Materialien aus der Bildungsforschung Nr. 65

Sigrid Wehner

EXPLORING AND VISUALIZING EVENT HISTORY DATA

Max-Planck-Institut für Bildungsforschung
Berlin 1999

GW ISSN 0173-3842
ISBN 3-87985-074-7

Materialien aus der Bildungsforschung

Bestellungen werden erbeten an die Institutsadresse. Der vorliegende Band 65 der Reihe kostet DM 10,– (Selbstkostenpreis, einschließlich 7% MwSt) zuzüglich DM 3,– Versandpauschale.

GW ISSN 0173-3842
ISBN 3-87985-074-7

TABLE OF CONTENTS:

Abstract / Zusammenfassung 7

Acknowledgements 8

1. Introduction 9
- 1.1 Why Exploratory Data Analysis? ... 9
- 1.2 The Special Characteristics of Event History Data ... 10

2. Data 13
- 2.1 The German Life History Study ... 13
- 2.2 The Complexity of the Total Data Set ... 15
- 2.3 Description of the Used Event History File ... 17

3. Basic Explorations: Describing Durations of the Event History Data Set 19
- 3.1 First View: The Inspection of Raw Data ... 20
- 3.2 The Job Duration Pattern ... 23
- 3.3 Censored versus Non-Censored Durations ... 26
- 3.4 Visualizing Multiple Episodes: Career Patterns over the Life Course ... 28

4. Deeper Explorations:
Visualizing Statistical Functions for Duration Processes 31
- 4.1 The Kaplan-Meier Product-Limit Estimator ... 31
- 4.2 Interactive Exploration of Survivor Curves ... 35
- 4.3 Inspecting Hazard Rates ... 39

5. Conclusion: Demands for Future Exploratory Software 43

References 45

Appendix

- **A** Codebook for the Used Data Set XPLOD
- **B** Source Code of the Developed Pascal Programmes
- **C** Figures and Tables

ABSTRACT

Exploration of a longitudinal data set in the social sciences means to look at the data dynamically over time. Graphic data presentations have to manage a high complexity in the data structure. Several exploratory methods will be applied to an event history data set of job careers which is part of the German Life History Study. Especially the high number of cases and the high proportion of censored episodes need particular attention. Different visualizations are developped to take into account the special properties of spell data. Concepts for interactive software to explore survival data are presented and discussed at the end.

ZUSAMMENFASSUNG

Exploration von sozialwissenschaftlichen Längsschnittdaten muß eine dynamische Betrachtungsweise über die Zeit erlauben. Dementsprechend haben grafische Darstellungen eine hohe Komplexität der Datenstrukturen zu bewältigen. Auf einen Ereignisanalysedatensatz von Erwerbsverläufen der deutschen Lebensverlaufsstudie werden verschiedene explorative Methoden angewendet. Insbesondere erfordern die hohe Fallzahl in diesen Daten und der große Anteil zensierter Episoden besondere Berücksichtigung. Es werden verschiedenene Visualisierungen für den speziellen Charakter von Spelldaten entwickelt und abschließend Konzepte für interaktive Software zur Exploration von Survivaldaten vorgestellt und diskutiert.

Acknowledgements

The following text is the research project report on explorations and visualizations of event history data and is submitted for the Diploma in Social Science Data Analysis at the University of Essex.

In the first place I want to thank Karl Ulrich Mayer, not only for his permission to use the life course data. He also supported and encouraged my studies at the Essex Summer School. During my years of work in his department at the Max Planck Institute for Human Development, I got to know his German Life History Study in so many aspects.

Eugène Horber was an inspiring instructor on data exploration and visualization. He showed the way to use fantasy, even in statistics.

I would also like to thank Ulrich Pötter for mathematical discussions, Heike Solga for always being open to sociological conversation, and Elinor Scarbrough for her advice during my summer school attendances.

Berlin, July 1998

... il faut réapprendre à <<VOIR>>. C'est peut-être la propriété essentielle de la Graphique.
(Jacques Bertin)

...one has to relearn to SEE. Perhaps this is the essential property of graphics.

1. Introduction

The purpose of this thesis is to apply exploratory methods to an event history data set which is part of the German Life History Study. In this study, many variables are observed over time so that a variety of complex interrelationships is given. Focussing on the history of job careers over life course, I will contribute to a better understanding of the data structure and its special properties. Different types of visualization for episode data and the advantages of interactivity are presented and discussed.

1.1 Why Exploratory Data Analysis?

Empirical researchers are convinced that not only theories can explain social phenomena in the world but that, in addition, a collection of data will raise the level of knowledge. The statistical methods for this research process can be devided into two main groups: the confirmatory and the exploratory methods.

Confirmatory statistics aim to draw inferences from a sample to a population. Formulated hypotheses have to be accepted or rejected, and levels of significance are the criteria about the judgement. Based on a set of assumptions, statistical models do or do not confirm a pre-formulated theory.

Exploratory methods do not judge about hypotheses. Beyond any statistical assumptions (as for example on special distributions), they attempt to clarify the puzzling jungle of data by simplifying and ordering them. This means to "try many things that do not work" but also "to notice what we never expected to see", as Tukey[1] states. Creating hypotheses is the result. Like Tukey, Erickson/Nosanchuk[2] compare exploration with a detective's work and confirmation with a "courtroom trial". Although these differences are important, both approaches should not be considered as separate schools of thinking but instead be used complementarily. The result of the data exploration phase will be a deeper insight into internal relationships inspired by

---

[1] Tukey, John W. (1977) Exploratory Data Analysis, Preface, pp. i-viii.
[2] Erickson,B.H./Nosanchuk,T.A. (1992) , Understanding Data/ Second Edition, Introductory chapter, pp. 3-11.

"scepticism and openness", as Hartwig/Dearing[3] (1979) formulate. The invention of boxplots (and many other plots) by John W. Tukey[4] amazes by its simplicity. Pictures help to grasp patterns better than lists of numbers do. Therefore the inspection of data should be fast and comfortable so that several tries are easy. Because "there is no abstract user", as Horber[5] (1987) writes, a variety of graphics is necessary to allow different perceptions of the same phenomenon. Since Tukey's days the computer technology has improved enormously, which allows looking for more and more suitable visualizations. Data exploration has been always stimulated by creativity, and it is an, if not the most, enjoyable part of the work with empirical data.

1.2 The Special Characteristics of Event History Data

Event history data observe changes of qualitative variables over time. For the understanding of the next chapters, the following terminology[6] is necessary:

The ***state space***	defines the categories of an observed qualitative variable over time. There may be multiple states, but often only a dichotomous state space is defined, e.g., being at school or not, being married or not.
The ***duration***	is the basic time concept. It is the time span from a given starting to an ending point during which the observed object stays in the same state. This duration is also called ***spell*** or ***episode.***
An ***event***	terminates an episode so that a state transition is observed, e.g., a "school-spell" ends when leaving school. Stemming from the terminology of medical statistics, the end of a spell is also called ***death***, ***exit***, or ***failure***.
Survivors	are spells which have not yet terminated at a given time. They are surviving.
Censored data	are incomplete observations because persons get lost from the study, or the interview date cuts off still continuing episodes.

[3] Hartwig, F. with Dearing,B.E. (1979), Exploratory Data Analysis, pp. 9-13.
[4] Tukey, John W. (1977), see concepts for box-and-whisker plots on pp. 39-56.
[5] Horber, Eugène (1987) Exploring Aggregate Data, Conceptual and Practical Aspects in: European Political Data Newsletter, Bergen, pp.30-36.
[6] Mostly, I use the vocabulary as given in Blossfeld/Hamerle/Mayer (1989), Event History Analysis, pp.14-30 and Blossfeld/Rohwer (1995), Techniques of Event History Modeling, pp.33-41. They also use examples of the German Life History Study.

Life-course data sets contain time-constant covariates (as gender or date of birth) as well as time-varying covariates (as marital status or income), associated to spell data. Many statistical packages have implemented a rich collection of models for survival analysis. But only a few attempts have been made to explore and visualize such data. Graphical displays have to solve a problem: snap-shot like plots are not any longer sufficient because they represent only one moment, and changes over time cannot be seen. The introduction of time adds a new dimension, but even a computer with a high-resolution monitor can indicate only an illusion of many dimensions on a two-dimensional screen. Thus, a presentation of process data necessarily has to reduce complexity.

For example, the correspondence analysis approach to event history data, as shown by Blasius[7] (1995), explains the association between several states of a person at two given time points. It is true that the graphical presentation of correspondence analysis results sums up groups of similarly changing persons, but the imagination of a time axis gets lost.

Another way of visualization is described by Francis/Fuller[8] (1996). They take Lexis diagrams to combine a historical time axis and an axis for individual age with graphical attributes for time-varying and time-constant covariates. A hyper-3D-Lexis diagram is the result which cannot be presented here in detail. This tool finds its limits by the number of cases. If N becomes large, a combination of too many graphical aspects overstrains the spectator's eye.[9]

The basic subject of the following chapters is the examination of durations. I will outline some visual representations for "long" and "short" spells and compare subgroups of the population. Available software[10] has been used and examined in terms of its applicability and results, and some new specifications for exploratory pictures can be proposed.

---

[7] Blasius,J. (1995). Visualizing Panel, Trend, and Event History Data with Correspondence Analysis. in: Faulbaum,F./Bandilla,W. (Eds.), Softstat '95, Advances in Statistical Software 5, Lucius & Lucius, Stuttgart, pp. 81-88.

[8] Francis,B./Fuller,M.(1996). Visualization of Event Histories. Journal of Royal Statistical Society, A, 159, pp.301-308.

[9] The same effect is already given by simple scatter-plots with a large number of cases.

[10] BMDP, SAS, SPSS, STATA, LIMDEP, TDA, X-LISP STAT, PASCAL. Unfortunately, I had no access to S-Plus which is rather expensive. Therefore, tests with the programming facilities of this package could not be done.

2. Data

2.1 The German Life History Study

The data used in this thesis are part of the German Life History Study (GLHS/ West and East Germany) directed by Karl Ulrich Mayer, and conducted at the Max Planck Institute for Human Development in Berlin. Nearly 8000 West and East German persons of different birth cohorts were interviewed retrospectively about several domains of life such as family, place of residence, education, vocational training, and employment. Separate parts of this study were conducted during different years, and two different methods of data collection were used: face-to-face interviews and computer assisted telephone interviews (CATI). The following table lists the sample sizes for the different parts of the study.

Table 1:

Sample Size for Different Birth Cohorts of the German Life History Study

Birth cohorts		Interview method	Number of persons	Time of interview
West German persons				
1919-21	a)	face-to-face	407	1985-86
	b)	telephone interview/CATI	1005	1987-88
1929-31, 1939-41, 1949-51		face-to-face	2171	1981-83
1954-56, 1959-61		telephone interview/CATI	2008	1989
East German persons				
1929-31,1939-41, 1951-53, 1959-61		face-to-face	2331	1991-92
		total:	7922	

These data are stored in a SIR data base. The GLHS is available for external scientific users at the Central Archive for Empirical Social Research in Cologne[11]. Detailed information about sampling, interviewing, and the data editing process are given in the GLHS methodological reports and codebooks[12] .

[11] ZOO/TDA archive files (developed by Götz Rohwer) are distributed for the West German data, the East German data are delivered as SPSS export files.
[12] Mayer, K.U. & Brückner, E. (1989); Brückner, E. (1993); Brückner, H. and Mayer, K.U. (1995); Wagner, M. (1996) ZA-Information 38, pp. 20-27; Solga, H. (1996) ZA-Information 38, pp. 28-38.

The cohort design[13] of the GLHS is used as a means to compare different historical periods and to find indicators of social change in Germany. In addition, an East/West comparison is possible. Thus, I will carry out most of my explorations for the grouping variables 'birth cohort' and 'East/West'. As a third distinction I will add 'gender' because I expect to find differences between men and women (at least when variables concerning education or employment are considered).

In principle, I focus on time-related variables. Since these data are collected retrospectively, I will briefly discuss their quality. It is often argued that retrospective data are less exact than panel data in the timing of events due to a lack of human memory. Auriat[14] (1996) states that the recall error in retrospective studies depends on which type of event is asked and recommends to guide the interview through past events of the respondent's life course. This was done in the GLHS by partitioning the internal sequence of the interview into several modules of different domains of life.

Other efforts for data quality have been done: interviewer training; computer-assisted data checking; intensive data editing. Summarizing the field experience with the 1919-21 birth cohorts, E. Brückner[15] (1993) reports that the main problem did not lie in the respondents' insufficient memory but in failures in the communication between interviewers and interviewed persons. H. Brückner[16] (1995) finds a declining rate of errors in event-related dates when more time is given for the interview.

Blossfeld/Rohwer[17] (1995) mention as advantages of a retrospective data collection, in contrast to panel studies, that they are based on a constant cultural and historical framework of coding and meaning of wording at time of interview. Moreover, they do not suffer from panel attrition, and they provide many time-related variables from a long period of the respondents' life immediately after a relatively short fieldwork.

---

[13] A file containing all cohorts therefore does not represent the proportionality in the population. Because my intentions are basic explorations instead of inferential statistics, I used a file with all cohorts together, even though most of the explorations have been done separately by cohort.
[14] Memory for timing of residential episodes is reported to be more reliable than for job shifts. Further discussion see in: Auriat,Nadia (1996), pp. 139-145 chapter 9, "La carrière professionnelle, les évènements démographiques et les erreurs de mémoire" and recommendations for retrospective surveys on pp.152-157.
[15] E. Brückner (1993,p. 27; for more details about CATI and data quality see the methodological report for birth cohorts 1919-21 on pp. 115-127; 178-183; 187-193).
[16] H. Brückner distinguishes "visible errors" like missing dates or implausible time sequences from "invisible errors" like the number of job spells before and after data editing. She estimates influences on both types of errors in Poisson regression models (pp. 28-33,54,73-78).
[17] Blossfeld/Rohwer (1995) discuss detailed methodological aspects of panel data versus retrospective data on pp. 11-26.

2.2 The Complexity of the Total Data Set

The total data set is a rich source for empirical research about historical changes and its interrelationships with individual life courses. The following three figures give a first visual impression of the logical organization of the data:

- figures 1 and 2 (lifeplots) show examples of individual life courses; figure 1 uses historical years as time axis, figures 2 takes years of age as time axis
- figure 3 (Lexis-boxplot) shows a data summary for two time-related variables in a combination of age and historical year.

The **'LIFEPLOT for Different Types of Episodes in Life over Historical Time'** *(figure 1 in appendix C)* presents the main episodes of the life of three West German respondents born in different years. The x-axis is the scale for historical years, the y-axis represents the different domains of life. Different episodes of life are presented by shorter or longer horizontal lines, distinguished by type and colour. Though Tufte[18] (1997) votes for being careful with colour displays and using it better as an information by itself like in geographical charts, I consider colour here to be necessary for a better visual contrast. The reader's eye should be able to find in one picture historical years, impression of durations, sequencing of events, and several types of events.

Example: Person 2

A woman, born in 1950 when the first place of residence begins (thin black line); she changes to her second place of residence in 1972 and to the third one in 1979. She reports one single school episode from 1956 to 1966 (red line), which is followed by a vocational training (open blue line). She starts her job career (bold blue line) in 1968, during this job she marries in 1972 (bold black line), and when getting her first child (open red line) in 1975 she interrupts her career (dotted blue line after the first job spell) and re-enters to labour market in 1977. The vertical dashed line at the right end marks the time of interview.[19]

This synopsis of individual lifeplots shows that episodes of one domain are ordered sequentially[20] (shown as a step function with current numbers) and that episodes of different domains may take place simultaneously. Furthermore, the historical position

---

[18] On pp. 76-77 in his latest graphical masterpiece "Visual Explanations", Tufte (1997) shows good and bad use of colour: a scale from light to dark blue as a graphical transformation of ocean depths is better than contrasting different areas below sea level in colours like yellow, red, and orange.

[19] The problem of censored data is important for survival models, I will discuss it in chapters 3 and 4.

[20] Jobs are defined as the predominant type of employment, no parallel spells are allowed; in the case of marriages, two episodes of being married at the same time cannot take place by law.

of the respondent's life can be seen easily. Taking into account the underlying cohort design which wants to separate cohort, age, and period effects (Glenn, 1977), I transported the lifeplots from the historical scale into years of age and marked specific dates like first marriage and household-formation.

The **'LIFEPLOT of Different Events and Episodes over Life Time'** *(figure 2 in appendix C)* scales the x-axis in years of age. Each person starts at age=0, so that age-specific life situations can be compared directly. The examples show the difficulty to summarize event sequences and durations simultaneously for the whole sample. Thus, figure 3 inspects only two time points in the life of women: the time of starting the first job, and the birth of the first child.

In the **'LEXBOX, Lexis Diagram Combined with Boxplots for Timing and Sequencing of Events'** *(figure 3 in appendix C),* I created a combination of the Lexis[21] diagram and the boxplot. For each birth cohort, a diagonal line connects the historical year (x-axis) with the corresponding age (y-axis). To this Lexis diagram, I have added diagonal boxplots to visualize a summary of the timing of two events. The box is positioned on the diagonal line where this event had occured for 25%, 50%, 75% of the women. Different box fill-patterns (full/stripes) denote the event types, different colours (blue/pink) contrast East and West. Obvious differences exist. In West Germany, the boxes for first job and first child are clearly separated. This time lag increases from cohort to cohort (looking at the median time). For East German women, the start of the first employment and the birth of the first child are closer together. The second difference is the different length of the boxes: the age variability is much greater in West Germany, especially for the birth of the first child.

This lexbox-plot visualizes the sequence of events and the variation in timing. It helps to decide whether further explorations on data quality have to follow or whether meaningful interpretations can be given. It finds its visual limits when the time boxes would overlap. Due to the high complexity of possible interrelations of time-related variables in the GLHS, in the following I will only focus on <u>one</u> specific event history variable, namely job spells and their sequence and duration.

---

[21] Lexis diagram invented by the German demographer Wilhelm Lexis (1875). Explanations in: Esser, H. (1996) Soziologie, Allgemeine Grundlagen, chapter17, pp.255-266. The concept of the Lexis diagram was also used by Francis/Fuller(1996).

2.3 Description of the Used Event History File

Job episodes of three birth cohorts are used for East and West Germany. The cohorts chosen are 1939-41, 1949-51 (West Germany)/1951-53 (East Germany), and 1959-61. In the following, they are only called the "1940", the "1950", and the "1960" cohort[22]. On the whole, 14,214 jobs of 4,119 East and West German persons have been observed. They display the historical existence of the two Germanies and can be used for a comparison of the two employment systems.

Appendix A is the codebook for the used file for explorations of durations, called XPLOD. The first set of variables contains episode-specific information: the person's current job number JOBNR; the censoring information CEN (a dummy variable); starting time TS and finishing time TF of the episode (calculated in months since 1900[23]); the job duration DUR in months; and other job-related information like income or size of firm.

The second set of variables refers to attributes of the respondent. They are always constant for identical CASID. The main grouping variables COHORT, SEX (gender), and EW (East or West Germany) are followed by other person-related variables as e.g. educational degree.

For complete information see the codebook in appendix A which also lists the percentage of missing data for each variable.[24]

[22] The exact date of birth is given in variable TBIRTH.

[23] This coding allows to calculate directly differences between dates.

[24] As a consequence of the interview techniques and data editing rules mentioned above, there are nearly no missing data for the time variables TS,TF and DUR. INCOME has a higher proportion of missings due to refusals and because only comparable monthly net income could be used. Variables for time of first marriage TMAR1 and birth of first child TKID1 have a high percentage of missing values because all persons who did not experience this event were coded as "does not apply".

3. Basic Explorations: Describing Durations of the Event History Data Set

Graphs are friendly. (John W. Tukey)

The next graphs intend to get a frame of reference for the variable "job duration". To guide through this chapter, the following overview summarizes their key characteristics including the already presented graphs.

Overview: Types of Exploratory Graphs

	Types of graphs	**shown in**	**sampling fraction**[25]
	already presented in chapter 2: understand life course data		
(1)	LIFEPLOT over historical time	figure 1	***single cases***
(2)	LIFEPLOT over age	figure 2	***single cases***
(3)	LEXBOX	figure 3	100%
	following in chapter 3: inspect a duration variable		
(4)	raw data listing	table 2	0.3%
(5)	job duration pattern	figure 4	1%
(6)	stacked stem-leaf diagram	figures 5a,b	100%
(7)	skeleton plots over historical time	figure 6	20%
(8)	skeleton plots over historical age	figure 7	20%
(9)	skeleton plots with slicing	figure 8	20%

What can be inspected by these graphs?	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
lifeplot/lexbox									
examples of individual life course	x	x							
sequencing of events and episodes	x	x							
cohort situation in historical years	x								
cohort situation at given age		x							
position and spread of event-related time points			x						
cohort versus period effects			x						
raw data listing/job duration pattern/stacked stem-leaf									
logic and structure of data organization				x	x				
proportion of data "holes"				x					
relation person/spells				x					
long versus short durations					x				
proportion of censored spells					x	x			
structure of distribution of censored episodes						x			
skeleton plots									
sequences of repeated spells							x	x	x
duration and timing of non-employment							x	x	x
labour force participation							x	x	x
period specific career patterns							x		
age specific career patterns								x	
high or low job mobility									x
termination of careers before interview date									x

[25] The reasons for subsampling and the advantages are discussed later on.

3.1 First View: The Inspection of Raw Data

Structuring a List

To get familiar with the logic of the data, I start with the inspection of some raw data lines.

Table 2, **'Raw Data Listing of Job Episodes'** *(see appendix C)*, shows a small part of file XPLOD. I have added horizontal lines to see all job spells belonging to the same person and vertical lines to separate episode varying variables[26] from person-related variables. DUR and CEN are marked because they are key variables in survival analysis. Missing values are left blank so that the proportion of "holes" in the data becomes evident. This simple list enriched with some structure allows a basic visual grasp of the event history data design.

Checking the Number of Cases: Persons versus Episodes and Censored Episodes

Survival data may be understood as single-episode or multi-episode data. In single-episode data, each line represents a single spell with assumed statistical independence[27] of the individual episodes. In the case of multi-episode data, all spells with the same identifier variable belong together. To see the relation person/spells in the XPLOD file, the number of jobs and the number of persons were counted. I did this by the "survival time commands" which are implemented in the software package STATA[28]. Consider the following user dialogue.

[26] See also chapter 2.3 and appendix A.

[27] Deeper explanations about single versus multiple episode data are given by Blossfeld/Rohwer (1995), pp. 33-50. Survival models can be done for both types of data.

[28] Details see STATA Release 5, Reference manual [R-Z], pp. 241-355. The group of st-commands allows to declare a data set to survival time data and to store the declared properties in the system file. Similarly, declarations for different sampling designs can be done by the svy-commands.

dialogue:
STATA commands for declaring XPLOD to multi-episode data

. stset XTF CEN, id(CASID) t0(TS) — *names the variables TS for entry time, XTF[29] for exit time, CEN for censoring information and CASID for the individual identification*

. stdes — *produces the screen[30]:*

```
failure time:     XTF
entry time:       TS
failure/censor:   CEN
id:  CASID
|-------------- per subject --------------|
Category                  total        mean          min      median        max
---------------------------------------------------------------------------------
no. of subjects            4112
no. of records            14201    3.453551            1           3         19

(first) entry time                 845.4086          617         849       1083
(final) exit time                  1040.204       687.99     1072.99    1115.99

subjects with gap          3533
time on gap if gap    81418.989    8.070075     .0100098    .0100098     364.01
time at risk          719577.85    174.9946     .9899902     142.975   477.9199
failures                  11033    2.683123            0           2         18
```

14,201 *non-missing* durations are read belonging to 4,112 subjects.[31]
The command **stvary** would give missing value diagnostics and a list of subject varying/constant variables.

