Soziale Bedingungen der Mortalität

Methodenbezogene Überlegungen anhand des SOEP

Götz Rohwer

Wolfgang Voges

Götz Rohwer, Universität Bremen, FB 8 (EMPAS), Postfach 330440, 28334 Bremen. Wolfgang Voges. Zentrum für Sozialpolitik. Parkallee 39, 28209 Bremen.

1 Einleitung

Die Vermutung, daß es soziale Bedingungen der Mortalität gibt, erscheint unmittelbar plausibel. Versucht man jedoch, dieser Vermutung mithilfe von für die Sozialforschung verfügbaren Daten empirisch nachzugehen, treten eine Reihe von Schwierigkeiten auf.

1. Zunächst ist bereits unklar, was genau mit sozialen Bedingungen der Mortalität gemeint sein soll. Dies betrifft sowohl die Indikatoren zur Erfassung potentiell relevanter Lebensbedingungen als auch und vor allem die Frage, mit welchen theoretischen Überlegungen ein möglicher Zusammenhang zwischen sozial definierten Lebensumständen und biologisch definierter Mortalität konzipiert werden kann.
2. Ein zweites Problem betrifft die verfügbaren Daten. Um der Frage nach möglichen sozialen Bedingungen von Mortalität nachgehen zu können, sind offenbar Längsschnittdaten erforderlich. Denn die Hypothese zielt, ganz unabhängig von ihrer Detailformulierung, darauf, daß das Sterblichkeitsrisiko von Menschen auch davon abhängt, wie sich ihr jeweils bisheriger Lebensverlauf entwickelt hat. Aber darüber hinaus sind Längsschnittdaten eines besonderen Typs erforderlich. Längsschnittdaten, die durch einmalige Retrospektivbefragungen gewonnen werden, sind offenbar unzureichend. Man benötigt vielmehr Daten, die sowohl Informationen über Sterbefälle als auch über die Lebensverläufe der gestorbenen Personen liefern. Der einzige zur Zeit verfügbare Datensatz, der diesen Anforderungen genügt, ist das Sozioökonomische Panel (SOEP), den wir deshalb in dieser Untersuchung zugrundelegen.
3. Ein drittes Problem betrifft die Beschaffenheit der statistischen Modelle, um Vermutungen über Zusammenhänge zwischen sozialen Lebensumständen und Mortalität empirisch zugänglich zu machen. Offenbar muß berücksichtigt werden, daß uns nur ein – gemessen an der durchschnittlichen Lebensdauer – sehr begrenzter Zeitraum zur Verfügung steht, in dem wir Sterbefälle beobachten können. Das SOEP liefert zur Zeit ein Beobachtungsfenster für 11 Jahre (1984 – 1994). Ein solches Beobachtungsfenster ist zum Beispiel ausreichend, um Arbeitslosigkeitsepisoden zu untersuchen; denn wir können innerhalb des gegebenen zeitlichen Rahmens sowohl Übergänge in die Arbeitslosigkeit als auch die Dauer der Arbeitslosigkeit und die schließlichen Übergänge in andere Folgezustände beobachten. Bei der Frage nach sozialen Bedingungen der Mortalität verhält es sich jedoch anders. Sowohl das Sterblichkeitsrisiko als auch die Herausbildung von Bedingungen der Sterblichkeit beginnt bei jedem Menschen mit der Geburt. Das verfügbare Beobachtungsfenster reicht jedoch nicht aus, um diesen Prozeß vollständig, d.h. beginnend mit der Geburt, beobachten zu können. Die verfügbaren Daten sind also „links abgeschnitten", und diese muß bei der Modellkonstruktion berücksichtigt werden.

Der weitere Text beschäftigt sich im wesentlichen mit diesen drei Problemen.

2 Todesursachen und soziale Bedingungen

Aus soziologischer bzw. sozialpsychologischer Sicht interessieren wir uns nicht unmittelbar für Todesursachen, sondern dafür, ob und ggf. wie ihr Vorhandensein und Wirksamwerden auch von sozialen Bedingungen abhängt. Dieser Unterscheidung zwischen potentiellen Todesursachen und sozialen Bedingungen für das Vorhandensein und Wirksamwerden potentieller Todesursachen kommt, wie wir glauben, für soziologische Mortalitätsuntersuchungen eine grundlegende Bedeutung zu. Sterben ist ein biologischer Vorgang; und jeder Versuch, Ursachen für diesen Vorgang herauszufinden, muß deshalb auf einer biologischen Ebene und ggf. bei äußeren Einwirkungen auf einem Sterblichkeitsrisiko ausgesetzten Organismen einsetzen.

