



Unemployment Duration in Germany –  
A comprehensive study with dynamic  
hazard models and P-Splines

Torben Kuhlenkasper and Max Friedrich Steinhardt

# Unemployment Duration in Germany – A comprehensive study with dynamic hazard models and P-Splines

TORBEN KUHLENKASPER

MAX FRIEDRICH STEINHARDT

Hamburg Institute of International Economics (HWWI)

Hamburg Institute of International Economics (HWWI)

Hamburg, Germany

Hamburg, Germany

& Centro Studi Luca d'Agliano, Italy

October 26, 2011

## Abstract

This paper makes use of data from the German socio-economic panel to gain new insights into the determinants of unemployment duration in Germany. Due to substantial differences with respect to labour market outcomes we follow a stratified approach with respect to gender and ethnicity. To analyze unemployment duration comprehensively, dynamic duration time models are used in which covariate effects are allowed to vary smoothly with unemployment duration and others enter the model in an a-priori unspecified functional form. We control for unobserved heterogeneity by following a modern frailty approach. As fitting routine we employ penalized spline smoothing effects using available software in R. We demonstrate with state-of-the-art regression models how effects of covariables change, either over duration time or within their domain and reveal substantial differences across gender and ethnicities for the German labour market. Among others we find large effects of family characteristics for women and a minor importance of formal qualifications for immigrants.

Keywords: Unemployment, Duration Time Models, Dynamic Effects, Penalized Splines, German Socio-Economic Panel, Ethnic Labour Market Segmentation

JEL classification: C14, C23, C41, F22, J16, J64, J71

*We would like to thank seminar participants at the Hamburg Institute of International Economics for useful comments. Financial support for this project from the NORFACE research program "Migration in Europe – Social, Economic, Cultural and Policy Dynamics" is gratefully acknowledged.*

# 1 Introduction

Long-term unemployment is a well-known problem in European labour markets. However, within recent years the US are also faced with growing shares of long-term unemployment (Aaronson et al. (2010)): according to the OECD (2011), the share of long-term unemployed individuals with respect to total unemployment has increased from 6% in 2000 to 29% in 2010. Although the situation in the European Union has slightly improved recently, the share of long-term unemployment is with 41% still substantially higher than in the US. For this reason, it is crucial to increase the understanding of the factors influencing the duration of unemployment. We address this issue with our current paper by a sophisticated methodological approach using data from the German Socio-Economic Panel.

In our analysis we extend the classical Cox model, which assumes proportional hazard rates and goes back to Cox (1972) by allowing for non-proportional hazards in the style of varying coefficients. These extensions have been suggested by Hastie and Tibshirani (1993), see also Gray (1994) and Therneau and Grambsch (2000). Our fitting routine employs penalized spline smoothing to estimate dynamic but also sufficiently smooth covariate effects as proposed by Kauermann (2005) and Kauermann and Khomski (2006). To capture unobserved heterogeneity and to control for serial correlation in the dataset we extend the non-proportional hazard model by including an individual latent factor as a frailty, which is assumed to be Gamma distributed. We therefore adapt the EM-algorithm of Klein (1992) to our varying coefficient models and obtain a consistent estimation framework. Applying the advanced estimation strategy to the datasets resulting from a stratified population we can graphically investigate the dynamics of the overall probability of returning into full time employment after unemployment. Our technique allows for an advanced and comprehensive analysis of unemployment in Germany with respect to gender and ethnicity compared to classical but in our case not sufficient proportional hazard models.

Initially, our analysis indicates that unemployment in Germany is characterized by a non-monotonic duration dependence. In particular, we find that hazard rates are increasing for all groups within the first spells of unemployment. After the fifth month, the likelihood of escaping unemployment declines with each additional spell of unemployment. With respect to the influence of observable characteristics, our paper highlights the important role of family characteristics for women. Notably, the presence of young children and older relatives in

the household reduces the probability of returning into employment. With respect to immigrants, our estimates show that formal qualifications are of minor importance for immigrants when leaving unemployment. Furthermore, our analysis has gained from the modern statistical approach used. In particular, we find two types of characteristics: variables with a constant effect over duration time and characteristics with time dependency. To the first category belong variables like education, while the second group mainly consists out of variables capturing an individual's previous labour market situation.

The paper is organized as follows: in Section 2 we give an detailed overview of the employed empirical database and some first analytics based on Kaplan-Meier estimators. In Section 3 we outline the statistical method used for the estimation. Section 4 analyzes both, the proportional and the non-proportional hazard models before we conclude in Section 5.

## 2 Data and Descriptive Statistics

The data used come from the German Socio-Economic Panel (GSOEP), which is a representative micro data set on persons, families and households in Germany. It contains a large array of socio-economic variables and is widely used by sociologists and economists. For a more detailed introduction to the GSOEP we refer to Haisken-DeNew and Frick (2005), Wagner et al. (2007) and Wagner et al. (2008). A main feature of the dataset is the provision of detailed information on respondents' immigration history like country of birth, citizenship and ethnicity. This allows us to identify different ethnic groups of immigrants and to distinguish between first and second generation immigrants. Furthermore, the GSOEP includes a number of variables describing the current employment status and the labour market experience of the interviewed persons. Due to the longitudinal structure of the data set, we are able to follow individuals throughout their employment biographies.

Our analysis in this paper is based on data from West-Germany covering the time period of January 1984 to December 2008 and therefore makes use of the entire GSOEP-history. As unemployment we define the period an individual is off the job and officially being registered unemployed in Germany. Due to spell data provided by the the GSOEP, our analysis is carried out on a monthly base and provides additional information about the employment situation for both, the period anteceding and following unemployment. The underlying population of our analysis consists therefore of adult men and women living in Germany, who have at

least one spell of unemployment between January 1984 and December 2008. However, due to well-known differences in the labour force behavior between men and women, we stratify the population with respect to gender. In addition to this, we decided to distinguish between natives and immigrants as well. The latter is motivated by a number of distinct features of immigrants with respect to labour market performance and access.<sup>1</sup> As a result of these, the German labour market is characterized by ethnic segmentation (see Steinhardt (2011)). In particular, natives and immigrants tend to work in different occupational segments despite having similar qualifications. The modeling exercise is therefore being carried out for each of the four above motivated strata.

As event for our duration time modeling, we consider the return into full time employment in the month immediately following the spell of unemployment. However, we extract people from the dataset who return into any kind of self-employment, since these transitions are of minor precision concerning the time spell. The maximum duration times of unemployment are limited to 36 months for each strata, which keeps at least 90% of the data in each stratified dataset. Duration times exceeding this threshold are treated as right-censored thereafter. In addition, observations are treated as right-censored if the corresponding individual leaves unemployment without starting to work full time or due to panel drop-out. This definition of censoring includes returning into part time employment as well as staying at home for being housewife or houseman and not being registered unemployed any more.

Besides defining conditions for censored observations, we exclude observations from the dataset if the individual is younger than 20 or older than 60 years of age by the first month of the unemployment spell. In addition, observations which are left-censored due to lacking information about the exact beginning of the spell have to be excluded, since an exact duration time modeling can not being carried out and information concerning the period anteceding unemployment is missing. As previously mentioned, observations indicating a transition into any kind of self-employment do not enter the estimation routine. Our approach of analyzing the unemployment duration with respect to gender and ethnicity restricts the datasets to individuals with undoubtful information provided to the GSOEP, which is especially of major

---

<sup>1</sup>In general, immigrants suffer from an inadequate transferability of skills from their home country. As a consequence, formal qualifications of natives and immigrants tend to have a different relevance for labour market outcomes. Furthermore, immigrants are likely to face discrimination by employers, which likely reduces chances to leave unemployment.

importance concerning the ethnic background. Taking these exclusions and the dropout of observations due to missing values in the variables into account, we will carry out the modeling exercise with 1239, 1081, 611 and 335 native men, native women, immigrant men and immigrant women, contributing 1708 (1045), 1405 (537), 853 (636) and 433 (239) observations (events), respectively.

Especially when modeling the economically important phenomenon of unemployment, the selection of the covariables is of crucial importance and therefore discussed in the following: In our analysis, we distinguish between three different categories of explanatory variables. Our first category contains variables on individual characteristics. Initially, we control for the age of a respondent, which is likely to influence the number of job offers as well as the individual search intensity. Both aspects are crucial for the likelihood of leaving unemployment. The corresponding covariable  $age_i$  captures the age of the respondent in observation  $i$  at the beginning of the spell, measured in years. For similar reasons we consider the educational attainment of an individual. In particular, we expect that hazard rates rise with education levels due to employer preferences for skilled workers. The GSOEP provides information about the educational attainment in the classification of the *International Standard Classification of Education (ISCED)* (see UNESCO Institute for Statistics (2006) for details), consisting of seven categories. With excluding the category 0, which indicates current school attendance, we define  $low.ISCED_i$  by the categories 1 and 2, take the categories 3 and 4 of the ISCED as reference category in the models and finally define  $high.ISCED_i$  by the levels 5 and 6.

Furthermore, we control for previous unemployment experience and disability, which are likely to reduce individual exit probabilities. The first is taken from the generated variables provided by the GSOEP for each survey participant and is defined as the entire unemployment experience, even before entering the panel. The corresponding covariable  $ue_i$  gives the years of experience at the beginning of the spell  $i$ . The latter characteristic leads to the binary coded covariable  $handicap_i$ , which takes the value of 1, if the person has an official reduction in work-capability in the year spell  $i$  starts. In addition, we capture whether a respondent is eligible for unemployment benefits in Germany, which tend to increase an individual's reservation wage. Since the eligibility is not directly addressed by the GSOEP questionnaires, the (binary) coded covariate  $no.ue.benefit_i$  is a proxy taking the value of 1 if the individual

has not received any unemployment benefit during the entire duration time  $i$  and is therefore likely not eligible for receiving the benefit. Since the duration of receiving unemployment benefits is also measured on a monthly base by the GSOEP, our proxy  $no.ue.benefit_i$  in the analysis seems to be a reliable (and valid) indicator for the aspired economic context. Finally, we create a binary coded variable  $previous.fulltime_i$  capturing the information about the time period anteceding the spell  $i$ . Since we have defined a return into *full* time employment as event in our duration time analysis, this corresponding variable takes the value of 1 if the person had to face a transition from full time employment into unemployment.