The next table shows the number of episodes and persons together with the proportion of censored episodes, differentiated by subgroups.
Table 3:

Number of Persons and Jobs in Subgroups of File XPLOD

COHORT	EW	GENDER	# persons	# jobs	# censored jobs	% censored	# jobs per person
1940	**EAST**	female	292	1374	222	16%	4.7
		male	294	1401	258	18%	4.8
	WEST	female	344	1021	160	16%	3.0
		male	375	1421	362	25%	3.8
1950	**EAST**	female	287	1254	233	19%	4.4
		male	292	1161	268	23%	4.0
	WEST	female	360	882	167	19%	2.5
		male	354	1063	326	31%	3.0
1960	**EAST**	female	306	1182	231	20%	3.9
		male	267	908	249	27%	3.4
	WEST	female	466	1155	279	24%	2.5
		male	482	1392	425	31%	2.9
	EAST	female	885	3810	686	18%	4.3
		male	853	3470	775	22%	4.1
total EAST			1738	7280	1461	20%	4.2
	WEST	female	1170	3058	606	20%	2.6
		male	1211	3876	1113	29%	3.2
total WEST			2381	6934	1719	25%	2.9
		total sample	**4119**	**14214**	**3180**	**22%**	**3.5**

[29] XTF has been recalculated instead of TF to allow short durations of exactly one month for cases with TS=TF.
[30] Original STATA output with added bold attributes.
[31] Originally, N=4119.

The average number of jobs per person is 3.5. The proportion of censored episodes is very high, 22% of the total sample. Therefore, visualizations of the durations should contrast terminated and unterminated observations.

Subsampling

Apart from the censored observations, the data set contains a second problem: the high number of episodes (in total 14,214 jobs). To allow quick explorations, I used random subsamples of suitable fractions. For the raw data table for example, I took the small fraction of 0.3%, so that a list would fit on one single sheet of paper. To preserve the life course feature of the data, the relation person/spells had to be kept. Therefore it was sampled first on the basis of all persons, and then all jobs of the selected persons were taken. Of course no statistical inference can be done by this mini-subset of data, but it is sufficient to get an impression about the data structure. In addition, systematic failures like controlling only first or last lines of a file can be avoided.

The underlying idea of the subsampling method is that tasting a part will give hints whether it is worth to inspect the whole data. Especially authors working on robust statistics with resampling methods agree with this point of view: Mooney/Duval[32] (1993) argue that each sample of a population builds a typical specimen of the true information. In most of the exploratory graphs of this chapter, I used subsamples.

[32] This argument is taken from their theoretical justification of bootstrapping methods on pp. 20-22.

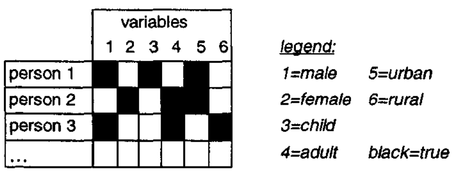
3.2 The Job Duration Pattern

The next step is to transform durations into a graphical presentation. A convenient visualization should allow to compare different lengths of time directly by eye. This is possible, if shorter and longer lines denote corresponding proportions of time, paying tribute to Tufte's [33](1997) argument :

> *"Visual representations of evidence should be governed by principles of reasoning about quantitative evidence."*

Keeping the usual view, the time axis will be ordered from the left to the right. Furthermore, the duration graph should simplify the data mass and, as Horber[34] (1990) emphasizes, be guided by the researcher's practical wishes instead of mathematical complexity. I constructed a duration diagram containing single data lines which build an overall impression of distribution.[35]

The visual arrangement refers to Jacques Bertin's[36] (1977) graphical data matrices. They take one dimension for the cases (e.g. cities or interviewed persons) and the other dimension for the collected information about observation units (e.g. age or salary). But instead of numbers, a Bertin data matrix contains little black and white fields as patterns of information. A short example will demonstrate the data organization:



Bertin's fundamental idea was that interchanging rows or columns would not destroy anything and keep the whole contents of information. Grouping together most similar lines could detect systematic relationships in the data optically. Bertin worked with sorted, folded sheets of paper. His book, published in 1977, still impresses with its variety of graphic ideas, especially since the possibilities of graphical computer software were still limited during that time. Because no statistical package contains the

[33] Tufte (1997), p.53.
[34] Horber, Eugène (1990) Analyse exploratoire des données et sciences sociales, pp. 3.14-3.20/chap. 3: Interrogations sur l'hétéronomie des sciences sociales.
[35] This principle is also used by Tukey's stem-leaf-plot. See:Tukey, John W. (1977) Exploratory Data Analysis, chap. 1.
[36] Bertin, Jacques (1977) La graphique et le traitement graphique de l'information, pp.32-89.

module *"create a Bertin matrix"* I constructed this by hand using EXCEL. The job durations were considered as single episode data. Several steps had to be done:

(1) It was subsampled again from the complete file. A fraction of 1% (=140 data lines) was convenient to handle.

(2) A subset of variables was chosen: the duration variable DUR, the censoring information CEN, and the main grouping variables EW (East/West German), GENDER (male/female), COHORT.

(3) Then the numerical values were transported into ASCII characters to enable a fast production of coded[37] and sorted lists by any editor and sorting routine. To indicate the historical sequencing , birth cohorts were coded '##----' for 1940, '--##--' for 1950, and '----##' for 1960. "Bertin like", the dichotomous group variables were coded into small black and white fields[38] with black for East German, followed by black for male persons, and black if a given job is censored. The duration itself was recoded into a string of varying numbers of '#' characters, each '#' indicating a 6-months-interval. The characters '...?' were added at the end to visualize an incomplete observation.

(4) After this coding step, all durations were grouped and sorted because ordering, similarity, and proportional representation are essential features of a graphic display (Bertin 1977, p.177). It was first grouped by EW, then by GENDER, and last by COHORT. Thus, East/West and male/female patterns can be compared; the separate inspection of cohorts is necessary by survey design because longer job spells are primarily observed for older cohorts. Finally, all durations within each subgroup were sorted by ascending length.

Figure 4 'Matrix for Job Duration Patterns of a 1% Random Subsample' *(appendix C)* shows the coded job duration pattern matrix for the selected subgroups.

Which information does the graph display?

Obviously, the number of cases differs within each subgroup. The sampling process reflects different sizes of groups according to their initial mixture. But the low

---

[37] Tukey (1977, pp. 382-389) demonstrates examples for coding residuals.

[38] First, ASCII 0/1 were produced; black and white attributes were added within the EXCEL table.

sampling fraction of 1% underlines the exploratory purpose which wants to look at a reasonable number of jobs.[39]

A cohort comparison here is trivial: longer durations belong to the older cohorts. More interesting is the comparison between men and women. A quick glance at the East German data does not deliver remarkable gender differences.[40] But in West Germany, the job spells for women are shorter than for men (mainly in the 1950 cohort). This gives some hints that West German women interrupt or finish their careers more often than men do, and also than East German women do.

The most interesting information beside the intuitive understanding of different time spans can be seen from the distribution of censored spells. In each group, several strings with '...?' appear. Probably they are concentrated on the longer episodes, particularly for men of the 1950 cohort. The percentage of incomplete observations looks relatively high so that a calculated median duration (see descriptives in appendix A) would severely underestimate the true length of jobs.

This outline of a job duration pattern matrix could need further improvements:

- higher graphic compression than I could do by EXCEL. More data lines will deliver more reliable patterns.
- user selection of different sampling fractions
- interactivity to select subgroups
- interactivity to interchange lines and rearrange them visually (the Bertin idea).

Horber[41] considers especially interactivity as essential and underlines that it does not mean the pure communication between software user and computer, but *"a creative interaction between the results obtained and the user's knowledge"*.

[39] One could inspect several subsamples.
[40] The 1960 cohort contains too few cases for interpretations. One should try another subsample.
[41] Horber, Eugène (1991), The Computer Aided Instructor, about Didactical Aspects of EDA (which is an interactive software package for exploration of aggregate data , developed by Horber).

3.3 Censored versus Non-Censored Durations

After the inspection of a random subsample, now censored and uncensored spells will be examined for the complete data set. Again the same subgroups are chosen. A convenient way to visualize distributions of continuous variables is the stem-leaf-display.

An example demonstrates the principle.

Let

{11,11,11,12,12,12,12,13,15,19,20,20,20,20,20,21,21,23,23,23,23,23,23,24,24,27,28,28,28,31,31,39,39,39,39}

be a given set of observations. The stem-and-leaf plot uses the numbers themselves to create a picture of the frequency distribution:

```
stem & leaf              displayed observations are
 1 * 11122223            {11,11,11,12,12,12,12,13}
 1 . 59                  {15,19}
 2 * 000001133333344     {20,20,20,20,20,21,21,23,23,23,23,23,23,24,24}
 2 . 7888                {27,28,28,28}
 3 * 11                  {31,31}
 3 . 9999                {39,39,39,39}
```

Taking ASCII characters, no special graphic device is needed. The stem-and-leaf plot combines single values to a global view. It allows to inspect the symmetry, the concentration of data, possible gaps, and the spread of the values. Emerson/Hoaglin[42] (1983) vote for experimenting with different designs. As a visual compression for the given mass of job spells, I introduce "tiny" multiple stem-leaf-displays.

In figures 5a and 5b, **'Stacked Stem-Leaf-Diagrams : East and West German Men and Women of Birth Cohort 1940; East and West German Men and Women of Birth Cohorts 1950 and 1960'** *(appendix C)*, a very small printer font is chosen, to see together as many groups as possible. Additionally, values for non-censored and censored durations are stacked in one plot and contrasted by colour attribute[43]. The terminated job episodes are presented in black, the censored ones in red, so that typical patterns can be detected. The time axis (here vertical, the "stem") is kept constant for all groups to allow a direct comparison. The varying lengths of the vertical axis indicate the varying observed life-spans for East and West cohorts. The horizontal lines indicate the frequencies within each time interval.

---

42 Emerson, J.D./Hoaglin, D.C. , Stem-and-Leaf Displays, in: Hoaglin/Mosteller/Tukey (1983), Understanding Robust and Exploratory Data Analysis, pp.7-30.

43 Interactive software could also do this by highlighting.

The plot confirms the different group sizes already seen in figure 4 (for the small 1% random sample). Evidently, the distributions are positively skewed[44]. For West German men, censored data are concentrated on longer durations. This cannot be seen in the West female groups. It could be suspected again, that many West German women were out of labour force at the time of interview. The East German groups do not signal these differences. The 1950 and 1960 East German cohorts show a concentration of censored job spells on short durations. Since the interview date was in 1991, presumably many East Germans from these cohorts started new jobs around the date of the German re-unification. In this year, 1990, new economical conditions began in the East.

But I stop here with detailed sociological interpretations, since I mainly want to present the applied exploratory method. Moreover, exploration does not necessarily mean to interpret significant differences. The first aim is to check the data and to get experience with the scale of the variables. Several tries with other subgroups may help to decide which research questions are worth doing more exploratory work.

[44] This is typical for count or duration variables which are bounded to zero.

3.4 Visualizing Multiple Episodes: Career Patterns over the Life Course

Emphasizing the life course aspect, now the sequencing of job durations in personal careers will be examined. This also means higher graphical complexity. An example for visualizing dynamic processes in survival data is given by Anne I. Goldman[45] (1992). Her article presents "eventcharts" to visualize survival times after transplantation. Horizontal lines per individual are ordered by date of begin and indicate the duration. Different types of events (infection, death, or censoring) are symbolized by a specific character plotted on the duration line.

Inspired by Goldman's idea, I first converted the job spell file XPLOD into a person-related file. Each line contains a sequence of all jobs for this person. Again the first variables are the grouping information COHORT, EW, and GENDER. Then a string with a varying number of characters symbolizes a time axis. Each character denotes a state in a 6-month-interval beginning with the first job. "_" indicates being out of labour force. It may be anything like education or being unemployed etc. Successive numbers denote length of episode for job 1,2,...9, A for job 10, B for job 11 ... The character "c" was added at the end of a person's career, if the last job was censored.

An example for some data lines is listed below:

```
##____ E M 111111111111111111111111111111111111111111111111111c   cohort 1940,East,male:
                                                                  one job,censored
__##__ E F 111122222___33333333333333333333333333333444c          cohort 1950,East,female:
                                                                  4 jobs,censored,interruption
____## W F 111111111111222222222222223                            cohort 1960,West,female:
                                                                  3 jobs, terminated
<-------->  <-------------------------------------------------->
 group                 job sequences
 variables
```

I wrote a short Turbo Pascal programme[46] to rearrange the spell data into such a sequence of ASCII characters. The result file is independent from a specific computer system and may be sorted by any simple sorting routine. Two other file versions were produced: with a time axis as historical time , and with a time axis as individual age[47]. To find stable patterns, the sampling fraction[48] should be higher than before: 20 % turned out to be acceptable. The sampled lines were read into EXCEL. Within each group, they were sorted by the job duration characters. The optical group and censoring attributes of figure 4 were kept.

---

[45] Goldman,Anne I. (1992), Eventcharts: Visualizing Survival and Other Time-Events Data.
[46] Source code XPLOPER.PAS: see appendix B.
[47] Starting at January 1951 and at age 11 was sufficient to show the earliest jobs.
[48] The sampling procedure was done by theTurbo Pascal programme XPLRAND.PAS in appendix B.

My last technique could be called "shrinking". It was done by the photocopy machine because inventing any high-resolution graphic programme would be a time consuming work for itself. I call the resulting figures "skeleton plots" because they remind of X-ray diagnosis.

Figure 6 shows **'Grouped Skeleton Plots for Career Patterns over Historical Time'** *(appendix C)*. The predominant message is given by the pattern of being in or out of labour force. Whereas careers for all East German cohorts appear stable, once they were started, the West cohorts show tremendous differences for men and women. Women interrupted and stopped their careers, only for the 1960 cohort this behaviour seems to have changed. A vertical comparison allows to inspect directly selected historical years: whether cohorts were labour market competitors in a given historical period[49], or whether the labour force participation had changed over years.

In figure 7, **'Grouped Skeleton Plots for Career Patterns over Age'** *(appendix C)*, only women are compared (on the time axis in years of age). The shrinkage factor for this plot is not so high, thus more details are comparable. First, we see a higher age at entry into first job from cohort to cohort. Second, we also find the proportion of censored jobs at the right end. Third, the plots display the early termination of careers for West German women born around 1940 and 1950. In the 1940 East cohort, many women re-entered into work again after longer periods of interruptions. This can be seen as holes in the career patterns for this group.

The last skeleton plot, figure 8 **'Skeleton Plots with Slicing for Career Patterns since First Job'** *(appendix C)*, focusses on the comparison of only two groups: East and West German women born around 1950. The time axis is now the time since starting the first job. The data lines are presented in a bigger format, so that group patterns can be inspected together with single cases. The lines are sorted in descending order by length of duration of the first job. It can be examined whether first jobs tend to be short or long, whether they are followed immediately by another episode or by interruptions. The coding of episodes by their sequence numbers gives a

[49] It could be interesting to compare pre or post 1989 years in the East; periods of conservative versus social-democratic governments in the West; years of economical welfare versus crisis of the employment system.

visual impression of job mobility over life. In contrast to that, job mobility tables[50] compare only two time points.

I added a time slicer to the plot. This tool helps to focus on a specific point on the x-axis[51]. Summary measures are listed at the bottom of the slice box: "percentage of persons staying still in the first job", and "percentage of non-employed persons". Thus, abstract numbers become visible.

To summarize the techniques of the skeleton plots, I combined coding, sorting, and image compression to a resulting pattern. Additional slicing was demonstrated. Further developments could be:

- a definition of a slicer in a user box as Cook/Weisberg[52] (1994) presented
- linking other variables[53] to the "skeleton" plot by highlighting.

---

[50] For the German and Norwegian Life History Study, career mobility tables between classes are discussed in: Mayer, K.U. et el. (1989). Class Mobility During Working Life: A Comparison of Germany and Norway. Transition matrices for men of the birth cohorts 1930,1940,1950 see pp.231-235.

[51] Two slices are marked: for the 6-months periods beginning one year and 10 years after start of first job.

[52] Cook, R.D./Weisberg, S. (1994) use a slice-window for regression diagnostics with the R-code in: An Introduction to Regression Graphics, pp.24-33,79-81.

[53] Examples: highlighting part-time versus full-time jobs, high versus low wages, or marking different educational levels.

4. Deeper Explorations: Visualizing Statistical Functions for Duration Processes

After basic descriptions of the job durations in chapter 3, I will now turn to statistical summaries for the total time-dynamic process. Graphical displays of the survival function as well as the hazard rate (or transition rate) function will assist the exploratory process. A design of a window-oriented user interface for exploring survival data demonstrates the importance of interactivity.

In this chapter, the data will be explored as single-episode/single-destination data. No different destination states are given, only *censored* and *terminated* observations are distinguished.

4.1 The Kaplan-Meier Product-Limit Estimator

Given a duration variable,

let **f(t)** be the ***density function*** of the episode duration

and $\mathbf{F(t) = Pr(T \leq t) = \int_0^t f(u)du}$ the (cumulated) ***distribution function***.

For continuously measured durations, the ***survivor function*** is now given by

$$\mathbf{S(t) = 1 - F(t)} \ .$$

In the XPLOD data, this is the proportion of still ongoing job spells at time t given that no exiting event occurred before.[54] At the beginning of the process, at $t_0=0$, all observed spells are still at risk to finish so that $S(t_0)=1.0$. Then with continuing time t, S(t) is monotonically falling down as more and more jobs get terminated. Due to the presence of censored values, the question is how to calculate an empirical survivor function, because it is uncertain how long incompletely observed observations will remain .

The ***Kaplan-Meier product limit method***[55] estimates the survivor function as

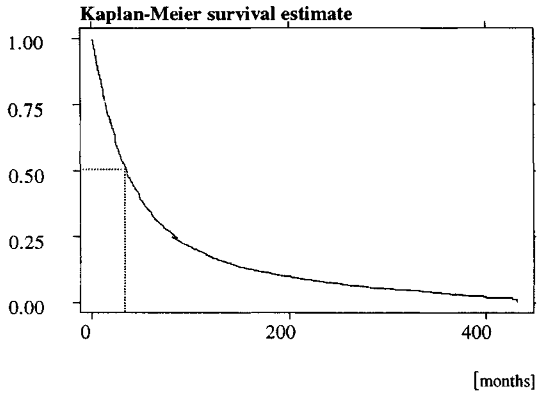
$\mathbf{S_{est}(t) = \prod_{j \mid t_j \leq t} \left((n_j - d_j) / n_j \right)}$	with:	n_t = number of episodes still alive at time t d_t = number of exits at time t

[54] Notation of formula according to Blossfeld/Hamerle/Mayer, Event History Analysis (1989, p. 31).

[55] Because the following examples were computed in STATA the formula is given in notation as in STATA Release 5, Reference manual [R-Z], p. 296. Mathematically equivalent but different in the interval limit, see the Kaplan-Meier estimator in Blossfeld/Hamerle/Mayer (1989, pp. 44-45) and Blossfeld/Rohwer (1995, pp.66-67).

The following figure shows the plot of the survivor function for all non-missing episodes in file XPLOD.

Figure 9: **Plot of the Survivor Function for Job Durations (N=14201) of East and West German Men and Women from Birth Cohorts 1940,1950,1960**



vertical axis: % of survivors
at 35 months: 50% of jobs are still surviving (estimated median duration)
source: GLHS / own calculations, 1998

The durations on the x-axis range from 0 to an observed maximum of more than 400 months. The steadily falling shape of the survival curve can be seen. Additionally, coordinates for the Kaplan-Meier estimate of <u>median duration</u> are marked. Here, the algorithm had calculated x_{medi}=35 months, where 50% of the origin set of durations are still "surviving".

Table 4 shows parts of the corresponding STATA programme listing. I highlighted the time and survivor function values where 25%, 50%, 75% of original number of episodes are still going on (Kaplan-Meier quartiles).

Table 4:

Calculating the Survivor Function by the Kaplan-Meier Product-Limit Estimator

failure time: **DUR**
failure/censor: **CEN**

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf.	Int.]	
1	14201	74	37	0.9948	0.0006	0.9935	0.9958	
2	14090	257	65	0.9766	0.0013	0.9740	0.9790	
3	13768	282	61	0.9566	0.0017	0.9532	0.9599	
4	13425	312	65	0.9344	0.0021	0.9302	0.9384	
5	13048	214	43	0.9191	0.0023	0.9145	0.9235	
6	12791	285	49	0.8986	0.0025	0.8935	0.9035	
7	12457	278	62	0.8786	0.0028	0.8730	0.8838	
8	12117	236	52	0.8614	0.0029	0.8556	0.8671	
9	11829	208	57	0.8463	0.0031	0.8402	0.8522	
10	11564	239	62	0.8288	0.0032	0.8224	0.8350	
11	11263	191	49	0.8147	0.0033	0.8082	0.8211	
12	11023	440	52	0.7822	0.0035	0.7752	0.7890	
13	10531	271	51	**0.7621**	0.0036	0.7549	0.7691	*75% still in job*
14	10209	206	56	**0.7467**	0.0037	0.7394	0.7539	*after 13.5 months (approx.)*
15	9947	191	64	0.7324	0.0038	0.7249	0.7397	
...	...	...	...		...	...	...	
35	6301	97	21	**0.5038**	0.0044	0.4952	0.5124	*median duration=35 months*
36	6183	220	22	**0.4859**	0.0044	0.4773	0.4945	
37	5941	110	14	0.4769	0.0044	0.4683	0.4855	
...	...	...	...		...	...	...	
83	2691	39	16	**0.2546**	0.0040	0.2468	0.2624	*25% still in job*
84	2636	56	9	**0.2492**	0.0040	0.2414	0.2570	*after 83.5 months (approx.)*
...	...	...	...	...	...	...	...	
436	2	0	1	0.0112	0.0053	0.0040	0.0260	
437	1	1	0	0.0000	.	.	.	

Table 5 demonstrates a by-hand-calculation of $S_{est}(t)$ for the first 4 months.

Table 5:

Estimating the Survivor Function by the Product-Limit Formula

time duration in months	**n_j** at begin still in job	**d_j** ending with event	cen sored	**$(n_j-d_j)/n_j$**	**product limit estimate**
1	14201	74	37	(14201-74)/14201 = .9948	**.9948**
2	14201-74-37 = 14090	257	65	(14090-257)/14090 = .9818	.9948 * .9818 = **.9767**
3	14090-257-65 = 13768	282	61	(13768-282)/13768 = 9795	.9766 * 9795 = **.9566**
4	13768-282-61 = 13425	312	65	(13425-312)/13425 = .9768	.9566 * .9768 = **.9344**
...	...	...	...	...	

It can be seen from the calculation procedure, that the algorithm keeps the censored episodes included in a given time interval, but in the following period, they are excluded from the set of still existing job spells.

The Kaplan-Meier product-limit estimator is non-parametric and only based on the observed data. It is often called an "exploratory" method. So Blossfeld/Rohwer (1995, p.51) argue that no distributional assumptions have to be made. In Atkinson's[56] (1995) article about graphics for exploratory survival analysis, the Kaplan-Meier estimator is used for dynamic displays of survival data. But the formula shows that in fact an assumption about the treatment of censored values is done. The influence of incompletely observed durations might be underestimated when a high proportion of censored data is given. (See also a discussion about this problem in Allison, 1984, pp.10-11.) To illustrate this argument, compare different median durations[57]: the Kaplan-Meier estimate is 35 months; the simply calculated median for all spells is 28 months; for terminated spells only, it is 26 months; and for the censored spells, it is 46 months.

Robust methods like simulations or resampling might help to clear up the uncertainty about censored data. With rising computer efficiency those methods become more and more practicable. For example, experiments with repeated simulations of artificially censored data might deliver interesting results for the 1920 cohort[58] of the GLHS. All job spells are completely terminated in this cohort. But the wide field of innovative methodology will be touched only briefly here.

