

Diskrete Hazardraten-Modelle in der Shell-Jugendstudie 1985: eine Anwendung des Programms GLAMOUR

Tutz, Gerhard; Georg, Werner

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Tutz, G., & Georg, W. (1991). Diskrete Hazardraten-Modelle in der Shell-Jugendstudie 1985: eine Anwendung des Programms GLAMOUR. *ZA-Information / Zentralarchiv für Empirische Sozialforschung*, 29, 81-93. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-202441>

Nutzungsbedingungen:

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

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

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

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



Diskrete Hazardraten-Modelle in der Shell-Jugendstudie 1985:
Eine Anwendung des Programms GLAMOUR

¹
von Gerhard Tutz und Werner Georg²

Abstract

Bei Modellierungsansätzen für diskrete Verweildauern mit wenigen Ausprägungen erweisen sich diskrete Hazardratenmodelle als adäquates Instrument. Ansätze dafür werden durch einen sequentiellen Mechanismus motiviert. Mit Hilfe des Programmsystems GLAMOUR wird ein Vergleich durchgeführt zwischen dem früheren Westdeutschland und Ungarn hinsichtlich des Merkmals "sexueller Erstkontakt von Jugendlichen".

Event history analysis in discrete time with few time categories may be based on discrete hazard models. Models of this type are motivated from a sequential approach. With the use of the program GLAMOUR a comparison is made between former West Germany and Hungary referring to the age teenagers have their first sexual contact.

1. Einleitung

Ziel der Verweildauer- oder Ereignisanalyse ist die Bestimmung der Dauer, die ein Individuum, eine Organisation oder allgemein eine statistische Einheit in einem Zustand verweilt. Dieser Zustand kann individuelle Arbeitslosigkeit ebenso sein, wie das Fehlen sexueller Erfahrungen. Letzteres ist ein Aspekt, der im Rahmen der Shell-Jugendstudie 1985 untersucht wird (vgl. Jugendwerk der Deutschen Shell 1985). Neben der Modellierung der reinen Verweildauer ist insbesondere die Identifizierung und Quantifizierung von Einflußgrößen, die den Verweildauerprozeß bestimmen, von zentralem Interesse.

In der stetigen Verweildaueranalyse ist das auf Cox (1972) zurückgehende 'proportional hazards'- oder Cox-Modell die dominierende Modellvorstellung. Die Hazard- oder Intensitätsfunktion, die die Chance bzw. das Risiko einer Zustandsänderung zum Zeitpunkt t wiedergibt, gegeben der Zustand hält bei t noch an, besitzt bei gegebenem Kovariablenvektor $\mathbf{x}' = (x_1, \dots, x_p)$ die Form

$$\lambda(t|\mathbf{x}) = \lambda_0(t)\exp(\mathbf{x}'\boldsymbol{\beta}), \quad (1.1)$$

wobei $\lambda_0(t)$ die sog. Baseline-Hazardrate darstellt. Diese Baseline-Hazardrate, der jedes Individuum der Population unterliegt, wird durch den Faktor $\exp(\mathbf{x}'\boldsymbol{\beta})$ in Abhängigkeit

1 Prof. Dr. Gerhard Tutz vertritt derzeit den Lehrstuhl für Statistik an der Universität Regensburg, Universitätsstr. 31, D-8400 Regensburg.

2 Dr. Werner Georg ist wissenschaftlicher Mitarbeiter im Fachbereich 2 (Sozialpädagogik) an der Universität - Gesamthochschule Siegen, Adolf-Reichwein-Str. 2, D-5900 Siegen.



von den Kovariablen modifiziert. Modelle von der Art (1.1) lassen sich auch in der vertraueneren Regressionsdarstellung als lineares Transformationsmodell $h(T) = x'\beta + \varepsilon$ formulieren, wobei h eine Transformation der Verweildauer T darstellt. (s. **Doksum & Gasko** 1990). Einen Überblick über stetige Verweildauermodelle liefern **Kalbfleisch & Prentice** (1980) sowie **Biossfeld, Mayer, Hamerle** (1989).

In vielen Problemstellungen läßt sich Verweildauer nur diskret messen. So lassen sich bestimmte Ereignisse einer Biographie häufig nur dem Jahr oder dem Monat nach datieren. Die exakte stetige Zeit ist nicht beobachtbar. Tritt für mehrere Beobachtungen dasselbe Ende der Verweildauer auf, führt das zu Problemen in einem stetigen Modell wie (1.1), da dieser Fall die Wahrscheinlichkeit Null besitzt. In derartigen Fällen ist es sinnvoller diskrete Hazardratenmodelle zu betrachten.

