Discussion Paper No. 04-26

New Estimates of the Duration and Risk of Unemployment for West-Germany

Ralf A. Wilke

ZEW
Zentrum für Europäische Wirtschaftsforschung GmbH
Centre for European Economic Research

Discussion Paper No. 04-26

New Estimates of the Duration and Risk of Unemployment for West-Germany

Ralf A. Wilke

Non–technical Summary

Detailed knowledge of the main micro- and macro-determinants for the length of individual unemployment periods is indispensable for the successful design of labor market policy measures. This paper presents detailed stylized facts about the risk of becoming unemployed and the distribution of unemployment duration. The IAB employment subsample 1975-1997 is used for the estimations. It is register data containing daily information about the employment and unemployment status of about 500.000 individuals from West-Germany. The analysis is based on a nonemployment proxy for unemployment since registered unemployment is not observed in the data. The huge number of observations allows to perform comprehensive nonparametric analysis for homogenous data segments, where the analysis is restricted to the main workforce aged 26-41. It is shown that the probability of remaining unemployed after a certain period varies significantly over many of the considered population segments. It is suggested that the effect of the macroeconomic variation differs across the segments and that there are general developments over the decades due to behavioral changes in the society. These effects are sometimes non-proportional over the calender time and over the duration time. The probability of becoming long-term unemployed is mainly between $20-50\%$ for the males and between $40-60\%$ for the females. Surprisingly, these shares did not increase from the beginning of the 1980ties until the mid 1990ties while the unemployment rate almost doubled during this period. Consequently, in many cases observable micro and macro variables cannot explain why unemployed exit to employment and why not. The important role of unobserved factors has to be taken into account when designing labor market policies such as further training measures.

The estimation results show that many unemployed leave unemployment during the first three months of the unemployment duration. The decline of the estimated survivor functions then decreases sharply and for long-term unemployed it is often almost zero. Once unemployed it does not become evident that higher educational degrees are a reliable protection against becoming long-term unemployed. However, individuals with higher educational degrees seem to experience a lower risk of unemployment given employment especially for males. Unskilled workers have the highest risk of unemployment, the longest unemployment durations and the largest fraction of long term unemployment. The business cycle mainly affects the unskilled workers, foreign nationals and females. Having a university degree seems to be a good protection against unemployment. Females have in general longer unemployment periods and a higher probability of long term unemployment but this seems to be (mainly) due to the married females. There is strong evidence that married females have the most favorable development over the two decades under consideration.

New Estimates of the Duration and Risk of Unemployment for West-Germany

Ralf A. Wilke*

March 2004

Abstract

This paper analyzes changes in the risk of unemployment and changes in the distribution of unemployment duration for the 26 to 41 years old working population in West-Germany during the 1980ties and 1990ties. The comprehensive IAB employment subsample 1975-1997 is used for the analysis. It contains employment and unemployment trajectories of about 500.000 individuals from West-Germany. The application of flexible nonparametric estimators yields results which are less sensitive to misspecification as it is often the case for parametric hazard rate models. By conditioning on several observable variables such gender, education, marital status etc. we identify significant differences in the first three quintiles of the distribution of the length of unemployment duration. A large share of long term unemployment with only few exits to employment is observed in almost any of the segments. The analysis also considers general evolutions over time and variations along the business cycle. The paper therefore provides a collection of detailed stylized facts about the distribution of unemployment durations in West-Germany during the past two decades.

Keywords: risk of unemployment, duration of unemployment, German register data, nonparametric analysis

JEL: C14, J64, J65

*I gratefully acknowledge financial support by the German Research Association (DFG) through the research project "Microeconometric modelling of unemployment durations under consideration of the macroeconomic situation". Thanks are due to Bernd Fitzenberger, Elke Lüdemann and Xuan Zhang for the joint work on the project. Thanks are also due to a number of employees in the employment offices at Duisburg, Freiburg, Gelsenkirchen, Mannheim, and Oberhausen for many useful hints and information. Thanks are also due to Melanie Arntz, Wolfgang Franz and Andrea Weber for giving me comments. ZEW Mannheim, P.O.Box 10 34 43, 68034 Mannheim, Germany, Email: wilke@zew.de

1 Introduction

The rising unemployment in Germany is becoming a more and more severe problem. Several policy changes and billions of Euros of public spending seem not to result in a turn around of this tendency. Obviously, a detailed knowledge of the main micro- and macro-determinants for the length of individual unemployment periods is indispensable for the successful design of policy measures. It is therefore of fundamental interest to explore the distribution of the length of individual unemployment periods in different macro environments given observable individual characteristics. This information helps us in examining how the business cycle has an impact on the length of individual unemployment periods and whether this change is the same for all individuals. Collecting detailed stylized facts may help in obtaining ideas about the main micro- and macroeconomic determinants of the risk of unemployment and the distribution of the length of individual unemployment periods. The analysis of this paper is restricted to the main workforce of mid aged individuals so that the results are not affected by several policy measures for young unemployed and by the early retirement issue for unemployed with extended entitlements for unemployment insurance (Fitzenberger and Wilke, 2004).

Recent unemployment duration studies for (West-)Germany are mainly based on the German Socio-Economic-Panel (GSOEP) using single spell hazard rate models, e.g. Hunt (1995), Steiner (2001) and Lauer (2003). The GSOEP is monthly interview data with a rather limited sample size but it provides a variety of explanatory variables. However, many of them are subject to measurement errors due to imperfect memory of the interviewed individuals or due to intentionally misleading replies. Schräpler (2002) analyzes the non-response behavior of the households. Another problem is the panel attrition. The limited sample size of the GSOEP does not allow for detailed exploratory analysis, since the sample size in the cells decreases rapidly while segmenting the data. Hunt (1995) provides limited nonparametric duration analysis by comparing individuals who are subject to a reform of the unemployment compensation system to other individuals. The specification of a common duration model is therefore the classical modelling approach when using interview data. Correctly specified models yield consistent estimates of the model coefficients. The above mentioned contributions apply a variety of (mixed) proportional hazard models or related frameworks in discrete time. Hunt (1995) uses the Cox-proportional hazard model, i.e. she ignores the possibility of unobserved heterogeneity and she does not specify the baseline hazard function. Steiner (2001) and Lauer (2003) use discrete time models with piecewise linear baseline hazard rates and a discrete distribution of unobserved heterogeneity. Simulations studies suggest that single spell approaches to (semi)- parametric duration models have several general drawbacks in

finite samples. Van den Berg (2001) gives a summary of the recent literature and concludes that "estimation results are sensitive to misspecification of the functional forms associated with the model determinants. Therefore, interpretations of those results are often unstable and should be performed with extreme caution." He also points out that an application of these models requires a deep prior knowledge of the main model determinants.

This paper aims at exploring the micro- and macro-determinants of the exit from unemployment with a nonparametric survival analysis using the IAB employment subsample. [1] Nonparametric estimates are less subject to misspecification and yield consistent estimates for a wide range of models. However, they do not allow for inference because the estimates might be affected by the compositions of the corresponding (sub-)samples in terms of other variables. The IAB employment subsample is comprehensive German register data. It provides enough information even if the data is segmented in several sub-samples by conditioning on observables. The obtained stylized facts provide information for the setup of a duration model and one can scrutinize whether duration models can explain stylized facts. This can for example be done by comparing the results of the recent contributions using the GSOEP.

Section 2 describes the data. Section 3 provides basic information about the macroeconomic situation. It also discusses the risk of unemployment given employment in the period of observation. Section 4 introduces the framework of the nonparametric survival analysis and section 5 presents the corresponding results. The last section summarizes the main findings.

2 Data and Description

The IAB employment subsample 1981-1997 -regional file- is used for the estimation. It is German register data and contains spell information of employment and unemployment trajectories of about 500.000 individuals from West-Germany. It is representative with respect to the socially insured working population. The data provides daily information about the starting and the ending of socially secured employment and of any receipt of unemployment compensation from the federal employment office (BA). Self-employment and employment as life-time civil servant (Beamte) are

[1] Fitzenberger and Wilke (2004) analyze with the same data the effects of the reform of the German unemployment compensation system in the 1980s. Plaßmann (2002) also analyzes this reform using similar data. She provides descriptive analysis and estimates a parametric proportional hazard model without unobserved heterogeneity. Her approach does not make use of the extreme richness of the data, she does not model the effects of the business cycle and she ignores the issue of early retirement.

not observed. The latter fact is not problematic for our analysis because life-time civil servants generally do not become unemployed. By not observing self-employment some useful information is lost because self-employment is often considered as eligible in order to leave unemployment. For further details about the data see Bender et al. (2000). Registered unemployment is not recorded and therefore one cannot precisely distinguish between unemployment and nonemployment periods because unemployment periods without receipt of unemployment compensation from the BA are not observed. For this reason we only consider two states: (socially secured) employment and nonemployment, where nonemployment is any time in which the individual is not (socially secured) employed and receives at least for one day some kind of unemployment compensation from the federal employment office. The latter condition ensures that at least a part of each nonemployment period overlaps with unemployment. A nonemployment spell is marked as right-censored if the last observed spell of an individual is the receipt of unemployment compensation. There is no left censoring by conditioning the sample on previous employment. This definition of unemployment using the IAB employment subsample is introduced by Fitzenberger and Wilke (2004) and is referred to simply as unemployment in what follows.

