BGPE Discussion Paper

No. 90

Determinants of Lifetime Unemployment. A Micro Data Analysis with Censored Quantile Regressions

Achim Schmillen
Joachim Möller

March 2010

ISSN 1863-5733

Editor: Prof. Regina T. Riphahn, Ph.D.
Friedrich-Alexander-University Erlangen-Nuremberg

B G P E Discussion Paper Series

DETERMINANTS OF LIFETIME UNEMPLOYMENT
—
A MICRO DATA ANALYSIS WITH CENSORED QUANTILE REGRESSIONS

Achim Schmillen[1] and Joachim Möller[2,3]

The empirical literature on unemployment almost exclusively focuses on the duration of distinct unemployment spells. In contrast, we use a large German administrative micro data set for the time span 1975–2004 to investigate individual *lifetime unemployment* (defined as the total length of all unemployment spells over a 25-year period). This new perspective enables us to answer questions regarding the long-term distribution and determinants of unemployment for West German birth cohorts 1950–1954. We find that lifetime unemployment is highly unevenly distributed and employ censored quantile regressions to show that, for men, pursuing a disadvantageous occupation early in the professional career leads to a significantly higher amount of lifetime unemployment.

Keywords: Lifetime unemployment, Censored quantile regressions, Occupation-specific human capital
JEL-Classification: J64, J24.

[1]Institute for Employment Research (IAB), Osteuropa-Institut Regensburg and University of Regensburg.
[2]Institute for Employment Research (IAB), IZA and University of Regensburg.
[3]Correspondence to: Joachim Möller, Institute for Employment Research (IAB), Regensburger Strasse 104, D-90478 Nürnberg, Germany; e-mail: joachim.moeller@iab.de.

1. INTRODUCTION

Prominent empirical literature examines occurrence, distribution or determinants of unemployment.[1] This literature almost completely focuses on distinct periods of unemployment over a relatively short overall time span. In contrast we use West German administrative micro data to assess what we call individual *lifetime unemployment*: the total length of all spells of unemployment over a 25-year period (from age 25 to 50). This new perspective enables us to answer questions regarding the long-term distribution of unemployment and the flexibility of the German labor market and its institutions.

So far only a few studies such as Kurtz and Scherl (2001) or Brooks (2005) have looked at individual unemployment from a similarly long-term perspective. These studies have tended to use survey data and were in any case confined to descriptive evidence.[2] Our administrative micro data set offers not only a large sample size but also very reliable information on unemployment and other variables of interest. This allows us not only to present descriptive statistics but also to analyze lifetime unemployment with the help of multivariate statistics. To the best of our knowledge, ours is the first study to use a rich and reliable administrative micro data set and statistical inference to analyze unemployment over a 25-year period.

As an illustration of how relying on the concept of lifetime unemployment can lead to new and interesting insights, our study presents descriptive statistics on how very unevenly lifetime unemployment is distributed for selected West German cohorts — more than 60% of individuals in our sample are not unemployed for a single day over the better part of their professional career while almost half of the total amount of unemployment falls upon 5% of the population.

This observation makes it highly relevant to find out which factors determine the individual amount of lifetime unemployment – not only for the "average" individual but in particular for those in the upper tail of the distribution of lifetime unemployment. Easily observable individual characteristics like low education could certainly be a reason for having an elevated amount of lifetime unemployment. Harder-to-observe characteristics like the state of health, the motivation of the individual etc. can be expected to play a role as well. What is more, wrong choices or simply bad luck at a young age may also influence the amount of lifetime unemployment. Recent studies have indeed shown that a job loss or a similarly incisive labor market event early in the professional career can have long-lasting effects [cp. for instance Kalwij (2004) and von Wachter and Bender (2006)].

Starting with the influential paper by Ljungqvist and Sargent (1998) a separate branch of literature has stressed the connection between human capital and

[1]Examples include Arulampalam, Booth and Taylor (2000) for Great Britain; Koenker and Bilias (2001) for the United States; Galiani and Hopenhayn (2003) for Argentina; and Lüdemann, Wilke and Zhang (2006) for Germany.

[2]An exception is the study by Kalwij (2004) which uses registry data from Britain's National Unemployment Benefits System to analyze unemployment among young British men.

unemployment. In this literature losing a job is seen as a sudden depreciation of human capital which (possibly together with other factors) might lead to long unemployment spells.

We link these two bodies of literature and investigate the causal effect of a specific investment in human capital early in the professional career on the amount of lifetime unemployment for selected quantiles of the distribution of lifetime unemployment. Following Kambourov and Manovskii (2009) we specifically focus on occupation-specific human capital and find that — at least for men — working in what *ex post* turned out to be a disadvantageous occupation early in the professional career (at age 25) is indeed connected to a significantly higher amount of lifetime unemployment.

Of course one might ask whether the *ex post* advantageousness of an occupation can really be exogenous in a regression of lifetime unemployment. Particularly one might worry that individuals with favourable personal characteristics *ex ante* sort into what they perceive as occupations with favourable characteristics. If the (perceived) *ex ante* advantageousness of an occupation is correlated with its (eventual) *ex post* advantageousness and also with lifetime unemployment this would mean there is an omitted variable problem if lifetime unemployment is regressed only on *ex post* advantageousness.

We would claim that it is very difficult, or even impossible, for individuals to make correct long-term occupational forecasts and therefore do not expect a strong omitted variable bias when regressing lifetime unemployment on *ex post* advantageousness only. Still we extend our baseline approach by controlling for the *ex ante* advantageousness of an occupation. Specifically we control for occupation-specific unemployment and wage rates at the beginning of the observation period. Additionally we use an elaborated long-run occupation-specific labor market forecast published by the German Federal Employment Agency, Blüm and Frenzel (1975), as a further indicator of the state of knowledge at the beginning of our observation period.

Even with all these control variables, our measures of the *ex post* advantageousness of the different occupations turn out to be significant. We argue that with the inclusion of the controls we solve the potential omitted-variable problem. Therefore our measures of the *ex post* advantageousness can be considered as exogenous and we identify a causal effect of the *ex post* advantageousness of the occupation chosen at young age on lifetime unemployment.[3]

The remainder of this paper is structured as follows: Section 2 deals with the theoretical basis of our study and Section 3 introduces our data set. Section 4 presents descriptive evidence while Sections 5 and 6 contain methods and results of our multivariate analyses. Section 7 provides conclusions.

[3] Henceforth we will mostly omit the "*ex post*" prefix when referring to the *ex post* advantageousness of an occupation but always keep the "*ex ante*" prefix when relating to the *ex ante* advantageousness.

2. THEORETICAL CONSIDERATIONS

A standard neoclassical labor market model can easily explain how an investment in a disadvantageous kind of human capital early in the professional career can henceforth reduce the individual's productivity and therefore depress his/her wages. However, in such a framework one would not expect this individual to exhibit an elevated amount of lifetime unemployment.

Radically different conclusions can be reached by models that allow for certain types of labor market imperfections. Prominent examples are general equilibrium search models that connect human capital and unemployment [pioneered by Ljungqvist and Sargent (1998)]. In the following we will outline a version of such a model where an investment in a disadvantageousness kind of human capital early in the professional career does indeed induce an elevated amount of lifetime unemployment.

The relevant models usually assume that at each point in time individuals are equipped with a level of human capital h. For each period $\mu_s(h, h')$ denotes the transition probability from human capital level h to h' where the subscript $s = \{u, e, l\}$ captures whether the individual is unemployed, employed or laid-off in the respective period. That is, an individual with human capital level h who is laid off faces a probability of $\mu_l(h, h')$ that his/her human capital level at the beginning of the next period is h'.

In the models a higher human capital depreciation at separation leads to more unemployment, especially in the presence of a welfare state. This happens because of two mechanisms. First, it is assumed that the welfare state pays unemployment benefits in proportion to past earnings. Individuals with highly depreciated human capital therefore have relatively high reservation wages and have difficulties in finding a new job that they prefer to their unemployment compensation. Second, it is assumed that job search is associated with disutility. So individuals with depreciated human capital "(...) reduce their search intensities to balance the small prospective gains from search against the utility costs associated with search" (Ljungqvist and Sargent, 1998, p. 535).

We now presume that the human capital depreciation rate of newly laid-off workers depends on specific individual or job characteristics. Specifically we assume that it depends on whether the human capital acquired in the previous job is still in demand at the time of a separation. If this is the case, lay-offs caused by technical change or shifting trade patterns should lead to an especially strong depreciation of specific human capital acquired in the previous job.