2. Diskrete Hazardraten-Modelle

Die Zeitachse sei zerlegt in die $q + 1$ Intervalle

$$[\alpha_0, \alpha_1), [\alpha_1, \alpha_2), \dots, [\alpha_{q-1}, \alpha_q), [\alpha_q, \infty),$$

wobei meist $\alpha_0 = 0$ als Anfangszeitpunkt gewählt wird. Die Verweildauer T ist nun eine Zufallsvariable mit den diskreten Werten $1, \dots, q+1$, wobei $T = t$ eine Zustandsänderung im Intervall $[\alpha_{t-1}, \alpha_t)$ bezeichnet. Die entsprechende *diskrete Hazard- oder Intensitätsfunktion* bei gegebenem Kovariablenvektor $\mathbf{x}' = (x_1, \dots, x_p)$ ist gegeben durch

$$\lambda(t|\mathbf{x}) = P(T = t | T \leq t, \mathbf{x}).$$

Die Hazardrate entspricht der Wahrscheinlichkeit eines Zustandswechsels zum Zeitpunkt t , gegeben das Intervall $[\alpha_{t-1}, \alpha_t)$ wird zumindest erreicht. Eine wichtige Kenngröße bei Verweildaueranalysen ist die *Survivorfunktion* $S(t|\mathbf{x}) = P(T \geq t|\mathbf{x})$, die die Wahrscheinlichkeit angibt, daß das Intervall $[\alpha_{t-1}, \alpha_t)$ erreicht wird. Der Zusammenhang zur Hazardfunktion ergibt sich durch

$$S(t|\mathbf{x}) = \prod_{k=1}^{t-1} (1 - \lambda(k|\mathbf{x})).$$

Als Wahrscheinlichkeit für einen Zustandsübergang im Intervall $[\alpha_{t-1}, \alpha_t)$ erhält man

$$P(T = t|\mathbf{x}) = \lambda(t|\mathbf{x}) \prod_{k=1}^{t-1} (1 - \lambda(k|\mathbf{x})).$$



2.1 Diskretisierung stetiger Modelle

Ein Zugang zur Parametrisierung des diskreten Hazards führt über die Diskretisierung von Modellen, die für stetige Zeit entwickelt wurden. Ausgehend vom stetigen Cox-Modell (1.1) führt die Annahme diskret erhobener Zeit zum diskreten Modell

$$\lambda(t|\mathbf{x}) = 1 - \exp(-\exp(\beta_{0t} + \mathbf{x}'\beta)), \tag{2.1}$$

wobei die Koeffizienten

$$\beta_{0t} = \int_{\alpha_{t-1}}^{\alpha_t} \lambda_0(u) du$$

die diskrete Form der Baseline-Hazardrate darstellen, die sich aus der stetigen Baseline-Hazardrate λ_0 des Cox-Modells ergibt. Zur Ableitung siehe **Kalbfleisch & Prentice** (1973). Das Verfahren der Diskretisierung läßt sich im Prinzip bei allen stetigen Verweildauermodellen durchführen. Allerdings sind die resultierenden diskreten Modelle meist nicht von einer mit (2.1) vergleichbar einfachen Form.

2.2 Sequentielle dichotome Modelle

Ein alternativer Ansatz führt über die vertrauten Regressionsmodelle für dichotome abhängige Variable. Man betrachtet *lokal* den möglichen Übergang vom Intervall $[\alpha_{t-1}, \alpha_t)$ zu $[\alpha_t, \alpha_{t+1})$. Die Übergangsvariable Y_t sei bestimmt durch

$$Y_t = \begin{cases} 1 & T = t | T \geq t \\ 0 & T > t | T \geq t. \end{cases}$$

Dabei bezeichnet $Y_t = 1$ das Ende des Prozesses (Zustandsänderung) und $Y_t = 0$ das Fortschreiten des Prozesses (keine Zustandsänderung) gegeben der Zustand besteht im Intervall $[\alpha_{t-1}, \alpha_t)$ noch. Als einfache Beziehung zwischen bedingter Übergangswahrscheinlichkeit und Hazardrate erhält man

$$P(Y_t = 1|\mathbf{x}) = \lambda(t|\mathbf{x}).$$

Die sequentielle Modellierung geht davon aus, daß der lokale Übergang Y_1, Y_2, \dots solange beobachtet wird bis ein Zustandswechsel eintritt. Der Mechanismus der auch als *continuation ratio-Modelle* (**Agresti** 1984) bezeichneten Modellklasse wird in **Tutz** (1991a) ausführlich behandelt.

Für die lokale dichotome Responsevariable gibt es eine Vielzahl von möglichen Parametrisierungen (z.B. **McCullagh & Neider** 1989, **Cox & Snell** 1989). Ein allgemeines binäres Regressionsmodell ist von der Form

$$P(Y_t = 1|\mathbf{x}) = F(\beta_{0t} + \mathbf{x}'\beta_t),$$

wobei F eine Verteilungsfunktion darstellt. Beachtenswert ist, daß der Gewichtsvektor in dieser Form von der Zeit t abhängt, d.h. der Einfluß der Kovariablen läßt sich problemlos als zeitabhängig modellieren. Für die logistische Verteilungsfunktion $F(u) = 1/(1 + \exp(-u))$ erhält man unmittelbar das *logistische Hazardratenmodell*

$$\lambda(t|x) = \exp(\beta_{0t} + x'\beta_t)/(1 + \exp(\beta_{0t} + x'\beta_t)). \quad (2.2)$$

