Arbeiten aus dem

OSTEUROPA-INSTITUT REGENSBURG

Wirtschaftswissenschaftliche Abteilung

Working Papers

Nr. 275 Oktober 2009

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

Achim SCHMILLEN* und Joachim MÖLLER**

* Osteuropa-Institut Regensburg, University of Regensburg and Institute for Employment Research (IAB). Correspondence to: Achim Schmillen, e-mail: schmillen@osteuropa-institut.de

** Institute for Employment Research (IAB), IZA and University of Regensburg.

OSTEUROPA-INSTITUT
REGENSBURG

Landshuter Str. 4
93047 Regensburg
Telefon: 0941 943 5410
Telefax: 0941 943 5427
E-Mail: oei@osteuropa-institut.de
Internet: www.osteuropa-institut.de

ISBN 978-3-938980-24-8

Contents

List of Tables

List of Figures

ABSTRACT

Building on a large administrative micro data set for the time span 1975–2004 we look at lifetime unemployment for West German birth cohorts 1950 to 1954. Descriptive evidence shows a highly uneven distribution of unemployment in West Germany — more than 60% of the individuals in our sample were not unemployed for a single day over the better part of their professional career while almost half of the total amount of unemployment fell upon 5% of the individuals covered.

We employ censored quantile regressions to explain the amount of individual lifetime unemployment. Explanatory variables are either characteristics of the individual (like education), or of the job (like the wage) or the employer (like the size of the firm) early in the professional career. A particular emphasis is placed on the importance of the occupation: we find that for men working in a disadvantageous occupation at age 25 *ceteris paribus* leads to a significantly higher amount of lifetime unemployment. Educational attainment or the wage earned at age 25 are also related to the amount of men's lifetime unemployment, amongst others. Some of these variables show very interesting patterns when looking at different quantiles. For women results are in general less clear-cut.

KEYWORDS: Lifetime unemployment, Censored-Quantile Regressions, Occupation-specific human capital

JEL-CLASSIFICATION: J64, J24.

1. INTRODUCTION

Descriptive evidence shows that lifetime unemployment in West Germany is highly unevenly distributed — more than 60% of individuals 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 know what distinguishes the individuals with a very high amount of lifetime unemployment from the rest of the population. Easily observable individual characteristics like low education could certainly be a reason for an increased likelihood of belonging to this group. 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 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 Bender and von Wachter (2006)].

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

The contribution our paper is to link these two bodies of literature and to investigate the causal effect of a specific investment in human capital early in the professional career on the amount of lifetime unemployment for selected West German cohorts. Following Kambourov and Manovskii (2009) we specifically focus on occupation-specific human capital and find that — at least for men — working in a disadvantageous occupation early in their professional career (at age 25) is indeed connected to a significantly higher amount of lifetime unemployment. Of course, this might be due to a selection effect where "good" workers select themselves in jobs with bright perspectives while "bad" workers" are left with the less advantageous rest.

To evaluate whether the advantageousness of the occupation held early in the professional career really has a causal effect on lifetime unemployment we put some effort in controlling for individual and job characteristics as well as the occupation-specific unemployment and wage rates in the mid 1975s — information that individuals could have had when they made their occupational choices at the beginning of our observation period (which runs from 1975 to 2004). Additionally we use an elaborated long-run occupation-specific labor market forecast published by the German Federal Employment Agency in 1975, Blüm and Frenzel (1975), as a further indicator of the state of knowledge in the mid-1970s.

Even with all these control variables our measures of the eventual advantageousness of the different occupations turns out to heavily influence lifetime unemployment. We argue that with the inclusion of the controls these measures can be considered as exogenous to the occupational choice at the early stage of a professional career and that we therefore identify a causal effect of the advantageousness of the occupation chosen at young age on lifetime unemployment.

Our study is related to the literature examining determinants of unemployment incidence, duration and distribution. 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. Due to widespread data limitations this literature almost completely focuses on distinct periods of unemployment over a relatively short overall time span. Only few studies look at the occurrence, distribution or determinants of unemployment over more than a couple of years. However, like Kurtz and Scherl (2001) or Brooks (2005), these tend to be confined to descriptive evidence.[1]

Most of the relevant literature is based on survey data. Instead we make use of an administrative micro data set that offers not only a larger sample size but also more reliable information on unemployment and other variables of interest. Even more importantly, in contrast to the majority of the literature we do not focus on the duration of distinct unemployment spells. Instead our data allow us to assess what we call lifetime unemployment: the total length of all unemployment spells 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. 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 such a long time span.

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 concludes.

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 her wages. However, in such a framework we 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)]. As will be outlined in the following, a modification of such a model where an investment in a disadvantageousness kind of

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

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 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 dependents on specific individual or job characteristics. Specifically it is assumed to depend on whether the human capital acquired in the previous job is still in demand at the time of a separation. If this is indeed an important factor in determining the human capital depreciation rate, 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 then 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.[2]

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 to spells of unemployment associated with the receipt of benefits.[3]

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

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

[3] 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 perform our descriptive and multivariate analyses separately for men and women. We are also very careful to compare the respective results.

from 1975 (when the individuals born 1950 turned 25) to 2004 (when the cohort of 1954 turned 50).

In order to ensure valid and undistorted results and to limit the impact of non-standard employment careers we additionally exclude a number of groups from our analyses and delete certain employment spells. Details on this can be found in appendix A.