Now it is crucial to ask which kind of specific human capital is lost at the time of a separation. While the majority of the relevant literature refers to job- or industry-specific human capital a recent study by Kambourov and Manovskii (2009) provides convincing evidence that it might be more appropriate to consider occupation-specific human capital instead.

We follow this approach and assume that the human capital depreciation rate of newly laid-off workers indeed depends on the type of human capital they attained in their previous job and that occupation-specific human capital is more important than job- or industry-specific human capital. Thus we predict that individuals who early in their professional career acquire occupation-specific human capital in an occupation that later becomes obsolete should experience an especially pronounced loss of human capital once they are laid off. On average this group is therefore expected to suffer from a comparatively high amount of unemployment.

This modification of a standard search model connecting human capital and unemployment will henceforth serve as a theoretical basis for our hypothesis that individuals who start their employment careers in what we call a "disadvantageous" occupation, that is, one with a disadvantageous kind of occupation-specific human capital, *ceteris paribus* face an elevated amount of lifetime unemployment. In the next section we will present our data set and explain — among other things — how we measure the advantageousness of an occupation.

3. DATA

The data set used in our study is the so-called IAB Employment Sample (IABS) of the Institute for Employment Research, Nuremberg (IAB). Its source is the German Employment Register which covers about 80% of Germany's total workforce. The IABS is a panel based on a 2% random sample of all German employees registered by the social security system and contains detailed longitudinal information exact to the day.[4]

The IABS contains all employment spells associated with the payment of social security contributions. Only employees not covered by social security, like civil servants or family workers and self-employed persons, are not included in the data. Spells during which workers receive unemployment benefits are added to the sample. Because records from the Employment Register are used to compute both social security contributions and unemployment benefits, the IABS data set is highly reliable.

The key variable for our analysis is what we call the individual amount of unemployment or — for the sake of brevity — lifetime unemployment. It is defined as the total length (in days) of all unemployment spells of an individual from age 25 to age 50. We restrict our sample to this range because of data limitations and because this procedure should limit distorting effects of (un-)employment patterns specific to particularly young or particularly old individuals (e.g. connected to tertiary education or early retirement).

About 90% of those registered as unemployed are eligible for unemployment relief or related benefits. Our data do not contain information on unemployed individuals who do not receive any unemployment benefits at all. The same applies to individuals who for some reason are not registered as unemployed but are still willing to take up a job. Thus we restrict our definition of unemployment

[4] A detailed description of the IABS can be found in Bender, Haas and Klose (2000).

to spells of unemployment associated with the receipt of benefits.[5]

There is one further consequence of using the receipt of unemployment benefits to define unemployment episodes: regulations concerning unemployment benefits have somewhat varied during the last decades. This makes it difficult to compare the length of unemployment periods from different points in time. This is why we limit our analysis to a number of selected cohorts. Specifically we focus on those individuals born between 1950 and 1954. Thus our study draws on data from 1975 (when the individuals born in 1950 turned 25) to 2004 (when the cohort of 1954 turned 50).[6]

Most explanatory variables for the multivariate analysis of Section 6 are constructed with the help of the IABS data set, too. Some are individual characteristics (like education). Others are taken from the job held by the individual on his/her 25th birthday or, if the individual was not employed on this date, from the first job taken up after the 25th birthday. We choose the 25th birthday on the one hand because most people aged 25 have finished education and entered the labor force. On the other hand they are still relatively early in their professional career.

A main aim of our study is to assess whether pursuing a "disadvantageous" occupation early in the career affects the amount of lifetime unemployment. After some aggregation and data cleansing (discarding occupations that are covered by our data only for certain years etc.) we are able to distinguish 56 two-digit occupations for which we have consistent data.

In order to decide the relative "advantageousness" or "disadvantageousness" of an occupation we first of all sum the total number of employment days for each year between 1975 and 2004 for each of the 56 different occupations. Next we use a Hodrick-Prescott-filter (with a smoothing parameter of 6.25) to determine trend and fluctuations of the employment series for the different occupations from 1975 to 2004. This gives us two measures for the relative advantageousness of all occupations contained in our data: first the trend employment growth rate between 1975 and 2004, second the standard deviation of the employment fluctuations over this period. An advantageous occupation is characterized by a relatively large (positive) employment growth rate together with a relatively small standard deviation of the employment fluctuations.[7]

A number of other variables are included in our multivariate analysis in Section 6 as controls and also because assessing their effect on the amount of lifetime

[5]This might slightly limit the informative value of our analysis. It might specifically distort the unemployment pattern of women, a comparatively large number of whom do not qualify for unemployment benefits. This is one reason why we show descriptive statistics separately for men and women. We are also very careful to compare the respective results.

[6]Details on further data cleansing can be found in Appendix A.

[7]Judging by the employment growth rate *natural scientists and humanists not elsewhere covered* hold the most advantageousness occupation while *spinners* work in the most disadvantageous. The employment fluctuations of *bankers and insurance specialists* have the smallest standard deviations, those of the *legal* professions the largest. For a more detailed overview of the most advantageous and disadvantageous occupations, see Appendix B.

unemployment might be interesting in itself:

- Education level. It is well-known that education is closely related to the occurrence of unemployment. Besides, education and occupation are strongly connected, so controlling for education is important. We do this by including five dummy variables that measure whether an individual holds a degree from vocational training but no high school diploma, a high school diploma but no degree from vocational training, a high school diploma and a degree from vocational training, a degree from a technical college or a university degree. The control group consists of those individuals that hold neither a high school diploma nor a degree from vocational training.
  We would expect that individuals with higher education and especially those with a tertiary degree (from a technical college or a university) are *ceteris paribus* faced with a lower amount of lifetime unemployment.
- Weekly wages earned at the age of 25. Because of the possible existence of efficiency wages and the like higher wages could be *ceteris paribus* associated with a lower amount of lifetime unemployment. At the same time, elevated wages in the beginning of the professional career could lead to higher reservation wages and ultimately to higher unemployment.
- Sector of the firm for which the individual worked on his/her 25th birthday. Many occupations are for the most part found in a specific sector of the economy (e.g. bricklayers will almost exclusively work in the construction sector). Even though Kambourov and Manovskii (2009) convincingly argue that occupation-specific human capital is much more important than the sector-specific kind we want to make sure that we do indeed measure the effect of the relative advantageousness of occupations and not that of sectors.
  We use dummy variables for six aggregated sectors: agriculture, energy and mining, manufacturing, services, construction as well as the public sector and other activities. *A priori* it is hard to derive hypotheses on the different sectors' roles in determining the amount of lifetime unemployment.
- Region where the job pursued at age 25 is based, measured by dummy variables for the 10 West German federal states (with the state of Schleswig-Holstein as the reference region and omitting Berlin). *A priori* we would expect that working in a state with a rather favorable economic development at age 25 should *ceteris paribus* be associated with a comparatively small amount of lifetime unemployment.
- The size of the establishment for which the individual worked when turning 25, measured by adding up the number of its employees. Since generally speaking in Germany the influence of labor unions is strongest in big companies this variable might be a signal for whether employees have some bargaining power that might lead to less lay-offs and a lower risk of unemployment. So we expect that individuals working for a larger firm at the beginning of their professional career *ceteris paribus* face a smaller amount of lifetime unemployment.

4. DESCRIPTIVE EVIDENCE

Before we turn to our multivariate analysis we first present some descriptive evidence on the interpersonal distribution of lifetime employment and unemployment with a particular emphasis on those individuals with a very high amount of lifetime unemployment. For this section our samples consist of 35,281 men and 29,953 women with the characteristics described in the last section.

We start with some summary statistics. For this purpose we distinguish three labor market states: *employed*, *unemployed* and *neither employed nor unemployed*. The first two states are defined as described in the last section. The remainder of the professional career is labeled *neither employed nor unemployed* even though strictly speaking it might encompass episodes of marginal employment and unemployment without receipt of unemployment benefits as well as self-employment and work as a civil servant (cp. Section 3).[8]

The top panel of Table I summarizes information on the three labor market states for all men in our sample. The top panel of Table II does the same for all the women. On average employment careers of men encompass 1.6 unemployment spells with an average length of 225 days. Women are on average 1.0 times unemployed with average unemployment episodes lasting for 227 days. For both genders the state *neither employed nor unemployed* plays on average a much greater role than unemployment. Men are on average counted as neither employed nor unemployed for almost a third of their prime age (2,821 of 9,497 days[9]). The average woman is even counted as neither employed nor unemployed during almost half of her prime age (on average 4,393 days out of 9,497 are spent neither in employment nor in unemployment and only 4,871 in employment).