Soziologische bzw. sozialpsychologische Hypothesen können demgegenüber nur bei der Frage ansetzen, wie gewissermaßen die Teilhabe der Menschen an potentiellen Todesursachen durch ihre sozialen Lebensumstände beeinflußt wird. Jeder Versuch, darüber spezifische Hypothesen zu formulieren, muß sich deshalb der Frage stellen, warum Menschen überhaupt sterben. Nur insoweit man diese Frage ansatzweise beantworten kann, kann man dann Hypothesen darüber formulieren, wie Todesursachen durch soziale Bedingungen für vorausgegangene Lebensverläufe beeinflußt werden können.

Versucht man, zumindest eine grobe Vorstellung über mögliche Todesursachen zu gewinnen, kann man zunächst mit folgenden Unterscheidungen beginnen: Menschen können Selbstmord begehen; sie können durch andere Menschen getötet werden; sie können tödliche Unfälle erleiden; sie können an Krankheiten sterben; und sie können schließlich auch auf eine Weise sterben, die sich mit keiner der zuvor genannten Möglichkeiten erklären läßt. In allen diesen Fällen, mit Ausnahme des letzten, ist die Todesursache meistens nicht strittig; und im zuletzt genannten Fall kann man vielleicht von „natürlichen“ Todesursachen sprechen. Wie auch immer; das Feststellen von (potentiellen) Todesursachen ist ein medizinisches Problem und fällt jedenfalls nicht in den Zuständigkeitsbereich soziologischer Untersuchungen. Die grobe Unterscheidung, die wir vorgenommen haben, ist gleichwohl wichtig. Denn je nachdem, welche potentiellen Todesursachen ins Auge gefaßt werden, stellt sich die Frage nach möglichen sozialen Bedingungen unterschiedlich.

Akzeptiert man diese Betrachtungsweise unseres Themas, stellt sich die Frage, ob mit den uns verfügbaren Daten überhaupt sinnvolle Hypothesen über soziale Bedingungen von Mortalität untersucht werden können. Wie bereits in der Einleitung erwähnt wurde, ist das SOEP zur Zeit der einzige Längsschnittdatensatz, mit dem Mortalität untersucht werden kann. Dieser Datensatz enthält jedoch keinerlei Informationen über die tatsächlichen Todesursachen der gestorbenen Personen. Man weiß zwar von einigen Personen, die an dieser Panelbefragung teilgenommen haben, daß sie während des bisherigen Untersuchungszeitraums gestorben sind, und man hat auch zahlreiche Informationen über ihren jeweils vergangenen Lebenslauf; aber man weiß nicht, warum sie gestorben sind, an Krankheiten, Unfällen, durch Tötungsdelikte, oder einfach eines „natürlichen“ Todes.

Es ist infolgedessen nicht möglich, spezifische Hypothesen über die soziale Bedingtheit von Todesursachen zu untersuchen. Unsere Untersuchung muß deshalb bescheidener konzipiert werden und kann nur einen deskriptiven Anspruch haben. Wir können nicht anstreben, Mortalität kausal zu erklären, sondern nur versuchen aufzuzeigen, in welcher Weise sich das Mortalitätsverhalten bei unterschiedlichen sozialen Lebensumständen unterscheidet.

3 Informationen aus dem SOEP

Welche Informationen liefert uns das SOEP, um mögliche soziale Bedingungen von Mortalität zu untersuchen? Zur Verfügung stehen uns die ersten 10 Wellen, d.h. das Beobachtungsfenster umfaßt den Zeitraum 1984 bis 1993.[1] Wir kennen die Gesamtheit der sog. Stammpersonen (einschließlich Kinder), die in der ersten Welle 1984 beobachtet worden

[1] Das 1990 für die neuen Bundesländer begonnene SOEP wird hier nicht einbezogen.