The analysis of this paper is restricted to unemployment spells of west-Germans [2] aged 26 to 41 which start between 1981 and 1995. The age restriction is chosen for the following reasons: the maximum entitlement for unemployment insurance for individuals above 41 years was subject to a reform between 1985 and 1987. Therefore we may expect changes in the distribution due to the policy change which are analyzed by Hunt (1995), Fitzenberger and Wilke (2004) and others. The observations of this analysis are restricted to a population that has a maximum of 12 months entitlements for unemployment insurance. Young people below 26 are not considered because it is expected that many of them are still in education and there are policy programs against youth-unemployment which may also systematically affect the distribution of the length of unemployment. [3] In the following analysis the data is segmented into homogenous sub-samples by conditioning on one or several explanatory variables such as gender and marital status that are available in the IAB-Employment sample. See table 4 in the appendix for getting an overview of the considered data segments with the respective sample sizes.

[2] In this analysis an individual is said to be west-German if the last employment period before unemployment was in West-Germany.

[3] Indeed, preliminary estimations suggest that the probability of becoming long-term unemployed of aged < 26 has significantly decreased during the observation period.

3 Macroeconomic Variation and Risk of Unemployment

This paper intents to explore the differences in the risk of unemployment given employment[4] and in the distribution of the length of unemployment durations for homogenous sub-populations taking into account macroeconomic variations such as the unemployment rate. Figure 1 presents the west-German unemployment rate in the period of interest. It is easy to see that it rose from 4% in the beginning of the eighties to more than 11% in 1997. There are periods of sharp increase, i.e. 1980-1983 and 1993-1997. 1984-1988 is a period of stagnation and the only period with an evidently decreasing unemployment rate is the time during and after the German reunification, i.e. 1989-1991. From Figure 1 it is also apparent that the average risk of unemployment given employment is related to the unemployment rate: the increase or decrease (Δ) of the current risk of unemployment is similar to the increase or decrease of the west-German unemployment rate two periods ahead and therefore the former may be used as a predictor of the latter (figure 1, right).[5]

Let us shed more light on the risk of unemployment given employment. It is interesting to see how this proportion varies for different segments of the data. Figures 2 and 3 present different functions, while conditioning on gender, education and citizenship. It is evident that unskilled (German) males have on average the highest risk of unemployment given employment, whereby males with a university degree have the lowest risk. It is also apparent from the figures that there is almost no variation over the educational groups for females. It seems only for males that education is the best insurance against unemployment. It becomes also clear that the average risk of unemployment for unskilled is more sensitive to the business cycle than for other educational groups. For individuals with university degree it is almost constant. Moreover, it is apparent that the risk of unemployment for foreign nationals surged during the recession of the nineties. This might be an indication that it is relatively more difficult for foreign nationals to keep their job in a weakening labor market. It is remarked that the findings are stylized facts and that the composition of the different (sub-)samples may affect the results.

The following nonparametric analysis of unemployment durations focuses on four different years (1981, 1985, 1990 and 1995), each of them in one of the above mentioned rather different

[4]This is defined as the ratio of the number of failures (number of individuals moving from employment into unemployment) and the number of observations at risk (number of employed individuals) in a specific period.

[5]This is a very simple relationship and of course there is space for improvements but this is out of scope of this paper.

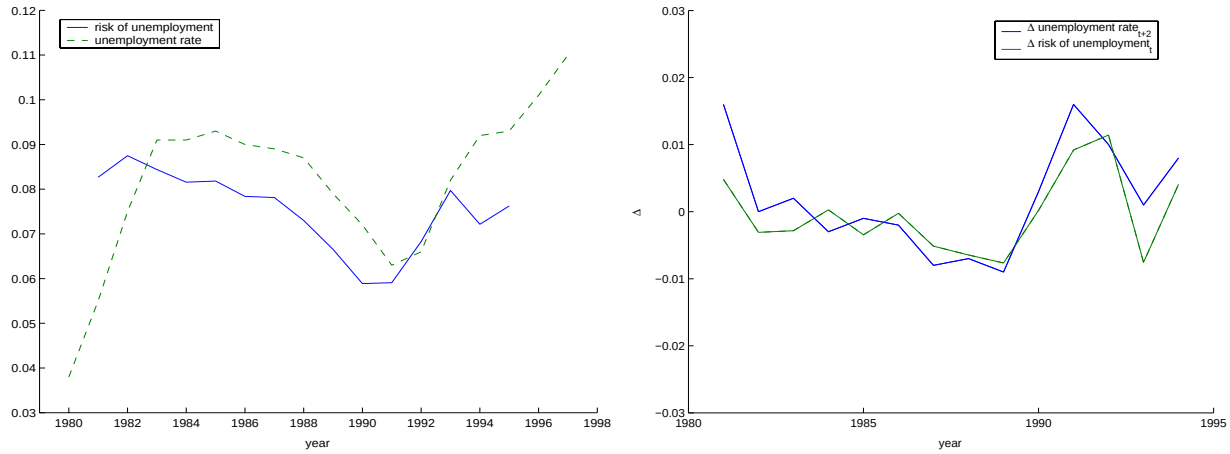

Figure 1: Yearly unemployment rate in West-Germany, the average risk of unemployment given employment (left) and how they are related (right)

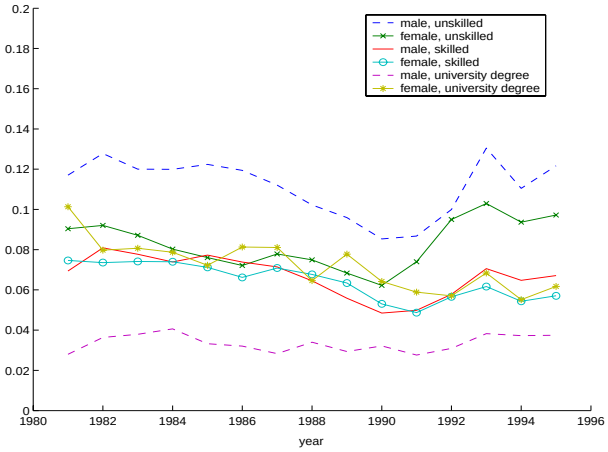

Figure 2: Average risk of unemployment given employment stratified by gender and education

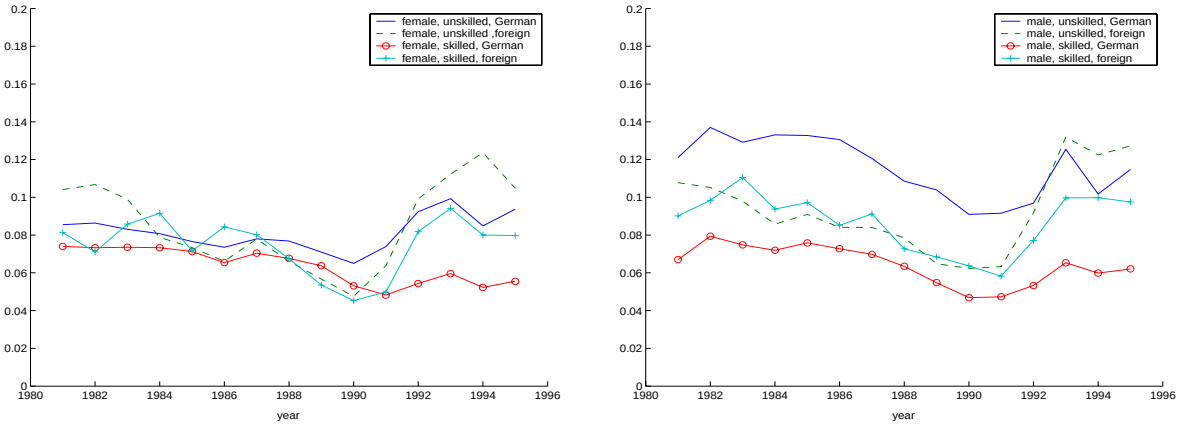

Figure 3: Average risk of unemployment given employment stratified by gender, education and citizenship

macroeconomic environments. This may allow us to capture the main evolution over the two decades and in addition it may provide us with information about the impact of the business cycle. All the gathered information can then be used for the setup of a duration model which allows one to make statistical inference. Let us now briefly describe the macroeconomic situation of the years under consideration. In general, job search theory suggests that a weak labor market, i.e. in periods of rising and high unemployment rate, yields on average in longer unemployment durations than a tight labor market, i.e. in periods of declining and low unemployment rate. This is simply because the competition for an existing open job is bigger in weak labor market. Given a competitive economy this results in a lower job offer arrival rate and in lower wages. The latter reduces the probability of the unemployed to accept the job offer given that he receives some form of unemployment compensation.

Year 1981 The beginning of the eighties is characterized by a quite low but sharply rising unemployment rate. The rise continues until 1983. We may therefore expect that this macroeconomic environment results in longer unemployment durations, since it is expected that companies hire less and lay off more in this and in the consecutive years.