The perhaps surprising importance of periods for which neither employment nor unemployment is reported can probably be explained not only by actual periods of inactivity but also by our relatively restrictive definitions of employment and unemployment. Not counting periods when individuals were neither employed nor unemployed we calculate average unemployment rates of 5.3% for men and 4.6% for women. While these figures cannot be directly compared with unemployment rates defined in a standard way they lie in a plausible range.

We now drop the categories *employed* and *neither employed nor unemployed* and focus solely on periods of unemployment. Here, we are especially interested in the long-term interpersonal distribution of unemployment. Our first step is to look at the fraction of the sample that was ever unemployed between age 25 and age 50. We find that "only" about 36% of men and 37% of women were unemployed for at least a single day during their prime age. Conversely more than 60% of the individuals in our sample were not personally affected by

[8]If, for an individual, information on employment or unemployment is only available some time after his/her 25th birthday or not right until his/her 50th birthday these gaps are also included in our notion of *neither employed nor unemployed*. Excluding them altogether would not qualitatively alter applicable results.

[9]While for all individuals we look at the time span from their 25th to their 50th birthday leap years have the effect that the total time span differs by up to two days for different cohorts.

TABLE I

Summary statistics on the three labor market states in the time period 1975 to 2004 for the 1950 to 1954 birth cohorts (men)

all men				
	employed	unemployed	neither employed nor unemployed	total
average total duration (in days)	6,318	357	2,821	9,497
average number of spells	4.2	1.6	4.5	10.2
average spell duration (in days)	1,516	225	630	928
in percent	66.5	3.8	29.7	100
in percent (not considering *neither employed nor unemployed*)	94.6	5.3		100
5% of men with the highest amount of lifetime unemployment				
	employed	unemployed	neither employed nor unemployed	total
average total duration (in days)	3,292	3,626	2,579	9,497
average number of spells	9.9	10.6	15.9	36.4
average spell duration (in days)	332	343	162	261
in percent	34.7	38.1	27.2	100
in percent (not considering *neither employed nor unemployed*)	47.6	52.4		100
all men excluding the 5% with the highest amount of lifetime unemployment				
	employed	unemployed	neither employed nor unemployed	total
average total duration (in days)	6,478	185	2,834	9,497
average number of spells	3.9	1.1	3.9	8.9
average spell duration (in days)	1,677	167	731	1,072
in percent	68.2	2.0	29.8	100
in percent (not considering *neither employed nor unemployed*)	97.2	2.8		100

TABLE II

Summary statistics on the three labor market states in the time period 1975 to 2004 for the 1950 to 1954 birth cohorts (women)

all women				
	employed	unemployed	neither employed nor unemployed	total
average total duration (in days)	4,871	238	4,393	9,497
average number of spells	3.6	1.0	3.9	8.5
average spell duration (in days)	1,342	227	1,131	1,112
in percent	51.3	2.5	46.3	100
in percent (not considering *neither employed nor unemployed*)	95.4	4.6		100

5% of women with the highest amount of lifetime unemployment				
	employed	unemployed	neither employed nor unemployed	total
average total duration (in days)	4,164	2,107	3,225	9,497
average number of spells	7.5	7.0	10.4	24.8
average spell duration (in days)	558	303	309	382
in percent	43.8	22.2	34.0	100
in percent (not considering *neither employed nor unemployed*)	66.4	33.6		100

all women excluding the 5% with the highest amount of lifetime unemployment				
	employed	unemployed	neither employed nor unemployed	total
average total duration (in days)	4,908	134	4,455	9,496
average number of spells	3.4	0.7	3.5	7.7
average spell duration (in days)	1,431	188	1,258	1,236
in percent	51.7	1.4	46.9	100
in percent (not considering *neither employed nor unemployed*)	97.3	2.7		100

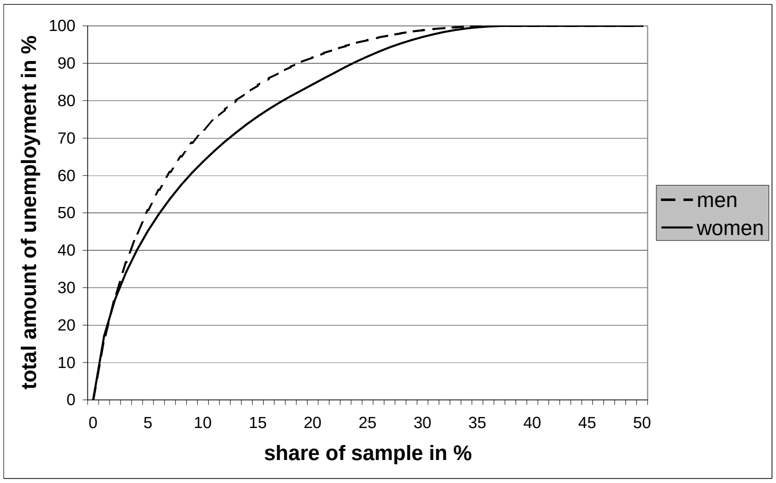

FIGURE 1.— Inverted Lorenz curves for the interpersonal distributions of the total amount of unemployment for men and women

unemployment between age 25 and 50 at all. This observation is a first indicator for a very uneven distribution of lifetime unemployment.

How concentrated lifetime unemployment is, becomes even more obvious when looking at Figure 1. The figure draws inverted Lorenz curves for the interpersonal distribution of the *total amount of unemployment* separately for men and women. The total amount of unemployment is defined as the sum of the amounts of lifetime unemployment for all individuals in our sample. Figure 1 shows two very uneven distributions. This result is again confirmed by the corresponding Gini-coefficients. Values of 0.851 for men and 0.816 for women signify a high concentration of the total amount of unemployment on some individuals.

For illustrative purpose one can also compare the fact that more than 60% of the individuals in our sample were not unemployed for a single day between age 25 and 50 with the observation that for men about half of the total amount of unemployment falls upon 5% of the sample. For women 6% of the sample are affected by roughly 50% of the total amount of unemployment.[10]

Table III again stresses the high concentration of lifetime unemployment by listing the amount of lifetime unemployment for selected percentiles of the distribution of the total amount of unemployment. Needless to say, for the 60th

[10]One might infer from Figure 1 that the total amount of unemployment is more unevenly distributed for men than for women. However, as was discussed in Section 3, such a comparison is problematic. The total amount of unemployment for women could in particular be less evenly distributed than shown by Figure 1 if a comparatively large number of women faced with high unemployment are not in fact registered as unemployed.

TABLE III

Lifetime unemployment in days for selected percentiles of the total amount of unemployment

percentile	men	women
60th	0	0
70th	123	182
80th	410	364
90th	1,101	668
95th	2,091	1,106
99th	4,671	2,727

percentile the amount of lifetime unemployment is zero. In contrast, the 5% of men with the highest amount of lifetime unemployment were unemployed for at least 2,091 days during their prime age. Men whose lifetime unemployment was in the top percentile were even unemployed for 4,671 days or more — that is more than 12 years! The corresponding column for women shows a pattern that is qualitatively similar but not as extreme.

The very uneven distribution of the total amount of unemployment leads to the following question: What variables determine the individual amount of lifetime unemployment? More specifically, it is especially relevant to know which attributes influence the amount of lifetime unemployment for those (say 5% or 10% of) individuals in the right tail of the distribution of lifetime unemployment. A method particularly suited to address this issue, (censored) quantile regression, is presented in the next section. Subsequently, results of its application to the interpersonal distribution of lifetime unemployment are discussed.

The middle panels of Tables I and II reproduce the top panels of these tables but focus exclusively on the 5% of men and women with the highest amount of lifetime unemployment. As mentioned above, about half of the total amount of unemployment falls upon members of this comparatively small group. For comparison, the bottom panels of Tables I and II report summary statistics on all individuals in our sample but the 5% with the highest amount of lifetime unemployment.

