— *October 31, 2012*—

THE SCARS OF YOUTH
—
EFFECTS OF EARLY-CAREER UNEMPLOYMENT ON FUTURE UNEMPLOYMENT EXPERIENCES

ACHIM SCHMILLEN[1,2] AND MATTHIAS UMKEHRER[1,3]

Does early-career unemployment cause future unemployment? We answer this question with German administrative matched employer-employee data that allow us to follow more than 800,000 individuals over 24 years. Using a censored quantile instrumental variable estimator and instrumenting early-career unemployment with local labor market conditions at labor market entry, we show that youth unemployment has significant and long-lasting scarring effects. These effects are especially pronounced in the right tail of the unemployment distribution where an additional day of youth unemployment increases prime-age unemployment by up to six and a half days.

KEYWORDS: Scarring, state dependence, censored quantile instrumental variable regressions.
JEL-CLASSIFICATION: J64, J62, C20.

1. INTRODUCTION

Over the past five years youth unemployment has risen considerably in the United States and most European countries. In 2011 the OECD-wide unemployment rate for 15- to 24-year-olds stood at 16.2 percent — that is one of the highest levels within the last 25 years. Figures for some of the larger OECD member countries were even more elevated, with the youth unemployment rate reaching 17.3 percent in the United States, 20 percent in the United Kingdom, 29.1 percent in Italy and a stunning 46.4 percent in Spain [source: OECD (2012)]. These worryingly high rates have stoked fears that "[t]he harm today's youth unemployment is doing will be felt for decades, both by those affected and by society at large" [The Economist (2011, p. 60)].

In order to decide if such fears are justified, one needs to distinguish between two conflicting notions of early-career unemployment: One contends that during the first years on the labor market an adjustment process takes place where job shopping enables individuals to offset disadvantageous initial conditions, gather heterogenous experiences and find their place in the professional world [cf. Freeman and Wise (1982) or Topel and Ward (1992)]. From this viewpoint, youth unemployment could be seen as nothing more than a temporary nuisance and

[1]Institute for Employment Research (IAB)
[2]Institute for East and Southeast European Studies (IOS)
[3]Correspondence to: Matthias Umkehrer, Institute for Employment Research (IAB), Regensburger Strasse 104, D-90478 Nürnberg, Germany, e-mail: matthias.umkehrer@iab.de, phone: +49 (911) 179-6211, fax: +49 (911) 179-3296.

any observed persistence in unemployment would be due to temporally correlated individual differences in the probability of experiencing unemployment. If, by contrast, early-career unemployment delayed the accumulation of productive skills and knowledge or prevented the formation of tight employer-employee matches, the picture would change dramatically: unemployment, in particular, might then exhibit *true state dependence*, i.e. unemployment early in the professional career might causally lead to more unemployment later in life.[1]

Ultimately, the question of which of the two conflicting views on early-career unemployment predominates, can only be answered empirically. This is exactly what our study attempts to achieve. The focus is on Germany, i.e. an economy that has recently made less news because of high youth unemployment than Spain, Greece and some other countries. Still, at 8.5 percent Germany's youth unemployment rate in 2011 was 50 percent higher than its general rate of joblessness [source: OECD (2012)]. This is one reason why, in our view, Germany provides a most interesting environment to study the (long-term) effects of early-career unemployment. Other reasons include that Germany is Europe's largest economy and that its dual education system and other institutions associated with school-to-work transitions have long attracted a lot of attention [cf. e.g. Heckman (1993) or Acemoglu and Pischke (1998)]. Moreover, Ryan's (2001, p. 49) assertion that "an economic mechanism as fundamental as state dependence (...) is unlikely to be nationally specific" makes our endeavor conceptually relevant for the study of the labor markets of developed economies more generally [cf. also von Wachter and Bender (2006)].

Our study documents that in Germany unemployment is highly persistent amongst a group of individuals. Even though we find some evidence in favor of an adjustment process during the first years on the labor market, we reach the conclusion that this persistence is (at least to a large extent) due to true state dependence: On average, every day of (involuntary) unemployment during the first eight years of the professional career induces two additional days of unemployment during the subsequent 16 years, other things being equal. What is more, this *scarring effect* of youth unemployment varies considerably across the (conditional) distribution of prime-age unemployment. In fact, scarring is strongest in the right tail of this distribution. While at its median an additional day of youth unemployment leads to an increase of prime-age unemployment by less than one day, at the 95$^{\text{th}}$ percentile another day of early-career unemployment induces almost six and a half days of prime-age unemployment, *ceteris paribus*. These high numbers imply that the long-term scarring effect of youth unemployment is not only statistically significant but also economically important.

We base our analysis on an administrative matched employer-employee data set that contains the universe of social security records in Germany. From these, we extract the complete employment biographies of all 827,114 men who grad-

[1] If "unemployment (...) alters preferences, prices or constraints that determine, in part, future unemployment" this is called true state dependence by Heckman and Borjas (1980, p. 247).

uated from Germany's dual education system between 1978 and 1980.[2] This system combines apprenticeships in a company and vocational education at a school in one course. In our view, it is an ideal institutional environment to study the effects of early-career unemployment. That is because a majority (around 60 percent) of young people enter the German labor market through the system, because apprentices constitute a fairly homogeneous group with regard to their experiences, training and background and because by focusing on its graduates we avoid, in large part, any problems caused by unobserved initial conditions [cf. Hoffmann (2010)].

