0.124	0.032
Betr-Größe	-0.037	-0.037	-0.034	-0.034
Öff-Dienst	0.196	0.196	0.449	0.257
Jahr 86	-0.191	-0.191	-0.136	-0.193
Jahr 87	0.038	0.038	0.110	0.051
Jahr 88	0.304	0.304	0.398	0.313
Jahr 89	0.910*	0.910*	1.059*	0.930*
Jahr 90	-0.033	-0.033	0.006	-0.045

* signifikant zum 0.05-Niveau

** signifikant zum 0.10-Niveau

*** t-Wert nahe dem 0.10-Niveau

Tabelle 8: Erwerbseinkommen von Jugendlichen nach Abschluß der betrieblichen Berufsausbildung

1. Spalte : OLS-Schätzung
2. Spalte : two-step mit univ. Probit für R_1
3. Spalte : two-step mit univ. Probit für R_2
4. Spalte : ML mit Probit für R_1
5. Spalte : ML mit Probit für R_2
6. Spalte : ML mit sequentielltem Probit

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Konstante	6.911*	6.926*	6.854*	6.914*	6.883*	6.899*
Alter	0.028*	0.028*	0.028*	0.028*	0.028*	0.027*
Nation (Ausl=1)	-0.023	-0.006	-0.053	-0.020	-0.046	-0.040
Geschlecht (M=1)	0.110*	0.070	0.137*	0.102**	0.131*	0.123*
Fam-stand (verh=1)	-0.006	-0.006	0.014	-0.006	0.008	0.017
Realabschluß	0.049	0.039	0.057	0.047	0.052	0.053
Ortsgr 1	-0.031	-0.049	-0.041	-0.035	-0.041	-0.043
Ortsgr 2	-0.099	-0.114	-0.084	-0.102	-0.088	-0.098
Ortsgr 3	-0.043	-0.046	-0.066	-0.044	-0.063	-0.063
Al-Quote	-0.013***	-0.021**	-0.013***	-0.015**	-0.013***	-0.014***
Arb-zeit	0.038	0.036	0.037	0.038	0.035	0.037
Arbeiter	0.057	0.048	0.060	0.054	0.052	0.052
Angest.	0.017	0.008	0.019	0.014	0.013	0.014
Br-Grundstoff	-0.087	-0.098	-0.098	-0.089	-0.102	-0.097
Br-Verbrauch	-0.164**	-0.153**	-0.157**	-0.162*	-0.161*	-0.162*
Br-Bau	0.070	0.042	0.049	0.063	0.049	0.045
Br-HandelGast	-0.156**	-0.167**	-0.159**	-0.159**	-0.163**	-0.164
Br-Dienst	-0.235*	-0.239*	-0.222*	-0.235*	-0.224*	-0.228*
Br-Übrige	-0.103	-0.100	-0.103	-0.103	-0.100	-0.100
Betr-Größe	0.082*	0.086*	0.080*	0.083*	0.081*	0.082*
Öff-Dienst	-0.003	0.022	0.005	0.002	-0.000	-0.003
Jahr 86	0.030	0.036	0.015	0.030	0.017	0.020
Jahr 87	-0.084	-0.086	-0.084	-0.085	-0.086	-0.088
Jahr 88	-0.000	-0.015	0.014	-0.003	0.011	0.006
Jahr 89	-0.066	-0.039	-0.028	-0.059	-0.036	-0.039
Jahr 90	0.077	0.064	0.074	0.074	0.075	0.071

Fortsetzung Tabelle 8

glbranche	0.091	0.095	0.101***	0.093	0.101**	0.101**
glbetrieb	-0.099*	-0.095*	-0.103*	-0.098*	-0.102*	- 0.101*
heckman	0.000	0.244	0.217	0.000	0.000	0.000
σ_E	0.32	0.32	0.32	0.31	0.32	0.32
R2	0.32	0.32	0.32	0.32	-	-
ρ_{1E}	-	-	-	0.16	-	0.11
ρ_{2E}	-	-	-	-	0.56*	0.45*
N	281	281	281	350	317	386

* signifikant zum 0.05-Niveau

** signifikant zum 0.10-Niveau

*** t-Wert nahe dem 0.10-Niveau

Soweit die erklärenden Variablen signifikant sind, entsprechen die Vorzeichen den Erwartungen. Geht man davon aus, daß Alter und Geschlecht individuell nicht beeinflußt werden können, können junge Erwerbstätige ihre Einkommenschancen also durch eine günstige Wahl der Region, der Betriebsgröße, der Branche und der beruflichen Mobilität beeinflussen.