It is interesting to note that the elevated amount of lifetime unemployment of the 5% of men and women with the highest amount of unemployment is on average to a large part due to a higher number of unemployment spells: the average number of unemployment spells for both men and women in this group exceeds the average number of spells for the rest of our sample by almost a factor of 10. At the same time the average duration of an unemployment spell — the main driving factor in the theoretical considerations in Section 2 — is still twice as long for men and about 60% longer for women.

In general, the employment careers of the 5% of individuals with the highest amount of lifetime unemployment are very unstable. On average, they exhibit not only an elevated number of unemployment spells but also more employment spells than the other individuals covered (even though their amount of lifetime

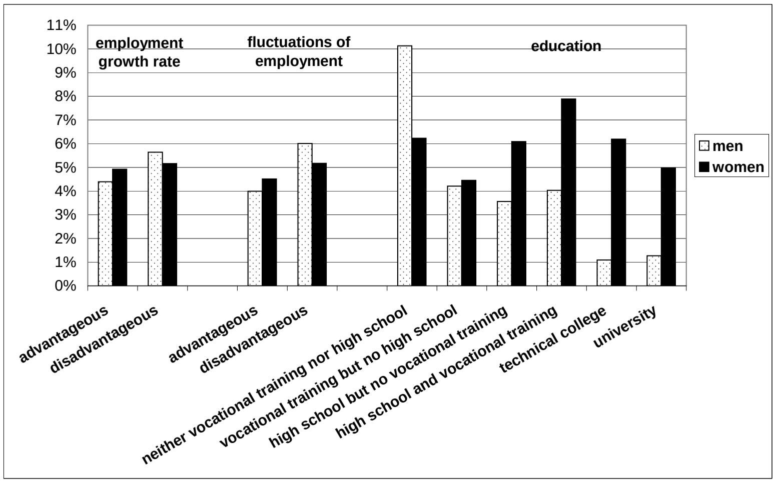

FIGURE 2.— Share of individuals with certain characteristics who are among the 5% of individuals with the highest amount of lifetime unemployment

employment is much smaller than the corresponding figure of the rest of the population). What is more, on average they also exhibit more periods with neither employment nor unemployment than the rest of our sample.

Figure 2 takes a closer look at the individuals with the highest amount of lifetime unemployment. It visualizes the share of individuals with certain characteristics who are among the 5% of our sample with the highest amount of lifetime unemployment. The focus is on education and the advantageousness of the occupation. Here we use two rather crude measures of the advantageousness of the occupation and label an occupation "advantageous" if its employment growth rate between 1975 and 2004 was stronger than the median employment growth rate over all occupations and/or if its fluctuations of employment were weaker than the median fluctuations of employment over all occupations.

If education and the advantageousness of the occupation were independent of the amount of lifetime unemployment one would expect all shares reported in Figure 2 to be close to 5%. For men, this seems clearly not to be the case. Rather it is obvious that men with a low educational level are faced with a much higher amount of lifetime unemployment than would be the case if education and lifetime unemployment were independent. For instance more than 10% of men with neither high school diploma nor vocational training are among the 5% of individuals with the highest amount of lifetime unemployment while only about 1% of men with a degree from a technical college or a university are among this group. For men, the advantageousness of the occupation (the variable at the center of this study) also seems to be strongly related to the amount of lifetime unemployment.

For women, results are in general less pronounced. Nevertheless for the advantageousness of the occupation they point in an intuitive direction. This is not necessarily the case for the education variable.

The aim of the following sections is to perform a multivariate analysis that clarifies whether the advantageousness of the occupation or one of the other variables introduced in Section 3 do indeed influence the amount of lifetime unemployment. These sections will concentrate on men because the data problems mentioned above as well as the descriptive evidence shown here make an analysis for women less promising.[11]

5. Methodology

For a multivariate analysis of the amount of lifetime unemployment it is important to recall that more than 60% of individuals in our sample were not unemployed between age 25 and age 50 at all. The rest of our sample exhibits a strictly positive amount of lifetime unemployment. That means we are faced with what is called *censoring* by most of the literature and somewhat more appropriately a *corner solution outcome* by Wooldridge (2002). Whatever the labeling an ordinary least square estimation of the amount of lifetime unemployment would lead to inconsistent results. The classical way to deal with censoring would be to use the Tobit estimator [proposed by Tobin (1958)]. We prefer, however, censored quantile regression (CQR), introduced by Powell (1986), as a more suitable alternative.

Compared to a Tobit estimator the CQR model offers several advantages: First, as shown by Powell (1986), it does not require homoscedasticity of the error terms. Second, it is consistent and asymptotically normal irrespective of the distribution of the error term as long as the conditional quantile of the error term is zero. Third, like the conventional quantile regression model introduced by Koenker and Bassett (1978), it allows marginal effects to differ between lower and higher conditional quantiles. This third point is especially important in the context of our study since we primarily focus on the role of occupation-specific human capital acquired early in the professional career and other factors in the upper tail of the distribution of lifetime unemployment.

In general, the CQR estimator for quantile θ assumes the following latent model:

$$y_i^* = x_i'\beta_\theta + \epsilon_{\theta i}, \tag{5.1}$$

where x_i is the vector of explanatory variables and $\epsilon_{\theta i}$ denotes the error term with a conditional quantile of zero, $\text{Quant}_\theta(\epsilon_{\theta i}|x_i) = 0$. y_i^* is the latent dependent variable.

[11]Results for women are available upon request.

When estimating the amount of lifetime unemployment we are faced with lower censoring at zero and no upper censoring.[12] In this case the following equation holds between the latent unemployment variable y_i^* and the observable amount of lifetime unemployment y_i:

$$y_i = \begin{cases} y_i^* & \text{if } y_i^* \geq 0 \text{ and} \\ 0 & \text{if } y_i^* < 0. \end{cases} \tag{5.2}$$

If lower censoring at zero is present, the conditional quantile of y is given by

$$\text{Quant}_\theta(y|x) = \max(0, x'\beta_\theta). \tag{5.3}$$

Powell (1986) showed that a consistent estimator for β_θ is obtained as a solution to minimizing

$$\frac{1}{N}\sum_{i=1}^{N}[[\theta - I(y_i < \max(0, x_i'\beta_\theta))][y_i - \max(0, x_i'\beta_\theta)]] \tag{5.4}$$

with respect to β_θ, where I is an indicator function that takes the value of unity when the expression holds and zero otherwise.

In Koenker and Bassett (1978)'s traditional quantile regression models linear programming is used to solve for the regression parameters. Because $\max(0, x_i'\beta_\theta)$ is not linear in β this is not possible for equation (5.4). Fortunately the literature suggests a number of ways to deal with this problem. The most prominent solutions are an iterative linear programming algorithm proposed by Buchinsky (1994) and a programming algorithm by Fitzenberger (1997). These approaches are, however, not without drawbacks: Fitzenberger (1997) and others point especially to the failure to reach asymptotic efficiency in practice, a high computational burden and a poor performance when a large proportion of the data is censored (as is the case in our application).

Therefore we make use of an improved estimator for censored quantile regressions introduced by Chernozhukov and Hong (2002) that overcomes many of the shortcomings of the more tested approaches. Chernozhukov and Hong (2002, p. 872) report that their "estimators are theoretically attractive (i.e., asymptotically as efficient as the celebrated Powell (...) estimator). At the same time, they are conceptually simple and have trivial computational expenses." In spite of these evident advantages they have not been widely used in the labor literature. Exceptions include Machado and Santos Silva (2008) and Ludsteck and Haupt (2007) who extend the method to censored panel data regressions.

The estimating procedure introduced by Chernozhukov and Hong (2002) consists of three steps. Now we briefly describe these steps and how we adjusted the procedure to our specific circumstances.

[12]Some of the studies on single unemployment episodes mentioned in Section 1 face not lower censoring but upper censoring. Koenker and Bilias (2001) and Lüdemann, Wilke and Zhang (2006) use CQR to approach this problem.

Step 1. Our first goal is to choose a subset of observations where the quantile line $x_i'\beta_\theta$ is above the censoring point. We start with a logit estimation of the model

$$\delta_i = \dot{x}_i'\gamma + \epsilon_{\gamma i} \tag{5.5}$$

where δ_i is an indicator of not-censoring and $\dot{x}_i$ is a transform of x_i. It is crucial that censoring is predicted as well as possible. Therefore we include a large number of explanatory variables in $\dot{x}_i$: a cubic polynomial in wage and establishment size, the advantageousness of the occupation, education and professional status dummies, three-digit occupation dummies as well as dummies for 326 West German administrative districts.