While section 4 mainly focuses on descriptive evidence on lifetime unemployment and its interpersonal distribution, section 6 contains a multivariate analysis of lifetime unemployment. As well as lifetime unemployment as the dependent variable all explanatory variables are constructed with the help of the IABS data set. Some of them are individual characteristics (like education). Others are taken from the job held by the individual on her 25th birthday or, if the individual was not employed at this date, at 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 cleaning (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.[4]

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 unemployment might be interesting in itself:

- **Education** level. It is well-known that education is closely related to the occurrence of unemployment. Since education and occupation are strongly connected as well, controlling for education is of outmost importance. We do this by including five dummy variables that measure whether an indi-

[4] 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.

vidual 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. Our 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 more 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.[5]

- Weekly **wages** earned at the age of 25. This variable might be interpreted as a proxy for unobserved individual characteristics and we expect that higher wages textitceteris paribus lead to a lower amount of lifetime unemployment.
- **Sector** of the firm for which the individual worked on 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 make statements of the different sectors' roles in determining the amount of lifetime unemployment.
- **Region** where the job pursued at age 25 was based. This variable might once again be construed as a proxy for unobserved personal or firm heterogeneity. It is measured by dummy variables for the 10 West German federal states ("*länder*"). A priori we would assume that working in a well-off state (like Bavaria) 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. This might indicate whether a company has (otherwise unobserved) positive or negative characteristics. Since generally speaking in Germany the influence of labor unions is strongest in big companies it might also be a signal for whether employees have some bargaining power that might lead to less lay-offs and a lower risk of unemployment. The size of the establishment is measured by simply adding up the number of its employees. We expect that individuals working for a larger firm at the beginning of their professional career *ceteris paribus* face a smaller amount

[5] While some information in our data set (for instances on the duration of employment or unemployment periods and on wages) is extremely reliable this is not always the case when it comes to education. We use the imputation mechanism suggested by Fitzenberger, Osikominu and Völter (2006) to obtain reliable education information.

of lifetime unemployment.

4. DESCRIPTIVE EVIDENCE

Before we turn to our multivariate analyses 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 section 3. For the regressions in section 6 this numbers reduce to 30,089 men and 27,589 women for whom we have information on all regressors.

We start by looking at 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 the self-employment and work as a civil servant (cp. section 3).[6]

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 1.59 unemployment spells with an average length of 225.28 days. Women are on average 1.03 times unemployed with average unemployment episodes lasting for 226.67 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 (2,821.20 of 9,496.60 days[7]). The average woman is even counted as neither employed nor unemployed during almost half of her prime age (on average 4,393.22 days out of 9,496.59 are spent neither in employment nor in unemployment and only 4,870.60 in employment).

The perhaps surprising importance of time spans 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 it is not feasible to compare these figures 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*

[6] If for an individual information on employment or unemployment is only available some time after her 25th birthday or not right until 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.

[7] 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)	6318.18	357.23	2821.20	9496.60
average number of spells	4.17	1.59	4.48	10.23
average spell duration (in days)	1516.59	225.28	629.76	928.17
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)	3291.50	3626.12	2578.97	9496.5
average number of spells	9.92	10.58	15.89	36.39
average spell duration (in days)	331.84	342.75	162.27	260.95
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)	6477.46	185.18	2833.94	9496.60
average number of spells	3.86	1.11	3.88	8.85
average spell duration (in days)	1676.69	166.48	730.57	1072.50
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)	4870.60	232.78	4393.22	9496.59
average number of spells	3.63	1.03	3.88	8.54
average spell duration (in days)	1341.77	226.67	1130.78	1111.75
relative occurrence of state	0.513	0.025	0.463	100
relative occurrence of state (not considering *neither employed nor unemployed*)	0.954	0.046		100
5% of women with the highest amount of lifetime unemployment				
	employed	unemployed	neither employed nor unemployed	total
average total duration (in days)	4164.24	2106.89	3225.45	9496.59
average number of spells	7.46	6.95	10.43	24.84
average spell duration (in days)	557.97	303.24	309.29	382.32
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)	4907.73	134.25	4454.61	9496.59
average number of spells	3.43	0.72	3.54	7.69
average spell duration (in days)	1431.46	187.58	1257.96	1235.69
in percent	51.7	1.4	46.9	100
in percent (not considering *neither employed nor unemployed*)	0.973	2.7		100

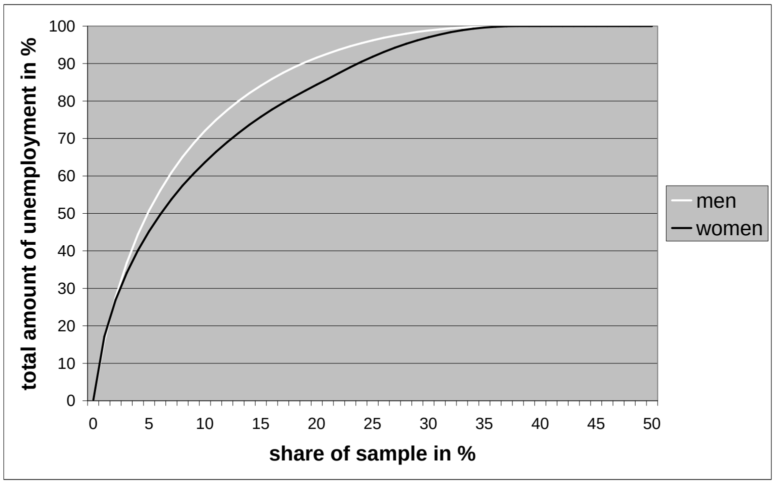

FIGURE 1.— Inverted Lorenz curves for the interpersonal distributions of the total amount of unemployment for men and women

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 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.[8]

[8]One might infer from figure 1 that the total amount of unemployment is more unevenly

TABLE III

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

percentile	men	women
50th	0	0
60th	0	0
70th	123	182
80th	410	364
90th	1101	668
95th	2091	1106
99th	4671	2727

Table III again stresses the high conecntration 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 50th and 60th 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 2091 days during their prime age. Men whose lifetime unemployment was in the top percentile were even unemployed for 4671 days or more — that is more than 12 years! Corresponding figures for women show 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 characterize those (say 5% or 10%) of individuals who are faced with very high 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.