Für die Extremwertverteilung $F(u) = 1 - \exp(-\exp(u))$ erhält man unmittelbar mit

$$\lambda(t|x) = 1 - \exp(-\exp(\beta_{0t} + x'\beta_t)) \quad (2.3)$$

eine Verallgemeinerung des diskreten Cox-Modells (2.1). Eine weitere Verallgemeinerung erhält man durch eine von **Aranda-Ordaz** (1983) benutzte Verteilungsfunktion

$$F_\alpha(u) = 1 - \exp(-(1 + \alpha u)^{1/\alpha}),$$

die von einem zusätzlichen Parameter $\alpha > 0$ abhängt. Als Grenzfall $\alpha \rightarrow 0$ ergibt sich daraus das diskrete Cox-Modell (2.3).

Die Zahl der für diese Modellklasse zu schätzenden Parameter hängt von der Anzahl der betrachteten Intervalle ab. Ist diese Anzahl im Verhältnis zum Stichprobenumfang gering, ist die Maximum Likelihood-Methode problemlos anwendbar. Für eine große Anzahl von Intervallen empfiehlt es sich mit $\beta_t = \beta$ von einer globalen, nicht intervall-spezifischen Wirkungsweise der Kovariablen auszugehen. Zu schätzen sind jedoch die der Baseline-Hazardrate entsprechenden Parameter $\beta_{01}, \dots, \beta_{0q}$. Um deren Zahl zu beschränken, läßt sich ein polynomialer Ansatz verwenden, wie ihn **Mantel & Hankey** (1978) für das logistische Modell vorschlagen. Mit

$$\beta_{0t} = \sum_{i=0}^r \gamma_i t^i, \quad r < q, \quad (2.4)$$

wird die Anzahl der zu schätzenden Koeffizienten $\gamma_0, \dots, \gamma_r$ auf $r + 1$ herabgesetzt. Der Grad des gewählten Polynoms bestimmt die Flexibilität des Baseline-Hazards. Efron (1988) betrachtet alternativ eine Spline-Lösung mit kubischen Splines der Form

$$\beta_{0t} = \gamma_0 + \gamma_1 t + \gamma_2 (t - c)_-^2 + \gamma_3 (t - c)_-^3, \quad (2.5)$$

wobei c eine Schranke darstellt und $(t - c)_- = \min\{0, t - c\}$. Dadurch wird erreicht, daß bis zur Schranke c ein Polynom dritten Grades und über der Schranke c ein linearer Ansatz Anwendung findet. Der Grund für größeres t linear zu parametrisieren liegt in der mit wachsender Zeit erfolgenden Ausdünnung der Daten. Einen interessanten Ansatz zur Beschränkung der Parameterzahl schlägt **Fahrmeir** (1991) vor. Basierend auf einer Verallgemeinerung des Kaimanfilters werden die Parameter β_{0t} durch einen



stochastischen Prozeß beschrieben und durch Filtern und Glätten hinreichend flexibel geschätzt. Nonparametrische Schätzverfahren auf der Basis von Kernglättern werden in Tutz (1991b) entwickelt.

3. Schätzung mit dem Programmsystem GLAMOUR

GLAMOUR (Fahrmeir et.al. 1990) ist ein in Regensburg entwickeltes Programmsystem zur Schätzung generalisierter linearer Modelle. Als interaktives menügesteuertes Programm, das auf PCs unter MS-DOS läuft, ist es besonders einfach handhabbar. Der gesamte Ablauf (Spezifikation der Einflußgrößen, gewünschtes Modell, Parameterisierung der Baseline-Hazardrate, Parameter- und Anpassungstests) wird durch Masken gesteuert und durch Information unterstützt.

Die Schätzung beruht auf dem Maximum Likelihood-Prinzip und der Einbettung der diskreten Hazardrate in die verallgemeinerten linearen Modelle (siehe Hamerle & Tutz 1989). Die Daten bei Verweildauermodellen besitzen die Form (t_i, x_i, δ_i) , wobei zum beobachteten Zeitpunkt t_i und dem Kovariablenvektor x_i noch der Zensuringsindikator δ_i hinzukommt. Durch

$$\delta_i = \begin{cases} 1 & \text{wenn Zustandswechsel beobachtet} \\ 0 & \text{wenn zum Zeitpunkt } t_i \text{ der Zustand noch besteht} \end{cases}$$

wird kodiert, ob ein Zustandswechsel tatsächlich beobachtet wird oder ob das beobachtete Individuum nur bis zum Zeitpunkt t_i beobachtet wurde ohne daß ein Zustandswechsel auftrat, d.h. daß sog. Rechtszensurierung auftrat. Der Likelihoodbeitrag der iten Beobachtung läßt sich dann durch

$$L_i = \lambda(t_i|x_i)^{\delta_i} \prod_{k=1}^{t_i-1} (1 - \lambda(k|x_i))$$

beschreiben. Details sind bei Hamerle & Tutz (1989) zu finden.