Year 1985 This year has the highest unemployment rate in the eighties. It is followed by several years of stagnating and declining unemployment rates. We may therefore expect here that it is a brightening environment for the unemployed.

Year 1990 This year is characterized by a tight west German labor market during the economic boom period after reunification. The unemployment rate is falling to the lowest level in the nineties (in 1991) and the lowest since 1982. We should therefore expect shorter unemployment durations for many individuals.

Year 1995 Due to a recession, the unemployment rate in the mid-nineties is at a high level and still rising. In 1995 the unemployment rate is almost back to the level in 1985 but it surges to the highest level ever in 1997. Therefore it should be a very difficult environment for unemployed and economic theory predicts us the longest unemployment durations in the period under consideration.

4 Nonparametric Survival Analysis

This section introduces the main tools for the nonparametric survival analysis which allow exploring the impact of macroeconomic and microeconomic observables. The probability of remaining unemployed after T days is

$$Prob(t \geq T) = 1 - F(T) = S(T),$$

where F is the cdf and S is the survivor function. The corresponding hazard rate is defined as $\lambda(T) = f(T)/S(T)$, where f is the pdf. The minimum unemployment spell-length with survival probability $\theta \in [0, 1]$ is given by

$$\inf \{T\}, \quad \text{s.t. } S(T) \leq \theta.$$

Note that $S(T)$ is weakly decreasing and therefore $T = S^{-1}(\theta)$ may not exist.

Suppose there is a sample of durations $t_{i=1,...,n}$ with distinct values $\tau_{j=1,...,J}$, where n is the number of observations and J is the number of distinct duration spell-lengths in the sample. The survivor function cannot be estimated by the empirical survivor function in the case of censoring.[6] Instead one may use

$$\hat{S}(T) = \prod_{\tau_j < T} (1 - \hat{\lambda}_{\tau_j}),$$

where $\hat{\lambda}_{\tau_j}$ is the Kaplan-Meier estimate of the hazard rate at time τ_j

$$\hat{\lambda}_{\tau_j} = a_j / r_j$$

where a_j is the number of uncensored durations of length τ_j, and r_j is the number of durations i with $t_i \geq \tau_j$. It is well known that

$$\sqrt{n}\left(\hat{S}(T) - S(T)\right) \sim N\left(0, \hat{S}(T)^2 \sum_j \frac{a_j}{r_j(r_j - a_j)}\right)$$

as $n \to \infty$. Using this we may obtain the corresponding $\underline{S}(T, \alpha)$ and $\overline{S}(T, \alpha)$ for any $\alpha \in [0, 1]$ such that $\text{Prob}(\hat{S}(T) \in \left[\underline{S}(T, \alpha), \overline{S}(T, \alpha)\right]) = 1 - \alpha$. Then we obtain confidence bands $\underline{T}_\theta$ and $\overline{T}_\theta$ for $\hat{T}_\theta$ by

$$\inf \{T\}, \quad \text{s.t. } \underline{S}(T, \alpha) \leq \theta$$
$$\inf \{T\}, \quad \text{s.t. } \overline{S}(T, \alpha) \leq \theta$$

[6] Suppose we observe t_i^* and not t_i, where $t_i^* = min\{t_i, C_i\}$ with C_i as the individual specific censoring time. The Kaplan-Meier estimator yields consistent estimates in the present framework of right-censoring.

In the following analysis $\hat{T}_\theta$ is estimated for the whole population and for several sub-populations in the years of interest. Using homogenous sub-populations corresponds to estimating conditionally on observable variables, i.e. the conditional survivor $S(T|x)$ is estimated, where x is a vector of explanatory variables. It is well known (e.g. Koenker and Geling (2001)) that common parametric frameworks of duration analysis such as the proportional hazard model, the accelerated failure time model and the proportional odds model induce that the parametric term yields parallel shifts of the quantile functions, i.e.

$$Quant_T(\theta|x) = x'\beta + F_T^{-1}(\theta),$$

where β is a vector of unknown parameters. This implies that the coefficients do not depend on the quantile and that the survivor functions cannot cross. Strong non-proportional shifts of the survivor functions may therefore indicate that the model specification of the above mentioned parametric frameworks is incorrect.

5 Estimation Results

Figure 4 presents the unconditional estimated survivor functions for the first two years of the unemployment duration in the four years of interest. It is evident that the magnitude of the slope of the survivor function is monotonically decreasing in the duration time. By looking at the shape of the estimated survivors, the first two years of duration can be decomposed into three intervals: the survivors are linearly decreasing in the first three months of the duration. From month three until the 12'th month this decrease is softened. After month 12 the survivor is again linearly decreasing at a decent rate. This suggests that the density of the distribution of unemployment spells is monotonically decreasing with the duration, in particular between month three and month 12 of the duration.[7] These findings do not suggest that many unemployed wait until their entitlements for unemployment insurance (which are typically 12 months) are exhausted. This observation should not be overestimated since the counterfactual outcome, i.e. a system without unemployment insurance, is not observed. Economic theory is roughly confirmed when comparing the estimated survivors in the years of interest. It appears that they are the lowest in years with constant or decreasing unemployment rates (1985, 1990) and higher in years with an increasing unemployment rate (1981,1995). It is also evident that the year 1995, which is in addition characterized by a high level of unemployment, shows the highest survival probabilities in unemployment.

[7] This is very evident for the years 1981, 1985 and 1990. In 1995 the decrease of the slope between month 3 and month 12 is less strong.

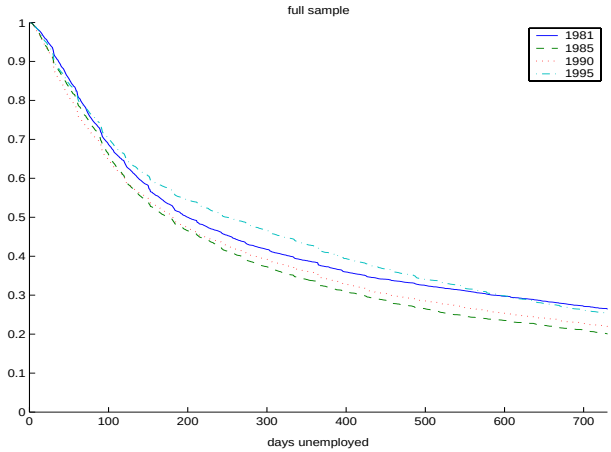

Figure 4: Kaplan-Meier-survival function estimates using full sample sizes.

Turning to a more detailed analysis, let us now consider the evolution of $\hat{T}_\theta$ over the four years and let us compare the homogeneous sub-populations within a respective year relative to the unconditional estimate. $\hat{T}_\theta$ is estimated for $\theta = \{0.8, 0.6, 0.4\}$, which corresponds to the lower three quintiles of the (conditional-)distribution of unemployment durations. Other quintiles are not considered because $\hat{T}_\theta$ and the respective confidence bands are simply too large for some data segments. Hence, the analysis is restricted to intervals, in which the survivor function is sufficiently decreasing.

The estimation results of $\hat{T}_\theta$ are presented in tables 5-7 (appendix). The comparison over the years using 1981 as a benchmark and the comparison of the sub-populations in the respective years are given in tables 1-3. The corresponding Kaplan-Meier survival function estimates for the first two years of duration are depicted in figures 5-10. Let us now turn to the main findings of the nonparametric survival analysis by exploring possible effects due to observable individual characteristics and due to the macroeconomic variation. Note that this analysis is not like true inference because the estimation results may be affected by compositional effects. However, detailed stylized facts help us in identifying some of the determinants of the length of unemployment.

Citizenship German males leave unemployment fastest, whereby female foreign nationals are the slowest to leave unemployment. This is true for all years and any considered quintile. The results for the foreigners are a stylized fact but probably this group of individuals appears worse due to compositional effects, e.g. educational status. Figure 5 shows that females exit unemployment at a much lower rate than their male counterparts, especially in the first three months of unemployment.

Education It is apparent for the considered quintiles that skilled[8] males leave unemployment fastest and unskilled females leave slowest. Interestingly, males with university degree[9] tend to stay longer unemployment than skilled males, whereby females with university degree leave unemployment faster than skilled or unskilled females. Skilled females experienced a favorable development over time. In the lower quintiles it is more difficult to observe a clear tendency over time for the educational groups of the males. It seems that the length of unemployment periods increases for skilled and unskilled males in the upper quintiles. Figure 6 shows that the slope change of the estimated survivor is less strong over the duration for individuals with university degree. The marginal probability of leaving unemployment does therefore decrease at a slower rate for academics compared to the skilled and unskilled. This is probably because recalls and seasonal effects are less common for academics and maybe the matching between employer and employee requires more time for many high skilled. It might also be due to a lower depreciation rate of human capital. Anyway, this observation is an indication for nonproportional effects.

Marital status There is clear evidence that married males leave unemployment fastest and married females tend to stay longest. This is the case for all quintiles of interest and in the years under consideration. At the same time it can be seen in the tables and in figure 7 that the gap between the two groups is sharply decreasing over time. While the group of married males is the one with the worst development over time, the group of unmarried females experienced the most favorable development over time in the first quintile and the group of married females in the second and in the third quintile. The slowdown of the married males is in particular during the nineties. This is an indication for a general change of the time allocation decision process within the households.

