Penalized Splines - Estimation with Longitudinal Unemployment Data

Analyses of Unemployment Durations and Unemployment Risks in Germany

Dissertation zur Erlangung des Grades
eines Doktors der Wirtschaftswissenschaften (Dr. rer. pol.)
der Fakultät für Wirtschaftswissenschaften
der Universität Bielefeld

vorgelegt von
Diplom-Kauffrau Nina Westerheide

Bielefeld, im April 2012

Dekan: Prof. Dr. Herbert Dawid
Gutachter: Prof. Dr. Göran Kauermann
Prof. Dr. Harry Haupt

Tag der mündlichen Prüfung: 20. August 2012

Gedruckt auf alterungsbeständigem Papier gemäß ISO 9706

Preface

Diese Dissertation entstand während meiner Arbeit als wissenschaftliche Mitarbeiterin am Lehrstuhl für Statistik der Universität Bielefeld.

Mein besonderer Dank gilt meinem Doktorvater Prof. Dr. Göran Kauermann für die hervorragende Betreuung und Unterstützung während der gesamten Bearbeitungszeit, die angenehme Arbeitsatmosphäre an seinem Lehrstuhl sowie die Anregung zur Beschäftigung mit meinem Dissertationsthema. Des Weiteren bedanke ich mich bei Prof. Dr. Harry Haupt für die Übernahme und zügige Erstellung des zweiten Gutachtens. Meinen Kollegen der Universität Bielefeld von Flur V9 danke ich für die angenehme Zusammenarbeit und die anregenden Fachgespräche und -diskussionen.

Schließlich gilt mein Dank Patrick Wolters sowie meinen Eltern Hans und Marianne Westerheide für ihre fortwährende Unterstützung während der Erstellung der Arbeit sowie Joachim und Marlies Wolters für die Unterstützung bei der Schlussredaktion.

Bielefeld, im September 2012

Nina Westerheide

Contents

1 Introduction

Unemployment is still a central problem in European economies and an abiding theme in labour market policy. In the 1990s, in the European Union, rising unemployment led to an European employment policy and a new section on employment was introduced into the Amsterdam Treaty signed in October 1997. In the same year, the European Employment Strategy (EES) was initiated at the Luxembourg Jobs Summit with the reduction of unemployment as its major target. The process of an European employment policy is continued to this day and the section on employment is now regulated by the Lisbon Treaty which came into force in December 2009. In 2010, new impulses on the European Employment Strategy for the next 10 years were given in context of the strategy 'Europe 2020' where one of the aims is to raise employment in all member states of the European Union, see for example Bergmann (2012) and European Commission (2010). Taking a look at Germany, after over a decade of high unemployment rates, extensive labour market reforms -named the Hartz reforms- took place in the early years of this millenium. The three main points of these reforms elaborated by an independent expert commission in 2002 were to ameliorate employment policy measures and services, to mobilize unemployed individuals, and to use labour market deregulations to encourage demand in the labour market, see Jacobi and Kluve (2006). To follow the outcomes of such policies and to identify differences and changes in labour markets, there is the need of labour market statistics and labour market research.

Frequently, unemployment and employment rates are used as a macroeconomic measure in order to compare and explain regional and national labour markets as well as to point out differences between gender, age, education, and duration, as performed, for example, in official statistics in OECD (2011), European Commission (2011a), European Commission (2011b), and Bundesagentur für Arbeit (2011). This kind of statistic is of course not the only way to analyse labour markets. Focusing on a country's situation of unemployment, beside the pure analysis and comparison of unemployment rates, the analysis of effects on the unemployment duration or the risk of getting unemployed is of outstanding interest in labour market research. Apart from those papers dealing with theoretical approaches to explain these aspects of unemployment (see by way of example for the duration of unemployment Mortensen, 1970), other papers in this field of research often analyse labour market data empirically with various methods. The different effects on the duration of unemployment are frequently investigated by making use of hazard models, see as an example Hunt (1995), Steiner (2001) or Lauer (2003). But also other methods are considered to analyse the unemployment duration, see Lüdemann, Wilke, and Zhang (2006) who used censored quantile regression or Fitzenberger and Wilke (2007) who utilized censored Box-Cox quantile regression. As aforementioned, another focal point of labour market research is the analysis of the risk of getting unemployed. For this topic, there also exist a lot of different approaches to analyse various effects. One approach is to use hazard models as did, for instance, Galiani and Hopenhayn (2003) or Covizzi (2008). Other approaches are to use logistic regression (see Thapa, 2004 or Arai and Vilhelmsson, 2004), Poisson regression models (see Hammer, 1997) or to simply consider specific unemployment rates (see Reinberg and Hummel, 2002, 2003, 2005).

The main contribution of this thesis to empirical labour market research on unemployment is two-fold: The first one is to show innovative flexible approaches to analyse labour market data concerning the duration of unemployment and the risk of getting unemployed with free available statistical software. The second contribution is to analyse and contrast different effects on the duration of unemployment and the risk of getting unemployed in Germany. For this purpose, three innovative analyses are presented in Chapter 4, 5, and 6, each based upon a paper, see Westerheide and Kauermann (2012a), Kauermann and Westerheide (2012), and Westerheide and Kauermann (2012b). The first two analyses deal with flexible modelling of unemployment duration using spline-based functional hazard models. In the third analysis, the chance of getting unemployed is analysed with a spline-based generalized additive model. Different longitudinal unemployment data are used in the presented models to demonstrate the flexibility and capacity of penalized spline smoothing. The intention of the analyses in Chapter 4 and 5 is to demonstrate penalized spline smoothing as estimation routine for modelling duration time data. The statistical model being used for both analyses is built upon the hazard rate model. While the classic model here is the Cox model, see Cox (1972), we allow for non-proportional hazards in the style of varying coefficients, see Hastie and Tibshirani (1993). To estimate smooth dynamic covariate effects penalized splines are used, see Kauermann (2004). In doing so, this contribution demonstrates how to make use of available software to easily fit rather complex functional duration time models after some data management. The non-proportional hazard model is applied to two examples. In the first analysis in Chapter 4, the unemployment behaviour in Germany and the UK between 1995 and 2005 is compared based on data from national panel studies, i.e. the German Socio-Economic Panel and the British Household Panel Survey. The dynamics of covariate effects are analysed. In particular it is investigated how individual effects as gender, age, education, and the professional history increase or decrease the chances of re-employment in the two countries and it is shown how these effects change over the length of unemployment. The focus of this analysis is to contrast the two economies. In the second analysis in Chapter 5, a non-proportional hazard model with competing risks is employed to investigate dynamic covariate effects and differences between competing job markets depending on the distance between former and recent working place. For this purpose a massive database, the Scientific Use File 'Regional File 1975 - 2004' of the IAB Employment Samples from the German Federal Employment Agency is used to analyse the unemployment behaviour in Germany between 2000 and 2004. The question whether unemployed individuals change their location to take up a new job and how this readiness of relocation changes with the length of unemployment is pursued. In addition, the spatial heterogeneity within Germany is explored. Here, the focus is to contrast the spatial, economic, and individual covariate effects of the competing job markets and to analyse their general influence on the unemployed's re-employment probabilities. The intention of the third analysis in Chapter 6 is to analyse the employment status of individuals or, to be more exact, to investigate which covariates influence the chance for an individual to get unemployed. As database the Scientific Use File 'Regional File 1975 - 2004' of the IAB Employment Samples from the German Federal Employment Agency is used and for the analysis the period between 2000 and 2004 is considered. The model employed for the analysis of unemployment risk is based on the log-linear Poisson model, see McCullagh and Nelder (1989). In this, grouped covariates are allowed to simplify the model in order to downsize the computational effort. The grouped covariates contain individual characteristics like gender, age, and education and are included as fixed effects. Beside these covariates regional as well as calendrical and economic information is considered and modelled by smooth functional effects as generalized additive model using a spline-based approach, see Hastie and Tibshirani (1990) or Wood (2006). Here, the focus is to contrast the results of this analysis on the unemployment risk to other findings on unemployment risks in Germany and to compare these to outcomes of analyses considering unemployment durations and unemployment rates.

The above-mentioned analyses form the main part of this thesis. In order to set the scene for the analyses in Chapters 4, 5, and 6, two things have to be done first. In Chapter 2 some light will be cast on terms and topics concerning unemployment to give some economic background for the interpretation made in Chapters 4, 5, and 6. Then, in

Chapter 3 an introduction of the statistical methods being used for modelling the data in Chapters 4, 5, and 6 will be given to equip the reader with the necessary statistical background. Finally, a conclusion to the analyses will be drawn in Chapter 7.

2 Economic Background

In this chapter, some notes on unemployment are given as economic background for the analyses in Chapters 4, 5, and 6. After a short overview about the macroeconomic problem of unemployment by a comparison of the unemployment rates in Germany and the UK between 1995 and 2005 as well as a comparison of the unemployment rates in the old West German states and the New Länder during that time, different types of unemployment are introduced. This is followed by an outline of, the job search theory, a labour market theory which tries to explain frictional unemployment. This theory is often used as theoretical background when applying hazard models to empirically analyse the duration of unemployment. Subsequently, an introduction to the German and British unemployment compensation system is given and an abstract of the different longitudinal data sets which allow to analyse unemployment behaviour and are used in the three analyses is presented.

2.1 Unemployment in Germany and the United Kingdom between 1995 and 2005

A comparison of the job markets in Germany and the UK shows numerous differences just by looking at pure numbers in official statistics. In both countries the unemployment rate[1] was at a similar level in 1995, i.e. 8% in Germany and 8.5% in the United Kingdom. While the German unemployment rate increased until 1997 to 9.1% and decreased afterwards until 2000 to 7.2%, the UK unemployment rate decreased steadily during that period to 5.4%. This downtrend continued to an unemployment rate of about 4.7% in 2005, compared to an increase in Germany to about 9.5% in 2005. Between 1996 and 2005, the unemployment rate of Germany was above the rate of the United Kingdom. Looking at the level of the unemployment rate of men and women in 2005, in Germany the unemployment rate of women (10.3%) was above the unemployment rate of men (8.9%), but is was vice versa in the United Kingdom where men had an unemployment

[1] Here, the unemployment rate mirrors the percentage of unemployed individuals of the labour force, where the latter is the total number of all employed and unemployed individuals.

rate of 5.1% and women a rate of 4.3%. In 2005, in both countries the unemployment rate for individuals under the age of 25 is considerably higher than for those above this age, i.e. in Germany (United Kingdom) individuals below the age of 25 years had a rate of unemployment of 15% (12.9%) and those above 25 had a rate of 8.6% (3.3%). In both countries better educated individuals had a lower unemployment rate compared to those with a lower education. Focusing on long-term unemployment, the scenario in Germany is less positive due to a higher percentage of long-term unemployed compared to the UK, for details and all official figures see Eurostat (2007).

Concerning unemployment in Germany, the discrepancy between the old West German states and the New Länder is of paticular note. A difference between both parts can be seen clearly in the considered period between 1995 and 2005. The unemployment rate[2] of the old West German states (except Berlin-West) was always lower than the rate of the New Länder (including Berlin). Since the German reunification, the unemployment rate in the New Länder strongly increased from 10.2% in 1991 and 14.8% in 1995 to 19.1% in 1997 and stayed thereafter at a high level between 18.5% (in 2000) and 20.6% (in 2005). In the old West German states the unemployment rate was at a level of 6.2% in 1991 and increased to a level of 9.1% in 1995 and 10.8% in 1997. Between 1997 and 2005 the unemployment rate ranged between 8% (in 2001) and 11% (in 2005). The unemployment rate of entire Germany was at a level of 7.3% in 1991, increased until 1995 to 10.4% and remained for the period of 1995 to 2005 between 10.3% (in 2001) and 13% (in 2005), see for further details and official figures Statistisches Bundesamt, Gesis-Zuma, and WZB (2008) and Statistisches Bundesamt (2012).

2.2 Types of Unemployment

When talking about unemployment, different kinds of unemployment can be distinguished which are subdivided due to heterogeneous causes. Thus, in the following we differentiate between four kinds of unemployment: frictional unemployment, seasonal unemployment, structural unemployment, and cyclical unemployment, and follow thereby the segmentation used, for example, in Sesselmeier, Funk, and Waas (2010), Ehrenberg and Smith (2012) and Stiglitz (1997). Other textbooks distinguish only between frictional unemployment, structural unemployment, and cyclical unemployment, see for instance Samuelson and Nordhaus (2005) and Reynolds, Masters, and Moser (1991). The subsequent explanations mainly follow Sesselmeier, Funk, and Waas (2010) as well

[2]Here, all unemployment rates refer to the dependent civilian labour force.

as Stiglitz (1997), Ehrenberg and Smith (2012), Samuelson and Nordhaus (2005), and Reynolds, Masters, and Moser (1991) where additional information can be found.

Frictional unemployment describes unemployment due to movements between jobs. On the labour market, information is imperfect, so that it is time-consuming to find a job or to fill a position. Hence, frictional unemployment can be caused on one side through employees who, for example, resign from a job and search for a new employment or individuals who generally access the labour force. On the other side it can be caused by employers who, for instance, need time to fill a vacancy or whose firms declare bankruptcy. This kind of unemployment would also exist in an economy with full employment since individuals generally do not change jobs without measurable delay. Seasonal unemployment originates from seasonal up- and downturns of demand and supply in some economic sectors. Typically affected sectors are tourism, the building and construction industry, and agriculture. Usually, the height of seasonal unemployment is not dependent on the macroeconomic situation of the labour market. This kind of unemployment can also be seen as part of frictional unemployment, see Sesselmeier, Funk, and Waas (2010) and Reynolds, Masters, and Moser (1991). If there is a mismatch on the labour market between supply and demand for the labour force, we talk about structural unemployment. This term enfoldes different forms of unemployment. It arises, for instance, in cases of differences between regional demand and supply, changes of structure in specific sectors, technological changes or a structural change in supplied and demanded skills. Often, structural unemployment comes along with a longer duration of unemployment. Commonly, structural and frictional unemployment are also subsumed under the term ‘natural unemployment’, see Sesselmeier, Funk, and Waas (2010) and Reynolds, Masters, and Moser (1991). Cyclical unemployment -also known as Keynesian or demand-deficient unemployment- increases when the overall demand of labour decreases. This kind of unemployment is dependent on the business cycle of an economy, therefore the duration can not be predicted. It decreases after a cyclical upturn and increases within an economic recession. Generally, all sectors of the economy are affected by cyclical unemployment.

2.3 Job Search Theory: A Possible Labour Market Theory to Explain the Duration of Unemployment

There exist numerous labour market theories which try to explain the problems surrounding unemployment. One of these theories is the job search theory which tries to explore frictional unemployment (see for example Ehrenberg and Smith, 2012) and is often used as theoretical background when applying hazard models to empirically analyse the duration of unemployment, especially when the unemployment duration is analysed concerning the duration of the compensation, see for instance Hunt (1995), Wurzel (1993), Steiner (1997, 2001), and Hujer and Schneider (1995). We will now take a closer look at this popular theory to show a potential theoretical approach to explain the observed effects in the examples of use in Chapters 4 and 5 and give an explanation of labour market behaviour and frictional unemployment. A theoretical framework for search markets has been formulated, for instance, by Peter A. Diamond, Dale T. Mortensen, and Christopher A. Pissarides who received 'The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2010' for their important contributions to search theory, see The Royal Swedish Academy of Science (2010). In general, the contributions of Stigler (1961, 1962) -who explored a nonsequential job search model- are regarded as the beginning of this research field, see for instance Sesselmeier, Funk, and Waas (2010) or Devine and Kiefer (1993). In the following, a short introduction to the basic sequential job search model is given. We follow thereby closely the original approach of Mortensen (1970) which is adopted, for example, in the textbooks of Sesselmeier, Funk, and Waas (2010) and Ehrenberg and Smith (2012) and include remarks of the latter two as well.

In contrast to neoclassical approaches, the job search theory, or simply search theory, assumes that there is no perfect information and the working places are not homogenious. Due to the fact that there is imperfect information about vacancies as well as the characteristics of the job applicant, a match between employee and employer can only be found with effort and time. There is as much the need for unemployed who seek employment to search for job offers as for employers or firms to search for suitable employees. In order to know the wage offer and the required skills for a certain job offer, the unemployed has to search for it. In this sequential job search model, it is assumed that an unemployed can only search one vacancy within a defined period. The individual is not aware of the vacancies' characteristics, hence a random sampling can be assumed for each search. The unemployed only knows all the offers' frequency distribution and the respective wage offers for each the skill level. For simplification, the skill level which may consist of different qualifications is summarized to one variable. The individual has to decide in advance for each vacancy under which conditions he or she is willing to accept the offer, i.e. for each offer a decision has to be made whether to accept the vacancy or whether to continue searching. Thus, the unemployed's economic problem is to determine the minimum wage that makes a job opportunity acceptable. A higher acceptance wage accounts for a longer expected search until he or she gets a suitable job offer, but it also leads to a higher expected wage when getting employed. Note, that the best acceptance wage is the wage 'which equates at the margin the value of time spent searching to the present value of future benefits attributable to search' (Mortensen, 1970, p. 849). We assume that s is the minimum tolerable level of skills for a certain employer and s° is the individual's skill level. The employer's relative wage offer, i.e. the money wage offer which is divided by the market's mean money wage offer, is denoted with w and the acceptable relative wage of the particular unemployed is indicated with w°. After a previous search, the individual gets hired when he or she gets a job offer where the individual's skill level s° is larger or equal to the minimum requested skill level s and the employer's relative wage offer w is larger or equal to the relative acceptance wage or reservation wage w° of the unemployed. Note, that job offers with an equal skill level have the identic wage and job offers with a higher payment need better skilled individuals. As soon as the required skill level is known, the wage offer is also known, i.e. the relative wage offers' frequency distribution mirrors that of the required skill of all job opportunities. We now imply that a continous density function $f(w)$ with the properties $f(w) > 0$ for all $1 > a < w < b > 1$, $f(w) = 0$ otherwise, and $\int_a^b wf(w)dw = \int_a^b f(w)dw = 1$ characterize the distribution of all relative wage offers in the market. With $\hat{w}$, we indicate the maximal possible relative wage offer for an individual with the skill level s°. Due to the latter, the individual can not get higher relative wage offers than $\hat{w}$, because the application for an employment will then be refused by the employer. Now, we can derive the probability α for a randomly selected job opportunity which is acceptable for the unemployed and for which he or she has the right skills, that is

$$\alpha = P(w^\circ \leq w \leq \hat{w}) = \int_{w^\circ}^{\hat{w}} f(w)dw. \tag{2.1}$$

The expectation of the relative wage offer achieved given $w^\circ \leq w \leq \hat{w}$ is

$$e = E(w|w^\circ \leq w \leq \hat{w}) = \frac{\int_{w^\circ}^{\hat{w}} wf(w)dw}{\int_{w^\circ}^{\hat{w}} f(w)dw}, \tag{2.2}$$

which is the mean of the area under the curve between w° and $\hat{w}$. An individual applying for a job has a probability of α of getting employed in each period. Thus, the expected duration of search, i.e. the number of periods, can be deduced by $1/\alpha$. Hence, a higher reservation wage results in a longer expected duration of search, although the advantage is, that a higher wage can be expected when getting employed. The individual's choice of the reservation wage is made in order to maximize the expected human wealth, H, that is his or her discounted further earnings. When employed in period t, the individual has an income of the height of the expected relative wage e in this period, whereas there is no income without employment, except the unemployed individual is entitled to unemployment benefits. We define with p_t the probability of participation in period t, where p_t is not dependent on the employment status and q_{t-1} denominates the probability of employment at the start of this period. Now, we can denote the expected human wealth at the beginning of the search with

$$H = m \sum_{t=1}^{\infty} \frac{p_t}{(1+i)^t} [q_{t-1} e + (1 - q_{t-1}) u], \tag{2.3}$$

where i is an interest rate at which the individual can borrow and u gives the ratio of the unemployment benefits to the average offered wage in the market m. The incidence of retirement is included stochastically and its probability for each period, δ, is used to include the workers age, i.e. the value of δ is higher for older individuals and assumed to be constant for a certain individual. The probability of being a participant in period t, p_t, is approximatively $(1-\delta)^t$, thus

$$\frac{p_t}{(1+i)^t} \cong \frac{1}{(1+\rho)^t} \tag{2.4}$$

with $\rho = \frac{i+\delta}{1-\delta}$ as discount rate. The employment probability at the end of t is denoted with $q_t = 1 - (1-\alpha)^t$, thus $q_{t-1} = 1 - (1-\alpha)^{t-1}$ is the employment probability at the end of $t-1$. Using the latter equation and (2.4), we get for H

$$H = m\ e \sum_{t=1}^{\infty} \left(\frac{1}{1+\rho} \right)^t - \frac{m(e-u)}{1-\alpha} \sum_{t=1}^{\infty} \left(\frac{1-\alpha}{1+\rho} \right)^t \tag{2.5}$$

which can also be written as

$$H = \frac{m}{\rho} - \left(\frac{\alpha(w^\circ, \hat{w})\ e(w^\circ, \hat{w}) + \rho u}{\rho + \alpha(w^\circ, \hat{w})} \right) \tag{2.6}$$

where (2.1) and (2.2) specify the functions $\alpha(w^\circ, \hat{w})$ and $e(w^\circ, \hat{w})$. As previously mentioned, the individual selects a reservation wage which maximizes his or her expected human wealth, so the relative reservation wage is that value of w° which maximizes H. The optimal choice of the relative reservation wage has to fulfil

$$\alpha(w^\circ, \hat{w})[e(w^\circ, \hat{w}) - w^\circ] = \rho[w^\circ - w], \tag{2.7}$$

where (2.7) can be derived from the derivative of H with respect to w°. When an individual gets an offer for a relative wage of w, he or she has two opportunities: to reject the offer which results in an income of u in period t or to accept the offer with an outcome of a subsequent lifelong income of w. This leads to the searching cost for an additional job offer, which is $w-u$. In case of a positive cost, the individual will only look for another job offer when he or she has the expectation of a higher wage than w with a probability > 0. Expecting a lifelong net stream of income which is equal to $e(w, \hat{w}) - w$ with the probability $\alpha(w, \hat{w})$ for the next search, the individual will only refuse the offer w for $\alpha(w, \hat{w})[e(w, \hat{w}) - w]/\rho$, i.e. the expected present value of the stream, higher than the searching cost $w - u$ of an additional search. Put into words, the reservation wage 'is that offer which equates the marginal cost of search to the present value of marginal expected gains from searching' (Mortensen, 1970, p. 851). Equation (2.7) is thereby the only reasonable solution for $\hat{w} \geq u$ and $\hat{w} \geq w^{\circ}$ what can be seen when using (2.1) and (2.2) to rewrite (2.7):

$$\int_{w^{\circ}}^{\hat{w}} (w - w^{\circ}) f(w) dw = \rho(w^{\circ} - u). \tag{2.8}$$

For all $u \geq \hat{w}$ there is no need to search for an individual because the individual's maximal possible relative wage offer is not higher than the relative unemployment benefit u which is provided in case of no search. Accordingly, an individual searches only in case of $\hat{w} > u$. From (2.8), we can conclude that $\hat{w} > w^{\circ} > u$, i.e. an individual will agree to a relative wage offer higher than the minimum he or she can expect and lower than the maximal possible relative wage, if he or she knows the nature of certain wage offers only imperfectly.

As pointed out in Ehrenberg and Smith (2012), the benefits of a countries unemployment insurance have an impact on the individual's unemployment cost. The lower the cost is, i.e. the higher the unemployment compensation is, the higher the individual's reservation wage gets, see also Mortensen (1970). An increasing reservation wage calls for an increase of the expected duration of unemployment as well as the expected wage rate after unemployment. This leads to the assumption that a generous unemployment compensation results in a higher reservation wage extending the unemployment duration which, ceteris paribus, raises the unemployment rate. Some studies support this conclusion, see for an overview, for instance, Ehrenberg and Smith (2012). Other studies found effects of the pure entitlement to unemployment compensation, see for an outline also Ehrenberg and Smith (2012). An example is, for instance, from Katz and Meyer (1990) who found an increased probability of accepting a job at the end of the benefit entitlement in the USA, see for a theoretical derivation also Mortensen (1977).

For Germany, a lot of different studies tried to examine the effects of different aspects of the unemployment compensation system on the duration of unemployment, see for instance Hunt (1995), Steiner (1997, 2001) or Hujer and Schneider (1995), but the results support the assumptions only in part. Hunt (1995) discovered in her analysis inter alia that an increase in the possible unemployment benefit duration leads to a rise of the duration of unemployment, especially for individuals between 44-48 years, but the effect of a cut in the compensation rate for unemployed without children remained ambiguous. Steiner (1997) found out that the eligibility to unemployment benefits increased the unemployment duration only for men, while having only a small influence on the duration for women. Furthermore, he could only identify a very small effect of marginal reductions of the income-replacement on the unemployment behaviour of men and women. Steiner (2001) concluded from his analysis that the eligibility to unemployment compensation has an impact on the unemployment duration and that -at least for men- the small changes in the replacement ratio are far less outstanding than the entitlement to compensation. In Hujer and Schneider (1995), effects of the length of unemployment compensation on the unemployment duration were found for men, but the findings for women were not all in line with the assumption and partly inconsistent with the search theory. Hence, the studies mentioned above leave a controversial image concerning the effects of different factors of the unemployment compensation system on the duration of unemployment, see also the explanatory notes in Steiner (1997) who gives an overview on further empirical studies analysing the effects of various aspects of the German unemployment compensation system. A critical discussion about the topic of unemployment compensation in the context of search theory can be found in Sesselmeier, Funk, and Waas (2010).

Further information about the job search theory is available in Mortensen (1986) who gives a detailed and formal introduction as well as in the literature surveys of Rogerson, Shimer, and Wright (2005) and Lippman and McCall (1976a, 1976b) or in Devine and Kiefer (1993) who give an overview about empirical labour economics concerning job search theory. In Woodbury and Davidson (2002) an introduction to job search theory is given and the development and the impact of job search theory on empirical work and public policy is shown. A comprehensible overview is available in Fitzgerald (1998) just as short presentations are given in Franz (2009), Layard, Nickell, and Jackman (2009) or Wurzel (1993). Last but not least, beside the application to the labour market, search theory is also used in other areas: monetary theory (see e.g. Kiyotaki and Wright, 1993) and marriage markets (see e.g. Mortensen, 1988 or Oppenheimer, 1988) are just two examples.

2.4 The German and British Unemployment Compensation System between 1995 and 2005

As it is apparent from Section 2.3, the unemployment compensation has a theoretical impact on the duration of unemployment in the job search context. Empirical effects of the benefit systems on the duration of unemployment could also be shown, even though they were not strictly in tune with the job search theory. To better understand the unemployment compensation systems of the countries considered in the analyses in Chapters 4, 5, and 6, i.e. Germany and the UK, and to be able to evaluate whether the effects in these analyses are influenced by the eligibility for benefits, a detailed introduction to these systems is given in the following. Both systems differ strongly from each other during the considered time between 1995 and 2005. While in Germany in the case of unemployment beneficiaries receive an income-related compensation, in the United Kingdom they only get a weekly flat rate.

The German Unemployment Compensation System

For the explanation of the German unemployment compensation system, we follow Clasen (2005), Werner and Winkler (2003), Plaßmann (2002), Franz (1996, 1999, 2006), Lampert (1996), Niesel (1998, 2002, 2005), Münder (2009), Steiner (1997), Jacobi and Kluve (2006) as well as the European Commission (2005a, 2005b) and Bundesministerium für Arbeit und Sozialordnung (1995, 1997). In Germany, it is obligatory for all employees to take part in the unemployment compensation system. This system changed twice during the period between 1995 and 2005, although most of the time (January 1998 - December 2004) it was covered by the Act to Reform Employment Promotion (Arbeitsförderungsreformgesetz) which reformed the Employment Promotion Act (Arbeitsförderungsgesetz) and was promulgated in March 1997. The changes made did not fundamentally touch the considered period. Until the end of December 2004, the German unemployment compensation system was built up of two parts: the contribution financed unemployment benefits (Arbeitslosengeld) and the tax financed unemployment assistance (Arbeitslosenhilfe) which were both based on a certain percentage of the former net earnings. Unemployed who received unemployment benefits could subsequently be entitled to unemployment assistance. Beginning with January 2005, unemployment compensation consists of unemployment benefits and the so-called 'Arbeitslosengeld II' which is a combination of the former unemployment assistance and social assistance and is paid as a flat rate. As aforementioned, until the end of December 1997 the 'Arbeitsförderungsgesetz' regulated the payment of benefits. Individuals who were registered as unemployed and available for a job placement were entitled to receive unemployment benefits when they completed the qualifying period of insured employment, i.e. they paid unemployment insurance contributions for at least 12 months in the last three years. Unemployed below the age of 42 years could receive benefits for 6 up to 12 months (156 and 312 weekdays without Sundays) depending on the duration of compulsory insurance coverage. Unemployed with a minimum age of 54 could receive unemployment benefits for up to 32 months (832 weekdays without Sundays) which was the maximum length of entitlement. The income-replacement ratio for unemployed with (without) children was fixed at 67% (60%) of the net income of the last 6 months of employment or the upper earnings limit. After the entitlement to unemployment benefits, unemployed could receive unemployment assistance depending on a means test. The income on which the assistance was calculated was reduced by 3% per year of entitlement. The income-replacement ratio was thereby 57% (53%) of the basis of contribution assessment (Beitragsbemessungsgrundlage) for unemployed with (without) children and paid limitless, except for some groups who e.g. did not pay any social contributions. For the latter, unemployment benefits were restricted to one year and could be followed by social assistance. For these or more information on this and further topics concerning the state of affairs during this period, see for example Steiner (1997), Franz (1996), Lampert (1996), Bundesministerium für Arbeit und Sozialordung (1995), and Plaßmann (2002). Since January 1998, the German Employment Promotion Law (Arbeitsförderungsrecht) was integrated in the Social Code (Sozialgesetzbuch), Book III, but apart from this only small changes have been made. Now, the net wages of the last 52 weeks were taken into account for assessment. The income-replacement ratio remained unchanged, but the duration of entitlement to unemployment benefits changed lightly. The entry age for a longer eligibility was increased, e.g. unemployed below 45 years received up to 12 months unemployment benefits depending on the duration of compulsory insurance coverage within an extension of the regular time frame of 3 years by 4 years. Now, for the maximum entitlement of 32 months, the unemployed had to be at least 57 years old. For these and more detailed information about the situation between January 1998 and December 2004, see for instance Niesel (1998, 2002), Werner and Winkler (2003), Plaßmann (2002), Franz (1999), and European Commission (2005b). Then, in the first half of the 2000s, the Hartz reforms took place in Germany resulting in four laws. The first two laws basically came into force in January 2003 and the third law came into force stepwise beginning with January 2004. Part of the third law were changes of the unemployment benefit which became effective in January 2005. The fourth law basically came into force in January 2005, covering amongst others the abolishment of the unemployment assistance and the introduction of 'Arbeitslosengeld II'. The income-replacement ratio of the unemployment benefits did not change. Due to a transitional regulation, changes concerning the duration of unemployment benefits, their qualifying period and time frame were not affected by the amendments for individuals who were entitled to unemployment benefits until January, 31st 2006. Thus, the previous regulations stayed in use for these individuals. Consequently, the changes of the unemployment benefits did not touch the considered period until April 2005. The tax-based unemployment assistance (Grundsicherung für Arbeitssuchende or Arbeitslosengeld II) was adopted in January 2005. It consists of a regular fixed benefit and may contain additional benefits for reasonable costs. In January 2005, the regular benefit for single persons in the old West German states was 345€ and in the newly formed German states 333€ per month with an additional reduced rate for family members, see European Commission (2005a). Individuals in need and able to work could draw on the latter, not only after the exhaustion of an entitlement to unemployment benefits. A means test is an obligatory condition to receive these benefits. For these and more information about the latter changes, see for instance Niesel (2005), Münder (2009), Jacobi and Kluve (2006), Clasen (2005), Franz (2006), and European Commission (2005a). In Germany, in the considered period and under certain conditions, older unemployed at the age of 60 years had the possibility to retire early after a previous period of unemployment of 52 weeks within the last 1 1/2 years. Since January 1997, the age limit increased from 60 years to 65 years with certain exceptions. For individuals born between January 1st 1937 and December 31st 1945, it was still possible to retire at the age of 60 years with a reduced retirement pension. For those who were born afterwards, the earliest possible entry age for early retirement was raised stepwise to 63 years with certain exceptions. Since January 2005, only individuals born before January 1st 1952 were entitled to this kind of retirement after unemployment, see also for further information Bundesministerium für Arbeit und Sozialordung (1995, 1997), European Commission (2005a, 2005b), Bundesministerium für Gesundheit und Soziale Sicherung (2005), Clasen (2005), Reinhardt (2006), and Clasen, Davidson, Ganßmann, and Mauer (2006). An overview of the German unemployment compensation system in the considered period can also be found in Clasen, Davidson, Ganßmann, and Mauer (2006) and partly in Plaßmann (2002) or the literature mentioned above.