Box 3.1 Basisinformationen aus der Teilstichprobe A des SOEP

	Insgesamt		Maenner		Frauen	
	Anzahl	%	Anahl	%	Anzahl	%
Teiln. 10 Wellen	4702	51.8	2203	50.9	2499	52.6
Verstorben	706	7.8	362	8.4	344	7.3
Umzug Ausland	84	0.9	36	0.8	48	1.0
Attritition	3584	39.5	1727	39.9	1857	39.1
Insgesamt	9076	100.0	4328	100.0	4748	100.0

sind. Während des dann folgenen Beobachtungszeitraums werden diese Personen älter; einige sterben, einige verlassen das Beobachtungsgebiet (Umzüge ins Ausland), und einige nehmen aus sonstigen Gründen nicht mehr an der Befragung teil (Attrition i.e.S.). Box 3.1 zeigt die Basisinformationen (nur Stammpersonen, die 1984 mindestens 16 Jahre alt waren).[2]

Für Mortalitätsuntersuchungen sind offenbar die während des Beobachtungsfensters feststellbaren Sterbefälle von entscheidender Bedeutung. Beschränken wir uns auf die Teilstichprobe A des SOEP, lassen also die Ausländerstichprobe (B) unberücksichtigt, lassen sich etwas mehr als 700 Sterbefälle ermitteln.

Um die verfügbare Information einzuschätzen, ist es zweckmäßig, von einer theoretischen Vorstellung auszugehen. Sei T eine Zufallsvariable für die Lebensdauer gemessen in Jahren, also eine diskrete Zufallsvariable. Was uns zunächst interessiert, ist die Verteilung dieser Zufallsvariablen, die durch die Wahrscheinlichkeitsfunktion $\Pr(T = t)$ beschrieben wird. Dabei ist $\Pr(T = t)$ die Wahrscheinlichkeit dafür, daß eine Person im Alter von t Jahren stirbt; bzw., wenn man sich auf eine endliche Grundgesamtheit bezieht, der Anteil der Personen aus dieser Grundgesamtheit, die im Alter von t Jahren sterben.

Welche Informationen liefern unsere Daten über diese Wahrscheinlichkeitsfunktion? Offenbar sind unsere Informationen begrenzt, denn über diejenigen Personen, die bereits vor dem Beginn unseres Beobachtungsfensters gestorben sind und die infolgedessen in der SOEP-Stichprobe gar

[2] Um die Sterbefälle zu ermitteln, wurde die von Infratest (1993) durchgeführte Untersuchung über Ausfälle im SOEP berücksichtigt.

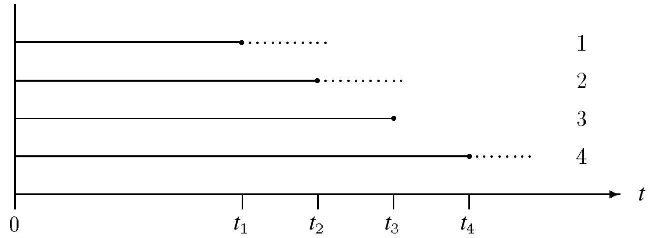

Abb. 3.1: Beobachtung der Lebensdauer von vier Personen.

nicht repräsentiert sein können, wissen wir nichts. Empirisch ermitteln lassen sich jedoch folgende bedingte Wahrscheinlichkeiten

$$r(t) = \Pr(T = t | T \geq t)$$

die wir als Sterberaten interpretieren können: $r(t)$ ist die Wahrscheinlichkeit, im Alter von t Jahren zu sterben, vorausgesetzt, daß man nicht schon vorher gestorben ist.

Diese Sterberaten lassen sich empirisch so ermitteln, daß man zunächst für jedes Alter t die sog. Risikomenge R_t feststellt, d.h. die Personen, die bis zu diesem Alter überlebt haben und infolgedessen noch dem Risiko zu sterben ausgesetzt sind. Der Anteil der Personen, die im Verlauf ihres t.ten Lebensjahrs sterben, ist dann eine sinnvolle Schätzung für $r(t)$. Abbildung 3.1 illustriert dies anhand von vier Personen, deren Lebensdauer jeweils bis zu den Zeitpunkten t_1, t_2, t_3 und t_4 beobachtet wird. Bei der dritten Person wissen wir, daß sie zum Zeitpunkt t_3 gestorben ist; bei den anderen Personen ist die Beobachtung rechts zensiert, d.h. wir wissen nicht, wie lange sie gelebt haben oder noch leben werden. Nur die dritte Person liefert in diesem Beispiel also eine Information zur Berechnung einer Sterberate für den Zeitpunkt t_3. Die Risikomenge besteht aus den Personen, die bis zu diesem Zeitpunkt überlebt haben, also aus den Personen 3 und 4.