Our data make it possible to identify the exact time and place of labor market entry for all 827,114 individuals and to track them for every day of the first 24 years of their professional career. Instead of relying on a traditional analysis of distinct unemployment spells focusing on durations or Markov transition rates, we examine whether an individual's total amount of unemployment during the eight years after graduation influences the overall length of unemployment spells in the subsequent 16 years. Compared to a period-to-period approach, this strategy is better able to capture medium- and long-run scarring effects of youth unemployment. It also provides more suitable measures of long-term labor market "success" or "failure".

A large proportion of individuals in our sample experience no or only short periods of unemployment during their professional career while others suffer from repeated and prolonged periods of joblessness. That is why estimations of the conditional mean function may leave unrevealed important aspects of the relationship between early-career and prime-age unemployment. Consequently, we make use of the innovative censored quantile instrumental variable (CQIV) estimator introduced by Chernozhukov, Fernández-Val and Kowalski (2011). This estimator not only takes account of the fact that almost 60 percent of the individuals in our sample experience no prime-age unemployment at all but also allows marginal effects to vary over the conditional distribution of prime-age unemployment. We are thus able to focus on those men who during their prime age are unemployed for a much longer time than most other individuals with comparable observable characteristics.

Drawing on Gregg (2001), we instrument involuntary youth unemployment with local labor market conditions at labor market entry. More specifically, the local unemployment rate right before graduation from the dual education system is used as an instrument for (involuntary) early-career unemployment. We consider this instrument to be (a) relevant because it influences the quality of initial matching of graduates to firms, (b) ignorably assigned because the choice of location at labor market entry can assumed to be exogenous given location-specific fixed effects and (c) excluded because time-varying patterns of economic conditions, the accumulation of skills and early matching processes prevent it

[2] We concentrate on men because data problems make an analysis for women conceptually difficult [cf. Appendix A].

from influencing prime-age unemployment through channels other than youth unemployment. Ultimately, the IV strategy allows us to do more than "just" to show that unemployment is highly persistent amongst a group of individuals: we argue that we capture a causal relationship.

This study contributes to the broader literature on true state dependence [partly surveyed in Ryan (2001)].[3] Early work by Heckman and Borjas (1980), Ellwood (1982) and Corcoran and Hill (1985) finds little evidence of scarring in American data. A more recent study for the United States by Mroz and Savage (2006) documents permanent wage losses due to early unemployment experiences but no significant unemployment effects. European research usually finds stronger evidence in favor of state dependence: Results by Nilsen and Reiso (2011) and Nordström Skans (2011) suggest that it exists for Norway and Arulampalam, Booth and Taylor (2000), Arulampalam (2001), Gregg (2001) and Gregg and Tominey (2005) find the same for Great Britain [Burgess, Propper, Rees and Shearer (2003) report negative effects of early-career unemployment only for unskilled but slightly positive effects for skilled individuals]. Concerning Germany, very little is known about the scarring effects of youth unemployment. The few relevant studies — most prominently Mühleisen and Zimmermann (1994), Schmelzer (2010) and Niedergesäss (2012) — tend to address state dependence more generally and universally confirm its existence.

The remainder of this paper is structured as follows: Conceptual considerations are discussed in the next section followed by a brief description of our matched employer-employee data set. In Section 4, we present descriptive evidence on the longer-term distribution of unemployment and unemployment dynamics over the professional career. Section 5 contains the results of our multivariate analysis, discusses their robustness with regard to variations of the empirical setup and investigates effects on different quantiles of the (conditional) distribution of prime-age unemployment. Section 6 concludes.

2. CONCEPTUAL CONSIDERATIONS

Theoretical explanations of scarring usually rely on one of the following three mechanisms: First, in many search and matching frameworks, unemployed individuals lower their reservation wages over time. As shown by Mortensen (1986), this behavior could on the one hand shorten the duration of unemployment periods. On the other hand, however, it could also mean that long-term unemployed

[3] More generally, this study aims to contribute to the literature on long-term effects of labor market events or decisions early in the professional career. Prominent examples include von Wachter and Bender (2006) who demonstrate that displacement leads to persistent wage losses for some groups of young workers while for others losses are substantial but drop to zero within five years and Raaum and Røed (2006), Stevens (2008) and Oreopoulos, von Wachter and Heisz (2012) who show that business cycle conditions at time of labor market entry have economically significant and long-lasting wage and employment effects. Cf. also the more structurally oriented literature on career dynamics like Keane and Wolpin (1997) or Hoffmann (2010).

individuals accept jobs that are not really a suitable match. For this reason they could be more likely to become unemployed again in the future.

Second, models by Pissarides (1992), Acemoglu (1995) and others stress the importance of human capital. They conjecture that valuable skills and/or knowledge are depreciated during unemployment. Indeed, Edin and Gustavsson (2008) show that in Sweden one year of non-employment is on average associated with a five-percentile move down the skill distribution. Such a loss of human capital lowers an individual's productivity and leads to persistently lower earnings and a higher risk to experience unemployment. Youth unemployment might be particularly harmful because the greatest investments in learning are usually made at the beginning of the professional career [cf. Ben-Porath's (1967) life-cycle human capital model]. Moreover, for young people forgone work experience during unemployment might mean that crucial skills are never even acquired.