Before turning to an econometric explanation of lifetime unemployment, we take a closer look at 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.

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. 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 remarkable that the elevated amount of lifetime unemployment of the 5% of men and women with the highest amount of unemployment is on average largely due to a higher number of unemployment spells and only to a lesser extent to an increased duration of these

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.

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 unemployment spell is "only" 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 have not only more unemployment spells than the other individuals covered but also more employment spells (even though their amount of lifetime 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, the advantageousness of the occupation and the wage received at age 25. Here we use a crude measure 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. Likewise the wage earned at 25 is called "high" if it is higher than the median of sample wages earned at that age.

If education, the advantageousness of the occupation and the wage at age 25 were independent of the amount of lifetime unemployment none of the shares reported in figure 2 should differ significantly from 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 university degree are among this group. For men, the wage earned at age 25 and the advantageousness of the occupation (the variable in the center of this study) also seem to be related to the amount of lifetime unemployment.

For women results are in general less pronounced. Nevertheless for both wages and the advantageousness of the occupation they point in an intuitive direction outlined in section 3. This is not necessarily the case for the education variable.[9]

To sum up our descriptive evidence, we find that lifetime unemployment is distributed highly unevenly in West Germany. This makes it very relevant to find the determinants of particularly high lifetime unemployment. Descriptive evidence indicates that both personal characteristics and characteristics of the job held at age 25 are connected with the amount of lifetime unemployment.

[9]Similar figures for the other explanatory variables introduced in section 3 are not shown here but available upon request. For men these other variables also seem to be related to the amount of lifetime unemployment. For women the connection sometimes is again less clear-cut.

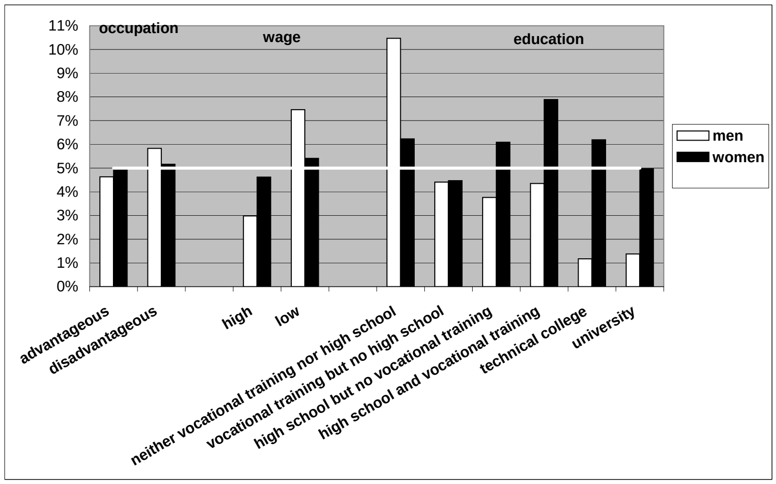

FIGURE 2.— Share of individuals with certain characteristics who are among the 5% of individuals with the highest amount of lifetime unemployment

Only an econometric analysis can clarify whether the advantageousness of the occupation or one of the other variables introduced in section 3 do indeed determine a very high amount of lifetime unemployment. This is the aim of the following sections.

5. METHODOLOGY

For a multivariate analysis of the amount of lifetime unemployment it is important to recall that for both men and women 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).[10] We will follow the majority of the literature and henceforth use the term *censoring* but whatever the labeling an ordinary least square estimation of the amount of lifetime unemployment would lead to biased results. The classical way to deal with censoring would be to use what is usually called a Tobit estimator [proposed by Tobin (1958)]. We prefer, however, censored quantile regression (CQR), introduced by Powell (1986), as a more suitable alternative.

[10] Wooldridge (2002, p. 517-519) reserves the term *censoring* for situations where "a data problem arises because [a latent dependant variable] is censored above or below some value; that is, it is not observable for part of the population." He also points out that *censored* and *corner solution outcomes* can be "transformed into the same statistical model"

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 whatever 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 relevant in the context of our study since we primarily want to find out whether occupation-specific human capital acquired early in the professional career and other factors are relevant for individuals with a very high amount 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, $\mathrm{Quant}_\theta(\epsilon_{\theta i}|x_i) = 0$. y_i^* is the latent dependent variable.

When estimating the amount of lifetime unemployment we are faced with lower censoring at zero and no upper censoring.[11] 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

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

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

[11] 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.

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.

Step 1. Our first goal is to chose 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 good 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 ("*Kreise*").

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 somewhat ad-hoc 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

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

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 is consistent and asymptotically efficient.

The next section summarizes our benchmark regressions. Quantile regressions were calculated with Stata and its *qreg*/*sqreg* commands. Because *qreg*'s analytical standard errors have frequently been criticized [e.g. by Koenker and Hallock (2001)] we relied on bootstrap standard errors with 200 replications obtained with the command *sqreg* for the quantile regressions in step 3.

6. RESULTS

Results of our benchmark regressions for men and women are summarized in tables IV and V, respectively. Additionally, results for the most interesting regressors are visualized in figures 3 and 4. We focus on higher ends of the distribution of the amount of personal unemployment because we are most interested in finding the factors that are associated with a very high amount of lifetime unemployment. Specifically we estimate CQR models for the 75th, 80th, 85th, 90th and 95th percentile.

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*.

Since one focus of our analysis is to assess whether pursuing an advantageous occupation early in the professional career leads to a significantly lower amount of lifetime unemployment, we discuss results concerning the advantageousness of the occupation in detail. Results on other variables are presented somewhat briefer. In the next subsection we look at the CQRs for men and then turn to the results for women.