Bemerkenswert ist, daß trotz der extrem homogenen Grundgesamtheit ca. 30 % der Einkommensvarianz durch den gewählten Ansatz erklärt werden können.

Abschließend ist zu fragen, wie die empirischen Ergebnisse durch die verschiedenen ökonometrischen Schätzansätze beeinflußt werden, d.h. durch die explizite Modellierung der Selektionsprozesse beim Übergang in Erwerbstätigkeit, bei der Beobachtung des Einkommens, der unbeobachteten Heterogenität sowie durch den Übergang von 2- step zu ML-Schätzung. Die Antwort ist außerordentlich klar und einfach: Die Koeffizientenwerte sind hochgradig stabil. Bei der wichtigen Variablen Geschlecht zeigt sich allerdings, daß die Heckman-Korrektur den Koeffizienten und den t-Wert stark - und in vermutlich übertriebener Weise - nach unten korrigiert, ein Reaktionsverhalten, das der Heckman-Korrektur in der Literatur verschiedentlich entgegengehalten wird.

Man kann das Ergebnis auf zweifache Weise formulieren. Entweder: Das einfache OLS-Verfahren liefert (für unsere Untersuchung) sehr gute Näherungswerte. Oder: Mittels der anderen, sehr viel raffinierteren Modelle kann

gezeigt werden, daß die Prozeßselektivität, die unbeobachtete Heterogenität und das Schätzverfahren für unser Untersuchungsproblem keine gravierende Rolle spielen.

Literatur

- [1] Allison, P. (1982), *Discrete-Time Methods for the Analysis of Event Histories*, in: S. Leinhardt (ed.), *Sociological Methodology*, San Francisco, Jossey-Bass
- [2] Blien, U., H. Reinberg, M. Tessaring (1990), *Die Ermittlung der Übergänge zwischen Bildung und Beschäftigung*, in: MittAB, H. 1, S. 181-204
- [3] Brandes, H., W. H. Brosi, A. Menk (1986), *Wege in die berufliche Bildung, Befragung von Lehrlingen in fünf Bundesländern 1983 und 1984*, in: MittAB, H. 2, S. 287-297
- [4] Butler, J., R. Moffitt (1982), *A Computationally Efficient Quadrature procedure of the One-Factor Multinomial Probit Model*, *Econometrica*, 50, 761-764
- [5] Chamberlain, G. (1980), *Analysis of covariance with qualitative data*, *Review of Economic Studies*, 47, 225-238
- [6] Chamberlain, G. (1984), *Panel Data*, in: *Handbook of econometrics*, Griliches, Z., Intriligator, M. (eds.), Elsevier, Amsterdam
- [7] Greene, W. (1988), *LIMDEP Version 5*
- [8] Heckman, J., R. Willis (1976), *Estimation of a Stochastic Model of Reproduction: An Econometric Approach*, in: *Household Production and Consumption*, Terleckyj, N. (ed.), National Bureau of Economic Research, NY u.a.
- [9] Hsiao, Ch. (1986), *Analysis of Panel Data*, Cambridge University Press, Cambridge
- [10] Johnson, X., S., Kotz (1972), *Continuous Multivariate Distributions in Statistics*, Wiley, New York
- [11] Palamidis, H., J. Schwarze (1989), *Jugendliche beim Übergang in eine betriebliche Berufsausbildung und in die Erwerbstätigkeit - Literaturüberblick und neuere empirische Ergebnisse*, in: MittAB, H. 1, S. 114 - 124

- [12] Stegmann, H., H. Kraft (1987), *Ausbildungs- und Berufswege von 23-24jährigen*, in: MittAB, H. 2, S. 142-163
- [13] Tessaring, M. U. Blien, G. Fischer, I. Hofmann, H. Reinberg (1990), *Bildung und Beschäftigung im Wandel*, BeitrAB 126, Nürnberg
- [14] Weißhuhn, G., G. Rohwer (1991), *Indikatoren zur Analyse des Arbeitsmarkterfolges von Absolventen der beruflichen Bildung (Duales System) in Westdeutschland* (Manuskript)