The British Unemployment Compensation System

We will now take a look at the British unemployment compensation system following mainly European Commission (2005a), Werner and Winkler (2003) or Clasen (2005) as well as partly Wikeley (1996) and Clasen, Davidson, Ganßmann, and Mauer (2006). In the United Kingdom, the unemployment compensation system was changed once at the beginning of the considered period. In 1995, the Jobseekers Act was passed in the United Kingdom and came completely into force in October 1996. It regulates the contribution-based jobseekers' allowance and the income-based jobseekers' allowance. The three most important changes were, in accordance with Wikeley (1996), the cutback on the eligibility to the contribution-based jobseekers' allowance from one year to six months, the lowering of the benefits for young individuals between 18 and 24 years and the condition that individuals who draw benefits sign a jobseeker's agreement with certain impositions. Between October 1996 and April 2005, the unemployment compensation system -which is obligatory for all employees and some of the self-employed- did barely change, see The Stationery Office (1995) and `http://www.legislation.gov.uk/ukpga/1995/18` (Last checked: 02/28/2012) for changes to legislation. To receive contribution-based jobseekers' allowance, unemployed need a certain qualifying period: In one of the two tax years before the year of the benefit claim unemployed had to pay at least 25 times the minimum contribution of the considered year and in both tax years the minimum contribution had to be paid at least 50 times. The contribution-based jobseekers' allowance is paid as flat-rate for 182 days per period of unemployment. In January 2005, £55.65 per week were paid for individuals with an age of 25 or older, individuals between 16 and 17 received £33.50 per week and for those between 18 and 24 £44.05 per week were paid, see European Commission (2005a). Before the introduction of the Jobseekers Act, unemployment benefits were paid for 312 days, what corresponds to one year excluding Sundays, see Werner and Winkler (2003). Subsequently to this benefit the income-based jobseekers' allowance could be received under further conditions relating to the savings and the working time of the individual's life partner. There are also special rules for claimants under 18 years. The income-based jobseekers' allowance is tax-financed and also paid as a flat-rate. For the income-based jobseekers' allowance there exists no qualifying period and it is paid as long as the unemployed is in need and his or her means test is positive. The basic levels of the income-based jobseekers' allowance for couples, where both partners are under 18 years, was £66.50 per week and for couples, where both partners are above 18 years, it was £87.30. For singles the income-based jobseekers' allowance equals the contribution-based jobseekers' allowance, see European Commission (2005a). In the United Kingdom there exists no early state pension, see for example European Commission (2005a), but individuals with an occupational or personal pension provision may retire early under certain conditions, though it does not seem to be very popular, see Clasen, Davidson, Ganßmann, and Mauer (2006). For further information on the British unemployment compensation system, see for instance European Commission (2005a) or Clasen (2005) or the literature mentioned above.

2.5 Data Sets for Empirical Analyses of Unemployment

To analyse the duration of unemployment or the risk of unemployment there is the need for data sets which include miscellaneous information on the individual's socio-demographic-related characteristics and employment history. Hence, useful data to analyse the behaviour of unemployment in Europe can be taken from national panel studies, see e.g. Sweden (Household Market and Nonmarket Activities (HUS)), the Netherlands (Dutch Socio-Economic Panel (SEP)), Luxembourg (Panel Socio-économique, Liewen zu Letzeberg (PSBE)), Italy (Indagine Longitudinale sulle Famiglie Italiane (ILFI)), and Switzerland (Swiss Household Panel (SHP)). There also exist European panels like the 'European Community Household Panel' (ECHP), the 'European Union Statistics on Income and Living Conditions' (EU-SILC), and the 'Consortium of Household Panels for European Socio-Economic Research' (CHER). For the analyses in Chapter 4 national panel data collections from Germany and the United Kingdom, namely the 'German Socio-Economic Panel' (GSOEP) and the 'British Household Panel Survey' (BHPS), are used and described in the following.

German Socio-Economic Panel

The German Socio-Economic Panel, established in 1984, is provided by the German Institute for Economic Research (Deutsches Institut für Wirtschaftsforschung (DIW)). Once a year, the latter releases a collection of different data sets which contain microdata on biographical information of individuals and their households up to the recent 'wave'. The households are representatively chosen and all members from the age of 16 are personally interrogated annually. Beside the regular questions of the interview, varying additional topics are included every three to six years in the questionary. The longitudinal study started in 1984 with a sample of 5921 households (thereof 4528 households with a head of the household not belonging to the main groups of foreigners in Germany and 1393 households with a head of the household belonging to the main groups of foreigners) with 12245 successfully interrogated individuals, see Haisken-DeNew and Frick (2005). In the following years a few more samples were introduced and contained, for instance, special households like households of immigrants or households in the former German Democratic Republic, other samples are supplementary samples. In 2007, 11689 households participated in the study and from the 22470 possible participants 21232 individuals were successfully interrogated, see von Rosenbladt (2008). This and more information is available in Wagner, Göbel, Krause, Pischner, and Sieber (2008), Haisken-DeNew and Frick (2005) and on the website `http://www.diw.de/de/soep` (Last checked: 02/28/2012) of the GSOEP. The webside also provides further downloadable documentations and an online help called 'SOEPinfo' (`http://panel.gsoep.de/soepinfo2010`, last checked: 02/28/2012) where information about the different GSOEP variables across all waves can be found. An introduction to the GSOEP can also be found in Hanefeld (1987). To generate the data used in the analysis in Chapter 4, the following GSOEP data sets are used: ARTKALEN, PPFAD (both wave W) as well as the $PGEN-files from wave L to wave V.

British Household Panel Survey

The British Household Panel Survey is conducted by the ESRC UK Longitudinal Studies Centre with the Institute for Social and Economic Research at the University of Essex and provided by the UK Data Archive. The BHPS contains socio-economic information on an individual and household level in Britain. From wave 11 on, the whole of the UK was included in the survey. Participants of the survey answer annually a questionnaire which consists of certain core themes and changing topics. In 1991, the sample started with wave one and 10264 individuals from Great Britain (aged 16 years or older) of 5505 households were interviewed for this survey, see Lynn (2006). Since the fourth wave also children between 11 and 15 are interviewed briefly. Additional samples have been included since the beginning of the survey and the total sample size now includes around 10000 households in the United Kingdom, see `http://www.iser.essex.ac.uk/bhps/about/sample` (Last checked: 03/20/2012). This and further information about the British Household Panel Survey can be found in Taylor, Brice, Buck, and Prentice-Lane (2010), Lynn (2006) and on the survey website `http://www.iser.essex.ac.uk/bhps` (Last checked: 02/28/2012). The latter also offers a comprehensive online documentation about all collected information, the so-called 'Volume B - the Codebook ', see `http://www.iser.essex.ac.uk/bhps/documentation/volb/index.html` (Last checked: 02/28/2012). The BHPS data used in Chapter 4 is based on the basis data set 'SN 5151 British Household Panel Survey; Wave 1-14, 1991-2005' and the additional data set 'SN 3954 British Household Panel Survey Combined Work-Life History Data, 1990-2005', see Halpin (2006).

The two panels allow to explore and investigate empirically the different economic situations in the two countries. In both data sets, there is a huge number of different covariates describing the individual- and household-specific socio-demographic information. It also gives information about the individual's employment history, while the number of observations is limited. Beside national panel data collections there exist numerous national administrative data sets which contain information about some socio-demographic-related characteristics and especially information about the employment history of individuals. One example is the enormous German administrative data set 'IAB Employment Samples Regional File 1975 - 2004' of which the analyses of Chapters 5 and 6 made use.

IAB Employment Sample Regional File 1975 - 2004

The 'IAB Employment Sample Regional File 1975-2004' is provided by the Research Data Centre (Forschungsdatenzentrum (FDZ)) of the German Federal Employment Agency (Bundesagentur für Arbeit (BA)) at the Institute for Employment Research (Institut für Arbeitsmarkt- und Berufsforschung (IAB)) which also offers other labour market-related data sets, see for an overview `http://fdz.iab.de/en/FDZ_Overview_of_Data.aspx` (Last checked: 03/20/2012). The 'IAB Regional File 1975-2004' is the fifth updated version and encompasses 2% of all employees who are subject to social security and have been working for at least one day between 1975 and 2004. It was sampled out of the 'Employee and Benefit Recipient History' of the IAB which consists of the 'Employment History' and the 'Benefit Recipient History' of the IAB. The first source covers information about employees and apprentices who are subject to social insurance contribution. The second source contains information about individuals getting wage replacement benefits from the German Federal Employment Agency. Hence, civil servants, self-employed persons and students are not included. Though this huge administrative data set does not contain as much socio-economic information as the panel studies mentioned above, the details concerning the employment history are much more precise. It consists of the employment history of more than 1,3 million individuals on a day-to-day basis. Therefore, we have 24,936,176 data rows including information about gender, age, education, profession, local information about the employer, type of employment or benefit, start and end date of employment, and benefit receipt period, etc. For the analyses in Chapters 5 and 6 the scientific use file was employed to generate the data sets. For researchers there is also the possibility to work with a weakly anonymous version of the IAB 'Regional File 1975-2004' which offers more detailed information, but can only be accessed on-site with a subsequently remote data access. This and a more detailed introduction to the data is given in Drews (2008) and general information can be found on the website of the FDZ, see `http://fdz.iab.de/en.aspx` (Last checked: 03/20/2012).

3 Statistical Background

The examples of use in Chapters 4 and 5 are based on non-proportional hazard models, while the example of use in Chapter 6 uses a generalized additive model or more specifically an additive Poisson model to analyse the data. In the following subsections a brief overview of the theoretical background for these models is given. Chapter 3.1 deals with hazard models which are used to model survival data. In this context the Cox proportional hazard model is introduced and subsequently extended to a non-proportional hazard model where smooth functional covariate effects are allowed to vary with time. In Chapter 3.2 it is shown how penalized spline smoothing (P-spline smoothing) is used as estimation routine for smooth unknown functions and can be utilized in generalized additive models. To begin with, generalized additive models are introduced in short, followed by an introduction of P-spline smoothing considering different spline bases and penalties. Then, the interference of generalized additive models via the representation as generalized linear mixed model is explained. Thereafter, it is shown how non-proportional hazard models can be linked with generalized additive models for Poisson-distributed variables.

3.1 Modelling Survival Data

In the following subsections a brief introduction of the hazard rate is given before the Cox proportional hazard model is presented. Then a functional hazard model which allows for smoothly time-varying covariate effects is motivated for contexts with and without competing risks.

3.1.1 The Hazard Rate

An essential component in survival analysis is the hazard rate $h(t)$ which is also known as hazard function or force of mortality. Following for example Collett (1996), Klein and Moeschberger (2003), Lawless (2003), and Tableman, Kim, and Portnoy (2004) it takes the form

$$h(t) = \lim_{\Delta t \to 0} \frac{P[t \leq T < t + \Delta t \mid T \geq t]}{\Delta t} \tag{3.1}$$

and describes the instantaneous rate of death or failure at time t, provided the individual outlives up to time t. The denominator Δt denotes a small time interval and T is a continuous non-negative random variable which represents the survival times of individuals in a population. The approximative probability of an event in $[t, t + \Delta t)$ under the condition that the individual had no event up to t is given with $h(t)\Delta t$. The hazard rate is restricted to be non-negative.

3.1.2 The Cox Proportional Hazard Model

A well-established and up to the present day applied approach to analyse duration time data is the proportional hazard model as introduced by Cox (1972). In the following a brief introduction of the Cox proportional hazard model which is also known as Cox model or proportional hazard model is given. Here we follow the standard works of Cox and Oakes (1984), Kalbfleisch and Prentice (2002), Klein and Moeschberger (2003), Lawless (2003), Collett (1996), Marubini and Valsecchi (2004), and Therneau and Grambsch (2004).

The hazard rate or hazard function $h(t, x_i)$ of a classical Cox proportional hazard model for an individual i with p covariates under investigation at time t can be written as

$$h(t, x_i) = h_0(t)c(x_i^T\beta) = h_0(t)\exp\big\{\sum_{j=1}^{p} x_{ij}\beta_j\big\} = \exp\big\{\beta_0(t) \ + \sum_{j=1}^{p} x_{ij}\beta_j\big\} \tag{3.2}$$

where $h_0(t) = \exp\big\{\beta_0(t)\big\}$ is the time-dependent baseline hazard, $x_i = (x_{i1}, \ldots, x_{ip})^T$ describes a set of p fix and time-independent covariates for an individual i with $i = 1, \ldots, n$ and $\beta = (\beta_1, \ldots, \beta_p)^T$ is a parameter vector with β_j fix and time-independent parameters, $j = 1, \ldots, p$. The hazard rate $h(t, x_i)$ has to be positive, therefore the exponential function is generally used as the known function $c(x_i^T\beta)$. Due to the fact that the baseline hazard, an unspecified non-negative function, is treated nonparametrically and the covariate effects are assumed to be parametrical, this is a semiparametric model. The baseline hazard can be seen as the intercept of the model. It describes the hazard rate of an individual i when all its covariates take the value zero. Taking the logarithm of (3.2) it can easily be seen that the result resembles the structure of a common linear model:

$$\log\big\{h(t, x_i)\big\} = \log\big\{h_0(t)\big\} + \sum_{j=1}^{p} x_{ij}\beta_j = \beta_0(t) + \sum_{j=1}^{p} x_{ij}\beta_j. \tag{3.3}$$

The observations in the Cox model consist of a data triplet (t_i, d_i, x_i) where t_i denotes the duration time for the ith individual, d_i, the censoring variable, states whether the event has occured ($d_i = 1$) or is right-censored ($d_i = 0$) and $x_i = (x_{i1}, \ldots, x_{ip})^T$ denotes a set of p exogenous covariates under investigation of the ith individual.

The term 'proportional hazard model' for the Cox model derives from the fact that the ratio of the hazard rates of two individuals i and k with different covariate vectors x_i and x_k is a constant:

$$\frac{h(t, x_i)}{h(t, x_k)} = \frac{h_0(t) \exp\{\sum_{j=1}^{p} x_{ij}\beta_j\}}{h_0(t) \exp\{\sum_{j=1}^{p} x_{kj}\beta_j\}} = \exp\{\sum_{j=1}^{p}(x_{ij} - x_{kj})\beta_j\}. \tag{3.4}$$

As seen in (3.4) the ratio can be reduced by the time-dependent baseline hazard and thus the hazard rates are proportional and constant over time. The ratio in (3.4) is known as the relative risk or hazard ratio. Because the proportionality of the hazard rates may cause difficulties as soon as it can not be maintained, a model with smooth time-varying covariate effects is introduced subsequently to this subsection.

A more profound introduction to the Cox model can be found, for example, in Cox and Oakes (1984), Kalbfleisch and Prentice (2002), Klein and Moeschberger (2003) or Collett (1996).

3.1.3 Functional Hazard Model

To allow smooth time-varying covariate effects, we assume the flexible hazard rate

$$h(t, x_i) = \exp\{\beta_0(t) + \sum_{j=1}^{p} x_{ij}\beta_j(t)\} \tag{3.5}$$

where $\beta_j(t)$ denotes a smooth function which varies in time t, see for instance Hastie and Tibshirani (1993) and Kauermann (2005). This functional shape extends the Cox model and incorporates non-proportional hazard behaviour. The flexible hazard rate can be extended by additional smooth functions

$$h(t, x_i, s_i, m_i) = \exp\{\beta_0(t) + \sum_{j=1}^{p} x_{ij}\beta_j(t) + \gamma(s_i) + \delta(m_i)\} \tag{3.6}$$

where $\gamma(s_i)$ and $\delta(m_i)$ describe additional smooth effects like calendar effects or effects of environmental quantities, see for an example of use Westerheide and Kauermann (2012a). Accounting for these effects, the hazard is allowed to vary, for instance, over the calendar time, season, unemployment rate, location, etc. The corresponding covariates are denoted in (3.6) with s_i and m_i. To achieve identifiability, we assume that the additional smooth functions $\gamma(s_i)$ and $\delta(m_i)$ integrate out to zero.

Extension to a Functional Hazard Model with Competing Risks

In some contexts we have competing risks, i.e. the individual has the possibility to fail due to one of K events with $K \geq 2$, see for instance Kalbfleisch and Prentice (2002), Lawless (2003) or Klein and Moeschberger (2003). In case an event occurs, we observe for each individual the time T and the kind of event given with $k \in \{1, 2, \ldots, K\}$. The overall hazard rate $h(t)$ then takes the form

$$h(t) = \sum_{k=1}^{K} h_k(t), \tag{3.7}$$

with $h_k(t)$ as cause-specific hazard rate for risk k of the form

$$h_k(t) = \lim_{\Delta t \to 0} \frac{P[t \leq T < t + \Delta t, d = k \mid T \geq t]}{\Delta t}, \tag{3.8}$$

for $k = 1, 2, \ldots, K$, see further details, for example, Kalbfleisch and Prentice (2002), Lawless (2003), Klein and Moeschberger (2003) or Marubini and Valsecchi (2004). To take competing risks into account we replace model (3.6) with a functional hazard model with competing risks:

$$h(t, x_i, s_i, m_i) = \sum_{k=1}^{K} h_k(t, x_i, s_i, m_i), \tag{3.9}$$

with $h_k(t, x_i, s_i, m_i)$ as additive competing hazard rate. In the case of competing risks the censoring variable $d_i \in \{0, 1, \ldots, K\}$ expresses which of the k events occured at time t or is right-censored ($d_i = 0$), see for an example of use Kauermann and Westerheide (2012).

3.2 Estimation of the Smooth Functional Components with P-Spline Smoothing

The basic idea of estimating the unknown smooth functional components such as $\beta_j(t)$ is to substitute each unknown function by a linear combination of high dimensional parametric basis functions, i.e.

$$\beta_j(t) = B_j(t) b_j, \tag{3.10}$$

where $B_j(t)$ are high dimensional basis functions and b_j are the corresponding basis coefficients, see for instance Wood (2006) or Fahrmeir, Kneib, and Lang (2009). Often B-splines (see de Boor, 2001) are used to build the high dimensional basis $B_j(t)$, see

Ruppert, Wand, and Carroll (2003). In contrast to classical spline smoothing where at each observed value of variables a knot is placed (see for instance Eubank, 1999), we make use of 'low rank smoothing' (see for instance Hastie, 1996 and Wood, 2003) which works with a reduced set of knots that is still large enough to contain the functional fit while lowering the computational effort. Additionally, a penalty is introduced to penalize too much variability of the fitted curves. This approach leads to 'P(enalized)-spline smoothing' which goes back to O'Sullivan (1986). Further details can be found, for example, in Ruppert, Wand, and Carroll (2003) or Fahrmeir, Kneib, and Lang (2009). The term 'P-splines' was introduced by Eilers and Marx (1996) who used a rather large number of knots and a difference penalty on the neighboured B-spline coefficients instead of the integral of a squared higher derivative of the fitted curve which was the standard approach to penalize splines until then (see O' Sullivan, 1986 or Wahba, 2006). Furthermore, Eilers and Marx (1996) demonstrated the link between both approaches for second-order differences.

We will now take a closer look at P-spline smoothing which is one of the possible methods to estimate the smooth functional components in a generalized additive model to which the functional hazard models in (3.6) and (3.9) can also be linked. In the following, generalized additive models are briefly presented in Section 3.2.1, before P-spline smoothing is introduced in Section 3.2.2. Subsequently, the interference of P-spline smoothing for a generalized additive model via the representation as generalized linear mixed model is pointed out in Section 3.2.3. The link between a functional hazard model and a generalized additive model for Poisson-distributed variables is pointed out extensively in Section 3.2.4. The following explanations are mainly based on Fahrmeir, Kneib, and Lang (2009), Ruppert, Wand, and Carroll (2003), Wood (2006), Krivobokova (2006), Eilers and Marx (1996) as well as Kauermann (2006, 2010), Wand (2003), Dierckx (1993), and Hastie and Tibshirani (1990).

3.2.1 Generalized Additive Models

Assume that we have independent distributed response variables y_i of the exponential family (see for instance McCullagh and Nelder, 1989 or Fahrmeir and Tutz, 2001) that is

$$y_i \sim \exp\left(\frac{y_i\theta_i - b(\theta_i)}{a(\phi, \omega_i)} + c(y_i, \phi, \omega_i)\right), \tag{3.11}$$

with the canonical parameter θ_i and functions $b(\theta_i)$, $c(y_i, \phi)$, and $a(\phi, \omega_i) = \frac{\phi}{\omega_i}$ with ϕ as the dispersion parameter and ω_i as a known weight, $i = 1, \ldots, n$. A generalized additive model (GAM) for the response variables y_i of the exponential family with the mean $E(y_i) = b'(\theta_i) = \mu_i$ and the variance $Var(y_i) = b''(\theta_i)a(\phi, \omega_i)$ has the following structure:

$$\begin{aligned} g(\mu_i) = \eta_i &= \sum_{j=1}^{p} f_j(x_{ij}) + \beta_0 + \beta_{1z} z_{i1} + \cdots + \beta_{kz} z_{ik} \\ &= \sum_{j=1}^{p} f_j(x_{ij}) + z_i^T \beta_z \end{aligned} \tag{3.12}$$

with $g(\mu_i)$ as link function and η_i as additive predictor. This model is also known as generalized semiparametric additive model. Through the response function $h(\eta_i)$ the additive predictor η_i is linked with the conditional mean $E(y_i) = \mu_i = h(\eta_i)$. The link function g is the inverse function of the response function h, i.e. $g = h^{-1}$. The smooth functions $f_j(x_j)$ describe non-linear effects of the metric covariates x_j, $j = 1, \ldots, p$ and interact additively. Here, they are to be estimated using P-spline smoothing. The linear part of a generalized additive model consists of a $k+1$-dimensional vector $z = (1, z_1, \ldots, z_k)^T$ with k predictor variables and the corresponding parameter vector $\beta_z = (\beta_0, \beta_{1z}, \ldots, \beta_{kz})^T$. However, the linear part (except the intercept) is not always included in a generalized additive model.

In particular, for Poisson-distributed response variables $y_i \sim Poisson(\lambda_i)$ with $E(y_i) = \lambda_i = \exp(\eta_i)$ the response function of an additive Poisson model is given by

$$h(\eta_i) = \exp(\sum_{j=1}^{p} f_j(x_{ij}) + z_i^T \beta_z). \tag{3.13}$$

The unknown non-linear smooth functions f_j in a generalized additive model are to be estimated through m_j basis functions

$$f_j(x_j) = \sum_{q=1}^{m_j} b_{jq} B_{jq}(x_j), \tag{3.14}$$

with B_{jq} as basis functions such as truncated polynomial or B-spline basis functions, x_j as value of the observed values $x_{1j}, \ldots, x_{nj}$ of the j-th covariate and the corresponding coefficients b_{jq}. This and further information can be found in the reference books of Fahrmeir, Kneib, and Lang (2009) and Ruppert, Wand, and Carroll (2003), Wood (2006) and Hastie and Tibshirani (1990).

3.2.2 P-Spline Smoothing

The smooth functional components $f_j(x_j)$ can be estimated using P-spline smoothing. As pointed out before, the idea of P-splines is to use a linear combination of high dimensional basis functions for the estimation of the unknown functions $f_j(x_j)$ and to introduce a penalty which penalizes too much variability of the estimation. The P-spline basis might exist of a truncated polynomial basis or a B-spline basis as well as of other bases. Different approaches for the penalty can be used as well. In the following, a short introduction is given to both types of P-spline bases mentioned above. Furthermore, different penalties are introduced. Most of the explanations are based on Fahrmeir, Kneib, and Lang (2009), Ruppert, Wand, and Carroll (2003) and Dierckx (1993) as well as Eilers and Marx (1996), Kauermann (2010), and Krivobokova (2006). For convenience, we only use one smooth functional component $f_j(x_j)$ and use $f(x) = \sum_{q=1}^{m} b_q B_q(x) = B(x)b$ as denotation, where B_q are basis functions and b_q the corresponding m coefficients.

Truncated Polynomial Basis

To represent the structure of the unknown smooth function f on an interval $[a, b]$, we define the knots $a = \kappa_0 < \cdots < \kappa_{d+1} = b$ with the interior knots $\kappa_1 < \cdots < \kappa_d$. A l-th degree spline model for f is given with

$$f(x) = \sum_{q=1}^{m} b_q B_q(x) = \beta_0 + \beta_1 x + \cdots + \beta_l x^l + \sum_{k=1}^{d} \tilde{b}_k (x - \kappa_k)_+^l, \tag{3.15}$$

where $m = l + 1 + d$, $\beta_i = b_{i+1}$ with $i = 0, \ldots, l$, $(x - \kappa_k)_+ = \max\{0, (x - \kappa_k)\}$, and the truncated polynomial basis function of degree l $(x - \kappa_k)_+^l$ -also known as truncated power basis function of degree l- has $l - 1$ continuous derivatives. Therefore, the higher the degree l, the smoother are the spline functions. The coefficients b in (3.15) are denoted with β and $\tilde{b}$. Thus, for all observed values $x_1, \ldots, x_n$ of covariate x (3.15) can be rewritten into the mixed model formulation

$$\boldsymbol{f} = Vb = X\beta + Z\tilde{b}, \tag{3.16}$$

with $\boldsymbol{f} = (f(x_1), \ldots, f(x_n))^T$, $V = [B_1(x_i), \ldots, B_m(x_i)]_{1 \leq i \leq n}$, $b = (b_1 \ldots, b_m)^T$, $X = [1, x_i, \ldots, x_i^l]_{1 \leq i \leq n}$, where x_i is the value of the ith individual of covariate x, $Z = [(x_i - \kappa_1)_+^l, \ldots, (x_i - \kappa_d)_+^l]_{1 \leq i \leq n}$, $\beta = (\beta_0, \ldots, \beta_l)^T$, and $\tilde{b} = (\tilde{b}_1, \ldots, \tilde{b}_d)^T$. This transcription is done for practical reasons which will be seen in Section 3.2.3.

This approach is easy and intuitively understandable, but because of numerical problems when applying this basis for P-spline smoothing, it is not often used. Instead, a B-spline basis is the preferred type of basis due to its numerical stability.

B-Spline Basis

Basing on a set of knots $a = \kappa_0 < \cdots < \kappa_{d+1} = b$, $d - l + 1$ linearly independent B-spline basis functions $B_q^l(x)$ of degree l can be built for $q = 0, \ldots, d - l$. A l-th degree

B-spline basis function $B_q^l(x)$ is built up of $(l+1)$ polynomial pieces of degree l which are $(l-1)$ times continuously differentiable connected at l interior knots. The B-spline basis functions are bounded above. In an interval between $l+2$ neighbouring knots $B_q^l(x) \geq 0$ and in other respects zero. A B-spline basis function overlaps, barring at the boundaries, with $2l$ neighbouring basis functions. The unknown smooth function f can be estimated through a linear combination of $m = l + 1 + d$ B-spline basis functions:

$$f(x) = \sum_{q=1}^{m} b_q B_q^l(x), \tag{3.17}$$

where $B_q^l(x)$ is the q-th B-spline basis function of degree l which is defined recursively (see also de Boor, 1978) through

$$B_q^l(x) = \frac{x - \kappa_q}{\kappa_{q+l} - \kappa_q} B_q^{l-1}(x) + \frac{\kappa_{q+l+1} - x}{\kappa_{q+l+1} - \kappa_{q+1}} B_{q+1}^{l-1}(x), \tag{3.18}$$

where $B_q^0(x) = \mathbb{1}_{[\kappa_q, \kappa_{q+1})}(x)$. To use this recursive definition for the construction of the full B-spline basis, we need beside the $d+2$ knots $\kappa_0 < \cdots < \kappa_{d+1}$ $2l$ further knots. Therefore, we have the knots $\kappa_{-l} < \kappa_{1-l} < \cdots < \kappa_{d+l} < \kappa_{d+1+l}$. At each point $x \in [a,b]$ one has $\sum_{q=1}^{m} B_q(x) = 1$. Another possibility to compute B-spline basis functions is the usage of differences of truncated polynomials, see for further details Eilers and Marx (2010).

In Wand and Ormerod (2008) it is shown that the spline representation $B(x)b$ of the smooth function $f(x)$ can be transformed to

$$B(x)b = \beta_0 + x\beta_1 + \tilde{B}(x)\tilde{b} = X(x)\beta + Z(x)\tilde{b} \tag{3.19}$$

when excerpting the linear slope, see also Kauermann (2010). Thus, $\boldsymbol{f} = X\beta + Z\tilde{b}$, where $X = [1, x_i]_{1 \leq i \leq n}$ and Z is a $n \times m - 2$ matrix. The reason for the transcription into the mixed model formulation will be seen in Section 3.2.3.

Penalisation

To hinder too much variability of the fitted curves, we introduce a penalty -sometimes also called roughness penalty- on the spline coefficients. Using a truncated polynomial basis, the basis consists of two different parts: the first part is composed of a global polynom of degree l with $l+1$ basis functions and the second part exists of d truncated polynomial basis functions of degree l $(x - \kappa_k)_+^l$. The latter are responsible if the fit gets too wiggly. Therefore, we introduce a penalty on the coefficients $\tilde{b}_k$ of the truncated polynomial basis functions. One possibility which is easy to apply is to construct the penalty using the sum of the squared coefficients $\tilde{b}_k$, so that coefficients with a large absolute value get penalized which results in a smoother fit. The penalty takes the form

$$\lambda \sum_{k=1}^{d} \tilde{b}_k^2 = \lambda b^T D b, \tag{3.20}$$

where $\lambda \geq 0$ is the smoothing parameter which controls the strength of the penalisation, $b = (\beta_0, \beta_1, \ldots, \beta_l, \tilde{b}_1, \ldots, \tilde{b}_d)^T$ are the coefficients, and $D = diag(0_{l+1}, 1_d)$ is the penalty matrix. If $\lambda \to 0$, the penalty has only a marginal influence and the penalized log-likelihood equals approximatively the log-likelihood. If $\lambda \to \infty$, the estimation is dominated by the penalty and the resulting fit is a polynom of degree l because the spline coefficients $\tilde{b}_k$ diminish towards zero.