6.1. *Results for Men*

Table IV and figure 3 show that for men the **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

TABLE IV

CENSORED QUANTILE REGRESSION RESULTS FOR MEN, DEPENDENT VARIABLE: LIFETIME UNEMPLOYMENT IN DAYS

variable	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)
vocational training but 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 but no vocational training	-220.07***	-297.29***	-478.83***	-727.62***	-1121.22***
	(44.23)	(60.79)	(98.66)	(166.73)	(182.82)
high school and vocational 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.

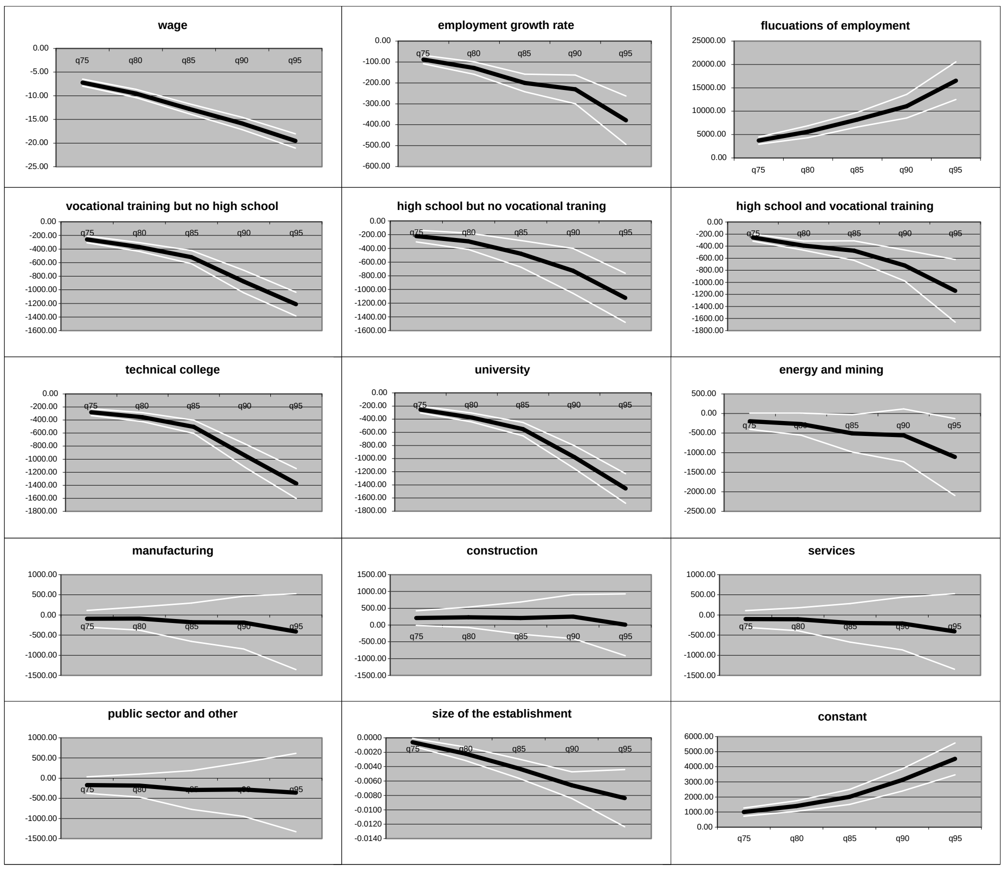

FIGURE 3.— Censored quantile regressions for men; dependant variable: lifetime unemployment in days; coefficients and 95% bootstrap confidence intervals

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 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. Our regressions control for a number of other factors which might be correlated with the advantageousness of the occupation (e.g. sector) or be interpreted as a measure of otherwise unobserved personal heterogeneity (e.g. wage). Therefore we can be confident that what we capture is indeed a causal connection between the advantageousness of the occupation and the amount of lifetime unemployment.

In subsection 6.3 we address this issue in greater detail and present what we think is even more compelling evidence that the advantageousness of the occupation does indeed influence the amount of lifetime unemployment. But for now turn to the coefficients of our various control variables. When estimating the amount of lifetime unemployment for men we find that most coefficients are statistically significant and have the expected sign.

This is by all means the case for the **wage** earned when turning 25. The wage coefficients do have the expected signs and are all statistically significant. Higher wages (which — as mentioned above — might be interpreted as a proxy for otherwise unobserved favorable personal characteristics) are associated with fewer days of unemployment over the professional career. This relationship is especially pronounced for higher quantiles of the amount of lifetime unemployment's distribution function.

The level of **education** is also strongly related to the amount of lifetime unemployment. Our 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 on average much better. For our education dummies all parameter estimates are statistically significant on the 1% level.

When it comes to the **region** the control group is the North German *land* of Schleswig-Holstein. As could be expected individuals who early in their professional career work in Schleswig-Holstein or other *länder* known for a poor economic performance tend to exhibit a comparatively high amount of lifetime unemployment. By contrast those who work in Bavaria or Baden-Württemberg, Germany's two most prosperous federal states on their 25th birthday ceteris paribus accumulate a significantly smaller number of unemployment days during their professional career. Once again this effect is strongest for higher quantiles of the distribution function of the amount of lifetime unemployment.

In contrast to the variables discussed so far 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 over all estimated 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 might have maybe assumed a bigger role for the sectoral 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: sectors information is of second order if one appropriately controls for occupations.

The effects of a number of explanatory variables are especially pronounced for higher quantiles of the amount of lifetime unemployment's distribution function. This is also the case for the variable representing **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. Amongst other things this meant to catch otherwise unobserved firm heterogeneity. As could be expected working in a large firm early in the professional career is *ceteris paribus* associated with a smaller amount of lifetime unemployment.