Third, if employers are unable to perfectly observe applicants' productivity when making hiring decisions, they may use previous unemployment spells as a screening device. They may thus prefer to hire workers with less unemployment experience. Such stigma effects of unemployment are prominently incorporated into the models of Vishwanath (1989), Lockwood (1991), Gibbons and Katz (1991) and Kroft, Lange and Notowidigdo (2012). Empirically, Gibbons and Katz (1991) find that at least part of the wage and employment consequences of job displacement appear to be due to stigma effects.[4]

Against this backdrop, we test whether there is a causal link between early-career unemployment and long-term labor market outcomes with the help of the following econometric model of prime-age unemployment:

$$m^*_{i,c,t2} = \overline{c} + \alpha d_{i,c,t1} + \gamma v_{i,c,t1} + \mathbf{x}'_{i,c,t0}\beta + \mu_i + \eta_r + \nu_c + u_{i,c,t2}, \tag{2.1}$$

where subscript $c = \{1978, 1979, 1980\}$ denotes the labor market entry cohort, subscript $i = \{1, ..., N\}$ the individual and subscript $r = \{1, ..., R\}$ the district of the training firm. $t = \{t0, t1, t2\}$ indicates whether a variable is measured prior to labor market entry, early in the professional career or during prime-age, respectively.

Prime-age unemployment (m^*_{t2}) is the dependent variable while regressors include a vector of control variables ($\mathbf{x}_{t0}$) and a constant ($\overline{c}$). All control variables (graduation age, daily remuneration, occupation, sector, size and average wage of the training firm) are measured before labor market entry and can arguably be considered exogenous. Besides, prime-age unemployment is determined by ef-

[4]Recent research by Kroft, Lange and Notowidigdo (2012) suggests that actual employer behavior towards unemployed job applicants might be more easily explained by stigma effects of unemployment than by a depreciation of their human capital. It should also be noted that a plethora of alternative explanations for the existence of state dependence exists. Underlying factors mentioned in the literature include contracts [Beaudry and diNardo (1991)], labor unions, hiring and firing costs, discouragement or habituation effects [Clark, Georgellis and Sanfey (2001)], the lack of physical capital after recessions or the different bargaining powers of insiders and outsiders; cf. Margolis, Simonnet and Vilhuber (2001) for an overview.

fects specific to the individual (μ_i), the district of the training firm (η_r) and the labor market entry cohort (ν_c) as well as an i.i.d. error term (u_{t2}).

The pivotal explanatory variable is unemployment early in the professional career. More specifically, we follow Ellwood (1982) and Mroz and Savage (2006) and explicitly distinguish between involuntary and voluntary early-career unemployment (d_{t1} and v_{t1}, respectively). The majority of the literature on state dependence makes such a distinction only implicitly. However, if one defines involuntary unemployment in the "Keynesian" sense [i.e. as "a situation where an unemployed worker is willing to work for less than the wage received by an equally skilled employed worker, yet no job offers are forthcoming", Shapiro and Stiglitz (1984, p. 434)] both mechanisms and political implications of state dependence differ widely between voluntary and involuntary unemployment.

Ultimately, we are mainly interested in estimating the size of α. This is challenging because not all variables included in equation 2.1 are in fact observable. First and foremost, our administrative data set covers all periods of *registered* unemployment but does not directly include a measure of involuntary early-career unemployment. "Total unemployment, however, is identically the sum of its involuntary and voluntary components. Isolating one of these components is sufficient to distinguish them empirically, since the other is identically the residual" [Mroz and Savage (2006, p. A13)].

More formally, we are interested in the scarring effects of d_{t1} but only observe the overall amount of early-career unemployment, $m_{t1} = d_{t1} + v_{t1}$. Solving this equation for d_{t1} and inserting into equation 2.1 yields

$$m^*_{i,c,t2} = \bar{c} + \alpha m_{i,c,t1} + (\gamma - \alpha) v_{i,c,t1} + \mathbf{x}'_{i,c,t0}\beta + \mu_i + \eta_r + \nu_c + u_{i,c,t2}. \tag{2.2}$$

As we observe neither μ_i, η_r nor v_{t1} directly, the probability limit of α from estimating equation 2.2 by ordinary least squares is given by

$$plim\alpha_{ols} = \alpha + \frac{cov(\mu, m_{t1})}{var(m_{t1})} + \frac{cov(\eta, m_{t1})}{var(m_{t1})} + (\gamma - \alpha)\frac{cov(v_{t1}, m_{t1})}{var(m_{t1})}. \tag{2.3}$$ [5]

Can we say anything about the likely direction of this bias? Starting with the term $\frac{cov(\mu, m_{t1})}{var(m_{t1})}$ on the right-hand side of equation 2.3, one might intuitively assume that omitting information on unobserved individual characteristics — e.g. an individual's ability or motivation — would upwardly bias simple OLS estimates of α. In fact, the omission of individual effects only leads to upwardly biased estimates if $cov(\mu, m_{t1}) < 0$. In the case where μ mainly captures a person's unobserved ability or motivation one might indeed be tempted to expect a negative relationship of this variable with the total durations of both prime-age and early-career unemployment. Yet, as pointed out by Neumark (2002), the returns to job hopping and/or the returns to search might be considerably higher

[5] Cf. von Wachter and Bender's (2006) exposition of the mechanisms of wage determination and theories of job mobility.

for individuals with high ability. Such mechanisms would generate a positive correlation between μ and m_{t1} in equation 2.3 and thus contributes to downward-biased OLS estimates. What is more, individual heterogeneity might also be due to differences in preferences or norms. These in turn might affect both early-career and prime-age unemployment and might bias OLS estimates of α in an unknown direction.