For a B-spline basis, we discuss two different approaches of penalties. One approach is to use the integral of the squared second derivative to construct the penalty

$$\int (f''(x))^2 dx = \sum_{q=1}^{m} \sum_{r=1}^{m} b_q b_r \int B_q''(x) B_r''(x) dx = b^T D b. \tag{3.21}$$

This leads to a in b quadratic penalty $\lambda \int (f''(x))^2 dx = \lambda b^T D b$ with the penalty matrix D whose entries are determined by derivatives of the basis functions. The usage of the second derivative is reasonable, because it serves as a measure for the function's curvature. This penalty got known after the usage in Reinsch (1967) and was used, for instance, in O'Sullivan (1986). Another approach is to use a difference penalty as a simple approximation of the integrated square of the derivatives as it was used, for instance, in Eilers and Marx (1996). The first derivative of a B-spline basis of degree l may be expressed in dependence on the first differences of the corresponding coefficients and the B-spline basis functions of degree $l-1$

$$\frac{\partial}{\partial x} \sum_q b_q B_q^l(x) = l \cdot \sum_q \frac{b_q - b_{q-1}}{\kappa_{q+l} - \kappa_q} B_q^{l-1}(x). \tag{3.22}$$

We now introduce a penalty based on the differences of the corresponding coefficients to reach a smooth functional fit as defined by the first derivative and to avoid too large values of the latter. There is also the possibility of differences of a higher order to achieve smooth functional fits as defined by higher derivatives. For reasons of simplification we use equidistant knots. The differences of the k-th order are denominated with Δ^k and are defined recursively through

$$\Delta^k b_q = \Delta^{k-1} b_q - \Delta^{k-1} b_{q-1} \tag{3.23}$$

with $\Delta^1 b_q = b_q - b_{q-1}$ and $\Delta^2 b_q = \Delta^1\Delta^1 b_q = \Delta^1 b_q - \Delta^1 b_{q-1} = b_q - 2b_{q-1} - b_{q-2}$. The difference penalty is given by

$$\lambda \sum_{q=k+1}^{m} (\Delta^k b_q)^2 = \lambda b^T C_k^T C_k b = \lambda b^T D_k b \tag{3.24}$$

where C_k is a difference matrix of dimension $(q-1) \times d$ and D_k is the penalty matrix.

For the estimation of the coefficients β_z and b in a generalized additive model as given in (3.12) with only one smooth function f, we use the penalized log-likelihood criterion

$$l_{pen}(\beta_z, b) = l(\beta_z, b) - \frac{1}{2}\lambda b^T D b, \tag{3.25}$$

where $l(\beta_z, b)$ is the log-likelihood of a generalized linear model with predictor $\eta = z^T\beta_z + f(x)$ and $\lambda b^T D b$ is a quadratic penalty which may be constructed as described above. The penalized maximum-likelihood estimator may be estimated through a Fisher scoring algorithm which is amended by the penalty.

3.2.3 P-Spline Smoothing of Generalized Additive Models via Generalized Linear Mixed Models

The predictors $\eta_i = z_i^T\beta_z + \sum_{j=1}^{p} f_j(x_{ij})$ of a generalized additive model with responsible variables y_i of the exponential family can be rewritten to

$$\boldsymbol{\eta} = \boldsymbol{X\beta} + \boldsymbol{Z\tilde{b}} \tag{3.26}$$

where $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_n)^T$. The vector $\boldsymbol{\beta}$ represents fixed effects while the vector $\tilde{\boldsymbol{b}}$ contains normally distributed random effects. This results in a generalized linear mixed model (GLMM) with the density

$$f(\boldsymbol{y}|\tilde{\boldsymbol{b}}) = \exp\left(\boldsymbol{y}^T(\boldsymbol{X\beta} + \boldsymbol{Z\tilde{b}}) - \boldsymbol{1}^T b(\boldsymbol{X\beta} + \boldsymbol{Z\tilde{b}}) + \boldsymbol{1}c(\boldsymbol{y})\right), \tag{3.27}$$

assuming that $\boldsymbol{y} = (y_1, \ldots, y_n)^T$, the dispersion parameter ϕ and the weights $\omega_1, \ldots, \omega_n$ are 1, and $\tilde{\boldsymbol{b}} \sim N(\boldsymbol{0}, G)$, where $\boldsymbol{0}$ is a null vector and G a covariance matrix which is defined later on. Due to the fact that the generalized additive model can be transcribed as a generalized linear mixed model, we can use techniques and software for generalized linear mixed models to fit generalized additive models. The estimated parameters $\hat{\boldsymbol{\beta}}$ and $\hat{\tilde{\boldsymbol{b}}}$ maximize the following expression

$$\boldsymbol{y}^T(\boldsymbol{X\beta} + \boldsymbol{Z\tilde{b}}) - \boldsymbol{1}^T b(\boldsymbol{X\beta} + \boldsymbol{Z\tilde{b}}) + \boldsymbol{1}c(\boldsymbol{y}) - \frac{1}{2}\tilde{\boldsymbol{b}}^T\Lambda\tilde{\boldsymbol{b}}, \tag{3.28}$$

where $\Lambda = \sigma_\epsilon^2 Cov(\tilde{\boldsymbol{b}})^{-1}$ with $\sigma_\epsilon^2 = \phi = 1$. Considering a generalized additive model with a truncated polynomial basis in matrix notation, the ith row of matrix $\boldsymbol{X}$ consists of z_i^T and the powers of degree 1 through l $x_{ij}, \dots, x_{ij}^l$ of all j smooth components. The matrix $\boldsymbol{X}$ represents the design matrix of the fixed effects of a generalized linear mixed model. The truncated polynomial basis functions $(x_{ij} - \kappa_k)_+^l$ of all j smooth components compose the ith row of matrix $\boldsymbol{Z}$ which forms the design matrix of the random effects of a generalized linear mixed model. The coefficients β_z of the linear predictors and the polynomial coefficients $\beta_{j1}, \dots, \beta_{jl}$ which form $\boldsymbol{\beta}$ remain unpenalized, while the coefficients $\tilde{\boldsymbol{b}}$ of the truncated polynomial basis functions get penalized. In the strict sense, the predictor η_i of a generalized additive model with a truncated polynomial basis takes the form

$$\begin{aligned}
\eta_i &= z_i^T \beta_z + \sum_{j=1}^{p} f_j(x_{ij}) \\
&= z_i^T \beta_z + \sum_{j=1}^{p} \sum_{q=2}^{m_j} b_{jq} B_{jq}(x_{ij}) \\
&= z_i^T \beta_z + \sum_{j=1}^{p} \left(\beta_{j1} x_{ij} + \dots + \beta_{jl} x_{ij}^l + \sum_{k=1}^{d_j} \tilde{b}_{jk} (x_{ij} - \kappa_k)_+^l \right)
\end{aligned} \tag{3.29}$$

and can be expressed as

$$\boldsymbol{\eta} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\tilde{\boldsymbol{b}}, \tag{3.30}$$

where $\boldsymbol{X} = [z_i^T, \boldsymbol{x}_i^T]_{1 \le i \le n}$, $\boldsymbol{x}_i = (\boldsymbol{x}_{ij}^T,\ j = 1, \dots, p)^T$, $\boldsymbol{x}_{ij} = (x_{ij}, \dots, x_{ij}^l)^T$, $\boldsymbol{Z} = [\underset{1 \le k \le d_1}{(x_{i1} - \kappa_k)_+^l}, \dots, \underset{1 \le k \le d_p}{(x_{ip} - \kappa_k)_+^l}]_{1 \le i \le n}$, $\boldsymbol{\beta} = (\beta_z^T, \beta_{j1}, \dots, \beta_{jl}, j = 1, \dots, p)^T$, $\tilde{\boldsymbol{b}} = (\tilde{b}_{j1}, \dots, \tilde{b}_{jd_j},$ $j = 1, \dots, p)^T$, and d_j as the number of truncated basis functions of f_j. The predictor in (3.30) has the form of a predictor of a generalized linear mixed model with $\tilde{\boldsymbol{b}} \sim N(\mathbf{0}, G)$, where $G = Cov(\tilde{\boldsymbol{b}})$ is a positive definite covariance matrix with structure $G = diag(\tau_1^2 \boldsymbol{I}, \dots, \tau_p^2 \boldsymbol{I})$. The penalized log-likelihood of a generalized linear mixed model can be expressed as

$$l_{pen}(\boldsymbol{\beta}, \tilde{\boldsymbol{b}}) = l(\boldsymbol{\beta}, \tilde{\boldsymbol{b}}) - \frac{1}{2} \tilde{\boldsymbol{b}}^T G^{-1} \tilde{\boldsymbol{b}}, \tag{3.31}$$

where $l(\boldsymbol{\beta}, \tilde{\boldsymbol{b}})$ is defined as log-likelihood for generalized linear models but with the extended predictor given in (3.30). For the maximization of (3.31) score functions with respect to $\boldsymbol{\beta}$ and $\tilde{\boldsymbol{b}}$ have to be constructed analogously to generalized linear models. To determine the penalized maximum-likelihood estimator, the roots of the score functions can be calculated iteratively by a Fisher scoring algorithm.

As it could already be seen in Section 3.2.2, the penalized log-likelihood criterion of a generalized additive model is defined by

$$\begin{aligned} l_{pen}(\beta_z, b_1, \ldots, b_p) &= l(\beta_z, b_1, \ldots, b_p) - \frac{1}{2}\sum_{j=1}^{p} \lambda_j b_j^T D_j b_j \\ &= l(\beta_z, \boldsymbol{b}) - \frac{1}{2}\boldsymbol{b}^T \boldsymbol{D}(\boldsymbol{\lambda})\boldsymbol{b}, \end{aligned} \tag{3.32}$$

with $\boldsymbol{D}(\boldsymbol{\lambda}) = diag(\lambda_j D_j)$ and where now $l(\beta_z, b_1, \ldots, b_p)$ is the log-likelihood of a generalized linear model for the predictor given in (3.29). Each penalty $\lambda_j b_j^T D_j b_j$ for the smooth functions f_j consists of the smoothing parameter λ_j and a penalty matrix D_j. Due to the possibility of rewriting a generalized additive model into a generalized linear mixed model, one can conclude that treating $\tilde{\boldsymbol{b}}$ as random coefficients in a generalized linear mixed model and using $\tilde{\boldsymbol{b}}$ and its covariance matrix G for penalization is equivalent to penalizing $\boldsymbol{b} = (b_j^T, j = 1 \ldots, p)^T$, where $b_j = (b_{j1}, \ldots, b_{jm_j})^T$, with the smoothing parameters $\boldsymbol{\lambda} = (\lambda_j, j = 1, \ldots, p)$ and the penalty matrix $\boldsymbol{D} = diag(D_1, \ldots, D_p)$. The smoothing parameters $\boldsymbol{\lambda}$ can be defined through G^{-1} and the term $\boldsymbol{b}^T \boldsymbol{D} \boldsymbol{b}$ can be transcribed to $\tilde{\boldsymbol{b}}^T \tilde{\boldsymbol{b}}$. Further information concerning the connection between P-spline smoothing and generalized linear mixed models can also be found in Kauermann (2005, 2010) and Eilers and Marx (2010) or in Fahrmeir, Kneib, and Lang (2004) for an Bayesian approach. The explanations above can be found in the reference books of Fahrmeir, Kneib, and Lang (2009), Ruppert, Wand, and Carroll (2003), and Wood (2006) as well as partly in Wand (2003) and Krivobokova (2006).

3.2.4 Link between Functional Hazard Models and Generalized Additive Models for Poisson-distributed Variables

So far it was explained how the unknown non-linear smooth functions f_j of a generalized additive model can be estimated via P-spline smoothing. Now, we link flexible hazard models to generalized additive models, more precise to additive Poisson models, see also Westerheide and Kauermann (2012a) and Kauermann and Westerheide (2012). We take the flexible hazard rate in (3.6) as a starting point, where instead of the smooth functions linear combinations of high dimensional parametric functions in the form of (3.10) are included. We assume that the observations $(t_i, d_i, x_i, s_i, m_i)$ for all n individuals are present, where t_i denotes the duration time, d_i is the censoring variable, and $x_i = (x_{i1}, \ldots, x_{ip})^T$, s_i and m_i are covariates of the ith individual, $i = 1, \ldots, n$. Assuming that all individual are independent, the log-likelihood for the flexible hazard function for the parameter vector $\boldsymbol{\theta} = (b_0^T, b_1^T, b_x^T, b_\gamma^T, b_\delta^T)$ with $b_x^T = (b_j^T, j = 1, \ldots, p)$ equals

$l(\boldsymbol{\theta}) = \sum_{i=1}^{n} l_i(\boldsymbol{\theta})$ where (see Cox and Oakes, 1984)

$$\begin{aligned} l_i(\boldsymbol{\theta}) = d_i \big\{ \boldsymbol{B}_i(t_i)\boldsymbol{b} + B_\gamma(s_i)b_\gamma + B_\delta(m_i)b_\delta \big\} \\ - \int_0^{t_i} \exp\big\{\boldsymbol{B}_i(t)\boldsymbol{b}\big\} \mathrm{d}t \ \exp\big\{B_\gamma(s_i)b_\gamma + B_\delta(m_i)b_\delta\big\} \end{aligned} \tag{3.33}$$

with $\boldsymbol{B}_i(t) = (x_{ij}B_j(t),\, j = 0, \dots, p)$ with $x_{i0} \equiv 1$, $\boldsymbol{b}^T = (b_j^T,\, j = 0, \dots, p)$ and $h(t, x_i, s_i, m_i) = \exp\{\boldsymbol{B}_i(t)\boldsymbol{b} + B_\gamma(s_i)b_\gamma + B_\delta(m_i)b_\delta\}$. The integral in (3.33) cannot be solved analytically, therefore it has to be calculated numerically. This can be done, for example, through a trapezoid or a Simpson approximation. A trapezoid approximation has been used, for instance, in Kauermann (2005) or Cai, Hyndman, and Wand (2002), but here a Simpson approximation which is numerically more accurate (see for example Gautschi, 1997 or Gil, Segura, and Temme, 2007) is applied and also used in the examples of use in Chapters 4 and 5. The Simpson' s rule to approximate an integral $\int_a^b f(x)dx$ of the interval $[a, b]$ is given with (see for instance Freund and Hoppe, 2007)

$$\frac{h}{3}(f_0 + 4f_1 + f_2) \tag{3.34}$$

with $h = (b - a)/2$ and $f_i := f(a + ih)$, $i = 0, 1, 2$. The interval $[0, t_i]$ of the ith individual is now divided into R equidistant subintervals $[T_{r-1}, T_r]$, where $T_0 = 0$ and $T_R = t_i$, $r = 1, \dots, R$. The subintervals may differ between the individuals, i.e. T_r depends on i, but this is omitted due to notational simplicity. When applying the standard Simpson approximation, the integral component in (3.33) can be approximated through the sum of the Simpson approximated subintervals $[T_{r-1}, T_r]$. After some simplifications this leads to

$$\sum_{r=0}^{R} \left[\frac{T_{r+1} - T_{r-1}}{6} \exp\big\{\boldsymbol{B}_i(T_r)\boldsymbol{b}\big\} + \frac{4(T_r - T_{r-1})}{6} \exp\Big\{\boldsymbol{B}_i\left(\frac{T_r + T_{r-1}}{2}\right)\boldsymbol{b}\Big\} \right] \tag{3.35}$$

with $T_{-1} = T_0$ and $T_{R+1} = T_R$. For further simplications of (3.35) the summation index is replaced by $\tilde{T}_j = T_{j/2}$ for j even and $\tilde{T}_j = (T_{(j+1)/2} + T_{(j-1)/2})/2$ for j odd, where $j = 0, \dots, 2R$. In so doing, all supporting points of the Simpson approximation are included in the index j as j odd. Equation (3.35) can now be rewritten to

$$\sum_{j=0}^{2R} \exp\{\boldsymbol{B}_i(\tilde{T}_j)\boldsymbol{b} + o_j\} \tag{3.36}$$

with o_j as so-called offset which is defined through $o_j = \log\big((T_{j/2+1} - T_{j/2-1})/6\big)$ for j even and $o_j = \log\big(4(T_{(j+1)/2} - T_{(j-1)/2})/6\big)$ for j odd. The integral in (3.33) is now replaced with (3.36). This leads to

$$\begin{aligned} l_i(\boldsymbol{\theta}) = d_i \Big\{ & \boldsymbol{B}_i(t_i)\boldsymbol{b} + B_\gamma(s_i)b_\gamma + B_\delta(m_i)b_\delta \Big\} \\ & - \sum_{j=0}^{2R} \exp\big\{\boldsymbol{B}_i(\tilde{T}_j)\boldsymbol{b} + o_j\big\} \ \exp\big\{B_\gamma(s_i)b_\gamma + B_\delta(m_i)b_\delta\Big\} \end{aligned} \tag{3.37}$$

which reveals the structure of a Poisson log-likelihood (see for instance Fahrmeir, Kneib, and Lang, 2009)

$$l(\boldsymbol{\theta}) = \sum_{i=1}^{n}(y_i \log(\lambda_i) - \lambda_i) = \sum_{i=1}^{n}(y_i\eta_i - \exp(\eta_i)) \tag{3.38}$$

with predictor $\eta_i = \log(\lambda_i)$. We define Y_{ij} artificial variables with values $Y_{ij} = 0$ for $j = 0, \ldots, 2R-1$ and $Y_{ij} = d_i$ for $j = 2R$. By neglecting the constant o_j in the first part of the equation, we get with

$$\begin{aligned} l_i(\boldsymbol{\theta}) = \sum_{j=0}^{2R} \Big\{ & Y_{ij}\big\{\boldsymbol{B}_i(t_i)\boldsymbol{b} + B_\gamma(s_i)b_\gamma + B_\delta(m_i)b_\delta\big\} \\ & - \exp\big\{\boldsymbol{B}_i(\tilde{T}_j)\boldsymbol{b} + o_j\big\} \ \exp\big\{B_\gamma(s_i)b_\gamma + B_\delta(m_i)b_\delta\big\}\Big\} \end{aligned} \tag{3.39}$$

the log-likelihood contribution of the log-likelihood $l(\boldsymbol{\theta}) = \sum_i^n l_i(\boldsymbol{\theta})$ for the independent Poisson distributed variables

$$Y_{ij} \sim \text{Poisson}\big(\lambda_{ij} = \exp\{\boldsymbol{B}_i(\tilde{T}_j)\boldsymbol{b} + B_\gamma(s_i)b_\gamma + B_\delta(m_i)b_\delta + o_j\}\big), \tag{3.40}$$

with $i = 1, \ldots, n$, $j = 0, \ldots, 2R$. From this it follows that after some simple data management as described above, the not analytical log-likelihood in (3.33) can be approximated by the standard likelihood resulting from an additive Poisson regression model as given in (3.13) with the independent Poisson distributed variables given in (3.40). Therefore, we only have to restructure the survival data at hand into Poisson distributed data to fit the flexible hazard model with standard software for generalized additive models. The restructuring of the data can easily be done by a small computer programme. A programme code written in the software language **R** to convert the data as described above can be found in the appendix.

To reach smooth functional fits of the unknown functions, the penalized log-likelihood of the additive Poisson regression model with the independent Poisson distributed variables given in (3.40) is used for modelling. It takes the form

$$l_{pen}(\boldsymbol{\beta}, \tilde{\boldsymbol{b}}) = l(\boldsymbol{\beta}, \tilde{\boldsymbol{b}}, \boldsymbol{\lambda}) = \sum_{i=1}^{n} l_i(\boldsymbol{\beta}, \tilde{\boldsymbol{b}}) - \frac{1}{2}\sum_{j=0}^{p}\lambda_j \tilde{b}_j^T \tilde{D}_j \tilde{b}_j - \frac{1}{2}\lambda_\gamma \tilde{b}_\gamma^T \tilde{D}_\gamma \tilde{b}_\gamma - \frac{1}{2}\lambda_\delta \tilde{b}_\delta^T \tilde{D}_\delta \tilde{b}_\delta \tag{3.41}$$

with l_i as log-likelihood contribution for the Poisson variables (3.40). When using e.g. a B-spline basis, we have $\boldsymbol{\beta} = \big((\beta_{0j}, \beta_{1j}), j = 0, \dots, p, (\beta_{0\gamma}, \beta_{1\gamma}), (\beta_{0\delta}, \beta_{1\delta})\big)$, $\tilde{\boldsymbol{b}} = (\tilde{b}_j^T, j = 0, \dots, p, \tilde{b}_\gamma^T, \tilde{b}_\delta^T)$, and an obvious definition for $\boldsymbol{\lambda} = (\lambda_j, j = 0, \dots, p, \lambda_\gamma, \lambda_\delta)$. The penalty matrix $\tilde{D}$ is of full rank and can be chosen to take the form of an indentity matrix, see for instance Kauermann (2010). The penalized log-likelihood in (3.41) can now be fitted with software for generalized additive models via a generalized linear mixed model approach like it is e.g. possible in the procedure `gam()` of the **R** package `mgcv` using a restricted maximum likelihood-based or marginal likelihood-based estimation, see Wood (2011a). This procedure is described in detail in Wood (2006, 2010, 2011b).

4 Flexible Modelling of Duration of Unemployment Using Functional Hazard Models and Penalized Splines: A Case Study Comparing Germany and the UK

The intention of this paper is to demonstrate the flexibility and capacity of penalized spline smoothing as estimation routine for modelling duration time data. We investigate the unemployment behaviour in Germany and the UK between 1995 and 2005 based on data from national panel studies. Functional duration time models are used to investigate the dynamics of covariate effects. The focus of our analysis is on contrasting the two economies. The statistical model being employed is built upon the hazard function, where we allow all covariate effects to vary smoothly with time. As result of the analyses we demonstrate that the most striking difference between the countries is that elderly unemployed in Germany have decreasing chances for re-employment compared to the UK.

4.1 Introduction

4.1.1 Analysis of Duration of Unemployment

Unemployment is a central problem in western economies as declared by the OECD (2009) and the European Commission (2009), see also Blanchard (2006), Ljungqvist and Sargent (1998) or Layard, Nickell, and Jackman (2006). Unemployment generally has two components, first the unemployment rate as macroeconomic index and secondly the duration of unemployment as job market characteristics, see e.g. Eurostat (2009) and Turnbull (1998). We focus here on functional analysis of duration of unemployment using flexible and highly structured statistical models. The paper aims thereby to provide two contributions to the field. First, we demonstrate how to use available software to model and easily fit rather complex duration time models, where effects are allowed to be dynamic, that is time dependent. Secondly, we analyse and compare the duration of unemployment in Germany and the UK. In particular we investigate how individual effects like gender, age, education, and the professional history increase or decrease the chances of re-employment in the two countries and show how these effects change over the length of unemployment.

The statistical model being used for the analysis is built upon the hazard rate or outflow rate respectively. The classic model here is the Cox model, see Cox (1972), but we allow for non-proportional hazards in the style of varying coefficients, see Hastie and Tibshirani (1993). The proportional hazard assumption in the Cox model has been under major investigation and numerous papers suggest extensions and testing procedures respectively. We refer exemplary to Gray (1994), Hess (1994) or Grambsch and Therneau (2003). Estimation of non-proportional hazards has been carried out with smoothing techniques like local likelihood techniques, see e.g. Fan, Gijbels, and King (1997) or Cai and Sun (2003), spline-based approaches, see e.g. Gray (1994) or Kooperberg, Stone, and Troung (1995) or Bayesian techniques, see e.g. Kneib and Fahrmeir (2007). Here we make use of penalized splines to estimate smooth dynamic covariate effects as proposed in Kauermann (2004). The idea of penalized spline smoothing is thereby simple and the method proves to be quite powerful. Originally introduced by Eilers and Marx (1996) the method has become quite fashionable over the last years, see Ruppert, Wand, and Carroll (2009). Instead of fitting a low dimensional parametric model, a high dimensional spline-based model is fitted, and in order to achieve a smooth and numerically stable fit a penalty is imposed on the high dimensional spline coefficients. The routine is implemented in **R** (see R Development Core Team, 2008 and Wood, 2006). The first purpose of this paper is to demonstrate how to make use of available and ready to use software to fit functional duration time models just after some simple data management.

The duration of unemployment is a central topic in economic research. Early references are, for instance, Nickell (1979), Narendranathan, Nickell, and Stern (1985) or Jackman and Layard (1991). A central data source when analysing duration of unemployment are national data panel collections which allow to employ hazard models to investigate the effect of individual and household specific covariates respectively. We refer exemplary to Hunt (1995) and Steiner (1997, 2001) using data from the German Socio-Economic Panel, Böheim and Taylor (2000) using data from the British Household Panel Survey and Bover, Arellano, and Bentolila (2002) using data from a rotating panel of the Spanish Labour Force Survey. The comparison of countries with respect to duration of unemployment behaviour has been pursued, for instance, by Kaiser and Siedler (2001) and Tatsiramos (2006). They made use of European data sets, more precisely they analysed data from the European Panel Analysis Group (EPAG) and the European Community Household Panel (ECHP) respectively. Dynamic changes in unemployment behaviour over calendar time have been studied, for instance, in Hunt (1995) and Steiner (1997, 2001). None of the cited papers investigate explicitly if and how the covariate effects vary over duration of unemployment. This is the contribution of our paper by allowing all covariate effects to be functional terms to capture time dynamics. To do so, we examine individual effects like gender, age, education, and the professional history as well as seasonal and calendar effects. As result, we contribute to the economic discussion by comparing unemployment behaviour in Germany and the UK using non-proportional hazard models. As will be seen the fundamental difference between the UK and Germany lies in the decreased chances of re-employment for elderly unemployed ($>$ 55 years) in Germany compared to the UK.

4.1.2 Data Base

We make use of data taken from the German Socio-Economic Panel (GSOEP) and the British Household Panel Survey (BHPS), see Haisken-DeNew and Frick (2005) and Taylor, Brice, Buck, and Prentice-Lane (2009). These two panels allow to empirically explore and investigate the effect of different economic and political situations in both countries on the effect of duration of unemployment. We analyse the duration of unemployment of 870 individuals from the GSOEP and 951 from the BHPS respectively. Panel data provide an informative source to assess individual effects on the re-employment probability. Our analysis is based on full-time employees becoming unemployed between January 1995 and April 2005 regardless of whether they are entitled or not to receive unemployment benefits. In fact, the latter information is only described imprecisely in the two databases and along with it is not comparable. We censor the maximum duration of unemployment at 36 months to restrict the analyis to short-term and medium-term unemployment. As event we define full-time re-employment. For individuals with more than one unemployment spell in the database, we randomly select one spell which maintains independence among the observations. For Germany we look at data from the former West Germany (including West-Berlin) only, and for both countries we include only unemployed with domestic nationality, i.e. Germans for the German database and UK citizens for the UK panel. As covariates we include information about gender, age, education, household matters, and former job history. The starting date of the unemployment spell and a seasonal component (month at the beginning of unemployment) are included as well to capture calendar time effects.

The paper is organised as follows. Section 2 introduces penalized spline smoothing and considers unobserved heterogeneity. Section 3 gives an overview about labour market policies, the construction of comparable covariates and shows some rough exploratory analysis based on Kaplan-Meier curves for both countries. Section 4 gives the data analysis before we conclude in Section 5.

4.2 Functional Hazard Model and Estimation

4.2.1 The Model

Let t denote the duration of unemployment for the ith individual and denote with $x_i = (x_{i1}, \ldots, x_{ip})$ the vector of the p exogenous variables under investigation. We restrict the analysis here to time constant effects, taking the value of x_{ij} at the point in time of entry in the unemployment spell. With s_i we denote the calendar time of the starting point of the unemployment spell and we write m_i for the season (month), where m_i ranges from 1 for January to 12 for December. We assume the flexible hazard function

$$h(t, x_i, s_i, m_i) = \underbrace{\exp\{\beta_0(t)\}}_{1} \underbrace{\exp\{\sum_{j=1}^{p} x_{ij}\beta_j(t)\}}_{2} \underbrace{\exp\{\gamma(s_i) + \delta(m_i)\}}_{3}, \tag{4.1}$$

where the model is assembled from three functional effect components. With $h_0(t) = \exp\{\beta_0(t)\}$ we denote the baseline hazard which builds the first component in (4.1). The baseline is modified by functional covariate effects, where $\beta_j(t)$ denotes smooth functions which vary in duration time t. Note that if all $\beta_j(t)$ are constant, i.e. $\beta_j(t) \equiv \beta_j$, the first two components in (4.1) mirror a classical Cox (1972) proportional hazard model. Hence the functional shape in the second component in (4.1) extends the Cox model by incorporating non-proportional hazard behaviour. The third component in (4.1) accounts for calendar effects in that the hazard is allowed to vary over time and season. The calendar effect $\gamma(s)$ is thereby a smooth function which mirrors economic changes in a country while $\delta(m)$ takes intra-annual variation into account. Apparently model (4.1) needs some constraints to be identifiable. We therefore assume that both $\gamma()$ and $\delta()$ integrate out to zero. Moreover, $\delta(m)$ is a periodic function which means that $\delta(12)$ smoothly connects to $\delta(1)$.

The hazard (4.1) is further modified to incorporate unobserved individual heterogeneity. This seems necessary since the length of unemployment is influenced by a number of unrecorded covariates so that the fitted model based on the available covariates can not fully explain individual behaviour. We therefore model the hazard function for the ith individual as

$$h_i(t, x_i, s_i, m_i) = h(t, x_i, s_i, m_i)v_i,$$

where v_i is random and unobserved and independent of the covariates with $E(v_i) = 1$ for identifiability reasons.

4.2.2 Estimation

We start the discussion by describing the estimation approach of the smooth, functional components in (4.1). We first represent each unknown function as a linear combination of thin plate spline basis terms (Wahba, 1990, pp. 30-34), with the popular cubic smoothing spline basis resulting as special case, see Wood (2006). We use the same approach for the seasonal effect $\delta(.)$ in (4.1), but with periodicity enforced on the basis terms. The functional components in (4.1) are therefore replaced by

$$\begin{aligned} \beta_0(t) &= B_0(t)b_0, \quad \beta_j(t) = B_j(t)b_j \\ \gamma(s) &= B_\gamma(s)b_\gamma, \quad \delta(m) = B_\delta(m)b_\delta \end{aligned} \tag{4.2}$$

with $B(.)$ as cubic smoothing splines. Classical spline smoothing is built on knots placed at the (unique) observed values of the variable one smoothes over. To reduce the computational burden we follow Hastie (1996) and Wood (2003) and employ so-called 'low rank smoothing'. For each function this involves to work with a reduced set of knots which is still large and enough to capture the functional shape but small enough to guarantee feasible computation. The idea has been coined by Eilers and Marx (1996) as penalized spline smoothing, see also Ruppert, Wand, and Carroll (2003). Ruppert, Wand, and Carroll (2009) provide an extensive survey of recent results and papers in this field demonstrating the popularity of the approach. Denoting with k the number of knots we follow Wood (2006, p. 161) and set $k = 20$ for the baseline, $k = 10$ for the covariate effects and $k = 4$ for the calendar functions. We fitted the model for larger values of k as well but observed the established fact that the choice of k has little influence on the fit; see Ruppert (2002) for a discussion.