Nevertheless, the estimated product-limit survival curve is useful to inspect, because the discussed assumptions are relatively low and the displayed axes are intuitively understandable. The x-axis represents the usual time continuum from the left to the right, the y-axis is the proportion of still existing jobs and ranges from 0 to 1.

---

[56] E. Neely Atkinson, Interactive Dynamic Graphics for Exploratory Survival Analysis in: The American Statistician, Vol. 49, No. 1, 1995. The presented graphics of this article will be used later.

[57] Sometimes it is called confusingly "average duration", but it is the time when 50% of the observed spells have not yet terminated.

[58] Persons were interviewed at age of 65 years so that they all had terminated their working life.This cohort is not included in the XPLOD file because data were only collected for West Germans.

4.2 Interactive Exploration of Survival Curves

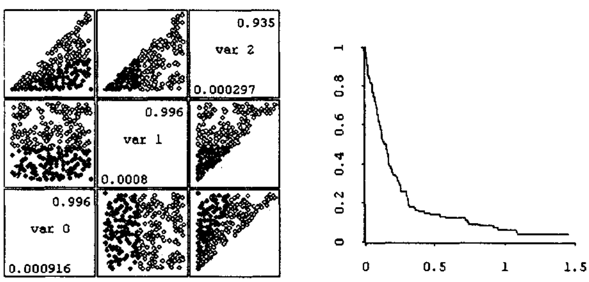
The shape of the survival curve in figure 9 leads to the question whether differences between groups exist, or whether influences of covariates can be detected. Linking a survival curve to other variables cannot simply be done by highlighting parts of the display like scatterplots do, because the survival curve has to be recalculated for each selected group.

E. Neely Atkinson[59] (1995) presents an XLISP-STAT[60] programme which allows to explore survival data interactively. Survival curves, censored box plots[61], and event charts can be linked with scatterplot matrices for categorical variables. In addition, a continuous variable can be dichotomized interactively by a slider, and two related survival curves will be displayed.

The following two figures show examples of the graphic screen as produced by Atkinson's programme[62] for arbitrarily generated test data. Figure 10a presents a survival curve which is calculated by the Kaplan-Meier product limit estimator, yet only for those cases highlighted in the linked scatterplot matrix.

Figure 10a

Interactive Graphics by E. Neely Atkinson:
Scatterplot Matrix Linked with Kaplan-Meier Survival Curve



programme: expsurv.lsp, test data *Survival curve is calculated only for selected highlighted cases.*

Figure 10b shows to the left the distribution of a continuous variable with a dichotomization cutpoint at 0.589. Two related survival curves for cases below and above this cutpoint are displayed.

---

59 See reference in 4.1.1, p.27.

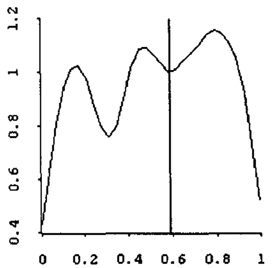
60 XLISP-STAT by Tierney, Luke (1988), A Statistical Environment on the XLISP Language.

61 Quartiles for the boxplot are calculated by the Kaplan-Meier estimator.

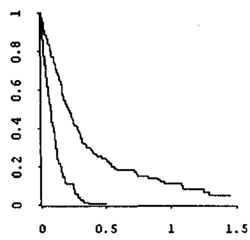
62 E.N. Atkinson, programme source code: *expsurv.lsp*

Figure 10b

Interactive Graphics by E. Neely Atkinson:
Dichotomized Variable Linked with two Corresponding Survival Curves



Cutpoint (here .589) can be moved in a slider.



Two corresponding curves, below and above cutpoint, are calculated.

programme: expsurv.lsp, test data

In contrast to clinical studies (which often have only a hundred patients), the job spell file XPLOD again turned out to be too large. Experiments with the *expsurv* programme were not satisfying concerning programme performance, so that again a random sample would have been necessary. In addition, XLISP commands[63] still give a feeling of early computer days. It is at least an open question, whether "statisticians who are comfortable with computer systems but no professional programmers ... implement such techniques in a reasonable amount of time", as Atkinson states. But nevertheless, Atkinson brought exploratory innovation to well known survival plots by introducing interactivity. Marking groups and seeing immediately how a survival curve changes can only be weakly replaced here by paper.

In the following figure, I designed an *interactive window* for exploring survival data. It was inspired by

- standardized windows for special tasks as realized in the R-Code by Cook and Weisberg
- user friendly easily feasible data handling
- an interactive combination of survival plots as innovated by Atkinson.

[63] An example: ***(defvar sckml (scat-km length status x y z))***

See figure 11 *(appendix C)*:

'Design of a Survival Data Exploration Window:

Examining the Time Span from Entry into Labour Market to Birth of First Child'

Explanation of the demonstrated SX window:

Underlying research question

A new duration variable will now be examined: DURKID1. This is the time span from the starting point of a person's first job until the birth of the first child.[64] If no child was born until interview date, the spell was taken as censored with a given duration from first job to interview date (censoring variable CENKID1=0).

Because the event of the birth of a child has a higher impact on female than on male careers, only women are considered here. To exclude the effect of different historical periods, only the 1960 cohort[65] was selected. East/West differences and possible influences of the level of education on careers will be examined.

User Windows

file	File name, information about type of survival data (single episode), and the names of the duration and censor/status variables are clicked by the user.
sampling	A sampling fraction can be chosen by a slider. Here the total sample is taken because the selected subgroup contains only 707 cases. The "sample once / repeated sampling" buttons indicate the possibility of integrated repeated sampling. This might be used to inspect by eye whether trends and results become stable.
groups	The user may click on a maximum of three variable names for subgroups, here COHORT, GENDER, and EW. To define a data filter, corresponding value labels are marked in red, here only women from the 1960 cohort. The value labels for the group comparison are marked in blue, here East versus West. I suggest this window[66] instead of scatterplot matrices for marking categorical variables.
covariates	Within the two selected groups, the influence of the level of education (given in variable SCHOOLR) will be examined. The user marks different value labels by different colours. They will be used for corresponding groups in the results window.
cut-point	No dichotomization of a variable is examined here. The window was designed to allow further useful explorations as done by Atkinson.

[64] To avoid complications, persons who had already a child before starting their first job, were excluded. This was the case for 8.5% of the women of the 1960 cohort.
[65] The data collection for this cohort took place in 1989 (West) and 1991 (East).
[66] A more advanced way for visualizing selections would be marking groups in a linked histogram.

Result Windows

The user-selected survival curves are presented in several windows. The first one, to the right of the group selection window, shows two survival curves for East and West German women. The curve is falling much more rapidly for the East group. Already 33 months after starting their first job, 50% of the East German women get their first child. This is the case not before 112 (!) months in the West German group. Obviously a big difference! Coordinates for median survival times in both groups are added in red (click on/off option). To the right of the covariates window, the results are further differentiated by level of education. The curves are marked by the same colour[67] as chosen in the covariate window. Also the values for median survival times are displayed in the colour belonging to the related educational level.

In general, also the earlier trend to a child in East Germany can be seen. But obviously, the influence of level of education is different. In West Germany, there is nearly no distinction between educational groups. All women seem to wait with a child, once they have started their careers. The three survival curves are even identical during the first 50 months with only slight differences later on. Not only women with high school degrees[68] wait, but also less educated persons. For younger women, West Germany does not seem to offer practicable compatibility of being employed, keeping the economic wealth of the household, and having children. Contrary to this, the survival curves for East German women clearly differ by education. The higher the school degree is, the earlier after starting the first job women get their first child. Evidently, longer education times do not lead to a delayed process of familiy founding. The former GDR seems to have offered better possibilities for young women to stay in the employment system. These findings are confirmed by Huinink[69] (1995): in the 1960 cohort, 65% of <u>West</u> German women at age of 26 years did not yet have born a child, but only 17% of the East German women. Huinink interprets different economic and institutional backgrounds as a reason for the decision to have a child.

---

[67] Only coloured labels in textfileds could be included. Unfortunately, the colour attribute of the curves got lost by transportation through several software products. Of course a future realization of the presented design should programme coloured curves.

[68] Here "Abitur", which is the German qualification for starting studying at a university.

[69] Huinink, J. (1995), Warum noch Familie? Zur Attraktivität von Partnerschaft und Elternschaft in unserer Gesellschaft; see cited tables on pp. 229-236.

The discussed research question is only one example for lots of interesting relationships. The survival data exploration window demonstrates how useful a tool can be to select groups and covariates easily, and to inspect directly the change of the corresponding survival plot.

4.3 Inspecting Hazard Rates

The last applied method is the exploration of hazard rates. Whereas survival curves visualize the view into the past, asking: "How many episodes have lasted how long?", the same data get another focus putting the question: "At what time is the risk higher or lower to terminate an episode?" The concept of the risk combines the knowledge about the past with the uncertainty about the future.

This risk , called hazard rate or transition rate, is given by $\lambda(t)$,

and is – roughly spoken – the ratio between those episodes terminating and those surviving at given time t. The mathematical connection[70] between density function, survivor function and hazard rate is :

$$\lambda(t) = f(t) / S(t)$$

The shape of this hazard function over time is a basic theoretic assumption in statistical models for survival data. Usually, parameters in this function are estimated by the maximum likelihood method. Some examples for shapes of transition rates[71] are:

- exponential models for a constant risk over time
- piecewise constant models for a constant risk within defined intervals of time
- the Weibull distribution for monotonically falling or rising risk (depending on a parameter)
- a log-logistic distribution for a rising risk at the beginning of a process and later on falling

[70] See Blossfeld/Hamerle/Mayer (1989, p.33).

[71] Mathematical formulas can be seen in Blossfeld/Hamerle/Mayer (1989, pp.31-42). An intense description of various parametric models is given by Blossfeld/Rohwer (1995, chapter 7, pp.162-197). The TDA programme by Rohwer allows the specifications of many transition functions.

- the Cox proportional hazards model with given baseline rate and different, but proportional risks in subgroups.

Will the choice of the correct model be done by a-priori thinking or by creating ideas from data exploration? Graphic inspections of the empirical hazard rate may help to decide. The empirical transition rate is easily calculated by the Life-Table method in the following way:

Let the time axis be devided into k intervals with the given numbers

d_k	number of exits during the k-th interval
w_k	number of censored episodes during the k-th interval
R_k	risk set of the k-th interval = number of episodes which had no event at the beginning of the k-th interval and have not been censored

The ***Life-Table method***[72] estimates the hazard rate as:

$$\boldsymbol{\lambda_k^{est} = d_k / (R_k - w_k/2)}$$

The values of the rates may be higher than 1 because they are no probabilities, but the ratio between terminating and surviving objects.

Similarly to the survival curves, also the hazard rates will be inspected in the exploration window.

See figure 12 *(appendix C)*

'Design of a Survival Data Exploration Window:

Examining Hazard Rates for Birth of First Child after Entry into Labour Market'

The same definitions are done in the user windows for the data set, sampling fraction, groups (COHORT, GENDER, EW), and covariates (3 categories of SCHOOLR). The results for the estimated hazard rates are presented in the lower part of the user window. As usual, the user should be able to move and resize all windows. It can be seen clearly in the West group for all 3 educational levels, that the hazard rate is very low for the first 50 months and later rising only quite moderately. Just a different pattern can be observed for the East German women: the curve of the rate is zig-

[72] Detailed derivation of this formula in Blossfeld/Hamerle/Mayer (1989, p.43).

zagging. The highest chance for the birth of the first child is observed after about 90 months in the group of the lowest educational level (which are only a few persons). In the group of the grade 10 degree (which was the usual case for most persons in the GDR), the highest transition rate is observed after 50 months. Then it is falling down and again rising to a second peak at about 125 months. Women with high school education tend to get their first child about 50 months after the start of their first job. Then the risk is instantaneously falling down. These differences indicate that a proportional hazards assumption is not appropriate for the educational variable as well as for the East/West comparison.

Thinking on theoretical assumptions about the nature of the hazard rate over time, this exploration might help to choose a model. Blossfeld/Hamerle/Mayer (1989) regard it as the "key concept" in the analysis of event history data. Of course, exploratory work challenges numerous repetitions. User windows are a comfortable interface for the researcher/computer dialogue.

5. Conclusion: Demands for Future Exploratory Software

The explorations of job spells of the German Life History Study were confronted with the problem: how to see many variables in the course of time and to compare groups in an overall view? Additionally, I had to cope with the problem of censored data and the data mass itself.

The presented graphics had been developed in two aspects. In the first part for a descriptive point of view: to inspect individuals and birth cohorts, to visualize a time axis, to describe censored and uncensored job durations, to contrast groups, and to describe persons vs. spells. I used matrix displays to raise dimensionality, combinations of plots, pattern matrices to visualize durations, and colour attributes to improve visibility. In the second part, the statistical *survivor* functions and *hazard rate* functions were considered and compared by groups.

All of the plots had been constructed "manually" by a bunch of different computer programmes. Whereas most software packages have implemented modules for survival models, there is a lack of exploratory tools for this kind of data. First steps are done by STATA with the commands for survival data, and by Atkinson with an XLISP-STAT programme. With my paper presentations of static plots, I can only contribute some concepts to the exploratory work. More effort on visual tools by software developers are desirable.

One of the most important tools, interactivity, is indicated by the design of a survival data exploration window (figures 11 and 12). The advantages of an interactive user interface are to change groups and variables very quickly, to combine different types of plots, and to see more dimensions by adding movement. A user-driven window design allows individual perception preferences. An additional option should integrate resampling methods. I used this mainly to build exploratory subsets of data, but it is also a technique for finding robust trends.

The computer development of the last 20 years teaches, that "large" data sets of today will get smaller and "slow" programmes soon will become fast. Therefore graphic software achieves better and better performance. Thus, data exploration will be possible even with complex data structures. Visualization assists human mind to step from concrete data to abstract ideas, to go back to the concrete, and to gain new insight.

References

Allison, Paul D. (1984) Event History Analysis. Regression for Longitudinal Event Data. Sage Publications, Newbury Park, pp.10-11.

Atkinson, E.Neely (1995) Interactive Dynamic Graphics for Exploratory Survival Analysis. in: The American Statistician, Vol. 49, No. 1, pp. 77-84.

Atkinson, E.Neely (1997) Interactive Dynamic Graphics for Exploratory Survival Analysis. programme source code: ftp://odin.mdacc.tmc.edu/pub/exp-surv/expsurv.lsp

Auriat, Nadia (1996) Les défaillances de la mémoire humaine. Aspects cognitifs des enquêtes rétrospectives. Éditions de l'Institut National d'Etudes Démographiques. Paris, pp. 139-145,152-157.

Bertin, Jacques (1977) La graphique et le traitement graphique de l'information. France. Flammarion, Paris, pp. 32-89,177.

Blasius, Jörg (1995) Visualizing Panel, Trend, and Event History Data with Correspondence Analysis. in: Faulbaum, F./Bandilla,W. (Eds.), Softstat '95, Advances in Statistical Software 5, Lucius & Lucius, Stuttgart, pp. 81-88.

Blossfeld, Hans-Peter /Hamerle, Alfred /Mayer, Karl Ulrich (1989) Event History Analysis. Lawrence Erlbaum, Hillsdale, New Jersey, pp. 14-45.

Blossfeld, Hans-Peter /Rohwer, Götz (1995) Techniques of Event History Modeling. Lawrence Erlbaum, Mahwah NJ, pp. 11-26,33-51,66-67,162-197.

Brückner, Erika (1993) Lebensverläufe und gesellschaftlicher Wandel - Konzeption, Design und Methodik der Erhebung von Lebensverläufen der Geburtsjahrgänge 1919-1921. Teile I-V. Berlin: Max-Planck-Institut für Bildungsforschung, Materialien aus der Bildungsforschung Nr. 44, pp. 27,115-127,178-193.

Brückner, Hannah (1995) Surveys Don't Lie, People Do? An Analysis of Data Quality in a Retrospective Life Course Study. Berlin: Max-Planck-Institut für Bildungsforschung, Materialien aus der Bildungsforschung Nr. 50, pp. 28-33,54,73-78.

Brückner, Hannah /Mayer, Karl Ulrich (1995) Lebensverläufe und gesellschaftlicher Wandel - Konzeption, Design und Methodik der Erhebung von Lebensverläufen der Geburtsjahrgänge 1954-56 und 1959-61. Teile I-III. Berlin: Max-Planck-Institut für Bildungsforschung, Materialien aus der Bildungsforschung Nr. 48.

Cook, Dennis R. /Weisberg, Sanford (1994) An Introduction to Regression Graphics. Wiley & Sons, New York, pp. 24-33,79-81.

Emerson, John D. /Hoaglin, David C. (1983) in: Hoaglin, David C. /Mosteller, Frederick /Tukey, John W. Understanding Robust and Exploratory Data Analysis. Wiley & Sons, New York, pp. 7-30.

Erickson,B.H. /Nosanchuk,T.A. (1992) Understanding Data. Second Edition, Open University Press, Buckingham, pp. 3-11.

Esser, Hartmut (1996) Soziologie. Allgemeine Grundlagen. Campus, Frankfurt, pp. 255-266.

Francis, Brian /Fuller, Mark (1996) Visualization of Event Histories. Journal of Royal Statistical Society, A, 159, pp.301-308.

Goldman, Anne I. (1992) Eventcharts: Visualizing Survival and Other Time-Events Data. The American Statistician, Vol.46, No.1, pp. 13-18.

Hartwig, Frederick with Dearing, Brian E. (1979) Exploratory Data Analysis. Sage Publications, Newbury Park, pp. 9-13.

Horber, Eugène (1987) Exploring Aggregate Data, Conceptual and Practical Aspects. in: European Political Data Newsletter, No. 63, Bergen, pp. 30-36.

Horber, Eugène (1990) Analyse exploratoire des données et sciences sociales. Vers une approche méthodologique pragmatique. Thèse de doctorat, Université de Genève, pp. 3.14-3.20.

Horber, Eugène (1991) The Computer Aided Instructor. in: Best, Heinrich /Mochmann, Ekkehard /Thaller, Manfred. Computers in the Humanities and the Social Sciences, Proceedings of the Cologne Computer Conference 1988. K.G. Saur, München, pp. 449-453.

Huinink, Johannes (1995) Warum noch Familie? Zur Attraktivität von Partnerschaft und Elternschaft in unserer Gesellschaft. Campus, Frankfurt New York, pp. 229-236.

Mayer, Karl Ulrich et al. (1989) Class Mobility During Working Life: A Comparison of Germany and Norway. In: Kohn, Melvin L. (ed.): Cross-national research in sociology. Sage Publications, Newbury Park, pp. 231-235.

Mayer, Karl Ulrich /Brückner, Erika (1989) Lebensverläufe und Wohlfahrtsentwicklung. Konzeption, Design und Methodik der Erhebung von Lebensverläufen der Geburtsjahrgänge 1929-1931, 1939-1941, 1949-1951. Teile I-III. Berlin: Max-Planck-Institut für Bildungsforschung, Materialien aus der Bildungsforschung Nr. 35.

Microsoft EXCEL 97 (1985-1997) Windows 95 Version, Microsoft Corporation.

Mooney, Christopher Z. /Duval, Robert D. (1993) Bootstrapping. A Nonparametric Approach to Statistical Inference. Sage Publications, Newbury Park, pp. 20-22.

Norusis, Marija J. /SPSS Inc. (1993) SPSS for Windows: Base System User's Guide, Release 6.0/ Version 7.5: online help.

SIR (1993) Version 3.2. Database Reference Manual, SIR Pty. Ltd, Sydney.

Solga, Heike (1996) Lebensverläufe und historischer Wandel in der ehemaligen DDR. ZA-Information 38, pp. 28-38.

STATA (1997) Release 5, User's Guide, Reference Manuals Volume I-III, STATA Press, College Station,Texas.

Tierney, Luke (1988) XLISP-STAT. A Statistical Environment on the XLISP Language (version 2.0). University of Minnesota, School of Statistics, Technical Report Number 528. The technical reports can be found in the internet: http://www.stat.umn.edu/PAPERS/papers.html

Tukey, John W. (1977) Exploratory Data Analysis. Addison-Wesley Publishing Company, Reading, Mass., pp. i-viii,1-25,39-56, 382-389.

Tufte, Edward R. (1997) Visual Explanations. Images and Quantities, Evidence and Narrative. Graphics Press, Cheshire, Connecticut. pp. 53,75-78.

Wagner, Michael (1996) Lebensverläufe und gesellschaftlicher Wandel: Die westdeutschen Teilstudien. ZA-Information 38, pp. 20-27.

APPENDIX

Appendix A: Codebook for the Used Data Set XPLOD

(1) Descriptives for file XPLOD

Variable	Label	Mean	Median	Min	Max	Valid N	% Missing
CASID	respondent's ID			10003	192038	14214	
job spell related variables, varying for each episode				*(100% are 14214 job spells)*			
JOBNR	current job nr in person's career	2,8	2	1	19	14214	
CEN	right censored?			0	1	14214	
LAST	last job?			0	1	14214	
TS	time starting job in months since 1900	903,53	921	617	1112	14207	
TF	time finishing job in months since 1900	953,27	982	644	1115	14207	
DUR	duration of job	50,68	28	1	437	14201	
LFX	labour force experience on entry to job	60,6	31	0	457	14201	
INCS	income starting job	963,7	800	19	20000	12230	???
INCF	income finishing job	1186,59	1000	25	39000	11369	????
HWEEK	weekly working hours	42,88	42	3	86	13192	??
FSIZE	size of firm	959,4	80	1	80000	12738	???
person related variables, constant for all episodes with same case id				*(100% are 14214 job spells)*			
TOTJOBS	total number of person's job spells	4,59	4	1	19	14214	
TBIRTH	time of birth in months since 1900			469	764	14214	
COHORT	birth cohort			1940	1960	14214	
SEX	gender			1	2	14214	
EW	east or west?			1	2	14214	
SCHOOLR	educational degree (recoded)	2,63	2	1	5	14049	?
VOCAT1R	degree 1. vocat.training (recoded)	3,98	5	1	6	13684	?
VOCAT2R	degree last vocat.training (recoded)	4,17	5	1	6	13505	?
TSCHOOL1	time finishing school	797,56	810	616	1002	14116	
TVOCAT1	time finish. 1.voc.train.	836,24	845	646	1073	13149	??
TVOCAT2	time finish. last voc.train.	887,18	907	650	1111	13170	??
TMAR1	time of 1. marriage	874,58	871	673	1107	12210	???
TKID1	time of birth 1.child	879,89	881	540	1099	11281	?????
descriptives for person related variables only counted for individuals				*(100% are 4119 persons)*			
CASID	respondent's ID			10003	192038	4119	
TOTJOBS	total number of person's job spells	3,45	3	1	19	4119	
TBIRTH	time of birth in months since 1900			469	764	4119	
COHORT	birth cohort			1940	1960	4119	
SEX	gender			1	2	4119	
EW	east or west?			1	2	4119	
SCHOOLR	educational degree (recoded)	2,7	2	1	5	4068	?
VOCAT1R	degree 1. vocat.training (recoded)	3,88	5	1	6	3940	?
VOCAT2R	degree last vocat.training (recoded)	4,07	5	1	6	3864	??
TSCHOOL1	time finishing school	809,95	819	616	1002	4094	
TVOCAT1	time finish. 1.voc.train.	849,62	854	646	1073	3788	??
TVOCAT2	time finish. last voc.train.	893,04	918	650	1111	3792	??
TMAR1	time of 1. marriage	882,53	880	673	1107	3407	????
TKID1	time of birth 1.child	887,94	892	540	1099	3087	??????