Kennt man die Sterberaten, kann daraus die Verteilung der Lebensdauervariablen T berechnet werden, zum Beispiel in Form einer Survivorfunktion $G(t) = \Pr(T > t)$. Offenbar gilt

$$1 - r(t) = 1 - \frac{\Pr(T = t)}{\Pr(T \geq t)} = \frac{\Pr(T > t)}{\Pr(T \geq t)} = \frac{G(t)}{G(t-1)}$$

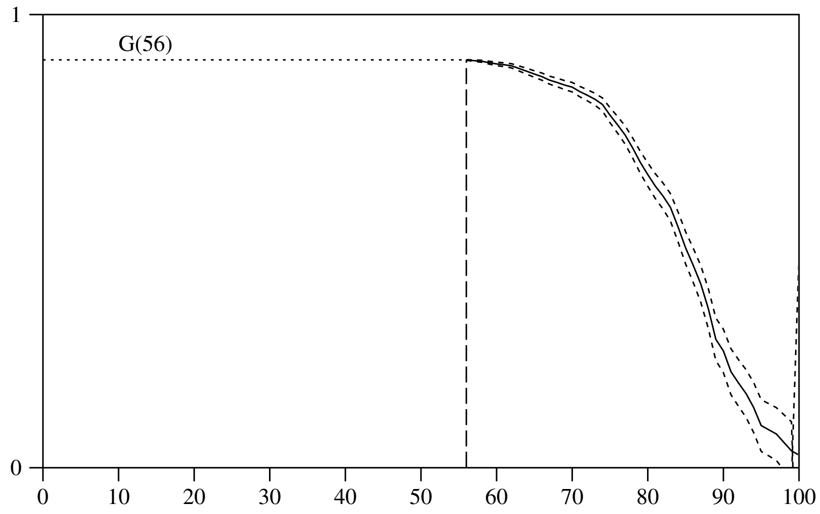

Abb. 3.2 Survivorfunktion für die Lebensdauer. Kaplan-Meier-Schätzung mit 2675 Stammpersonen des SOEP, die vor 1930 geboren wurden. Abszisse: Lebensalter in Jahren. Gestrichelt: 95 % Konfidenzintervall.

woraus dann folgender Ausdruck für die Survivorfunktion resultiert:

$$G(t) = \prod_{l=1}^{t} (1 - r(l))$$

Allerdings muß berücksichtigt werden, daß wir nur während unseres Beobachtungsfensters Informationen über Sterberaten gewinnen können; d.h., im Hinblick auf die Zeitachse für die Lebensdauervariable T, nur für einen Zeitraum von t_a bis t_b, wobei t_a das Sterbealter der am jüngsten gestorbenen Person und t_b das Sterbealter der am ältesten gestorbenen Person ist.

In den uns verfügbaren Daten ist $t_a = 21$ und $t_b = 103$. Allerdings können die Sterberaten nicht für diesen gesamten Altersbereich verläßlich geschätzt werden, da bis zu einem Alter von etwa 55 Jahren nur sehr wenige Sterbefälle auftreten. Wir beschränken deshalb unsere weiteren Untersuchungen auf Personen, die vor 1930 geboren wurden, d.h. zum Beginn unseres Beobachtungsfensters mindestens 55 Jahre alt sind. Dies sind insgesamt 2675 Personen (1124 Männer und 1551 Frauen), von denen 600 während des Beobachtungsfensters gestorben sind.