In GLAMOUR unmittelbar abrufbar ist die Schätzung für das logistische Modell, das diskrete Cox-Modell und die Erweiterung nach Aranda-Ordaz (1983). Verfügbar sind sowohl der polynomiale Ansatz (2.4) als auch der Spline-Ansatz (2.5). GLAMOUR kodiert automatisch kategoriale Variable in Dummy-Variablen um (z.B. Hamerle, Kemeny, Tutz 1984) und liefert Signifikanztests mit den zugehörigen Überschreitungswahrscheinlichkeiten. Ein besonderer Vorteil ist die automatisierte Variablenselektion, die als Vorwärts- oder Rückwärtsselektion anwählbar ist. Damit wird, basierend auf Teststatistiken (Fahrmeir & Frost 1989), die Anzahl der zur Verfügung stehenden Einflußgrößen auf die relevanten reduziert. Für Verweildauern von Interesse ist der Sterbefatelschätzer, der eine erste Veranschaulichung der Daten ermöglicht. Verfügbare Graphiken für Hazardrate und Survivorfunktion illustrieren die Modelle und erleichtern die Interpretation

von Ergebnissen. Details der Analysemöglichkeiten finden sich im GLAMOUR-Handbuch (Fahrmeir et al 1990).

4. Empirisches Beispiel eines gruppierten Cox-Modells: Ein Ländervergleich zwischen Westdeutschland und Ungarn hinsichtlich des Zeitpunkts des ersten sexuellen Kontakts

4.1 Daten

In der Shell-Jugendstudie 1985 (ZA Nr. 1438) (vgl. Jugendwerk der Deutschen Shell 1985) beschäftigte sich eine Frage mit der retrospektiven altersbezogenen Datierung von bedeutsamen Lebensereignissen. Ein für den Übergang von der Kindheit zur Jugend zentrales Lebensereignis, an dem sich gesellschaftliche Modernisierungsprozesse im Hinblick auf die Jugendphase gut nachweisen lassen, ist der Zeitpunkt des ersten sexuellen Kontakts (vgl. Fuchs 1985, Molnar 1991).

Hierbei hat in der Jugendforschung die Unterscheidung zwischen einer unterschichtspezifischen Jugendbiografie mit frühen sexuellen Erfahrungen und einer mittelschichtstypischen Ausformung, bei der die Aufgabe des Bildungserwerbs mit Triebaufschub („deferred gratification pattern“) einhergeht, eine bis in die 20er Jahre zurückreichende Tradition (vgl. Bernfeld 1973, Roth 1961, Spranger 1966). Diesem klassischen, dichotomen Modell von Jugendbiografie kann eine modernisierte Variante an die Seite gestellt werden, derzufolge ein gesellschaftlicher Individualisierungsschub (vgl. Beck 1983, 1986) seit den 60er Jahren zu Entstrukturierung des Lebenslaufs führte (vgl. Kohli 1986, Fuchs 1983, 1985), die sich in einer Entsynchronisierung der Sequenzen von Lebensereignissen und einer Lockerung des Zusammenhangs zwischen sozialer Schichtzugehörigkeit und Lebenslauf ausdrückt. Für interkulturelle Jugendforschung ist nun eine empirische Prüfung der Gültigkeit des klassischen oder modernisierten Modells von Jugendbiografie von Interesse, zumal beide nur über eine schwache empirische Fundierung (insbesondere auf der Ebene von Massenumfragen) verfügen (vgl. Fuchs 1985).

Im Jahr 1985 replizierte eine ungarische Forschungsgruppe (vgl. Molnar 1991, Gabor 1991) die Shell-Jugendstudie in wesentlichen Teilen. Diese Replikation ermöglicht einen Vergleich der Datierung erster sexueller Kontakte in beiden Gesellschaften in Abhängigkeit von relevanten sozio-demografischen Variablen.

Bei der methodischen Umsetzung dieses Vorhabens stößt man jedoch auf das Problem, daß noch nicht alle Befragten sexuelle Kontakte hatten. Eine Möglichkeit, die auch von Fuchs (1985) hinsichtlich der Datenanalyse der Shell-Jugendstudie besprochen wird, ist nun eine Begrenzung auf die Gruppe, bei der das jeweilige Lebensereignis bereits eingetreten ist. Dieses Vorhaben beinhaltet jedoch, wie Fuchs selbst feststellt, ein Problem:

„Bei aller Einfachheit gegenüber Schätzverfahren für zensierte Daten hat das hier gewählte Vorgehen auch Nachteile: Mit der Beschränkung auf die Teilstichprobe von 577 Jugendlichen im Alter zwischen 21 und 24 Jahren geht ein Informationsverlust einher; die Differenzierung in Untergruppen stößt rasch an die Grenze kleiner Zahlen. Und indem wir statt mit der Gesamtstichprobe der Jugendlichen mit dieser Teilstichprobe arbeiten, unterstellen wir .. Strukturgleichheit der Teilstichprobe mit der Gesamtstichprobe der **Jugendlichen**" (Fuchs 1985, S. 201).