Within the second category, we control for family characteristics like the presence of children and elderly relatives within the household. In both cases, we expect a negative effect on the supply side, in particular for women. The covariate  $children.under.6_i$  takes therefore the value of 1 if the individual has at least one child younger than six years at the time point of unemployment begin. The threshold of six years is chosen due to an obligation of sending children to school thereafter. The latter family characteristic leads to the covariable  $multi.hh_i$ , taking the value 1, if the individual lives in an multi-generation household (this information is generated by the GSOEP) or has to take care of at least on person with intensive care needed (this information is gained from the annual questionnaires). In addition to this, we take into account that leaving unemployment is likely to depend on the presence of a working partner and the overall household income. Instead of focusing on the usual marital status of the respondent, we look at possible life-partners and their employment situation at the beginning of the spell  $i$ :  $working.partner_i$  is binary coded and takes the value 1 if the partner works a the labour market when the unemployment spell of our observed individual begins. The overall household income is provided as an (potentially adjusted) net household income by the GSOEP. After adjusting the income to inflation with respect to the German Consumer Price Index, we set  $ahinc.high_i$  to 1, if the corresponding household belongs to the fourth quartile of all households in the strata. Obversely,  $ahinc.low_i$  captures those households belonging to the first quartile and suffering from low income. The 50% of remaining households belonging to the second and the third quartile are taken as reference category.

Our third category is made of variables capturing regional and calendar effects. This includes

the dummy *south.germany<sub>i</sub>* for the southern federal German states<sup>2</sup>, characterized by low unemployment and dynamic labour markets. We also create dummies for quarters of the corresponding year: By taking the second quarter of a year as reference category, *first.quarter<sub>i</sub>* captures the month of January to March, *third.quarter<sub>i</sub>* takes obversely the value of 1 for the months of July to September and *fourth.quarter<sub>i</sub>* corresponds to October to December. In addition, we include the calendar year as a metrical effect (*year<sub>i</sub>*) and ensure therefore that we control for both seasonal and business cycle effects on individual exit probabilities (see van den Berg and van der Klaauw (2001)). Finally, in the case of immigrants, we include dummies for source country regions. By this, we account for the possibility that hazard rates vary systematically across immigrant groups due to cultural and legal differences. A detailed list of the different immigrant groups is provided in Table 1.

If the individual has an indirect migrational background and is therefore born in Germany, the binary coded variable *second.generation<sub>i</sub>* takes the value 1.

Before proceeding to our empirical models, we will provide some descriptive evidence based on simple Kaplan-Meier curves. The first graph in figure 1 presents overall survivor curves for our four groups of interest. It becomes obvious that German and immigrant men are characterized by very similar Kaplan-Meier curves. In particular, within the first ten month the unconditional probability of remaining unemployed is almost identical for immigrant and native men. A similar pattern can be observed for foreign and German women who both exhibit higher curves than their male counterparts, indicating higher chances to remain unemployed. In the subsequent months, the difference in the curves of immigrants and natives increases. Relative high numbers of immigrants leaving unemployment into part-time employment might drive latter.

In the following, we will show separate figures for each of our subpopulations: the first and second column present Kaplan-Maier-curves for native men and women, while the third and fourth column display survivor curves for immigrant men and women. First, we look at the influence of education. As expected, having a low educational attainment is likely to increase the duration of unemployment. However, the figures point into the direction that the role of education is less important for immigrants. With respect to individual employment biographies, the Kaplan-Meier-curves indicate a higher chance of unemployment exit after

---

<sup>2</sup>Bavaria, Baden-Württemberg and Hessen

covariable (binary coded)	countries of origin
$south.europ.guest_i$	Italy, Greece, Portugal, Spain
$Eastern.Europe_i$	Bulgaria, Czech-Republic, Hungary, Poland, Rumania, Slovakia, other eastern europe countries
$form.ussr_i$	Armenia, Azerbaijan, Belarus, Estonia, Georgia, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Moldova, Russian Federation, Tajikistan, Turkmenistan, Ukraine, Uzbekistan
$form.Yugoslavia_i$	Albania, Bosnia and Herzegovina, Croatia, Kosovo, Macedonia, Serbia and Montenegro, Slovenia
$OECD_i$	Australia, Austria, Belgium, Denmark, Finland, France, Iceland, Ireland, Japan, Luxembourg, New Zealand, Sweden, Switzerland, United Kingdom, United States
$rest_i$	all other countries
(Turkey)	(reference category)

Table 1: generated dummy variables of ethnic groups

previous full time employment. Surprisingly, this effect is only pronounced for native men. Furthermore, it is noteworthy that the effect of having a child under the age of six seems to differ across gender – being positive for men and negative for women. The Kaplan-Meier estimates for immigrants by country of birth exhibit huge differences across ethnic groups. While for example 80% of Eastern European men have returned to employment after 10 months, the corresponding share of Turkish immigrants is only about 60%. The figures further show substantial differences in the survival curves of immigrants from Eastern Europe and the former Soviet Union. Overall, the Kaplan-Meier estimates support our empirical approach, which explicitly models the heterogeneity across ethnic groups. Finally, the Kaplan-Meier curves emphasize the necessity of an empirical model allowing for non-proportional hazard

rates. In particular, the curves for education of native women (non-parallel curves) and immigrants (crossing curves) indicate the presence of non-linear relationships and a violation of the underlying assumptions of the proportional hazard models. Therefore, the descriptive analysis supports our decision for a non-parametric model, which allows effects to change with the duration of unemployment.

### 3 Nonparametric Hazard Models and Penalized Spline Smoothing

We denote with  $h(t, x, z)$  the hazard function which mirrors the probability of returning to full time employment after  $t$  months of being registered as unemployment. To be more specific,  $t_i$  captures the duration of unemployment contributed by the  $i$ -th observation in the corresponding dataset of the stratum. The hazard depends on a set of covariates  $x_i = (x_{i1}, \dots, x_{ip})$  with  $p = 14$  for the strata of the native men and women and  $p = 21$  for the strata of the immigrant men and women. The covariates of  $x_i$  capture individual and socio-economic indicators, which are introduced in Section 2 and are binary coded. In addition we denote with  $z_i = (age_i, ue_i, year_i)$  the set of metrically scaled covariables capturing the age of the respondent at the beginning of the unemployment spell, the previous unemployment experience and the starting year of the corresponding spell. With  $d_i$  we denote the censoring indicator, stating whether the true but unobserved duration is larger than  $t_i$ , see Section 2 for details about the definition for the employed data. The resulting typical Cox model takes therefore the form

$$h(t, x, z) = \exp\{\beta_0(t)\} \exp\left\{\sum_{j=1}^p x_{ij}\beta_j + \sum_{l=1}^{q=3} z_{il}\beta_l\right\}, \quad (1)$$

where  $h_0(t) = \exp\{\beta_0(t)\}$  is the baseline hazard and  $\beta_j$  and  $\beta_l$  give the covariate effects for the binary coded and metrically scaled covariables, respectively (See Cox (1972)). However, the effects expressed in  $\beta_j$  are assumed to be constant over time, so that model (1) implies proportional hazards. Looking at the KM-estimators in Figures 1 to 3 and referring to the previous section the proportionality assumption seems questionable since some of the Kaplan-Meier curves reveal a dynamic behavior and therefore do not mirror proportionality. We therefore allow covariate effects of  $x_i$  to vary with the duration of unemployment. This

interaction of effects is incorporated in the model in a functional and therefore non-parametric form where  $\beta_j(t)$  are smooth effects to be fitted from the data and have been coined 'varying coefficients' by Hastie and Tibshirani (1993). In addition, we ease restrictions on the assumed effects of the covariables summarized in  $z_i$  by allowing for non-linearities with functional and a-priori unspecified forms. We therefore define  $\gamma(\text{age}_i)$ ,  $\delta(\text{ue}_i)$  and  $\phi(\text{year}_i)$  as smooth and sufficiently differentiable functions, estimated by the data for each of the covariables in  $z_i$ , respectively. The hazard function as introduced in (1) for the unemployment duration therefore changes to

$$\begin{aligned}
 h(t, x_i, \text{age}_i, \text{ue}_i, \text{year}_i) &= \exp\{\beta_0(t)\} \\
 &\cdot \exp\left\{\sum_{j=1}^p x_{ij}\beta_j(t)\right\} \\
 &\cdot \exp\{\gamma(\text{age}_i) + \delta(\text{ue}_i) + \phi(\text{year}_i)\}. \tag{2}
 \end{aligned}$$

Apparently, model (2) with its functional components  $\gamma(\cdot)$ ,  $\delta(\cdot)$  and  $\phi(\cdot)$  is itself not identifiable since an offset can go into any of the three latter functions. We therefore need the further constraint that  $\gamma(\cdot)$ ,  $\delta(\cdot)$  and  $\phi(\cdot)$  each integrates out to zero with respect to the (empirical) distributions of  $\text{age}_i$ ,  $\text{ue}_i$  and  $\text{year}_i$ , respectively.