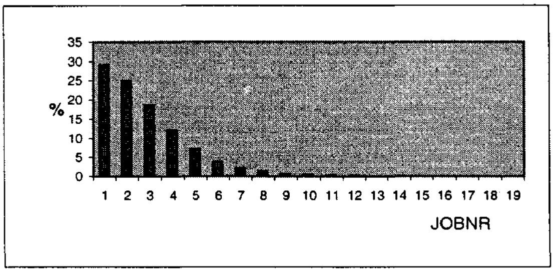
means and medians are only listed if adequate to type of variable

% missing is shown as follows:

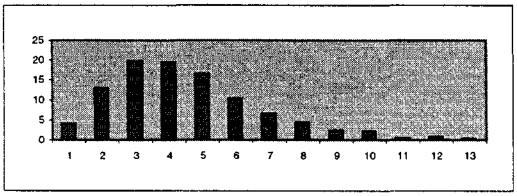
blank=	less than 1%	????=	15% - under 20%
?=	1% - under 5%	?????=	20% - under 25%
??=	5% - under 10%	??????=	25% - under 30%
???=	10% - under 15%		

(2) Codes and Frequencies for file XPLOD (N=14214 job spells)

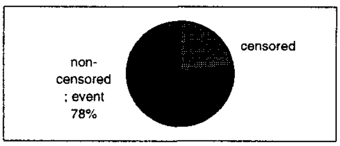
Variable	Value Label	Value	Frequ	%	Valid %
JOBNR current job nr in person's caree		1	4119	29	29
		2	3539	24,9	24,9
		3	2627	18,5	18,5
		4	1700	12	12
		5	1012	7,1	7,1
		6	543	3,8	3,8
		7	301	2,1	2,1
		8	168	1,2	1,2
		9	91	0,6	0,6
		10	54	0,4	0,4
		11	23	0,2	0,2
		12	16	0,1	0,1
		13	7	0	0
		14	4	0	0
		15	4	0	0
		16	2	0	0
		17	2	0	0
		18	1	0	0
		19	1	0	0
		Total	14214	100	100



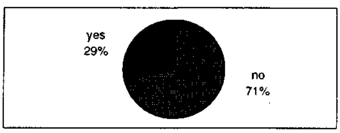
Variable	Value Label	Value	Frequ	%	Valid %
TOTJOBS total number of person's job spell		1	580	4,10	4,1
		2	1824	12,8	12,8
		3	2781	19,6	19,6
		4	2752	19,4	19,4
		5	2345	16,5	16,5
		6	1452	10,2	10,2
		7	931	6,5	6,5
		8	616	4,3	4,3
		9	333	2,3	2,3
		10	310	2,2	2,2
		11	77	0,5	0,5
		12	108	0,8	0,8
		13	39	0,3	0,3
		15	30	0,2	0,2
		17	17	0,1	0,1
		19	19	0,1	0,1
		Total	14214	100	100



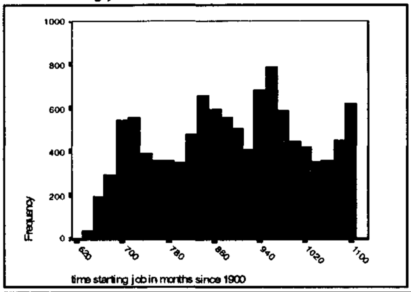
Variable	Value Label	Value	Frequ	%	Valid %
CEN right censored?	censored	0	3180	22,4	22,4
	non-censored; even	1	11034	77,6	77,6
		Total	14214	100	100



Variable	Value Label	Value	Frequ	%	Valid %
LAST last job?	no	0	10095	71	71
	yes	1	4119	29	29
		Total	14214	100	100

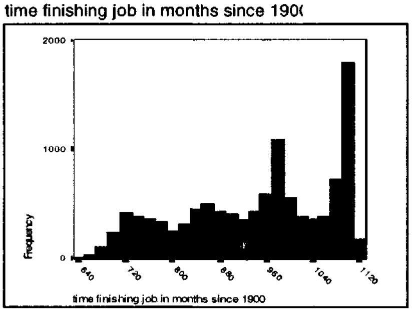


TS
time starting job in months since 190(



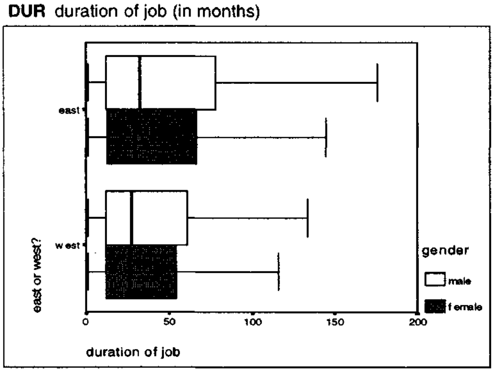
May 51	617	1	0	0
	633	2	0	0
	640	3	0	0
	... etc...			
	1111	2	0	0
Aug 92	1112	1	0	0
missing	-9	7	0	Miss.
	Total	14214	100	100

TF
time finishing job in months since 190(



Aug 53	644	1	0	0
	646	2	0	0
	648	1	0	0
	... etc...			
	1114	13	0,1	0,1
Nov 92	1115	1	0	0
missing	-9	7	0	Miss.
	Total	14214	100	100

DUR duration of job (in months)



	1	111	0,8	0,8
	2	322	2,3	2,3
	3	343	2,4	2,4
	4	377	2,7	2,7
	... etc...			
	436	1	0	0
	437	1	0	0
missing	-9	13	0,1	Miss.
	Total	14214	100	100

extremes and outliers are left out

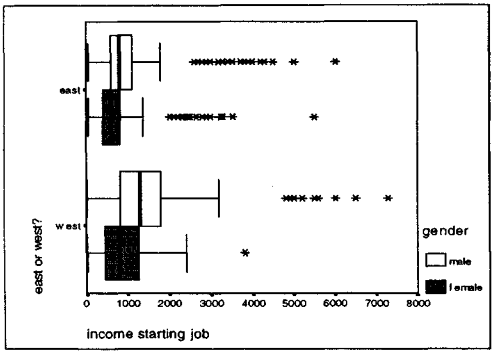
LFX
months of labour force experience on entry to jc

	0	4112	28,9	29
	1	25	0,2	0,2
	2	112	0,8	0,8
	3	118	0,8	0,8
	4	134	0,9	0,9
	... etc...			
	444	1	0	0
	457	1	0	0
missing	-9	13	0,1	Miss.
	Total	14214	100	100

extremes and outliers are left out

INCS income when starting job

monthly net income measured in DM for West German jobs and GDR mark for East German jobs

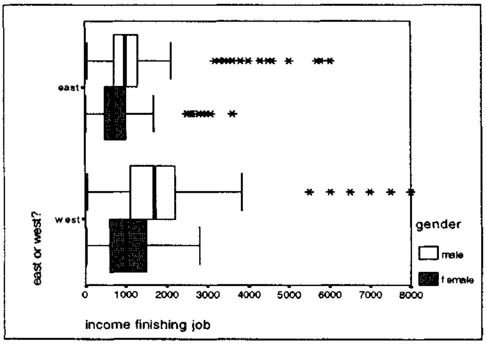


	19	1	0	0
	20	3	0	0
	25	7	0	0,1
	30	13	0,1	0,1
	35	4	0	0
	40	12	0,1	0,1
	... etc...			
	5600	1	0	0
	6000	4	0	0
	6500	1	0	0
	7300	1	0	0
	10000	1	0	0
	12000	1	0	0
	12500	1	0	0
	20000	1	0	0
missing	-9	1984	14	Miss.
	Total	14214	100	100

only extremes; outliers are left out

INCF income when finishing job

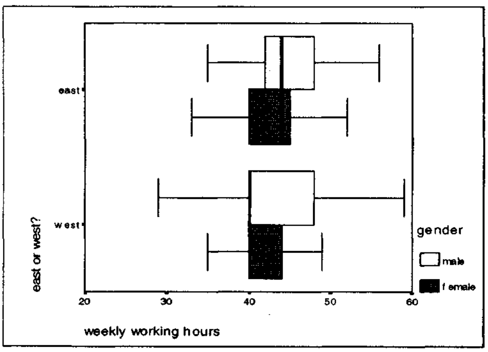
monthly net income measured in DM for West German jobs and GDR mark for East German jobs



	25	3	0	0
	30	8	0,1	0,1
	35	2	0	0
	40	5	0	0
	50	12	0,1	0,1
	... etc...			
	7500	1	0	0
	8000	2	0	0
	12000	2	0	0
	14500	1	0	0
	20000	1	0	0
	39000	1	0	0
missing	-9	2845	20	Miss.
	Total	14214	100	100

only extremes; outliers are left out

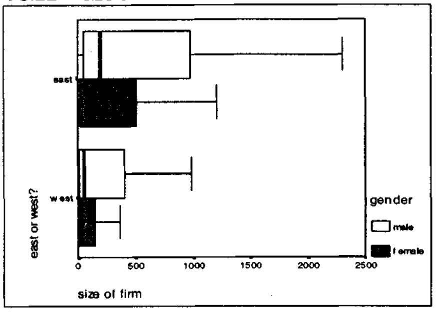
HWEEK weekly working hours



	3	2	0	0
	4	2	0	0
	5	6	0	0
	6	12	0,1	0,1
	7	5	0	0
	8	19	0,1	0,1
	9	13	0,1	0,1
	10	25	0,2	0,2
	... etc...			
	81	1	0	0
	84	22	0,2	0,2
	85	8	0,1	0,1
	86	1	0	0
missing	-9	1022	7,2	Miss.
	Total	14214	100	100

extremes and outliers are left out

FSIZE size of firm



	1	207	1,5	1,6
	2	317	2,2	2,5
	3	389	2,7	3,1
	4	374	2,6	2,9
	5	372	2,6	2,9
	... etc...			
	65000	1	0	0
	66666	1	0	0
	70000	1	0	0
	71000	3	0	0
	78000	1	0	0
	80000	1	0	0
missing	-9	1476	10,4	Miss.
	Total	14214	100	100

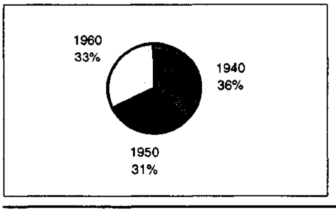
extremes and outliers are left out

TBIRTH time of birth in months since 190(

Jan 39	469	153	1,1	1,1
	470	120	0,8	0,8
	471	167	1,2	1,2
	472	143	1	1
	... etc...			
	744	131	0,9	0,9
	750	3	0	0
Aug 63	764	3	0	0
	Total	14214	100	100

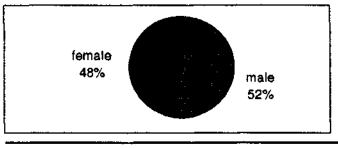
COHORT birth cohort

1939-41	1940	5217	36,7	36,7
1949-51west / 1951-53 eas	1950	4360	30,7	30,7
1959-61west / 1959-63 eas	1960	4637	32,6	32,6
	Total	14214	100	100



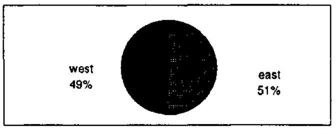
SEX gender

male	1	7346	51,7	51,7
female	2	6868	48,3	48,3
	Total	14214	100	100



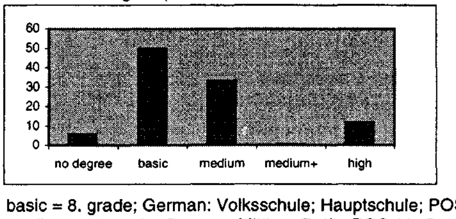
EW east or west?

east	1	7280	51,2	51,2
west	2	6934	48,8	48,8
	Total	14214	100	100



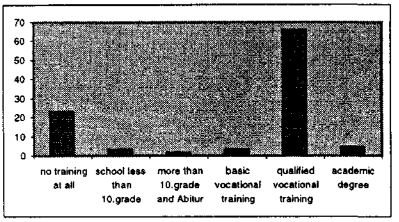
SCHOOLR
educational degree (recodec

no degree	1	800	5,6	5,7
basic	2	6953	48,9	49,5
medium	3	4613	32,5	32,8
medium+	4	58	0,4	0,4
high	5	1625	11,4	11,6
missing	-9	165	1,2	Miss.
	Total	14214	100	100



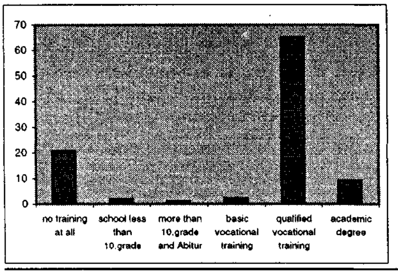
basic = 8. grade; German: Volksschule; Hauptschule; POS(Eas
medium =10.grade; German: Mittlere Reife; POS 10. Grade (Eas
medium+ = more than 10. Grade, but no Abitu
high = Abitur (entrance degree to universit)

VOCAT1R
degree of 1. vocational training (recode(



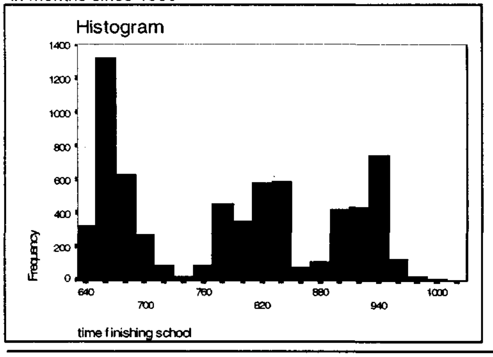
no training at all	1	3104	21,8	22,7
school less than 10.grade	2	428	3	3,1
more than 10.grade and Abitur	3	181	1,3	1,3
basic vocational training	4	421	3	3,1
qualified vocational training	5	8958	63	65,5
academic degree	6	592	4,2	4,3
missing or unclear	-9	530	3,7	Miss.
	Total	14214	100	100

VOCAT2R
degree of last vocational training (recode(



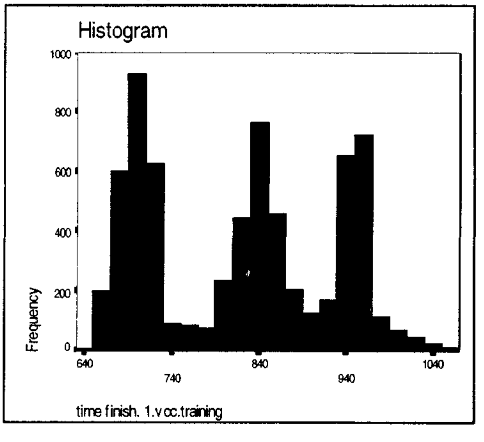
no training at all	1	2780	19,6	20,6
school less than 10.grade	2	256	1,8	1,9
more than 10.grade and Abitur	3	155	1,1	1,1
basic vocational training	4	275	1,9	2
qualified vocational training	5	8792	61,9	65,1
academic degree	6	1247	8,8	9,2
missing or unclear	-9	709	5	Miss.
	Total	14214	100	100

TSCHOOL1 time when finishing school
in months since 190(



Apr 51	616	3	0	0
	631	3	0	0
	632	14	0,1	0,1
	638	2	0	0
	... etc...			
	991	2	0	0
	1000	3	0	0
	1001	1	0	0
Jun 83	1002	2	0	0
missing	-9	98	0,7	Miss.
	Total	14214	100	100

TVOCAT1 time when finishing 1.
vocational training in months since 190

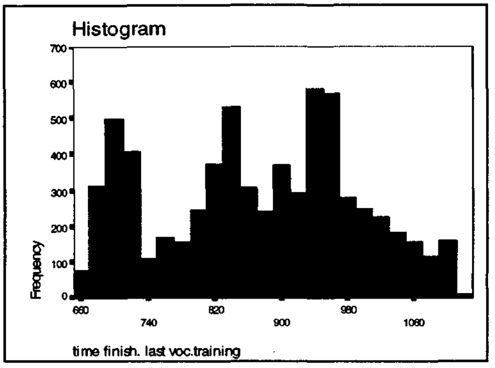


Oct 53	646	1	0	0
	650	7	0	0,1
	652	3	0	0
	653	3	0	0
	... etc...			
	1068	3	0	0
	1069	3	0	0
	1072	2	0	0
May 89	1073	2	0	0
missing	-9	1065	7,5	Miss.
	Total	14214	100	100

TVOCAT2 time when finishing last

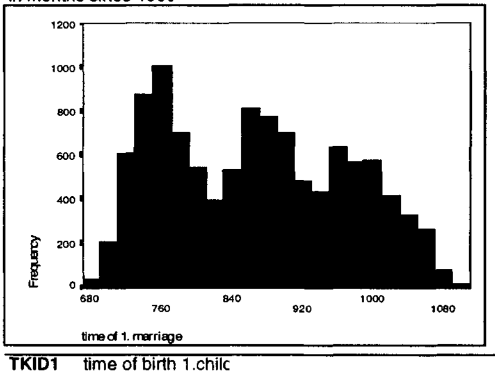
Feb 54	650	7	0	0,1

vocational training in months since 190



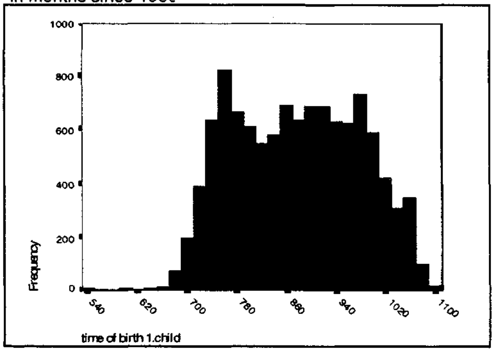
	653	3	0	0
	655	11	0,1	0,1
	656	2	0	0
	... etc...			
	1109	18	0,1	0,1
	1110	9	0,1	0,1
Jul 92	1111	14	0,1	0,1
missing	-9	1044	7,3	Miss.
	Total	14214	100	100

TMAR1 time of 1. marriage
in months since 190C



Jan 56	673	2	0	0
	676	1	0	0
	677	1	0	0
	... etc...			
	1102	2	0	0
	1103	6	0	0
Mar 92	1107	3	0	0
never	-9	2004	14,1	Miss.
	Total	14214	100	100