Assume now that $(t_i, d_i, x_i, s_i, m_i)$ denote the observations for the ith individual, $i = 1, \ldots, n$. With d_i we denote the usual censoring variable indicating whether t_i is the observed duration time ($d_i = 1$) or a censored version ($d_i = 0$). Assuming for the moment that the unobserved individual effects v_i are fixed (and known) the log-likelihood for parameter vector $\boldsymbol{\theta} = (b_0^T, b_x^T, b_\gamma^T, b_\delta^T)^T$ with $b_x^T = (b_j^T,\ j = 1, \ldots, p)$ equals $l(\boldsymbol{\theta} \mid v) = \sum_{i=1}^n l_i(\boldsymbol{\theta} \mid v_i)$ where (see Cox and Oakes, 1984)

$$\begin{aligned} l_i(\boldsymbol{\theta} \mid v_i) = & d_i \big\{ \boldsymbol{B}_i(t_i)\boldsymbol{b} + B_\gamma(s_i)b_\gamma + B_\delta(m_i)b_\delta + \log(v_i) \big\} \\ & - v_i \int_0^{t_i} \exp\big\{\boldsymbol{B}_i(t)\boldsymbol{b}\big\} \mathrm{d}t \ \exp\big\{B_\gamma(s_i)b_\gamma + B_\delta(m_i)b_\delta\big\} \end{aligned} \tag{4.3}$$

with $\boldsymbol{B}_i(t) = (x_{ij}B_j(t),\ j = 0, \ldots, p)$ with $x_{i0} \equiv 1$ and $\boldsymbol{b}^T = (b_j^T,\ j = 0, \ldots, p)$. Note that (4.3) is not analytical due to the integral which requires to be calculated numerically. Cai, Hyndman, and Wand (2002) and Kauermann (2004) make use of trapezoid approximation. Numerically more accurate, however, is a Simpson approximation, see e.g. Gautschi (1997) or Gil, Segura, and Temme (2007), which results as follows. For the ith individual we divide the interval $[0, t_i]$ into R equidistant subintervals $[T_{r-1}, T_r]$, say, with $r = 1, \ldots, R$ where $T_0 = 0$ and $T_R = t_i$. Note that the subintervals are different for each individual, i.e. T_r depends on i, which however is omitted for the sake of notational simplicity. Then applying the standard Simpson approximation we can approximate the integral component in (4.3) through

$$\sum_{r=0}^{R} \left[\frac{T_{r+1} - T_{r-1}}{6} \exp\big\{\boldsymbol{B}_i(T_r)\boldsymbol{b}\big\} + \frac{4(T_r - T_{r-1})}{6} \exp\Big\{\boldsymbol{B}_i\left(\frac{T_r + T_{r-1}}{2}\right)\boldsymbol{b}\Big\} \right] \tag{4.4}$$

with $T_{-1} = T_0$ and $T_{R+1} = T_R$. We substitute the summation index in (4.4) by defining $\tilde{T}_j = T_{j/2}$ for j even and $\tilde{T}_j = (T_{(j+1)/2} + T_{(j-1)/2})/2$ for j odd, where $j = 0, \ldots, 2R$. This allows to rewrite (4.4) to

$$\sum_{j=0}^{2R} \exp(\boldsymbol{B}_i(\tilde{T}_j)\boldsymbol{b} + o_j) \tag{4.5}$$

where o_j is defined as so-called offset through $o_j = \log\big((T_{j/2+1} - T_{j/2-1})/6\big)$ for j even and $o_j = \log\big(4(T_{(j+1)/2} - T_{(j-1)/2})/6\big)$ for j odd. Replacing the integral in (4.3) with (4.5) reveals the structure of a Poisson likelihood. In fact defining with Y_{ij} the artificial variables with values $Y_{ij} = 0$ for $j = 0, \ldots, 2R-1$ and $Y_{ij} = d_i$ for $j = 2R$ we get with (4.3) the log-likelihood for the independent Poisson distributed variables

$$Y_{ij} \mid v_i \sim \mathrm{Poisson}\big(\lambda_{ij} = \exp\{\boldsymbol{B}_i(\tilde{T}_j)\boldsymbol{b} + B_\gamma(s_i)b_\gamma + B_\delta(m_i)b_\delta + o_j\}v_i\big), \tag{4.6}$$

with $i = 1, \ldots, n$, $j = 0, \ldots, 2R$. In other words, after some simple data management we can approximate the non-analytic likelihood (4.3) by the standard likelihood resulting from a Poisson regression model (4.6).

The next step is to impose a penalty on the spline coefficients to achieve smooth functional fits. Apparently, the model is high dimensional which implies that the Maximum Likelihood estimate based on the log-likelihood for (4.6) will yield wiggled fitted curves. Following Eilers and Marx (1996) and Ruppert, Wand, and Carroll (2003) we therefore impose a penalty on the coefficients. As demonstrated in Wand and Ormerod (2008) using a singular value decomposition of the entire basis matrix we can rewrite the spline representation in (4.2) by extracting the linear slope yielding

$$\beta_j(t) = B_j(t)b_j = \beta_{0j} + t\beta_{1j} + \tilde{B}_j(t)\tilde{b}_j$$

where $\tilde{B}_j$ is now the reduced rank basis with the linear slope extracted by orthogonal projection, $j = 0, \ldots, p$. Similarly we obtain for $\gamma(s)$ and $\delta(m)$ reduced basis matrices $\tilde{B}_\gamma(s)$ and $\tilde{B}_\delta(m)$ respectively. We now impose a quadratic penalty on the spline coefficient in the form $\lambda_j \tilde{b}_j^T \tilde{D}_j \tilde{b}_j$. It can be shown that this is equivalent to penalizing squared second order derivatives of the function, (see O'Sullivan, 1986 or Wahba, 1990), or second (or higher) order differences of the spline coefficients b_j, (see Eilers and Marx, 1996). We use the implemented version in the `mgcv` package in **R** (see the end of this section) which penalizes the squared second order derivatives. Note that the second order derivative is easily calculated by differentiating the basis $B_j(t)$, say. The extact form of the penalty matrix can be found e.g. in Wood (2006). The parameter λ_j plays thereby the role of a smoothing parameter with $\lambda_j \to \infty$ leading to a linear fit. The complete penalized likelihood takes then the form

$$l(\boldsymbol{\beta}, \tilde{\boldsymbol{b}}, \boldsymbol{\lambda} \mid v) = \sum_{i=1}^{n} \tilde{l}_i(\boldsymbol{\beta}, \tilde{\boldsymbol{b}} \mid v_i) - \frac{1}{2}\sum_{j=0}^{p} \lambda_j \tilde{b}_j^T \tilde{D}_j \tilde{b}_j - \frac{1}{2}\lambda_\gamma \tilde{b}_\gamma^T \tilde{D}_\gamma \tilde{b}_\gamma - \frac{1}{2}\lambda_\delta \tilde{b}_\delta^T \tilde{D}_\delta \tilde{b}_\delta \tag{4.7}$$

with $\tilde{l}_i$ as log likelihood for the Poisson variables (4.6), $\boldsymbol{\beta} = \big((\beta_{0j}, \beta_{1j}),\ j = 0, \ldots, p;\ (\beta_{0\gamma}, \beta_{1\gamma}), (\beta_{0\delta}, \beta_{1\delta})\big)^T$, analogous definition for $\tilde{\boldsymbol{b}}$ and obvious definition for $\boldsymbol{\lambda} = (\lambda_j, j = 0, \ldots, p, \lambda_\gamma, \lambda_\delta)^T$.

Finally, inference can be drawn following standard asymptotic arguments as outlined in Ruppert, Wand, and Carroll (2003) or Wood (2006), see also Kauermann, Krivobokova, and Fahrmeir (2009). In fact, let $\boldsymbol{\theta} = (\boldsymbol{\beta}^T, \tilde{\boldsymbol{b}}^T)^T$ denote the complete parameter vector we define with $F(\boldsymbol{\theta}, \boldsymbol{\lambda})$ the Fisher matrix $-E(\partial^2 l(\boldsymbol{\theta}, \boldsymbol{\lambda} \mid v)/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T)$. It can generally be shown (Wood, 2006) that

$$\mathrm{Var}(\hat{\boldsymbol{\theta}} \mid v) = F^{-1}(\boldsymbol{\theta}, \boldsymbol{\lambda}) F(\boldsymbol{\theta}, \boldsymbol{\lambda} = \mathbf{0}) F^{-1}(\boldsymbol{\theta}, \boldsymbol{\lambda}).$$

Note that (4.7) can be easily fitted with software for generalized additive models, see Hastie and Tibshirani (1990), where we use the `gam()` procedure in **R** (package mgcv) which is extensively described in Wood (2006). In fact, the only thing which is numerically necessary in order to fit the model with available software is to restructure the data to obtain the Poisson model (4.6). The required code is available on request from the authors. The computational effort needed to fit the model with the present data corresponds to state-of-the-art computer technology and took about an hour for each database, including the unobserved heterogeneity discussed in the next section.

4.2.3 Unobserved Heterogeneity

It is a common feature that hazard rates or outflow rates decrease with duration of unemployment. A plausible cause for this is that long-term unemployed in general have worse chances of getting reemployed. This is also called ‚true duration dependence‘ or ‚negative duration dependence‘ respectively and discussed, for instance, in Machin and Manning (1999) or Steiner (2001). Another explanation for a decreasing hazard is that different individuals have different unobservable hazard rates in the style of unobserved heterogeneity. This implies that, individuals with low hazard rates are over-represented in the group of long-term unemployed which leads to negative duration dependence, see Machin and Manning (1999). For models ignoring such unobserved heterogeneity, the estimated duration dependence of the hazard function is generally smaller (and hence biased) than for models including unobserved heterogeneity, see Van den Berg (2001). Following the above arguments, we therefore account for unobserved population heterogeneity in the model by including an unobserved latent effect v_i for each individual. There are two common strategies to model the unobserved individual effect v_i namely either using a discrete mixture or alternatively a Gamma distribution, see e.g. Heckman and Singer (1984) or Lancaster (1990). As Abbring and Van den Berg (2007) point out, there is no argument for prefering the one or the other and preference can therefore be based on numerical feasibility or convergence statements. We feel attracted by the Gamma distribution approach here since the Poisson approximation suggests to use the Gamma distribution as conjugate prior for v_i so that the likelihood remains analytically trackable. We assume that

$$v_i \sim \text{Gamma}\left(\frac{1}{\alpha}, \alpha\right) = \frac{1}{(\frac{1}{\alpha})^\alpha \Gamma(\alpha)} v_i^{\alpha-1} \exp(-\alpha\, v_i)$$

such that $E(v_i) = 1$ and $\text{Var}(v_i) = 1/\alpha$. Parameter α can be easily estimated following straight forwardly the EM algorithm outlined in Klein (1992). In practice this means that one fits the model (4.6) for fixed v_i, $i = 1, \ldots, n$ and maximises the resulting likelihood with respect to α.

4.3 Data Description

We will subsequently entertain the above model to analyse and compare German and UK unemployment data. Before we start, however, we give a short overview of the different labour market policies and benefit schemes in Germany and the UK for the observed period of time.

4.3.1 Differences in Labour Market Policies

In Germany dismissed employees receive 60% (or 67% in case of children living in the family) of their former salary during the first year of unemployment. Elderly unemployed can expect a longer period of support depending on their age and previous working time. From the second half of the 1990s onwards the maximum length was 32 months for unemployed aged 57 and older. For unemployed younger than 45 years the maximum duration was 12 months. After this period further but reduced support (53% respectively 57% in case of children living in the family) could be provided, depending on the financial situation of the unemployed and his/her family. In December 2004 the system was ammended, but this is only at the margin of the range of our database. In the data period it was possible for individuals in the age group between 55 and 64 years to retire early after a previous period of unemployment of at least 52 weeks. Depending on their date of birth, they could access early retirement benefits from an age of 60 to 63 years, see Reinhardt (2006). We will see that in particular for elderly unemployed the length of unemployment is prolonged compared to the UK. In the UK the contribution-based Jobseeker's allowance is a fixed support provided to the unemployed independent of the height of his or her previous salary. This benefit is provided for six months. The fixed and more restricted support provided afterwards, the income-based Jobseeker's allowance, is comparable to the German system and is only eligible to households which depend on social welfare. In the UK there exist no early pension like in Germany. For further information about both countries labour market policies and benefit schemes, see Clasen (2005) and European Commission (2005b).

The benefit schemes of Germany and the UK have been amended and modified several times in the period between 1995-2005. None of these modifications were fundamental in a way that the structure of the support has completely changed. However, we have a calendar effect in our model to compensate for changes over time, see also Steiner (2001) or Hunt (1995). Nonetheless the unemployment support schemes of both countries differ in a number of ways and such differences are likely to lead to different behaviour in the duration of unemployment, see also Røed and Zhang (2003), Tatsiramos (2006) or Heer (2006). This will be the focus of our analysis.

4.3.2 Covariates and Kaplan-Meier Curves

The two data panels provide different information which requires to make covariate comparable. Before starting our comparative study we give a short review of the construction of the comparable covariates and the corresponding Kaplan-Meier curves. The Kaplan-Meier curves of the data to which we refer are shown in Figure 4.1.

The overall survivor curve for the German data decreases less than the curve of the UK data. The probability that an individual remains in the state of unemployment after 6 months is approximately 60% in Germany while it is 55% in the UK. The difference increases with duration of unemployment. Concerning gender, the Kaplan-Meier curves of women are always above the Kaplan-Meier curves of men, meaning that in both countries males have higher chances to return to full-time employment.

The individuals are arranged in three age groups: between 25 and 44 years (taken as reference), between 45 and 54, and from 55 to 64 years. The Kaplan-Meier curves for the different age groups exhibit a substantially different pattern in both countries. The UK data show for the first two years a similar pattern in all age groups and afterwards they differ barely. In Germany unemployed aged 55 to 64 have worst chances to get reemployed among the three groups.

Considering education level, we need to bring this covariate on a comparable scale in both countries. Though an university degree might be comparable in both countries, vocational training, like in Germany, does not have an unique comparable analogon in the UK. Our comparison is therefore restricted to three groups following the ISCED-97 classification, see UNESCO (2003) or OECD (1999). We group the unemployed with ISCED-97 levels 5 and 6 (higher education) as one group, individuals with ISCED-97 levels 3 and 4 as second group (intermediate level), taken as reference, and the third group is composed of unemployed with ISCED-97 levels 0-2 (lower education). This information is available for the GSOEP data. To classify the educational achievement of the BHPS data we use information out of an ISCED level generating algorithm created by Malcom Brynin (Institute for Social and Economic Research, UK) and provided by John Brice (Institute for Social and Economic Research, UK). Looking at the educational differences shown in the Kaplan-Meier curves, it becomes clear that the better the education, the higher the probability leaving the state of unemployment.

Information about the professional status of the last job is given by the Goldthorpe Category of the former job, see Goldthorpe (1987) or Gazeboom and Treiman (2003).

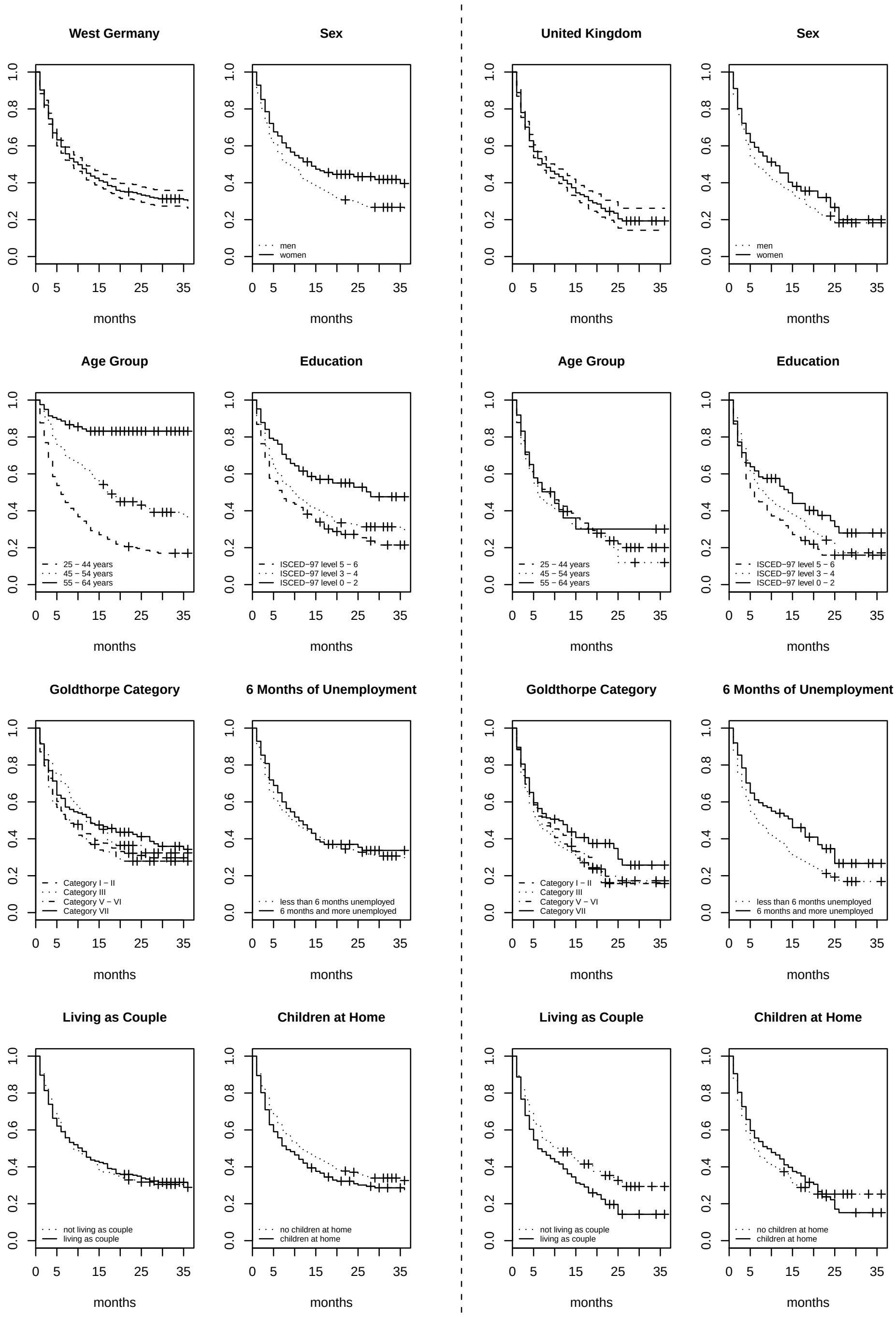

Figure 4.1: Kaplan-Meier curves for German GSOEP (first and second column) and for UK BHPS data (third and fourth column).

We summarise the available Goldthorpe Categories of both data sets in four categories: the first group combines Goldthorpe Category I and II (higher and lower managerial and professional workers), the second group is Goldthorpe Category III (routine clerical, service and sales work respectively), taken as reference, furthermore we group Goldthorpe Category V and VI (manual supervisors and skilled manual workers), and the fourth group is Goldthorpe Category VII (semi- and unskilled manual workers and agricultural labour respectively). Individuals with a former professional status of Goldthorpe Category IV (small self-employed with or without employees, self-employed farmers) are not taken into account. Looking at the Kaplan-Meier curves of the four different Goldthorpe Categories, the curves show only small differences in the two countries. More information is available by looking at the job history given here as binary covariate denoting whether the individual was at least six months unemployed in the last three years before the recent start of unemployment. Looking at the Kaplan-Meier curves for this covariate, we find only weak differences for the German data whereas in the UK data there is clearly an effect. Individuals with less than six months of unemployment show better re-employment probabilities than individuals with a duration of six or more months of unemployment in the past three years.

Two further binary variables give information about household matters. The first variable considers if the individual lives as a couple or not. This variable is only related to the status of partnership of the individual and takes not into account if there are children in the household. The latter is considered separately by the second variable which informs about children in the household or not. While in Germany it makes nearly no differences whether the individual is living as a couple or not, in the UK the Kaplan-Meier curve for individuals living as a couple decreases stronger over the observed 36 months than the curve of single individuals. In Germany the fact that there are children at home has a small positive effect on the decrease of the Kaplan-Meier curve. In the UK the trend of the curve is opposite for the first 2 years. Afterwards the chance of leaving unemployment for individuals with children at home is better.

For further information about the data we refer to Tables 4.1 and 4.2. Subsequently we will model these data using non-proportional hazard effects.

4.4 Data Analysis

In Figures 4.2 and 4.3 we show the resulting fit of model (4.1) for Germany and the UK in comparison. The left two columns show the effects for Germany, the right two columns present the fitted effects for the UK. The plots show the fitted effects and corresponding confidence intervals. As dotdashed horizontal line we also include simple non-dynamic parametric effects based on a proportional hazard assumption, that is, we fitted model (4.1) assuming $\beta_j(t) \equiv \beta_j$ for $j = 1, \ldots, p$. The resulting estimates are listed in Tables

German Socio-Economic Panel	Event	No Event	Σ	$\hat{\beta}_j$	Std. Error
Men	321	226	547	reference	-
Women	135	188	323	-0.40115	0.10944
Age 25-44	365	218	583	reference	-
Age 45-54	73	93	166	-1.29825	0.13312
Age 55-64	18	103	121	-3.06343	0.24974
ISCED-97 Level 5-6	128	92	220	0.44984	0.12172
ISCED-97 Level 3-4	282	243	525	reference	-
ISCED-97 Level 0-2	46	79	125	-0.89878	0.16372
Goldthorpe Category I-II	154	136	290	0.12650	0.15036
Goldthorpe Category III	78	86	164	reference	-
Goldthorpe Category V-VI	114	80	194	0.48531	0.15847
Goldthorpe Category VII	110	112	222	0.06408	0.15707
6 Months and more Unemployed	83	85	168	-0.34120	0.12429
Less than 6 Months Unemployed	373	329	702	reference	-
Living as Couple	330	309	639	0.06302	0.11218
Living as Single	126	105	231	reference	-
Children at Home	238	189	427	0.21265	0.10116
No Children at Home	218	225	443	reference	-

Table 4.1: Distribution of covariates, estimated non-dynamic parametric effects $\hat{\beta}_j$ and standard errors for German GSOEP data (870 individuals).

4.1 and 4.2. Note that the model can be seen as benchmark and apparently the dynamic behaviour of the fitted curves $\hat{\beta}_j(t)$ indicate the superiority of including dynamics instead of relying on proportional hazards. The variance of the individual effect v_i was estimated with 1.826 (α=0.5475186) for the German data and 1.812 (α=0.5519524) for the UK data, respectively.

Baseline and Gender

We first look at the baseline effect $\hat{\beta}_0(t)$ in Figure 4.2 which mirrors the reference categories. The baseline of the German data has altogether an increasing positive effect for the first 19 months of unemployment, so that the chance of getting reemployed in Germany increases continuously. This positive trend possesses local peaks at regular quarterly intervals. After a duration of 19 months the curve stabilises and further peaks are within the confidence band width. Such positive duration dependence for men was also observed in other German panel studies, see e.g. Steiner (2001) and for a survey Machin and Manning (1999). The baseline of the UK data has also a positive effect which is however less strong for the first two years compared to Germany. Within this

British Household Panel Survey	Event	No Event	Σ	$\hat{\beta}_j$	Std. Error
Men	333	294	627	reference	-
Women	134	190	324	-0.42038	0.11174
Age 25-44	329	352	681	reference	-
Age 45-54	106	90	196	0.08243	0.11252
Age 55-64	32	42	74	-0.31782	0.18825
ISCED-97 Level 5-6	219	198	417	0.35840	0.10957
ISCED-97 Level 3-4	161	180	341	reference	-
ISCED-97 Level 0-2	87	106	193	-0.01107	0.13820
Goldthorpe Category I-II	130	143	273	-0.37256	0.13895
Goldthorpe Category III	98	96	194	reference	-
Goldthorpe Category V-VI	103	83	186	-0.29144	0.15306
Goldthorpe Category VII	136	162	298	-0.35054	0.14261
6 Months and more Unemployed	87	100	187	-0.53329	0.12748
Less than 6 Months Unemployed	380	384	764	reference	-
Living as Couple	356	342	698	0.48801	0.11346
Living as Single	111	142	253	reference	-
Children at Home	262	284	546	-0.41420	0.09879
No Children at Home	205	200	405	reference	-

Table 4.2: Distribution of covariates, estimated non-dynamic parametric effects $\hat{\beta}_j$ and standard errors for UK BHPS data (951 individuals).

positive effect there are again some local peaks, the most remarkable around two, five, and fourteen months. Afterwards the effect is not estimated in significant order. Panel studies from the UK mostly have negative duration dependence, see for an overview Machin and Manning (1999). In our analysis there is however no evidence for a negative duration dependence characterising the duration of unemployment in the baseline.

Concerning gender, the effect in both countries is different and changes with duration of unemployment. In Germany the effect for females is negative and becomes stronger over the length of unemployment. In the UK there is a negative effect in the first 18 months which fades away thereafter. Generally, it appears that in Germany it is more difficult for women to find a job after a period of unemployment. Other analyses of German data came to a similar result for women compared to men, see Steiner (2001) and for a survey Machin and Manning (1999).

Human Capital

Most striking in our analysis is how the effect of age differs in the two countries. In the UK the age group 45-54 years has a slightly positive effect between 5 and 20 months.

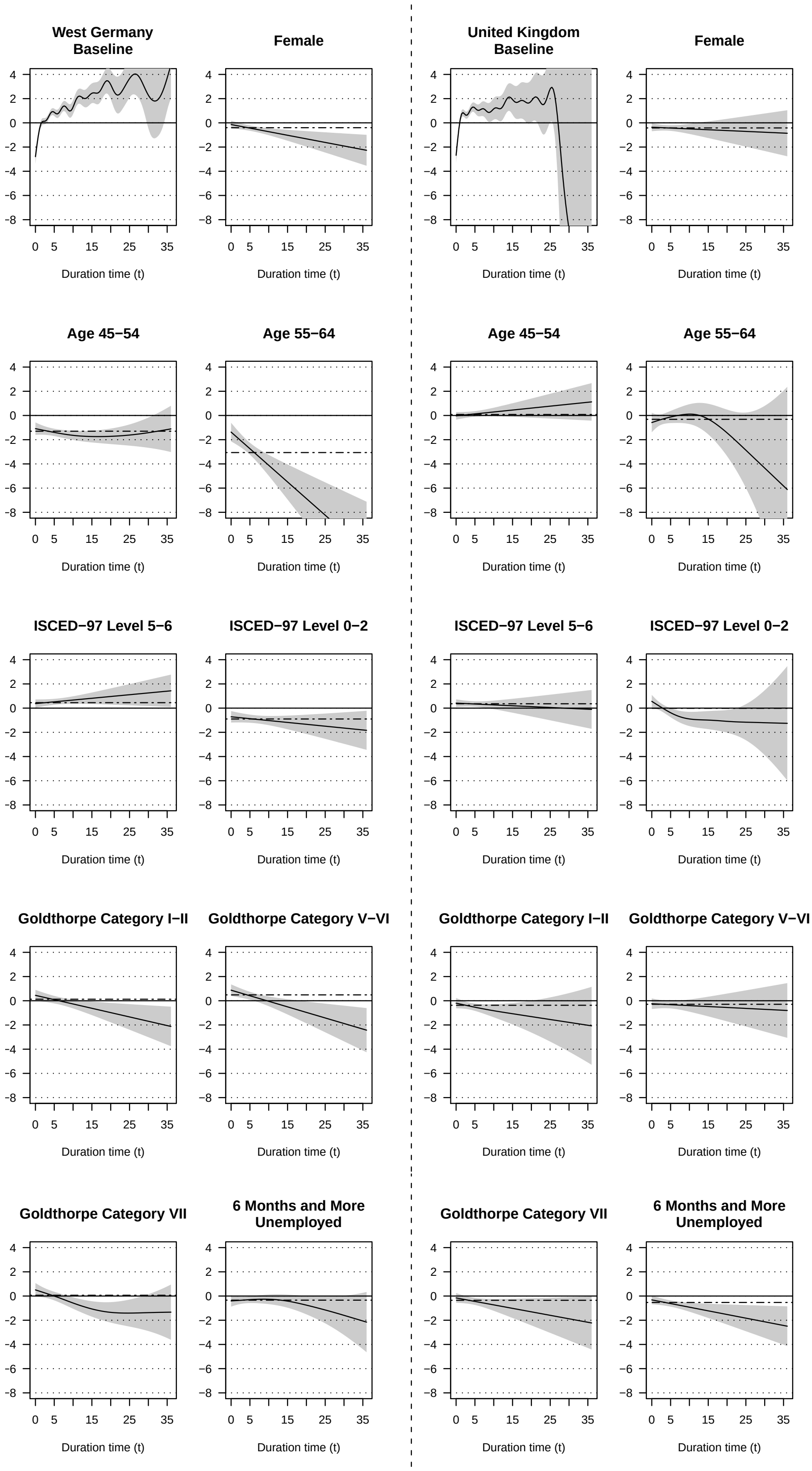

Figure 4.2: Fitted dynamic effects with confidence intervals for duration time of unemployment (in months) for German GSOEP (first and second column) and UK BHPS data (third and fourth column). Fitted parametric effects (dotdashed lines) are added.

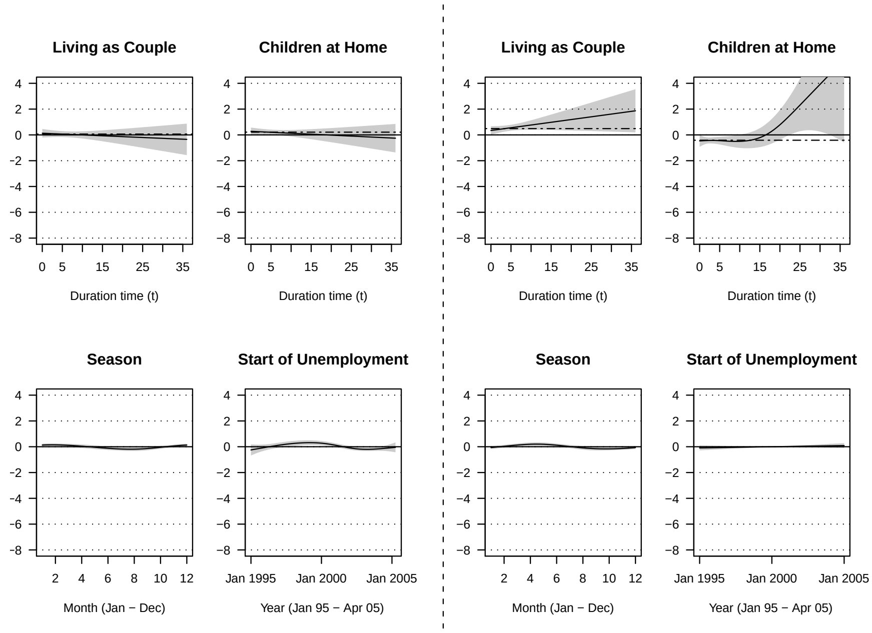

Figure 4.3: Fitted dynamic effects with confidence intervals for duration time of unemployment (in months), months (January - December) and year (January 1995 - April 2005) for German GSOEP (first and second column) and UK BHPS data (third and fourth column). Fitted parametric effects (dotdashed lines) are added.

The remaining months have no significant effect. In contrast, in Germany this age group shows a negative effect for the first 32 months of unemployment but fades away thereafter. In the UK the age group between 55 and 64 has a slightly negative effect for the first five months which fade away afterwards. However, more importantly is the effect for elderly unemployed in Germany. Individuals between 55-64 show a strong, decreasing negative effect which differs significantly compared to the UK. This is one of the central comparative findings in our analyses. The results mirror findings of previous studies, see for example Hunt (1995) and Hujer and Schneider (1995) for German data. One reason for the decreasing chance of elderly unemployed Germans to get reemployed might be the longer duration of unemployment benefits as discussed in Hunt (1995) and the possibility for older employees to retire early by bridging the time gap between employment and retirement with unemployment benefits, see Rein and Jacobs (1993) and Knuth and Kalina (2002). Following Fitzenberger and Wilke (2004) this option is the reason for longer unemployment durations among elderly unemployed in Germany who did not search for a job but used this way to end their working life or whose companies used this subsidised path to discard their old employees. Therefore the negative effect might contribute to the fact that unemployment is used as a fluent passage to early retirement.

Looking at the educational differences, it becomes clear that in both countries higher educated individuals have better chances to get reemployed. In Germany individuals with ISCED-97 level 5 or 6 have a continuous positive effect and therefore better chances to find a job compared to individuals with ISCED-97 level 3 or 4. In the UK this effect is positive for the first years, but fades away thereafter. The effect for less educated individuals (ISCED-97 level 0-2) is constantly negative for the German data. In the UK this effect is positive for the first two months, followed by a negative effect for a duration up to two years and fades away afterwards. Looking at UK data, Begum (2004) came to similar results in her analysis: individuals with no qualification had highest likelihood for having a longer length of unemployment among the reviewed groups. Lauer (2003) drew the conclusion that higher educated individuals have better probabilities of getting reemployed. In her survey, German individuals with a tertiary education level had the best chances to leave unemployment and take up employment again. Overall, education seems to have a general positive effect on reentering the job market, regardless which country you live in, quod vide Eurostat (2007).