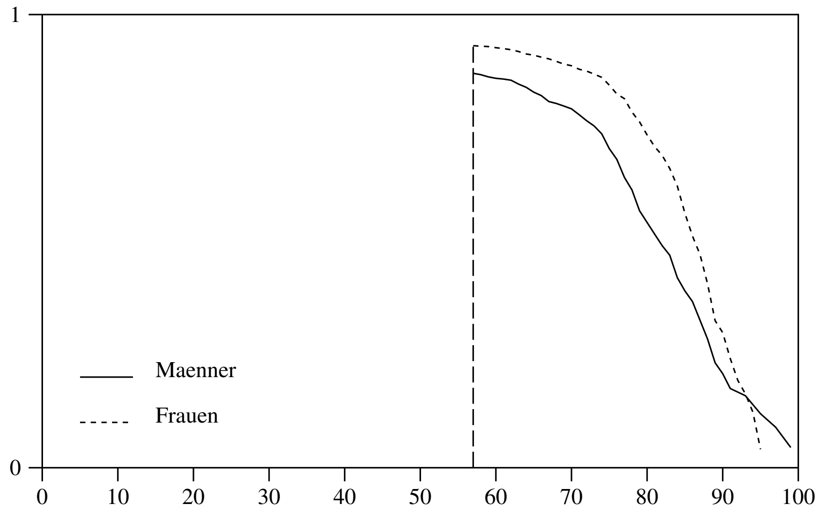

Abb. 3.3 Survivorfunktion für die Lebensdauer, differenziert nach Männern und Frauen. Kaplan-Meier-Schätzung mit 2675 Stammpersonen des SOEP, die vor 1930 geboren wurden. Abszisse: Lebensalter in Jahren.

Mit diesen Informationen kann jetzt eine Survivorfunktion für die Lebensdauer geschätzt werden. Wir verwenden dafür das Kaplan-Meier-Verfahren und berücksichtigen, wie oben dargestellt, daß es sich um links abgeschnittene Daten handelt.[3] Abbildung 3.2 zeigt die geschätzte Survivorfunktion. Der erste beobachtbare Sterbefall tritt in einem Alter von 56 Jahren auf; und eine Schätzung der Survivorfunktion ist infolgedessen erst ab diesem Alter möglich, d.h. die von uns geschätzte Funktion ist

$$G(t) = G(55) \prod_{l=1}^{t} (1 - r(l)) \quad t \geq 56$$

Da wir nicht wissen, wieviele Personen bereits bis zum Alter von 56 Jahren gestorben sind, kann $G(55)$ nicht ermittelt werden;[4] und infolgedessen können auch keine Quantile der Lebensdauerverteilung geschätzt

[3] Kaplan-Meier-Schätzungen von Surviforfunktionen mit links abgeschnittenen Daten können direkt mit dem Computerprogramm TDA durchgeführt werden; vgl. dazu Blossfeld und Rohwer 1995, Abschnitt 3.2.

[4] Ein Näherungswert kann jedoch aus der amtlichen Statistik ermittelt werden. Die Survivorfunktion aus der amtlichen Sterbetafel für 1986–88 (vgl. Meyer und Paul 1991) zeigt, daß bis zum Alter von 57 Jahren noch 87.2 % der Männer und 93.3 % der Frauen überlebt haben.

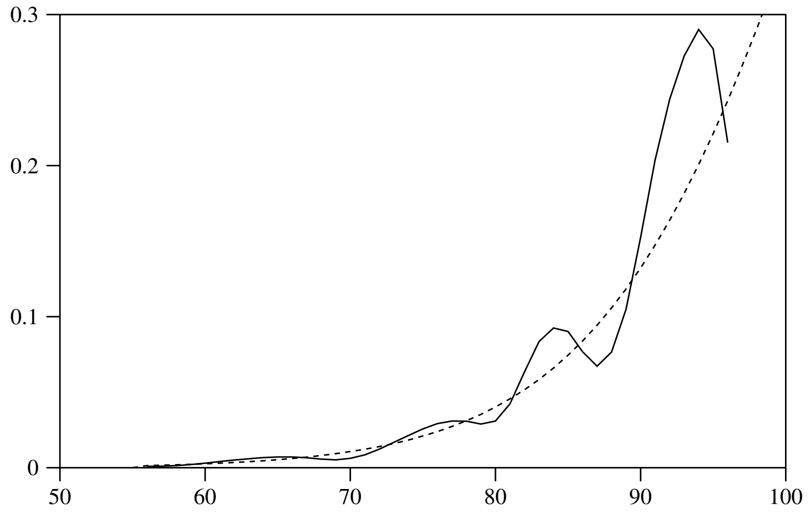

Abb. 3.4 Geglättete Sterberaten, geschätzt für 2675 Stammpersonen des SOEP, die vor 1930 geboren wurden. Gestrichelt: Baselinerate geschätzt mit einem logistischen Regressionsmodell, wobei die Zeitabhängigkeit durch ein Polynom 2. Grades definiert wird. Abszisse: Lebensalter in Jahren.

werden (z.B. die mittlere Lebensdauer).[5] Aber der Verlauf des mit unseren Daten schätzbaren Abschnitts der Survivorfunktion ist jedenfalls proportional zur tatsächlichen Gesamtsurvivorfunktion; und insofern erhält man für den Zeitabschnitt ab 56 Jahren eine brauchbare Schätzung.[6]