TKID1 time of birth 1.chilc
in months since 190C



Dec 45	540	5	0	0
	596	4	0	0
	638	2	0	0
	... etc...			
	1096	6	0	0,1
	1098	3	0	0
Jul 91	1099	4	0	0
no kids	-9	2933	20,6	Miss.
	Total	14214	100	100

```
(***********************************************************************)
(* PASCAL programme  XPLOPER.PAS                                       *)
(* ----------------------------                                        *)
(* reads the job-spell data set and builds a person related file       *)
(* with career patterns                                                *)
(*                                                                     *)
(* Turbo Pascal Programme                                              *)
(* written by:          Sigrid Wehner                                  *)
(*                      Max Planck Institute for Human Development     *)
(* stand:               28.Nov 1997                                    *)
(***********************************************************************)
(* infile: XPLOD.DAT                                                   *)
(*                                                                     *)
(* data vector:                      (1) CASID                         *)
(*                                   (2) JOBNR                         *)
(*                                   (3) TOTJOBS                       *)
(*                                   (4) CEN                           *)
(*                                   (5) LAST                          *)
(*                                   (6) TS                            *)
(*                                   (7) TF                            *)
(*                                   (8) DUR                           *)
(*                                   (9) TBIRTH                        *)
(*                                  (10) COHORT                        *)
(*                                  (11) SEX                           *)
(*                                  (12) EW                            *)
(*                                  (13) TSCHOOL1                      *)
(*                                  (14) TMAR1                         *)
(*                                  (15) TKID1                         *)
(*                                                                     *)
(* outfiles: XPLOPER.AGE, XPLOPER.DUR, XPLOPER.HST                     *)
(*                                                                     *)
(***********************************************************************)

PROGRAM XPLOPER;

(* ********************************)
(* declarations and initializations *)
(* ********************************)

       CONST
              tmin = 613;
              (* tc = '123456789ABCDEFGHIJK'; *)
              err0 = 'Missing data (-9)';
              err1 = 'Start of job before birth';
              err2 = 'End of job before start of job';
              fnDat = 'XPLOD.DAT';
              fnAge = 'xploper.age';
              fnDur = 'xploper.dur';
              fnHst = 'xploper.hst';
              fnErr = 'xploper.err';

       VAR
              casid, jobnr, totjobs, cen, last, ts, tf, dur, tbirth, cohort: longint;
              sex, ew, tschool1, tmar1, tkid1: longint;
              i, j, iAge, iDur, iHst, iCase: longint;
              vars: ARRAY[1..15] OF longint;
              striew, strisex: CHAR;
              tic: ARRAY[1..500] OF CHAR;
              stricoh: STRING[6];
              fDat, fAge, fDur, fHst, fErr: text;
              isTest: boolean;
        tc: STRING[20];

       PROCEDURE OpenFiles;
       BEGIN
              Assign(fDat, fnDat);
              Reset(fDat);
              Assign(fAge, fnAge);
              Rewrite(fAge);
              Assign(fDur, fnDur);
              Rewrite(fDur);
              Assign(fHst, fnHst);
              Rewrite(fHst);
              Assign(fErr, fnErr);
              Rewrite(fErr);
       END; (* OpenFiles *)

       PROCEDURE WriteErrs (VAR fOut: text; id, mBirth, nJob, mStart, mStop: longint);
```

```
        BEGIN
                IF (mStart < 0) OR (mStop < 0) THEN
                        writeln(fOut, 'ERROR in case ', id, ' job ', nJob : 2, ': ', err0)
                ELSE IF mStart < mBirth THEN
                        writeln(fOut, 'ERROR in case ', id, ' job ', nJob : 2, ': ', err1)
                ELSE IF mStop < mStart THEN
                        writeln(fOut, 'ERROR in case ', id, ' job ', nJob : 2, ': ', err2)
        END; (* WriteErrs *)

        PROCEDURE WriteVars (VAR fOut: text);
        BEGIN
                write(fOut, stricoh, ' ');
                write(fOut, striew, ' ');
                write(fOut, strisex, ' ');
        END; (* WriteVars *)

        PROCEDURE WriteJobs (VAR fOut: text; VAR iMonth: longint; mBirth, nJob, mStart, mStop:
                longint);
        BEGIN
                WHILE iMonth < mStart DO
                BEGIN
                        write(fOut, '_');
                        iMonth := iMonth + 6;
                END;
                WHILE iMonth <= mStop DO
                BEGIN
                        write(fOut, tc[nJob]);
                        iMonth := iMonth + 6;
                END;
        END; (* WriteJobs *)

BEGIN
        isTest := false;
        iCase := 0;
        tc := '123456789ABCDEFGHIJ';

        Writeln('Welcome to XPLOPER...');

        OpenFiles;

        WHILE NOT eof(fDat) DO
        BEGIN
                FOR i := 1 TO 15 DO
                        Read(fDat, vars[i]);
                readln(fDat);

                iCase := iCase + 1;

                casid := vars[1];
                jobnr := vars[2];
                totjobs := vars[3];
                cen := vars[4];
                last := vars[5];
                ts := vars[6];
                tf := vars[7];
                dur := vars[8];
                tbirth := vars[9];
                cohort := vars[10];
                sex := vars[11];
                ew := vars[12];
                tschool1 := vars[13];
                tmar1 := vars[14];
                tkid1 := vars[15];

(* **************************************)
(* build pattern for personal information)
(* **************************************)

                IF (cohort = 1940) THEN
                        stricoh := '##____';
                IF (cohort = 1950) THEN
                        stricoh := '__##__';
                IF (cohort = 1960) THEN
                        stricoh := '____##';
                IF (ew = 1) THEN
                        striew := 'E'
                ELSE
                        striew := 'W';
```

```
        IF (sex = 1) THEN
                strisex := 'M'
        ELSE
                strisex := 'F';

        IF isTest THEN
        BEGIN
                WriteErrs(fErr, casid, tBirth, jobnr, ts, tf);
                write(casid : 7, ' ');
                WriteVars(output);
                write(totjobs : 3);
                write(ts : 5);
                write(tf : 5);
        END
        ELSE
        BEGIN
                Write('.');
                IF iCase MOD 50 = 0 THEN
                BEGIN
                        write(iCase : 5);
                        writeln;
                END;
        END;

        iAge := tBirth + 120;
        iDur := ts;
        iHst := tmin;

        WriteVars(fAge);
        WriteVars(fDur);
        WriteVars(fHst);

        WriteJobs(fAge, iAge, tbirth, jobnr, ts, tf);
        WriteJobs(fDur, iDur, tbirth, jobnr, ts, tf);
        WriteJobs(fHst, iHst, tbirth, jobnr, ts, tf);

        FOR i := 2 TO totjobs DO
        BEGIN
                FOR j := 1 TO 15 DO
                        Read(fDat, vars[j]);
                readln(fDat);
                jobnr := vars[2];
                ts := vars[6];
                tf := vars[7];

                IF isTest THEN
                BEGIN
                        write(ts : 5);
                        write(tf : 5);
                END;

                WriteJobs(fAge, iAge, tbirth, jobnr, ts, tf);
                WriteJobs(fDur, iDur, tbirth, jobnr, ts, tf);
                WriteJobs(fHst, iHst, tbirth, jobnr, ts, tf);
        END;

         cen := vars[4];
        IF cen = 0 THEN
        BEGIN
                write(fAge, 'c');
                write(fDur, 'c');
                write(fHst, 'c');
                IF isTest THEN
                        write(' C');
        END;

        IF isTest THEN
                writeln;

        writeln(fAge);
        writeln(fDur);
        writeln(fHst);
    END;

    IF iCase MOD 50 <> 0 THEN
        writeln(iCase : 5);
    writeln;
END.
```

```
(***********************************************************************)
(* PASCAL programme  XPLRAND.PAS                                      *)
(* -----------------------------                                      *)
(***********************************************************************)

PROGRAM XPLRAND;

(* ********************************)
(* declarations and initializations *)
(* ********************************)

       CONST
              fnAge = 'xploper.age';
              fnDur = 'xploper.dur';
              fnHst = 'xploper.hst';

       VAR
              line: STRING;
               fnRndAge, fnRndDur, fnRndHst: STRING;
              fAge, fDur, fHst: TEXT;
              fRndAge, fRndDur, fRndHst: TEXT;
              nCases, iSel, nSel, Percentage, i: LONGINT;
        selected: ARRAY[1..10000] OF BOOLEAN;
        isTest: BOOLEAN;

       PROCEDURE OpenInputFiles;
       BEGIN
            Assign(fAge, fnAge);
             Reset(fAge);
            Assign(fDur, fnDur);
             Reset(fDur);
             Assign(fHst, fnHst);
             Reset(fHst);
        END; (* OpenInputFiles *)

        PROCEDURE InsertPercentage(p: LONGINT; VAR fn: STRING);
        BEGIN
           fn[7] := CHR(ORD('0') + p DIV 10);
           fn[8] := CHR(ORD('0') + p MOD 10);
        END; (* InsertPercentage *)

        PROCEDURE OpenOutputFiles(p: LONGINT);
        BEGIN
            fnRndAge := 'xplrndxx.age';
             InsertPercentage(p, fnRndAge);
             Assign(fRndAge, fnRndAge);
             Rewrite(fRndAge);
            fnRndDur := 'xplrndxx.dur';
             InsertPercentage(p, fnRndDur);
             Assign(fRndDur, fnRndDur);
             Rewrite(fRndDur);
            fnRndHst := 'xplrndxx.hst';
             InsertPercentage(p, fnRndHst);
             Assign(fRndHst, fnRndHst);
             Rewrite(fRndHst);
       END; (* OpenFiles *)

       PROCEDURE CountCases(VAR nCases: LONGINT);
       BEGIN
          nCases := 0;
           WriteLn('counting cases...');
          WHILE NOT EOF(fDur) DO
           BEGIN
               nCases := nCases + 1;
                IF nCases MOD 10 = 0
                THEN Write('.');
                IF nCases MOD 500 = 0
                THEN WriteLn(nCases:6);
               ReadLn(fDur);
          END;
           Reset(fDur);
           WriteLn;
       END; (* CountCases *)


BEGIN
     isTest := FALSE;
     WriteLn;
```

```
      WriteLn('Welcome to XPLRAND ...');
      WriteLn;

      OpenInputFiles;
      CountCases(nCases);
      WriteLn('Number of Cases          : ', nCases: 5);

      Write  ('Percentage               : ');
      ReadLn(Percentage);
      nSel := (nCases * Percentage) DIV 100;
      WriteLn('Cases to be selected     : ', nSel:5);

      FOR i := 1 TO nCases DO
          selected[i] := FALSE;

      Randomize;
      iSel := 0;
      WHILE iSel < nSel DO
      BEGIN
           i := RANDOM(nCases);
           IF isTest
           THEN Write(i:5);
           IF NOT selected[i]
           THEN
           BEGIN
                iSel := iSel + 1;
                selected[i] := TRUE;
           END;
      END;
      WriteLn('Random numbers generated');

      WriteLn('Selecting cases...');
      OpenOutputFiles(Percentage);
      iSel := 0;
      FOR i := 1 TO nCases DO
      BEGIN
         IF selected[i]
         THEN
         BEGIN
              iSel := iSel + 1;
              IF isTest
              THEN Write(i:5,'s')
              ELSE
              BEGIN
                  IF iSel MOD 10 = 0
                  THEN Write('.');
                  IF iSel MOD 500 = 0
                  THEN WriteLn(iSel:6);
              END;
              ReadLn(fDur, line);
              WriteLn(fRndDur, line);
              ReadLn(fAge, line);
              WriteLn(fRndAge, line);
              ReadLn(fHst, line);
              WriteLn(fRndHst, line);
         END
         ELSE
         BEGIN
              IF isTest
              THEN Write(i:5,'n');
              ReadLn(fDur);
              ReadLn(fAge);
              ReadLn(fHst)
         END;
      END;

      WriteLn;
      WriteLn(isel:1, ' cases selected');
      WriteLn('Files ', fnRndAge, ', ', fnRndDur, ' and ', fnRndHst, ' created');
END.
```

Appendix C: Figures and Tables

Table 1: *Sample Size for Different Birth Cohorts of the German Life History Study (see page 13)*

Figure 1: LIFEPLOT - Different Types of Episodes in Life over Historical Time
3 Individuals of Different Birth Cohorts of the GLHS

Figure 2: LIFEPLOT - Different Events and Episodes over Life Time
3 Individuals of Different Birth Cohorts of the GLHS

Figure 3: LEXBOX – Lexis Diagram Combined with Boxplots
Timing and Sequencing of Events: Starting the First Job and Birth of the First Child
A Comparison of East and West German Women of Three Birth Cohorts

Table 2: Raw Data Listing for Job Episodes of a 0.3% Random Sample of Persons

Table 3: *Number of Persons and Jobs in Subgroups of File XPLOD (see page 21)*

Figure 4: Matrix for Job Duration Pattern of a 1% Random Subsample

Figure 5a: Stacked Stem-Leaf Diagrams for Job Durations
East and West German Men and Women of Birth Cohort 1940

Figure 5b: Stacked Stem-Leaf Diagrams for Job Durations
East and West German Men and Women of Birth Cohorts 1950 and 1960

Figure 6: Grouped Skeleton Plots for Career Patterns over Historical Time
East and West German Women and Men of 3 Different Birth Cohorts –
20 % Random Subsample

Figure 7: Grouped Skeleton Plots for Career Patterns over Age
East and West German Women of 3 Different Birth Cohorts –
20 % Random Subsample

Figure 8: Skeleton Plots with Slicing for Career Patterns since First Job
East and West German Women of Birth Cohort 1950 – 20% Random Subsample

Figure 9: *Plot of the Survivor Function for Job Durations (see page 32)*
Table 4: *Calculating the Survivor Function by the Kaplan- Meier Product-Limit Estimator (see page 33)*
Table 5: *Estimating the Survivor Function by the Product-Limit Formula (see page 33)*
Figure 10a: *Interactive Graphics by E. Neely Atkinson: Scatterplot Matrix Linked with Kaplan-Meier Survival Curve (see page 35)*
Figure 10b: *Interactive Graphics by E. Neely Atkinson: Dichotomized Variable Linked with two Corresponding Survival Curves (see page 36)*

Figure 11: Design of a Survival Data Exploration Window
Examining the Time Span from Entry into Labour Market to Birth of First Child/
Comparing East and West German Women of Birth Cohort 1960

Figure 12: Design of a Survival Data Exploration Window
Examining Hazard Rates for Birth of First Child after Entry into Labour Market/
Comparing East and West German Women of Birth Cohort 1960

Figure 1: LIFEPLOT
Different Types of Episodes in Life over Historical Time
3 Individuals of Different Birth Cohorts of the GLHS

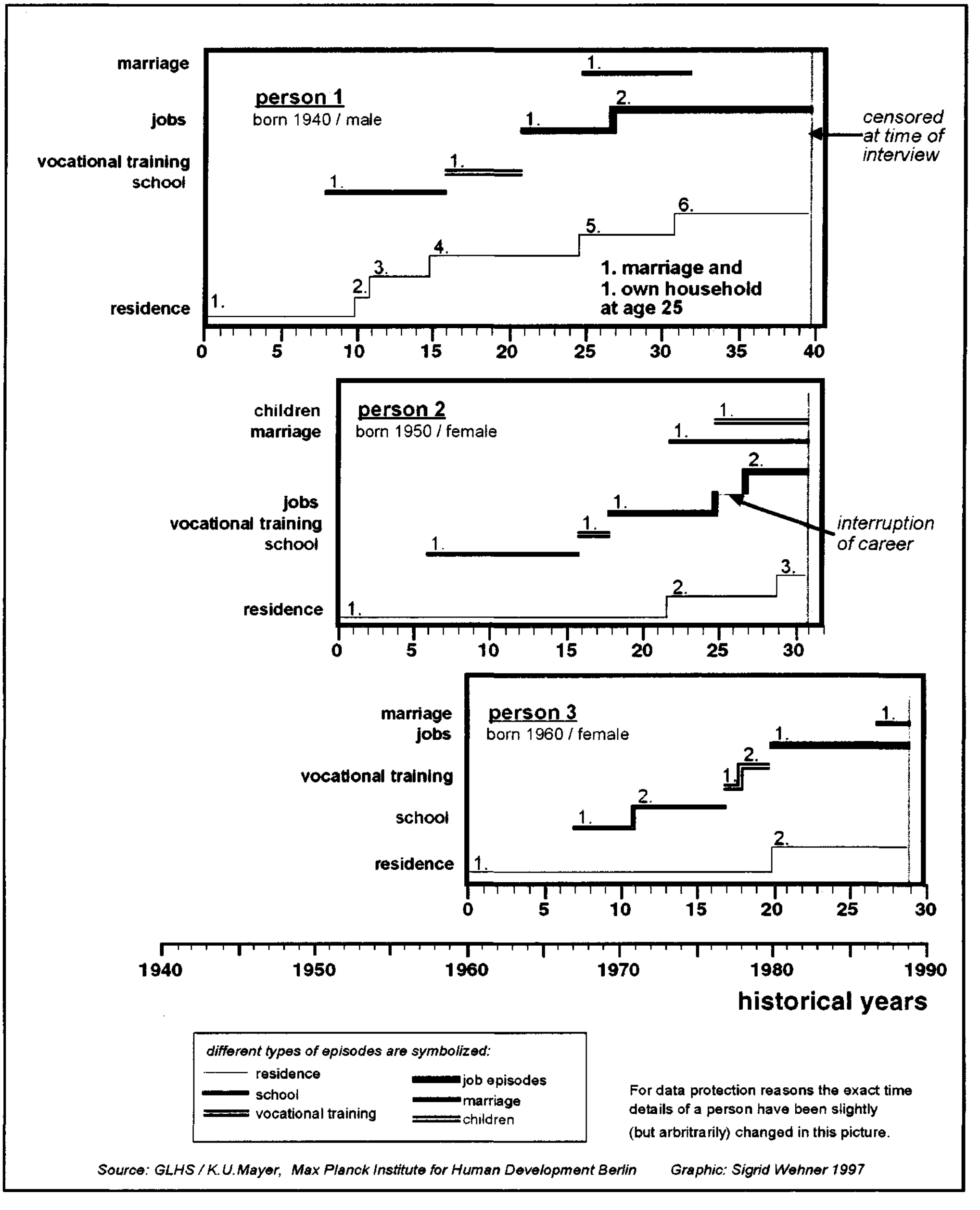


Figure 2: LIFEPLOT
Different Events and Episodes over Life Time
3 Individuals of Different Birth Cohorts of the GLHS

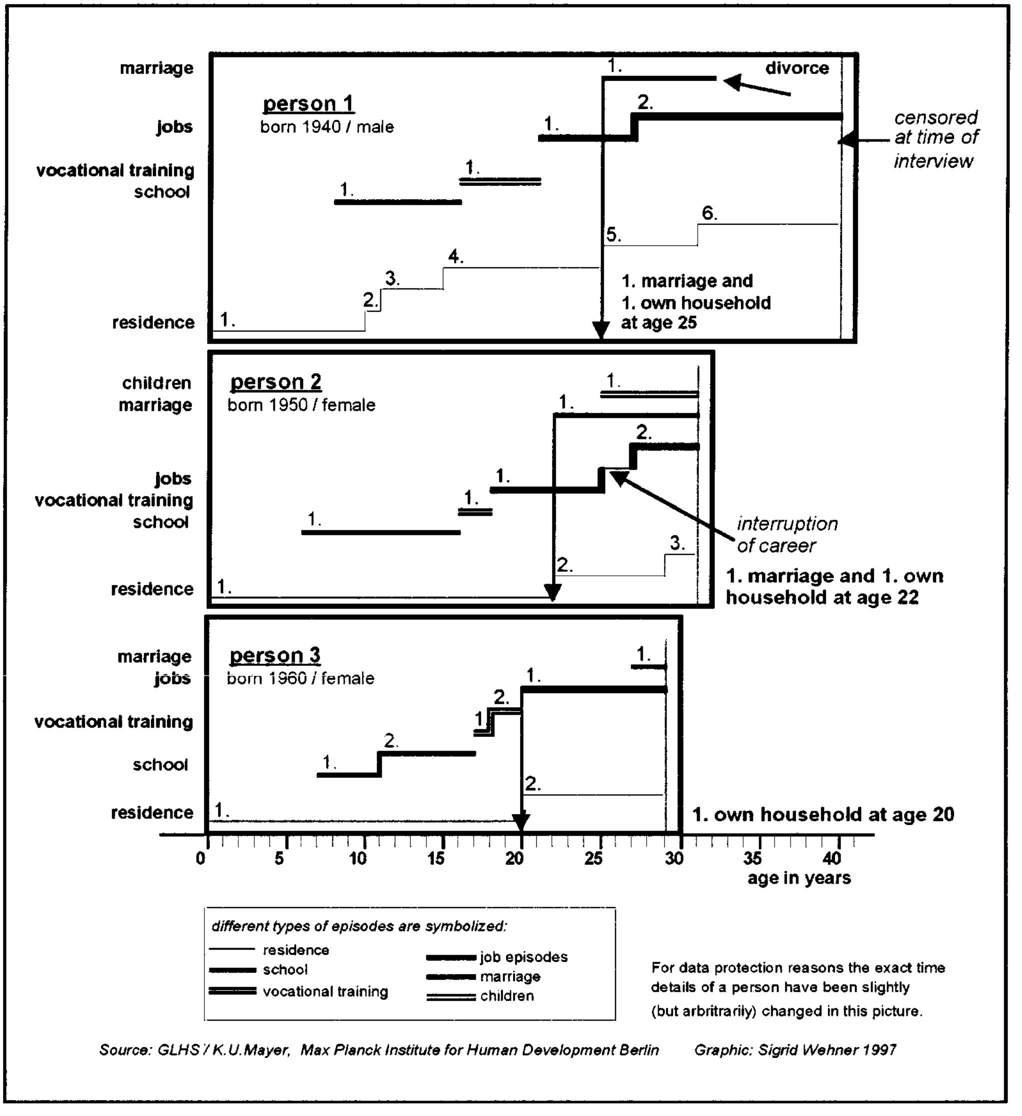
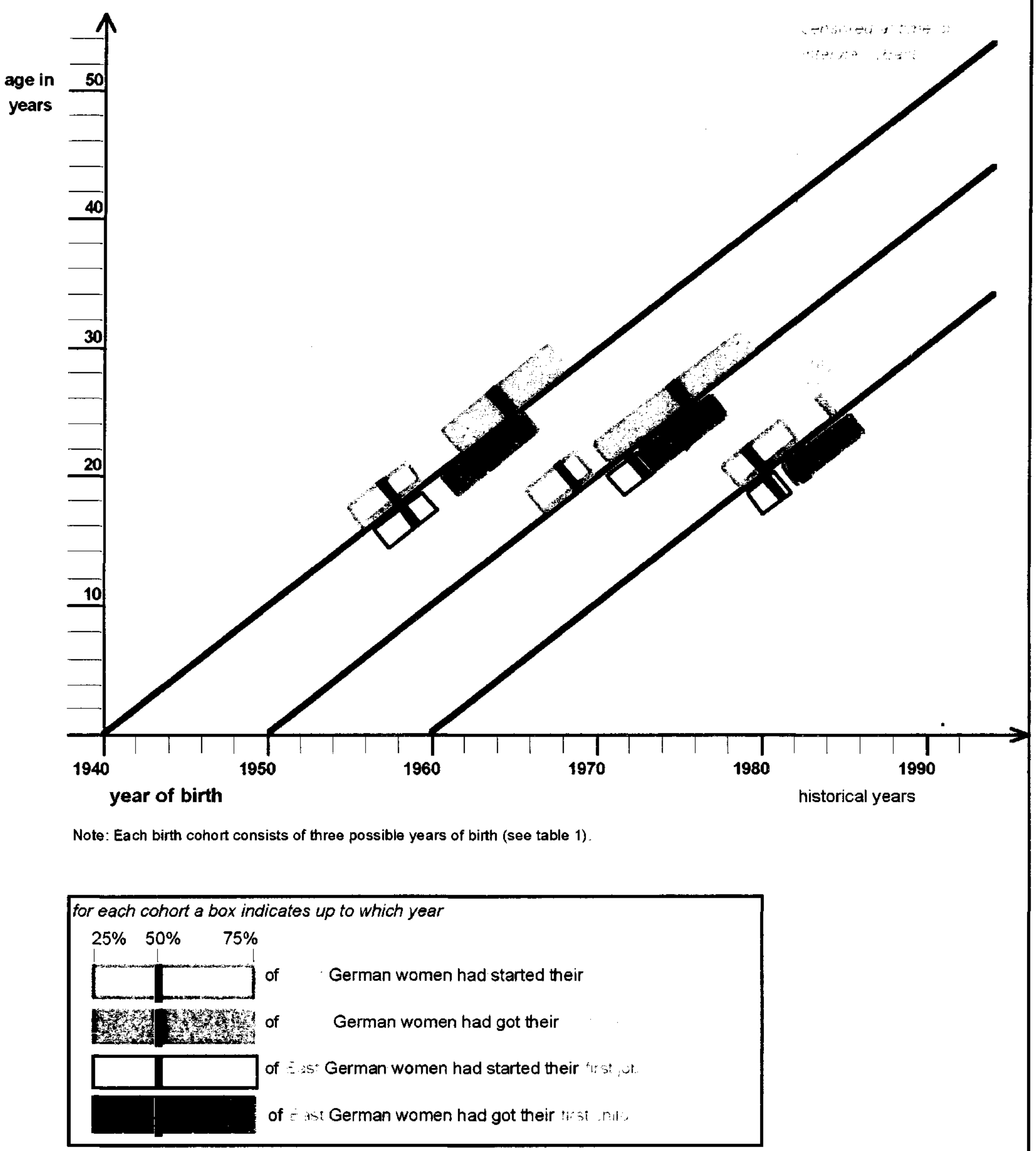


Figure 3: LEXBOX
Lexis Diagram Combined with Boxplots

Timing and Sequencing of Events: Starting the First Job and Birth of the First Child
A Comparison of East and West German Women of Three Birth Cohorts



Source: GLHS / K.U. Mayer, Max Planck Institute for Human Development Berlin *Graphic: Sigrid Wehner 1998*

table 2:

Raw Data Listing for Job Episodes of a 0.3% Random Sample of Persons

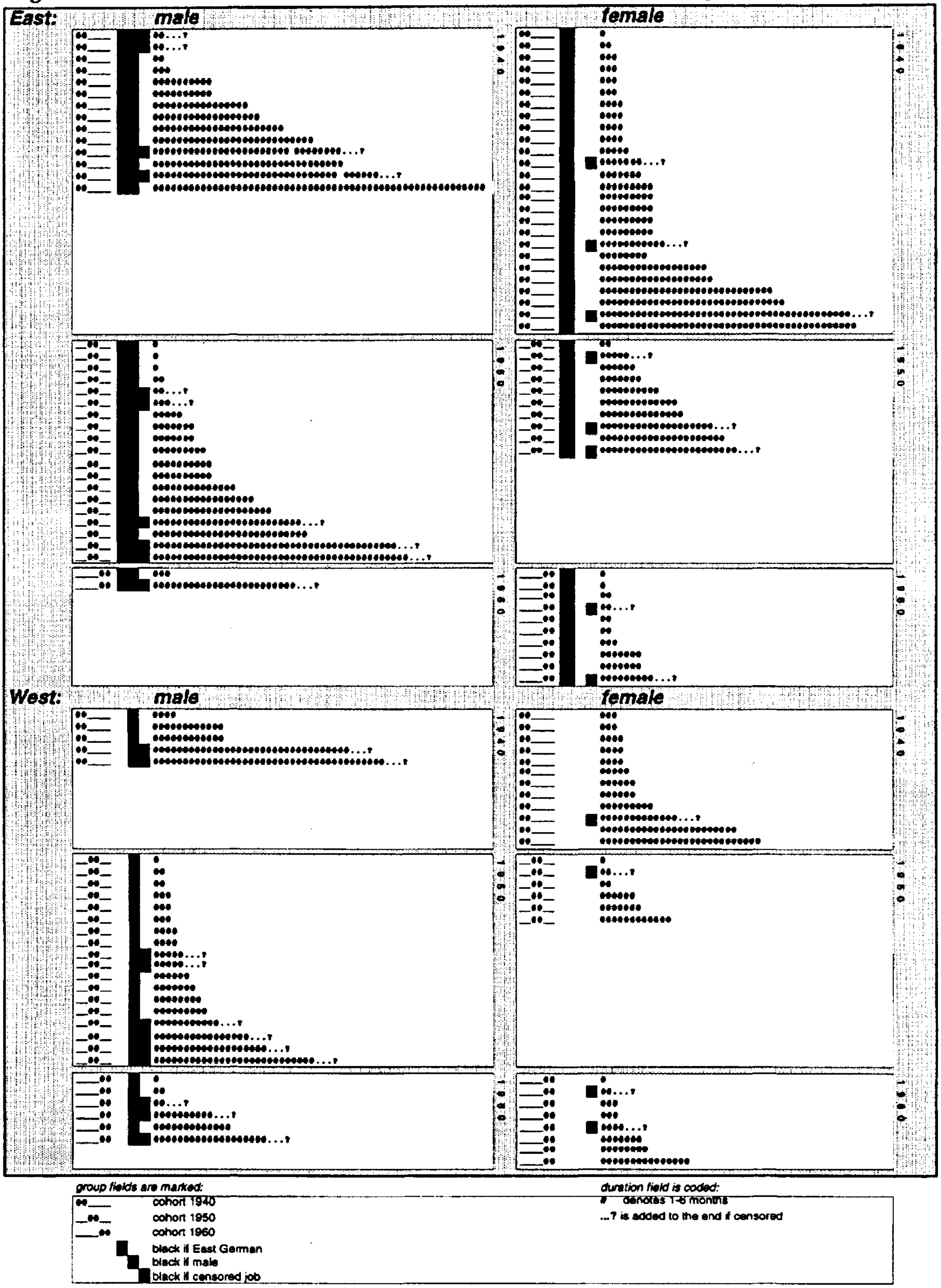
variables varying per job-spell													*variables varying per person (constant for all jobs within the same person)*											
CASID	JOBNR	TOTJOBS	CEN	LAST	TS	TF	DUR	LFX	INCS	INCF	HWEEK	FSIZE	TBIRTH	COHORT	SEX	EW	SCHOOLR	VOCAT1R	VOCAT2R	TSCHOOL1	TVOCAT1	TVOCAT2	TMAR1	TKID1
10492	1	1	1	1	823	869	47	0	600	1000	40		620	1950	2	2	2	5	5	795	820	820	865	871
10950	1	3	1	0	820	843	24	0					591	1950	1	2	2	2	5	762	783	819	868	838
10950	2	3	1	0	844	883	40	24	2500	2500	40	20	591	1950	1	2	2	2	5	762	783	819	868	838
10950	3	3	0	1	884	984	101	64	3500	4500	72	2	591	1950	1	2	2	2	5	762	783	819	868	838
11523	1	1	0	1	838	984	147	0	2200		40	600	604	1950	1	2	2	1	1	795	837	837	862	922
11688	1	4	1	0	688	733	46	0			48	50	482	1940	2	2	2	5	5	651	687	687	743	763
11688	2	4	1	0	734	769	36	46			42	50	482	1940	2	2	2	5	5	651	687	687	743	763
11688	3	4	1	0	770	816	47	82			20	40	482	1940	2	2	2	5	5	651	687	687	743	763
11688	4	4	0	1	817	989	173	129			12	40	482	1940	2	2	2	5	5	651	687	687	743	763
11787	1	1	1	1	809	926	118	0	400	1400	40	2000	610	1950	2	2	1	5	5	783	808	808	862	928
45136	1	3	1	0	942	966	25	0	1000	1400	40	30	729	1960	1	2	2	1	1	907	941	989	953	958
45136	2	3	1	0	990	1026	37	25	2200	2500	50	8000	729	1960	1	2	2	1	1	907	941	989	953	958
45136	3	3	0	1	1027	1070	44	62	2950	2950	50	200	729	1960	1	2	2	1	1	907	941	989	953	958
45171	1	3	1	0	962	972	11	0	1300	1300	40	5	738	1960	1	2		1	1	919	961	961	1046	
45171	2	3	1	0	988	1054	67	11	1500	1650	40	20	738	1960	1	2		1	1	919	961	961	1046	
45171	3	3	0	1	1055	1070	16	78	1800	2000	40	15	738	1960	1	2		1	1	919	961	961	1046	
45187	1	1	1	1	986	1034	49	0	2208	2208	48	500	716	1960	1	2	5			942	993	1070		
45347	1	2	1	0	970	972	3	0	1000	1000	40		734	1960	1	2	5	1		954	1039	1072		
45347	2	2	1	1	973	996	24	3	1200	2000	35	5	734	1960	1	2	5	1		954	1039	1072		
46362	1	2	1	0	953	960	8	0	360	360	9	11	711	1960	2	2	3	4	2	907	915	952	941	942
46362	2	2	0	1	1065	1071	7	8	440	440	10	30	711	1960	2	2	3	4	2	907	915	952	941	942
142330	1	2	1	0	691	798	108	0	310	540	48	300	471	1940	2	1	2	5	5				762	773
142330	2	2	0	1	871	1104	234	108	350	1350		1050	471	1940	2	1	2	5	5				762	773
151529	1	4	1	0	871	874	4	0	800		43	30	644	1950	1	1	5	6	6	870	966	966	1000	993
151529	2	4	1	0	911	944	34	4	800	1300	50		644	1950	1	1	5	6	6	870	966	966	1000	993
151529	3	4	1	0	967	1026	60	38	1400	1700	60		644	1950	1	1	5	6	6	870	966	966	1000	993
151529	4	4	0	1	1027	1110	84	98	1700	1700	40		644	1950	1	1	5	6	6	870	966	966	1000	993
152379	1	4	1	0	909	930	22	0	785	785		40	633	1950	2	1	5	6	6	872	908	908	904	931
152379	2	4	1	0	944	976	33	22	1000	1000		35	633	1950	2	1	5	6	6	872	908	908	904	931
152379	3	4	1	0	990	1100	111	55	1000	1000		30	633	1950	2	1	5	6	6	872	908	908	904	931
152379	4	4	0	1	1101	1110	10	166	2000			25	633	1950	2	1	5	6	6	872	908	908	904	931
161533	1	2	1	0	968	976	9	0	650	1100	40	2000	739	1960	1	1	3	5	5	944	967	967		
161533	2	2	0	1	995	1105	111	9	1100			2000	739	1960	1	1	3	5	5	944	967	967		

Source: German Life History Study / K.U. Mayer, Max Planck Institute for Human Development Berlin

JOBNR current job nr
TOTJOBS total number of person's jobs
CEN right censored?
LAST last job?
TS time starting job
TF time finishing job
DUR duration of job in months
LFX labour force experience
INCS income starting job
INCF income finishing job
HWEEK weekly working hours
FSIZE size of firm

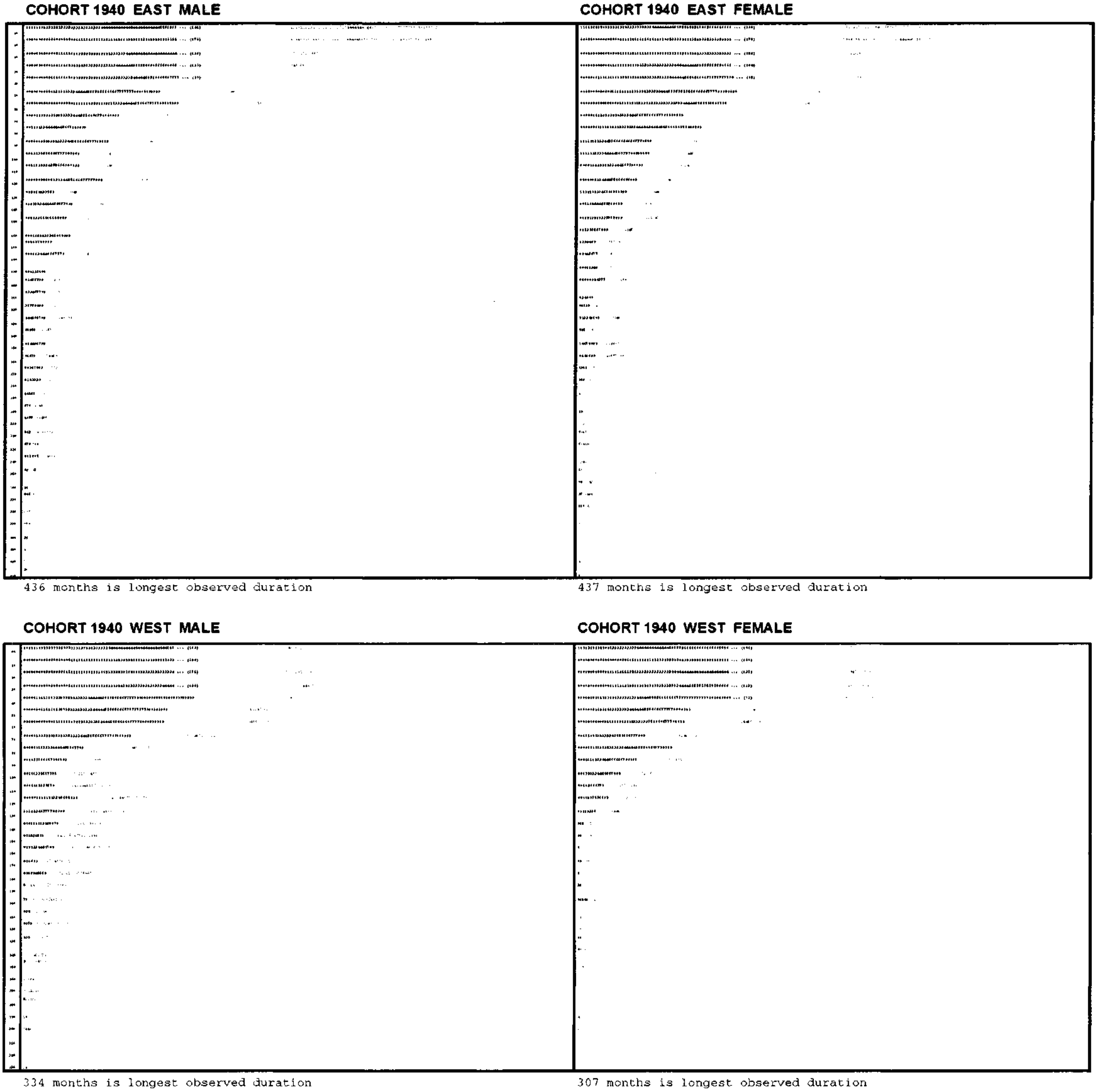
TBIRTH time of birth in month since 1900
COHORT birth cohort
SEX gender
EW east or west?
SCHOOLR educational degree (recoded)
VOCAT1R degree 1. vocational training (recoded)
VOCAT2R degree last vocational training (recoded)
TSCHOOL1 time finishing school
TVOCAT1 time finishing 1. vocational training
TVOCAT2 time finishing last vocational training
TMAR1 time of 1. marriage
TKID1 time of birth of 1. child

Figure 4: Matrix for Job Duration Pattern of a 1% Random Subsample



Source: German Life History Study / K.U. Mayer, Max Planck Institute for Human Development Berlin *Graphic: Sigrid Wehner 1998*

Figure 5a: Stacked Stem-Leaf Diagrams for Job Durations East and West German Men and Women of Birth Cohort 1940



black digits: non-censored durations
[illegible]
... (nnn) denotes shortening of nnn digits

Source: German Life History Study / K.U. Mayer, Max Planck Institute for Human Development Berlin *Graphic: Sigrid Wehner 1998*

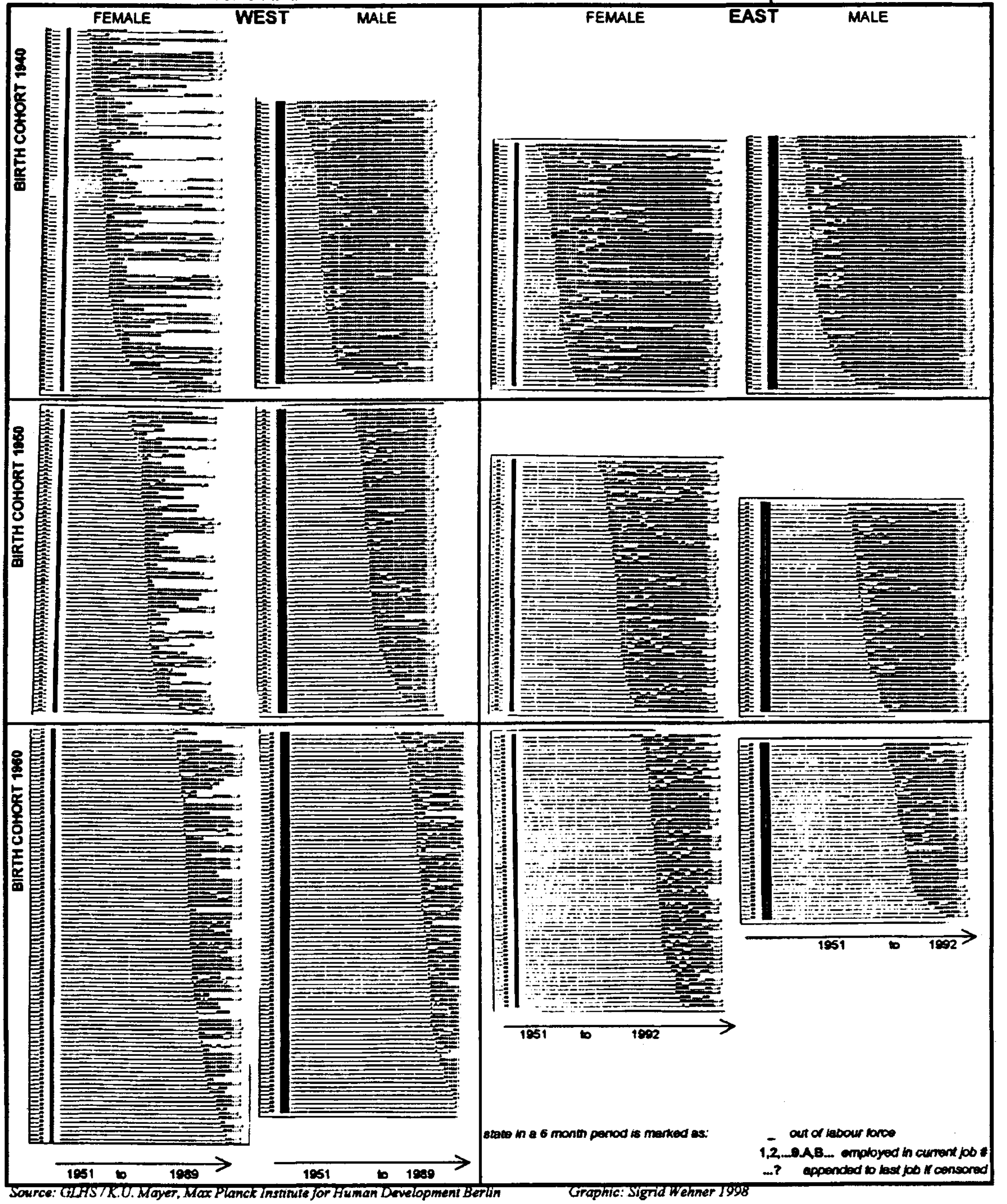
Figure 5b: Stacked Stem-Leaf Diagrams for Job Durations
East and West German Men and Women of Birth Cohorts 1950 and 1960

black digits: non-censored durations
[illegible]
... (nnn) denotes shortening of nnn digits

Source: German Life History Study / K.U. Mayer, Max Planck Institute for Human Development Berlin *Graphic: Sigrid Wehner 1998*

Figure 6: Grouped Skeleton Plots for Career Patterns over Historical Time

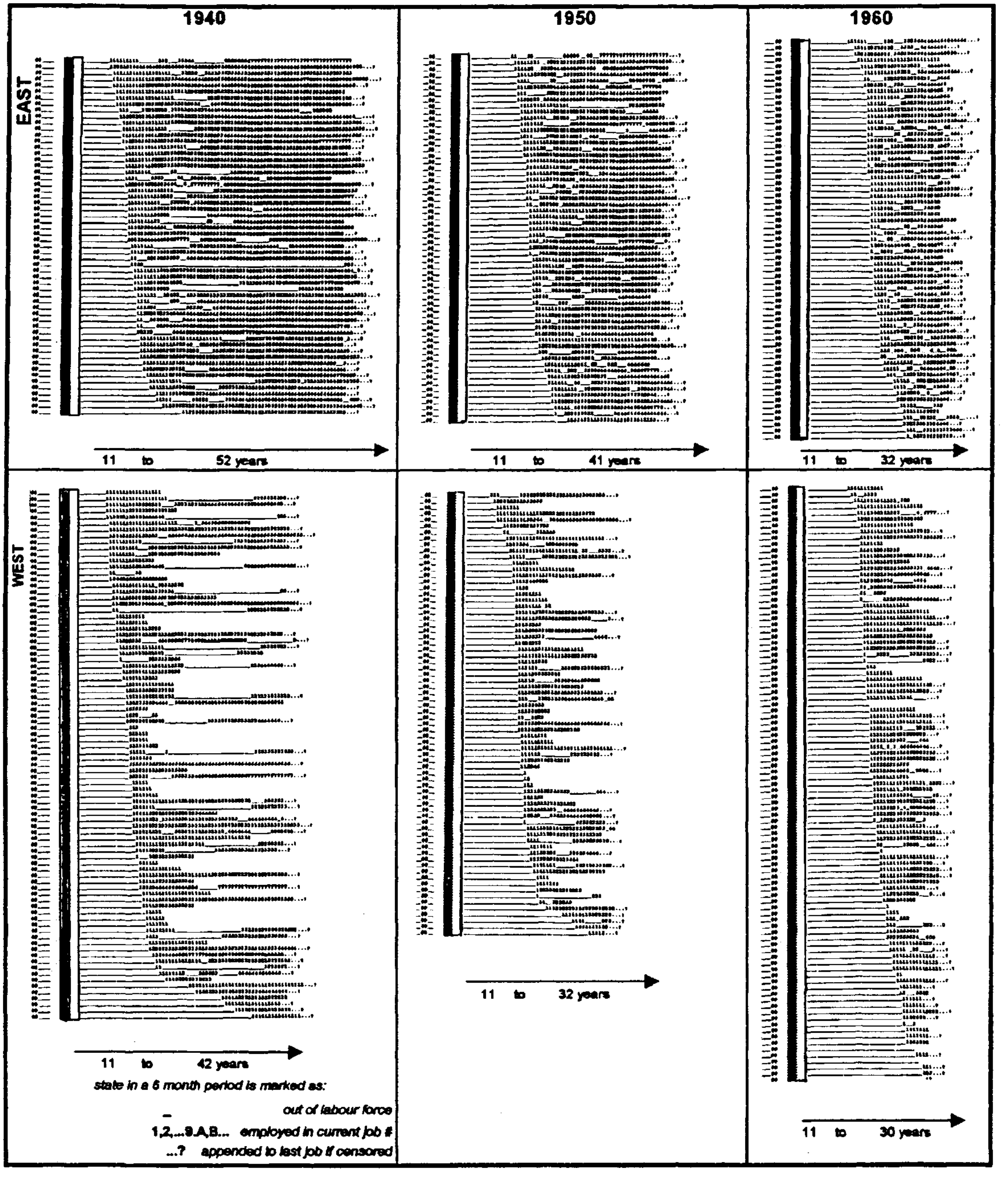
East and West German Women and Men of 3 Different Birth Cohorts - 20 % Random Subsample



Source: GLHS / K.U. Mayer, Max Planck Institute for Human Development Berlin *Graphic: Sigrid Wehner 1998*

Figure 7: Grouped Skeleton Plots for Career Patterns over Age

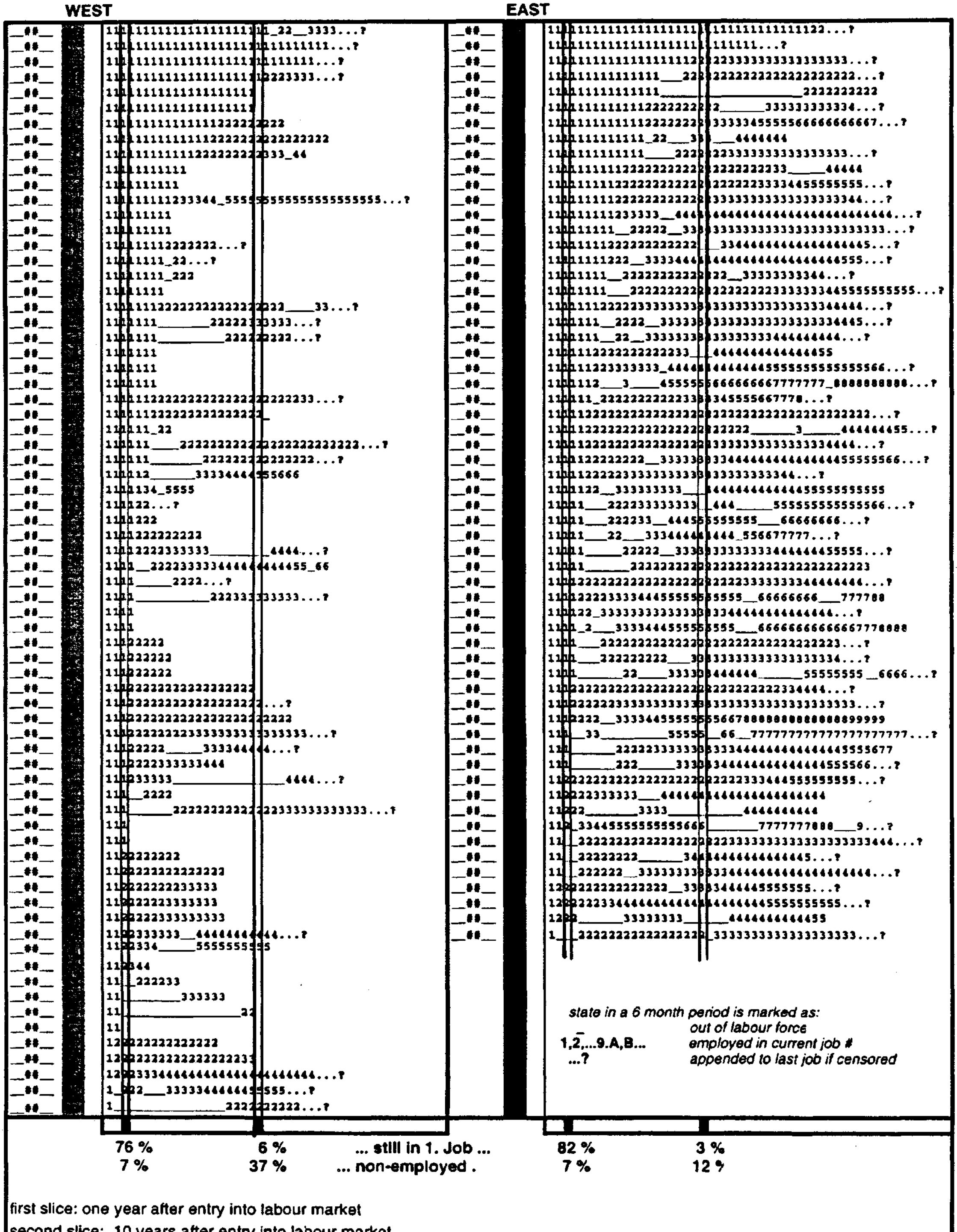
East and West German Women of 3 Different Birth Cohorts - 20 % Random Subsample



Source: GLHS / K.U. Mayer, Max Planck Institute for Human Development Berlin *Graphic: Sigrid Wehner 1998*