6.2. *Results for Women*

While the coefficients estimated for men usually hold the expected sign and are statistically significant table V and figure 4 show that this is not always the case for the CQRs for women. Here, our results do not imply a statistically significant relationship between lifetime unemployment and some of the explanatory variables.

Regarding the **advantageousness of the occupation** held early in the professional career results from the estimations for men are to some extent confirmed when looking at the data for women. That is, we again find consistently negative coefficients associated with our first measure of the advantageousness of the occupation held when turning 25 (the trend employment growth rate between 1975 and 2004). However, none of the five estimated coefficients is statistically significant at the 10% level. For our second measure of the advantageousness of the occupation (the standard deviation of the employment fluctuations) results are even less assuring. Coefficients are positive for some, but negative for other percentiles. None of them is statistically different from zero, at least not on the 5% significance level.

All in all we would argue that also for women there might still be some evidence that the advantageousness of the occupation held early in the professional career influences the amount of lifetime unemployment. However this evidence is not as strong for women as it is for men. Therefore subsection 6.3, which assesses whether we can really call this influence causal, will focus exclusively on men.

As has already been mentioned we obtain relatively few significant and some outright counterintuitive results for some of the other explanatory variables when estimating the amount of lifetime unemployment for women. This is for instance the case for **wages**: for women wages seem not to play a statistically significant

TABLE V

Censored quantile regression results for women, dependant variable: lifetime unemployment in days

variable	75th percentile	80th percentile	85th percentile	90th percentile	95th percentile
employment growth	-12.56	-7.56	-4.19	-14.26	-47.96
rate	(21.50)	(8.93)	(11.91)	(18.92)	(40.94)
fluctuations of	-583.08	-426.41	-99.23	662.25	2496.74*
employment	(494.84)	(278.01)	(513.21)	(613.56)	(1343.36)
wage	0.04	-0.20	-0.82	-1.96**	-0.94
	(0.56)	(0.34)	(0.60)	(0.78)	(1.67)
vocational training	-12.96	-16.43	-51.85***	-91.48***	-228.10***
but no high school	(20.21)	(11.79)	(17.19)	(26.99)	(55.51)
high school but no	-10.87	18.18	117.70	121.19	-49.59
vocational training	(72.66)	(93.47)	(85.93)	(118.15)	(180.12)
high school and	77.24**	101.83**	69.75	129.11	132.97
vocational training	(38.13)	(43.68)	(62.12)	(102.30)	(203.71)
technical college	124.23***	142.25***	157.21***	156.95**	61.70
	(45.19)	(42.81)	(60.61)	(89.51)	(134.81)
university	-97.72**	-110.00***	-96.69***	-88.73	-186.41***
	(38.41)	(38.53)	(36.34)	(65.10)	(93.60)
Hamburg	0.08	-10.16	-0.71	-13.64	84.69
	(48.34)	(51.70)	(60.05)	(91.14)	(148.52)
Lower Saxony	-17.24	-54.50	-45.50	-44.24	-30.11
	(39.16)	(39.99)	(47.28)	(68.70)	(134.12)
Bremen	-0.87	-34.10	45.41	250.88	1024.81***
	(33.44)	(75.47)	(137.44)	(160.03)	(382.63)
North	-84.98**	-102.62***	-149.69***	-177.20***	-116.45
Rhine-Westphalia	(42.13)	(36.83)	(43.06)	(57.49)	(116.52)
Hesse	-125.49***	-150.86***	-211.56***	-296.11***	-389.33***
	(34.64)	(38.85)	(46.20)	(59.53)	(119.34)
Rhineland-	-163.98***	-159.73***	-221.94***	-284.65***	-337.46***
Palatinate	(44.80)	(39.72)	(45.50)	(69.48)	(127.17)
Baden-	-246.62***	-234.89***	-272.07***	-368.09***	-453.13***
Württemberg	(42.42)	(41.75)	(46.42)	(58.60)	(119.96)
Bavaria	-112.91***	-130.43***	-205.25***	-305.31***	-410.49***
	(34.82)	(38.53)	(44.37)	(58.96)	(116.37)
Saarland	-45.64	-92.99**	-161.61***	-210.04**	13.99
	(52.58)	(38.40)	(55.59)	(96.27)	(231.69)
energy and mining	-53.53	-99.57	-183.23	-671.95*	-952.22
	(68.94)	(141.08)	(290.74)	(382.19)	(988.04)
manufacturing	11.78	-50.63	-91.63	-543.78	-797.04
	(53.76)	(136.02)	(280.70)	(357.79)	(980.63)
construction	-54.68	-64.54	-101.00	-582.36	-933.06
	(59.03)	(139.93)	(291.74)	(360.53)	(988.52)
services	-71.42	-113.3	-191.69	-648.90*	-923.49
	(53.44)	(133.64)	(279.69)	(356.03)	(980.92)
public sector and	-157.01**	-173.19	-244.22	-718.93**	-969.55
other	(54.64)	(136.10)	(280.29)	(360.34)	(978.76)
size of the	-0.0045	-0.0034***	-0.0050***	-0.0059***	-0.0082**
establishment	(0.0041)	(0.0010)	(0.0012)	(0.0017)	(0.0035)
constant	478.35***	621.73***	862.64***	1591.28***	2328.25**
	(57.10)	(140.44)	(289.59)	(361.32)	(988.60)

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.