Professional History

In Germany individuals of Goldthorpe Categories I-II and V-VI have a similar effect compared to Goldthorpe Category III (reference). For the first six months the effect is positive before it turns negative. Individuals of Goldthorpe Category VII have a positive effect for the first six months which is followed by a negative effect as well, but after two years this effect fades away. In the UK the individuals of Goldthorpe Categories I-II have a negative effect which gets to zero after two years. In Goldthorpe Categories V-VI the effect is negative for the first year and disappears thereafter. Individuals of Goldthorpe Category VII (semi- and unskilled manual or agricultural workers) have a constant negative effect. These observed effects go along with the empirical findings in Collier (2005) where qualified workers have shorter durations and unskilled workers can look towards longer durations of unemployment.

Being at least six months unemployed in the three years before this spell of unemployment has a negative effect on the re-employment chances in both countries. This effect is even stronger in the UK. Similar results for British individuals who were previously affected by unemployment were observed and discussed, for instance, in Böheim and Taylor (2000).

Household Variables

Looking at the two household variables, there are some differences between the two countries. In Germany there is no significant effect of whether the individual lives as a couple or not. In contrast, in the UK this effect is positive. Individuals living as couple have an increased hazard to get reemployed. In the UK having a full-time employed partner, individuals are not entitled to further benefits after they received the contribution-based Jobseeker's allowance for six months due to the entrance requirement for the income-based Jobseeker's allowance. To receive further benefits the partner might only work for at least 24 hour a week, see European Commission (2005b). Thus they might be forced to find a job more quickly.

In Germany the fact that there are children at home causes a marginal positive effect for the first year, subsequently the effect fades away. In the UK this effect is negative for the first year and then turns into a positive one which fades away in the last few months. Böheim and Taylor (2000) found reduced hazard rates into full-time employment for British women with dependent children compared to childless men or women. In our analysis it seems that for unemployed in the UK it might be generally more difficult to find a suitable full-time job at the beginning of unemployment when children are present. Therefore it might take longer before they are successful in finding a job.

Calendar Effects

Generally, calendar effects are small compared to covariate effects. In Germany the seasonal effect between June and October is slightly negative while from November to April there is a marginal positive effect. This could be explained through weather-related higher seasonal unemployment in winter and better re-employment chances thereafter. Such seasonal changes in the unemployment rate are reported for Germany e.g. in Rudolph (1998). In the UK there is a slightly positive effect between March and July and a marginal negative effect between August and January. This negative effect might be explained through more competition caused by graduates who enter the labour market in summer and students who use their long vacation to work and replace other individuals during that time, see National Institute of Economic and Social Research (1986).

In the UK the calendar effect is not very distinct. Barely it can be seen that in the first five years the effect was negative and turns then into a slight positive effect. It might be explained through the decreasing of long-term unemployment during that period which can be deduced through the declining proportion of long-term unemployed, see Begum (2004) or Office for National Statistics (2006). In contrast in Germany there is a positive effect between 1997 and 2001 which turns in the second half of 2001 to a negative effect lasting until the end of the observed time. The negative effect might be explained through a long phase of economic weakness after a peak in a German business cycle at the beginning of the 2000's, see Schirwitz (2009).

4.5 Conclusion

In this paper we have analysed the duration of unemployment making use of penalized spline smoothing. We investigated how individual effects vary over the duration of unemployment in Germany and the UK. Seasonal and calendar effects were also taken into account. Beside a macroeconomic different behaviour of unemployment in Germany and the UK, we realised a quite substantially different behaviour on an individual level. Most dominant is the effect of age drawing a quite negative picture for elderly unemployed in Germany, while in the UK elderly unemployed have not so pronounced decreasing chances for re-employment. Other important influences for taking up employment are gender, the educational level of the unemployed individual, and former unemployment. Men have better chances to leave unemployment just as individuals with a higher educational level have better re-employment prospects. The fact of being unemployed for six months and more during the last three years before the recent spell of unemployment has especially in the UK a negative effect. We are reluctant to explain the different performances solely by different benefit schemes for the unemployed, even though it seems plausible for some effects that this contributes to it. Our analysis therefore ends with the exploratory message based on our data analysis but does not go deeper into political explanation. The latter might also not be possible using the data at hand. The analysis however demonstrates the flexibility and capacity of penalized spline smoothing as estimation routine for functional data. Given that the software is available and the analysis did not require extensive extra implementation, it seems inviting to make use of the non-proportional hazard model in other settings as well.

This chapter is based upon the following publication:
Westerheide, N. and Kauermann, G. (2012): Flexible Modelling of Duration of Unemployment Using Functional Hazard Models and Penalized Splines: A Case Study Comparing Germany and the UK. *Studies in Nonlinear Dynamics & Econometrics*, 16(1), Article 5.

5 To Move or Not to Move to Find a New Job: Spatial Duration Time Model with Dynamic Covariate Effects

The aim of this paper is to show the flexibility and capacity of penalized spline smoothing as estimation routine for modelling duration time data. We analyse the unemployment behaviour in Germany between 2000 and 2004 using a massive database from the German Federal Employment Agency. To investigate dynamic covariate effects and differences between competing job markets depending on the distance between former and recent working place, a functional duration time model with competing risks is used. It is build upon a competing hazard function where some of the smooth covariate effects are allowed to vary with unemployment duration. The focus of our analysis is on contrasting the spatial, economic, and individual covariate effects of the competing job markets and on analysing their general influence on the unemployed's re-employment probabilities. As a result of our analyses, we reveal differences concerning gender, age, and education. We also discover an effect between the newly formed and the old West German states. Moreover, the spatial pattern between the considered job markets differs.

5.1 Introduction

Unemployment, and especially its duration, is a core theme in economic research. Early references on this topic are, for example, Nickell (1979), Narendranathan, Nickell, and Stern (1985) or Narendranathan and Stewart (1993), who analysed the unemployment duration of men in Great Britain. For Germany, this subject is also of general interest. Since the German reunification in 1990, the German unemployment rate increased from 7.3% in 1991 to 12.7% in 1997 and decreased afterwards to 10.3% in 2001. Looking at the time span between 2001 and 2005, the German unemployment rate increased again from 10.3% to 13.0%. A regional discrepancy between the Old Federal States and the New Länder is distinguishable. The rate of unemployment of the old West German states during the considered period was always lower than the unemployment rate for all-Germany, that is, 11.0% in 2005, see for further details Statistisches Bundesamt, Gesis-Zuma, and WZB (2008). A central data source for analysing the duration of unemployment in Germany is the German Socio-Economic Panel (GSOEP) which allows us to employ hazard models to investigate the effect of individual- and household-specific covariates. We refer to Hunt (1995), Hujer and Schneider (1995), and Steiner (1997, 2001) using this panel for analyses with hazard rate models. Though national panel data like the GSOEP provide an information source for statistical analyses, the limited number of observations does not allow for complex models investigating local heterogeneity of a national job market. A more extensive (not to say massive) data source is available with the administrative data set of the German Federal Employment Agency provided by the Research Data Centre (Forschungsdatenzentrum (FDZ)) at the Institute for Employment Research (Institut für Arbeitsmarkt- und Berufsforschung (IAB)). In this paper, we make use of the Scientific Use File 'Regional File 1975 - 2004' of the IAB Employment Samples. The database includes information about employment biographies of employees covered by social security and of benefit recipients in Germany on a day-to-day basis. It is based on a 2% random sample taken from all employees. Previous analyses based on a comparable data set are found in Fahrmeir, Lang, Wolff, and Bender (2003), Arntz (2005), Lüdemann, Wilke, and Zhang (2006), Fitzenberger and Wilke (2007) or Arntz and Wilke (2009), for instance. Lüdemann, Wilke, and Zhang (2006) used censored quantile regression and Fitzenberger and Wilke (2007) applied censored Box-Cox quantile regression to analyse the duration of unemployment in West Germany using the IAB employment subsample from 1981-1997 or 1975-2001, respectively. Arntz (2005) took the migration behaviour of West German unemployed into account and made use of a competing risk model of leaving unemployment for a local job or a job further away. In Arntz and Wilke (2009), a semi-parametric duration model was applied which considered three different exit states: regional employment, non-regional employment, and subsidized employment. Only Fahrmeir, Lang, Wolff, and Bender (2003) included beside individual and economic effects spatial effects for their semi-parametric Bayesian time-space analysis. Except Arntz and Wilke (2009), all papers mentioned above only used data of the Old Federal States. For our analysis we use data of full-time employees all over Germany becoming unemployed between January 2000 and June 2004. This constitutes a period without major legal modifications of the unemployment compensation, see Jacobi and Kluve (2006). All in all, we analyse the duration of unemployment of 111154 individuals and explore individual, economic, and regional effects on the re-employment probabilities. Extensive data like these allow to look more closely into spatial heterogeneity of the job market and the migration of job seekers between regions. Regional mobility of unemployed in Germany has been studied before in Arntz and Wilke (2009) or Arntz (2005), for example, and for other countries, we refer to Kettunen (2002), Dockery (2000) or Détang-Dessendre (1999). In this article, we pursue the question whether unemployed individuals change their location to take up a new job and how this readiness of relocation changes with the length of unemployment. Hence we explore spatial heterogeneity within Germany. To do so, we take the distance between the former and the new working place into account.

The statistical model being used for our analysis is built upon the hazard rate or outflow rate. The classic model here is the Cox model, see Cox (1972), but we allow for non-proportional hazards in the style of varying coefficients, see Hastie and Tibshirani (1993). The proportional hazard assumption in the Cox model has been under major investigation and numerous papers suggest extensions and testing procedures. We refer the reader specifically to Gray (1994), Hess (1994) or Grambsch and Therneau (2003). Estimation of non-proportional hazards has been carried out with different smoothing techniques where we refer to spline-based approaches, see e.g. Gray (1994) or Kooperberg, Stone, and Troung (1995) or Bayesian techniques, see e.g. Kneib and Fahrmeir (2007). Hazard models with spatial effects have been proposed, for instance, by Kneib (2006) and Kneib and Fahrmeir (2007). Here we make use of penalized splines to estimate smooth dynamic covariate effects as proposed in Kauermann (2004), see also Kuhlenkasper and Kauermann (2010) or Westerheide and Kauermann (2012a). The idea of penalized spline smoothing is thereby simple and the method proves to be quite powerful. Originally introduced by Eilers and Marx (1996), the method has become quite fashionable over the last years, see Ruppert, Wand, and Carroll (2009). Instead of fitting a low dimensional parametric model, a high dimensional spline-based model is fitted, and in order to achieve a smooth and numerically stable fit, a penalty is imposed on the high-dimensional spline coefficients. The routine is implemented in **R** (see R Development Core Team, 2008 and Wood, 2006) and we demonstrate how to make use of it after simple data processing.

The paper is organised as follows. Section 2 introduces the statistical model and describes the estimation routine. Section 3 gives more detailed information about the database and the covariates being used. Section 4 gives a detailed data analysis before we conclude in Section 5.

5.2 Functional Hazard Model with Competing Risks

5.2.1 Statistical Model

Let t_i denote the duration of unemployment for the ith individual and denote with $x_i = (x_{i1}, \ldots, x_{ip})$ a set of covariates under investigation. These are individual characteristics, like age or education of the employee. With s_i we denote the location of the former working place of the ith individual at the point in time of losing the job. The location is thereby the centroid of one of 343 districts in Germany. With u_i we denote the (average) unemployment rate in the district and address with c_i the calendar time at the beginning of the unemployment spell. Finally, we denote with d_i the censoring variable stating whether the true (unobserved) duration time is larger than t_i. We assume now that the hazard function for the duration of unemployment decomposes to

$$h(t, x_i, s_i, c_i, u_i) = \exp\{\underbrace{\beta_0(t)}_{1} + \underbrace{\sum_{j=1}^{p} x_{ij}\beta_j(t)}_{2} + \underbrace{\gamma(s_i)}_{3} + \underbrace{\delta(c_i)}_{4} + \underbrace{\phi(u_i)}_{5}\}. \tag{5.1}$$

The first component expresses the baseline hazard assuming that $\beta_0(t)$ is a smooth function in time t. The second component gives the covariate effects which may vary with duration time, where again $\beta_j(t)$ are smooth functions to be fitted from the data. Note that assuming constant effects $\beta_j(t) \equiv \beta_j$ and $h_0(t) = \exp(\beta_0(t))$ we get with the first two components in Equation (5.1) a classical Cox proportional hazard model. The third component does now capture spatial heterogeneity while the fourth component exhibits the temporal structure. Finally, the fifth component gives the influence of the regional unemployment rate. We leave $\gamma(s)$, $\delta(c)$ and $\phi(u)$ unspecified and estimate its shape through the data. We assume and postulate, however, that $\gamma(s)$, $\delta(c)$ and $\phi(u)$ are smooth functions, that is, there are no abrupt changes or jumps. Smoothness can also be interpreted as sufficient differentiability, if in fact the covariate is metrically scaled. To achieve identifiability we additionally postulate that $\gamma(s)$, $\delta(c)$ and $\phi(u)$ integrate out to zero.

Model (5.1) assumes that the chances of re-employment depend on the location of the former job but not on the spatial heterogeneity of the job market in general. An unemployed person in region s_i, say, participates in the job market not only in or close to location s_i but also nation-wide (or even internationally). That is to say that an unemployed person has chances of getting a new job locally or, somewhat competitively, further away from his/her original employer. We model this using a competing hazard model and assume that the hazard function (5.1) depends on the distance between the location of the potential new job place and the location of the lost job. The chances of getting a job are competing, that is, if a job is taken locally, the unemployed person is no longer a potential employee for the job market further away and vice versa. To incorporate such spatial heterogeneity of the job market, we discretize the problem and consider the competing hazards of taking up a job (a) up to 50 km of the last work location ($k = 1$), (b) over 50 up to 150 km ($k = 2$) and (c) beyond 150 km ($k = 3$). The hazard (5.1) is therefore replaced by the additive competing hazards

$$h(t, x_i, s_i, c_i, u_i) = \sum_{k=1}^{K} h_k(t, x_i, s_i, c_i, u_i), \tag{5.2}$$

where $K = 3$ in our setting and $h_k(.)$ decomposes like (5.1) but with functional effects being dependent on k, the distance to the former work location, that is,

$$h_k(t, x_i, s_i, c_i, u_i) = \exp\left\{ \sum_{j=0}^{p} x_{ij}\beta_{jk}(t) + \gamma_k(s_i) + \delta_k(c_i) + \phi_k(u_i) \right\}, \tag{5.3}$$

where $x_{i0} \equiv 1$. Having competing chances, we express with $d_i \in \{0, 1, \ldots, K\}$ the censoring information stating whether at time t_i the job taken was up to 50 km of the original work location ($d_i = 1$), over 50 but up to 150 km ($d_i = 2$) or above 150 km ($d_i = 3$). Censored observations are notated as $d_i = 0$.

5.2.2 Estimation

We start the discussion on estimation by describing the fitting of the smooth, functional components in Equation (5.3). We first represent each unknown function as a linear combination of thin plate spline basis terms (Wahba, 1990, pp. 30-34), with the popular cubic regression spline basis resulting as a special case, see Wood (2006). We use the same approach for the calendar effect $\delta(.)$ and the unemployment effect $\phi(.)$. For the spatial effect $\gamma(.)$ we use thin plate regression splines. The functional components in Equation (5.3) are therefore replaced by

$$\begin{aligned} &\beta_{0k}(t) = B_0(t)b_{0k}, \quad \beta_{jk}(t) = B_j(t)b_{jk} \\ &\gamma_k(s) = B_\gamma(s)b_{\gamma k}, \quad \delta_k(c) = B_\delta(c)b_{\delta k}, \quad \phi_k(u) = B_\phi(u)b_{\phi k} \end{aligned} \tag{5.4}$$

with $B(.)$ as spline bases. Classical spline smoothing is built upon knots placed at the (unique) observed values of the variables. To reduce the computational burden, we follow Hastie (1996) and Wood (2003) and employ so-called low-rank smoothing. For each function, this involves to work with a reduced set of knots which is still large enough to capture the functional shape but small enough to guarantee feasible computation. The idea has been coined as P(enalized)-spline smoothing by Eilers and Marx (1996), see also Ruppert, Wand, and Carroll (2003). Ruppert, Wand, and Carroll (2009) provide an extensive survey of recent results and papers in this field demonstrating the popularity of the approach. Denoting with q the number of knots, we follow Wood (2006, p. 161) and set $q = 10$ for the baseline, covariate, and calendar effects and $q = 50$ for the spatial effect functions. We fitted the model for larger values of q as well but observed the established fact that the choice of q has little influence on the fit; see Ruppert (2002) or Kauermann and Opsomer (2011) for a discussion.

Assume now that $(t_i, d_i, x_i, s_i, c_i, u_i)$ denote the observations for the ith individual, $i = 1, \dots, n$. Assuming independence of the individuals, the log-likelihood for parameter vector $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^T, \dots, \boldsymbol{\theta}_K^T)^T$ with $\boldsymbol{\theta}_k = (b_{0k}^T, b_{xk}^T, b_{\gamma k}^T, b_{\delta k}^T, b_{\phi k}^T)^T$ and $b_{xk}^T = (b_{jk}^T,\ j = 1, \dots, p)$ equals $l(\boldsymbol{\theta}) = \sum_{i=1}^n l_i(\boldsymbol{\theta})$ where (see Cox and Oakes 1984)

$$
\begin{aligned}
l_i(\boldsymbol{\theta}) = \sum_{k=1}^{K} \Bigg[& 1_{\{d_i=k\}} \left\{ \boldsymbol{B}_i(t_i)\boldsymbol{b}_k + B_\gamma(s_i)b_{\gamma k} + B_\delta(c_i)b_{\delta k} + B_\phi(u_i)b_{\phi k} \right\} \\
& - \int_0^{t_i} \exp\left\{ \boldsymbol{B}_i(t)\boldsymbol{b}_k \right\} \mathrm{d}t\ \exp\left\{ B_\gamma(s_i)b_{\gamma k} + B_\delta(c_i)b_{\delta k} + B_\phi(u_i)b_{\phi k} \right\} \Bigg], \qquad (5.5)
\end{aligned}
$$

with $\boldsymbol{B}_i(t) = (x_{ij}B_j(t),\ j = 0, \dots, p)$ with $x_{i0} \equiv 1$ and $\boldsymbol{b}_k^T = (b_{jk}^T,\ j = 0, \dots, p)$. Note that (5.5) is not available analytically due to the integral component which requires to be calculated numerically. Cai, Hyndman, and Wand (2002) and Kauermann (2004) make use of trapezoid approximation. Numerically more accurate, however, is a Simpson approximation, see e.g. Gautschi (1997) or Gil, Segura, and Temme (2007), which results as follows. For the ith individual, we divide the interval $[0, t_i]$ into R equidistant subintervals $[T_{r-1}, T_r]$, say, with $r = 1, \dots, R$, where $T_0 = 0$ and $T_R = t_i$. Then, the integral component in (5.5) can be approximated through

$$
\sum_{r=0}^{R} \left[\frac{T_{r+1} - T_{r-1}}{6} \exp\bigl\{ \boldsymbol{B}_i(T_r)\boldsymbol{b}_k \bigr\} + \frac{4(T_r - T_{r-1})}{6} \exp\Bigl\{ \boldsymbol{B}_i\left(\frac{T_r + T_{r-1}}{2} \right) \boldsymbol{b}_k \Bigr\} \right], \qquad (5.6)
$$

with $T_{-1} = T_0$ and $T_{R+1} = T_R$. Defining $Y_{ijk} = 0$ for $j = 0, \dots, 2R - 1$ and $Y_{ijk} = 1_{\{d_i=k\}}$ for $j = 2R$ we can approximate each summand in (5.5) using the Simpson approximation (5.6). This in turn leads to the log-likelihood contributions of independent Poisson distributed variables

$$
Y_{ijk} \sim \text{Poisson}\bigl(\lambda_{ij} = \exp\{ \boldsymbol{B}_i(\tilde{T}_j)\boldsymbol{b}_k + B_\gamma(s_i)b_{\gamma k} + B_\delta(c_i)b_{\delta k} + B_\phi(u_i)b_{\phi k} + o_j \} \bigr), \qquad (5.7)
$$

with $i = 1, \dots, n$, $j = 0, \dots, 2R$, where $\tilde{T}_j = T_{j/2}$ for j even and $\tilde{T}_j = (T_{(j+1)/2} + T_{(j-1)/2})/2$ for j odd. With o_j we define the offset $o_j = \log\bigl((T_{j/2+1} - T_{j/2-1})/6 \bigr)$ for j

even and $o_j = \log\big(4(T_{(j+1)/2} - T_{(j-1)/2})/6\big)$ for j odd. In other words, after some simple data management, we can approximate the likelihood contributions in Equation (5.5) by the standard likelihood resulting from a Poisson regression model (5.7).

The next step is to impose a penalty on the spline coefficients to achieve smooth functional fits. Apparently, the model is high-dimensional which implies that the Maximum Likelihood estimate based on the log-likelihood for Equation (5.7) will yield wiggled fitted curves. Following Eilers and Marx (1996) and Ruppert, Wand, and Carroll (2003), we therefore impose a penalty on the coefficients. As demonstrated in Wand and Ormerod (2008), we can rewrite the spline representation in Equation (5.4) by extracting the linear slope, that is,

$$\beta_{jk}(t) = B_j(t) b_{jk} = \beta_{0jk} + t\beta_{1jk} + \tilde{B}_j(t)\tilde{b}_{jk}$$

where $\tilde{B}_j$ is now the reduced rank basis with intercept and linear slope extracted, $j = 0, \ldots, p$. Similarly, we obtain for $\gamma(s)$, $\delta(c)$ and $\phi(u)$ reduced basis matrices $\tilde{B}_\gamma(s)$, $\tilde{B}_\delta(c)$ and $\tilde{B}_\phi(u)$, respectively. We now impose a quadratic penalty on the spline coefficient, for example, $\lambda_{jk}\tilde{b}_{jk}^T \tilde{D}_{jk}\tilde{b}_{jk}$ which penalizes squared second-order derivatives of the function (see O'Sullivan, 1986 or Wahba, 1990), or second (or higher)-order differences of the spline coefficients b_{jk} (see Eilers and Marx, 1996). The parameter λ_{jk} thereby plays the role of a smoothing parameter with $\lambda_{jk} \to \infty$ leading to a linear fit. The complete penalized likelihood takes then the form

$$\begin{aligned} l(\boldsymbol{\beta}, \tilde{\boldsymbol{b}}, \boldsymbol{\lambda}) = \sum_{k=1}^{K} \Bigg\{ & \sum_{i=1}^{n} \tilde{l}_i(\boldsymbol{\beta}_k, \tilde{\boldsymbol{b}}_k) - \frac{1}{2}\sum_{j=0}^{p} \lambda_{jk}\tilde{b}_{jk}^T \tilde{D}_{jk}\tilde{b}_{jk} \\ & - \frac{1}{2}\lambda_{\gamma k}\tilde{b}_{\gamma k}^T \tilde{D}_{\gamma k}\tilde{b}_{\gamma k} - \frac{1}{2}\lambda_{\delta k}\tilde{b}_{\delta k}^T \tilde{D}_{\delta k}\tilde{b}_{\delta k} - \frac{1}{2}\lambda_{\phi k}\tilde{b}_{\phi k}^T \tilde{D}_{\phi k}\tilde{b}_{\phi k} \Bigg\} \end{aligned} \tag{5.8}$$

with $\tilde{l}_i$ as log likelihood for the Poisson variables (5.7), $\boldsymbol{\beta}_k = \big((\beta_{0jk}, \beta_{1jk}),\ j = 0, \ldots, p;$ $(\beta_{0\gamma k}, \beta_{1\gamma k}), (\beta_{0\delta k}, \beta_{1\delta k}), (\beta_{0\phi k}, \beta_{1\phi k})\big)^T$, analogous definition for $\tilde{\boldsymbol{b}}_k$ and obvious definition for $\boldsymbol{\lambda}_k = (\lambda_{jk}, j = 0, \ldots, p, \lambda_{\gamma k}, \lambda_{\delta k}, \lambda_{\phi k})^T$, $k = 1, \ldots, K$. Note that Equation (5.8) can be easily fitted with software for generalized additive models, see Hastie and Tibshirani (1990). We use the `bam()` procedure in **R** (package `mgcv`) which extends the `gam()` procedure to work with large data sets, see also Wood (2010). In fact, the only thing which is numerically necessary in order to fit the model with the available software is to restructure the data to obtain the Poisson model (5.7). The smoothing parameters $\boldsymbol{\lambda}_k$ can be selected following a generalized cross validation, as implemented in the `bam()` procedure. We employed REstricted Maximum Likelihood (REML) estimation which is also implemented in `bam()`.

Finally, inference can be drawn following standard asymptotic arguments as outlined in Ruppert, Wand, and Carroll (2003) or Wood (2006), see also Kauermann, Krivobokova, and Fahrmeir (2009). In fact, with $\boldsymbol{\theta}_k = (\boldsymbol{\beta}_k^T, \tilde{\boldsymbol{b}}_k^T)^T$ denoting the complete parameter vector, we define with $F(\boldsymbol{\theta}_k, \boldsymbol{\lambda}_k)$ the Fisher matrix. It can generally be shown (Ruppert, Wand, and Carroll, 2003) that

$$\text{Var}(\hat{\boldsymbol{\theta}}_k) = F^{-1}(\boldsymbol{\theta}_k, \boldsymbol{\lambda}_k) F(\boldsymbol{\theta}_k, \boldsymbol{\lambda}_k = 0) F^{-1}(\boldsymbol{\theta}_k, \boldsymbol{\lambda}_k).$$

5.3 Data Description

For our data analysis we use the Scientific Use File ‘Regional File 1975 - 2004’ of the IAB, where a detailed description of the entire database is provided in Drews (2008). Beside socio-demographic-related characteristics, the database also includes employment and regional characteristics. Thus, it allows us to empirically explore and investigate individual effects on the duration of unemployment as well as regional and economic effects. We analyse the unemployment duration of 111154 former full-time employed individuals who became unemployed between January 2000 and June 2004. The observation period ends in December 2004. The maximum duration of unemployment is censored at 1095 days, that is, 3 years, to restrict the analysis to medium-term unemployment. As event we define full-time re-employment dependent on the distance between the new and the former working place classified into up to 50 km, between 50 km, and 150 km and over 150 km. As distance, we use the Euclidean distance between the centroids of 343 defined regions given in the data. For individuals with more than one unemployment spell in the database, we randomly select one spell for our analyses which maintains the independence of our observations. As individual covariates we include gender, age, education, and an East/West indicator. Self-employed and civil servants are not included in the database. Table 5.1 shows the event rate (re-employment rate) broken down by distance and covariates.

We distinguish between male and female unemployed and differentiate among five age groups: below 25 years, between 25 and 34 years (taken as reference), between 35 and 44 years, from 45 to 54 years, and over 54 years. The educational level during the last period of full-time employment is divided into four different categories: individuals without vocational training, individuals who attended a secondary general school or intermediate secondary school and completed successfully a vocational training (taken as reference), individuals with A-levels and with or without vocational training, and graduates from university or universities of applied science. The East/West indicator gives information about the location of the individual’s former working place (eastern

IAB Employment Samples Regional File 1975 - 2004	Event <= 50 km		Event > 50 and < 150 km		Event > 150 km		Σ
	all	1st year	all	1st year	all	1st year	
Overall	49.64%	45.10%	4.62%	3.99%	4.64%	4.06%	111154 (100%)
Men	54.01%	49.44%	5.15%	4.45%	5.09%	4.45%	76038 (100%)
Women	40.17%	35.71%	3.45%	2.98%	3.65%	3.19%	35116 (100%)
Age < 25	55.01%	52.17%	4.93%	4.51%	4.51%	4.18%	18546 (100%)
Age 25-34	52.28%	47.88%	5.45%	4.74%	5.93%	5.28%	29055 (100%)
Age 35-44	52.75%	47.67%	5.13%	4.43%	5.34%	4.62%	29358 (100%)
Age 45-54	49.92%	43.92%	4.27%	3.50%	4.03%	3.28%	22205 (100%)
Age $\geq$ 55	25.57%	23.32%	1.47%	1.18%	1.08%	0.96%	11988 (100%)
Without Vocational Training	48.92%	43.35%	3.45%	2.87%	2.89%	2.36%	21280 (100%)
With Vocational Training	51.49%	47.14%	4.69%	4.06%	4.66%	4.08%	81450 (100%)
A-Levels	37.44%	33.23%	6.21%	5.67%	7.50%	6.71%	3560 (100%)
University	30.70%	27.28%	7.34%	6.52%	9.85%	9.09%	4864 (100%)
West Indicator	49.07%	44.96%	4.32%	3.81%	4.37%	3.84%	78664 (100%)
East Indicator	51.01%	45.44%	5.33%	4.43%	5.28%	4.57%	32490 (100%)

Table 5.1: Distribution of events and covariates for the IAB Regional File (111154 individuals) for all job returns and job returns during the first year. The last column presents the total sum of all individuals of the corresponding row.

German states versus former western German states). As spatial information, we use the centroid of the unemployed's former working place. To account for the economic environment, we include the unemployment rate of the region at the time point, that is, year, of entry into unemployment of an individual and as calendar time we use the day when the individual became unemployed.

5.4 Data Analysis

In Figures 5.1-5.3, we show the resulting fit of model (5.3). The left column of the first two figures shows the effects for unemployed getting a full-time job close to their former working place, that is, less than 50 km distance. The middle column shows the fitted effects for the chance of returning to full-time re-employment in a distance between 50 and 150 km from the former job place and the right column displays the fitted effects for a distance beyond 150 km, respectively. The plots show the fitted effect and the corresponding confidence intervals as shaded areas. As dot-dashed lines we also include simple non-dynamic parametric effects based on a Cox proportional hazard model, that is, we fitted model (5.3) assuming $\beta_{jk}(t) \equiv \beta_{jk}$. The resulting estimates $\hat{\beta}_{jk}$ are listed in Table 5.2 for completeness. The proportional hazard model can be seen as benchmark.

IAB Employment Samples Regional File 1975 - 2004	k=1 $\hat{\beta}_{jk}$ (Std. Error)	k=2 $\hat{\beta}_{jk}$ (Std. Error)	k=3 $\hat{\beta}_{jk}$ (Std. Error)
Intercept	-5.854 (0.013)	-8.263 (0.039)	-7.954 (0.038)
Men	reference	reference	reference
Women	-0.381 (0.010)	-0.572 (0.033)	-0.535 (0.032)
Age < 25	0.175 (0.013)	0.145 (0.042)	0.011 (0.043)
Age 25-34	reference	reference	reference
Age 35-44	-0.075 (0.011)	-0.178 (0.036)	-0.233 (0.035)
Age 45-54	-0.232 (0.013)	-0.471 (0.041)	-0.622 (0.042)
Age ≥ 55	-1.180 (0.020)	-1.878 (0.080)	-2.295 (0.091)
Without Vocational Training	-0.194 (0.011)	-0.456 (0.041)	-0.650 (0.045)
With Vocational Training	reference	reference	reference
A-Levels	-0.220 (0.028)	0.352 (0.070)	0.473 (0.064)
University	-0.383 (0.027)	0.584 (0.056)	0.803 (0.050)
West Indicator	reference	reference	reference
East Indicator	0.150 (0.029)	0.716 (0.081)	-0.076 (0.086)

Table 5.2: Estimated non-dynamic parametric effects $\hat{\beta}_{jk}$ and standard errors of the IAB Regional File (111154 individuals).

Evidently, the dynamic behaviour of the fitted curves $\hat{\beta}_{jk}(t)$ indicates the superiority of a model including dynamics in comparison to proportional hazards, but some of the effects show only a weak dynamic behaviour. Therefore, we make use of a backward selection excluding successively the estimated dynamic effects $\hat{\beta}_{jk}(t)$ until the best model is found. As selection criteria, we use a revised Bayesian information criterion (BIC) recommended by Volinsky and Raferty (2000). The results are shown in Table 5.3.

The estimated dynamic effects of the best model found through this selection are highlighted in Figures 5.1 and 5.2 with bold frames. The estimated spatial effects for all three competing distances are shown in Figure 5.3.