Figure 8: Skeleton Plots with Slicing for Career Patterns since First Job

East and West German Women of Birth Cohort 1950 - 20% Random Subsample

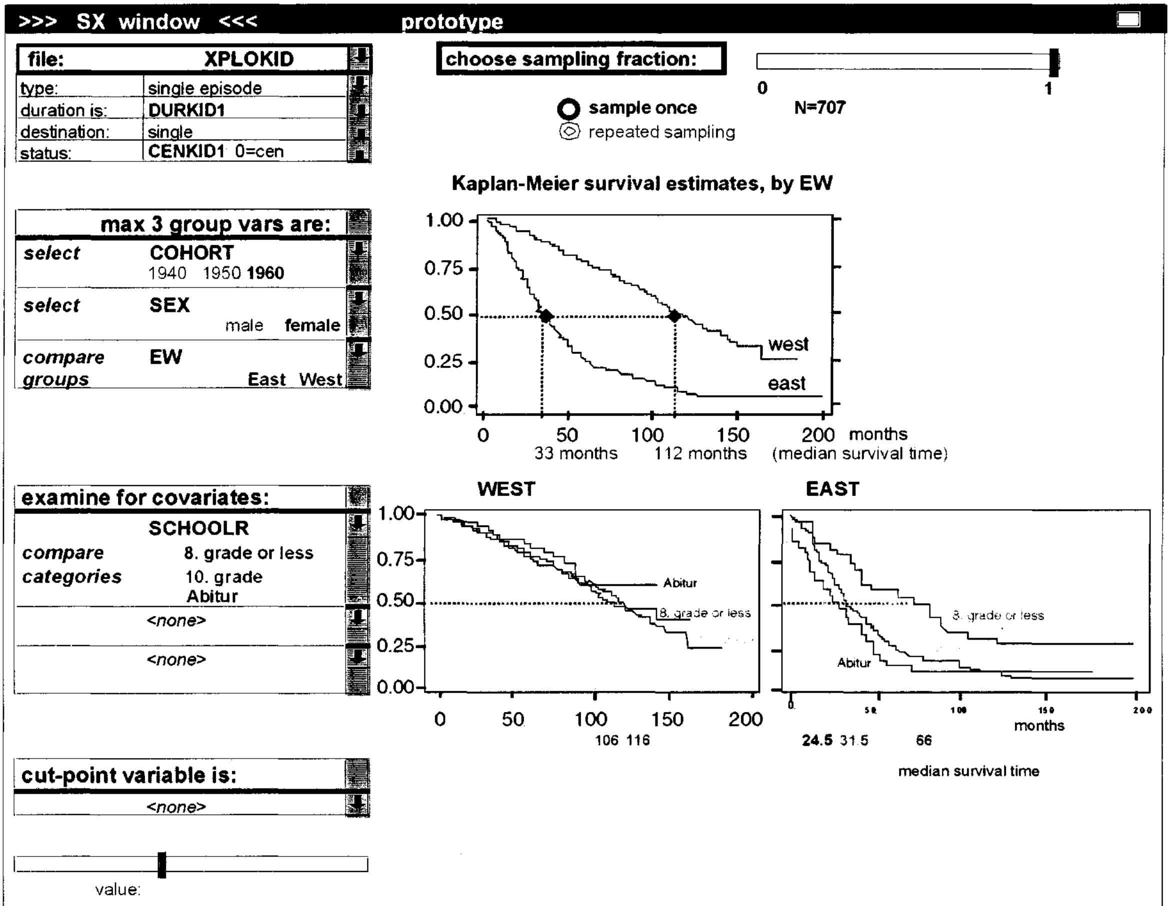


Source: GLHS / K.U. Mayer, Max Planck Institute for Human Development Berlin *Graphic: Sigrid Wehner 1998*

Figure 11

Design of a Survival Data Exploration Window

Examining the Time Span from Entry into Labour Market to Birth of First Child / Comparing East and West German Women of Birth Cohort 1960



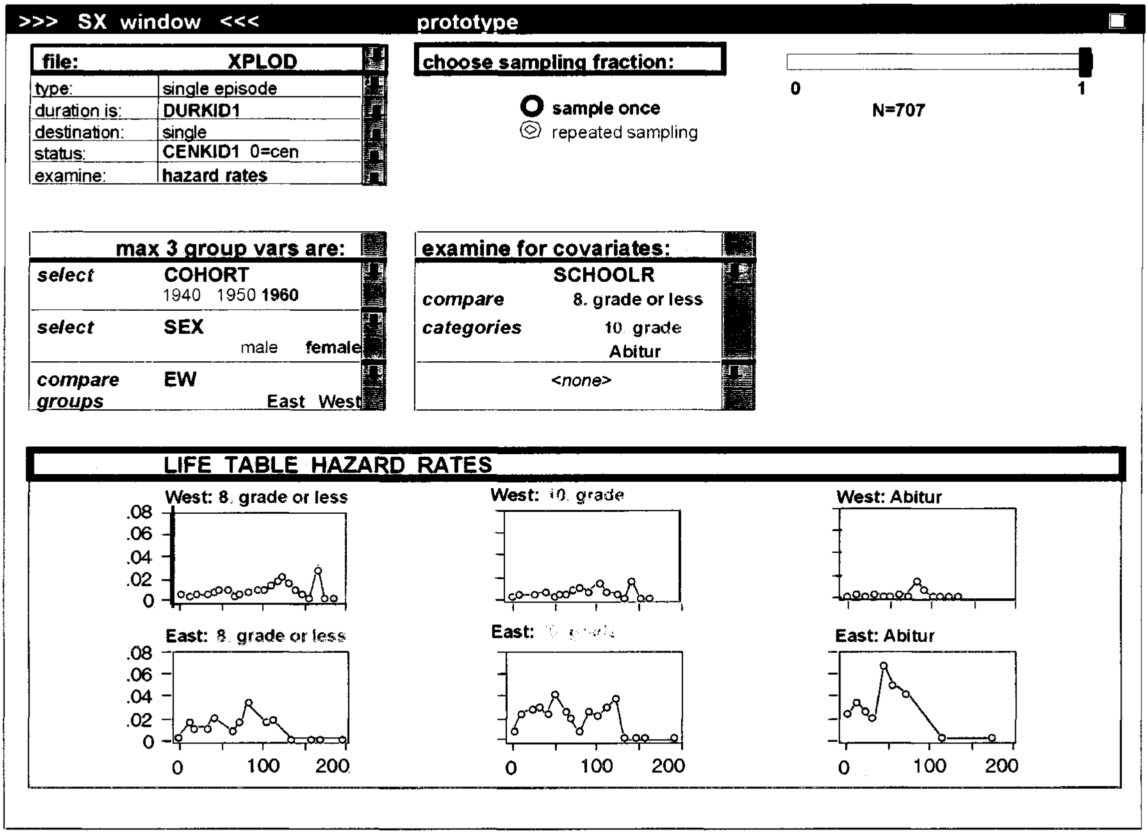
Data Source: *German Life History Study / K.U. Mayer, Max Planck Institute for Human Development Berlin*

Design by: *Sigrid Wehner 1998* *Kaplan-Meier estimates calculated by: STATA*

Figure 12

Design of a Survival Data Exploration Window

Examining Hazard Rates for Birth of First Child after Entry into Labour Market Comparing East and West German Women of Birth Cohort 1960



Data Source: German Life History Study / K.U. Mayer, Max Planck Institute for Human Development Berlin

Design by: Sigrid Wehner 1998 hazard rates calculated by: STATA

I. Reihe STUDIEN UND BERICHTE

Beim Max-Planck-Institut für Bildungsforschung erhältliche Bände (nicht über den Buchhandel beziehbar; Preise zuzüglich Versandpauschale)

66 Ute Kunzmann
Being and Feeling in Control.
Two Sources of Older People's Emotional Well-Being
242 S. Erschienen 1999.
ISBN 3-87985-072-0 DM 29,–

65 Susanne A. Böhmig
Leistungspotentiale wert-relativierenden Denkens.
Die Rolle einer wissensaktivierenden Gedächtnisstrategie.
231 S. Erschienen 1998.
ISBN 3-87985-068-2 DM 27,–

64 Jürgen Baumert, Wilfried Bos und Rainer Watermann
TIMSS/III: Schülerleistungen in Mathematik und den Naturwissenschaften am Ende der Sekundarstufe II im internationalen Vergleich.
Zusammenfassung deskriptiver Ergebnisse.
140 S. Erschienen 1998.
ISBN 3-87985-067-4 DM 10,–

63 Ursula Henz
Intergenerationale Mobilität.
Methodische und empirische Untersuchungen.
354 S. Erschienen 1996.
ISBN 3-87985-059-3 DM 32,–

62 Andreas Maercker
Existentielle Konfrontation.
Eine Untersuchung im Rahmen eines psychologischen Weisheitsparadigmas.
170 S. Erschienen 1995.
ISBN 3-87985-045-3 DM 19,–

61 Alexandra M. Freund
Die Selbstdefinition alter Menschen.
Inhalt, Struktur und Funktion.
251 S. Erschienen 1995.
ISBN 3-87985-057-7 DM 17,–

60 Klaus Schömann
The Dynamics of Labor Earnings over the Life Course.
A Comparative and Longitudinal Analysis of Germany and Poland.
188 S. Erschienen 1994.
ISBN 3-87985-056-9 DM 13,–

59 Frieder R. Lang
Die Gestaltung informeller Hilfebeziehungen im hohen Alter – Die Rolle von Elternschaft und Kinderlosigkeit.
Eine empirische Studie zur sozialen Unterstützung und deren Effekt auf die erlebte soziale Einbindung.
177 S. Erschienen 1994.
ISBN 3-87985-055-0 DM 13,–

58 Ralf Th. Krampe
Maintaining Excellence.
Cognitive-Motor Performance in Pianists Differing in Age and Skill Level.
194 S. Erschienen 1994.
ISBN 3-87985-054-2 DM 14,–

57 Ulrich Mayr
Age-Based Performance Limitations in Figural Transformations.
The Effect of Task Complexity and Practice.
172 S. Erschienen 1993.
ISBN 3-87985-053-4 DM 13,–

56 Marc Szydlik
Arbeitseinkommen und Arbeitsstrukturen.
Eine Analyse für die Bundesrepublik Deutschland und die Deutsche Demokratische Republik.
255 S. Erschienen 1993.
ISBN 3-87985-052-6 DM 21,–

55 Bernd Schellhas
Die Entwicklung der Ängstlichkeit in Kindheit und Jugend.
Befunde einer Längsschnittstudie über die Bedeutung der Ängstlichkeit für die Entwicklung der Kognition und des Schulerfolgs.
205 S. Erschienen 1993.
ISBN 3-87985-051-8 DM 13,–

Max-Planck-Institut für Bildungsforschung
Lentzeallee 94 14195 Berlin (Dahlem)
Tel. 030/824 06-0 Fax 030/824 99 39

I. Reihe STUDIEN UND BERICHTE (Fortsetzung)

54 Falk Fabich
Forschungsfeld Schule: Wissenschaftsfreiheit, Individualisierung und Persönlichkeitsrechte.
Ein Beitrag zur Geschichte sozialwissenschaftlicher Forschung.
235 S. Erschienen 1993.
ISBN 3-87985-050-X DM 22,–

53 Helmut Köhler
Bildungsbeteiligung und Sozialstruktur in der Bundesrepublik.
Zu Stabilität und Wandel der Ungleichheit von Bildungschancen.
133 S. Erschienen 1992.
ISBN 3-87985-049-6 DM 10,–

52 Ulman Lindenberger
Aging, Professional Expertise, and Cognitive Plasticity.
The Sample Case of Imagery-Based Memory Functioning in Expert Graphic Designers.
130 S. Erschienen 1991.
ISBN 3-608-98257-4 DM 11,–

51 Volker Hofmann
Die Entwicklung depressiver Reaktionen in Kindheit und Jugend.
Eine entwicklungspsychopathologische Längsschnittuntersuchung.
197 S. Erschienen 1991.
ISBN 3-608-98256-6 DM 14,–

50 Georgios Papastefanou (vergriffen)
Familiengründung im Lebensverlauf.
Eine empirische Analyse sozialstruktureller Bedingungen der Familiengründung bei den Kohorten 1929–31, 1939–41 und 1949–51.
185 S. Erschienen 1990.
ISBN 3-608-98255-8 DM 15,–

49 Jutta Allmendinger
Career Mobility Dynamics.
A Comparative Analysis of the United States, Norway, and West Germany.
169 S. Erschienen 1989.
ISBN 3-608-98254-X DM 13,–

48 Doris Sowarka
Weisheit im Kontext von Person, Situation und Handlung.
Eine empirische Untersuchung alltagspsychologischer Konzepte alter Menschen.
275 S. Erschienen 1989.
ISBN 3-608-98253-1 DM 20,–

47 Ursula M. Staudinger
The Study of Live Review.
An Approach to the Investigation of Intellectual Development Across the Life Span.
211 S. Erschienen 1989.
ISBN 3-608-98252-3 DM 19,–

46 Detlef Oesterreich
Die Berufswahlentscheidung von jungen Lehrern.
115 S. Erschienen 1987.
ISBN 3-608-98251-5 DM 9,–

45 Hans-Peter Füssel
Elternrecht und Schule.
Ein Beitrag zum Umfang des Elternrechts in der Schule für Lernbehinderte.
501 S. Erschienen 1987.
ISBN 3-608-98249-3 DM 22,–

44 Diether Hopf
Herkunft und Schulbesuch ausländischer Kinder.
Eine Untersuchung am Beispiel griechischer Schüler.
114 S. Erschienen 1987.
ISBN 3-608-98248-5 DM 8,–

43 Eberhard Schröder
Entwicklungssequenzen konkreter Operationen.
Eine empirische Untersuchung individueller Entwicklungsverläufe der Kognition.
112 S. Erschienen 1986.
ISBN 3-608-98247-7 DM 13,–

II. Reihe MATERIALIEN AUS DER BILDUNGSFORSCHUNG

Beim Max-Planck-Institut für Bildungsforschung erhältliche Bände (nicht über den Buchhandel beziehbar; Preise zuzüglich Versandpauschale)

65 Sigrid Wehner
Exploring and Visualizing Event History Data.
76 S. Erschienen 1999
ISBN 3-87985-074-7 DM 10,–

64 Gundel Schümer
Basic Data on the Educational System of Berlin.
62 S. Erschienen 1999.
ISBN 3-87985-073-9 DM 8,–

63 Martin Lages
Algebraic Decomposition of Individual Choice Behavior
201 S. Erschienen 1999.
ISBN 3-87985-070-4 DM 25,–

62 Jürgen Baumert, Wilfried Bos u. a. (Hrsg.)
Testaufgaben zu TIMSS/III
Mathematisch-naturwissenschaftliche Grundbildung und voruniversitäre Mathematik und Physik der Abschlußklassen der Sekundarstufe II. (Population 3).
140 S. Erschienen 1999.
ISBN 3-87985-069-0 DM 15,–

61 Jürgen Baumert, Rainer Lehmann u. a. (Hrsg.)
Testaufgaben Naturwissenschaften TIMSS
7./8. Klasse (Population 2).
111 S. Erschienen 1998.
ISBN 3-87985-066-6 **(vergriffen)** DM 13,–
Internet:www.mpib-berlin.mpg.de: TIMSS/II Testaufgaben

60 Jürgen Baumert, Rainer Lehmann u. a. (Hrsg.)
Testaufgaben Mathematik TIMSS
7./8. Klasse (Population 2).
131 S. Erschienen 1998.
ISBN 3-87985-065-8 **(vergriffen)** DM 15,–
Internet:www.mpib-berlin.mpg.de: TIMSS/II Testaufgaben

59 Todd D. Little and Brigitte Wanner
The Multi-CAM: A Multidimensional Instrument to Assess Children's Action-Control Motives, Beliefs, and Behaviors.
194 S. Erschienen 1997.
ISBN 3-87985-064-X DM 13,–

58 Christine Schmid
Geschwister und die Entwicklung soziomoralischen Verstehens.
Der Einfluß von Altersabstand und Geschlecht jüngerer und älterer Geschwister im Entwicklungsverlauf.
121 S. Erschienen 1997.
ISBN 3-87985-062-3 DM 10,–

57 Kurt Kreppner und Manuela Ullrich
Familien-Codier-System (FCS).
Beschreibung eines Codiersystems zur Beurteilung von Kommunikationsverhalten in Familiendyaden.
94 S. Erschienen 1996.
ISBN 3-87985-061-5 DM 10,–

56 Rosmarie Brendgen
Peer Rejection and Friendship Quality.
A View from Both Friends' Perspectives.
194 S. Erschienen 1996.
ISBN 3-87985-060-7 DM 21,–

55 Siegfried Reuss und Günter Becker
Evaluation des Ansatzes von Lawrence Kohlberg zur Entwicklung und Messung moralischen Urteilens.
Immanente Kritik und Weiterentwicklung.
112 S. Erschienen 1996.
ISBN 3-87985-048-8 DM 13,–

54 Beate Krais und Luitgard Trommer
Akademiker-Beschäftigung.
Sonderauswertung aus der Volkszählung 1987.
324 S. Erschienen 1995.
ISBN 3-87985-047-X DM 33,–

53 Marianne Müller-Brettel
Frieden und Krieg in der psychologischen Forschung.
Historische Entwicklungen, Theorien und Ergebnisse.
296 S. Erschienen 1995.
ISBN 3-87985-046-1 DM 32,–

52 Harald Uhlendorff
Soziale Integration in den Freundeskreis.
Eltern und ihre Kinder.
130 S. Erschienen1995.
ISBN 3-87985-044-5 DM 15,–

Die nicht aufgeführten Bände sind vergriffen, bzw. nur noch in Restexemplaren erhältlich.

II. Reihe MATERIALIEN AUS DER BILDUNGSFORSCHUNG (Fortsetzung)

51 Peter M. Roeder und Bernhard Schmitz
Der vorzeitige Abgang vom Gymnasium.
Teilstudie I: Schulformwechsel vom Gymnasium in den Klassen 5 bis 10.
Teilstudie II: Der Abgang von der Sekundarstufe I.
159 S. Erschienen 1995.
ISBN 3-87985-043-7 **(vergriffen)** DM 18,–

50 Hannah Brückner
Surveys Don't Lie, People Do?
An Analysis of Data Quality in a Retrospective Life Course Study.
86 S. Erschienen 1995.
ISBN 3-87985-042-9 DM 7,–

49 Todd D. Little, Gabriele Oettingen, and Paul B. Baltes
The Revised Control, Agency, and Means-ends Interview (CAMI).
A Multi-Cultural Validity Assessment Using Mean and Covariance Structures (MACS) Analyses.
97 S. Erschienen 1995.
ISBN 3-87985-041-0 DM 8,–

48 Hannah Brückner und Karl Ulrich Mayer
Lebensverläufe und gesellschaftlicher Wandel.
Konzeption, Design und Methodik der Erhebung von Lebensverläufen der Geburtsjahrgänge 1954–1956 und 1959–1961.
Teil I, Teil II, Teil III.
169 S., 224 S., 213 S.
Erschienen 1995.
ISBN 3-87985-039-9 DM 48,–

46 Ursula M. Staudinger, Jacqui Smith und Paul B. Baltes
Handbuch zur Erfassung von weisheitsbezogenem Wissen.
89 S. Deutsche Ausgabe
Manual for the Assessment of Wisdom-Related Knowledge.
83 S. Englische Ausgabe Erschienen 1994.
ISBN 3-87985-037-2 DM 10,–

45 Jochen Fuchs
Internationale Kontakte im schulischen Sektor.
Zur Entwicklung und Situation des Schüleraustausches sowie von Schulpartnerschaften in der BRD.
174 S. Erschienen 1993.
ISBN 3-87985-035-6 DM 19,–

44 Erika Brückner
Lebensverläufe und gesellschaftlicher Wandel.
Konzeption, Design und Methodik der Erhebung von Lebensverläufen der Geburtsjahrgänge 1919–1921.
Teil I, Teil II, Teil III, Teil IV, Teil V.
235 S., 380 S., 200 S., 230 S., 141 S.
Erschienen 1993.
ISBN 3-87985-033-X DM 84,–

43 Ernst-H. Hoff und Hans-Uwe Hohner
Methoden zur Erfassung von Kontrollbewußtsein.
Textteil; Anhang.
99 S. und 178 S. Erschienen 1992.
ISBN 3-87985-032-1 DM 25,–

42 Michael Corsten und Wolfgang Lempert
Moralische Dimensionen der Arbeitssphäre.
Literaturbericht, Fallstudien und Bedingungsanalysen zum betrieblichen und beruflichen Handeln und Lernen.
367 S. Erschienen 1992.
ISBN 3-87985-031-3 DM 20,–

41 Armin Triebel
Zwei Klassen und die Vielfalt des Konsums.
Haushaltsbudgetierung bei abhängig Erwerbstätigen in Deutschland im ersten Drittel des 20. Jahrhunderts. Teil I, Teil II.
416 S., 383 S. Erschienen 1991.
ISBN 3-87985-030-5 DM 48,–

39 Gundel Schümer
Medieneinsatz im Unterricht.
Bericht über Ziel, Anlage und Durchführung einer Umfrage in allgemeinbildenden Schulen.
230 S. Erschienen 1991.
ISBN 3-87985-025-9 DM 24,–

38 Clemens Tesch-Römer
Identitätsprojekte und Identitätstransformationen im mittleren Erwachsenenalter.
312 S. Erschienen 1990.
ISBN 3-87985-026-7 DM 25,–

37 Helmut Köhler
Neue Entwicklungen des relativen Schul- und Hochschulbesuchs.
Eine Analyse der Daten für 1975 bis 1978.
138 S. Erschienen 1990.
ISBN 3-87985-024-0 DM 10,–

II. Reihe MATERIALIEN AUS DER BILDUNGSFORSCHUNG (Fortsetzung)

36 Wilfried Spang und Wolfgang Lempert
Analyse moralischer Argumentationen.
Beschreibung eines Auswertungsverfahrens.
Textteil: Grundlagen, Prozeduren, Evaluation.
Anhang: Interviewleitfaden, Tonbandtranskript und Auswertungsbeispiele.
102 und 191 S. Erschienen 1989. DM 29,–

35 Karl Ulrich Mayer und Erika Brückner
Lebensverläufe und Wohlfahrtsentwicklung.
Konzeption, Design und Methodik der Erhebung von Lebensverläufen der Geburtsjahrgänge 1929–1931, 1939–1941, 1949–1951.
Teil I, Teil II, Teil III.
261 S., unpaginiert, 175 S.
Erschienen 1989. DM 39,–

34 Christoph Droß und Wolfgang Lempert
Untersuchungen zur Sozialisation in der Arbeit 1977 bis 1988.
Ein Literaturbericht.
204 S. Erschienen 1988. DM 12,–

32 Friedrich Edding (Hrsg.)
Bildung durch Wissenschaft in neben- und nachberuflichen Studien.
Tagungsbericht.
157 S. Erschienen 1988. DM 11,–

29 Ulrich Trommer
Aufwendungen für Forschung und Entwicklung in der Bundesrepublik Deutschland 1965 bis 1983.
Theoretische und empirisch-statistische Probleme.
321 S. Erschienen 1987. DM 32,–

III. Einzelpublikationen

Beim Max-Planck-Institut für Bildungsforschung erhältliche Titel (nicht über den Buchhandel beziehbar; Preise zuzüglich Versandpauschale)

Max-Planck-Institut für Bildungsforschung (Hrsg.)
In Memoriam Dietrich Goldschmidt
Reden auf der Akademischen Trauerfeier
am 16. Oktober 1998.
ISBN 3-87985-071-2

Max-Planck-Institut für Bildungsforschung (Hrsg.)
Reden zur Emeritierung von Wolfgang Edelstein
118 S. Berlin: Max-Planck-Institut für Bildungsforschung, 1997.
ISBN 3-87985-063-1

Max-Planck-Institut für Bildungsforschung (Hrsg.)
Pädagogik als empirische Wissenschaft.
Reden zur Emeritierung von Peter Martin Roeder.
90 S. Berlin: Max-Planck-Institut für Bildungsforschung, 1996.
ISBN 3-87985-058-5

Ingo Richter, Peter M. Roeder, Hans-Peter Füssel (Eds.)
Pluralism and Education.
Current World Trends in Policy, Law, and Administration.
345 S. Berkeley: University of California/USA, 1995.
DM 25.–

Max-Planck-Institut für Bildungsforschung (Hrsg.)
Bekenntnis und Dienst.
Reden zum 80. Geburtstag von Dietrich Goldschmidt.
96 S. Berlin: Max-Planck-Institut für Bildungsforschung, 1995.
ISBN 3-87985-040-2

Max-Planck-Institut für Bildungsforschung (Hrsg.)
Abschied von Hellmut Becker.
Reden auf der Trauerfeier am 18. Januar 1994.
47 S. Berlin: Max-Planck-Institut für Bildungsforschung, 1994.
ISBN 3-87985-036-4

Max-Planck-Institut für Bildungsforschung (Hrsg.)
Bildungsforschung und Bildungspolitik.
Reden zum 80. Geburtstag von Hellmut Becker.
98 S. Berlin: Max-Planck-Institut für Bildungsforschung, 1993.
ISBN 3-87985-034-8

Wolfgang Schneider and Wolfgang Edelstein (Eds.)
Inventory of European Longitudinal Studies in the Behavioral and Medical Sciences.
A Project Supported by the European Science Foundation.
557 S. Munich: Max Planck Institute for Psychological Research, and Berlin: Max Planck Institute for Human Development and Education, 1990.
ISBN 3-87985-028-3
DM 58,–

Max-Planck-Institut für Bildungsforschung (Hrsg.)
Entwicklung und Lernen.
Beiträge zum Symposium anläßlich des 60. Geburtstages von Wolfgang Edelstein.
98 S. Berlin: Max-Planck-Institut für Bildungsforschung, 1990.
ISBN 3-87985-023-2

Max-Planck-Institut für Bildungsforschung (Hrsg.)
Normative Voraussetzungen und ethische Implikationen sozialwissenschaftlicher Forschung.
Beiträge zum Symposium anläßlich des 75. Geburtstages von Dietrich Goldschmidt.
108 S. Berlin: Max-Planck-Institut für Bildungsforschung, 1990.
ISBN 3-87985-027-5

Max-Planck-Institut für Bildungsforschung (Hrsg.)
25 Jahre Max-Planck-Institut für Bildungsforschung.
Festvorträge.
48 S. Berlin: Max-Planck-Institut für Bildungsforschung, 1989.