Estimation is carried out using penalized splines. We follow thereby closely Kauermann and Khomski (2006) and Kuhlenkasper and Kauermann (2010). The basic idea is to replace the smooth functions in (2) by some high dimensional spline bases and to achieve smoothness a penalty component is added to the likelihood. In a first step of estimation, the components in (2) can therefore be replaced by

$$\begin{aligned}
 \beta_0(t) &= B_0(t)u_0 \\
 \beta_j(t) &= B_j(t)u_j \\
 \gamma(\text{age}) &= B_\gamma(\text{age})u_\gamma \\
 \delta(\text{ue}) &= B_\delta(\text{ue})u_\delta \\
 \phi(\text{year}) &= B_\phi(\text{year})u_\phi
 \end{aligned}$$

with  $B(\cdot)$  as high dimensional but formally strict parametric functions and  $u$ . as the corresponding coefficient vector. To be more specific, in this analysis a cubic smoothing spline basis

is chosen, see Wahba (1978) for details. By carrying out an (unpenalized) spline smoothing estimation with knots placed at the unique observed data points, the computational burden in this case would be enormous and the resulting fit will be poor unless we handle the wiggliness of the fit by imposing a penalty in the coefficient vectors  $u$ . We therefore employ so called "low ranked smoothers", which reduce the amount of knots placed in the domains of the variables and therefore allow for feasible computation but are still flexible enough to capture existing non-linearities in the data. These "low ranked smoothers" have been proven to be reliable and stable and are therefore commonly implemented in the available statistical software like R, see Wood (2006). The underlying idea has been coined as (P-)enalized-spline smoothing, see Eilers and Marx (1996). As Ruppert (2002) and Kauermann and Opsomer (2011) have shown, the exact a-priori choice of the amount of knots is of minor importance and not discussed here.

It can be shown that the likelihood can be approximated by a Poisson type Mixed Model: For simplicity of notation and presentation of the penalized spline idea we ignore the other covariates  $x$  in model (2) for the moment and demonstrate the estimation strategy only for the baseline hazard rate  $\beta_0(t)$ . However, this technique can easily be transferred to the other covariates  $x$  and the three metrically scaled covariables  $age_i$ ,  $ue_i$  and  $year_i$ : After having replaced the unknown smooth functions by the high dimensional basis representation, imposing a penalty on the coefficient vector  $u_0$  is necessary. This guarantees that the resulting fitted curve  $\hat{\beta}_0(t) = B_0(t)\hat{u}_0$  is smooth. This is achieved by adding the penalty component  $\lambda_{0u}u_0^T D_{0u}u_0$  to the log likelihood, with  $D_{0u}$  as penalty matrix and  $\lambda_0$  as penalty parameter steering the amount of smoothness.

Denote now with  $(t_i, d_i)$  the observations (again omitting covariates for simplicity of presentation), where  $t_i$  is the length of unemployment and  $d_i$  the censoring indicator. The penalized likelihood for coefficients  $u_0$  results now with classical theory, see Cox and Oakes (1984), to

$$\ell(u_0, \lambda_{0u}) = \sum_{i=1}^n \left\{ d_i B_0(t_i) u_0 - \int_0^{t_i} \exp(B_0(z) u_0) dz \right\} - \lambda_{0u} u_0^T D_{0u} u_0. \quad (3)$$

For estimation two further aspects have to be considered. First, one has to numerically solve the integral in (3) resulting from the integrated hazard function. A simple and numerically feasible way to do so is to use a trapezoid approximation, see Kuhlenkasper and Kauermann (2010) for an application in a comparable economic context. However, numerically more precise is an approximation following Simpson's rule, which allows to approximate definite

integrals and is widely used in many areas of science and research, see Atkinson (2007) and Süli and Mayers (2003) for details.

This approximation method boils down to discretizing the continuous time scale of unemployment by dividing each observed interval  $[0, t_i]$  into  $M$  (equidistant) subintervals  $[T_{m-1}, T_m]$  with  $m = 1, \dots, M$  as well as  $T_0 = 0$  and  $T_M = t_i$ . The integral in (3) becomes now with the application of Simpson's rule

$$\begin{aligned} \int_0^{t_i} \exp\{B_0(z)U_0 dz\} &\approx \sum_{m=0}^M \frac{T_{m+1} - T_{m-1}}{6} \left[ \exp\{B(T_{m-1})u_0\right. \\ &+ 4 \cdot \exp\left\{B\left(\frac{T_{m-1} + T_{m+1}}{2}\right)u_0\right\} \\ &\left. + \exp\{B(T_{m+1})u_0\} \right], \end{aligned} \quad (4)$$

with  $T_{-1} = T_0$  and  $T_{M+1} = T_M$ . Replacing the integral in (3) by this sum yields a penalized likelihood for the artificial random variables  $Y_{ik}$  with

$$Y_{ik} = \begin{cases} 0 & , \quad k \leq 2M - 1 \\ d_{ik} & , \quad k = 2M \end{cases} \quad (5)$$

and having the Poisson distribution

$$Y_{ik}|u_0 \sim \text{Poisson}\left(\lambda_{ik} = \exp\left\{B_0(\tilde{T}_k)u_0 + o_{ik}\right\}\right), \quad (6)$$

with  $\tilde{T}_k = T_{k/2}$  if  $k$  is even and  $\tilde{T}_k = (T_{(k+1)/2} + T_{(k-1)/2})$  if  $k$  is odd. Note, that  $o_{ik}$  is the known offset after the approximation with  $o_{ik} = \log((T_{j/2+1} - T_{j/2-1})/6)$  or  $o_{ik} = \log(4(T_{(j+1)/2} - T_{(j-1)/2})/6)$  for  $k$  being even or odd, respectively.

The next aspect is to select the smoothing parameter  $\lambda_{0u}$  appropriately, that is data driven. This can be done by comprehending the penalty as a priori normality imposed on the coefficient. In this case  $\lambda_0$  becomes a parameter which can be estimated by maximizing the corresponding likelihood, which leads to

$$u_0 \sim N(0, \lambda_{0u}^{-1} D_{0u}^-) \quad (7)$$

with  $D^-$  as (generalized) inverse. With (6) and (7) we obtain a Generalized Linear Mixed Model (GLMM) and the smoothing or penalty parameter becomes an a priori variance com-

ponent which could be estimated following the likelihood principle. This idea has proved to be quite powerful, both in theory as well as in its numerical performance. For further details we refer to Wand (2003) and Kauermann (2005). The model can now be fitted using available software for GLMMs in the style of Breslow and Clayton (1993). The idea is to treat spline coefficient  $u_{0t}$  as random so that the likelihood to be maximized results by integrating out the random terms. The latter can be done by Laplace approximation. Clearly, the idea of penalized splines and its connection to GLMMs extends to model (2), that is for fitting the smooth covariates effects  $\beta_j(t), j = 1, \dots, p$  and to fit the functional effects  $\gamma(\text{age}), \delta(\text{ue})$  and  $\phi(\text{year})$ .

A user-friendly implementation to fit the model is provided in R, see R Development Core Team (2010). As Hastie and Tibshirani (1990) show, the model can easily be fitted by employing software for generalized additive models. We follow Wood (2006) and use the `gam()`-Function in the package `mgcv`, see Wood (2010) for details about the package. The selection of the penalty parameter  $\lambda$  is carried out data driven by a generalized cross validation. Moreover, using standard asymptotic arguments, one can derive variance formulae from the estimates, making use of asymptotic normality statements. This allows not only to fit functional shapes but also to provide confidence bands for the functional effects.

For our stratified data set, which appears to be an unbalanced panel data set with multiple observations per individual we have to supplement (2) by introducing an individual specific frailty effect. This latent random effect enters the model in a multiplicative way and takes on the one hand unobserved heterogeneity in the data into account and controls for serial correlation on the other hand. We therefore extend (2) so that the  $i$ -th observation corresponds to the specific hazard

$$h_i(t, x_i) = h(t, x_i)v_i \tag{8}$$

with  $v_i$  as unobserved latent effect with  $E(v_i) = 1$  to maintain identifiability. The latent effect can either be modeled by a finite mixture of mass points and weights or by assuming a Gamma distribution for  $v_i$ . The first is described by Heckman and Singer (1984) and is applied by Bover et al. (2002) and Kuhlenkasper and Kauermann (2010). The modeling however has to be carried out employing time consuming bootstrapping. We therefore follow Klein (1992) with the latter modeling strategy and use the conjugate distribution to Poisson which allows for numerically simple estimation of our model. We thereby extend model (6)

by assuming  $\lambda_{ik}$  to depend on some unobservable heterogeneity as well. We replace (6) by

$$Y_{ik}|u_0, v_i \sim \text{Poisson}(\lambda_{ik}v_i) \tag{9}$$

with

$$v_i \sim \text{Gamma}\left(\frac{1}{\alpha}, \alpha\right) = \frac{1}{\left(\frac{1}{\alpha}\right)^\alpha \Gamma(\alpha)} v_i^{\alpha-1} \exp(-\alpha v_i). \quad (10)$$

Note again that  $E(v_i) = 1$  and  $Var(v_i) = \frac{1}{\alpha}$ . The Gamma distribution is the conjugate distribution for the Poisson distribution so that  $v_i | (d_{ij}, j = 1, \dots, 2M)$  is again Gamma distributed. This allows to easily apply an EM algorithm with updating the previously motivated offset  $o_{ik}$  in each step and carrying out a maximization with Fisher Scoring. For comparableness we follow Klein (1992) and also extend the proportional hazard model (1) with respect to gamma frailties. The variance of  $v_i$  for model (8) is each estimated from the data and takes value 0.33, 0.32, 0.19 and 0.27 in our data example for native men, native women, immigrant men and immigrant women, respectively.