Profession Four characteristics of this variable are considered for males only. It is apparent that unemployed males with a profession related to agriculture leave unemployment fastest in three quintiles of interest, whereby technical professions are the slowest. For agricultural professions this is probably due to seasonal reasons because many individuals loose their job during the winter-period and are immediately reemployed in spring. Technical professions maybe require most time for the job match process due to highly specialized skills. When looking at the time path it is evident that manufacturing professions perform relatively best in the first two quintiles, technical professions in the third quintile and agricultural professions perform worst in the first two quintiles. Figure 9 shows that the probability of leaving unemployment decreases sharply after a duration

[8]Individuals with completed apprenticeship are marked as skilled workers.

[9]This includes individuals with a degree from a university or from a university of applied sciences

of six months for professions related to agriculture and that this decrease is slowest for technical professions.

Part time The differences between female full-time and part-time workers seem to be small.[10] Female part-time workers have the tendency to leave unemployment faster in the first quintile. There is no clear difference in the other quintiles. The results suggest that female full-time workers experience a relatively more favorable development over the years than the female part-time workers. For the part time workers the changes over time appear to be quite disproportional (figure 8).

Recall This variable is defined as if the unemployed individual experienced a recall to the former employer at the end of his last unemployment period. [11] This can only be the case if the unemployed was at least once unemployed in the past. The estimated survivors for these groups of individuals are presented in figure 10. From the tables it is apparent that perviously recalled unemployed stay shorter periods in unemployment than the average unemployed. This difference is in particular evident in the lower quintiles where the previously recalled unemployed stay only one third or half of the time in unemployment than the average unemployed. Future recall is defined as if the current unemployment period ends due to a recall to the former employer. The Kaplan-Meier estimator coincides is this case with the empirical distribution function. The estimated survivors are presented in figure 11. It is apparent that 90% of the recalls for the males and 70% of the recalls for the females arrive within 200 days. The distribution for males seems not to be affected by the business cycle. The change in the year 1990 might be due to an exceptional situation after the German reunification. The distribution for the females is monotonically shifted to the left over the years. Moreover, an obvious kink after one year of duration emerges over the time period under consideration. This means that more and more recalls arrive after one year. It should be investigated in more detail whether this is somehow related to the unemployment compensation system. For further descriptive results about recalls see Plaßmann (2002).

The business cycle It is observed that the variation of the distribution of unemployment periods over time is greater for females and for unskilled workers. In particular this is the case in the lower quintiles. These groups possess relatively better chances in getting a job in boom

[10] Part time is not considered for males because of a lack of observations, i.e. there are only very few unemployed male part time workers.

[11] In fact it has to be a recall to the same branch. Other recalls are not identifiable from the data and therefore the recall variable underreports the total amount of recalls.

periods compared to situations of economic slowdown. In particular the unskilled workers are the big losers of the mid nineties recession. Whether this is due to the globalization, technical progress or also reinforced by illegal employment and due to (legal) cheap manpower from eastern European countries has to be examined in more detail. In contrast, unemployed with university degree seem to have a lower variation in their survival probabilities in unemployment.

Evolution over time From the beginning of the eighties until the end of the nineties the labor market participation rate of the females in West-Germany has risen from 33.8% to 39.6%. This fact is important because the presented results are based on the nonemployment definition of unemployment which consists of unemployment periods plus an eventual period where the respective individual is out of the labor market. A reduction of the out of the labor market periods can therefore yield a reduction of the unemployment durations. It seems that this fact affects the estimation results for the (married) females and helps in explaining that the classical gap between married males and married females is reduced by 50% over the two decades. It also becomes apparent that the female foreign nationals stay longest in unemployment and that they did not experience a favorable development over the period of observation. In particular the group of skilled married males have increasing unemployment durations over the period under consideration.

Long term unemployment If an unemployment period lasts for more than 12 months, the corresponding unemployed individual is said to be long term unemployed. The above described findings suggest that the marginal probability of leaving unemployment is low for long term unemployed. The nonparametric analysis cannot explain why it is low but it can show us for which data segments the probability of becoming long term unemployed is greater and for which data segments it is smaller. Figure 4 tells us that the unconditional probability of becoming long term unemployed is between 30% (in 1985) and 45% (in 1995) depending on the year. Interestingly, this share did not increase from 1981 to 1995 despite a doubling of the unemployment rate during this period. Since the data does not contain registered unemployment the resulting probabilities can deviate from the official statistics. The chosen definition of unemployment generates unemployment spells which may contain out of the labor force periods. This systematically increases the length of unemployment periods. At the same time unemployment periods are not considered if the respective individuals do not receive, at least for a short period, some form of unemployment compensation from the BA. It is not sure how this selection affects the results but it is expected that especially unskilled workers and females may not meet this requirement. When looking at the specific data segments (figures 5-10) it is observed that unemployed females have a higher

probability in becoming long term unemployed and in particular unskilled unemployed have a higher probability of becoming long term unemployed. Briefly speaking, there is only one data segment in which the probability of becoming long term unemployed is less than 10%: males getting a recall or who already got a recall in the past (exception: 1995). In the other segments this probability varies between 20% and 60%. Unobserved heterogeneity is therefore a very important determinant why some individuals leave and others do not leave unemployment. Steiner (2001) focuses on the question whether the low re-employment probabilities for long term unemployed are due to a negative sorting effect over the duration time or due to negative duration dependence. His results are mixed and they have to be considered as a first benchmark. Due to the weak finite sample performance of single spell proportional hazard models with unobserved heterogeneity, the limited sample size of the GSOEP and the measurement errors in the data there is still a lot of room for improvement. It remains therefore for future research to find more stable explanation for the low re-employment rates of long term unemployed and in addition for the high probability of becoming long-term unemployed in Germany.

			1981	1985	1990	1995	1981	1985	1990	1995
All			100	100	100	100	100	94	80	98
Citizenship	male	German	82	84	88	94	100	96	87	113
		foreign	106	98	119	116	100	87	90	107
	female	German	142	148	117	113	100	98	66	78
		foreign	178	152	177	211	100	80	79	116
Education	male	unskilled	102	95	104	113	100	88	82	109
		skilled	78	79	85	92	100	94	86	116
		university	83	105	106	95	100	119	102	113
	female	unskilled	160	164	135	147	100	96	67	90
		skilled	142	148	115	111	100	98	65	77
		university	109	130	113	122	100	111	83	110
Marital Status	male	unmarried	100	100	104	98	100	94	83	97
		married	77	77	75	94	100	94	78	120
	female	unmarried	118	105	102	95	100	83	69	79
		married	162	197	125	148	100	114	62	90
Profession	male	agriculture	49	38	60	91	100	72	97	181
		manufacturing	86	84	87	95	100	91	80	109
		technical	94	126	113	141	100	126	97	148
		services	91	98	102	97	100	102	90	105
Part Time	female	(yes)	145	134	100	144	100	87	55	98
		(no)	142	151	119	117	100	100	67	82
Recall	male	(yes)	74	72	71	75	100	92	77	100
	female	(yes)	80	74	69	72	100	87	69	88

Table 1: $\hat{T}_{0.8}$ relative to all observations (left) and relative to 1981 (right)

			1981	1985	1990	1995	1981	1985	1990	1995
All			100	100	100	100	100	90	88	114
Citizenship	male	German	74	81	84	80	100	99	100	123
		foreign	107	105	102	99	100	88	83	105
	female	German	159	198	127	147	100	113	70	105
		foreign	187	191	146	210	100	92	69	128
Education	male	unskilled	90	91	101	99	100	91	98	125
		skilled	72	76	77	77	100	96	95	122
		university	88	111	101	95	100	114	101	124
	female	unskilled	180	199	135	199	100	100	66	126
		skilled	172	201	124	139	100	106	63	92
		university	128	157	127	138	100	111	87	123
Marital Status	male	unmarried	101	99	101	94	100	89	88	107
		married	70	75	76	77	100	98	96	127
	female	unmarried	113	124	103	118	100	99	80	119
		married	193	262	152	190	100	123	69	112
Profession	male	agriculture	45	48	69	68	100	97	134	170
		manufacturing	77	78	81	80	100	91	92	118
		technical	102	125	103	138	100	111	89	154
		services	90	100	101	99	100	100	98	125
Part Time	female	(yes)	152	151	135	180	100	90	79	135
		(no)	167	202	127	144	100	109	67	98
Recall	male	(yes)	58	62	55	65	100	97	85	112
	female	(yes)	67	80	71	69	100	108	94	102

Table 2: $\hat{T}_{0.6}$ relative to all observations (left) and relative to 1981 (right)