Jugendliche, für die das Zielereignis noch nicht eingetreten ist, können als rechtszensierte Beobachtungen in die Verweildaueranalyse eingehen. Um der Zensierung der Daten Rechnung zu tragen, könnte man auf das klassische Cox-Modell ausweichen (vgl. **Kalbfleisch & Prentice** 1980, **Blossfeld, Hamerle, Mayer** 1986, **Diekmann & Mitter** 1985, **Andreß** 1985), jedoch liegt im vorliegenden Fall eine Verletzung der Modellannahmen vor, denn die Datierung erfolgt nach Jahren, was zu einer Häufung von Ereignissen zu einem Zeitpunkt führt; das klassische Cox-Modell unterstellt jedoch, daß nur ausnahmsweise Ereignishäufungen („ties“) zu einem Zeitpunkt stattfinden. Beide Probleme, das der Zensierung und das der Ereignishäufungen können jedoch durch das gruppierte Cox-Modell gelöst werden.

4.2 Verwendete Variablen

Jugend unterliegt im gesamteuropäischen Maßstab seit Ende der 50er Jahre einem Modernisierungsschub, der sich zentral auf den Wechsel der Kontrollinstitutionen bezieht (vgl. **Zinnecker** 1991). Waren diese in den 50er Jahren noch durch die Arbeit, die Familie und das soziale Milieu der Herkunftsgruppe beschreibbar, so fand seitdem ein Übergang in Richtung auf Bildungsinstitutionen, Gruppen von Gleichaltrigen und marktbezogene Kontrollinstitutionen statt. Einher mit diesem Wechsel ging eine größere kulturelle Autonomie der Jugendphase und ihre zeitliche Verlängerung. Dieser Modernisierungsschub wirkte sich nicht ungebrochen auf die osteuropäischen Länder aus. Vielmehr kann man hier von einer selektiven Modernisierung ausgehen (vgl. **Zinnecker** 1991), die sich vor allem auf den Arbeitsbereich und den Bildungserwerb bezog und jene Individualisierungsprozesse, die Teil der marktbezogenen Modernisierung in Westeuropa waren, ausklammerte. Hinsichtlich der oben diskutierten Modelle der Jugendbiografie kann deshalb für Ungarn von einer stärkeren Gültigkeit des oben beschriebenen klassischen Modells ausgegangen werden, das Bildungsunterschiede für den Zeitpunkt des ersten sexuellen Kontakts unterstellt, während für Westdeutschland eine Angleichung in dieser Hinsicht angenommen werden kann.

Für den Vergleich wurden zwei gruppierte Cox-Modelle angepaßt, die folgende soziodemografischen Variablen enthielten: Geschlecht, Bildung der Eltern, Bildung der Befragten (niedrig/mittel/hoch) und Wohnortgröße (Dorf, Kleinstadt, mittlere Stadtgröße, Großstadt). Um die Tragfähigkeit der Interpretation dieser Modelle angemessen einschätzen

zu können, muß man sich vergegenwärtigen, daß das Alter des ersten sexuellen Kontakts retrospektiv erfragt wurde. Während diese Tatsache für das Geschlecht und die Bildung der Eltern keine Probleme aufwirft, gehen wir bei der Bildung des Befragten implizit davon aus, daß er im Alter des ersten sexuellen Kontakts den gleichen Schultyp besuchte, den er auch zum Erhebungszeitpunkt besuchte (was relativ unproblematisch ist, da der Wechsel zur Sekundarstufe im Alter von zehn Jahren erfolgt) und daß er zu beiden Zeitpunkten in einem Wohnort gleicher Größe wohnte (was sich nicht überprüfen läßt und sicher nur für einen Teil der Befragten zutrifft).

4.3 Ergebnisse

Alle Modelle wurden mit dem Programm „GLAMOUR“ (vgl. Fahrmeir et al. 1990) berechnet, das neben diskreten Modellen der Ereignisdatenanalyse eine größere Anzahl von Untermodellen des Verallgemeinerten Linearen Modells umfaßt. Tabelle 1 enthält die Parameterschätzungen für Westdeutschland, wobei die Kovariablen in Effektkodierung vorliegen, d.h. die Effekte summieren sich bei jeder Variable zu Null, angegeben sind nur die Effekte bis zur vorletzten Kategorie.

Die Modelle wurden jeweils mit einem polynomialen Ansatz dritten Grades geschätzt. Die diskrete Zeitdauer in den Ausprägungen 1,2,...,9 entspricht dem Ereignis „erster sexueller Kontakt“ im Alter von 12,13,...,20 und älter. Neben den Parameterschätzungen für die Kovariablen werden auch die geschätzten Gewichte des Polynoms nach (2.4) angegeben, wobei sich zeigt, daß der Koeffizient für die dritte Potenz für Deutschland wenig relevant ist.