## 4 Empirical Analysis

### 4.1 Proportional Hazard Model

Table A.2 presents the results from our proportional hazard model for natives and immigrants, split by gender. Initially, the estimates confirm our previous observation: education matters for the duration of unemployment. However, the effect differs across gender and ethnicity. In particular, we find that only native men profit from having of a high qualification, while having a negative qualification decreases the exit probabilities for all groups. However, with respect to latter the effect for immigrants is almost not significant. In contrast to this, we find a positive and highly significant effect of previous full time employment for immigrants and natives of both genders. Therefore, our results suggest that formal qualifications are of minor importance for immigrants while full time work experience increases chances for leaving unemployment. This is a reasonable finding, since immigrants are often not able to make use of their formal qualifications due to an insufficient transferability of human capital or a missing recognition of foreign credentials. Moreover, the German labour market is characterized by occupational segmentation between natives and immigrants, which further weakens the role of formal education (Steinhardt (2011)). Furthermore, we are finding a strong negative disability effect for men. Finally, the estimates from the proportional hazard model indicate that being non-eligible for unemployment benefits reduces the duration of

unemployment for native women and male immigrants. This finding is in line with other empirical studies who find a negative relation between the receipt of unemployment benefits and the likelihood of leaving unemployment (among others Hunt (1995), Bover et al. (2002)). With respect to the family channel, our estimates suggest that the presence of children under the age of six has a negative effect on individual exit probabilities of women. On the other hand, we find that native men return earlier into employment if they have young children within the household. This implies that they either increase their search intensity or reduce their reservation wage as a reaction to the presence of young children. Our results further show that women return significantly later to employment if they live in a household with older relatives. For all other groups living in a multigenerational household seems to have no influence on the hazards of leaving unemployment. Both effects highlight the persistence of traditional role models in German households with women as primary caregivers. With respect to the influence of a working life-partner within the household, we yield surprising results. In particular, we find a positive effect of a having a working partner for native men and immigrant women. However, as we see in the next section, this result does not hold if we allow the effects to change over time. Having a high household income plays no role for any of our groups, while belonging to a household at the lower end of the income distribution has a negative effect on the unemployment duration of native women.

Regarding regional differences within Germany, we find that natives of both genders living in one of the Southern states are characterized by higher hazard rates than their counterparts in the rest of Germany. This finding reflects the dynamic job market situation in Germany's South. However, the regional dummy is not significant for immigrants. This implies that immigrants do not experience better job and wage offers in Southern Germany than in the rest of Germany. Similar differences between natives and immigrants are found with respect to seasonal influences. Natives who enter unemployment during the winter (first and fourth quarter of the year) are likely to leave unemployment quicker than individuals who become unemployed in spring, while immigrants seem to experience no substantial seasonal effects.

<sup>3</sup> In other words, getting unemployed in a period of high seasonal labour demand (spring) reduces the chances of quick reemployment for natives. On the other hand, entering unemployment in winter increases chances to leave unemployment with the next peak in seasonal

---

<sup>3</sup>With the exception of women who seem to experience a positive effect when entering unemployment in the third quarter.

labour demand. <sup>4</sup>We will pick up this point in the next section.

As to our metrically scaled explanatory variables, we find negative effects for age and year of entry into unemployment. However, we refrain from a detailed interpretation at this point since we expect to gain more insight from the non-proportional hazard model in the next section. In particular, we expect both effects to be non-linear. With respect to unemployment duration of immigrants, we find substantial differences in hazard rates across ethnic groups for men. For women we do not find any significant differences across nationality groups after controlling for observable characteristics and unobservable heterogeneity. In particular, our estimates for men suggest that immigrants from Turkey have the lowest hazard rates among all ethnic groups. The highest likelihood of return into employment is observed for immigrants from OECD and Eastern-European countries. Both groups are characterized by high shares of individuals with tertiary education working in white-collar occupations. Therefore, they are likely to enjoy high job offer rates within the German economy. Potential explanations for the inferior performance of Turks are insufficient language skills and low human capital at the supply side. In addition to this, recent evidence from a field experiment emphasizes that Turks are faced with discrimination in the German labour market (Kaas and Manger (2011)). Therefore, the underperformance of Turkish immigrants might be also driven by lower job offer arrival rates of German employers. Finally, we find that children of immigrants who were born in Germany exhibit higher hazard rates than all immigrants born abroad. This intuitive result emphasizes that country-specific human capital and socialization matter for job search.

## 4.2 Non-Proportional Hazard Model

In the Figures 4 to 7 we present the resulting fit of the models corresponding to model (8) with pointwise 95% confidence bands. The columns in each of the figures display the results for native men, native women, immigrant men and immigrant women in the first, second, third and fourth column respectively. In addition to the functional effects and the confidence bands, we integrate the information gained from the Kaplan-Meier curves and the proportional Cox model for the binary coded covariates. The first is achieved by adding a

---

<sup>4</sup>For example, van den Berg and van der Klaauw (2001) find for France the highest positive seasonal effect on individual exit probabilities in spring.

shaded region at the bottom of the graph corresponding to the KM graphs: from *black* to *white* the shades indicate the percentage of individuals having the specific characteristic and are still observed to be unemployed up to this timepoint. The thick mark corresponds to a 50% estimated probability of extending the unemployment duration beyond this threshold. The dashed lines represent the proportional Cox-effects being described in the previous subsection. The first line presents the baseline effects for our four groups of interest. Initially, it becomes obvious that the baseline hazard rates are not constant over time. In particular, we find an increase in predicted hazard rates within the first months with a peak around the fifth month. This pattern can be observed for all groups. In the following months, the probability of leaving unemployment decreases slightly for men, while the hazard rates of native women drop sharply. The hazard rates of male immigrants are almost constant until 15 months of duration when they start to decline. The estimates for immigrant women indicate that the probability of re-entering full time employment continues to rise after month 8. However, the confidence band widens strongly due to the fact, that most of the individuals have either returned into full time employment or have dropped out of the analysis before. This phenomenon can be observed at many displayed dynamic effects near the end of the defined timescale of 36 months. Overall, the baseline estimates from our non-proportional hazard mode with unobserved heterogeneity suggest the existence of a non-monotonic duration dependence. Within the first months, the hazard rates increases with unemployment duration (positive duration dependence). After the fifth month, the probability of leaving unemployment is declining with its duration (negative duration dependence). Potential explanations for the latter are deterioration of human capital, discouragement and decreasing job offers through negative signaling (Machin and Manning (1999)). Our results are in contrast to the evidence of Steiner (2001) who finds for Germany a positive duration dependence after controlling for unobserved population heterogeneity. Hazard rates with very similar shapes to ours are found by Bover, Arellano, and Bentolila (2002) for Spain.

Looking at the educational attainment both, native men and women reveal an almost constant positive effect when having achieved a high ISCED level. However, low educated native females suffer even more as the unemployment duration continues. While 50% of highly educated native men are likely to return into full time employment within six months, low educated native females are not likely to return within three years. As previously mentioned,

the weaker effects of immigrant might be due to inadequate transferability of their skills. While handicapped people have to face very low reentering probabilities almost constantly throughout the four strata, the effect of previous full time employment shows an interesting dynamic pattern: while the effect is rather strong at the beginning of each unemployment duration, the effect weakens almost linear throughout the strata as the unemployment duration continues. However, native and immigrant men return earlier into full time employment than the females. The effects of being non eligible of receiving unemployment benefits in Germany is almost similar for native and immigrant women. While it has a (rather weak) positive effect for the first five months, it turns to be strong negative beyond this time point. As expected, native and immigrant men return much earlier to full time employment if they do / can not receive the benefits. The effect loses importance after about one year. In contrast to the proportional hazard models, the effect of a working life-partner reveals a clear dynamical pattern for females: the effect is strongly negative in the first year of unemployment with an increasing likelihood of returning into full time employment thereafter. Furthermore, the shaded bar shows that 50% of the native born females have not returned within three years. For immigrant women however, the effect is weakly positive throughout the entire analyzed duration period and starts even increasing after two years. For males, especially natives, the effect is of minor importance. The findings for natives support the often discussed male-breadwinner model, especially in the first year of the female's unemployment. The presence of children younger than six years affects the females' return into full time employment as being described in the previous section. On the other hand, many native fathers of young children return to the job rather quickly. However, this effect diminishes with ongoing duration time. An interesting dynamic effect can be found by looking at native women living in a multi-generation household: while the likelihood of returning is rather low in general, the effect is strongly negative at the beginning of the spell and changes the sign as duration time continues. This might indicate a financial pressure on the females living in these household structures due to the necessity of a financial contribution to the income of a large household. The duality between taking care of other family members and working however remains since at least 50% of these native females have not returned into full time employment within 36 months. While the overall household income is of minor importance and reveals no clear pattern throughout the strata, region where the unemployed

individuals live seems to affect the hazard of returning to the labour market in a dynamical pattern: in contrast to the Cox-effects it seems to be important for immigrant females living in the three southern German states to return to employment within the first five months of unemployment. If they have not returned up to this threshold, the effect declines and stays negative thereafter. Looking at the effects of the quarters within the year of unemployment beginning, it can be observed that these effects vanish as the duration of unemployment continues. In contrast to the findings by the proportional hazard models, the effects only seem to be important in the very short run. These findings are not surprising and underline the argument that seasonal unemployment focusses only on the short run. As the duration continues, other socio-economic effects are more important for explaining the phenomenon of unemployment.