Da es bekanntlich deutliche Unterschiede in der Mortalität bei Männern und Frauen gibt, ist es zweckmäßig, getrennte Survivorfunktionen zu schätzen. Abbildung 3.3 zeigt die Schätzergebnisse beginnend mit einem Alter von 57 Jahren.[7] Um die geschätzten Survivorfunktionen besser vergleichen zu können, haben wir aus den Angaben der amtlichen Sterbetafel für 1986–88 (Meyer und Paul 1991) als Wert der Survivorfunktionen im Alter von 57 Jahren für Männer den Wert 87.2 und für Frauen den Wert 93.3 angenommen.

Für das Alter ab 56 (bzw. bei Frauen ab 57) Jahren können auch die Sterberaten geschätzt werden. Da das Kaplan-Meier-Verfahren nicht un-

[5] Allerdings können *bedingte* Lebenserwartungen geschätzt werden.

[6] Abb. 3.2 zeigt zusätzlich noch ein 95 % Konfidenzband, berechnet aus den üblichen asymptotischen Standardfehlern. Da es sich um links abgeschnittene Daten handelt, werden diese Standardfehler allerdings unterschätzt.

[7] Da in unserem Datensatz keine Frau vorkommt, die im Alter von 56 Jahren stirbt.

mittelbar die Raten selbst, sondern nur kumulierte Raten liefert, müssen die Sterberaten aus einer numerischen Differentiation der kumulierten Raten gewonnen werden. Da die so berechenbaren Raten infolge von Unstetigkeiten der kumulierten Raten stark fluktuieren, nehmen wir vorher eine Glättung der kumulierten Raten mithilfe von Splinefunktionen vor, woraus sich dann die Raten selbst durch analytische Differentiation berechnen lassen. Abbildung 3.4 zeigt die so gewonnene Schätzung von Sterberaten für ein Alter ab 56 Jahren.

4 Modellschätzungen

Da sich mit den verfügbaren Daten die Sterberaten für einen gewissen Altersbereich sinnvoll schätzen lassen, ist es möglich, Modelle zu konzipieren, um zu untersuchen, wie diese Sterberaten von gewissen Merkmalen der Lebensverläufe der Personen abhängen. Eine Möglichkeit besteht darin, sog. Übergangsratenmodelle zu verwenden, d.h. Modelle der Art

$$r(t) = \Pr(T = t | T \geq t) = g(t, x_1, \ldots, x_n)$$

Dabei ist g eine noch genauer zu bestimmende Funktion, durch die die Rate $r(t)$ von der Lebensdauer t und einer Reihe von Kovariablen $(x_1, \ldots, x_n)$ abhängig gemacht wird. Wir verwenden hier zur Illustration der Vorgehensweise ein sog. logistisches Regressionsmodell, bei dem die Abhängigkeit der Sterberate vom Alter durch ein Polynom spezifiziert wird:

$$r(t) = \frac{\exp(X\alpha + \beta_0 + t\beta_1 + t^2\beta_2 + t^3\beta_3 + \cdots)}{1 + \exp(X\alpha + \beta_0 + t\beta_1 + t^2\beta_2 + t^3\beta_3 + \cdots)}$$

X ist ein Zeilenvektor mit Kovariablen und α der zugehörige Spaltenvektor mit Koeffizienten; die Koeffizienten $\beta_0, \beta_1, \beta_2, \ldots$ definieren das Polynom für die Altersabhängigkeit der Sterberate.

Es handelt sich um ein proportionales Übergangsratenmodell. Abbildung 3.3 zeigt jedoch, daß die Sterberaten für Männer und Frauen nicht proportional sind. Wir nehmen die Modellschätzung deshalb getrennt für Männer und Frauen vor. Außerdem messen wir, um die Modellschätzung zu vereinfachen, das Alter erst beginnend mit dem 55.ten Lebensjahr, d.h. $t = 1$ entspricht einem Alter von 56 Jahren.

Box 4.1 Bildungsniveau (1. Welle) bei Stammpersonen des SOEP, die vor 1930 geboren wurden (ohne Fehlwerte).