Tabelle 2 enthält die entsprechenden Schätzungen für Ungarn, wobei hier die acht Ausprägungen der Zeitdauer dem Lebensalter 12,13,...,19 und älter entsprechen. Die hinsichtlich der letzten Kategorie von Deutschland abweichende Kategorienbildung ergab sich aus der empirischen Verteilung.

Im Gegensatz zu unseren hypothetischen Annahmen beschleunigt in beiden Ländern eine niedrige bzw. mittlere Bildung der Befragten den Prozeß. Während in Ungarn eine niedrige Bildung der Mutter einen reduzierenden Effekt auf die Hazard-Rate hat, beeinflußt eine mittlere Bildung die Rate in positiver Hinsicht. In Westdeutschland hat hingegen eine mittlere Bildung der Mutter eine verringernde Wirkung auf die Übergangswahrscheinlichkeit.

Geschlechtsunterschiede, die in Ungarn auf eine frühere Bereitschaft der Jungen zu sexuellen Kontakten verweisen, erweisen sich im früheren Westdeutschland als nicht signifikant. Ortseffekte haben tendenziell in beiden Ländern die gleiche Wirkung: Je größer der Wohnort, desto wahrscheinlicher sind frühere sexuelle Kontakte. Allerdings ist der Effekt für Ungarn ausgesprochen schwach ausgeprägt. Die Bildung des Vaters hat weder in Westdeutschland noch in Ungarn eine nachweisbare Wirkung.



Bei einer Rückwärts-Selektion der Variablen auf dem 5% -Signifikanzniveau wurde übereinstimmend in beiden Ländern die Bildung des Vaters ausgeschlossen. In Westdeutschland wurde zusätzlich das Geschlecht entfernt, während in Ungarn die Wohnortgröße ausgeschlossen wurde. Der Vergleichbarkeit wegen werden die Schätzungen in Tabelle 1 und 2 für das Grundmodell angegeben.

Tabelle 1: Gruppirtes Cox-Modell für das frühere Westdeutschland

Parameter-Name	Schätzung	Überschreitungs- wahrscheinlichkeit
1	-5.342	0.000
POLY [1]	1.525	0.000
POLY [2]	-0.139	0.014
POLY [3]	0.003	0.320
Geschlecht (männlich)	-0.011	0.716
Bildung Vater (niedrig)	-0.037	0.477
Bildung Vater (mittel)	-0.037	0.477
Bildung Mutter (niedrig)	-0.028	0.635
Bildung Mutter (mittel)	-0.126	0.030
Ort (Dorf)	-0.127	0.013
Ort (Kleinstadt)	-0.071	0.232
Ort (mittl. Stadt)	0.174	0.010
Bildung Befragter (niedrig)	-0.012	0.781
Bildung Befragter (mittel)	0.108	0.015

Die Gemeinsamkeiten zwischen Westdeutschland und Ungarn liegen also insbesondere im Bereich der Bildung der Jugendlichen und ihrer Mütter: Während in Ungarn eine niedrige und in Westdeutschland eine mittlere Bildung der Mutter eine deutlich verlangsamende Wirkung auf den Prozeß hat, wird dieser, insbesondere in Ungarn, aber auch in Westdeutschland, durch eine niedrige oder mittlere Bildung des Befragten beschleunigt. Der zentrale Unterschied ist in der Existenz geschlechtsspezifischer Unterschiede in Ungarn und einer diesbezüglichen Angleichung in Westdeutschland zu sehen.

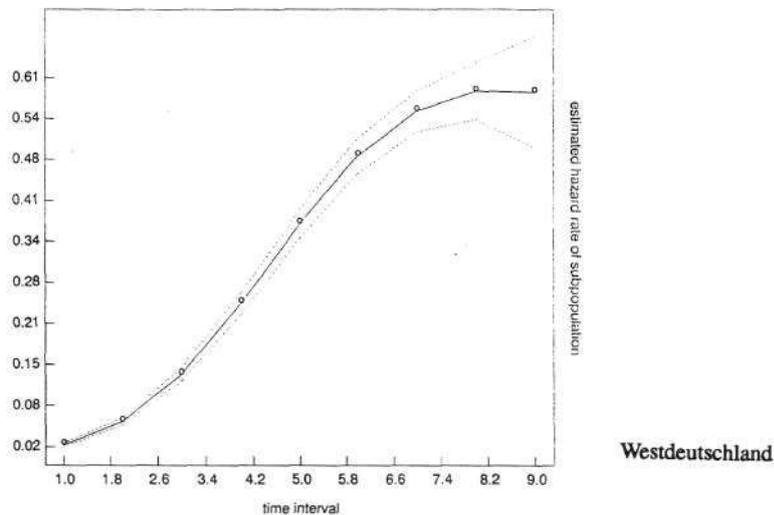
Unterschiede der Länder zeigen sich hinsichtlich der Form der Hazardraten. Dieser Effekt wird deutlich aus den Sterbetafelschätzern ebenso wie bei der Betrachtung von Subpopulationen. Abbildung 1 zeigt die geschätzten Hazardraten der beiden Länder für ein männliches Individuum mit niedriger Bildung, dessen Eltern beide niedrige Bildung besitzen und das in einer Großstadt lebt. Die gepunkteten Linien geben die von GLAMOUR berechnete Schätzgenauigkeit wieder, ausgedrückt durch Standardabweichungen nach unten und oben. Die Hazardrate im früheren Westdeutschland zeigt einen wachsenden Verlauf, der spätestens mit dem neunzehnten Lebensjahr abflacht. In Ungarn hingegen