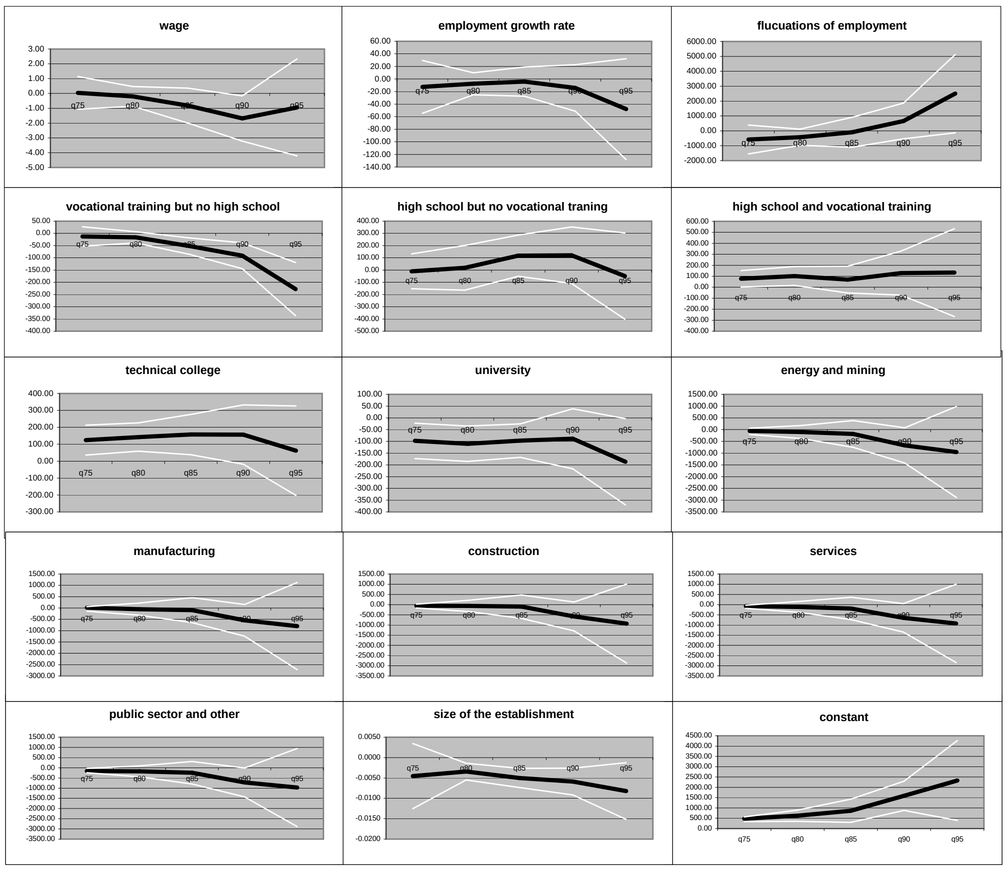

FIGURE 4.— Censored quantile regressions for women; dependant variable: lifetime unemployment in days; coefficients and 95% bootstrap confidence intervals

role in determining the amount of lifetime unemployment. Apart from the 90th percentile coefficients are all very small and confidence intervals pretty large.

At first glance, some of the results for **education** seem puzzling: while for all quantiles university graduates face a smaller probability of being affected by high unemployment than individuals with neither high school diploma nor professional education, this is not true for graduates from technical colleges. Even though these individuals also hold a tertiary degree our CQRs consistently say that they are more likely to be faced with a high amount of lifetime unemployment than members of the control group (though this difference is only significant at the 5% level for the 75th, 80th and 85th percentiles). Similarly surprising, a high school diploma does not bring a statistically significant decline in the amount of lifetime unemployment.

But even though individuals with a high school diploma and vocational training do not face significantly lower lifetime unemployment than the control group, women without high school diploma but with professional training do (at least when looking at the 85th to 95th percentile). A likely explanation for the puzzling results for women with high school diploma but no vocational training, high school diploma and vocational training and a degree from a technical college is that all these groups are rather small in our sample. Altogether only 1160 women fall in one of these three categories, that is less than 5% of our sample.

The dummies indicating the **region** where the individual works on her 25th birthday are associated with more intuitive results than the education dummies. Women who early in their professional career work in a *land* that is economically well-off like Bavaria and Baden-Württemberg tend to face a comparatively small amount of lifetime unemployment. In contrast those employed in Schleswig-Holstein, Bremen or other *länder* with a more difficult economic environment accumulate a significantly larger number of unemployment days during their professional career. All in all regional effects are qualitatively similar for men and women.

In a way this can also be said about our **sector** variables. For men only one sector exhibits coefficients that differ significantly from those of our reference category (agriculture) over all estimated quantiles. For women results are even more sobering: for all 5 sector variables only three coefficients differ from zero at the 10% significance level. Apparently the low explanatory power of our sector variables (already mentioned above) combined with the more general problems of our regressions for women (also discussed above) leads to a very weak link between sector variables and women's amount of lifetime unemployment.

Apart from the constant, the last variable in our regression is again the **size of the establishment**. Here, our results from the regressions for women are almost as clear as those for men. For women working at a large firm at the age of 25 is *ceteris paribus* associated with a statistically significantly reduction in the amount of lifetime unemployment. This is the case for all but the 75th percentile of the distribution function of the amount of lifetime unemployment used for our estimations and especially pronounced for the 90th and 95th percentile.

6.3. *An Extended Approach*

In subsection 6.1 we argued that at least for men our regressions established a causal link between the advantageousness of the occupation held early in the professional career and the amount of lifetime unemployment. We claimed that the causal interpretation of the coefficients was valid not only because of the covariates included in the regression, but also because from our point of view the individuals in our sample could not know in 1975 which occupations would develop advantageously in the following 25 years.

Still, a word of caution might be appropriate: so far we have used relatively crude measures to control for individual or job characteristics. More importantly, our regressions in subsection 6.1 cannot really guarantee that occupation-specific human capital gained early in the professional career causally affects the amount of 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 are in fact caused by unobserved personal heterogeneity: It could be the case that individuals with unobservable characteristics rewarded by the labor market but not captured by control variables like the wage earned when turning 25 were more likely to get into what they perceived as advantageous occupations in the first place and that these occupations really turned out to be advantageous.