Next we select the sample

$$J_0(c) = \{i : \dot{x}_i'\hat{\gamma} > 1 - \theta + c\} \tag{5.6}$$

with c strictly between 0 and θ. We choose c such that $\#J_0(c)/\#J_0(0) = 0.9$. According to Chernozhukov and Hong (2002) this rule works well in simulations.

Step 2. Now we obtain an initial estimator $\hat{\beta}_\theta^0$ by running an ordinary quantile regression

$$y_i = x_i'\beta_\theta^0 + \epsilon_{\theta i}^0 \tag{5.7}$$

on the sample J_0. Chernozhukov and Hong (2002) show that the resulting estimator is consistent and useful for building up the efficiency of the last step. For Step 3 we calculate a sample with the properties

$$J_1(k) = \{i : x_i'\hat{\beta}_\theta^0 > k\} \tag{5.8}$$

where k plays a similar role as c did in Step 2. Much of the literature sets $k = 0$. We follow this approach but also make sure [as suggested by Gustavsen, Jolliffe and Rickertsen (forthcoming)] that $\#J_1/\#J_0 > 0.66$ and $\#\{J_0 \not\subset J_1\}/\#J_1 < 0.1$ in order to avoid using too small a sample and to ensure robustness.

Step 3. Finally we run another ordinary quantile regression

$$y_i = x_i'\beta_\theta^1 + \epsilon_{\theta i}^1 \tag{5.9}$$

using observations from J_1 this time. Chernozhukov and Hong (2002) show that the resulting estimator $\hat{\beta}_\theta^1$ not only works well in their simulations but is also consistent and asymptotically efficient.[13]

[13] Quantile regressions were calculated with Stata and its *qreg*/*sqreg* commands. For the regressions in Step 3 we relied on bootstrap standard errors with 200 replications.

6. RESULTS

6.1. *Baseline Regressions*

Results of our benchmark regressions are summarized in Table IV. Additionally, results for the most interesting regressors are visualized in Figure 3. We focus on the upper tail of the distribution of lifetime unemployment because we are most interested in finding factors important for those individuals with a very high amount of lifetime unemployment. Specifically we estimate CQR models for the 75th, 80th, 85th, 90th and 95th percentile of the conditional distribution of lifetime unemployment.[14]

The dependent variable of all our regressions is the amount of lifetime unemployment (measured in days). That means a negative sign of an explanatory variable's coefficient implies this variable is *ceteris paribus* associated with a smaller amount of lifetime unemployment and *vice versa*.

One focus of our study is to assess whether we can find evidence of the connection between occupation-specific human capital and lifetime unemployment outlined in Section 2. This would be the case if pursuing an advantageous occupation early in the professional career actually led to a significantly lower amount of lifetime unemployment. Therefore we discuss results concerning the advantageousness of the occupation in detail. Results on other variables are presented somewhat briefer.

Table IV and Figure 3 show that for men the (*ex post*) advantageousness of the occupation held early in the professional career is clearly related to the amount of lifetime unemployment. The more advantageous the occupation held on the 25th birthday the smaller the expected amount of lifetime unemployment. This is true for both measures of the advantageousness of the occupation, the trend employment growth rate between 1975 and 2004 and the standard deviation of the employment fluctuations over this period. The corresponding coefficients are always statistically significant at the 1% level and especially pronounced for higher quantiles of the conditional distribution of lifetime unemployment.

This result lends support to the hypothesis that occupation-specific human capital plays a role in determining the amount of lifetime unemployment. In the next subsection we address this hypothesis in greater detail and especially focus on the question whether there might be an omitted variable problem in our baseline regression (cp. Section 1). But for now we turn to the coefficients of our various control variables.

For the wage earned when turning 25 all coefficients are statistically significant and have negative signs. This relationship is especially pronounced in the right tail. Thus concerning the influence of the wage earned early in the professional career on lifetime unemployment efficiency wages and the like seem to be more important than rising reservation wages.

[14] Our sample is now reduced to 30,089 men for whom we have information on all regressors.

TABLE IV

Censored quantile regression results for men (baseline approach)

	lifetime unemployment				
	75th percentile	80th percentile	85th percentile	90th percentile	95th percentile
employment growth rate	-88.30***	-128.97***	-200.10***	-230.98***	-378.37***
	(10.45)	(15.29)	(21.55)	(35.02)	(59.09)
fluctuations of employment	3711.72***	5604.97***	8198.27***	11079.29***	16524.06***
	(388.41)	(650.54)	(795.96)	(1291.17)	(2054.79)
wage	-7.21***	-9.51***	-12.74***	-15.85***	-19.54***
	(0.38)	(0.45)	(0.55)	(0.66)	(0.76)
voc. training; no high school	-258.13***	-371.04***	-520.04***	-878.19***	-1210.36***
	(25.58)	(32.85)	(49.63)	(84.60)	(89.55)
high school; no voc. training	-220.07***	-297.29***	-478.83***	-727.62***	-1121.22***
	(44.23)	(60.79)	(98.66)	(166.73)	(182.82)
high school & voc. training	-257.79***	-389.91***	-474.65***	-720.38***	-1141.12***
	(30.11)	(38.79)	(82.68)	(131.07)	(264.62)
technical college	-279.73***	-355.12***	-504.37***	-945.26***	-1374.49***
	(26.96)	(34.63)	(51.70)	(91.65)	(117.22)
university	-253.28***	-373.82***	-551.04***	-977.56***	-1453.92***
	(27.17)	(35.25)	(51.96)	(89.12)	(115.09)
Hamburg	-89.90	-185.68**	-224.63*	-553.08***	-375.83
	(70.88)	(90.36)	(134.27)	(159.24)	(288.04)
Lower Saxony	-188.27***	-265.24***	-292.27**	-584.27***	-434.60*
	(62.31)	(92.17)	(125.83)	(151.46)	(245.64)
Bremen	-36.35	16.83	18.17	-75.89	-336.52
	(130.88)	(141.07)	(214.04)	(230.29)	(331.23)
North Rhine-Westphalia	-229.19***	-318.95***	-365.98***	-648.03***	-478.18**
	(60.37)	(82.69)	(119.54)	(146.07)	(218.79)
Hesse	-292.83***	-422.44***	-541.55***	-975.64***	-1256.73***
	(60.36)	(84.76)	(119.22)	(146.08)	(234.82)
Rhineland-Palatinate	-325.45***	-466.83***	-605.22***	-1025.77***	-1317.49***
	(63.31)	(82.63)	(120.69)	(145.27)	(227.44)
Baden-Württemberg	-376.77***	-541.26***	-693.38***	-1204.45***	-1528.42***
	(59.54)	(81.35)	(116.45)	(139.74)	(224.03)
Bavaria	-300.47***	-454.55***	-575.35***	-1031.89***	-1389.87***
	(60.28)	(84.21)	(116.25)	(141.57)	(220.60)
Saarland	-268.93***	-382.41***	-395.39**	-571.71***	-853.80***
	(70.83)	(103.41)	(164.49)	(189.49)	(279.20)
energy and mining	-197.47*	-270.77*	-508.77**	-557.67	-1111.84**
	(106.20)	(144.10)	(243.73)	(344.36)	(500.60)
manufacturing	-94.52	-87.90	-179.40	-189.74	-413.29
	(105.83)	(146.64)	(243.30)	(334.32)	(480.65)
construction	208.03*	230.01	207.65	252.07	8.31
	(110.23)	(154.46)	(244.90)	(335.50)	(470.69)
services	-101.21	-106.94	-196.94	-211.68	409.34
	(105.87)	(145.31)	(243.89)	(333.60)	(477.35)
public sector and other	-171.62	-185.06**	-292.14	-279.65	-357.27
	(104.90)	(145.28)	(245.50)	(340.98)	(495.61)
size of the establishment	-0.0006**	-0.0022***	-0.0043***	-0.0066***	-0.0084***
	(0.0003)	(0.0005)	(0.0007)	(0.0009)	(0.0020)
constant	1000.38***	1401.02***	2002.95***	3123.42***	4524.06
	(138.35)	(177.85)	(254.17)	(374.85)	(537.54)

Notes: Bootstrap standard errors in parentheses. *, (**), (***) indicates significance at the 10, (5), (1) per cent level. For a detailed description of variables used, see Section 3.