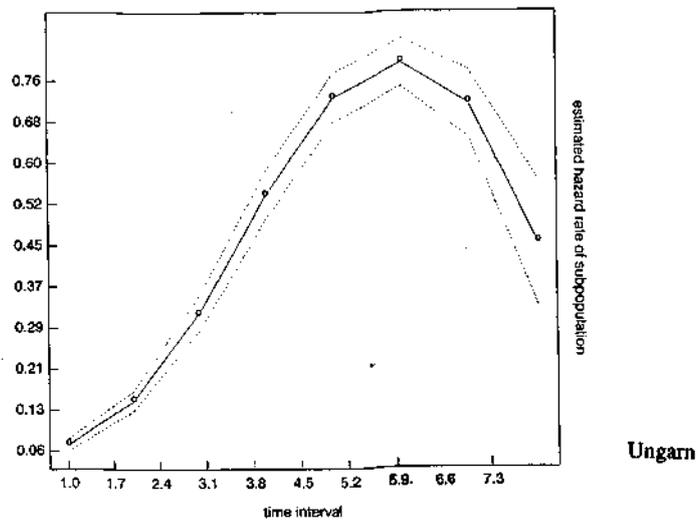
kulminiert die wachsende Hazardrate im siebzehnten Lebensjahr und zeigt dann ein deutliches Abfallen, d.h. die Wahrscheinlichkeit der ersten Kontaktnahme sinkt, wenn sie bis zu diesem Lebensalter nicht stattgefunden hat.

Tabelle 2: Gruppieretes Cox-Modell für Ungarn

Parameter-Name	Schätzung	Überschreitungs- wahrscheinlichkeit
1	-3.999	0.000
POLY [1]	0.677	0.031
POLY [2]	0.094	0.266
POLY [3]	-0.016	0.016
Geschlecht (männlich)	0.191	0.000
Bildung Vater (niedrig)	0.081	0.215
Bildung Vater (mittel)	0.015	0.757
Bildung Mutter (niedrig)	-0.177	0.004
Bildung Mutter (mittel)	0.096	0.078
Ort (Dorf)	-0.075	0.192
Ort (Kleinstadt)	-0.034	0.625
Ort (mittl. Stadt)	-0.034	0.621
Bildung Befragter (niedrig)	0.360	0.000
Bildung Befragter (mittel)	0.059	0.258

Abb. 1: Geschätzte Hazardraten für das frühere Westdeutschland und Ungarn





5. Schlußbemerkung

In der quantitativen Biografieforschung gibt der Zeitpunkt und die Sequenz spezifischer Lebensereignisse Auskunft über gesellschaftliche Modernisierungsprozesse. Im europäischen Maßstab geht der Übergang von industriellen zu post-industriellen Gesellschaften einher mit einer Akzeleration im Hinblick auf die sexuelle Entwicklung, eine Verlängerung der Jugend und eine Entsynchronisierung dessen, was für industrielle Gesellschaften als „Normalbiografie“ (Kohli 1986,1988), als klar definierte, sequentielle Abfolge von Lebensereignissen, bezeichnet wurde.

Retrospektive Erhebungen von Lebensereignissen im internationalen Vergleich stoßen jedoch häufig auf methodische Probleme, die in der Zensierung der Daten und der Art der Datierung zu sehen sind. Differenziert man etwa die Datierung relativ „weicher“ Lebensereignisse im Hinblick auf Monate und Jahre, läuft man Gefahr, zu Artefakten zu kommen, die nicht der wirklichen Gedächtnisleistung entsprechen. Häufig wird man deshalb zu einer Datierung nach Lebensalter kommen, die wegen der Häufung von Ereignissen während eines Intervalls eine Schätzung nach dem klassischen Cox-Modell nicht mehr unverzerrt ermöglicht. In diesem Fall kann das gruppierte Cox-Modell als Verfahren der „weichen“ Datenstruktur und dem doppelten Dilemma der Zensierung und Ereignishäufung gerecht werden.