In a similar way, not controlling for location-specific fixed effects would induce a bias in OLS estimates of α because of the term $\frac{cov(\eta,m_{t1})}{var(m_{t1})}$ (which can be interpreted as reflecting initial sorting of individuals). More specifically, it would lead to a downward bias if $cov(\eta, m_{t1}) > 0$ and to an upward bias if $cov(\eta, m_{t1}) < 0$.

Finally, the existence of voluntary unemployment probably contributes to a downward bias in OLS estimates of α. That is because of the last term on the right-hand side of equation 2.3, $(\gamma - \alpha)\frac{cov(v_{t1},m_{t1})}{var(m_{t1})}$: voluntary and overall youth unemployment are positively related [i.e. $cov(v_{t1}, m_{t1}) > 0$] and the average costs incurred by voluntary unemployment can be expected to be smaller than the costs incurred by involuntary unemployment ($\gamma < \alpha$).

Because of these diverse sources of bias we will pursue an identification strategy where we will sequentially remove one source of bias after the other. We will begin with a simple OLS estimation that will already control for quite a number of socio-demographic and firm-related variables. Next, we will draw on Heckman and Borjas (1980) as well as Gregg (2001) and Neumark (2002) and will instrument involuntary early-career unemployment with local labor market conditions prevailing at the training firm's location right before graduation.[6] Finally, we will control for initial sorting of individuals by including fixed effects for the training firms' districts.

More specifically, the local unemployment rate right before graduation will be used as an instrument, where locations will be defined by the administrative districts of Germany's Federal Employment Agency. We can distinguish 141 such districts with unemployment rates varying considerably from 0.9 (the district of Nagold in 1979) to 8.2 percent (the district of Saarbrücken in 1978). In our view the local unemployment rate at graduation is a suitable instrument because we consider it to be relevant, ignorably assigned and excluded.

The instrument is relevant because the conditions that prevail just before labor market entry have an effect on whether an individual becomes unemployed after graduation from the dual education system and, if this is the case, on the duration of the resulting unemployment spell. In fact, Raaum and Røed (2006) use Norwegian data to show that individuals entering the labor market have larger difficulties in establishing themselves on this market if local unemployment rates are high. Besides, conditions at labor market entry affect the quality of initial matching of apprentices to firms [cf. Bowlus (1995)]. In turn, the quality of

[6]Local labor market conditions will be captured on June 30$^{\text{th}}$ of the graduation year if the apprenticeship is completed on or after that day and on June 30$^{\text{th}}$ of the year prior to graduation if the graduation happens earlier.

the initial match is important for early-career employment stability, adjustment processes and, in particular, involuntary unemployment.

We consider the instrument to be ignorably assigned because, following the reasoning in Gregg (2001), the choice of location at labor market entry can assumed to be exogenous. That is because individuals are on average 16.8 years old when they begin training. At that age most individuals still reside with their parents and do not have the means to move to another region. Indeed, we observe that 97.6 percent of individuals in our sample do not change districts during their apprenticeship.[7] Still, we cannot completely be sure that no initial sorting into districts by the individuals before the start of their apprenticeship takes place (in Germany it is for instance not uncommon for apprentices to live in dedicated boarding houses which might be located relatively far away from their original place of residence). Alternatively, sorting might occur even earlier by the individuals' parents. This is why we exploit the repeated cross-sectional design of our data set and control for geographical sorting by including fixed effects for the training firms' districts.

We argue that the instrument is excluded because time-varying patterns of economic conditions, the accumulation of skills and the dynamism of matching processes early in the professional career prevent it from influencing prime-age unemployment through channels other than youth unemployment. In any case, we will follow Gregg (2001) and control for local unemployment rates eight years after graduation. Thereby, we hope to capture any possible correlations between the instrument and the error term of equation 2.3.[8]

An additional concern might be that variation in the local unemployment rate right before graduation could influence not only involuntary but also voluntary unemployment. While it is well known that search and matching models (where excessive reservations wages lead to what arguably constitutes voluntary unemployment) predict the unemployment rate to vary over the business cycle, it is equally well-known that the predicted fluctuations are much smaller than those actually observed (this is the so-called "Shimer puzzle"). Yet, Wesselbaum (2010) shows how the introduction of efficiency wages — and thus involuntary unemployment as defined here — into a matching model with endogenous separations can plausibly increase the variability of workers' effort over the business cycle to such an extent that this inconsistency disappears. [cf. also Uhlig and Xu (1996)]. By implication, involuntary unemployment fluctuates more over the

[7] 35.7 percent of graduates do not stay at their training firm after graduation. Of these, 40 percent change districts between graduation and their first job subject to social security contributions. So the location of the first employment or unemployment spell has to be considered as endogenous. This is why our identification strategy relies on the local labor market conditions right before graduation.

[8] Our empirical approach does not allow us to use individual-specific fixed effects. However, as mentioned above and documented in Section 4, the early years of the professional career are often seen as providing opportunities for adjustments and for finding a productive employer-employee match. So in the context of the youth labor market controlling for unobserved time-invariant individual heterogeneity would be of little use [cf. von Wachter and Bender (2006)].

business cycle than voluntary unemployment and should be primarily affected by variation in the local unemployment rate.

Finally, a regression of prime-age unemployment on early-career unemployment poses a somewhat more technical challenge: as will be shown in Section 4, nearly 60 percent of the individuals in our sample are not unemployed for a single day during prime age. Thus we are faced with the typical case of *censoring* or rather a *corner solution outcome* as described by Wooldridge (2002). As a consequence, OLS or even simple IV estimates would be biased and inconsistent because of a correlation between the regressors and the error term.