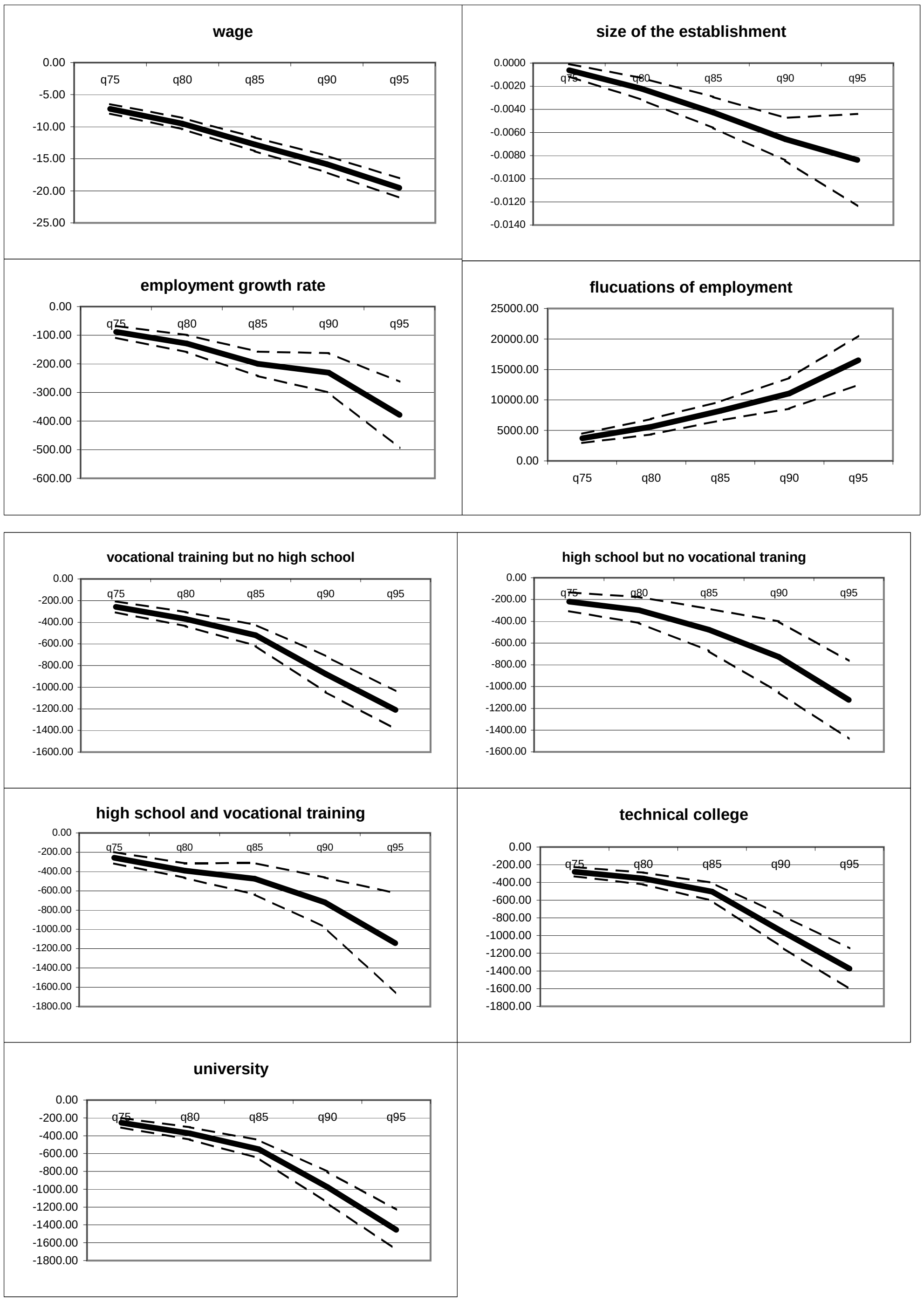

FIGURE 3.— Coefficients and 95% bootstrap confidence intervals for selected explanatory variables for censored quantile regressions (men; dependant variable: lifetime unemployment in days)

Level of education is also strongly related to the amount of lifetime unemployment. The control group consists of individuals with neither high school diploma nor vocational training. Its members have by far the highest expected amount of lifetime unemployment. Individuals with vocational training and particularly those with a tertiary degree do much better on average. For our education dummies all parameter estimates are statistically significant on the 1% level.

When it comes to the region, individuals who work early in their professional career in a federal state known for its poor economic performance are clearly faced with a comparatively high amount of lifetime unemployment. In contrast, results for the sector variable are not very clear-cut. The reference category is the agricultural sector and now few of the coefficients differ statistically significantly from zero. Only for the energy and mining sector do we find coefficients that are more or less statistically different from zero across quantiles: individuals engaged in the energy/mining sector at age 25 can expect a comparatively small amount of lifetime unemployment. For other sectors (almost) all coefficients are insignificant at the 5% level.

A priori one could have expected a bigger role for the sector variables. A reason for their apparent low importance might be that the results of Kambourov and Manovskii (2009) are valid not only for the United States but also for West Germany: sector information is of second order if one appropriately controls for occupations.

The effects of a number of explanatory variables are especially pronounced in the right tail. This is for instance the case for the variable representing the size of the establishment. As detailed above, this variable counts the number of employees of the firm for which the individual worked at age 25. As expected working in a large firm early in the professional career is *ceteris paribus* associated with a smaller amount of lifetime unemployment.

6.2. *An Extended Approach*

A potential weakness of our regressions in the last subsection is that the measures for the *ex post* advantageousness of the occuopation pursued at age 25 might not be exogenous. As detailed in Section 1 the corresponding coefficient could be biased if the *ex ante* advantageousness of an occupation was correlated with its *ex post* advantageousness and also with lifetime unemployment. While we would argue that the individuals in our sample could not know in 1975 which occupations would develop advantageously in the following 25 years, we cannot completely rule out that our results suffer from unobserved variable bias.

In order to address this issue and to strengthen our claim of a causal link still further, we will now present the results of a number of regressions which include not only our measures of the *ex post* advantageousness of the 56 different occupations but also a range of proxies for what could have been *ex ante* inferred about their advantageousness.

The first variable we introduce to capture the *ex ante* advantageousness of the

TABLE V

Censored quantile regression results for men (extended approach)

	(1)	(2)	(3)	(4)	(5)
	lifetime unemployment				
employment growth	-378.37*** (59.09)	-372.87*** (55.20)	-267.19*** (64.19)	-176.03** (69.70)	-119.65* (67.45)
fluctuations of employment	16524.06*** (2054.79)	17649.01*** (2172.42)	16421.33*** (2186.01)	9528.57*** (2380.80)	14919.68*** (1923.18)
median wage	-	222.23* (121.02)	1106.67*** (129.19)	689.88*** (156.66)	1096.51*** (135.41)
unemployment rate	-	-	351.69*** (26.32)	296.36*** (28.20)	331.29*** (22.33)
forecast I	-	-	-	297.96*** (50.81)	-
forecast II	-	-	-	-	-674.26*** (113.16)

Notes: 95th percentile; Bootstrap standard errors in parentheses. *, (**), (***) indicates significance at the 10, (5), (1) per cent level. Constant, education, region and sector dummies as well as variables for the wage and the size of the establishment at age 25 not displayed. For a detailed description of variables used, see Section 3.

56 different occupations are the coefficients obtained from adding occupation dummies to a Mincer-type wage equation for the years 1975 to 1977 (with education and year dummies as well as age and age squared included as additional regressors). Column (2) of Table V repeats the censored quantile regression for the 95th percentile of the distribution of lifetime unemployment from subsection 6.1 but also includes the occupation-specific wage effects from the Mincer-type wage equation.[15] For comparison, column (1) replicates the corresponding results from the benchmark regressions in subsection 6.1.

Economic theory is ambiguous concerning the expected sign of the occupation-specific wage variable: On the one-hand high wages could mean that productivity in these occupations was very high. If this was the case one would probably expect the coefficient in the quantile regressions to have a negative sign. On the other hand generally high wages for a specific occupation might lead the individuals to develop high reservation wages which might ultimately lead to more lifetime unemployment. It turns out that this latter effect seems to be the prevailing economic force. Individuals who early in their professional career pursue an occupation that *ceteris paribus* pays higher wages are faced with a higher amount of lifetime unemployment.