We first look at the baseline effect $\hat{\beta}_{0k}(t)$ for all three competing distances in Figure 5.1. Note that the confidence intervals get larger to the end because of less events in the end of the observation time. Generally, the baseline is first increasing with a maximum at around 50 days and then decreasing afterwards. The latter mirrors the fact that longer unemployment decreases the chance for re-employment. For distances above 50 km, we observe an interesting and clearly exposed second peak at around 180 days. This peak can be explained by changing benefit schemes. Unemployed whose duration of compulsory insurance coverage is between 12 and 16 months are only entitled to unemployment insurance benefits (Arbeitslosengeld) for a duration of six months which is by the way the shortest duration of unemployment insurance benefits being paid, see European Commission (2005b). Thus, it seems that after 180 days this group of

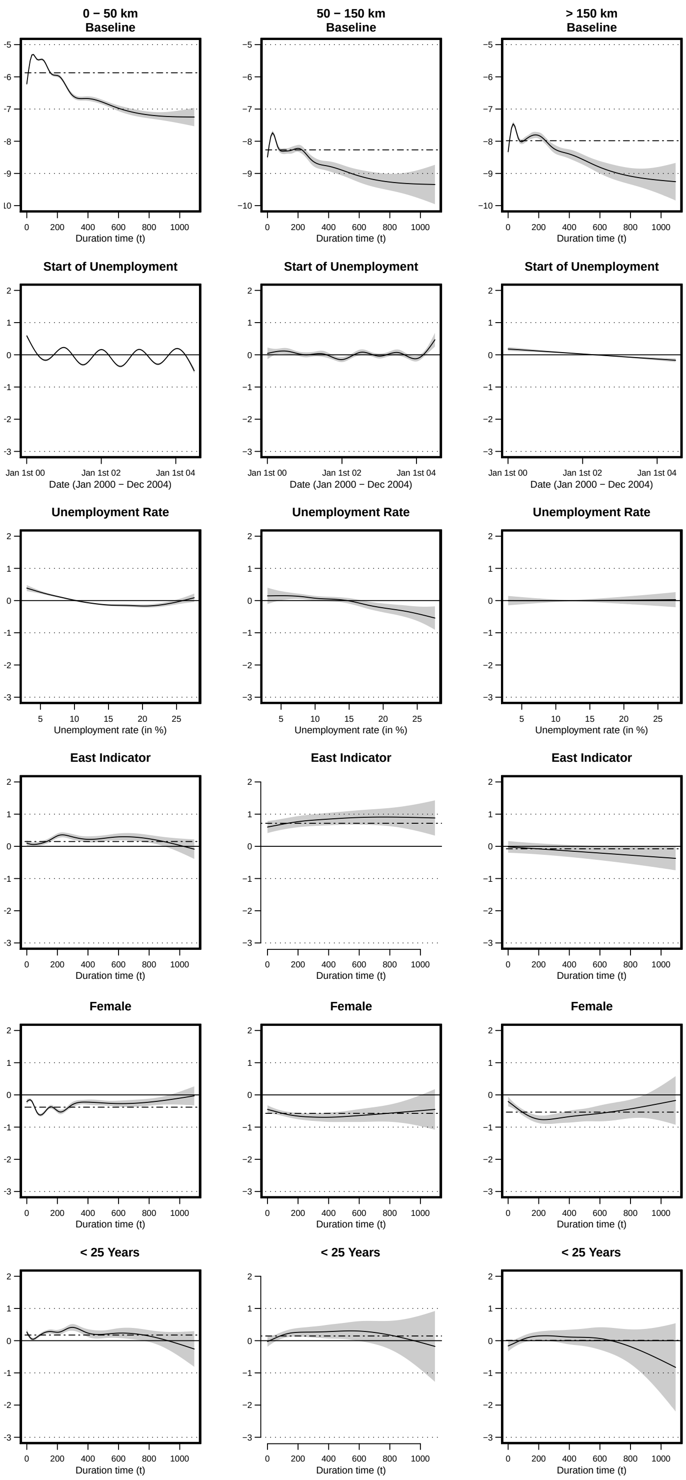

Figure 5.1: Fitted dynamic effects with confidence intervals for duration time of unemployment (in days), start of unemployment (in days), and unemployment rate (in %) for a distance up to 50 km (first column), between 50 and 150 km (second column), and over 150 km (third column) of the IAB Regional File. Fitted parametric effects (dot-dashed lines) are added. Effects selected as dynamic effects by the model selection are drawn with bold frames.

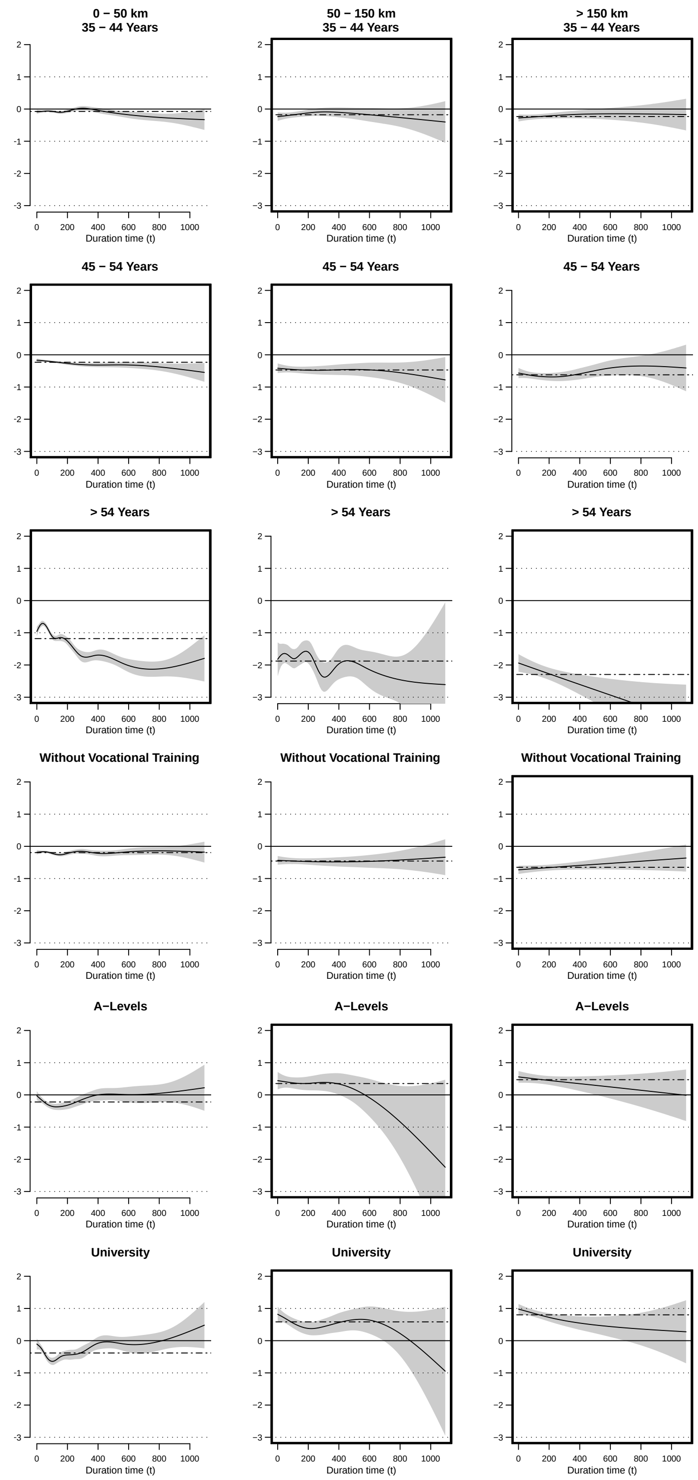

Figure 5.2: Fitted dynamic effects with confidence intervals for duration time of unemployment (in days) for a distance up to 50 km (first column), between 50 and 150 km (second column), and over 150 km (third column) of the IAB Regional File. Fitted parametric effects (dot-dashed lines) are added. Effects selected as dynamic effects by the model selection are drawn with bold frames.

Revised BIC for k=1 with q=log(55172)	
Model with all dynamic effects	625117.5
- Without Vocational Training	625081.0
- Age 35-44	625070.0
- A-Levels	625058.4
- University	625057.7

Revised BIC for k=2 with q=log(5130)	
Model with all dynamic effects	81553.52
- Age $<$ 25	81502.27
- Without Vocational Training	81500.98
- East Indicator	81489.72
- Age $\geq$ 55	81489.45

Revised BIC for k=3 with q=log(5153)	
Model with all dynamic effects	81491.90
- Age 45-54	81485.27

Table 5.3: Starting with the competing hazards given in (5.3) the steps of a backward selection using the revised BIC $= -2$ log-likelihood $+ pq$ by Volinsky and Raferty (2000), where p is the number of parameters in the fitted model and q is the logarithm of the number of uncensored events, are shown.

unemployed shows more willingness to accept a new job farther away from their former working location.

Economic Effects

Looking now at calendar and economic effects, we see a clear seasonal pattern for unemployed which are reemployed locally, that is, within 50 km of the former job location. The functional curve shows increased chances for re-employment during the winter, meaning that individuals losing their job during winter months are more likely to get rehired compared to those getting unemployed during the summer. This mirrors the effect of workers employed in a seasonal business, for example, construction industry. Similar cyclic effects like this are found in Fahrmeir, Lang, Wolff, and Bender (2003) and Lüdemann, Wilke, and Zhang (2006). The seasonal pattern is not observed for non-local re-employment, that is, for distances above 50 km. For large distances, that is, above 150 km, we see a slightly decreasing effect over time. Looking now at the effect of the district-specific unemployment rate, we see a general pattern. Regions with a low unemployment rate provide higher chances for local re-employment, that is, in a distance of up to 150 km from the former working place. This effect might be caused by a higher economic strength of the region where the former working place is situated. In contrast, the chance of finding a job in a distance above 150 km is not influenced by the height of the local unemployment rate of the region of the individual's former employer. In Arntz and Wilke (2009), it is pointed out that when using competing-risk Cox-proportional hazard estimates, regional factors such as the unemployment rate usually only have a weak influence on the unemployment duration compared to individual characteristics. This is similar to our findings.

Next we look at the East/West indicator which is natural to be included in a model given the different economic conditions in East and West Germany. The effect for having the former working place in the newly formed eastern German states is slightly positive for individuals getting reemployed in a distance between 0 and 50 km and this effect is even stronger for a distance between 50 and 150 km. For unemployed who find a job in a region more than 150 km away of their former working place, there is no effect evident in the data. The effect for all three distances is thereby nearly constant over the observed unemployment duration. This implies that unemployed who worked in the newly formed German states have increased chances of getting reemployed, in particular if they are mobile in their job search and find a job in a distance between 50 and 150 km away from their former working place. Arntz and Wilke (2009) found out that conditional durations of unemployment are similar in both parts of Germany. They argue that in the 2000s there was only a small difference left between the conditional unemployment duration for individuals from the eastern and western part of Germany, respectively. They differentiate between local and non-local regular employment. The probability of finding a local permanent appointment for most unemployed in East Germany is slightly lower than that for unemployed from the old West German states, but East German unemployed have higher migration rates which Arntz and Wilke (2009) explain by intense pull factors from the old West German states. This effect is mirrored in our results.

Spatial Effects

Before interpreting the individual-specific effects we investigate the fitted spatial effects $\gamma_k(s)$ shown in Figure 5.3 for the three competing distances. For a better visual interpretation, we include the fitted East/West indicator fitted in a non-dynamic form, that is, using the dot-dashed lines in Figure 5.1 (row 4). Spatial heterogeneity of the job market in Germany is seen in all three graphs and it differs between the three distances. Since the job market between 50 and 150 km shows the strongest pattern, it is primarily highlighted here and interpreted. First, in the western part of Germany, there is a clus-

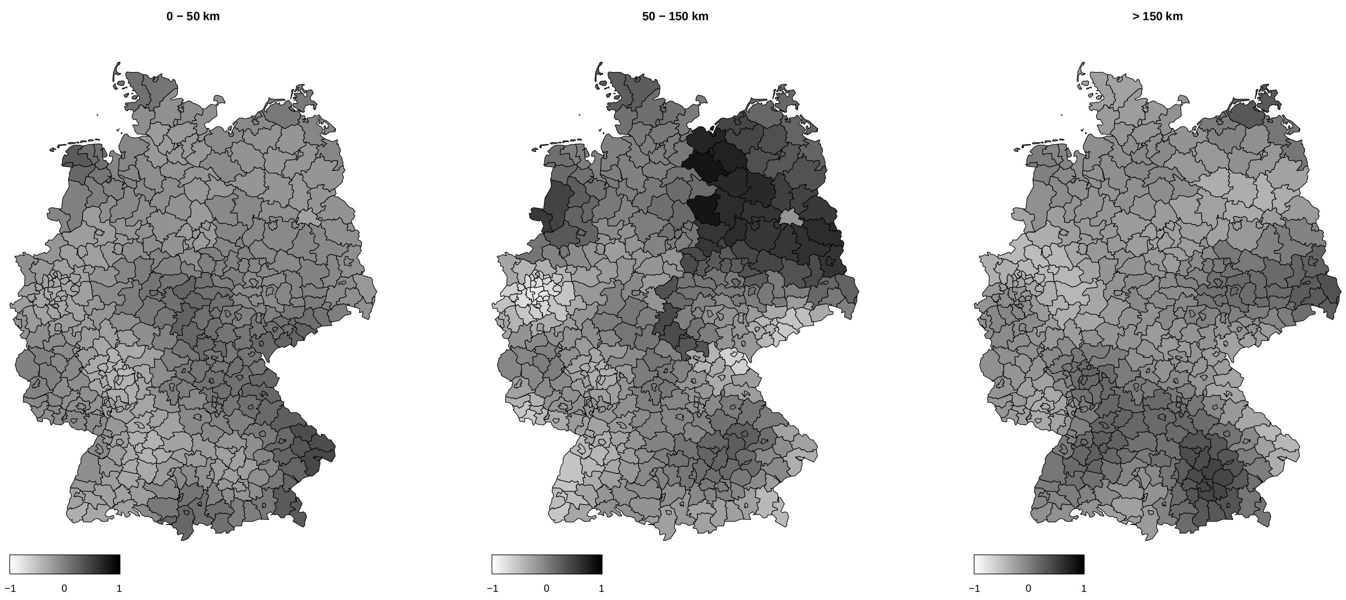

Figure 5.3: Fitted spatial effects for a distance up to 50 km (first map), between 50 and 150 km (second map), and over 150 km (third map) of the IAB Regional File. Dark-coloured districts correspond to a positive effect and light-coloured districts correspond to a negative effect on the re-employment chances.

ter where reduced re-employment chances are observed for a distance between 50 and 150 km. These districts centre around the Ruhr region and in the far southwest around Saarland and the Black Forest. In the eastern part of Germany, only near the border to the Czech Republic reduced re-employment chances are visible. Economically these regions are mostly (old) industrialised regions sometimes in agglomeration areas which do not have other city regions in the further surrounding area. Looking at regions with increased re-employment chances for job places 50-150 km away from the former job the structure in the newly formed German states in the east of Germany is remarkable. The northern part exhibits increased re-employment chances. This area is mostly rural and the effect mirrors the fact that if employees lose their job in this area they are in need of finding a new job far away from their former local job. Note that the spatial structure shown is a partial effect since the district-wise unemployment rate is included in the model. Insofar we can explain the difference to Fahrmeir, Lang, Wolff, and Bender (2003), who found in their analysis particular low employment hazards in the Saar and the Ruhr areas and higher employment hazards in the southern part of West Germany.

Effects of Individual Covariates

Next we investigate the effect of individual covariates. Considering gender, the effect for females is slightly negative compared with males as reference for all three distances, and it remains negative over the duration of unemployment. The effect for women taking up a job locally, that is, up to 50 km, shows small peaks straight at the beginning, at around 180 and 360 days. The latter two peaks might be due to ending unemployment insurance benefits after 6 and 12 months respectively; as already discussed above, see for details European Commission (2005b). Generally, it seems that females compared with males have worse re-employment chances regardless of the distance. The difficulty in finding a job for women compared with men was observed in other analyses of German data sets as well, see for example Steiner (2001).

Next we look at the effect of age. With increasing age the re-employment chances generally decrease. This negative effect intensifies with an increasing distance to the former job location. Looking at a maximum distance of 50 km for re-employment, the effect for age group 1 (< 25 years) compared to age group 2 (25-34 years) is positive for up to 2 years. A similar but weaker pattern can be seen for the other two competing distances. The age effect for unemployed between 35 and 44 years compared to age group 2 is slightly negative for most of the duration time and a bit more pronounced the second half for the local job market, that is, up to 50 km, and for distances beyond 50 km it is at first a little more negative but vanishing. This negative effect might trace back to the restrained mobility in the first year of unemployment. For individuals between 45 and 54 and over 54 years, the negative age effect is more pronounced and lasts over the entire time interval. In Fahrmeir, Lang, Wolff, and Bender (2003) a similar pattern has been observed. Arntz and Wilke (2009) discovered in their study that usually the chances of finding a permanent appointment in local or non-local areas decreases with increasing age. Worse re-employment probabilities for older unemployed were also noticed in Hunt (1995), Hujer and Schneider (1995), and Steiner (2001).

Looking now at the educational effects, it becomes obvious that for better educated unemployed, the re-employment chances are increased, in particular for the non-local job market. While individuals with A-Levels and graduates have a negative effect for the local job market during the first year; this effect is positive for the distances further away. Unemployed with an university degree even show in the first 2 years a positive effect for finding a job in a distance further than 50 km away from the former working place. Thus, it seems that better educated unemployed participate in the national job market and are likely to take a job even in a distance away from their former job, especially in the first year. In contrast, for less educated unemployed (without vocational training), the effect is negative compared to individuals with vocational training, and this negative effect intensifies with increasing distance. Similar findings were observed in Arntz (2005). She found out that education impinges on the mobility of unemployed, that is, higher education leads to an increasing probability of being mobile. One reason for better re-employment chances in the distance for better educated unemployed might be the migration out of a rural district to find a job as Détang-Dessendre (1999) concluded for educated young unemployed in rural France. Lauer (2003) reasons that higher educated individuals generally have a better re-employment probability. In Fahrmeir, Lang, Wolff, and Bender (2003) and Lüdemann, Wilke, and Zhang (2006), education has only a weak influence on the re-employment probability.

5.5 Conclusion

In this paper, we have analysed the duration of unemployment making use of penalized spline smoothing. To investigate how spatial, economic, and individual effects behave and how they compete in view of different distances between former and new working place we used a functional hazard model with competing risks. Generally, the re-employment chances for women are worse than for men. Moreover, less educated and older (over 44 years) unemployed tend to have reduced re-employment probabilities whatever distance is regarded. In addition to that, for unemployed over 54 years the negative effect intensifies with duration of unemployment. Besides a different behaviour over the duration of unemployment, we realised a substantially different behaviour between the three considered distances, especially for age and education but also for regional and temporal effects. The effect of education draws a positive picture for the job chances for graduates and unemployed with A-Levels in a distance above 50 km from the former working place, particularly during the first year. Another highlighted effect exists for unemployed between 35 and 44 years who have worse re-employment chances for a distance above 50 km during the first year. Individuals whose former working place has been in the newly formed German states have far better re-employment probabilities in a distance between 50 and 150 km compared with both other distances and unemployed working before in the old West German states. A cyclic temporal effect is only visible for a distance up to 50 km. It shows better re-employment chances for individuals who got unemployed during wintertime. Spatial effects for the location of the former working place differ between both the considered regions and distances between former and recent working place. Most striking is the spatial pattern for a distance between 50 and 150 km.

Although it seems plausible to explain all noticed effects by the individual's mobility behaviour, the German benefit scheme or regional characteristics, we do this with caution. Our analysis leaves an exploratory message based on the analysis of our data, but we do not go deeper into labour economic interpretation. However, our analysis demonstrates the flexibility and capacity of penalized spline smoothing as estimation routine for a massive database. Given that the software is available and the analysis did not require extensive extra implementation, it seems inviting to make use of the non-proportional hazard model with competing risks in other settings as well.

This chapter is based upon the following publication:
Kauermann, G. and Westerheide, N. (2012): To move or not to move to find a new job: spatial duration time model with dynamic covariate effects. *Journal of Applied Statistics*, 39(5), 995-1009.

6 Getting Unemployed: Factors Influencing the Risk of Unemployment in Germany

The intention of this paper is to investigate which covariates influence the risk of getting unemployed in Germany. For our analysis, we use the massive database from the German Federal Employment Agency (Scientific Use File 'Regional File 1975 - 2004') to model the risk of an individual to become unemployed between 2000 and 2004 in Germany. As individual covariates we include gender, age, and education as fixed effects in our model. Beside these individual characteristics, regional as well as calendrical and economic information is considered and included as smooth functional effects in the model. As result of our data analysis we uncover strong educational and age effects as well as a dominating calendrical effect on the individual's risk of getting unemployed. Surprisingly and interestingly though, we find that neither the rate of unemployment nor the region has a strong influence on the risk of getting unemployed.

6.1 Introduction

A well-known problem in economies and a focal point in economic research is unemployment, see for example Layard, Nickell, and Jackman (2009), Blanchard (2006) or Ljungqvist and Sargent (1998). Often the unemployment rate is used as a macroeconomic measure to explain changes in regional and national labour markets, as done, for instance, in official statistics in European Commission (2009), Eurostat (2009) or OECD (2009). Moreover, the duration of unemployment is of utmost interest to explain the unemployment behaviour of individuals for different points of focus, see for example Narendranathan and Stewart (1993), Böheim and Taylor (2000), Bover, Arellano, and Bentolila (2002), Røed and Zhang (2003), Lauer (2003), Tatsiramos (2009) or Westerheide and Kauermann (2012a). Analyses regarding only the unemployment duration in Germany include Hunt (1995), Steiner (1997, 2001) or Fahrmeir, Lang, Wolff, and Bender (2003). Looking at unemployment in Germany, the differences between the old West German states and the newly formed Eastern German states are of particular interest. After the German reunification in 1990, a regional difference between the unemployment rates of the Old and New Länder is clearly noticeable. Between 1991 and 2008, the unemployment rate of the old West German states was always lower than the unemployment rate of entire Germany, see for further details Statistisches Bundesamt, Gesis-Zuma, and WZB (2008) and OECD (2010). Regarding analyses of German unemployment durations, small regional differences could be found between both parts of Germany, for example, in Arntz and Wilke (2009) or Kauermann and Westerheide (2012). Beside the unemployment rate and the duration of unemployment mentioned above, the risk of getting unemployed is also of high interest. The latter is defined and analysed in different ways and in different contexts. Galiani and Hopenhayn (2003), for instance, analysed the risk of unemployment in Argentina between 1989 and 1998 making use of hazard models. Covizzi (2008) determined the unemployment risk of Swiss individuals concerning union dissolution, health, and gender with Cox proportional hazard models. Thapa (2004) and Arai and Vilhelmsson (2004) explored the unemployment risk of immigrants to natives in Australia and Sweden, respectively. Both used a logistic regression model. Hammer (1997) utilised logistic and Poisson regression models to investigate the unemployment risk of young Norwegian individuals. Fieldhouse (1996) looked at social and geographical factors to investigate the unemployment risk in Great Britain using logistic regression models after looking at factor-specific unemployment rates. Regarding the different papers analysing the risk of unemployment in Germany, we conclude that similar methods and topics are considered. Reinberg and Hummel (2002, 2003, 2005) used qualification-specific unemployment rates to analyse the unemployment risk in different educational groups in Germany. Arrow (1996) analysed the impact of health on the unemployment risk by using, amongst others, a Cox's proportional hazard model. Wilke (2004) analysed -beside the unemployment duration- the risk of unemployment given employment in Germany, that means he looked at the ratio of the number of individuals getting unemployed and the number of employed individuals in a defined period and compared the results with the unemployment rate. Lauer (2003) analysed the influence of education on the risk of getting unemployed and reemployed in a cross-national study with a discrete time competing risks hazard rate model based on the data of the German Socio-Economic Panel Study for Germany and the Emploi survey for France. Lurweg (2010) used a pooled logistic regression to analyse amongst others the impact of international trade on the risk of getting unemployed. Some of the papers mentioned above include regional information in their analysis as, for example, Fieldhouse (1996) who used geographical factors concerning different British regions or Thapa (2004) whose analysis contains Australian regions. However, none of the papers above that analyse the unemployment risk in Germany include regional or spatial information apart from differentiating between the old West German states and the newly formed German states, see Reinberg and Hummel (2002, 2003, 2005) or Lurweg (2010). In addition, often only data of the Old Länder is used for analysing the risk of unemployment in Germany, see for instance Wilke (2004) or Lauer (2003).

With our analysis we aim to contribute to the discussion in two aspects. The first contribution of our paper is to analyse the influence of different covariate effects -including individual, spatial, calendrical, and economic effects- on the unemployment risk in all of Germany between 2000 and 2004. Beside the analysis of the different unemployment risks, we want to compare our results with other research findings on unemployment risks and contrast these results with the conclusions of studies investigating the duration of unemployment or analyses interpreting unemployment rates. Our second contribution is to demonstrate how to use available software to model and easily fit an additive Poisson model with fixed grouped individual covariate effects and smooth dynamic covariate effects of spatial, calendrical, and economic information after restructuring the likelihood of a log-linear Poisson model.

As database we use the Scientific Use File 'Regional File 1975-2004' which is an administrative data set of the German Federal Employment Agency and provided by the Research Data Centre (Forschungsdatenzentum (FDZ)) at the Institute for Employment Research (Institut für Arbeitsmarkt- und Berufsforschung (IAB)). It contains information about the employment biographies of employees covered by social security and of benefit recipients in Germany on a day-to-day basis. Furthermore, it includes spatial information about 343 defined regions. The database is a 2% random sample out of the Employee and Benefit Recipient History of the IAB.

The statistical model being used for our analysis is built upon the log-linear Poisson model, see McCullagh and Nelder (1989). We allow for grouped covariates to simplify the model and to downsize the computational effort. Beside grouped covariates with individual information like gender, age, and education, we include smooth functional effects as proposed in Hastie and Tibshirani (1990) for generalized additive models. The additive Poisson model is fitted with software for generalized additive models in **R**, see R Development Core Team (2009) and Wood (2006).

The paper is organized as follows. Section 2 introduces the statistical model being used. Section 3 gives more detailed information of the database and the utilised covariates. In Section 4, a detailed data analysis is given before we draw our conclusions in Section 5.

6.2 Statistical Model

Let the random variables Y_{ti} denote the employment status of the ith individual in interval t, $t \in \{1, \dots, T\}$ with $i = 1, \dots, N_t$. With $Y_{ti} = 1$ we denote an individual which is unemployed in period t, but has been working in the previous period $t - 1$. Otherwise we set $Y_{ti} = 0$. In other words, $Y_{ti} = 1$ indicates individuals getting unemployed from period $t-1$ to period t. We assume that Y_{ti} mirror a Poisson process, that is Y_{ti} are independent and identically Poisson-distributed with intensity parameter $\lambda_{ti} = \exp(\eta_{ti})$. The linear predictor η_{ti} depends on a number of covariates x_{ti}, say, and a set of parameters θ to be specified later. Thus, the log-likelihood contribution for time point t can be written as (see McCullagh and Nelder, 1989)

$$l_t(\theta) = \sum_{i=1}^{N_t} \big[\underbrace{Y_{ti} \log(\lambda_{ti})}_{1} - \underbrace{\lambda_{ti}}_{2} \big]. \tag{6.1}$$

In our example the number of observations N_t at each time point is rather large, in the order of 500,000 observations, summing up to 29,978,674 observations for all time points. In contrast, the number of events, that are observations with $Y_{ti} = 1$, is comparably small, about 4,000 observations for each time point summing up to 237,507 observations for all time points. Hence, about 1% of the individuals become unemployed. To handle the data in a numerically efficient way, we restructure the likelihood by grouping observations with respect to their covariate values. Thus, we group the covariate age into $J = 4$ groups, the educational level is grouped into $L = 4$ categories and for gender we have $K = 2$ groups. Moreover, in our database we have $T = 60$ time intervals, each representing a month, which run from January 2000 to December 2004. Let now N_{tjkl} denote the total number of observations in the specified group categories and let n_{tjkl} be the number of events in age group j, $j = 1, \dots, J$, gender k, $k = 1, 2$, and educational group l, $l = 1, \dots, L$ in interval t, $t = 1, \dots, T$. Within the particular groups we assume that all individuals follow a homogeneous Poisson process. Let I_t be the index set of individuals becoming unemployed in t, that is $I_t = \{i : 1 \le i \le N_t, Y_{ti} = 1\}$. We define with $o_{ti} = \log(N_{tj_ik_il_i}/n_{tj_ik_il_i})$ the offset for $i \in I_t$, where j_i, k_i and l_i denote the category level of individual i. Then, the log-likelihood (6.1) can be simplified to

$$l_t(\theta) = \sum_{i \in I_t} \big[\log(\lambda_{ti}) - \lambda_{ti} \exp(o_{ti}) \big]. \tag{6.2}$$

Note that the likelihood now consists only of the individuals for which we observe the event of getting unemployed and hence it is numerically manageable. The implicit assumption is that covariates not included in the grouping have the same effect amongst all individuals in the groups. Beside the grouped covariates mentioned above we include further covariates in our model which are not on an individual level, but involve regional, calendrical, and economic information. These are the location of the former working place, the corresponding regional unemployment rate, the entry date into unemployment and the duration of employment at the last working place during the last year as well as the duration of unemployment during the last year. The effects of these covariates will be modeled by smooth functions while the grouped covariates will be included as fixed effects in our model. This leads us to a generalized additive model (see Hastie and Tibshirani, 1990 or Wood, 2006), more precisely an additive Poisson model. Let therefore the predictor $\eta_{ti} = \log(\lambda_{ti})$ in (6.2) take the form

$$\eta_{ti} = \boldsymbol{x}_{ti}^T\boldsymbol{\beta} + \gamma(r_{ti}) + \delta(t) + \phi(s_{ti}) + \zeta(c_{ti}) + \xi(u_{ti}), \tag{6.3}$$

where $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_p)^T$ are the parameters to be estimated and the corresponding covariates are $\boldsymbol{x}_{ti} = (1, x_{ti1}, \ldots, x_{tip})^T$, $r = 1, \ldots, p$. Moreover, $\gamma(.)$ is the smooth effect of the regional unemployment rate, $\delta(.)$ is the smooth calendrical effect of the start of unemployment, $\phi(.)$ describes a smooth spatial effect and finally $\zeta(.)$ and $\xi(.)$ specify the smooth effects of the duration of former employment and unemployment during the last year, respectively. Note that only the smooth calendrical effect $\delta(.)$ is dependent on t, the parametric and the remaining nonparametric effects are not directly influenced by different t. Further information concerning the model estimation can be found in the Appendix.

6.3 Data Description

For our analysis we use the Scientific Use File ‘Regional File 1975-2004’. A detailed description of the entire database is provided in Drews (2008). We use data from 5 years from January 2000 to December 2004 and analyse the risk of getting unemployed for 91625 men (146548 events in all time intervals out of 16715859 observations from 383769 men in all time intervals) and 66609 women (90959 events in all time intervals out of 13262815 observations from 317066 women in all time intervals) who became unemployed during the considered time. More information is shown in Table 6.1.

The covariate age is grouped into: up to 30 years, between 30 and 39 years (reference category), between 40 and 49 years, and 50 years of age and over. The educational background during the last period of employment is categorized into four levels: individuals

Year	Men		Women	
	Events	Observations	Events	Observations
2000	25718	3406426	16707	2623568
2001	25842	3426328	16719	2679894
2002	30863	3366198	18447	2685901
2003	32096	3283558	19724	2644449
2004	32029	3233349	19362	2629003
Σ	146548	16715859	90959	13262815

Table 6.1: Distribution of the events and the total amount of observations for the IAB 'Regional File 1975-2004' for men and women separated by the year the individual became unemployed.

without vocational training, individuals who attended a secondary general school or intermediate secondary school and completed successfully a vocational training (reference category), individuals with A-levels and with or without vocational training, and graduates from university or universities of applied science. The data set also contains local information with the region of the workplace. All in all, there are 343 defined regions in Germany in the data set. We use the centroid of the corresponding region as spatial information. Considering the economic environment, we include the unemployment rate of the individual's employment region in the year of entry into unemployment. The unemployment rates are based on annual rates of administrative districts in Germany on which the average rates of unemployment for the defined regions in the data set are calculated. As calendar time we use the date (month and year) when the individual became unemployed. The duration of former employment (in months) at the last working place during the last year before unemployment is included as well as the duration of unemployment (in months) during the last year before unemployment.