Looking at the metrically scaled covariables and their effects, especially the effect of age reveals an interesting dynamic pattern in contrast to the findings of the Cox models: Throughout all four strata, the shape of the effect is almost identically concave, which states a rather constant effect up to a certain threshold. Due to identifiability reasons of the model (8), all of the four curves have to cross the zero-line. However, the threshold is quite different for the strata: while native men have to face negative effects of the age not before 50 years, immigrant females have a positive age-effect only up to 40 years. Native females and immigrant men reveal an almost similar pattern with a common threshold around 45 years. The previous unemployment experience seems to be of minor importance. If individuals throughout the strata have accumulated more than one year of unemployment experience, they have to face strengthening negative effects on the returning likelihood. With respect to business cycle effects, we find a clear cyclical effect on the hazard rates for native men. On the one hand, we observe a positive effect on hazard rates for years of strong economic growth (1985-1990). Becoming unemployed in periods of economic weakness, like the years after the bursting of the dot-com-bubble in 2000, has a negative effect on the probabilities on return. In contrast to this, we do not find cyclical influences for hazard rates of women. An interesting case is the time effects on hazard rates of immigrant men. From the mid of the nineties the shape of the effect is similar to the one of native men reflecting the business cycle. However, between 1985 and the mid of the nineties the estimates suggest a substantial negative time effect. Potential explanations are worsening job opportunities through immigration. From 1985 on,

Germany was experiencing large annual increases in net-immigration from Eastern Europe and the Soviet Union. In addition to this, West-German regions attracted substantial numbers of East-Germans after the reunification. Both effects implied increased job-competition for low skilled immigrants already living in Germany. Finally, we provide estimates for the dummies, which we have incorporated to control for systematic differences across nationalities within the immigrant population. Overall, it becomes obvious that the hazard rates of most immigrant groups are constant or slightly increase over time. Our results therefore suggest that the influence of group-specific influences on the exit probability does not diminish with the duration of unemployment. This implies that first generation-immigrants from Turkey have lower exit probabilities than immigrants from other countries even after long periods of unemployment. In addition, the figures highlight substantial gender differences within all groups regarding the duration of unemployment. While 50% of the men return within 5 months, the majority of women tend to remain unemployed much longer.

## 5 Conclusions

The present paper makes use of data from the German Socio-Economic Panel to gain new insights into the determinants of unemployment duration in Western societies. In our analysis, we distinguish between women and men, as well as between natives and immigrants. Latter is motivated by the fact that the German labour market is characterized by occupational segmentation of immigrants. The paper is based on a new methodological approach which extends the established statistical model and allows to analyze unemployment duration sophisticatedly.

Initially, we find evidence for a non-monotonic duration dependence. Within the first months of unemployment, we observe increasing hazard rates for all groups. In the following months, hazard rates are characterized by negative duration dependence. In other words, the likelihood of escaping unemployment declines with each additional spell of unemployment. This pattern is typical for situations of deteriorating human capital, discouragement and decreasing job offers through negative signaling. Furthermore, our results highlight the important role of family characteristics for women. We find that the likelihood of returning into employment is significantly reduced by the presence of young children and older relatives in the household. The estimates for immigrants have revealed large disparities across groups of

different origin. For example immigrants from Eastern Europe and OECD countries tend to return to employment quickly, whereas Turkish immigrants are faced with long durations of unemployment. In addition, we find that formal qualifications are of minor importance for immigrants. This result is driven by an insufficient transferability of foreign human capital and ethnic segmentation within the German labour market.

While some of the results gained from our modern statistical technique go hand in hand with previous findings in the literature and the results from the classical model, some effects reveal a clear dynamical pattern and justify the use of the modern statistical methods. We provide therefore a new and deeper insight into the phenomenon of unemployment in Germany, revealing two types of characteristics: the first consists of effects being almost identical to those found with the classical model: these effects are constant with the duration time, like the educational attainment of natives. Secondly, we found time-dependent effects, like the previous employment situation, which are likely to change and lose their importance as the unemployment goes on. Being able to analyze unemployment duration without restrictive a-priori knowledge demonstrates therefore the flexibility and the applicability of non-proportional hazard models being estimated with P-Splines in an economic context. The statistical method used could easily be applied in other fields of labour market research. Our findings have clear implications for social and labour market policies. Our results of a negative duration dependence imply that policy makers are well advised to implement measures which prevent long-term unemployment. Moreover, policy makers should continue their efforts to improve the compatibility of family and work. This includes for example an extended provision of external childcare for unemployed women.

## References

- Aaronson, D., B. Mazumder, and S. Schechter (2010). What is behind the rise in Long-Term unemployment? *Economic Perspectives Vol.34*(No.2).
- Atkinson, K. E. (2007). *An Introduction to Numerical Analysis* (2. ed.). Chichester: Wiley.
- Bover, O., M. Arellano, and S. Bentolila (2002). Unemployment Duration, Benefit Duration and the Business Cycle. *The Economic Journal* 112(479), 223–265.
- Breslow, N. E. and D. G. Clayton (1993). Approximate inference in Generalized Mixed

- Model. *Journal of the American Statistical Association*. 88, 9–25.
- 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 & Hall.
- Eilers, P. and B. Marx (1996). Flexible smoothing with b-splines and penalties. *Statistical Science* 11(2), 89–121.
- Gray, R. (1994). Spline-based tests in survival analysis. *Biometrics* 50, 640–652.
- Haisken-DeNew, J. P. and J. R. Frick (2005, December). *DTC – Desktop Companion the the German Socio-Economic Panel (SOEP)* (8. ed.). Essen, Berlin: RWI Essen & DIW Berlin.
- Hastie, R. and R. Tibshirani (1990). *Generalized Additive Models*. London: Chapman & Hall.
- Hastie, T. and R. Tibshirani (1993). Varying-coefficient models. *Journal of the Royal Statistical Society, Series B* 55, 757–796.
- Heckman, J. and B. Singer (1984). A method for minimizing the distributional assumptions in econometric models for duration data. *Econometrica* 52(2), 271–320.
- Hunt, J. (1995). The effect of unemployment compensation on unemployment duration in germany. *Journal of Labor Economics* 13(1), 88–120.
- Kaas, L. and C. Manger (2011). Ethnic discrimination in germany’s labour market: A field experiment. *German Economic Review* forthcoming.
- Kauermann, G. (2005). Penalized spline smoothing in multivariable survival models with varying coefficients. *Computational Statistics & Data Analysis* 49, 169–186.
- Kauermann, G. and P. Khomski (2006). Additive two way hazards model with varying coefficients. *Computational Statistics & Data Analysis* 51(3), 1944–1956.
- Kauermann, G. and J. Opsomer (2011). Data-driven selection of the spline dimension in penalized spline regression. *Biometrika* 98(1), 225–230.
- Klein, J. P. (1992). Semiparametric Estimation of Random Effects Using the Cox Model based on the EM Algorithm. *Biometrics* 48(3), 795–806.

- 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, 466–473.
- Machin, S. and A. Manning (1999). The causes and consequences of longterm unemployment in europe. In O. C. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume Volume 3, Part 3, Chapter Chapter 47, pp. 3085–3139. Elsevier.
- OECD (2011). *OECD.StatExtracts*. URL: <http://stats.oecd.org/Index.aspx>.
- R Development Core Team (2010). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0, version: 2.11.1.
- Ruppert, D. (2002). Selecting the number of knots for penalized splines. *Journal of Computational and Graphical Statistics* 11, 735–757.
- Steiner, V. (2001, February). Unemployment persistence in the west german labour market: Negative duration dependence or sorting? *Oxford Bulletin of Economics and Statistics* 63(1), 91–113.
- Steinhardt, M. F. (2011). The wage impact of immigration in germany - new evidence for skill groups and occupations. *The B.E. Journal of Economic Analysis & Policy Contributions* 11(1).
- Süli, E. and D. Mayers (2003). *An Introduction to Numerical Analysis*. Cambridge: Cambridge University Press.
- Therneau, T. M. and P. M. Grambsch (2000). *Modeling Survival Data – Extending the Cox Model*. New York: Springer.
- UNESCO Institute for Statistics (2006). *International Standard Classification of Education ISCED 1997*. Montreal, Canada. ISBN 92-9189-035-9.
- van den Berg, G. J. and B. van der Klaauw (2001, June). Combining micro and macro unemployment duration data. *Journal of Econometrics* 102(2), 271–309.
- Wagner, G. G., J. Frick, and J. Schupp (2007). The German Socio-Economic Panel Study (SOEP) – Scope, Evolution and Enhancements. *Schmollers Jahrbuch (Journal of Applied Social Science Studies)* 127(1), 139–169.

- Wagner, G. G., J. Göbel, P. Krause, R. Pischner, and I. Sieber (2008). Das Sozio-  
oekonomische Panel (SOEP): Multidisziplinäres Haushaltspanel und Kohortenstudie  
für Deutschland. *Wirt Sozialstat Archiv* 2, 301–328.
- Wahba, G. (1978). Improper priors, spline smoothing, and the problem of guarding against  
model errors in regression. *Journal of the Royal Statistical Society B* 40, 364–372.
- Wand, M. (2003). Smoothing and mixed models. *Computational Statistics* 18, 223–249.
- Wood, S. (2006). *Generalized Additive Models : An Introduction with R*. Boca Raton:  
Chapman & Hall/CRC.
- Wood, S. (2010). *mgcv: GAMs with GCV/AIC/REML smoothness estimation and  
GAMMs by PQL*. R package version 1.6-2.