The introduction of occupation-specific wages leaves the coefficients of the variables measuring the *ex post* advantageousness of the 56 different occupations

[15]In both the current and the next subsection we focus exclusively on regressions for the 95th percentile. Almost all coefficients for other percentiles have the same signs. Generally they are statistically significant on even higher significance levels (as is obvious from Figure IV confidence intervals tend to be rather large for the 95th percentile).

largely unchanged.

As a second measure of the *ex ante* advantageousness of the 56 occupations in 1975, column (3) of Table V introduces the occupation-specific unemployment rate in 1977 as an additional regressor.[16] Our conjecture is that the higher the occupation-specific unemployment rate, the lower the *ex ante* advantageousness of the respective occupation. Thus we would expect a positive coefficient for the occupation-specific unemployment rate. This is indeed what we find — the corresponding coefficient is positive and highly significant.

The introduction of the occupation-specific unemployment rate in 1977 lowers the absolute value of the coefficient of the first variable measuring the *ex post* advantageousness of the 56 different occupations but leaves it highly significant. The coefficient of the second measure of the *ex post* advantageousness of an occupation, the employment stability over the business cycle, is left largely unchanged.

While the occupation-specific wage effects and unemployment rates in the mid-1970s might very well be strongly associated with the *ex ante* advantageousness of different occupations, they are static measures.[17] For the long term perspective of this study it might be relevant to also add measures of what occupations where *ex ante* perceived to develop in an advantageous way in the long run.

Data on individual expectations of occupational advantageousness from 1975 do not exist. However, as an indicator for what individuals could have known at that time, one can draw upon professional forecasts made by economists during that period. One such forecast is the study by Blüm and Frenzel (1975), a complex and detailed work that forecasts (from the perspective of 1975) the labor supply and demand separately for the 56 occupations covered by our analysis for the subsequent 15 years. Because the study was published by the research institute of Germany's Federal Employment Agency it can be assumed that it had a major influence on the Federal Employment Agency's occupational guidance policy and on individuals' expectations.

Columns (4) and (5) of Table V once again report censored quantile regressions with our measures of the *ex post* advantageousness of the 56 occupations. As in columns (2) and (3), the occupation-specific wage effects and unemployment rates in 1975 are also included. Besides, two distinct variables obtained from the study by Blüm and Frenzel (1975) are added. In column (4) this is the predicted occupation-specific ratio of labor demand to labor supply in 1990. A value of this variable greater than one signifies a predicted excess demand for labor for this occupation. The higher the variables' value the higher its *ex ante* advantageousness.

As an alternative, column (5) focuses exclusively on the demand side of the

[16]We use the rate for 1977 because the unemployment information in the IABS for 1975 and 1976 does not appear to be completely reliable.

[17]A third such measure could be based on whether occupation-specific employment exhibits a strong seasonal pattern. Including the standard deviation of occupation-specific employment for all days of 1977 as an additional regressor leaves the variables of interest largely unchanged.

labor market. Here, the occupation-specific ratio of predicted labor demand in 1990 to the actual number of employment relationships in 1970 is included as an explanatory variable. Higher values of this variable signify a higher growth rate of the demand for labor in the corresponding occupation. Such an occupation could therefore *ex ante* have been perceived to be more advantageous.

In contrast to the occupation-specific median wage the two measures for the *ex ante* predicted advantageous of the 56 relevant occupations do indeed lower the absolute values of the coefficients of our measures of the *ex post* advantageousness of these occupations. This lends some support to the hypothesis that our initial estimates of the *ex post* advantageousness of occupations suffer from unobserved variable bias. However the variables measuring the *ex post* advantageousness of an occupation still stay significant, though the coefficient of the employment growth rate is only significant on the 10% level in column (5) of Table V. The coefficient corresponding to the fluctuations of employment growth remains highly significant.

We included the two measures for the *ex ante* predicted advantageous of the 56 relevant occupations in order to assess whether they affected the coefficients of our measures of the *ex post* advantageousness of these occupations. Nevertheless it is illuminating to take a look at these coefficients themselves. While both of them are statistically highly significant, it is interesting to note that they exhibit opposite signs. Column (5) shows that the predictions made by Blüm and Frenzel (1975) about the demand side of the labor turned out to be reasonably good. Individuals who early in their professional career had chosen an occupation with a predicted rise in demand tend to have a lower amount of lifetime unemployment compared to the average.

However, the opposite is true for those individuals who early in their professional career worked in an occupation with an advantageous forecast of the gap between future labor supply and demand. If the individuals' *ex ante* perception of the advantageousness of a given occupation in 1975 in fact followed these recommendations to a significant extent, one might conjecture that this led so many young people to join occupations with a perceived future excess labor demand that eventually labor was in excess supply. Hence the forecast could have produced a so-called pork cycle [cp. Chavas and Holt (1991)].

6.3. *Robustness*

Before coming to our conclusion we will now report the outcomes of a number of checks that evaluate whether the results of the last subsection are robust to variations of the empirical setup. The starting point are the regression results reported in column (4) of Table V, that is, those from the extended approach with all control variables including the predicted occupation-specific ratio of labor demand to labor supply in 1990 based on Blüm and Frenzel (1975).[18]

[18] Qualitatively results are almost identical when using the occupation-specific ratio of Blüm and Frenzel (1975)'s predicted labor demand in 1990 to the actual number of employment

TABLE VI

CENSORED QUANTILE REGRESSION RESULTS FOR MEN (ROBUSTNESS)

	(1)	(2)	(3)	(4)
	lifetime unemployment		lifetime nonemployment	lifetime unemployment
employment	-176.03**	-118.46	-294.29*	-95.66
growth	(69.70)	(64.19)	(169.08)	(72.32)
employment	9528.57***	6828.78***	12501.86*	13263.63***
fluctuations	(2380.80)	(2186.01)	(2380.80)	(1923.18)
	equivalent to column (4) of table V	*including foreigners*		*1980 – 2004*

Notes: 95th percentile; Bootstrap standard errors in parentheses. *, (**), (***) indicates significance at the 10, (5), (1) per cent level. Constant, education, region and sector dummies, variables for the wage and the size of the establishment at age 25, occupation-specific wage coefficients for 1975–1977 and unemployment rates for 1977 as well as the predicted occupation-specific ratio of labor demand to labor supply in 1990 based on Blüm and Frenzel (1975) not displayed. For a detailed description of variables used, see Sections 3 and 6.2.

Table VI summarizes the coefficients for our measures of the advantageousness of an occupation for the extended approach from the last subsection as well as for a number of alternative specifications. Column (1) repeats column (4) of Table V while in columns (2) to (4) results are reported for the following alternative specifications:

First we check whether results are robust to the inclusion of foreigners working in West Germany. As noted in Appendix A throughout our study we have focused on individuals with a German passport. Foreigners (who are strongly overrepresented among individuals with a very high amount of lifetime unemployment) have so far been excluded.

Next we evaluate if altering the definition of unemployment from Section 3 changes our results. As an alternative definition we make use of the concept of *nonemployment* introduced by Fitzenberger and Wilke (forthcoming). Here all time periods not spent in employment that follow an employment spell and contain at least one spell of receiving unemployment benefits are counted as nonemployment. Modeled on lifetime unemployment's definition we define *lifetime nonemployment* as the total length (in days) of all nonemployment spells of an individual from age 25 to age 50 and use it as an alternative dependant variable.

Finally the IABS does not reliably cover all spells of unemployment benefit receipt for some of the years before 1980. That is why we reestimate our CQRs for 20-year periods starting at the earliest in 1980 instead of the 25-year periods beginning in 1975 or later used throughout the paper.

relationships in 1970 as control variable.

As Table VI shows, our results are qualitatively robust to the alternative specifications presented here. No coefficients ever change their signs and their orders of magnitude stay broadly constant as well. What is more the variable capturing the occupation-specific fluctuations of employment stays statistically significant throughout all the alternative specifications. The same is not always true for the employment growth variable which is at best weakly significant.

7. CONCLUSIONS

In this paper we showed that lifetime unemployment is very unevenly distributed in West Germany. Looking at selected birth cohorts we found that more than 60% of the individuals in our sample were not affected by unemployment between the age of 25 and 50 at all. On the other hand, half of the total amount of unemployment fell upon 5% of the men and 6% of the women in our sample.

Using censored quantile regressions we documented that for men pursuing what *ex post* turned out to be an advantageous occupation early in the professional career negatively affected the amount of lifetime unemployment. This relationship was especially strong for the upper tail of the distribution of lifetime unemployment. Even when we controlled for what could *ex ante* have been perceived to be an advantageous occupation the influence of *ex post* occupational advantageousness on lifetime unemployment remained strong.