Friedrich Edding
Mein Leben mit der Politik.
126 S. Berlin: Max-Planck-Institut für Bildungsforschung, 1989.

Max-Planck-Institut für Bildungsforschung (Hrsg.)
Gewerbliche Unternehmen als Bildungsträger.
Beiträge zum Symposium anläßlich des 80. Geburtstages von Friedrich Edding.
126 S. Berlin: Max-Planck-Institut für Bildungsforschung, 1989.

IV. Buchveröffentlichungen bei Verlagen (nach dem Erscheinungsjahr geordnet, nur lieferbare Titel; nur über den Buchhandel zu beziehen)

Friedrich Edding und Emile J. Lorang (Hrsg.)
Wege aus der Arbeitskrise.
Bericht über ein Colloquium
der Stiftung PRO VITA SANA.
221 S. Luxembourg: Editions Saint-Paul, 1999.

Paul B. Baltes and Karl Ulrich Mayer (Eds.)
The Berlin Aging Study.
Aging from 70 to 100.
552 pp. Cambridge/UK:
Cambridge University Press, 1999.

Gerd Gigerenzer, Zeno Swijtink, Theodor Porter, Lorraine Daston, John Beatty und Lorenz Krüger
Das Reich des Zufalls.
Wissen zwischen Wahrscheinlichkeiten, Häufigkeiten und Unschärfen.
374 S. Heidelberg, Berlin: Spektrum Akademischer Verlag, 1999.

Christa Händle, Detlef Oesterreich und Luitgard Trommer
Aufgaben politischer Bildung in der Sekundarstufe I.
Studien aus dem Projekt Cicic Education.
281 S. Opladen: Leske + Budrich, 1999.

Christa Händle
Lehrerinnen in System und Lebenswelt.
Erkundungen ihrer doppelten Sozialisation.
336 S. Opladen: Leske + Budrich, 1998.

Michael Corsten
Die Kultivierung beruflicher Handlungsstile.
Einbettung, Nutzung und Gestaltung von Berufskompetenzen.
209 S. Frankfurt a.M. / New York: Campus, 1998.

Hermann Avenarius, Jürgen Baumert, Hans Döbert und Hans-Peter Füssel (Hrsg.)
Schule in erweiterter Verantwortung.
Positionsbestimmungen aus erziehungswissenschaftlicher, bildungspolitischer und verfassungsrechtlicher Sicht.
166 S. Neuwied: Luchterhand, 1998.

Matthias Grundmann
Norm und Konstruktion.
Zur Dialektik von Bildungsvererbung und Bildungsaneignung.
231 S. Opladen: Leske + Budrich, 1998.

Tobias Krettenauer
Gerechtigkeit als Solidarität.
Entwicklungsbedingungen sozialen Engagements im Jugendalter.
267 S. Weinheim: Deutscher Studien Verlag 1998.

Michael Wagner und Yvonne Schütze
Verwandtschaft.
Sozialwissenschaftliche Beiträge zu einem vernachlässigten Thema.
281 S. Stuttgart: Enke, 1998.

Kai Schnabel
Prüfungsangst und Lernen.
Empirische Analysen zum Einfluß fachspezifischer Leistungsängstlichkeit auf schulischen Lernfortschritt.
201 S. New York, München, Berlin: Waxmann, 1998.

Olaf Köller
Zielorientierungen und schulisches Lernen.
216 S. New York, München, Berlin: Waxmann, 1998.

Michael Wagner
Scheidung in Ost- und Westdeutschland.
Zum Verhältnis von Ehestabilität und Sozialstruktur seit den 30er Jahren.
355 S. Frankfurt a.M. / New York: Campus, 1997.

Gero Lenhardt und Manfred Stock
Bildung, Bürger, Arbeitskraft.
Schulentwicklung und Sozialstruktur in der BRD und der DDR
253 S. Frankfurt a.M. Suhrkamp Taschenbuch Verlag, 1997.

Michael Corsten und Wolfgang Lempert
Beruf und Moral.
Exemplarische Analysen beruflicher Werdegänge, betrieblicher Kontexte und sozialer Orientierungen erwerbstätiger Lehrabsolventen.
200 S. Weinheim: Beltz-Deutscher Studien Verlag, 1997.

Jürgen Baumert und Rainer Lehmann u.a.
TIMSS - Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich.
Deskriptive Befunde.
242 S. Opladen: Leske + Budrich, 1997.

IV. Buchveröffentlichungen bei Verlagen (Fortsetzung)

Gabriele Oettingen
Psychologie des Zukunftsdenkens.
Erwartungen und Phantasien.
452 S. Göttingen/Bern/Toronto/Seattle: Hogrefe, 1996.

Detlef Oesterreich
Flucht in die Sicherheit.
Zur Theorie des Autoritarismus und der autoritären Reaktion.
250 S. Opladen: Leske + Budrich, 1996.

Karl Ulrich Mayer und Paul B. Baltes (Hrsg.)
Die Berliner Altersstudie.
(Ein Projekt der Berlin-Brandenburgischen Akademie der Wissenschaften)
672 S. Berlin: Akademie Verlag, 1996.

Paul B. Baltes and Ursula M. Staudinger (Eds.)
Interactive Minds.
Life-Span Perspectives on the Social Foundation of Cognition.
457 pp. New York: Cambridge University Press, 1996.

Monika Keller
Moralische Sensibilität: Entwicklung in Freundschaft und Familie.
259 S. Weinheim: Psychologie Verlags Union, 1996.

Martin Diewald, Karl Ulrich Mayer (Hrsg.)
Zwischenbilanz der Wiedervereinigung.
Strukturwandel und Mobilität im Transformationsprozeß.
352 S. Opladen: Leske + Budrich, 1996.

Johannes Huinink, Karl Ulrich Mayer u.a.
Kollektiv und Eigensinn.
Lebensverläufe in der DDR und danach.
414 S. Berlin: Akademie Verlag, 1995.

Johannes Huinink
Warum noch Familie?
Zur Attraktivität von Partnerschaft und Elternschaft in unserer Gesellschaft.
385 S. Frankfurt a.M./New York: Campus, 1995.

Heike Trappe
Emanzipation oder Zwang?
Frauen in der DDR zwischen Beruf, Familie und Sozialpolitik.
242 S. Berlin: Akademie Verlag, 1995.

Heike Solga
Auf dem Weg in eine klassenlose Gesellschaft?
Klassenlagen und Mobilität zwischen Generationen in der DDR.
265 S. Berlin: Akademie Verlag, 1995.

Lothar Krappmann und Hans Oswald
Alltag der Schulkinder.
Beobachtungen und Analysen von Interaktionen und Sozialbeziehungen.
224 S. Weinheim/München: Juventa, 1995.

Freya Dittmann-Kohli
Das persönliche Sinnsystem.
Ein Vergleich zwischen frühem und spätem Erwachsenenalter.
402 S. Göttingen/Bern/Toronto/Seattle: Hogrefe, 1995.

Hartmut Zeiher und Helga Zeiher
Orte und Zeiten der Kinder.
Soziales Leben im Alltag von Großstadtkindern.
223 S. Weinheim/München: Juventa, 1994.

Christiane Lange-Küttner
Gestalt und Konstruktion.
Die Entwicklung der grafischen Kompetenz beim Kind.
242 S. Bern/Toronto: Huber, 1994.

Jutta Allmendinger
Lebensverlauf und Sozialpolitik.
Die Ungleichheit von Mann und Frau und ihr öffentlicher Ertrag.
302 S. Frankfurt a.M./New York: Campus, 1994.

Wolfgang Lauterbach
Berufsverläufe von Frauen.
Erwerbstätigkeit, Unterbrechung und Wiedereintritt.
289 S. Frankfurt a.M./New York: Campus, 1994.

Arbeitsgruppe Bildungsbericht am Max-Planck-Institut für Bildungsforschung
Das Bildungswesen in der Bundesrepublik Deutschland.
Strukturen und Entwicklungen im Überblick.
843 S. Reinbek: Rowohlt, 1994 (4., vollständig überarbeitete und erweiterte Neuausgabe).

IV. Buchveröffentlichungen bei Verlagen (Fortsetzung)

Hellmut Becker und Gerhard Kluchert
Die Bildung der Nation.
Schule, Gesellschaft und Politik vom Kaiserreich zur Weimarer Republik.
538 S. Stuttgart: Klett-Cotta, 1993.

Rolf Becker
Staatsexpansion und Karrierechancen.
Berufsverläufe im öffentlichen Dienst und in der Privatwirtschaft.
303 S. Frankfurt a.M./New York: Campus, 1993.

Wolfgang Edelstein und
Siegfried Hoppe-Graff (Hrsg.)
Die Konstruktion kognitiver Strukturen.
Perspektiven einer konstruktivistischen Entwicklungspsychologie.
328 S. Bern/Stuttgart/Toronto: Huber, 1993.

Wolfgang Edelstein, Gertrud Nunner-Winkler und Gil Noam (Hrsg.)
Moral und Person.
418 S. Frankfurt a.M.: Suhrkamp, 1993.

Lothar Lappe
Berufsperspektiven junger Facharbeiter.
Eine qualitative Längsschnittanalyse zum Kernbereich westdeutscher Industriearbeit.
394 S. Frankfurt a.M./New York: Campus, 1993.

Detlef Oesterreich
Autoritäre Persönlichkeit und Gesellschaftsordnung.
Der Stellenwert psychischer Faktoren für politische Einstellungen – eine empirische Untersuchung von Jugendlichen in Ost und West.
243 S. Weinheim/München: Juventa, 1993.

Marianne Müller-Brettel
Bibliographie Friedensforschung und Friedenspolitik:
Der Beitrag der Psychologie 1900–1991.
(Deutsch/Englisch)
383 S. München/London/New York/Paris: Saur, 1993.

Paul B. Baltes und Jürgen Mittelstraß (Hrsg.)
Zukunft des Alterns und gesellschaftliche Entwicklung.
(= Forschungsberichte der Akademie der Wissenschaften zu Berlin, 5.)
814 S. Berlin/New York: De Gruyter, 1992.

Matthias Grundmann
Familienstruktur und Lebensverlauf.
Historische und gesellschaftliche Bedingungen individueller Entwicklung.
226 S. Frankfurt a.M./New York: Campus, 1992.

Karl Ulrich Mayer (Hrsg.)
Generationsdynamik in der Forschung.
245 S. Frankfurt a.M./New York: Campus, 1992.

Erika M. Hoerning
Zwischen den Fronten.
Berliner Grenzgänger und Grenzhändler 1948–1961.
266 S. Köln/Weimar/Wien: Böhlau, 1992.

Ernst-H. Hoff
Arbeit, Freizeit und Persönlichkeit.
Wissenschaftliche und alltägliche Vorstellungsmuster.
238 S. Heidelberg: Asanger Verlag, 1992 (2. überarbeitete und aktualisierte Auflage).

Erika M. Hoerning
Biographieforschung und Erwachsenenbildung.
223 S. Bad Heilbrunn: Klinkhardt, 1991.

Max-Planck-Institut für Bildungsforschung
Traditions et transformations.
Le système d'éducation en République fédérale d'Allemagne.
341 S. Paris: Economica, 1991.

Dietrich Goldschmidt
Die gesellschaftliche Herausforderung der Universität.
Historische Analysen, internationale Vergleiche, globale Perspektiven.
297 S. Weinheim: Deutscher Studien Verlag, 1991.

Uwe Henning und Achim Leschinsky (Hrsg.)
Enttäuschung und Widerspruch.
Die konservative Position Eduard Sprangers im Nationalsozialismus. Analysen – Texte – Dokumente.
213 S. Weinheim: Deutscher Studien Verlag, 1991.

Ernst-H. Hoff, Wolfgang Lempert und Lothar Lappe
Persönlichkeitsentwicklung in Facharbeiterbiographien.
282 S. Bern/Stuttgart/Toronto: Huber, 1991.

IV. Buchveröffentlichungen bei Verlagen (Fortsetzung)

Karl Ulrich Mayer, Jutta Allmendinger und Johannes Huinink (Hrsg.)
Vom Regen in die Traufe: Frauen zwischen Beruf und Familie.
483 S. Frankfurt a.M./New York: Campus, 1991.

Maria von Salisch
Kinderfreundschaften.
Emotionale Kommunikation im Konflikt.
153 S. Göttingen/Toronto/Zürich: Hogrefe, 1991.

Paul B. Baltes and Margret M. Baltes (Eds.)
Successful Aging: Perspectives from the Behavioral Sciences.
397 pp. Cambridge: Cambridge University Press, 1990.

Paul B. Baltes, David L. Featherman and Richard M. Lerner (Eds.)
Life-Span Development and Behavior.
368 pp. Vol. 10. Hillsdale, N.J.: Erlbaum, 1990.

Achim Leschinsky and Karl Ulrich Mayer (Eds.)
The Comprehensive School Experiment Revisited: Evidence from Western Europe.
211 pp. Frankfurt a.M./Bern/New York/Paris: Lang 1990.

Karl Ulrich Mayer (Hrsg.)
Lebensverläufe und sozialer Wandel.
467 S. Opladen: Westdeutscher Verlag, 1990.
(= Sonderheft 31 der KZfSS).

Karl Ulrich Mayer and Nancy Brandon Tuma (Eds.)
Event History Analysis in Life Course Research.
320 pp. Madison, Wis.: The University of Wisconsin Press, 1990.

Hans J. Nissen, Peter Damerow und Robert K. Englund
Frühe Schrift und Techniken der Wirtschaftsverwaltung im alten Vorderen Orient.
Informationsspeicherung und -verarbeitung vor 5000 Jahren.
Katalog zur gleichnamigen Ausstellung Berlin-Charlottenburg, Mai–Juli 1990.
222 S. Bad Salzdetfurth: Franzbecker, 1990.
(2. Aufl. 1991).

Peter Alheit und Erika M. Hoerning (Hrsg.)
Biographisches Wissen.
Beiträge zu einer Theorie lebensgeschichtlicher Erfahrung.
284 S. Frankfurt a.M./New York: Campus, 1989.

Arbeitsgruppe am Max-Planck-Institut für Bildungsforschung
Das Bildungswesen in der Bundesrepublik Deutschland.
Ein Überblick für Eltern, Lehrer und Schüler.
Japanische Ausgabe: 348 S. Tokyo: Toshindo Publishing Co. Ltd., 1989.

Hans-Peter Blossfeld
Kohortendifferenzierung und Karriereprozeß.
Eine Längsschnittstudie über die Veränderung der Bildungs- und Berufschancen im Lebenslauf.
185 S. Frankfurt a.M./New York: Campus, 1989.

Hans-Peter Blossfeld, Alfred Hamerle and Karl Ulrich Mayer
Event History Analysis.
Statistical Theory and Application in the Social Sciences.
297 pp. Hillsdale, N.J.: Erlbaum, 1989.

Erika M. Hoerning und Hans Tietgens (Hrsg.)
Erwachsenenbildung: Interaktion mit der Wirklichkeit.
200 S. Bad Heilbrunn: Klinkhardt, 1989.

Johannes Huinink
Mehrebenensystem-Modelle in den Sozialwissenschaften.
292 S. Wiesbaden: Deutscher Universitäts-Verlag, 1989.

Kurt Kreppner and Richard M. Lerner (Eds.)
Family Systems and Life-Span Development.
416 pp. Hillsdale, N.J.: Erlbaum, 1989.

Bernhard Schmitz
Einführung in die Zeitreihenanalyse.
Modelle, Softwarebeschreibung, Anwendungen.
235 S. Bern/Stuttgart/Toronto: Huber, 1989.

IV. Buchveröffentlichungen bei Verlagen (Fortsetzung)

Eberhard Schröder
Vom konkreten zum formalen Denken.
Individuelle Entwicklungsverläufe von der Kindheit zum Jugendalter.
328 S. Bern/Stuttgart/Toronto: Huber, 1989.

Michael Wagner
Räumliche Mobilität im Lebensverlauf.
Eine empirische Untersuchung sozialer Bedingungen der Migration.
226 S. Stuttgart: Enke, 1989.

Paul B. Baltes, David L. Featherman and Richard M. Lerner (Eds.)
Life-Span Development and Behavior.
338 pp. Vol. 9. Hillsdale, N.J.: Erlbaum, 1988.

Paul B. Baltes, David L. Featherman and Richard M. Lerner (Eds.)
Life-Span Development and Behavior.
337 pp. Vol. 8. Hillsdale, N.J.: Erlbaum, 1988.

Lothar Krappmann
Soziologische Dimensionen der Identität.
Strukturelle Bedingungen für die Teilnahme an Interaktionsprozessen.
231 S. Stuttgart: Klett-Cotta, 7. Aufl., 1988
(= Standardwerke der Psychologie).

Detlef Oesterreich
Lehrerkooperation und Lehrersozialisation.
159 S. Weinheim: Deutscher Studien Verlag, 1988.

Michael Bochow und Hans Joas
Wissenschaft und Karriere.
Der berufliche Verbleib des akademischen Mittelbaus.
172 und 37 S. Frankfurt a.M./New York: Campus, 1987.

Hans-Uwe Hohner
Kontrollbewußtsein und berufliches Handeln.
Motivationale und identitätsbezogene Funktionen subjektiver Kontrollkonzepte.
201 S. Bern/Stuttgart/Toronto: Huber, 1987.

Bernhard Schmitz
Zeitreihenanalyse in der Psychologie.
Verfahren zur Veränderungsmesung und Prozeßdiagnostik.
304 S. Weinheim/Basel: Deutscher Studien Verlag/Beltz, 1987.

Margret M. Baltes and Paul B. Baltes (Eds.)
The Psychology of Control and Aging.
415 pp. Hillsdale, N.J.: Erlbaum, 1986.

Paul B. Baltes, David L. Featherman and Richard M. Lerner (Eds.)
Life-Span Development and Behavior.
334 pp. Vol. 7. Hillsdale, N.J.: Erlbaum, 1986.

Hans-Peter Blossfeld, Alfred Hamerle und Karl Ulrich Mayer
Ereignisanalyse.
Statistische Theorie und Anwendung in den Wirtschafts- und Sozialwissenschaften.
290 S. Frankfurt a.M./New York: Campus, 1986.

Axel Funke, Dirk Hartung, Beate Krais und Reinhard Nuthmann
Karrieren außer der Reihe.
Bildungswege und Berufserfolge von Stipendiaten der gewerkschaftlichen Studienförderung.
256 S. Köln: Bund, 1986.

Ernst-H. Hoff, Lothar Lappe und Wolfgang Lempert (Hrsg.)
Arbeitsbiographie und Persönlichkeitsentwicklung.
288 S. Bern/Stuttgart/Toronto: Huber, 1986.

Klaus Hüfner, Jens Naumann, Helmut Köhler und Gottfried Pfeffer
Hochkonjunktur und Flaute: Bildungspolitik in der Bundesrepublik Deutschland 1967–1980.
361 S. Stuttgart: Klett-Cotta, 1986.

Jürgen Staupe
Parlamentsvorbehalt und Delegationsbefugnis.
Zur „Wesentlichkeitstheorie“ und zur Reichweite legislativer Regelungskompetenz, insbesondere im Schulrecht.
419 S. Berlin: Duncker & Humblot, 1986.

Hans-Peter Blossfeld
Bildungsexpansion und Berufschancen.
Empirische Analysen zur Lage der Berufsanfänger in der Bundesrepublik.
191 S. Frankfurt a.M./New York: Campus, 1985.

Christel Hopf, Knut Nevermann und Ingrid Schmidt
Wie kamen die Nationalsozialisten an die Macht.
Eine empirische Analyse von Deutungen im Unterricht.
344 S. Frankfurt a.M./New York: Campus, 1985.

IV. Buchveröffentlichungen bei Verlagen (Fortsetzung)

John R. Nesselroade and Alexander von Eye (Eds.)
Individual Development and Social Change: Explanatory Analysis.
380 pp. New York: Academic Press, 1985.

Michael Jenne
Music, Communication, Ideology.
185 pp. Princeton, N.J.: Birch Tree Group Ltd., 1984.

Gero Lenhardt
Schule und bürokratische Rationalität.
282 S. Frankfurt a.M.: Suhrkamp, 1984.

Achim Leschinsky und Peter Martin Roeder
Schule im historischen Prozeß.
Zum Wechselverhältnis von institutioneller Erziehung und gesellschaftlicher Entwicklung.
545 S. Frankfurt a.M./Berlin/Wien: Ullstein, 1983.

Max Planck Institute for
Human Development and Education
Between Elite and Mass Education.
Education in the Federal Republic of Germany.
348 pp. Albany: State University of New York Press, 1983.

Margit Osterloh
Handlungsspielräume und Informationsverarbeitung.
369 S. Bern/Stuttgart/Toronto: Huber, 1983.

Knut Nevermann
Der Schulleiter.
Juristische und historische Aspekte zum Verhältnis von Bürokratie und Pädagogik.
314 S. Stuttgart: Klett-Cotta, 1982.

Gerd Sattler
Englischunterricht im FEGA-Modell.
Eine empirische Untersuchung über inhaltliche und methodische Differenzierung an Gesamtschulen.
355 S. Stuttgart: Klett-Cotta, 1981.

Christel Hopf, Knut Nevermann und Ingo Richter
Schulaufsicht und Schule.
Eine empirische Analyse der administrativen Bedingungen schulischer Erziehung.
428 S. Stuttgart: Klett-Cotta, 1980.

Diether Hopf
Mathematikunterricht.
Eine empirische Untersuchung zur Didaktik und Unterrichtsmethode in der 7. Klasse des Gymnasiums.
251 S. Stuttgart: Klett-Cotta, 1980.

Max-Planck-Institut für Bildungsforschung
Projektgruppe Bildungsbericht (Hrsg.)
Bildung in der Bundesrepublik Deutschland.
Daten und Analysen.
Bd. 1: Entwicklungen seit 1950.
Bd. 2: Gegenwärtige Probleme.
1404 S. Stuttgart: Klett-Cotta, 1980.

Dietrich Goldschmidt und Peter Martin Roeder (Hrsg.)
Alternative Schulen?
Gestalt und Funktion nichtstaatlicher Schulen im Rahmen öffentlicher Bildungssysteme.
623 S. Stuttgart: Klett-Cotta, 1979.