			1981	1985	1990	1995	1981	1985	1990	1995
All			100	100	100	100	100	89	89	133
Citizenship	male	German	64	66	75	80	100	92	104	166
		foreign	90	84	90	97	100	83	90	144
	female	German	195	195	137	128	100	89	62	87
		foreign	205	216	154	188	100	94	67	121
Education	male	unskilled	82	85	94	114	100	91	102	183
		skilled	60	56	69	71	100	83	102	158
		university	80	99	89	90	100	110	99	149
	female	unskilled	194	201	145	171	100	92	67	117
		skilled	221	216	138	122	100	87	55	73
		university	141	145	125	114	100	92	79	107
Marital Status	male	unmarried	100	96	92	92	100	85	82	121
		married	55	55	64	73	100	88	102	175
	female	unmarried	119	116	100	112	100	87	75	125
		married	242	277	172	157	100	102	63	86
Profession	male	agriculture	33	44	56	43	100	121	153	177
		manufacturing	63	57	74	75	100	80	105	158
		technical	104	117	92	102	100	101	79	131
		services	86	96	88	91	100	100	91	141
Part Time	female	(yes)	194	162	145	138	100	74	66	94
		(no)	198	204	134	133	100	92	60	89
Recall	male	(yes)	34	38	37	29	100	99	97	115
	female	(yes)	43	55	52	53	100	114	107	161

Table 3: $\hat{T}_{0.4}$ relative to all observations (left) and relative to 1981 (right)

6 Summary

This paper delivers detailed stylized facts about the distribution of unemployment duration for a variety of homogenous sub-samples of the IAB employment subsample. The estimation results indicate that the probability of remaining unemployed after a certain period varies significantly over many of the considered population segments. They also suggest that the variation due to the macroeconomic environment differs across the segments and that there are general developments over time due to behavioral changes in the society, e.g. the reduction of the nonemployment periods of married females. It is also observed that these variations are not always proportional over time and over the considered quintiles. This might be due to compositional effects of the compared samples but it might also be due to a violation of the proportionality assumption that is required for the correct specification of proportional hazard models. However, the latter question requires further inquiries.

Many unemployed leave unemployment during the first three months of the unemployment duration. The decline of the estimated survivor function then decreases sharply in many data segments. Some of the estimated survivor functions are almost constant after a duration of 12 months which corresponds to the period of long term unemployment. In particular the probability for an unemployed of becoming long term unemployed has increased for the males during the two decades under consideration whereby the contrary is observed for the females. On average it is not observed that a doubling in the unemployment rate had strong effect on the length of unemployment duration. A high probability of becoming long term unemployed ($20\% - 60\%$) is observed in most of the considered population segments. The only exception are male unemployed who got previously a recall to the former employer. The performed analysis is not able to provide an explanation for this well known phenomenon. Once unemployed it does not seem that higher educational degrees are a reliable protection against becoming long-term unemployed. It is therefore not apparent that the educational degree or the profession of an individual are striking characteristics why unemployed exit to employment and do not become long term unemployed. This has to be taken into account when designing further training measures and selecting possible participants.

However, higher educational degrees seem to experience a lower risk of unemployment given employment especially for males. Unskilled workers have the highest risk of unemployment, the longest unemployment durations and the largest fraction of long term unemployment. The business cycle mainly affects the unskilled workers, foreign nationals and females. Having a university degree seems to be a good protection against unemployment. However, if once unemployed,

males with completed apprenticeship leave unemployment fastest and therefore faster than the individuals with university degrees. The specific educational degree (apprenticeship completed or university degree) seems to be less important for the length of unemployment periods of the females. Females have in general longer unemployment periods and a higher probability of long term unemployment but this seems to be (mainly) due to the married females. However, there is strong evidence that married females have the most favorable development over the two decades under consideration. It seems that this is mainly due to a change in the willingness to work, since at the same time the labor market participation rate of the females is rising and the birth rate is declining.

References

[1] Bender, S., Haas, A., and Klose, C. (2000) The IAB Employment Subsample 1975–1995. Schmollers Jahrbuch, 120, 649–662.

[2] Fitzenberger, B. and Wilke, R. (2004). Unemployment Durations in West-Germany Before and After the Reform of the Unemployment Compensation System during the 1980ties. *Mimeo*. ZEW Mannheim.

[3] Hunt, J. (1995). The effect of the Unemployment Compensation on Unemployment Duration in Germany. *Journal of Labor Economics*. Vol.13.1, 88–120

[4] IAB Employment Subsample 1975-1997 - regional file -, Handbook, IAB Nürnberg (in German).

[5] Koenker, R. and Geling, O. (2001). Reapprising Medlfy Longevity: A Quantile Regression Survival Analysis. *Journal of the American Statistical Association*. Vol. 96, No. 454, 458–468

[6] Lauer, C. (2003). Education and Unemployment: A French-German Comparison. *ZEW Discussion Paper*. No. 03-34

[7] Plaßmann, G. (2002). Der Einfluss der Arbeitslosenversicherung auf die Arbeitslosigkeit in Deutschland. *Beiträge zur Arbeitsmarkt– und Berufsforschung*, 255, Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit (IAB) Nürnberg.

[8] Schräpler, J.P (2002). Respondant Behavior in Panel Studies - A Case Study for Income-Nonresponse by Means of the German Socio-Economic Panel. *DIW Discussion Paper*. No. 299

[9] Steiner, V. (2001). Unemployment Persitence in the West-German Labor Market: Negative Duration Dependence or Sorting? *Oxford Bulletin of Economics and Statistics*. No. 63.1, 91–113

[10] Van den Berg, G. (2001). Duration Models: Specification, Identification and Multiple Durations. in *Handbook of Econometrics*, edited by J.Heckman and E.Leamer, Vol. 5, North-Holland

Appendix: Tables and Figures

			1981	1985	1990	1995
All			7.978	7.410	6.459	9.349
Citizenship						
	male	German	4.057	4.153	3.249	4.482
		foreign	901	561	398	994
	female	German	2.505	2.398	2.459	3.040
		foreign	427	230	176	343
Education						
	male	unskilled	1.492	1.218	991	1.406
		skilled	2.729	2.836	2.046	3.092
		university	139	166	222	282
	female	unskilled	919	645	617	788
		skilled	1.602	1.613	1.553	1.984
		university	136	130	182	225
Marital Status						
	male	unmarried	1.927	2.034	2.128	3.138
		married	3.057	2.737	1.671	2.716
	female	unmarried	779	929	1.128	1.618
		married	2.129	1.710	1.532	1.877
Profession	male					
		agriculture	204	228	193	254
		manufacturing	2.941	2.881	2.070	3.270
		technical	145	150	128	279
		services	1.707	1.489	1.388	2.017
Part Time						
	female	(yes)	542	442	535	749
		(no)	2.396	2.197	2.125	2.746
Recall						
	male	(yes)	685	1.252	719	1.057
	female	(yes)	200	349	323	398

Table 4: Sample sizes of the (sub-)samples. Note that the sum in each data segment does not have to coincide with total amount of observations due to missings.

			1981	1985	1990	1995
All			65 (62 – 67)	61 (59 – 62)	52 (49 – 55)	64 (62 – 68)
Citizenship	male	German	53 (51 – 58)	51 (48 – 54)	46 (43 – 49)	60 (57 – 61)
		foreign	69 (61 – 79)	60 (55 – 72)	62 (50 – 74)	74 (63 – 82)
	female	German	92 (88 – 96)	90 (84 – 97)	61 (55 – 62)	72 (62 – 79)
		foreign	116 (104 – 141)	93 (75 – 120)	92 (64 – 117)	135 (107 – 165)
Education	male	unskilled	66 (60 – 81)	58 (51 – 64)	54 (46 – 61)	72 (65 – 81)
		skilled	51 (48 – 54)	48 (45 – 51)	44 (40 – 48)	59 (55 – 61)
		university	54 (45 – 61)	64 (56 – 75)	55 (46 – 62)	61 (53 – 69)
	female	unskilled	104 (95 – 114)	100 (88 – 117)	70 (61 – 84)	94 (89 – 120)
		skilled	92 (90 – 106)	90 (82 – 100)	60 (51 – 62)	71 (61 – 77)
		university	71 (61 – 90)	79 (62 – 91)	59 (48 – 75)	78 (61 – 99)
Marital Status	male	unmarried	65 (61 – 71)	61 (57 – 66)	54 (49 – 59)	63 (61 – 69)
		married	50 (47 – 53)	47 (44 – 51)	39 (35 – 44)	60 (56 – 62)
	female	unmarried	77 (69 – 89)	64 (61 – 77)	53 (46 – 61)	61 (59 – 70)
		married	105 (95 – 116)	120 (104 – 125)	65 (61 – 77)	95 (91 – 112)
Profession	male	agriculture	32 (28 – 44)	23 (15 – 28)	31 (25 – 42)	58 (42 – 74)
		manufacturing	56 (53 – 60)	51 (48 – 55)	45 (41 – 48)	61 (57 – 63)
		technical	61 (39 – 78)	77 (58 – 98)	59 (36 – 91)	90 (61 – 92)
		services	59 (53 – 61)	60 (54 – 65)	53 (47 – 61)	62 (60 – 69)
Part Time	female	(yes)	94 (78 – 112)	87 (62 – 105)	52 (44 – 62)	92 (73 – 117)
		(no)	92 (90 – 104)	92 (88 – 100)	62 (59 – 67)	75 (65 – 86)
Recall	male	(yes)	48 (43 – 53)	44 (39 – 47)	37 (34 – 43)	48 (43 – 52)
	female	(yes)	52 (40 – 64)	45 (39 – 59)	36 (32 – 51)	46 (37 – 55)