These findings have several important implications: First they lend support to the finding of Kambourov and Manovskii (2009) that occupation-specific human capital is highly relevant. Second they are consistent with theories by Ljungqvist and Sargent (1998) and others that stress the connection between human capital and unemployment. They are also consistent with the theoretical considerations of Section 2: An explanation for our findings could indeed be that individuals with a disadvantageous investment in occupation-specific human capital experience an especially pronounced human capital depreciation at separation. Therefore, they exhibit a higher reservation wage and a reduced search effort. Third they are in line with the result by von Wachter and Bender (2006) that having good or bad luck early in the professional career can have significant and long-lasting consequences. Fourth — and maybe most importantly — they hint at a certain inflexibility of the German labor market. As outlined in Section 2, in a perfectly flexible labor market one would not expect the advantageousness of the occupation pursued early in the professional career to have a causal influence on lifetime unemployment.

To the best of our knowledge, ours is the first study to use a rich and reliable administrative micro data set and multivariate statistics to analyze unemployment over a 25-year period. This new perspective offers vast opportunities for future research. A focus on lifetime unemployment opens the door for investigations of the long-term distribution of unemployment, of labor market flexibility and of long-ranging effects of education and vocational training, amongst many others.

ACKNOWLEDGEMENTS

We thank Philipp vom Berge, Wolfgang Dauth, Martin Dietz, Tobias Pickelmann, Helmut Rudolph and Ulrich Wenzel as well as conference and seminar participants in Bonn, Duisburg, Nuremberg, Mannheim and Tallinn for helpful comments and suggestions. The usual disclaimer applies.

REFERENCES

Arulampalam, Wiji; Booth, Alison and Taylor, Mark (2000). Unemployment Persistence. *Oxford Economic Papers,* **52** 24–50.

Bender, Stefan; Haas, Anette and Klose, Christoph (2000). IAB Employment Sample 1975-1995. *Schmollers Jahrbuch: Journal of Applied Social Science Studies,* **120** 649–662.

Blüm, Adalbert and Frenzel, Udo (1975). *Quantitative und qualitative Vorrausschau auf den Arbeitsmarkt der Bundesrepublik Deutschland - Stufe 3.*, IAB, Nürnberg.

Brooks, Bradley (2005). Chronic Unemployment: A Statistical Profile. *Statistics Canade, Analysis in Brief 31.*

Buchinsky, Moshe (1994). Changes in the U.S. Wage Structure 1963-1987: Application of Quantile Regression. *Econometrica* **62** 405–458.

Chavas, Jean-Paul and Holt, Matthew (1991). On Nonlinear Dynamics: The Case of the Pork Cycle. *American Journal of Agricultural Economics* **73** 819–828.

Chernozhukov, Victor and Hong, Han (2002). Three-Step Censored Quantile Regression and Extramarital Affairs. *Journal of the American Statistical Association* **97** 872–882.

Fitzenberger, Bernd (1997). A Guide to Censored Quantile Regression. In Madalla, G.S. and Rao, C.R. (Eds.). *Handbook of Statistics*, Elsevier, Amsterdam.

Fitzenberger, Bernd and Wilke, Ralf (forthcoming). Unemployment Durations in West-Germany Before and After the Reform of the Unemployment Compensation System During the 1980s. *German Economic Review.*

Galiani, Sebastian and Hopenhayn, Hugo (2003). Duration and Risk of Unemployment in Argentina. *Journal of Development Economics* **71** 199–212.

Gustavsen, Geir Waehler; Jolliffe, Dean and Rickertsen, Kyrre (forthcoming). Censored Quantile Regression and Purchases of Ice Cream. *Food Economics.*

Kalwij, Adriaan (2004). Unemployment Experiences of Young Men: On the Road to Stable Employment? *Oxford Bulletin of Economics and Statistics* **66** 205–237.

Kambourov, Gueorgui and Manovskii, Iourii (2009). Occupational Specificity of Human Capital. *International Economic Review* **50** 63–115.

Koenker, Roger and Bassett, Gilbert (1978). Regression Quantiles. *Econometrica* **46** 33–50.

Koenker, Roger and Bilias, Yannis (2001). Quantile regression for duration data: A reappraisal of the Pennsylvania Reemployment Bonus Experiments. *Empirical Economics* **26** 199–220.

Kurtz, Beate and Scherl, Hermann (2001). Zur interpersonalen Verteilung von Arbeitslosigkeit in kohortenbezogener langfristiger Betrachtung: Untersuchung am Beispiel männlicher Arbeitnehmer der Jahrgänge 1925 bis 1930. *Mitteilungen aus der Arbeitsmark- und Berufsforschung* **34** 127–136.

Ljungqvist, Lars and Sargent, Thomas (1998). The European Unemployment Dilemma. *Journal of Political Economy* **106** 514–550.

Lüdemann, Elke; Wilke, Ralf and Zhang, Xuan (2006). Censored Quantile Regressions and the Length of Unemployment Periods in West Germany. *Empirical Economics* **31** 1003–1024.

Ludsteck, Johannes and Haupt, Harry (2007). An Empirical Test of the Reder Hypothesis and Specific Human Capital Against Standard Wage Competition. *University of Munich, Dicussion Papers in Economics 1977.*

Machado, Jose and Santos Silva, Joao (2008). Quantiles for Fractions and Other Mixed Data. *University of Essex, Economics Discussion Paper 656.*

Powell, James (1986). Censored Regression Quantiles. *Journal of Econometrics* **32** 143–155.

Tobin, James (1958). Estimation of Relationships for Limited Dependent Variables. *Econometrica* **26** 24–36.

Wachter, Thill von and Bender, Stefan (2006). In the Right Place at the Wrong Time: The Role of Firms and Luck in Young Workers' Careers. *The American Economic Review* **96** 1679-1705.

Wooldridge, Jeffrey (2002). *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge.

APPENDIX A: DATA CLEANSING

In order to ensure valid and undistorted results and to limit the impact of non-standard employment careers we exclude the following groups from our analysis:

- East Germans because they have only been included in our data since the early 1990s.[19]
- Individuals who were employed with coverage by the social security system or recipients of some form of unemployment benefit for the very first time after their 30th birthday.
- Foreigners, i.e. individuals that at the end of their career history did not hold a German passport.

Additionally, it is important to identify meaningful employment spells. When an individual holds multiple jobs at the same time we delete all of these but the one with the highest wage. Employment spells with the following characteristics are also discarded:

- Spells of marginal employment that have only been covered by our data since 1999.
- Employment spells with a wage below the marginal part-time income threshold. We believe that for these employment spells the wage information is corrupt (in fact many of them indicate a daily wage of zero).
- Spells during which the individual was in an apprenticeship. These spells are arguably not comparable to "regular" employment episodes.

APPENDIX B: ADVANTAGEOUS AND DISADVANTAGEOUS OCCUPATIONS

Table VII lists the ten most advantageous and the ten most disadvantageous occupations. As could be expected many of the most disadvantageous occupations are associated with manual tasks while advantageous occupations often involve the provision of services or the knowledge of new technologies.

[19]We label all individuals "East German" whose first employment or unemployment spell registered by the social security system took place in East Germany.

TABLE VII

The 10 most advantageous and disadvantageous occupations

	employment growth rate	fluctuations of employment growth
10 most advantageous occupations		
1	Natural scientists and humanists n.e.c.	Bankers and insurance specialists
2	Social workers	Technical specialists
3	Lawyers	Chemical workers
4	Helpers not elsewhere covered	Miners
5	Teachers	Printers
6	Health professional n.e.c.	Paper makers and processing operatives
7	Security guards	Technicians
8	Cleaners	Precision fitters, assemblers
9	Engineers	Alimentary occupations
10	Cooks	Painters

	employment growth rate	fluctuations of employment growth
10 most disadvantageous occupations		
1	Spinners	Lawyers
2	Miners	Cleaners
3	Textile processing operatives	Ground transport occupations n.e.c.
4	Leather makers and processing operatives	Teachers
5	Textile makers	Water and air transport occupations
6	Building laborer, general	Security guards
7	Bricklayers, concrete workers	Farmers
8	Wood preparers and product makers	Helpers n.e.c.
9	Machinists	Guest attendants, housekeepers
10	Construction material makers	Warehouse and transport workers