6.4 Data Analysis

We estimate a separated model for men and women and include interaction for the parametric effects between age and educational groups. The estimated intercept $\hat{\beta}_0$ differs slightly between the model for men ($\hat{\beta}_0 = -4.606$) compared to the model for women ($\hat{\beta}_0 = -4.987$), i.e. the risk of getting unemployed in the reference category (30-39 years old individuals with vocational training) is slightly lower for women compared to men. The estimated parametric effects $\hat{\beta}_r$ including the interactions are presented in Figure 6.1 and show a similar tendency for both models. However, the effects for women

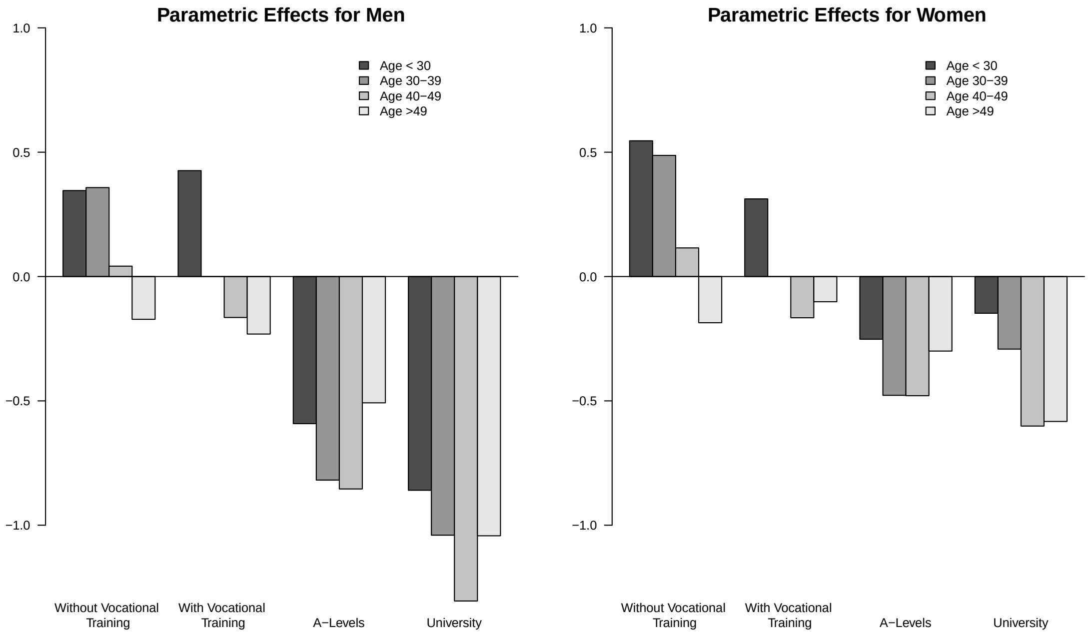

Figure 6.1: Estimated parametric effects $\hat{\beta}_r$ (including interaction) for getting unemployed for the IAB 'Regional File 1975-2004' for men and women, $r = 1, \ldots, p$.

compared to those for men vary less strongly, i.e. the different parametric effects for women do not influence the risk of getting unemployed as much as do the parametric effects for men. In Figure 6.2 we show the resulting fit of the smooth effects in Equation (6.3) which will be discussed later on. Surprisingly, the spatial effects are very small with a range of 0.0236 for men and a range of 0.0112 for women and hence omittable. Therefore, we do not show maps of the spatial fitted effects here. The model has been evaluated using an approximative Akaike Information Criterion (AIC) (see Wood, 2006, p. 230) and dropping effects from the model increased the AIC value. All effects are now discussed and interpreted in detail.

6.4.1 Parametric Effects

Effects for Men

Looking at the educational effects, it becomes clear that men with a higher education such as A-levels or an university degree have a lower risk of getting unemployed compared to individuals with a lower education in all age groups: the effects are negative for all age groups and much more pronounced than the other effects. Men with university degrees have the lowest risk to lose their jobs compared to all the other educational and age groups. Overall, graduates between 40 and 49 years have the strongest negative effect (-1.305). Looking only at men with vocational training, individuals from 40 years on have -compared to those between 30 and 39 years- a lower risk of getting unemployed. In contrast, individuals up to 30 years with vocational training lose their job faster than the older ones and overall have the highest risk to lose their job (0.426). Men without vocational training up to 49 years have a higher risk of getting unemployed than men above 49 years of the same educational level and those of the reference category (30-39 year old men with vocational training). With increasing education the risk of getting unemployed decreases in each age group, except for under 30-year-old individuals with vocational training. Men of this age group with vocational training have a slightly higher risk to lose their job compared to men without vocational training of the same age group. However the age effects within a certain educational level differ between the educational groups. In general, it can be said that a higher educational level reduces the risk of losing the job, while the age effects depend on the educational level.

Effects for Women

Women with a higher education, i.e. holding A-levels or an university degree, have a lower risk of getting unemployed compared to women with a lower education in the same age group. Comparing women with an university degree to women with A-levels, women up to 39 years have a slightly higher risk to lose their job while for women of an age of 40 years or older it is vice versa. Altogether, women with an university degree between 40 and 49 years have the lowest risk of getting unemployed (-0.603). Women with vocational training or a higher education in the age group between 40 and 49 years have the lowest risk to lose their job compared to the other age groups. Looking only at women without vocational training, women over 49 years have the lowest risk of getting unemployed compared to the other age groups. Women younger than 30 years have the highest risk of getting unemployed in all different educational groups while women without vocational training in this age group have the highest risk of all (0.545). Overall it can be seen that also for women a higher level of education reduces the risk of getting unemployed, but these effects do not behave as strict as it could be seen for men. Within each educational group the age effects act differently, similar to the observations for men.

Similar results concerning education and age were found in Reinberg and Hummel (2002, 2003, 2005) who analysed the unemployment rates in different qualification groups for Germany: higher educated employees have a distinct lower risk of getting unemployed than lower educated men and women in East and West Germany. This is also true for older employees with a higher education which have lower unemployment rates than younger less educated individuals. In 2004, higher educated employees between 55 and 64 years had the lowest unemployment rates compared to the younger age groups, see Reinberg and Hummel (2005). This result seems to stand in contrast to the unemployment behaviour in different age groups as analysed, for instance, in Hunt (1995), Hujer and Schneider (1995), Westerheide and Kauermann (2012a) or Kauermann and Westerheide (2012) where older unemployed generally had worse chances of getting reemployed. Note that this effect can be caused by different reasons and does not have to mirror worse labour market conditions for older unemployed. One of the reasons might be the possibility of a longer duration of unemployment benefits for older unemployed as argued in Hunt (1995), for more information about the German benefit scheme see for instance Clasen (2005) or European Commission (2005b). Another reason might be the possibility of early retirement for older unemployed and accordingly the usage of unemployment as passage between employment and retirement, see for further information Rein and Jacobs (1993), Knuth and Kalina (2002) or Fitzenberger and Wilke (2004).

Looking again at educational effects, Steiner and Schmitz (2010) concluded that an investment in education reduces the risk of unemployment. Wilke (2004) found out that on the one hand education has a high impact on a lower risk of unemployment especially for men, on the other hand he found only very small variation for women. Regarding personal characteristics, Lurweg (2010) observed in her analysis that an increase in education lowers the chance of getting unemployed. The results of Lauer (2003) concerning the risk of getting unemployed differ somewhat. She found out that individuals without vocational training have the highest risk of getting unemployed while individuals with vocational qualifications of an intermediate level have the lowest risk. University graduates have a higher risk of getting unemployed than individuals with an intermediate qualification level. In addition, she found out that women have a higher risk of getting unemployed in all educational groups compared to men. Generally, it can be inferred that a better education reduces the risk of getting unemployed. This result matches with our analysis. A higher education seems to have a positive impact on the individual's labour market conditions. Looking at educational effects in papers analysing the unemployment duration, similar results can be found. Typically, education improves the re-employment probability, see for instance Lauer (2003), Westerheide and Kauermann (2012a) or Kauermann and Westerheide (2012).

6.4.2 Smooth Effects

Effects for Men

Referring to the left panel in Figure 6.2, we find that the duration of unemployment during the last year, the unemployment rate of the region where the individual has worked before as well as the spatial information of the region of the former working place (not shown) have no effect on the risk of getting unemployed. The range of the spatial effect for men is only 0.0236 and we could not detect any difference between the old West German states and the newly formed German states. The duration of the former employment at the last working place during the last year has a weak effect on the risk of losing the job: for men who had been employed for up to 8 months there is a very low negative effect on the risk of getting unemployed while it is vice versa for men who worked for a longer duration. Individuals seem not to be employed for only some weeks. All in all, it seems that all these covariates do not strongly influence the risk of getting unemployed. Only the calendar time has a distinct effect and a regular pattern can be observed. The effect always shows a high peak in January and a lower peak in June/July, i.e. the risk of getting unemployed is the highest in January and is still more pronounced in June/July than in the surrounding months. During spring men have the lowest risk of losing their jobs. This effect behaves similar to observed unemployment data of other years in Germany, see Rudolph (1998) or Institut für Arbeitsmarkt- und Berufsforschung (2009). Beside these seasonal effects one can find a slightly decreasing trend between January 2000 and January 2002 and a slightly increasing trend of getting unemployed between January 2002 and December 2004. This trend goes along with the German business cycle, see Schirwitz (2009).

Effects for Women

We now look at the right hand panel of Figure 6.2. The effects of the duration of unemployment or employment during the last year, the unemployment rate of the region where the individual has worked before as well as the spatial information for women barely differ from the effects of men and are again negligible. The range of the spatial effect for women is even smaller (0.0112) compared to men. Only the effect of former employment is slightly more pronounced. The calendrical effect is distinct, but it does not show such a seasonal pattern as it could be seen for the calendrical effect for men. One can identify a peak during winter and summer over the observed period, but the peaks in wintertime are not so pronounced. Similarly, the trend of getting unemployed over the observed period is not so distinct for women.

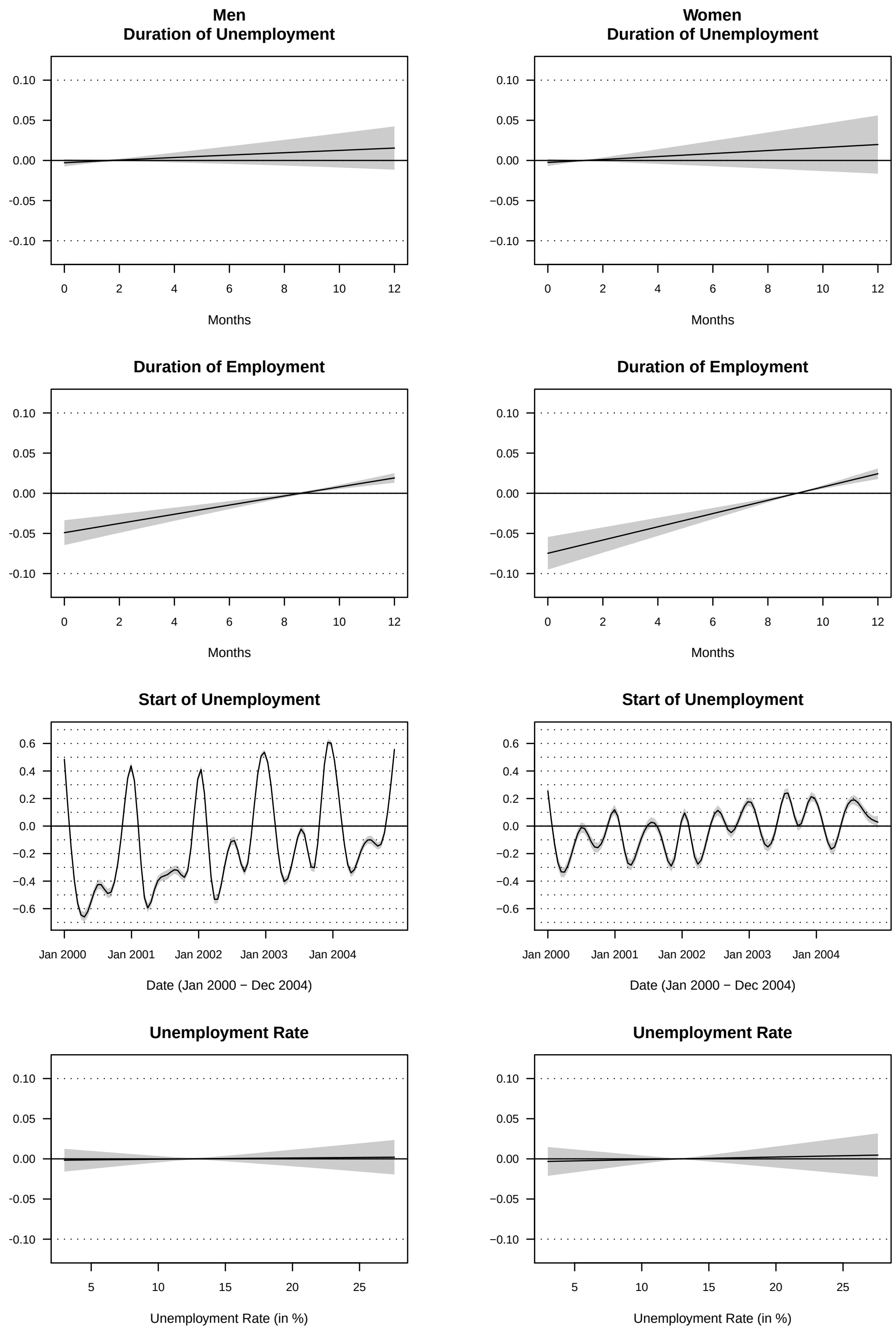

Figure 6.2: Fitted smooth effects with 95%-confidence intervals for the IAB 'Regional File 1975-2004' for men (left column) and women (right column).

It is remarkable that neither the unemployment rate nor the region of the former working place influence the individual's risk of losing the job. Generally, the unemployment rates differ between the old West German states and the newly formed German states, see Statistisches Bundesamt, Gesis-Zuma, and WZB (2008) and OECD (2010). Reinberg and Hummel (2005) found differences between the New and Old Länder concerning the unemployment rates in different qualification groups for Germany. Lurweg (2010) discovered a higher risk of unemployment for East German households compared to West German households. Looking at analyses of the unemployment duration, spatial effects as well as effects of the local unemployment rate are clearly found. Kauermann and Westerheide (2012), who explored the chance of getting reemployed in Germany using also the IAB 'Regional File 1975-2004', found out that these covariates have a significant influence on the individual's re-employment chances. Arntz and Wilke (2009) detected only small differences between the unemployment durations in West and East Germany. Analysing the unemployment duration in West Germany, Fahrmeir, Lang, Wolff, and Bender (2003) found spatial heterogeneity, too. In our analyses, though, we do not find an effect for the included spatial information and the regional unemployment rate, respectively, i.e. both effects do not influence the risk of losing the job. We may conclude, that spatial heterogeneity of the job market occurs not due to different chances of losing the job but of spatial heterogeneity with respect to finding a new job when being unemployed. We think that this is an interesting result of the paper.

6.5 Conclusion

In this paper we have analysed the risk of getting unemployed in Germany with an additive Poisson model. We studied fixed individual covariate effects of men and women of different age and educational groups as well as smooth flexible covariate effects of calendrical, economic or spatial information. Looking at the educational effects, we can conclude that differences in the level of education strongly influence the risk of getting unemployed. The higher the educational level, the lower the risk of unemployment. This rule is true for all but one educational effect of the different age groups respecting men and also for most educational effects of women. A higher education is one of the essentials to be successful on the labour market. This conclusion goes along with analyses of qualification-specific unemployment rates and unemployment durations as well as with other studies concerning the risk of unemployment. The risk of getting unemployed is lower for older better educated individuals (men and women) than for lesser educated younger individuals and can be found in analyses of age- and qualification-specific unemployment rates, too. Looking at the smooth flexible effects, only the calendrical effect has a high influence on the unemployment risk of both genders. For men a regular cyclical pattern can be seen with the highest risk in wintertime and the lowest risk during spring. This effect is associated with the seasonal unemployment rate. For women this pattern is not so regular, but similar weaker risks are visible. Regarding the smooth effect of the unemployment rate and the region of the former working place, there is no influence on the individual's risk of getting unemployed. These findings stand in contrast to conclusions made by analysing the unemployment rate or the unemployment duration, so it may be concluded from this analysis that the region and the local unemployment rate influence only the chance of finding a job but not the risk of losing a job.

Following Reinberg and Hummel (2002, 2003, 2005) that lower unemployment rates indicate a lower unemployment risk, we get similar results for educational effects but not for regional effects. Our analysis shows that it is not always sufficient to analyse pure unemployment rates or other macroeconomic measurements to gain information about the risk of getting unemployed. However, conclusions drawn from analyses of unemployment duration can also not be taken to make an impact on the individual's risk of unemployment. As it could be seen in our analysis, the usage of an additive Poisson model seems to be a good way to obtain more detailed information about the influence of covariate effects on the unemployment risk.

This chapter is based upon the following working paper:
Westerheide, N. and Kauermann, G. (2012, April): Getting Unemployed: Factors Influencing the Risk of Unemployment in Germany.

7 Conclusion

After a brief introduction in Chapter 1, a short overview of terms and topics regarding unemployment was given in Chapter 2. Chapter 3 covers a basic introduction of the statistical methods being used in the applications in Chapters 4, 5, and 6. In the main part of this thesis, i.e. in Chapters 4, 5, and 6, three applications for longitudinal unemployment data based on different spline-based models were presented. Thereby, the flexibility and capacity of penalized spline smoothing as estimation routine for longitudinal data could be shown for all examples of use, and its easy application to free available statistical software for generalized additive models was demonstrated.

Taking a look at the results of the analyses of different effects on the duration of unemployment and on the risk of unemployment, we may conclude the following: In the context of the analyses concerning the duration of unemployment in Chapters 4 and 5 worse re-employment probabilities were found for elderly unemployed in Germany but more favourable re-employment chances for men and better educated individuals. Additionally, it was shown in Chapter 4 that former unemployment degrades the re-employment chances, especially in the United Kingdom. Moreover, individuals living as a couple have better re-employment chances in the United Kingdom. Furthermore, in Chapter 5 it was demonstrated that the age and educational effects differ between the job markets in the three considered distances as well as in the regional and temporal effects. For individuals over 54 years, the negative effect intensifies with the duration of unemployment. In particular, during the first year better educated unemployed have better re-employment chances at a distance of more than 50 km from their former working place while individuals between 35 and 44 years have worse re-employment chances at the same distance. Unemployed with a former working place in the newly formed German states have far better re-employment probabilities at a distance between 50 and 150 km in comparison with the other considered distances and unemployed who have been working in the Old Länder. This distance also has the most remarkable spatial pattern. For a distance of up to 50 km, a cyclic temporal effect, denoting that during winter time unemployed have better re-employment chances, is clearly visible. Moreover, in regions with a lower local unemployment rate better re-employment chances could be found for a distance of up to 150 km from the former working place. At around 180 days, the baseline effect for distances above 50 km shows a peak similar to the effect for women in a distance of up to 50 km.

Relating the results of these two analyses to the job search theory presented in Chapter 2.3, the following can be concluded: Although the decreasing chance for elderly unemployed Germans to get reemployed might be justified rather by a passage via unemployment to early retirement than by a longer duration of unemployment compensation, this effect can not be explained through the job search theory, though other effects support this theory: Similar to the results of Katz and Meyer (1990), increasing chances to get reemployed were found in some cases at the end of the length of entitlement to unemployment benefits for the baseline and for women in the analysis in Chapter 5. These effects might be caused by more willingness of the individual and consequently an increased probability to accept a job due to the absence of benefits or their subsequent reduction, i.e. the lower rate of unemployment assistance compared to the unemployment benefits.

Throwing a glance at the outcomes of the analysis of various effects on the risk of getting unemployed in Chapter 6, the following can be subsumed: the higher the educational level, the lower the unemployment risk. Thus, a higher education seems to be a key asset to be successful on the labour market. This outcome goes along with the results of the two analyses concerning the duration of unemployment mentioned above and analyses of qualification-specific unemployment rates, see also Chapter 2.1. Furthermore, it can be concluded that for both genders older better educated individuals have a lower risk of getting unemployed than younger lesser educated individuals. In addition, for men a regular cyclical pattern going along with the seasonal unemployment rate is visible where the highest risk of unemployment is during winter and the lowest risk during spring. For women the cyclical pattern is not so regular, but similar weaker effects can be seen. In contrast, the local unemployment rate and the region of the former working place have no effects on the risk of unemployment. These findings are in opposition to the outcomes of the analyses concerning unemployment durations or local unemployment rates. Finally, it may be concluded from the analyses in Chapter 5 and 6 that the region and the local unemployment rate influence only the chance of finding a job, but not the risk of losing a job.

Apparently, various statements for the German and British labour market could be made utilizing the introduced spline-based models and their application to longitudinal unemployment data. The remarkable and extensive possibilities of interpretation of the models' smooth functional effects and spatial effects are an enrichment for analyses in labour market research. Furthermore, it could be shown that the approaches are suitable for massive databases. Hence, further applications of the presented models in other settings or fields of research seem to be inviting due to the easy realization with free available statistical software after some data management.

A Appendix

A.1 R-Code

The function `survival.to.poisson.simpson` (R-Code) converts survival data to poisson distributed data with a simpson approximation. The following arguments are necessary:

- `time`: duration time
- `status`: observation censored? 1: no or 0: yes
- `x`: data frame containing all covariates
- `number.int`: number of integration points for simpson approximation/ number of borders of subintervals, default: 5 ($T_0 = 0$ and $T_R = t_i$, with $R = 4$)

R-Code

```
survival.to.poisson.simpson <- function(time = time, status = status,
        x = NULL, number.int=5) {

        T <- lapply(seq(time), FUN = function(i)
                   (round(quantile(c(0,time[i]), prob = seq(0, 1,
                   length = number.int)), max(time + 1))))

        Yt.list <- lapply(seq(time), FUN = function(i)
                   c(rep(0,2*(number.int-1)), status[i]))

        t.minus.1 <- lapply(seq(time), FUN = function(i)
                     T[[i]][-1])
        t.minus.time <- lapply(seq(time), FUN = function(i)
```

```
             T[[i]][-number.int])

m <-lapply(seq(time), FUN = function(i)
           (t.minus.1[[i]]+t.minus.time[[i]])/2)

T.m <- lapply(seq(time), FUN = function(i)
       sort(c(T[[i]],m[[i]])))

t.plus <- lapply(seq(time), FUN = function(i)
       c(0,T[[i]],time[i]))

o.ki.1 <- lapply(seq(time), FUN = function(i)
  log(((t.plus[[i]][-c(1,2)])-
  (t.plus[[i]][-c((length(t.plus[[i]])-1):
  (length(t.plus[[i]])))]))/6))

o.ki.2 <- lapply(seq(time), FUN = function(i)
  log(4*(t.minus.1[[i]]-t.minus.time[[i]])/6))

insert <- function(a,b){
  res <- rbind (a,c(b,0))
  dim(res) <-NULL
  res[-length(res)]
  }

o.ki.12 <- lapply(seq(time), FUN = function(i)
                  insert(o.ki.1[[i]],o.ki.2[[i]]))

index <- lapply(seq(time), FUN = function(i)
                rep(i,length(T.m[[i]])))

x.multi <- lapply(seq(time), FUN = function(i)
                  kronecker(matrix(1,length(T.m[[i]]), 1),
                  t(matrix(as.matrix(x[i, ])))))
x.comb  <- eval(parse(text = paste("rbind(", paste("x.multi[[",
    1:length(x.multi), "]]", sep = "", collapse = ","), ")",
```

```
            sep = "")))
        x.data <- as.data.frame(x.comb)
        names(x.data) <- names(x)


        return(data.frame(index = unlist(index), Y = unlist(Yt.list),
               x.data,grid = unlist(T.m), offset = unlist(o.ki.12)))
    }
```

Example

```
# to provide the code in R:
source("survival.to.poisson.simpson.r")

# required data:
time <- c(19,36,1)
stat <- c(1,0,1)
x.datax <- data.frame(sex=c(1,0,1), edu= c(1,5,1), agegp=c(2,3,1),
                      area=c(123,586,354))

# usage of the function:
survival.to.poisson.simpson(time=time, status=stat, x=x.datax,
                            number.int=5)
```

A.2 Estimation

First, we describe the fitting of the smooth, functional components in (6.3). The unknown functions are represented by a linear combination of thin plate spline basis terms, see Wahba (1990, pp. 30-34) with the popular cubic regression spline basis resulting as special case, see Wood (2006). This approach is used for all smooth functions except for the spatial effect $\phi(.)$. For the smooth function $\phi(.)$ we use thin plate regression splines, see Wood (2006). We now replace the functional components in (6.3) by

$$\begin{aligned}&\gamma(r) = B_\gamma(r)b_\gamma, \quad \delta(t) = B_\delta(t)b_\delta, \quad \phi(s) = B_\phi(s)b_{\phi k}\\&\zeta(c) = B_\zeta(c)b_\zeta, \quad \xi(u) = B_\xi(u)b_\xi\end{aligned} \tag{A.1}$$

with $B(.)$ as spline bases. We follow Hastie (1996) and Wood (2003) and use so-called 'low rank smoothing', i.e. each function works with a reduced set of knots. This set of knots is still large enough to capture the functional shape but small enough to guarantee feasible computation. This concept has been characterized by Eilers and Marx (1996) as P(enalized)-spline smoothing, see also Ruppert, Wand, and Carroll (2003, 2009). The number of knots is denoted with q. Following Wood (2006, p. 161), we set $q = 30$ for the calendar effect and $q = 60$ for the spatial effect functions, respectively. For the remaining smooth functions we set $q = 10$. The model was also fitted for larger values of q but the choice of q has only small influence on the fit, see also Ruppert (2002) or Kauermann and Opsomer (2011). Suppose that $(\boldsymbol{x}_{ti}, r_{ti}, t, s_{ti}, c_{ti}, u_{ti})$ denote the observations for the ith individual in interval t, where $i \in I_t$, that is individual i becomes unemployed in period t. Assuming that the individuals are independent the log-likelihood in (6.2) for parameter vector $\boldsymbol{\theta} = (\beta_0^T, \boldsymbol{\beta}_x^T, b_\gamma^T, b_\delta^T, b_\phi^T, b_\zeta^T, b_\xi^T)^T$ with $\boldsymbol{\beta}_x^T = (\beta_r^T,\ r = 1, \ldots, p)$ can be expressed for all t as $l(\boldsymbol{\theta}) = \sum_{i \in I_t} \sum_{t=1}^T l_{ti}(\boldsymbol{\theta})$ where

$$\begin{aligned}l_{ti}(\boldsymbol{\theta}) = Y_{ti}\big[&\boldsymbol{x}_{ti}^T\boldsymbol{\beta} + B_\gamma(r_{ti})b_\gamma + B_\delta(t)b_\delta + B_\phi(s_{ti})b_\phi\\&+ B_\zeta(c_{ti})b_\zeta + B_\xi(u_{ti})b_\xi\} - \exp\{\boldsymbol{x}_{ti}^T\boldsymbol{\beta} + B_\gamma(r_{ti})b_\gamma\\&+ B_\delta(t)b_\delta + B_\phi(s_{ti})b_\phi + B_\zeta(c_{ti})b_\zeta + B_\xi(u_{ti})b_\xi + o_{ti}\big]\end{aligned} \tag{A.2}$$

Next, we establish a penalty on the spline coefficients to obtain a smooth functional fit. The model is high dimensional which implies that the Maximum-Likelihood estimate will produce wiggled fitted curves. Hence, we use a penalty on the coefficients as described in Eilers and Marx (1996) and Ruppert, Wand, and Carroll (2003). Following Wand and Ormerod (2008), we rewrite the spline representation in (A.1) by extracting the intercept and the linear slope, i.e.

$$\gamma(r) = B(r)b_\gamma = \beta_{\gamma_0} + r\beta_{\gamma_1} + \tilde{B}_\gamma(r)\tilde{b}_\gamma \tag{A.3}$$

where $\tilde{B}_\gamma(r)$ is the reduced rank basis with intercept and linear slope extracted. For $\delta(t)$, $\phi(s)$ $\zeta(c)$ and $\xi(u)$, we receive the reduced basis matrices $\tilde{B}_\delta(t)$, $\tilde{B}_\phi(s)$, $\tilde{B}_\zeta(c)$ and $\tilde{B}_\xi(u)$. In the following a quadratic penalty on the spline coefficient is imposed, e.g. $\lambda_\gamma \tilde{b}_\gamma^T \tilde{D}_\gamma \tilde{b}_\gamma$. It can be demonstrated that it is equivalent to penalize with squared second order derivatives of the function (see O'Sullivan, 1986 or Wahba, 1990), or second (or higher) order differences of the spline coefficient b_γ (see Eilers and Marx, 1996). Here we make use of derivatives to penalize because this approach is implemented in the software we use for fitting the data (see end of this section). The parameter λ_γ is thereby a smoothing parameter which leads to a linear fit with $\lambda_\gamma \to \infty$. This yields to the penalized log-likelihood

$$\begin{aligned} l(\boldsymbol{\beta}, \tilde{\boldsymbol{b}}, \boldsymbol{\lambda}) = & \sum_{i \in I_t} \sum_{t=1}^{T} \tilde{l}_{ti}(\boldsymbol{\beta}, \tilde{\boldsymbol{b}}) - \frac{1}{2}\lambda_\gamma \tilde{b}_\gamma^T \tilde{D}_\gamma \tilde{b}_\gamma - \frac{1}{2}\lambda_\delta \tilde{b}_\delta^T \tilde{D}_\delta \tilde{b}_\delta \\ & - \frac{1}{2}\lambda_\phi \tilde{b}_\phi^T \tilde{D}_\phi \tilde{b}_\phi - \frac{1}{2}\lambda_\zeta \tilde{b}_\zeta^T \tilde{D}_\zeta \tilde{b}_\zeta - \frac{1}{2}\lambda_\xi \tilde{b}_\xi^T \tilde{D}_\xi \tilde{b}_\xi \end{aligned} \tag{A.4}$$

with $\tilde{l}_{ti}$ as log-likelihood for the Poisson distributed variables and $\boldsymbol{\beta} = \big(\beta_0, \beta_1, \ldots, \beta_p; (\beta_{0\gamma}, \beta_{1\gamma}), (\beta_{0\delta}, \beta_{1\delta}), (\beta_{0\phi}, \beta_{1\phi}), (\beta_{0\zeta}, \beta_{1\zeta}), (\beta_{0\xi}, \beta_{1\xi})\big)^T$, analogous definition for $\tilde{\boldsymbol{b}}$, and obvious definition for $\boldsymbol{\lambda} = (\lambda_\gamma, \lambda_\delta, \lambda_\phi, \lambda_\zeta, \lambda_\xi)^T$. The penalized log-likelihood can be fitted with standard software for generalized additive models, see Hastie and Tibshirani (1990). The only additional step which has to be done before modelling the data is to group the data to calculate the offsets. This can be easily done with simple data management as described above. For fitting our data we use the `bam()` procedure in **R** of the package `mgcv`. This procedure extends the `gam()` procedure and is helpful when working with large data sets, see Wood (2010). The smoothing parameters $\boldsymbol{\lambda}$ can be selected using a generalized cross validation which is embedded in the `bam()` procedure. We made use of REML estimation which is also implemented in this procedure. In the end, the inference for the model can be drawn. We follow thereby standard asymptotic arguments as presented in Ruppert, Wand, and Carroll (2003), Wood (2006) or Kauermann, Krivobokova, and Fahrmeir (2009). Denoting with $\boldsymbol{\theta} = (\boldsymbol{\beta}^T, \tilde{\boldsymbol{b}}^T)^T$ the complete parameter vector, the Fisher matrix can be determined with $F(\boldsymbol{\theta}, \boldsymbol{\lambda})$ and it can be demonstrated that

$$\mathrm{Var}(\hat{\boldsymbol{\theta}}) = F^{-1}(\boldsymbol{\theta}, \boldsymbol{\lambda}) F(\boldsymbol{\theta}, \boldsymbol{\lambda} = 0) F^{-1}(\boldsymbol{\theta}, \boldsymbol{\lambda}),$$

see for further information e.g. Ruppert, Wand, and Carroll (2003).