Table 5: Estimation results of $\hat{T}_{0.8}$, $\alpha = 0.05$, $\underline{T}_\theta$ and $\overline{T}_\theta$ in brackets

			1981	1985	1990	1995
All			135 (130 – 140)	122 (120 – 125)	119 (112 – 122)	154 (152 – 165)
Citizenship	male	German	100 (95 – 105)	99 (95 – 102)	100 (94 – 105)	123 (121 – 130)
		foreign	145 (131 – 162)	128 (115 – 144)	121 (101 – 145)	152 (133 – 170)
	female	German	215 (195 – 243)	242 (215 – 258)	151 (136 – 166)	226 (202 – 243)
		foreign	253 (221 – 331)	233 (171 – 365)	174 (132 – 228)	324 (246 – 409)
Education	male	unskilled	122 (112 – 133)	111 (103 – 119)	120 (110 – 131)	153 (141 – 175)
		skilled	97 (92 – 103)	93 (90 – 97)	92 (90 – 99)	118 (110 – 123)
		university	119 (101 – 136)	136 (123 – 153)	120 (102 – 132)	147 (130 – 165)
	female	unskilled	243 (214 – 285)	243 (188 – 287)	161 (135 – 191)	307 (252 – 365)
		skilled	232 (212 – 273)	245 (228 – 273)	147 (126 – 166)	214 (191 – 240)
		university	173 (151 – 199)	192 (151 – 243)	151 (122 – 182)	213 (179 – 245)
Marital Status	male	unmarried	136 (126 – 148)	121 (115 – 129)	120 (110 – 126)	145 (136 – 153)
		married	94 (90 – 98)	92 (89 – 95)	90 (81 – 93)	119 (110 – 123)
	female	unmarried	153 (141 – 182)	151 (134 – 173)	122 (112 – 139)	182 (155 – 196)
		married	261 (240 – 300)	320 (285 – 365)	181 (161 – 200)	293 (264 – 324)
Profession	male	agriculture	61 (53 – 76)	59 (45 – 75)	82 (68 – 96)	104 (91 – 121)
		manufacturing	104 (100 – 109)	95 (92 – 99)	96 (91 – 102)	123 (117 – 129)
		technical	138 (106 – 212)	153 (120 – 184)	123 (95 – 181)	212 (147 – 277)
		services	122 (111 – 131)	122 (116 – 132)	120 (108 – 126)	152 (137 – 165)
Part Time	female	(yes)	205 (181 – 257)	184 (160 – 234)	161 (122 – 212)	277 (236 – 335)
		(no)	226 (212 – 250)	246 (231 – 273)	151 (136 – 166)	222 (199 – 243)
Recall	male	(yes)	78 (74 – 83)	76 (72 – 82)	66 (62 – 72)	87 (83 – 91)
	female	(yes)	90 (75 – 107)	97 (84 – 111)	84 (73 – 93)	92 (82 – 110)

Table 6: Estimation results of $\hat{T}_{0.6}$, $\alpha = 0.05$, $\underline{T}_\theta$ and $\overline{T}_\theta$ in brackets

			1981	1985	1990	1995
All			307 (291 – 327)	273 (259 – 283)	273 (255 – 289)	407 (392 – 427)
Citizenship	male	German	196 (184 – 211)	181 (169 – 191)	204 (189 – 219)	326 (305 – 349)
		foreign	275 (249 – 317)	229 (203 – 273)	247 (203 – 318)	396 (352 – 457)
	female	German	600 (546 – 669)	532 (487 – 607)	374 (365 – 405)	519 (486 – 550)
		foreign	629 (500 – 863)	591 (449 – 717)	421 (334 – 651)	764 (655 – 927)
Education	male	unskilled	253 (225 – 274)	231 (210 – 256)	257 (220 – 289)	462 (407 – 525)
		skilled	184 (173 – 199)	153 (145 – 163)	188 (178 – 208)	290 (274 – 311)
		university	245 (207 – 288)	270 (237 – 302)	242 (203 – 283)	366 (317 – 410)
	female	unskilled	595 (523 – 729)	548 (468 – 652)	396 (344 – 462)	696 (621 – 799)
		skilled	679 (589 – 813)	591 (516 – 669)	376 (359 – 420)	495 (457 – 547)
		university	432 (313 – 539)	396 (352 – 528)	342 (287 – 424)	464 (387 – 605)
Marital Status	male	unmarried	307 (272 – 335)	261 (235 – 273)	252 (227 – 275)	373 (352 – 406)
		married	170 (160 – 181)	149 (139 – 153)	174 (153 – 186)	297 (274 – 331)
	female	unmarried	365 (303 – 426)	318 (281 – 365)	273 (243 – 304)	456 (397 – 490)
		married	742 (654 – 863)	756 (669 – 831)	470 (422 – 529)	639 (578 – 714)
Profession	male	agriculture	100 (88 – 129)	121 (106 – 176)	153 (112 – 191)	177 (146 – 313)
		manufacturing	194 (183 – 210)	155 (147 – 166)	203 (184 – 219)	306 (282 – 340)
		technical	318 (232 – 396)	320 (212 – 365)	252 (185 – 382)	415 (366 – 493)
		services	263 (240 – 281)	262 (239 – 284)	239 (208 – 266)	371 (347 – 414)
Part Time	female	(yes)	596 (467 – 791)	443 (364 – 546)	395 (348 – 456)	561 (479 – 639)
		(no)	607 (548 – 679)	558 (516 – 638)	367 (352 – 411)	541 (493 – 580)
Recall	male	(yes)	104 (100 – 110)	103 (100 – 108)	101 (92 – 108)	120 (114 – 128)
	female	(yes)	133 (121 – 151)	151 (132 – 179)	141 (120 – 160)	214 (176 – 275)

Table 7: Estimation results of $\hat{T}_{0.4}$, $\alpha = 0.05$, $\underline{T}_\theta$ and $\overline{T}_\theta$ in brackets