**Literatur**

- Agresti, A.** (1984): Analysis of ordinal categorical data. New York: Wiley.
- Andreß, H.-J.** (1985): Multivariate Analyse von Verlaufsdaten. Mannheim
- Aranda-Ordaz, F. J.** (1983): An extension of proportional-hazards-model for grouped data. *Biometrics* 39, 110-118.
- Beck, U.** (1983): Jenseits von Stand und Klasse? In: Kreckel (Hrsg.): Soziale Ungleichheiten. Soziale Welt Sonderband 2. 35-74.
- Beck, U.** (1986): Risikogesellschaft. Frankfurt am Main.
- Bernfeld, S.** (1923): Über eine typische Form der männlichen Pubertät. *Imago* 9, 169-188.
- Blossfeld, H.-P., Hamerle, A., Mayer, K. U.** (1986): Ereignisanalyse. Frankfurt: Campus-Verlag.
- Cox, D.R.** (1972): Regression models and life tables (with discussion). *J.Roy. Statist. Soc. B*, 34, 187-220.
- Cox, D.R., Snell, E.J.** (1989): Analysis of binary data. London: Chapman & Hall.
- Diekmann, A., Mitter, P.** (1984): Methoden zur Analyse von Zeitverläufen. Stuttgart.
- Doksum, K.A., Gasko, M.** (1990): On a correspondence between models in binary regression analysis and in survival analysis. *International Statistical Review* 58, 243-252.
- Efron, B.** (1988): Logistic regression, survival analysis and the Kaplan-Meier curve. *J. Am. Stat. Ass.* 414-425.
- Fahrmeir, L.** (1991): Extended Kalman filtering for dynamic generalized linear models and survival data. *JASA* (to appear).
- Fahrmeir, L., Frost, H.** (1989): Stepwise model and variable selection by Wald and score tests in non-normal regression and time series models. *Regensburger Beiträge*.
- Fahrmeir, L., Frost, H., Hennevogel, W., Kaufmann, H., Kranert, T., Tutz, G.** (1990): GLAMOUR-User and Example Guide. University of Regensburg.
- Fuchs, W.** (1983): Jugendliche Statuspassage oder individualisierte Jugendbiographie? In: *Soziale Welt* 34, 341-371.
- Fuchs, W.** (1985): Jugend als Lebenslaufphase. In: *Jugendwerk der Deutschen Shell* (Hrsg.): Jugendliche und Erwachsene 1985. Generationen im Vergleich, Bd. 1. Leverkusen, 195-264.



- Gabor, K.** (1991): Jugendkultur und gesellschaftliche Orientierungsmuster in Ungarn. Ein Vergleich zwischen Oberschülern, Berufsschülern und Studenten. In: W. Melzer, et. al. (Hrsg.): Osteuropäische Jugend im Wandel. Weinheim und München. 161-169.
- Hamerle, A., Kemeny, P., Tutz, G.** (1984): Kategoriale Regression In: L. Fahrmeir, A. Hamerle (Hrsg.), Multivariate statistische Verfahren. Berlin: De Gruyter.
- Hamerle, A., Tutz, G.** (1989): Diskrete Modelle zur Analyse von Verweildauer und Lebenszeiten. Campus, Frankfurt.
- Jugendwerk der Deutschen Shell** (Hrsg.) (1985): Jugendliche und Erwachsene 1985. Generationen im Vergleich. 5 Bände. Leverkusen
- Kalbfleisch, J.D., Prentice, R. L.** (1980): The Statistical analysis of failure time data. New York.
- Kohli, M.** (1986): Gesellschaftszeit und Lebenszeit. Der Lebenslauf im Strukturwandel der Moderne. In: J. Berger (Hrsg.): Die Moderne - Kontinuitäten und Zäsuren. Soziale Welt. Sonderband 4, 183-208.
- Kohli, M.** (1988): Normalbiographie und Individuum. Zur institutionellen Dynamik des gegenwärtigen Lebenslaufsregimes. In: H.G. Brose, B. Hildebrand (Hrsg.): Vom Ende des Individuums zur Individualität ohne Ende. 33-54. Opladen.
- Mantel, N., Hankey, B. F.** (1978): A logistic regression analysis of responsetime data where the hazard function is time dependent. Comm. Statist. A7, 333-347.
- McCullagh, P., Neider, J.** (1989): Generalized linear models (second edition). London: Chapman and Hall.
- Molnar, P.** (1991): Jugend als Statuspassage in verschiedenen Geburtskohorten in Ungarn und Westdeutschland. In: W. Melzer et. al. (Hrsg.): Osteuropäische Jugend im Wandel. Weinheim und München. 137-148.
- Roth, H.** (1961): Jugend und Schule zwischen Reform und Restauration. Berlin
- Spranger, E.** (1966): Psychologie des Jugendalters. Heidelberg
- Tutz, G.** (1991a): Sequential models in categorical regression. Computational Statistics & Data Analysis 11, 275-295.
- Tutz, G.** (1991b): Glättung mit diskreten Daten. In: G. Seeber, Ch. Minder (Hrsg.), Multivariate Modelle, 54-88. Heidelberg: Springer Verlag.
- Zinnecker, J.** (1987): Jugendkultur zwischen 1940 und 1985. Opladen