Accordingly, we interpret m^*_{t2} as the latent amount of prime-age unemployment as opposed to the actually observed amount of prime-age unemployment, m_{t2}. It holds that

$$m_{i,t2} = \begin{cases} m^*_{i,t2} & \text{if } m^*_{i,t2} \geq 0 \text{ and} \\ 0 & \text{if } m^*_{i,t2} < 0. \end{cases} \tag{2.4}$$

To address the issue of a corner solution outcome in practice, we will supplement our OLS regressions by simple Tobit models in the tradition of Tobin (1958). Correspondingly, all estimations involving instruments will be done both with the standard IV estimator and Smith and Blundell's (1986) conditional maximum likelihood estimator for a Tobit model with continuous endogenous regressors.

3. DATA

We rely on matched employer-employee data created by the merger of two data sets: first, the Integrated Employment Biographies [IEB, cf. Oberschachtsiek, Scioch, Seysen and Heining (2009)] and, second, the Establishment History Panel (BHP). Both are administrative data sets provided by the Institute for Employment Research in Nuremberg, Germany.

The IEB contain the universe of all individuals who received unemployment benefits and/or were employed subject to social security contributions in the Federal Republic of Germany at least once between 1975 and 2008. Only spells of employment not covered by social security — like those of civil servants or family workers — and of self-employment are not in the data. All in all, the IEB cover about 80 percent of Germany's total workforce and encompass detailed longitudinal information on employment status, wages, socio-demographic and firm characteristics exact to the day. Because Germany's social security agencies use the underlying administrative data to compute both social security contributions and unemployment benefits, they are highly reliable. In the context of our study, another important advantage of not relying on survey but on administrative data is that we need not worry about panel mortality or non-response.

For the purpose of this study, the IEB are matched with establishment data from the BHP. For June 30th of any given year, the BHP encompasses all German establishments that on this day employ at least one worker subject to social security contributions. As described in Hethey-Maier and Seth (2010), variables contained in the data set include an establishment's sector and its geographic location. Information on the number of employees and their average wage is included, too. The different cross sections of the BHP can be combined to form a panel.

This study focuses on those individuals that start their professional career after graduating from Germany's dual education system. This system combines apprenticeships in a company and vocational education at a school in one course and is the way through which around 60 percent of young people enter the labor market. While Germany's system is often described as *the* model dual education system, similar regimes play an important role in many economies (e.g. in Austria, Switzerland or on the Balkans). In others countries, including the United States and the United Kingdom, there has long been a discussion about whether to strengthen the importance of education programs that combine vocational training in a company and learning at a school [cf. for instance Heckman (1993) or Neumark (2002)].

Access to Germany's dual education system is not formally linked to a specific school certificate; most individuals enter after grades nine or ten, a few after graduating from high school. The period of training is usually two to three years and the system is organized around more than 300 different occupations (ranging from doctor's assistants to opticians to oven builders). Limiting our sample to individuals going through the system implies that we can concentrate on a fairly homogeneous group of individuals that is at the same time central to the German labor market. Moreover, apprentices have to pay social security contributions and so periods in the dual education system are listed in our matched employer-employee data set. As a consequence, our data set contains detailed information related to individuals' socio-demographic characteristics, the type of their training and the nature of their training firm. This information is available for the time of individuals' graduation from the dual education system, i.e. right before their actual labor-market entry. Thus, we avoid (in large part) any problems that might be caused by unobserved initial conditions [cf. Hoffmann (2010)].[9]

[9] The institutional setup of Germany's dual education system is described in detail by Hippach-Schneider, Krause and Woll (2007). Concerning our specific sample, Table X in Appendix C shows individuals are on average a little less than 19 years old when they graduate from the dual education system. It should also be noted that in our sample the initial apprenticeship lasts on average 793 days while its median duration is 876 days. After graduation, 61 percent of graduates stay with their training firm. For those who do not stay there, the first employment subject to social security contributions is on average recorded 433 days after graduation. The time between graduation and the first job might not only encompass periods of unemployment and job search but also self-employment, military service or further education. Also, half of those individuals that do not stay with their training firm after graduation enter an employment relationship subject to social security contributions within 50 days and 70 percent take at most one year to do so.

This study's two key variables are *early-career unemployment* — defined as the total length in days of all unemployment spells of an individual in the eight years after finishing the first apprenticeship — and *prime-age unemployment*, the overall length of unemployment spells in the subsequent 16 years. While the latter is our dependent variable, the former is the key regressor. Section 4 will explain the rationale behind dividing the professional career into exactly these two intervals.[10]

About 90 percent of individuals registered as unemployed are eligible for unemployment relief or related benefits. Our data only contain information on individuals officially registered as job-seeking who do not receive any unemployment benefits from the year 2000 onwards; individuals who for some reason are not registered as unemployed but still willing to take up a job are not covered at all. That is why our benchmark definition of unemployment encompasses exactly those spells of unemployment that are associated with the receipt of benefits. In addition to that, in Section 5.2 we will test whether our main results are robust to an alternative definition of unemployment frequently found in the literature.