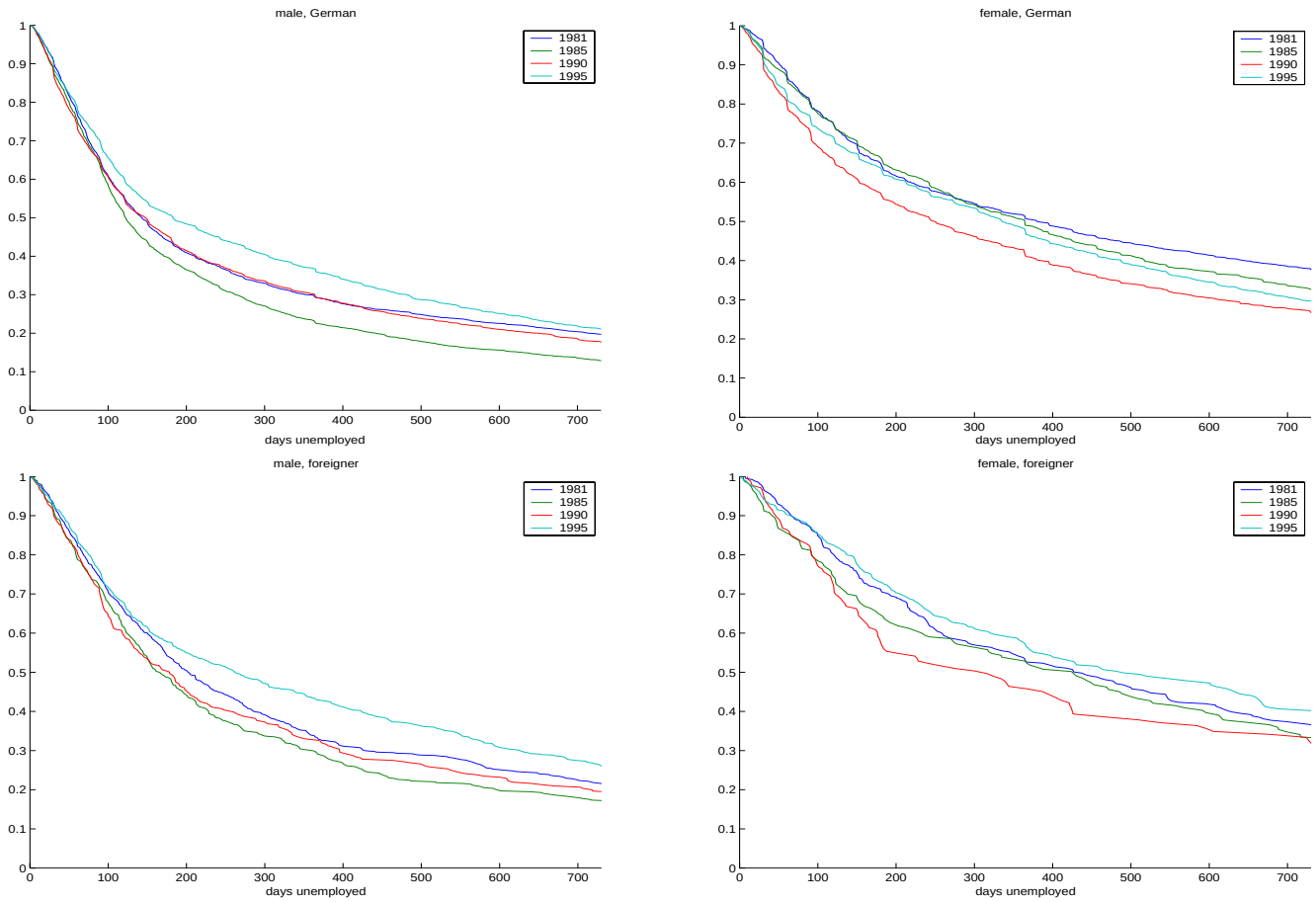

Figure 5: Kaplan-Meier-survival function estimates

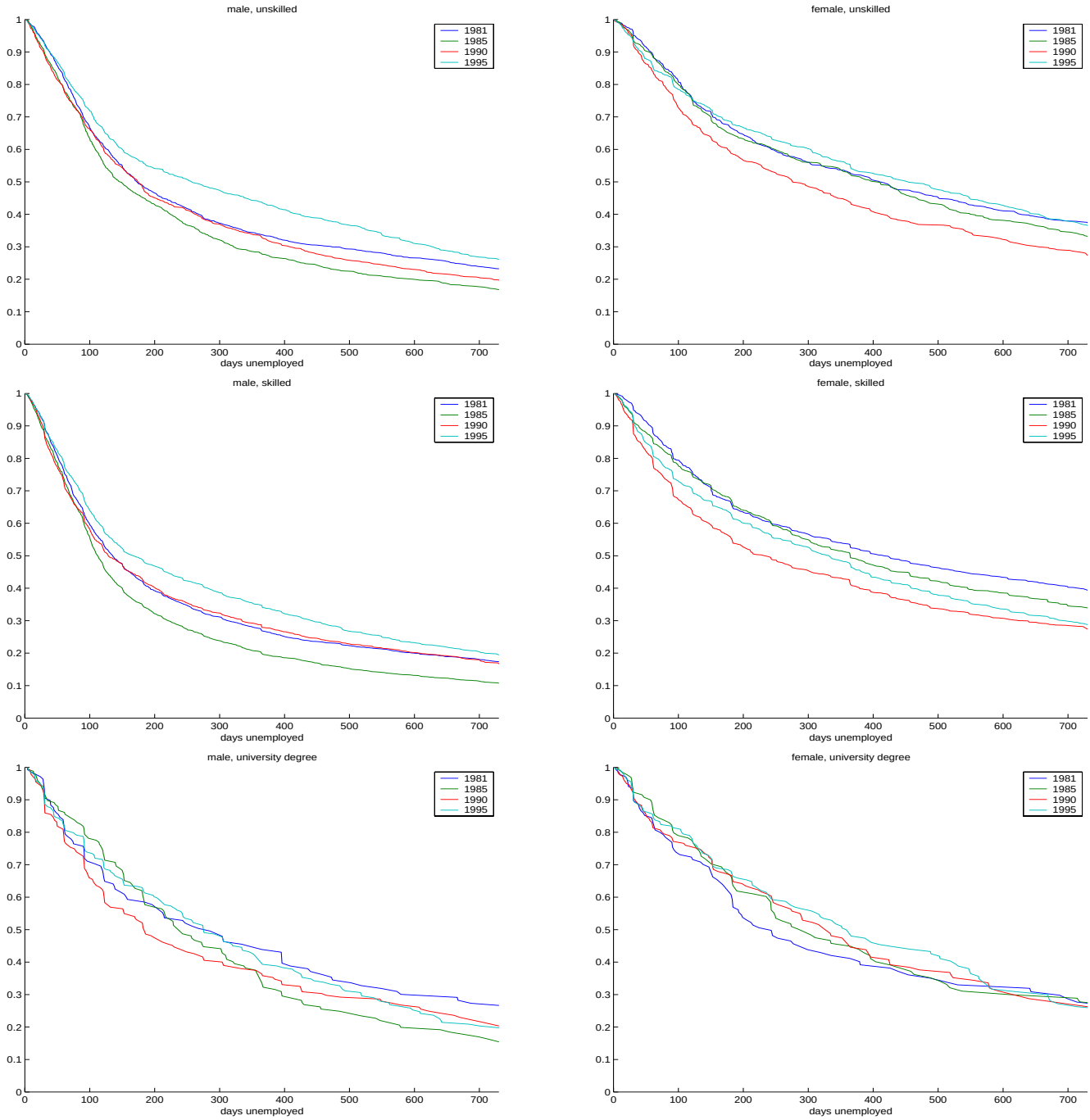

Figure 6: Kaplan-Meier-survival function estimates

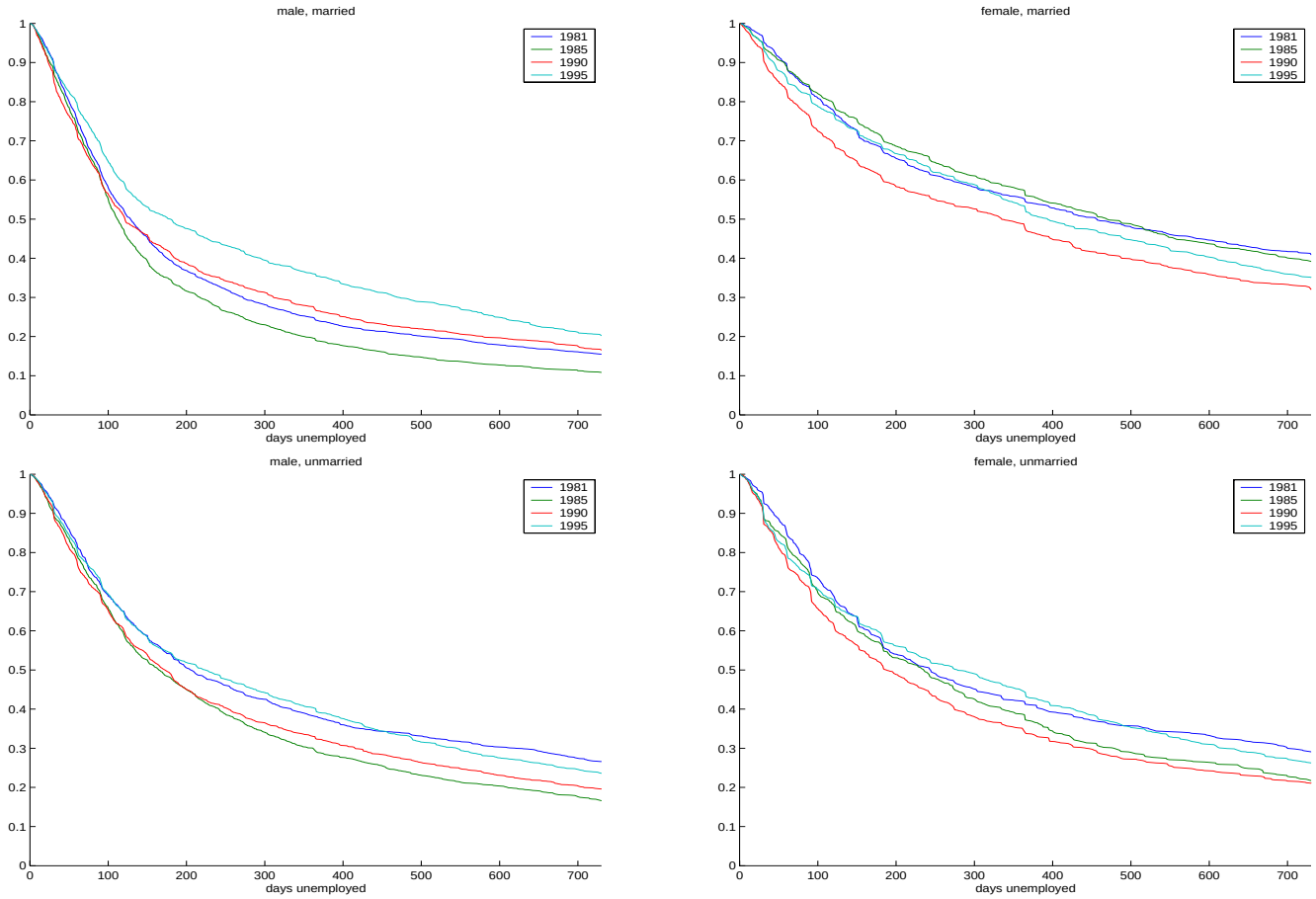

Figure 7: Kaplan-Meier-survival function estimates

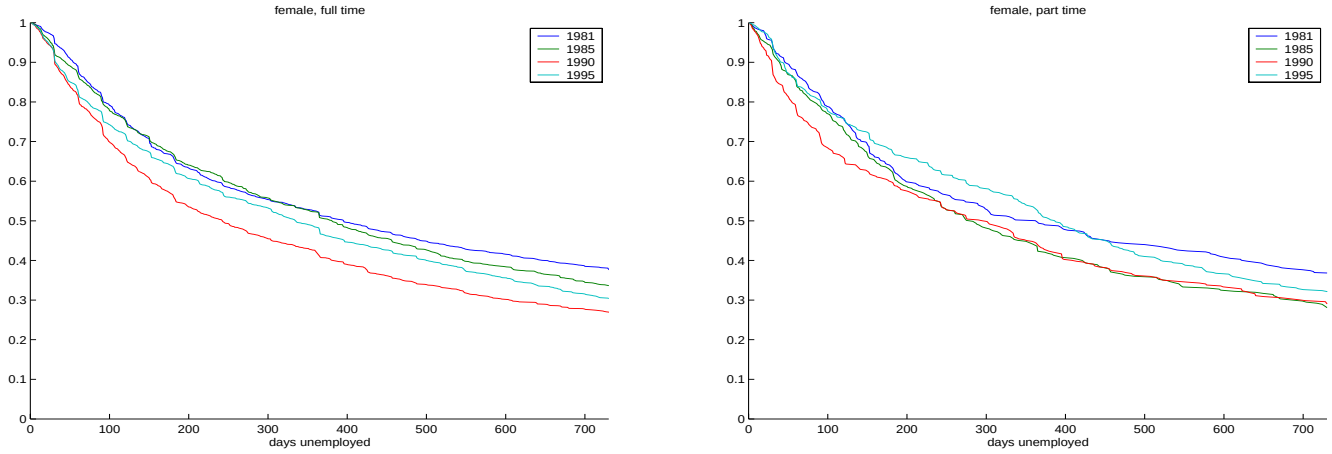