# A Estimation Results

## A.1 Kaplan-Meier curves

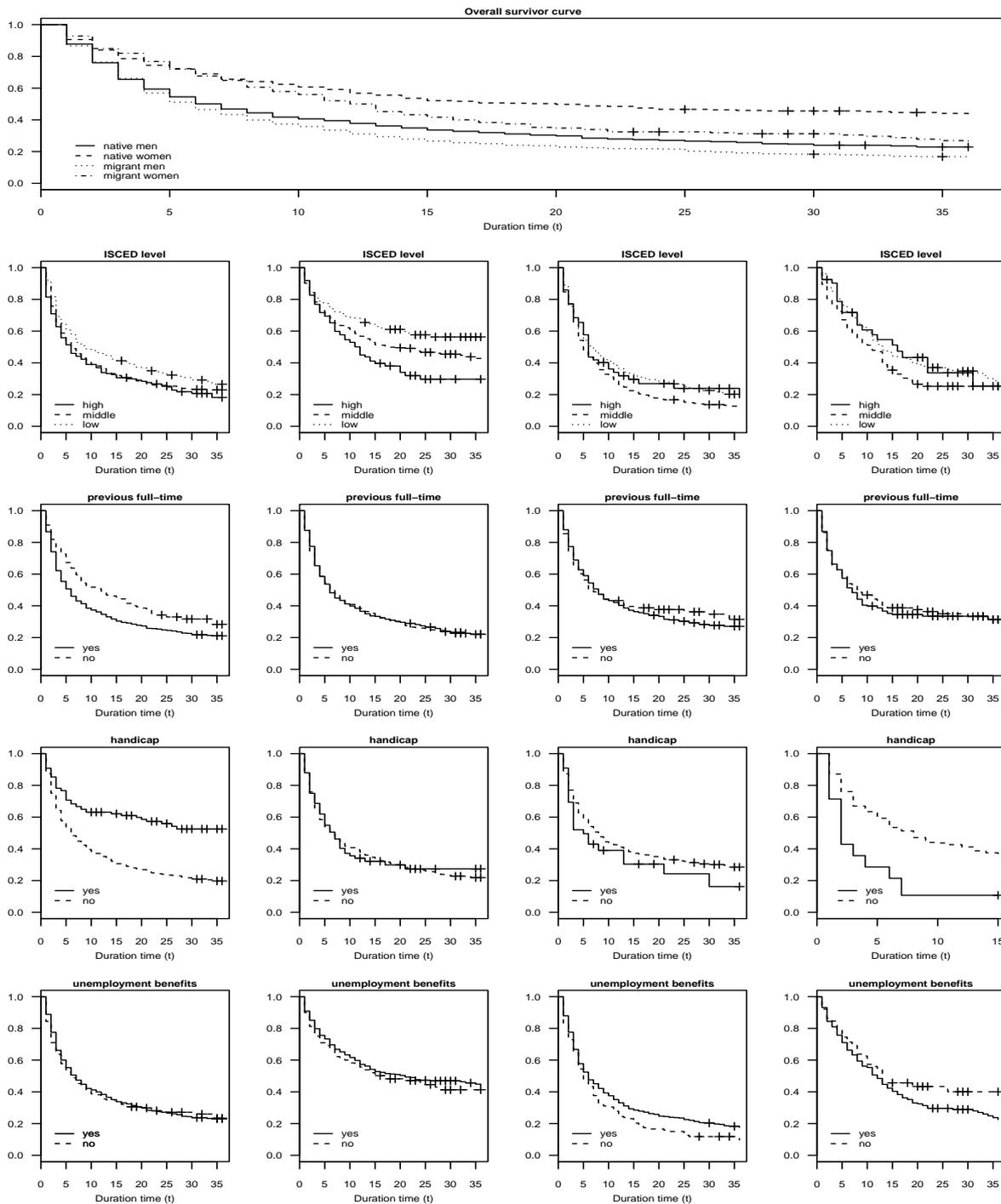


Figure 1: Kaplan-Meier curves for duration of unemployment for native men, native women, immigrant men and immigrant women (first, second, third and fourth column respectively).

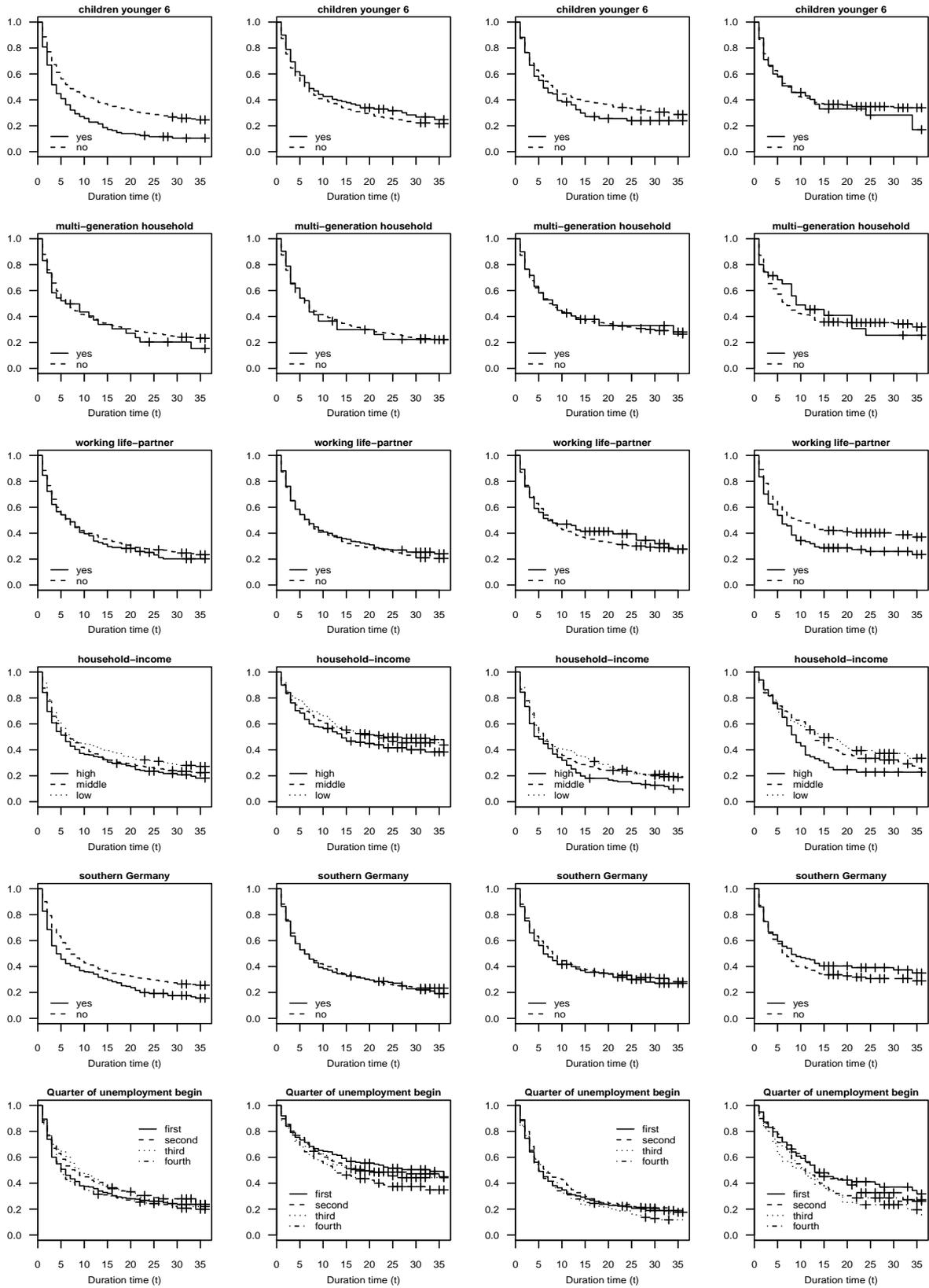


Figure 2: Kaplan-Meier curves for duration of unemployment for native men, native women, immigrant men and immigrant women (first, second, third and fourth column respectively), cont.