Using the receipt of unemployment benefits to define unemployment episodes has one important consequence: because regulations concerning unemployment benefits have somewhat varied during the last decades, it is difficult to compare the length of unemployment periods from different points in time. To circumvent this issue and to be sure that results are not driven by cohort effects, we restrict our analysis to three consecutive labor market entry cohorts. More precisely, we focus on those individuals that finished their first apprenticeship in 1978, 1979 or 1980.[11]

The following variables are included in the multivariate analysis of Section 5 as controls and also because assessing their effects on prime-age unemployment might be interesting in themselves (unless otherwise noted all variables are extracted from the last spell before graduation from the dual education system; for summary statistics cf. Table X in Appendix C):

Labor market entry cohort. Cohort dummies are meant to capture business cycle conditions at labor market entry or differences in size between labor market entry cohorts. They also control for longer-term trends, e.g. related to the quality of the German education system, that might influence on prime-age unemployment.

Graduation age. Graduation age might be a measure of time spend in education

[10] According to our data, 62 percent of the sample entered the labor market on December 31st. This seems unlikely and may be an artifact caused by employers that reported changes in employment status only at the end of the calendar year (which was legal during the late 1970ies). The actual time of graduation might therefore lie before the one we use. However, our main explanatory variable — the duration of early-career unemployment — is not affected by this issue because unemployment always induces a report by the social security agencies.

[11] Details on further data cleansing can be found in Appendix A. Because changes in regulations concerning unemployment benefits occurred during our sample frame for unemployment observations, they might still affect the observed pattern in unemployment over time. We have no reason to believe that this biases our results in a particular way and therefore disregard it.

or training that is not directly covered in our data set. Therefore, a negative relationship between this variable and prime-age unemployment might exist.

Daily remuneration. In Germany's dual education system, apprentices receive a remuneration from their training firm. Even though the rates of this remuneration are to a large extent regulated by collective bargaining agreements, a higher rate could still be a sign of high ability and thus be associated with lower prime-age unemployment. At the same time, it could lead to a higher reservation wage and ultimately to higher unemployment.

Occupation. Schmillen and Möller (2012) document long-term unemployment effects of the occupation pursued early in the professional career. We control for the initial occupation with dummy variables for nine occupation categories based on Blossfeld's (1987) classification: agricultural occupations (the reference category), unskilled manual occupations, skilled manual occupations, technicians and engineers, unskilled services occupations, skilled services occupations, semiprofessions and professions, unskilled commercial occupations, skilled commercial occupations and managers.

Sector of the training firm. Dummy variables for ten aggregated sectors are included: energy and mining, manufacturing, construction, trade, transport and communication, financial intermediation, other services, non-profits and households and public administration. The agricultural sector serves as the reference category.

Size and average wage of the training firm. Size is captured by the number of employees subject to social security contributions (in 1000) and the average wage is given by the median daily wage of this group. Both variables might be a signal for whether a firms' employees and apprentices have some bargaining power. Such bargaining power might among other things be associated with more productive training conditions. It might also mean that more apprentices stay at their training firm after graduation or return to it later.

Local unemployment at the transition from youth to prime-age. Following Gregg (2001), county-specific unemployment rates are used to capture local labor demand at the transition from youth to prime-age. In the benchmark regressions, the appropriate county is determined by the location of the last pre-transition employment spell.

4. Descriptive Evidence

As noted by Schmillen and Möller (2012), the empirical literature on unemployment almost exclusively focuses on the duration of distinct unemployment spells. In contrast, little is known about the longer-term distribution of unemployment and even less about the unemployment dynamics over the professional career. Against this backdrop this section will characterize the distributions of early-career, prime-age and lifetime unemployment followed by a description of the short- and long-run unemployment dynamics. The goals are to see whether there is evidence of an adjustment process during the first years on the labor

TABLE I

SUMMARY STATISTICS ON EARLY-CAREER, PRIME-AGE AND LIFETIME UNEMPLOYMENT

	lifetime unemployment	early-career unemployment	prime-age unemployment
mean	497	188	308
s.d.	900	334	701
min	0	0	0
max	8,754	2,922	5,844
p35	0	0	0
p40	32	0	0
p45	70	0	0
p50	118	15	0
p55	178	44	0
p60	251	78	28
p65	338	121	84
p70	439	175	162
p75	588	244	272
p80	760	331	406
p85	1,023	438	633
p90	1,460	615	990
p95	2,339	894	1,745

Notes: *Early-career unemployment* is the total length in days of all unemployment spells of an individual in the eight years after finishing the first apprenticeship while *prime-age unemployment* is the overall length of all unemployment spells in the subsequent 16 years. Early-career and prime-age unemployment sum to *lifetime unemployment* [cf. Schmillen and Möller (2012)].

market and to evaluate if unemployment is persistent over the professional career [which is often seen as a necessary condition for the existence of true state dependence, cf. Heckman and Borjas (1980)].

Table I provides summary statistics on early-career, prime-age and lifetime unemployment as defined in the last section. It shows that the average individual in our sample suffers from 188 days of unemployment during the first eight years of the professional career and from 308 days of unemployment over the subsequent 16 years. The mean amount of *lifetime unemployment* — defined as the sum of youth unemployment and prime-age unemployment, cf. Schmillen and Möller (2012) — is 497 days. Its distribution is highly skewed to the right: More than 35 percent of individuals in the sample are never registered as unemployed over the entire observation period. At the same time, 20 percent are registered as unemployed for at least 760 days and five percent for six and a half years or longer.