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 advantageousness of the 56 different occupations but also a range of proxies for what could have been inferred about this advantageousness by the individuals in 1975. Since the results for women were less conclusive than those for men we concentrate on men here. Results of corresponding regressions for women are available upon request.

The first variable we introduce to capture the perceived advantageousness of the 56 different occupations are occupation dummies from 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 VI repeats the censored quantile regression for the 95th percentile of the distribution function of the amount of lifetime unemployment from subsection 6.1 but also includes the occupation-specific median wages in 1975.[12] For comparison, column (1) of table VI 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 jobs in this occupations are very productive. If this was the case one would 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 working in this occupation to develop high reservation wages which might ultimately

[12] In the current subsection we focus exclusively on regressions for the 95th percentile.

TABLE VI

Censored quantile regressions for men; 95th percentile; dependent variable: lifetime unemployment in days

variable	(1)	(2)	(3)	(4)	(5)
employment	-378.37***	-372.87***	-267.19***	-176.03**	-119.65*
growth	(59.09)	(55.20)	(64.19)	(69.70)	(2835.41)
fluctuations of	16524.06***	17649.01***	16421.33***	9528.57***	14919.68***
employment	(2054.79)	(2172.42)	(2186.01)	(2380.80)	(1923.18)
median wage	-	222.23*	1106.67***	689.88***	1096.51***
		(121.02)	(129.19)	(156.66)	(135.41)
unemployment	-	-	351.69***	296.36***	331.29***
rate			(26.32)	(28.20)	(22.33)
forecast I	-	-	-	297.96***	-
				(50.81)	
forecast II	-	-	-	-	-674.26***
					(113.16)

Notes: 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.

lead to more lifetime unemployment. It turns out that the latter effect seems to be the prevailing economic force and that higher occupation-specific wages prevailing at the beginning of an individual career are associated with higher lifetime unemployment.

At the same time the introduction of occupation-specific wages leaves the coefficients of the variables measuring the advantageousness of the 56 different occupations largely unchanged. Whatever the effect of a high wage level in a given occupation on the perceived advantageousness of this occupation is, it is apparently independent of the actual advantageousness' impact on lifetime unemployment.

As a second measure of the perceived advantageousness of the 56 occupations in 1975, column (3) of table VI introduces the occupation-specific unemployment rate in 1977 as an additional regressor.[13] Our conjecture is that the higher the occupation-specific unemployment rate, the lower the perceived 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 highly significant and positive. At the same time the introduction of the occupation-specific unemployment rate in 1977 lowers the absolute value of the coefficient of the first variable measuring the actual advantageousness of the 56 different occupations but leaves it highly significant. The coefficient of the second measure of the advantageousness of an occupation, the employment stability over the business cycle, is left largely unchanged.

While the occupation-specific median-wage and unemployment rate in the

[13]We use the rate for 1977 because the unemployment information at the beginning of our sample period is less reliable.

mid-1970s might very well be strongly associated with the perceived advantageousness of different occupations they are static measures. For the long term perspective adopted in our study it might be relevant to also add measures of what occupations where perceived to develop in an advantageous way in the years following 1975.

Data on individual expectations of occupational advantageousness from 1975 do not exist. However, as an indicator what individual could have known at that time, one could employ professional forecasts made by researchers in that period. Fortunately, there is the study by Blüm and Frenzel (1975), a complex and detailed work that tries do predict (from the perspective of 1975) the labor supply and demand separately for the 56 occupations covered by our analysis for the subsequent 15 years.[14] Because it was published by the research institute of Germany's Federal Employment Agency it can be assumed that it had major influence on the Federal Employment Agency's occupational guidance policy.

Columns (4) and (5) of table VI once again report censored quantile regressions with our measures of the advantageousness of the 56 occupations. As in columns (2) and (3) the occupation-specific median wage and unemployment rate 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. So 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 the perceived advantageousness.

As an alternative, column (5) focuses exclusively on the demand side of the 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 in 1975 have been perceived to be more advantageous.

In contrast to the occupation-specific median wage the two measures for the predicted advantageous of the 56 relevant occupations do indeed lower the absolute values of the coefficients of our measures of the actual advantageousness of these occupations. This lends some support to the hypothesis that our measures of the advantageousness of occupations in part only capture a sorting of individuals with unobserved positive characteristics. However the variables measuring the unexpected future 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 VI. The coefficient corresponding to the fluctuations of employment growth remains highly significant.

With the battery of control variables from subsection 6.1 as well as the occupation-specific median wage and unemployment rate in 1975 still included as additional

[14]While our study extends further than 1990 we still use the results obtained by Blüm and Frenzel (1975) as a proxy of perceived long-term advantageousness of different occupations around 1975.

regressors we feel confident that the regressions reported in columns (4) and (5) of table VI capture a causal effect of the unexpected future advantageousness of the occupation pursued early in the professional career on lifetime unemployment.

We included the two measures for the predicted advantageous of the 56 relevant occupations in order to assess whether they affected the coefficients of our measures of the actual 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 reverse 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 a professional career worked in an occupation with an advantageous forecast of the gap between future labor supply and demand. If the individuals' 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)].

7. CONCLUSION

In this paper we first of all 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. On the other hand, half of the total amount of unemployment falls upon 5% of the men and 6% of the women in our sample. This result makes it politically highly relevant to find out more about the factors that are associated with very high lifetime unemployment.