Bildung	Maenner	Frauen
BILD__8_10	168	847
BILD_11_12	766	614
BILD_13_15	69	40
BILD_17_19	119	47
Insgesamt	1122	1548

Es bleibt die Frage, welche Kovariablen sinnvollerweise in einem solchen Modell verwendet werden können. Wie in Abschnitt 2 ausgeführt worden ist, können wir mit unseren Daten nicht anstreben, Mortalität durch eine Bezugnahme auf Todesursachen kausal zu erklären, denn über Todesursachen haben wir keinerlei Information. Das Ziel der Modellbildung kann daher überhaupt nicht darin liegen, Mortalität kausal zu erklären. Sie soll vielmehr dazu dienen, zu zeigen, daß es je nach sozialen Lebensumständen eine unterschiedliche Art von Mortalität gibt.

Kovariablen, um soziale Bedingungen des Mortalitätsrisikos zu erfassen, sind typischerweise zeitabhängig. Die grundlegende Idee besteht darin, das Mortalitätsrisiko im Alter t, also $r(t)$, von sozialen Umständen abhängig zu machen, wie sie zu diesem Zeitpunkt, also im Alter t, bestehen. Dabei muß allerdings berücksichtigt werden, daß sich zumindest einige wesentliche soziale Bedingungen von Mortalität während des Lebensverlaufs kumulativ entwickeln. Für die allgemeine Modellspezifikation ist dies zwar unerheblich, denn der allgemeine Ansatz, $r(t)$ von Kovariablenwerten zum Zeitpunkt t abhängig zu machen, kann durchaus auch dazu verwendet werden, um sich im Lebensverlauf kumulierende soziale Bedingungen zu erfassen. Ein wesentlicher Unterschied besteht jedoch im Hinblick auf die Konstruktion der Kovariablen. Zum Beispiel kann man das jeweils sich von Jahr zu Jahr ändernde Einkommen verwenden, oder (stattdessen) versuchen, einen sinnvollen Indikator für sich im Lebenslauf akkumulierende Einkommenspositionen zu finden. Offenbar werden dadurch in ganz unterschiedlicher Weise mögliche soziale Bedingungen von Mortalität erfaßt.

Da es in diesem Papier nur um eine Illustration von Methoden zur Untersuchung sozialer Bedingungen von Mortalität geht, beschränken wir uns auf eine einzige Kovariable, die in sehr allgemeiner Weise Lebensverläufe

Box 4.2 Logistische Regressionsmodelle für die Sterberate mit einem Polynom 2. Grades für die Altersabhängigkeit

Variable	Maenner Koeffizient	Maenner t-Stat	Frauen Koeffizient	Frauen t-Stat
Beta_0	-6.376	-22.19	-6.869	-23.23
Beta-1	0.182	6.98	0.114	3.92
Beta-2	-0.002	-2.42	0.001	1.40
BILD_11_12	-0.107	-0.69	0.078	0.62
BILD_13_15	-0.114	-0.40	-0.284	-0.67
BILD_15_17	-0.671	-2.38	-0.756	-1.29
Beobachtungen	1122		1548	
Sterbefaelle	299		299	

unterscheidbar macht: das Bildungsniveau. Da wir uns auf Personen beschränken, die vor 1930 geboren wurden, kann angenommen werden, daß das höchste Bildungsniveau bereits in der ersten Welle des SOEP (1984) erreicht worden ist. Box 4.1 zeigt die hier unterschiedenen Niveaus: 8 – 10, 11 – 12, 13 – 15, 17 – 19 Jahre. Für die Modellschätzung behandeln wir sie als Dummyvariablen und verwenden `BILD_8_10` als Referenzkategorie.

Box 4.2 zeigt die Ergebnisse einer Schätzung des oben formulierten logistischen Regressionsmodells für die Sterberaten von Männern und Frauen, wobei die Altersabhängigkeit durch ein Polynom 2. Grades spezifiziert wurde.[8] Offenbar besteht eine gewisse Abhängigkeit vom Bildungsniveau; jedenfalls haben Männer mit einem hohen Bildungsniveau ein deutlich niedrigeres Mortalitätsrisiko. Bei Frauen ist der Effekt nicht signifikant, vermutlich wegen der sehr geringen Fallzahlen in den höheren Bildungsniveaus (man erinnere sich, daß wir nur Personen betrachten, die vor 1930 geboren wurden).

5 Ein alternativer Modellansatz

Ratenmodelle werden üblicherweise auf einer Prozeßzeitachse formuliert, die die Dauer bis zum Eintreten eines Ereignisses erfaßt. So sind wir

[8] Die Modellschätzung wurde mit dem Computerprogramm TDA vorgenommen.