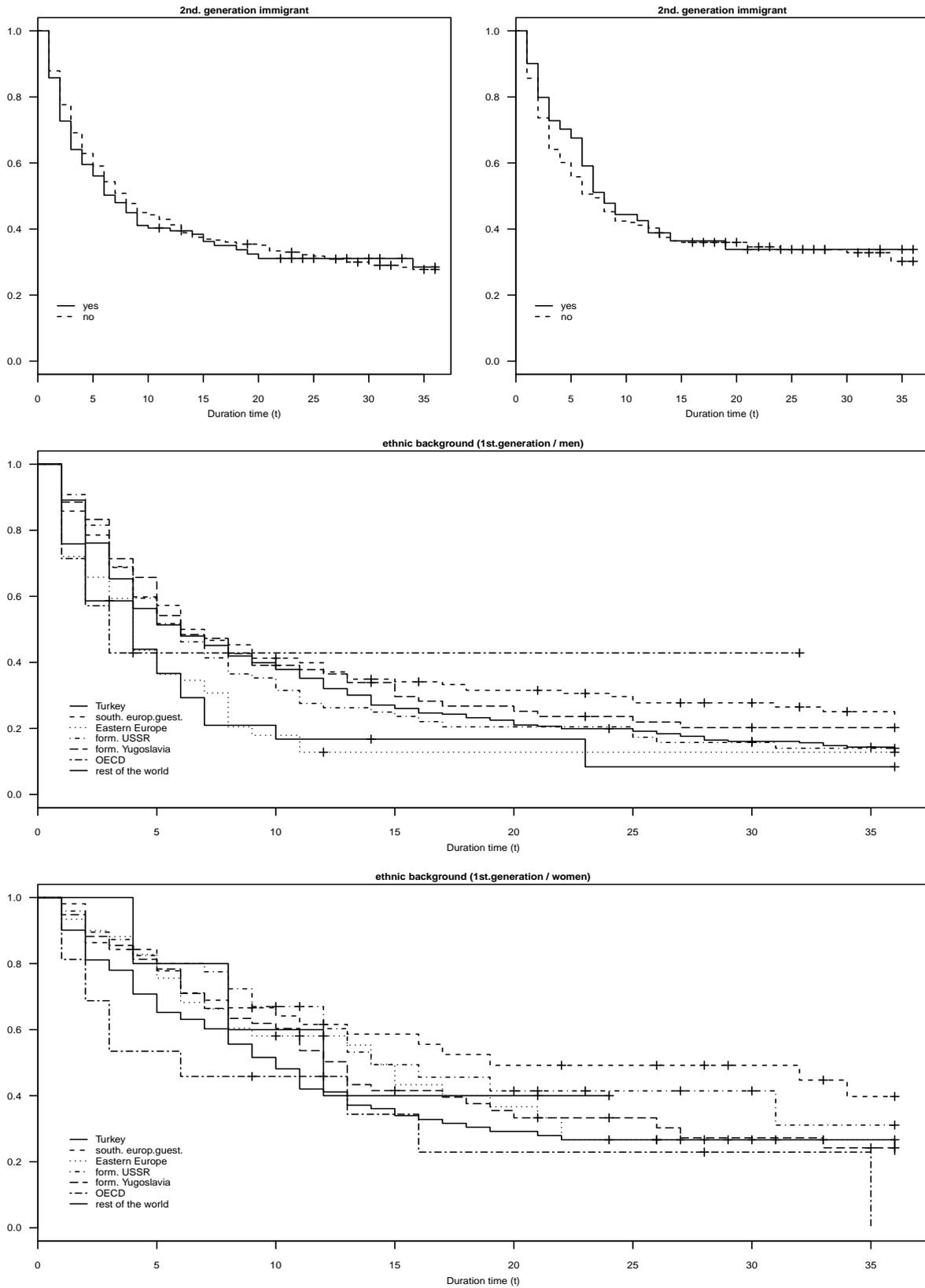


Figure 3: Kaplan-Meier curves for duration of unemployment for immigrant men and immigrant women.

## A.2 Proportional Hazard Models

effect	native men		native women		immigrant men		immigrant women	
	$\hat{\beta}_j$	p-value	$\hat{\beta}_j$	p-value	$\hat{\beta}_j$	p-value	$\hat{\beta}_j$	p-value
<i>low.ISCED</i>	-0.49	< 0.01	-0.45	< 0.01	-0.19	0.07	-0.32	0.08
<i>high.ISCED</i>	0.35	< 0.01	0.12	0.39	0.03	0.88	-0.06	0.84
<i>previous.fulltime</i>	0.76	< 0.01	0.73	< 0.01	0.57	< 0.01	0.69	< 0.01
<i>handicap</i>	-0.35	0.03	-0.29	0.25	-0.49	0.05	0.5	0.34
<i>no.ue.benefit</i>	0.11	0.26	0.21	0.07	0.3	< 0.01	-0.26	0.17
<i>children.under.6</i>	0.44	< 0.01	-1.06	< 0.01	0.06	0.66	-0.55	0.05
<i>multi.hh</i>	0.24	0.23	-0.58	0.03	-0.11	0.5 1	0.29	0.28
<i>working.partner</i>	0.23	0.03	-0.16	0.16	0.15	0.22	0.46	0.02
<i>ahinc.low</i>	-0.02	0.84	-0.27	0.04	0.15	0.21	0.19	0.36
<i>ahinc.high</i>	0.06	0.52	0.01	0.99	0.19	0.09	0.01	0.98
<i>south.germany</i>	0.41	< 0.01	0.34	< 0.01	-0.03	0.75	0.07	0.69
<i>first.quarter</i>	0.3	< 0.01	0.15	0.3	0.19	0.1	0.03	0.9
<i>third.quarter</i>	-0.08	0.55	0.15	0.3	0.11	0.44	0.52	0.03
<i>fourth.quarter</i>	0.22	0.05	0.4	< 0.01	0.22	0.11	0.23	0.32
<i>age</i>	-0.06	< 0.01	-0.07	< 0.01	-0.05	< 0.01	-0.07	< 0.01
<i>ue</i>	-0.07	0.06	-0.06	0.2	-0.01	0.7	-0.03	0.58
<i>year</i>	-0.03	< 0.01	-0.01	0.19	-0.03	< 0.01	-0.03	0.08
<i>second.generation</i>	–	–	–	–	0.28	0.07	0.23	0.44
<i>south.europ.guest</i>	–	–	–	–	0.27	0.06	-0.37	0.23
<i>Eastern.Europe</i>	–	–	–	–	1.0	< 0.01	0.27	0.39
<i>form.ussr</i>	–	–	–	–	0.48	0.12	0.18	0.59
<i>form.Yugoslavia</i>	–	–	–	–	0.36	0.04	0.19	0.48
<i>OECD</i>	–	–	–	–	1.08	0.05	0.51	0.26
<i>rest</i>	–	–	–	–	1.07	< 0.01	-0.76	0.28
$Var(\hat{v}_i)$	0.4	–	0.38	–	0.22	–	0.29	–

Table 2: Results of the proportional hazard models

### A.3 Non-Proportional Hazard Models

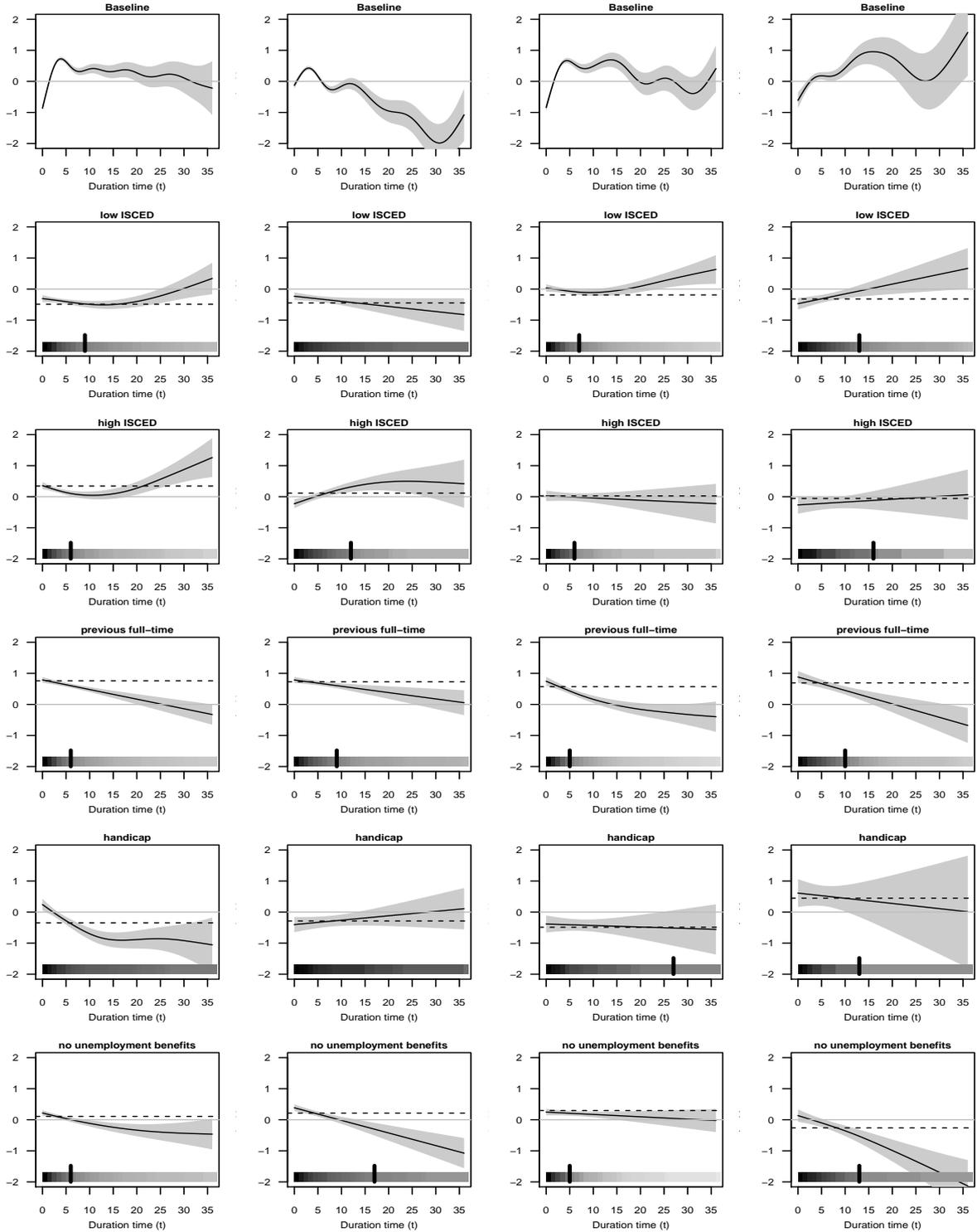


Figure 4: Fitted dynamic effects for duration of unemployment for native men, native women, immigrant men and immigrant women (first, second, third and fourth column respectively).