References

Abbring, J. H. and G. J. Van den Berg (2007). The unobserved heterogeneity distribution in duration analysis. *Biometrika 94*(1), 87–99.

Arai, M. and R. Vilhelmsson (2004, July). Unemployment-risk differentials between immigrant and native workers in Sweden. *Industrial Relations 43*(3), 690–698.

Arntz, M. (2005). The Geographical Mobility of Unemployed Workers: Evidence from West Germany. ZEW Discussion Papers 05-34, ZEW - Zentrum für Europäische Wirtschaftsforschung / Center for European Economic Research.

Arntz, M. and R. Wilke (2009). Unemployment duration in Germany: Individual and regional determinants of local job finding. *Regional Studies 43*(1), 43–61.

Arrow, J. (1996). Estimating the influence of health as a risk factor on unemployment: A survival analysis of employment durations for workers surveyed in the german socio-economic panel (1984 -"1990). *Social Science & Medicine 42*(12), 1651 – 1659.

Begum, N. (2004, April). Characteristics of the short-term and long-term unemployed. *Labour Market Trends 112*, 139–144.

Bergmann, J. M. (Ed.) (2012). *Handlexikon der Europäischen Union*. Baden-Baden: Nomos.

Blanchard, O. (2006). European unemployment: the evolution of facts and ideas. *Economic Policy 21*(45), 5–59.

Bover, O., M. Arellano, and S. Bentolila (2002). Unemployment duration, benefit duration and the business cycle. *Economic Journal 112*(479), 223–265.

Bundesagentur für Arbeit (2011). *Sondernummer der Amtlichen Nachrichten der Bundesagentur für Arbeit (ANBA): Arbeitsmarkt 2010 Arbeitsmarktanalyse für Deutschland, West- und Ostdeutschland*. Nuremberg: Bundesagentur für Arbeit.

Bundesministerium für Arbeit und Sozialordung (Ed.) (1995). *Übersicht über das Sozialrecht*. Bonn: Bundesministerium für Arbeit und Sozialordung.

Bundesministerium für Arbeit und Sozialordung (Ed.) (1997). *Übersicht über das Sozialrecht*. Bonn: Bundesministerium für Arbeit und Sozialordung.

Bundesministerium für Gesundheit und Soziale Sicherung (Ed.) (2005). *Übersicht über das Sozialrecht*. Nürnberg: BW Bildung u. Wissen.

Böheim, R. and M. P. Taylor (2000, July). Unemployment duration and exit states in Britain. CEPR Discussion Paper No. 2500, Centre for Economic Policy Research, London.

Cai, T., R. Hyndman, and M. Wand (2002). Mixed model-based hazard estimation. *Journal of Computational and Graphical Statistics 11*, 784 – 798.

Cai, Z. and Y. Sun (2003). Local linear estimation for time-dependent coefficients in Cox's regression models. *Scandinavian Journal of Statistics 30*, 93 – 111.

Clasen, J. (2005). *Reforming European Welfare States*. Oxford: Oxford University Press.

Clasen, J., J. Davidson, H. Ganßmann, and A. Mauer (2006). Non-employment and the welfare state: the United Kingdom and Germany compared. *Industrial Law Journal 16*(2), 134–154.

Collett, D. (1996). *Modelling survival data in medical research*. London: Chapman & Hall.

Collier, W. (2005). Unemployment duration and individual heterogeneity: a regional study. *Applied Economics 37*(2), 133–154.

Covizzi, I. (2008). Does union dissolution lead to unemployment? a longitudinal study of health and risk of unemployment for women and men undergoing separation. *European Sociological Review 24*(3), 347–361.

Cox, D. R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society, Series B 34*, 187–220.

Cox, D. R. and D. Oakes (1984). *Analysis of Survival Data*. London: Chapman and Hall.

de Boor, C. (1978). *A Practical Guide to Splines*. Berlin: Springer.

de Boor, C. (2001). *A Practical Guide to Splines, Revised Edition*. Berlin: Springer.

Devine, T. J. and N. M. Kiefer (1993). The empirical status of job search theory. *Labour Economics 1*(1), 3 – 24.

Dierckx, P. (1993). *Curve and surface fitting with splines*. Oxford: Clarendon Press.

Dockery, A. M. (2000). Regional unemployment rate differentials and mobility of the unemployed: An analysis of the facs longitudinal data set. *International Journal of Manpower 21*(5), 400–423.

Drews, N. (2008). Das Regionalfile der IAB-Beschäftigtenstichprobe 1975-2004: Handbuch-Version 1.0.2. FDZ Datenreport. Documentation on Labour Market Data 2008/02(DE), Institut für Arbeitsmarkt- und Berufsforschung (IAB), Nürnberg [Institute for Employment Research, Nuremberg, Germany].

Détang-Dessendre, C. (1999). Reciprocal link between exit from unemployment and geographical mobility. *Environment and Planning A 31*(8), 1417–1431.

Ehrenberg, R. G. and R. S. Smith (2012). *Modern Labor Economics: Theory and Public Policy*. Boston: Prentice Hall.

Eilers, P. H. C. and B. D. Marx (1996). Flexible smoothing with B-splines and penalties. *Statistical Science 11*(2), 89–121.

Eilers, P. H. C. and B. D. Marx (2010). Splines, knots, and penalties. *Wiley Interdisciplinary Reviews: Computational Statistics 2*(6), 637–653.

Eubank, R. (1999). *Nonparametric regression and spline smoothing*. New York: Dekker.

European Commission (2005a). *MISSOC - Social protection in the Member States or the European Union, of the European Economic Area and in Switzerland - Situation on 1 January 2005*. Luxembourg: Office for Official Publications of the European Communities. http://ec.europa.eu/employment_social/missoc/missoc2_en.pdf, http://ec.europa.eu/employment_social/missoc/missoc4_en.pdf.

European Commission (2005b). *MISSOC - Social protection in the Member States or the European Union, of the European Economic Area and in Switzerland - Situation on 1 May 2004*. Luxembourg: Office for Official Publications of the European Communities. ISBN 92-894-8469-1, http://ec.europa.eu/employment_social/missoc/missoc2004_may_en.pdf.

European Commission (2009). *Employment in Europe 2009*. Luxembourg: Office for Official Publications of the European Communities.

European Commission (2010, December). The european employment strategy: Working to improve employment in Europe. Technical Report KE-31-10-900-EN-C, European Commission. ISBN 978-92-79-17844-3, http://ec.europa.eu/social/BlobServlet?docId=6416&langId=en, Last checked: 03/11/2012.

European Commission (2011a). *Europe in figures - Eurostat yearbook 2011.* Luxembourg: Publications Office of the European Union.

European Commission (2011b). *Eurostat regional yearbook 2011.* Luxembourg: Publications Office of the European Union.

Eurostat (2007). *Europe in Figures - Eurostat Yearbook 2006-07.* Luxembourg: Office for Official Publications of the European Communities.

Eurostat (2009). *Europe in Figures - Eurostat Yearbook 2009.* Luxembourg: Office for Official Publications of the European Communities.

Fahrmeir, L., T. Kneib, and S. Lang (2004). Penalized structured additive regression for space-time data: A bayesian perspective. *Statistica Sinica 14*(3), 731–761.

Fahrmeir, L., T. Kneib, and S. Lang (2009). *Regression.* Berlin: Springer.

Fahrmeir, L., S. Lang, J. Wolff, and S. Bender (2003). Semiparametric bayesian time-space analysis of unemployment duration. *Allgemeines Statistisches Archiv 87*, 281–207.

Fahrmeir, L. and G. Tutz (2001). *Multivariate Statistical Modelling Based on Generalized Linear Models* (2 ed.). New York: Springer Verlag.

Fan, J., I. Gijbels, and M. King (1997). Local likelihood and local partial likelihood in hazard regression. *Annals of Statist. 25*, 1661–1690.

Fieldhouse, E. A. (1996). Putting unemployment in its place: Using the samples of anonymized records to explore the risk of unemployment in Great Britain in 1991. *Regional Studies 30*(2), 119–133.

Fitzenberger, B. and R. A. Wilke (2004). Unemployment durations in West-Germany before and after the reform of the unemployment compensation system during the 1980s. ZEW Discussion Papers 04-24, ZEW - Zentrum für Europäische Wirtschaftsforschung / Center for European Economic Research.

Fitzenberger, B. and R. A. Wilke (2007, February). New insight on unemployment duration and post unemployment earnings in germany: Censored box-cox quantile regression at work. IZA Discussion Papers No. 2609, Institute for the Study of Labor (IZA), Bonn.

Fitzgerald, T. J. (1998). An introduction to the search theory of unemployment. *Economic Review* (Q III), 2–15.

Franz, W. (1996). *Arbeitsmarktökonomik.* Berlin: Springer.

Franz, W. (1999). *Arbeitsmarktökonomik.* Berlin: Springer.

Franz, W. (2006). *Arbeitsmarktökonomik*. Berlin: Springer.

Franz, W. (2009). *Arbeitsmarktökonomik*. Berlin: Springer.

Freund, R. W. and R. H. W. Hoppe (2007). *Stoer/Bulirsch: Numerische Mathematik 1*. Berlin, Heidelberg: Springer.

Galiani, S. and H. A. Hopenhayn (2003, June). Duration and risk of unemployment in Argentina. *Journal of Development Economics 71*(1), 199–212.

Gautschi, W. (1997). *Numerical Analysis: An Introduction*. Boston: Birkhäuser.

Gazeboom, H. B. and D. J. Treiman (2003). Three internationally standardised measures for comparative research on occupational status. In J. H. Hoffmeyer-Zlotnik and C. Wolf (Eds.), *Advances in Cross-National Comparison: A European Working Book for Demographic and Socio-Economic Variables*, Chapter 9. New York: Kluwer Academic Publishers.

Gil, A., J. Segura, and N. M. Temme (2007). *Numerical methods for special functions*. Philadelphia, Pa.: SIAM.

Goldthorpe, J. H. (1987). *Social mobility and class structure in modern Britain*. Oxford: Clarendon Pr.

Grambsch, P. and T. Therneau (2003). *Modeling Survival Data: Extending the Cox Model*. Springer.

Gray, R. J. (1994). Spline-based tests in survival analysis. *Biometrics 50*, 640–652.

Haisken-DeNew, J. P. and J. R. Frick (Eds.) (2005). *DTC - Desktop Companion to the German Socio-Economic Panel Study (SOEP)* (8 ed.). Berlin: Deutsches Institut für Wirtschaftsforschung.

Halpin, B. (2006, November). British household panel survey combined work-life history data, 1990-2005 (SN: 3954) [computer file]. 5th Edition. University of Essex. Institute for Social and Economic Research, Colchester, Essex: UK Data Archive, http://dx.doi.org/10.5255/UKDA-SN-3954-1.

Hammer, T. (1997). History dependence in youth unemployment. *European Sociological Review 13*(1), pp. 17–33.

Hanefeld, U. (1987). *Das Sozio–ökonomische Panel: Grundlagen und Konzeption*. Frankfurt/Main: Campus Verlag.

Hastie, T. (1996). Pseudosplines. *J. Roy. Statist. Soc. Ser. B 58*, 379–396.

Hastie, T. and R. Tibshirani (1990). *Generalized Additive Models*. London: Chapman and Hall.

Hastie, T. and R. Tibshirani (1993). Varying–coefficient models. *J. Roy. Statist. Soc. Ser. B 55*, 757–796.

Heckman, J. and B. Singer (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica 52*(2), 271–320.

Heer, B. (2006). Should unemployment benefits be related to previous earnings? *FinanzArchiv: Public Finance Analysis 62(4)*, 530–550.

Hess, K. R. (1994). Assessing time-by-covariate interactions in proportional hazards regression models using cubic spline functions. *Statistics in Medicine 13*, 1045–1062.

Hujer, R. and H. Schneider (1995). Institutionelle und strukturelle Determinanten der Arbeitslosigkeit in Westdeutschland: Eine mikroökonometrische Analyse mit Paneldaten. In B. Gahlen, H. Hesse, and H. J. Ramser (Eds.), *Arbeitslosigkeit und Möglichkeiten ihrer Überwindung*, Volume 25 of *Wirtschaftswissenschaftliches Seminar Ottenbeuren*, pp. 53–76. Tübingen: J. C. B. Mohr.

Hunt, J. (1995). The effect of unemployment compensation on unemployment duration in Germany. *Journal of Labor Economics 13*(1), 88–120.

Institut für Arbeitsmarkt- und Berufsforschung (Ed.) (2009). *IAB-Jahresbericht 2008.* Nürnberg: Institut für Arbeitsmarkt- und Berufsforschung (IAB) der Bundesagentur für Arbeit.

Jackman, R. and R. Layard (1991). Does long-term unemployment reduce a person's chance of a job? A time-series test. *Econometrica 58*(229), 93–106.

Jacobi, L. and J. Kluve (2006). Before and after the Hartz Reforms: The performance of active labour market policy in Germany. RWI: Discussion Papers 41, Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI), Essen.

Kaiser, L. C. and T. Siedler (2001). Die Dauer von Arbeitslosigkeit in Deutschland und Großbritannien: Ein internationaler Vergleich (1990-1995). *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung 34*, 402–418.

Kalbfleisch, J. D. and R. L. Prentice (2002). *The statistical analysis of failure time data.* Hoboken, NJ: Wiley-Interscience.

Katz, L. F. and B. D. Meyer (1990). Unemployment insurance, recall expectations, and unemployment outcomes. *The Quarterly Journal of Economics 105*(4), 973–1002.

Kauermann, G. (2004). A note on smoothing parameter selection for penalised spline smoothing. *Journal of Statistical Planing and Inference 127*, 53–69.

Kauermann, G. (2005). Penalised spline smoothing in multivariable survival models with varying coefficients. *Computational Statistics and Data Analysis 49*, 169–186.

Kauermann, G. (2006). Nonparametric models and their estimation. *Allgemeines Statistisches Archiv 90*(1), 137–152.

Kauermann, G. (2010). Penalized splines, mixed models and Bayesian ideas. In T. Kneib and G. Tutz (Eds.), *Statistical Modelling and Regression Structures*, pp. 45–58. Heidelberg: Physica-Verlag.

Kauermann, G., T. Krivobokova, and L. Fahrmeir (2009). Some asymptotic results on generalized penalized spline smoothing. *Journal of the Royal Statistical Society, Series B 71*(2), 487–503.

Kauermann, G. and J. Opsomer (2011). Data-driven selection of the spline dimension in penalized spline regression. *Biometrika 98*(1), 225–230.

Kauermann, G. and N. Westerheide (2012). To move or not to move to find a new job: spatial duration time model with dynamic covariate effects. *Journal of Applied Statistics 39*(5), 995–1009.

Kettunen, J. (2002). Labour mobility of unemployed workers. *Regional Science and Urban Economics 32*(3), 359 – 380.

Kiyotaki, N. and R. Wright (1993). A search-theoretic approach to monetary economics. *The American Economic Review 83*(1), 63–77.

Klein, J. P. (1992). Semiparametric estimation of random effects using the Cox model based on the EM algorithm. *Biometrics 48*(3), 795–806.

Klein, J. P. and M. L. Moeschberger (2003). *Survival analysis*. New York: Springer.

Kneib, T. (2006). Mixed model-based inference in geoadditive hazard regression for interval censored survival times. *Computational Statistics and Data Analysis 51*(2), 777–792.

Kneib, T. and L. Fahrmeir (2007). A mixed model approach for geoadditive hazard regression. *Scandinavian Journal of Statistics 34*, 207–228.

Knuth, M. and T. Kalina (2002). Early exit from the labour force between exclusion and privilege: Unemployment as a transition from employment to retirement in West Germany. *European Societies 4 (4)*, 393–418.

Kooperberg, C., C. Stone, and Y. Troung (1995). Hazard regression. *Journal of the American Statistcal Association 90*, 78–94.

Krivobokova, T. (2006). *Theoretical and practical aspects of penalized spline smoothing.* Bielefeld (Germany): Bielefeld University. http://pub.uni-bielefeld.de/publication/2301835.

Kuhlenkasper, T. and G. Kauermann (2010). Duration of maternity leave in Germany: A case study of nonparametric hazard models and penalized splines. *Labour Economics 17*(3), 466–473.

Lampert, H. (1996). *Lehrbuch der Sozialpolitik.* Berlin: Springer.

Lancaster, T. (1990). *The econometric analysis of transition data.* Econometric Society monographs ; 17. Cambridge Univ. Pr.

Lauer, C. (2003). Education and unemployment : A French-German comparison. ZEW Discussion Papers 03-34, ZEW - Zentrum für Europäische Wirtschaftsforschung / Center for European Economic Research.

Lawless, J. F. (2003). *Statistical models and methods for lifetime data.* New York: Wiley.

Layard, P. R. G., S. J. Nickell, and R. Jackman (2006). *Unemployment.* Oxford: Oxford Univ. Press.

Layard, P. R. G., S. J. Nickell, and R. Jackman (2009). *Unemployment.* Oxford: Oxford Univ. Press.

Lippman, S. A. and J. J. McCall (1976a). The economics of job search: A survey: Part i. *Economic Inquiry 14*(2), 155–189.

Lippman, S. A. and J. J. McCall (1976b). The economics of job search: A survey: Part ii. *Economic Inquiry 14*(3), 347–368.

Ljungqvist, L. and T. J. Sargent (1998). The European unemployment dilemma. *Journal of Political Economy 106*(3), 514–550.

Lurweg, M. (2010, May). Perceived job insecurity, unemployment risk and international trade - a micro-level analysis of employees in german service industries. SOEPpapers 300, DIW Berlin, The German Socio-Economic Panel (SOEP).

Lynn, P. (2006, February). Quality profile: British household panel survey version 2.0: Waves 1 to 13: 1991-2003. Technical report, Institute for Social and Economic Research, University of Essex. http://www.iser.essex.ac.uk/files/bhps/quality-profiles/BHPS-QP-01-03-06-v2.pdf (Download: 01/31/2012).

Lüdemann, E., R. A. Wilke, and X. Zhang (2006). Censored quantile regressions and the length of unemployment periods in West Germany. *Empirical Economics 31*(4), 1003–1024.

Machin, S. and A. Manning (1999). The causes and consequences of long-term unemployment in Europe. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics* (1 ed.), Volume 3, Part C, Chapter 47, pp. 3085–3139. Elsevier.

Marubini, E. and M. G. Valsecchi (2004). *Analysing survival data from clinical trials and observational studies*. Chichester: Wiley.

McCullagh, P. and J. A. Nelder (1989). *Generalized Linear Models* (second ed.). New York: Chapman and Hall.

Mortensen, D. T. (1970). Job search, the duration of unemployment, and the Phillips curve. *The American Economic Review 60*(5), 847–862.

Mortensen, D. T. (1977). Unemployment insurance and job search decisions. *Industrial and Labor Relations Review 30*(4), 505–517.

Mortensen, D. T. (1986, October). Job search and labor market analysis. In O. Ashenfelter and R. Layard (Eds.), *Handbook of Labor Economics*, Volume 2, Chapter 15, pp. 849–919. Elsevier.

Mortensen, D. T. (1988). Matching: Finding a partner for life or otherwise. *American Journal of Sociology 94*, S215–S240.

Münder, J. (Ed.) (2009). *Sozialgesetzbuch II: Grundsicherung für Arbeitsuchende; Lehr- und Praxiskommentar*. Baden-Baden: Nomos-Verlagsgesellschaft.

Narendranathan, W., S. Nickell, and J. Stern (1985, June). Unemployment benefits revisited. *Economic Journal 95*(378), 307–29.

Narendranathan, W. and M. B. Stewart (1993). Modelling the probability of leaving unemployment: Competing risks models with flexible base-line hazards. *Journal of the Royal Statistical Society. Series C (Applied Statistics) 42*(1), 63–83.

National Institute of Economic and Social Research (1986). Seasonal patterns in the British economy. *National Institute Economic Review 117*(1), 33–42.

Nickell, S. J. (1979, March). The effect of unemployment and related benefits on the duration of unemployment. *Economic Journal 89*(353), 34–49.

Niesel, K. (Ed.) (1998). *Sozialgesetzbuch, Arbeitsförderung SGB III, Kommentar*. München: Beck.

Niesel, K. (Ed.) (2002). *Sozialgesetzbuch, Arbeitsförderung SGB III, Kommentar.* München: Beck.

Niesel, K. (Ed.) (2005). *Sozialgesetzbuch, Arbeitsförderung SGB III, Kommentar.* München: Beck.

OECD (1999). *Classifying Educational Programmes: Manual for ISCED-97 Implementation in OECD Countries* (1999 Edition ed.). Paris: OECD.

OECD (2009). *OECD Employment Outlook: Tackling the Jobs Crisis* (2009 Edition ed.). Paris: OECD Publications.

OECD (2010, March). *OECD Economic Surveys: Germany 2010*, Volume 2010/9. Paris: OECD Publishing.

OECD (2011). *OECD Employment Outlook 2011*. Paris: OECD Publishing.

Office for National Statistics (2006, December). Tables: C.1 unemployment by age and duration. *Labour Market Trends 114*(12), S42.

Oppenheimer, V. K. (1988). A theory of marriage timing. *American Journal of Sociology 94*(3), 563–591.

O'Sullivan, F. (1986). A statistical perspective on ill-posed inverse problems (c/r: P519-527). *Statistical Science 1*, 502–518.

Plaßmann, G. (2002). *Der Einfluss der Arbeitslosenversicherung auf die Arbeitslosigkeit in Deutschland - Eine mikroökonomische und empirische Untersuchung*, Volume 255 of *Beiträge zur Arbeitsmarkt- und Berufsforschung*. Nürnberg: Bundesanstalt für Arbeit.

R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.

R Development Core Team (2009). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.

Rein, M. and K. Jacobs (1993). Aging and employment trends: a comparative analysis for OECD countries. In P. Johnson and K. F. Zimmermann (Eds.), *Labour markets in an ageing Europe*, pp. 53 –76. Cambridge: Cambridge University Press.

Reinberg, A. and M. Hummel (2002). Arbeitslosigkeit: Qualifikation bestimmt Position auf dem Arbeitsmarkt. IAB Kurzbericht 15, Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.

Reinberg, A. and M. Hummel (2003). Geringqualifizierte: In der Krise verdrängt, sogar im Boom vergessen. IAB Kurzbericht 19, Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.

Reinberg, A. and M. Hummel (2005). Vertrauter Befund: Höhere Bildung schützt auch in der Krise vor Arbeitslosigkeit. IAB Kurzbericht 9, Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.

Reinhardt, H. (Ed.) (2006). *Sozialgesetzbuch VI.* Baden-Baden: Nomos Verlagsgesellschaft.

Reinsch, C. (1967). Smoothing by spline functions. *Numerische Mathematik 16*, 177–183.

Reynolds, L. G., S. H. Masters, and C. H. Moser (1991). *Labor economics and labor relations.* Englewood Cliffs, NJ: Prentice Hall.

Rogerson, R., R. Shimer, and R. Wright (2005). Search-theoretic models of the labor market: A survey. *Journal of Economic Literature 43*(4), 959–988.

Rudolph, H. (1998). Alle Jahre wieder: Saisoneffekte in der Arbeitslosigkeit. IAB Kurzbericht 12, Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.

Ruppert, D. (2002). Selecting the number of knots for penalized splines. *Journal of Computational and Graphical Statistics 11*, 735–757.

Ruppert, D., M. Wand, and R. Carroll (2003). *Semiparametric Regression.* Cambridge University Press.

Ruppert, D., M. Wand, and R. Carroll (2009). Semiparametric regression during 2003-2007. *Electronic Journal of Statistics 3*, 1193–1256.

Røed, K. and T. Zhang (2003). Does unemployment compensation affect unemployment duration? *Economic Journal 113*, 190–206.

Samuelson, P. A. and W. D. Nordhaus (2005). *Economics.* Boston: McGraw-Hill Irwin.

Schirwitz, B. (2009). A comprehensive German business cycle chronology. *Empirical Economics 37*(2), 287–301.

Sesselmeier, W., L. Funk, and B. Waas (2010). *Arbeitsmarkttheorien: Eine ökonomisch-juristische Einführung.* Berlin: Physica-Verlag.

Statistisches Bundesamt (2012, March). Registrierte Arbeitslose, Arbeitslosenquote nach Gebietsstand. Online. https://www.destatis.de/DE/ZahlenFakten/

Indikatoren/LangeReihen/Arbeitsmarkt/lrarb003.html?nn=55254&cms_gtp=151844_list%253D1&https=1, Last checked: 03/14/2012.

Statistisches Bundesamt, Gesis-Zuma, and WZB (Eds.) (2008). *Datenreport 2008 - Ein Sozialbericht für die Bundesrepublik Deutschland*. Bonn: Bundeszentrale für politische Bildung.

Steiner, V. (1997). Extended benefit entitlement periods and the duration of unemployment in West Germany. ZEW Discussion Papers 97-14, ZEW - Zentrum für Europäische Wirtschaftsforschung / Center for European Economic Research.

Steiner, V. (2001). Unemployment persistence in the West German labour market: Negative duration dependence or sorting? *Oxford Bulletin of Economics and Statistics 63*(1), 91–113.

Steiner, V. and S. Schmitz (2010). Hohe Bildungsrendieten durch Vermeidung von Arbeitslosigkeit. *Wochenbericht des DIW Berlin 77*(5), 2–8.

Stigler, G. J. (1961). The economics of information. *Journal of Political Economy 69*(3), pp. 213–225.

Stigler, G. J. (1962). Information in the labor market. *Journal of Political Economy 70*(5), pp. 94–105.

Stiglitz, J. E. (1997). *Economics*. New York: Norton.

Tableman, M., J. S. Kim, and S. Portnoy (2004). *Survival analysis using S*. Boca Raton: Chapman & Hall/CRC.

Tatsiramos, K. (2006, August). Unemployment insurance in Europe: unemployment duration and subsequent employment stability. IZA Discussion Papers No. 2280, Institute for the Study of Labor (IZA), Bonn.

Tatsiramos, K. (2009). Unemployment insurance in Europe: Unemployment duration and subsequent employment stability. *Journal of the European Economic Association 7*(6), 1225–1260.

Taylor, M. F., J. Brice, N. Buck, and E. Prentice-Lane (2009). *British Household Panel Survey User Manual, Volume A: Introduction, Technical Report and Appendices.* Colchester: University of Essex.

Taylor, M. F., J. Brice, N. Buck, and E. Prentice-Lane (2010). *British Household Panel Survey User Manual, Volume A: Introduction, Technical Report and Appendices.* Colchester: University of Essex.

Thapa, P. J. (2004, June). On the risk of unemployment: a comparative assessment of the labour market success of migrants in Australia. *Australian Journal of Labour Economics 7*(2), 199–229.

The Royal Swedish Academy of Science (2010, October). The prize in economic science 2010. Online. Press release, http://www.nobelprize.org/nobel_prizes/economics/laureates/2010/press.pdf, downloaded: 14/02/2012.

The Stationery Office (1995). *Jobseekers Act 1995: Elizabeth II Chapter 18.* London: Stationery Office Books.

Therneau, T. M. and P. M. Grambsch (2004). *Modeling survival data.* New York: Springer.

Turnbull, K. (1998). Unemployment: analysis of age and duration. *Labour Market Trends 106*(5), 243–247.

UNESCO (2003). International standard classification of education, ISCED 1997. In J. H. Hoffmeyer-Zlotnik and C. Wolf (Eds.), *Advances in Cross-National Comparison: A European Working Book for Demographic and Socio-Economic Variables*, Chapter 10. New York: Kluwer Academic Publishers.

Van den Berg, G. J. (2001). Duration Models: Specification, Identification and Multiple Durations. In J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics* (1 ed.), Volume 5, Chapter 55, pp. 3381–3460. Elsevier.

Volinsky, C. T. and A. E. Raferty (2000). Bayesian information criterion for censored survival models. *Biometrics 56*, 256–262.

von Rosenbladt, B. (2008). Datenerhebung im SOEP: die ersten 25 Jahre. *Vierteljahrshefte zur Wirtschaftsforschung / Quarterly Journal of Economic Research 77*(3), 142–156.

Wagner, G., J. Göbel, P. Krause, R. Pischner, and I. Sieber (2008). Das Sozio-oekonomische Panel (SOEP): Multidisziplinäres Haushaltspanel und Kohortenstudie für Deutschland - Eine Einführung (für neue Datennutzer) mit einem Ausblick (für erfahrene Anwender). *AStA Wirtschafts- und Sozialstatistisches Archiv 2*(4), 301–328.

Wahba, G. (1990). *Spline models for observational data.* CBMS NSF regional conference series in applied mathematics ; 59. Philadelphia: Society for Industrial and Applied Mathematics.

Wahba, G. (2006). *Spline models for observational data.* CBMS NSF regional conference series in applied mathematics ; 59. Society for Industrial and Applied Mathematics.

Wand, M. (2003). Smoothing and mixed models. *Computational Statistics 18*, 223–249.

Wand, M. and J. Ormerod (2008). On semiparametric regression with O'Sullivan penalised splines. *Australian and New Zealand Journal of Statistics 50*, 179 – 198.

Werner, H. and W. Winkler (2003). Systeme des Leistungsbezugs bei Arbeitslosigkeit: ein zwischenstaatlicher Vergleich. IAB Werkstattbericht 4, Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg. http://doku.iab.de/werkber /2003/wb0403.pdf.

Westerheide, N. and G. Kauermann (2012a). Flexible modelling of duration of unemployment using functional hazard models and penelized splines: A case study comparing Germany and the UK. *Studies in Nonlinear Dynamics and Econometrics 16*(1). Article 5.

Westerheide, N. and G. Kauermann (2012b, April). Getting unemployed: Factors influencing the risk of unemployment in Germany. Working paper.

Wikeley, N. (1996, March). The jobseekers act 1995: What the unemployed need is a good haircut... *Industrial Law Journal 25*(1), 71–76.

Wilke, R. A. (2004). New estimates of the duration and risk of unemployment for West-Germany. ZEW Discussion Papers 04-26, ZEW - Zentrum für Europäische Wirtschaftsforschung / Center for European Economic Research, Mannheim.

Wood, S. N. (2003). Thin-plate regression splines. *Journal of the Royal Statistical Society (B) 65*(1), 95–114.

Wood, S. N. (2006). *Generalized Additive Models: An Introduction with R.* Boca Raton, FL: Chapman & Hall/CRC.

Wood, S. N. (2010, April). Package 'mgcv'. Online. Version 1.6-2, http://cran.at.r-project.org/web/packages/mgcv/mgcv.pdf, downloaded: 08/06/2010.

Wood, S. N. (2011a). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 73*(1), 3–36.

Wood, S. N. (2011b, December). Package 'mgcv'. Online. http://cran.at.r-project.org/web/packages/mgcv/mgcv.pdf, downloaded: 01/07/2010.

Woodbury, S. A. and C. Davidson (Eds.) (2002). *Search Theory and Unemployment.* Boston/Dordrecht/London: Kluwer Academic Publishers.

Wurzel, E. (1993). *An econometric analysis of individual unemployment duration in West Germany.* Heidelberg: Physica-Verlag.