Box 5.1 Logitmodelle für die Sterbewahrscheinlichkeit.

Variable	Maenner Koeffizient	Maenner t-Stat	Frauen Koeffizient	Frauen t-Stat
Beta_0	-6.376	-22.19	-6.869	-23.23
Beta-1	0.182	6.98	0.114	3.92
Beta-2	-0.002	-2.42	0.001	1.40
BILD_11_12	-0.107	-0.69	0.078	0.62
BILD_13_15	-0.114	-0.40	-0.284	-0.67
BILD_15_17	-0.671	-2.38	-0.756	-1.29
Konstante	-4.247	-14.06	-4.748	-15.03
Alter	0.099	3.47	0.033	1.05
Alter2	-0.002	-0.34	0.020	2.68
BILD_11_12	-0.340	-2.14	-0.096	-0.76
BILD_13_15	-0.516	-1.80	-0.458	-1.07
BILD_17_19	-0.968	-3.39	-0.892	-1.51

auch im vorangegangenen Abschnitt vorgegangen: die Prozeßzeitachse entsprach dem Lebensalter. Eine Alternative besteht darin, den Modellansatz unmittelbar in der Kalenderzeit vorzunehmen und die Altersabhängigkeit des Mortalitätsrisikos durch Kovariablen zu erfassen.

In diesem Abschnitt soll kurz ein solcher Modellansatz skizziert werden. Ausgangspunkt ist die Kalenderzeitachse $t = 1985, \ldots, 1993$ entsprechend den bisherigen 10 Wellen des SOEP.[9] Die abhängige Variable ist Y_{it}, der Zustand des i.ten Individuums zum Zeitpunkt (Jahr) t; es handelt sich um eine binäre Variable mit den Werten 0, wenn die Person noch lebt, und 1, wenn sie nicht mehr lebt. Als Modell verwenden wir ein einfaches Logitmodell:

$$\Pr(Y_{i,1985}, \ldots, Y_{i,1993}) = \prod_{t=1985}^{1993} \frac{\exp(X_{i,t}\beta)}{1 + \exp(X_{i,t}\beta)}$$

Als Kovariablen betrachten wir wie im vorausgegangenen Abschnitt drei Dummyvariablen für das Bildungsniveau, sowie das Alter (linear und quadriert) in der jeweiligen Welle. Die Modellformulierung nimmt an,

[9] Wir beginnen in 1985, da natürlich in 1984 noch kein Sterbefall beobachtet werden kann.

daß – konditional auf diese Kovariablen – die jahresspezifischen Sterbewahrscheinlichkeiten unabhängig sind.[10]

Dieses Modell kann im Prinzip auf einfache Weise als ein gepooltes Logitmodell geschätzt werden. Es muß allerdings bedacht werden, daß einige Personen bereits vor dem Ende des Beobachtungsfensters ausscheiden, sei es weil sie sterben oder infolge von Attrition i.e.S. Dies kann dadurch berücksichtigt werden, daß für die Formulierung der Likelihood des Modells nur die tatsächlich vorhandenen Beobachtungen verwendet werden.

Box 5.1 zeigt die Schätzergebnisse. Um die Berechnung zu vereinfachen, wurde das Alter ab dem 55.ten Lebensjahr gemessen und das quadrierte Alter durch 10 geteilt. Die Ergebnisse entsprechen offensichtlich im wesentlichen denen des logistischen Regressionsmodells in Box 4.2.

Literatur

Blossfeld, H.-P., Rohwer, G. (1995). Techniques of Event History Modeling. Mahwah, NJ: Lawrence Erlbaum.

Infratest (1993). SOEP 92 Testerhebungen, bearbeitet von B. von Rosenbladt, M. Haberkorn, F. Stutz. München: Infratest Sozialforschung.

Meyer, K., Paul, C. (1991). Allgemeine Sterbetafel 1986/88. Wirtschaft und Statistik, Heft 6, 371 – 381, 234* – 241*.

[10] Da wir sicherlich nicht alle möglicherweise relevanten Kovariablen erfassen, ist diese Annahme natürlich problematisch. Um die Ausführungen nicht zu komplizieren, verzichten wir jedoch darauf, zusätzliche Kontrollen für eine verbleibende serielle Korrelation im Zeitablauf in die Modellformulierung einzubauen.