Figure 8: Kaplan-Meier-survival function estimates

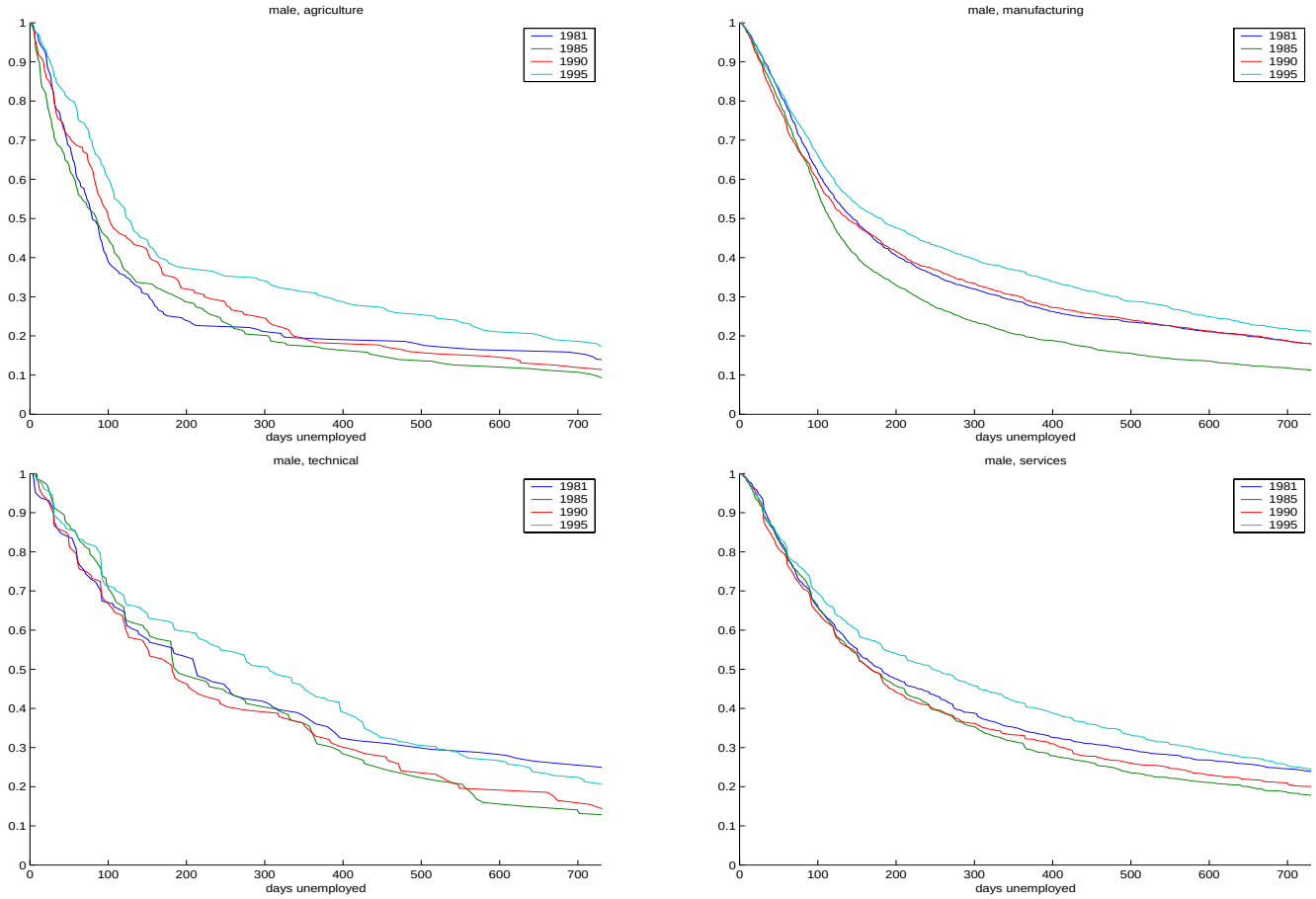

Figure 9: Kaplan-Meier-survival function estimates

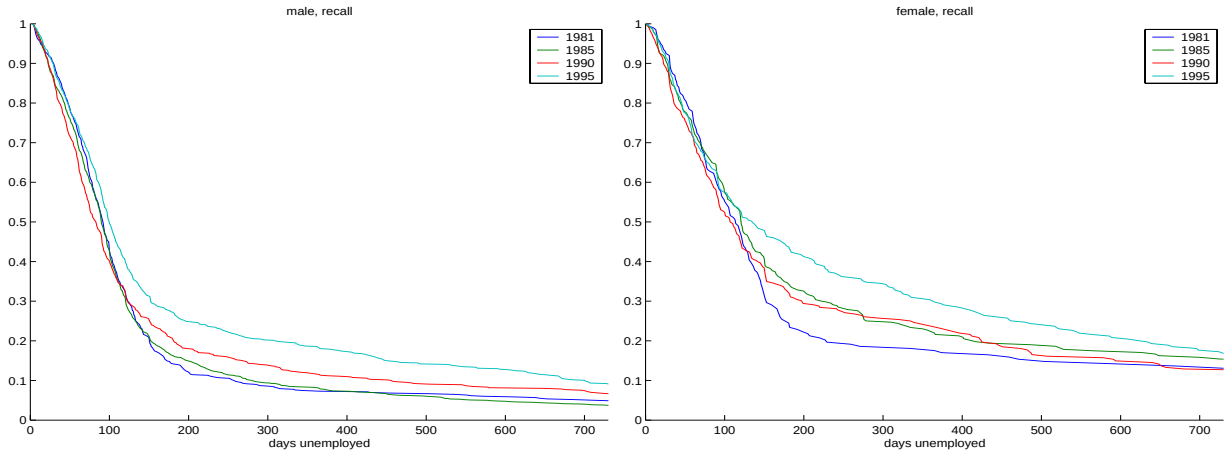

Figure 10: Kaplan-Meier-survival function estimates

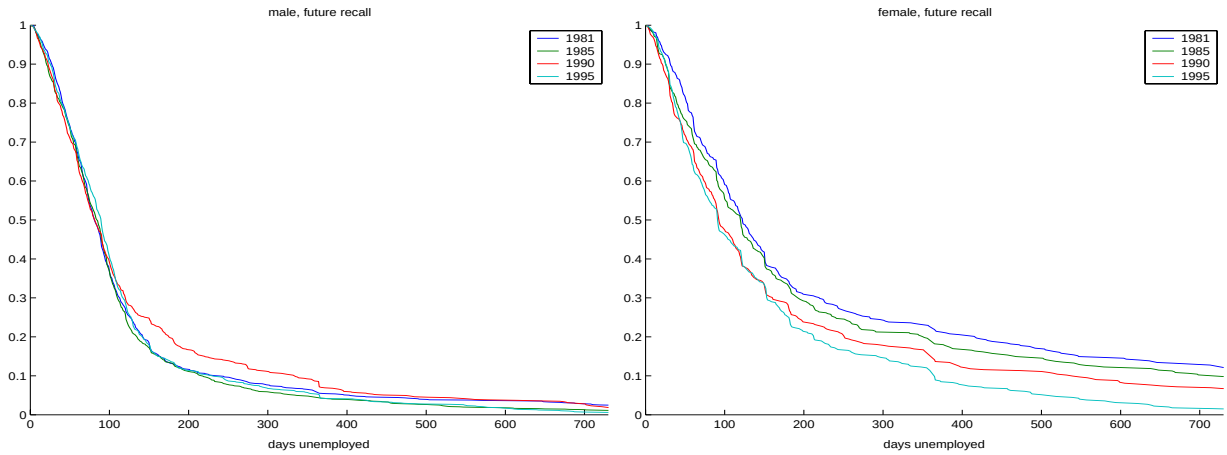

Figure 11: Kaplan-Meier-survival function estimates