The econometric analysis leads to several insights. Using censored quantile regressions we found that education and several characteristics of the job held when turning 25 have a statistically significant effect on the amount of lifetime unemployment. We also documented that pursuing an advantageous occupation early in the professional career is negatively correlated with the amount of lifetime unemployment. This relationship is especially strong for higher quantiles of the distribution of lifetime unemployment.

If this finding is really caused by interpersonal differences in occupation-specific human capital — and we argue that this is indeed the case — it has two important implications: It firstly lends support to theories by Ljungqvist and Sargent (1998) and others that stress the connection between human capital and unemployment. Secondly it has direct policy implications. If some individuals are faced with high unemployment because they invested in disadvantageous occupation-specific human capital in the past it might be advisable to publicly fund re-training programs that provide these individuals with more advantageous occupation-specific human capital.

ACKNOWLEDGEMENTS

We thank Martin Dietz, Helmut Rudolph, Philipp vom Berge and Ulrich Wenzel as well as conference and seminar participants in Bonn, Nuremberg, Mannheim and Tallinn for helpful comments and suggestions.

REFERENCES

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

Bender, Stefan and Wachter, Thill von (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.

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; Osikominu, Aderonke and Völter, Robert (2006). Imputation Rules to Improve the Education Variable in the IAB Employment Subsample. *Schmollers Jahrbuch: Journal of Applied Social Science Studies* **126** 405–436.

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 (2008). Censored Quantile Regression and Purchases of Ice Cream. *Food Economics* forthcoming.

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.

Koenker, Roger and Hallock, Kevin (2001). Quantile Regression. *Journal of Economic Perspectives,* **15** 143-156.

Kruppe, Thomas; Müller, Eva; Wichert, Laura and Wilke, Ralf (2008). On the Definition of Unemployment and its Implementation in Register Data. *Schmollers Jahrbuch: Journal of Applied Social Science Studies* **128** 461–488.

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 Arbeitsmarkt- 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.

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 additionally exclude the following groups from our analyses:

- East Germans because they are only included in our data since the early 1990s.[15]
- 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. Also employment spells with the following characteristics are 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: OCCUPATIONS

B.1. *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.

[15]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

10 MOST ADVANTAGEOUS AND DISADVANTAGEOUS OCCUPATIONS

	10 most advantageous occupations	
	employment growth rate	fluctuations of employment growth
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
	10 most disadvantageous occupations	
	employment growth rate	fluctuations of employment growth
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

B.2. *Occupational Mobility*

Concerning occupational mobility about 40% of the men and more than 50% of the women in our sample never change their occupation between age 25 and 50. For another 20% of both men and women we record one change of occupation. 95% of the men in our sample change occupations seven times our less while 95% of women change it at most four times.

If we focus on the number of distinct occupations held from age 25 to age 51 and therefore omit all the occupational changes by which an individual returns to an occupation she has already pursued earlier in her professional career we of course find less occupational mobility: 95% of the men in our sample hold five distinct occupations our less during their prime age while 95% of women hold at most four occupations.

For histograms of the number of prime-age occupation changes in our sample see figure 5 for men and figure 6 for women. The left panels lists all changes while the right panels omit occupation returning to previously pursued occupations. Because a number of individuals in our sample exhibit a large number of occupation changes (especially when looking at all changes, irrespective of whether the individual had worked in the occupation before) the histograms do not depict the number of occupation changes for the percentile of individuals with the most occupation changes.

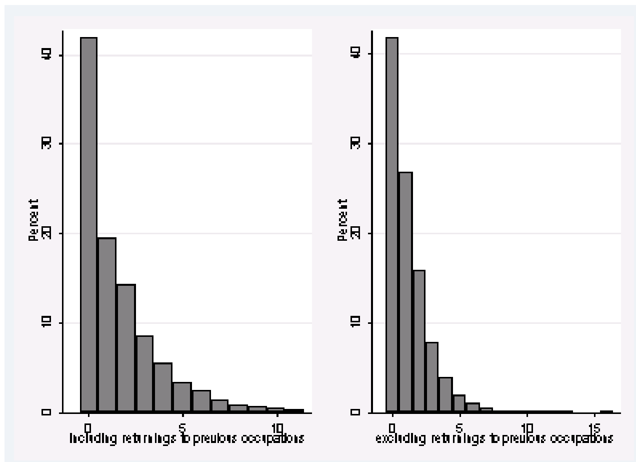

FIGURE 5.— Number of occupation changes for men

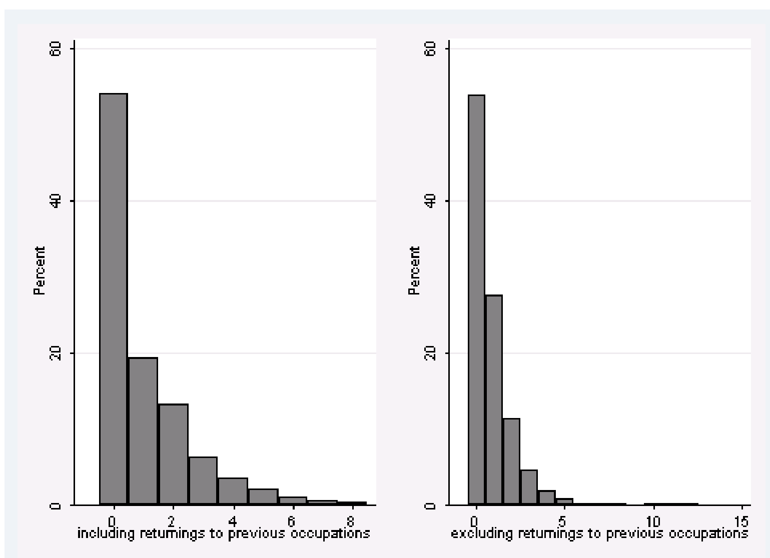

FIGURE 6.— Number of occupation changes for women