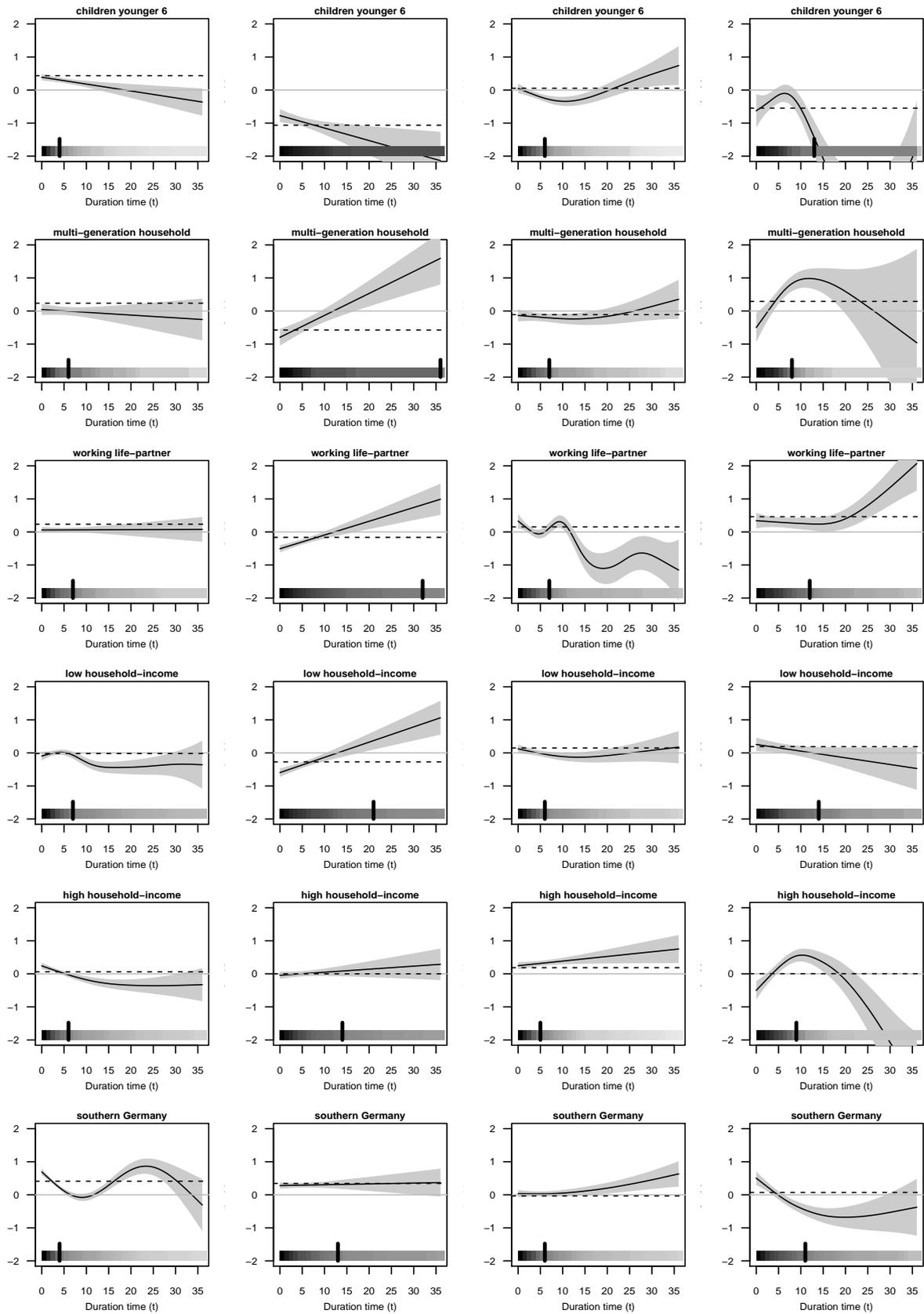


Figure 5: Fitted dynamic effects for duration of unemployment for native men, native women, immigrant men and immigrant women (first, second, third and fourth column respectively), cont.

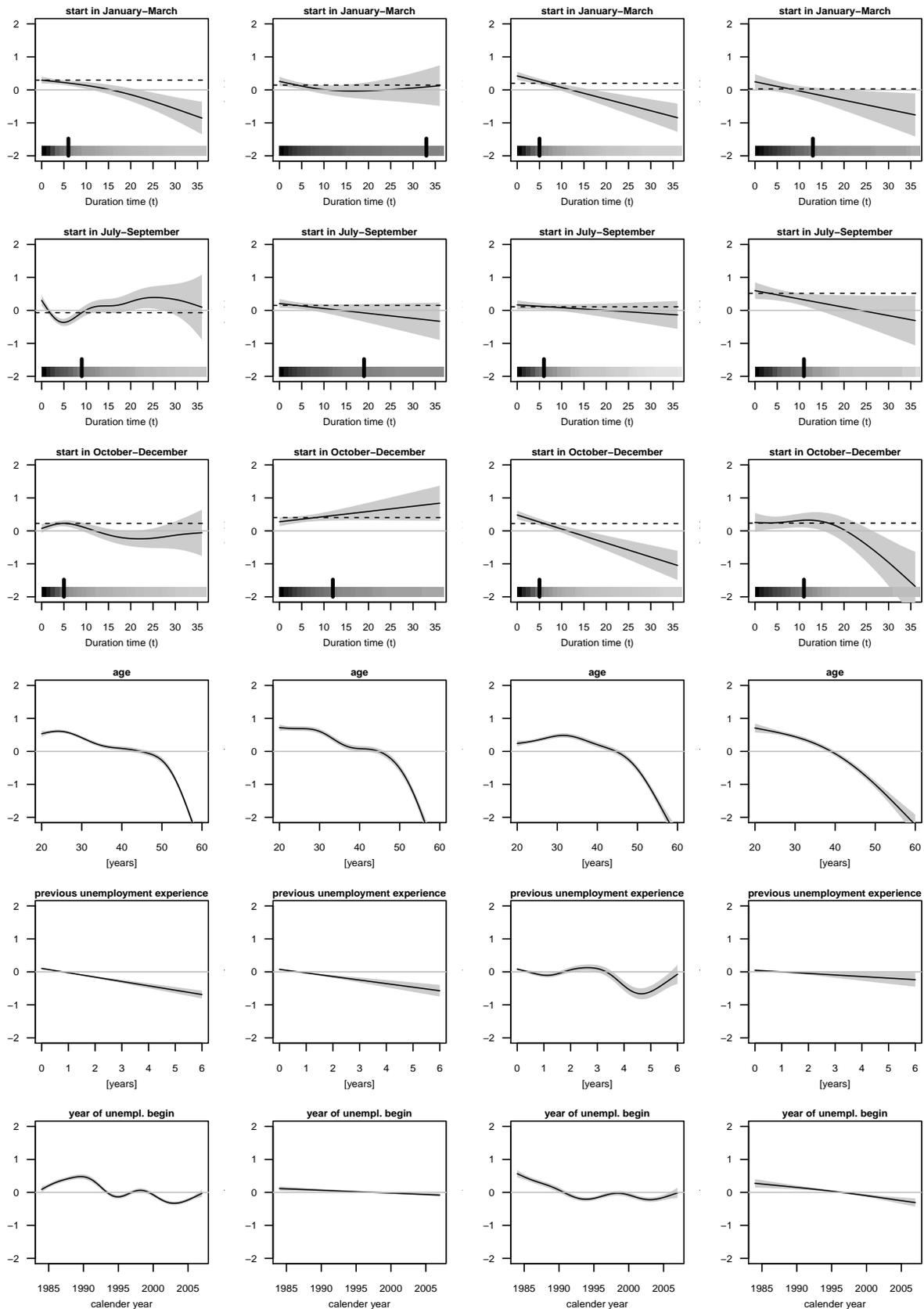


Figure 6: Fitted dynamic effects for duration of unemployment for native men, native women, immigrant men and immigrant women (first, second, third and fourth column respectively), cont.

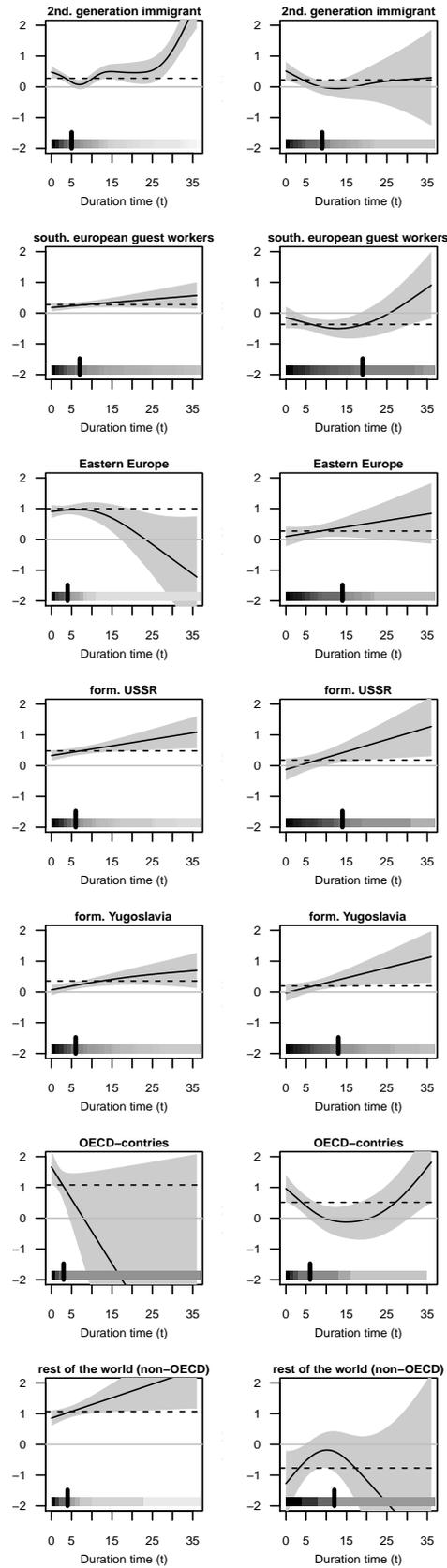


Figure 7: Fitted dynamic additional effects for duration of unemployment immigrant men and immigrant women (first and second column respectively).