The distributions of early-career and prime-age unemployment are even more skewed to the right. The median of the distribution of early-career unemployment is 15 days, its 65$^{\text{th}}$ percentile four months and its 95$^{\text{th}}$ percentile 894 days. At the same time, almost 60 percent of the individuals in the sample experience no unemployment at all during prime age.[12]

The highly skewed distributions of early-career, prime-age and lifetime unemployment explain why we think that estimations of the conditional mean function might provide only an incomplete picture of the relationship between youth and prime-age unemployment. In particular, they might not be fully indicative of the size or nature of effects on the upper tail of the prime-age unemployment distribution. Furthermore, the underlying unemployment remain hidden. That is why Figure 1 shifts the attention to the short-run distribution of unemployment. Here, the goal is to determine whether the first years on the labor market can really be viewed as a time where job shopping enables individuals to offset disadvantageous initial conditions, gather heterogenous experiences and find their place in the professional world.

The figure displays the proportion of individuals in our sample that are not registered as unemployed during any given year of our observation period. It shows that throughout the professional career unemployment is concentrated on a comparatively small proportion of our sample (in some years more than 90 percent of individuals are not registered as unemployed at all). However, this concentration is much smaller during the first years on the labor market.

A similar picture emerges if one uses Gini coefficients of total annual unemployment for each year of our observation period to characterize the short-term unemployment inequality. This is done in Table 4. Its third column shows a Gini coefficient of 0.92 in the first year after graduation. Two years later, the coefficient drops to 0.89. It arrives at its minimum value of 0.87 when the individuals in our sample have been on the labor market for five years. Afterwards, the Gini coefficient rises again and reaches 0.93 in the tenth year on the labor market. From that point on, it stays more or less constant.[13]

Two mechanisms explain the Gini coefficients' trajectory: First, at every point in time a high amount of unemployment will tend to be distributed more evenly than a low volume. Second, for any given amount of unemployment, the distribution appears to become more and more uneven over the course of the professional career. The first mechanism would dominate the second if the Gini coefficients for years with an equal amount of overall unemployment were identical. Clearly, this is not the case: E.g., one may compare the Gini coefficients for the third

[12]Figure 4 in Appendix C contains a quantile-quantile plot that plots the probability distributions of early career and prime-age unemployment against each other. The figure shows that comparatively small proportions of unemployment during the early career are plotted against even shorter proportions of prime-age unemployment. At the same time, unemployment proportions higher than 40 percent of the early career — as experienced by less than five percent of the sample — are plotted against even higher proportions of unemployment later in life. This confirms that the distribution of prime-age unemployment is even more skewed to the right than that of early-career unemployment.

[13]The Gini coefficients for lifetime, early-career and prime-age unemployment are 0.74, 0.75 and 0.82, respectively. This confirms Bönke, Corneo and Lüthen's (2011) result that annual measures of inequality overestimate inequality as compared to measures based on a lifetime perspective [cf. also Hoffmann (2010)].

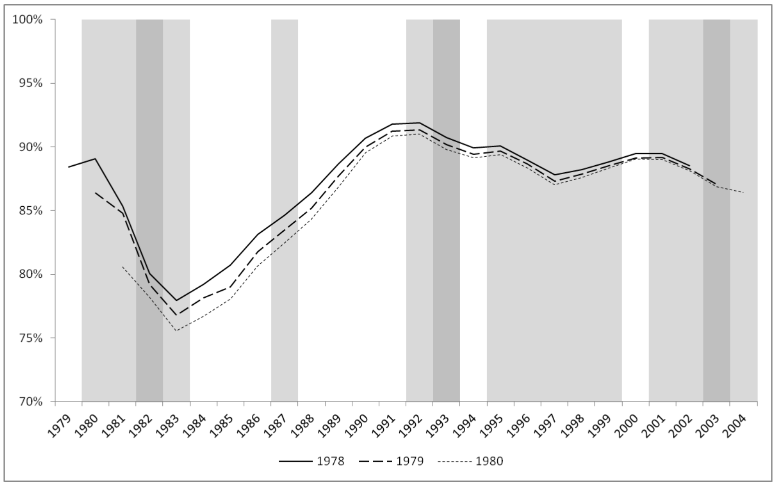

FIGURE 1.— Proportion of individuals not registered as unemployed by labor market entry cohort and year.

Notes: *1978*, *1979* and *1980* give the year of labor market entry. Dark-shaded areas denote years with negative GDP growth and grey-shaded areas those with positive GDP growth not exceeding 2 %.

and the 18$^{\text{th}}$ year on the labor market, two years with a roughly equal amount of overall unemployment (given by column 2 of Table 4).

So, again, our conclusion would be that unemployment appears to be quite unevenly distributed but much less so during the first years of the professional career.

Table 4 also shows that — at least in the short run — mobility in the distribution of annual unemployment is pretty low. Its fourth column displays the values from Spearman's rank correlation coefficients between the unemployment distributions of subsequent years (where a higher value indicates a higher immobility in the distribution). As the table also makes clear, the mobility in the distribution of annual unemployment becomes smaller and smaller over the course of the professional career; the correlation coefficients increase from an already high value of 0.37 in the first year to 0.7 in the 23$^{\text{rd}}$ year on the labor market.