8 Anhang, Variablenübersicht

Alter	Alter in Jahren zum Zeitpunkt des Bildungsabschlusses
Nation	Nationalität: 1, wenn Person türkischer, jugoslawischer, italienischer, spanischer oder griechischer Nationalität ist, sonst 0.
Geschlecht	Geschlecht: 1, wenn Person männlich, sonst 0.
Fam-stand	Familienstand: 1, wenn Person verheiratet, sonst 0.
Realabsschluß	Schulabschluss: 1, wenn Realschulabschluss, sonst 0.
ORT 0	Nichttrandzonen mit 500 000 und mehr Einwohnern (Referenzgruppe).
Ort 1	Ortsgröße: 1, wenn 100 000 bis 500 000 Einwohner, sonst 0.
Ort 2	Ortsgröße: 1, wenn 20 000 bis 100 000 Einwohner, sonst 0.
Ort 3	Ortsgröße: 1, wenn unter 20 000 Einwohner, sonst 0.
Azu-Quote	Quotient aus: neu abgeschlossenen Ausbildungsverträge und Schulabgängern aus allgemein bildenden Schulen nach Bundesländern (Quellen: Statistisches Bundesamt: Fachserie 11, Reihe 1, Allgemeinbildende Schulen und Reihe 3, Berufliche Bildung, Jahrgänge 1984 bis 1989, die Zahlen für 1990 sind geschätzt).
Al-Quote	Arbeitslosenquote nach Bundesländern (Quelle: Amtliche Nachrichten der Bundesanstalt für Arbeit (39), Arbeitsstatistik 1990 - Jahreszahlen 1991).
Jahr86	Zeiteffekt: 1, wenn Abschluß im Jahr 1986, sonst 0.
Jahr87	Zeiteffekt: 1, wenn Abschluß im Jahr 1987, sonst 0.
Jahr88	Zeiteffekt: 1, wenn Abschluß im Jahr 1988, sonst 0.
Jahr89	Zeiteffekt: 1, wenn Abschluß im Jahr 1989, sonst 0.
Jahr90	Zeiteffekt: 1, wenn Abschluß im Jahr 1990, sonst 0. (Referenzjahr ist 1985)
VaterBni	Beruf des Vaters: 1, wenn ungelernter/angelernter Arbeiter, Angestellter mit einfacher Tätigkeit, Beamter im einfachen Dienst, sonst 0.

VaterBho	Beruf des Vaters: 1, wenn Angestellter mit hochqualifizierter Tätigkeit, Leitungsfunktion, oder umfassenden Führungsaufgaben, Beamter im höheren Dienst, Selbständiger in freiem Beruf, mit 10 oder mehr Angestellten oder selbständiger Akademiker, sonst 0. Referenzgruppe: alle anderen.
Vater-Al	Vater arbeitslos.
HH-Einkom	Prokopfeinkommen im Haushalt des Vaters geteilt durch 100.
HH-Einkom 2	Prokopfeinkommen im Haushalt des Vaters geteilt durch 100, zum Quadrat.
Br-Grundstoff	Branche des Ausbildungsbetriebes: 1, wenn Grundstoffindustrie, sonst 0.
Br-Verbrauch	Branche des Ausbildungsbetriebes: 1, wenn Verbrauchsgüterindustrie, sonst 0.
Br-Bau	Branche des Ausbildungsbetriebes: 1, wenn Baubereich, sonst 0.
Br-HandelGast	Branche des Ausbildungsbetriebes: 1, wenn Handel und Gaststätten, sonst 0.
Br-Dienst	Branche des Ausbildungsbetriebes: 1, wenn Dienstleistungen, sonst 0.
Br-Übrige	Branche des Ausbildungsbetriebes: 1, wenn restliche Branchen, sonst 0. (Referenzgruppe: Investitionsgüterindustrie)
Öff-Dienst	Öffentlicher Dienst: 1, wenn Ausbildungsbetrieb zum öffentlichen Dienst gehört, sonst 0.
Betr-Größe	Größe des Ausbildungsbetriebes nach Zahl der Beschäftigten. 4 Klassen aufsteigend.
Arb-zeit	Tatsächliche Arbeitszeit in Stunden (log)
glbranche	Gleiche Branche: 1, wenn Branche des Betriebs gleich Branche des Ausbildungsbetriebs, sonst 0.
glbetrieb	Gleiche Betriebsgröße: 1, wenn Größe des Betriebs gleich Größe des Ausbildungsbetriebs, sonst 0.