While this section has so far been concerned with unemployment, Topel and Ward (1992) and others who see the first years on the labor market as an adjustment period usually focus on the related but not identical phenomenon of job mobility. That is why Figure 2 plots annual job mobility rates. These are defined as the ratio of individuals who experience at least one change of employer to the total number of individuals who are employed for at least one day in any particular year. The figure distinguishes between two forms of job mobility: direct and indirect changes of employer. Direct changes of employer are defined as changes with an interruption of employment of less than three weeks. If the interruption lasts longer and the worker is not recalled by his/her former employer, then it is counted as an indirect change. Such indirect changes of employer are especially pronounced in the early years of the professional career. In the first employment year, the average rate of such changes rises steeply to 38 percent and falls continuously to merely ten percent ten years later. From there on it remains almost constant. In contrast, the rate of direct changes of employer does not appear to be particularly high in the early years of the professional career.[14]

Table III shifts the attention to the long-term dynamics of unemployment. It lists the transition probabilities between certain positions in the distributions of early-career and prime-age unemployment. The table divides these distributions into cells of equal size (five percent of our sample) as well as one larger cell that mostly contains individuals with no unemployment at all in the respective period.

If an individual's youth and prime-age unemployment were independent, one would expect roughly five percent of individuals from each early-career unemployment cell to transition into every prime-age unemployment cell (apart from the larger cells containing those with zero unemployment). Table III demonstrates that this is not what is actually happening. In contrast, almost all transition probabilities are statistically significantly different from five percent. In the table's lower left corner they stay below five percent but are all much larger in the lower right corner. Strikingly, the probability for individuals whose amount of early-career unemployment exceeds the 95$^{\text{th}}$ percentile to belong to the five percent of individuals with the highest amount of prime-age unemployment is almost a third. And hardly anybody from this group suffers from no prime-age unemployment at all.

The general picture that emerges from Table III is that unemployment is indeed persistent over the whole professional career. High youth unemployment almost constitutes a necessary condition for having a very elevated amount of prime-age unemployment. In contrast, those who experience no unemployment during the first years of their professional career often exhibit no prime-age unemployment either (even though there are a few individuals who manage to transition from a youth characterized by high unemployment to relatively low unemployment levels later in their career or who experience no early-career but

[14]Unsurprisingly, direct changes of employer are more pronounced in years with favorable economic conditions, as indicated by the areas in Figure 2 that are not shaded in grey or black. The opposite is true for indirect changes. It should also be noted that over the entire observation period, 13 percent of individuals continually stay with their initial employer. About 79 percent experience at least one direct change of employer.

TABLE II

Inequality and immobility in the distribution of annual unemployment

	unemployment		
year on labor market	total sum (m days)	inequality (Gini coef.)	immobility (Spearman's ρ)
1	8.1	0.9211	0.3731
2	11.7	0.9152	0.4385
3	18.5	0.8904	0.5002
4	24.0	0.8731	0.5605
5	26.2	0.8729	0.5947
6	24.7	0.8818	0.5980
7	22.2	0.8939	0.6101
8	20.1	0.9056	0.6154
9	17.8	0.9171	0.6062
10	14.8	0.9314	0.5882
11	12.2	0.9431	0.5789
12	10.8	0.9496	0.5739
13	11.4	0.9483	0.6047
14	13.2	0.9433	0.6258
15	14.7	0.9389	0.6436
16	15.9	0.9351	0.6574
17	17.2	0.9303	0.6773
18	18.5	0.9262	0.6932
19	18.5	0.9263	0.6964
20	17.3	0.9308	0.6864
21	16.4	0.9338	0.6815
22	16.7	0.9327	0.6868
23	18.5	0.9268	0.7049
24	20.7	0.9195	—

Notes: *Year on labor market* indicates the number of years since labor market entry. For every year, *total sum (in million days)* adds up the days of registered unemployment over all individuals in the sample. *Inequality* reports Gini coefficients of total annual unemployment. These Gini coefficients include all zeros and are computed with the Stata command *ineqdec0*. *Immobility* reports Spearman's ρ as a measure of the rank correlation between the distributions of total annual unemployment between consecutive years on the labor market.

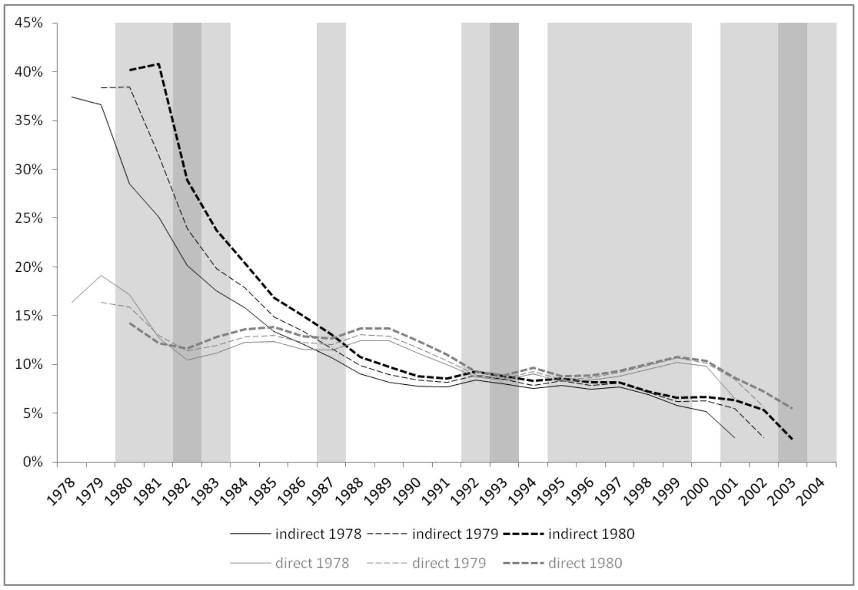

FIGURE 2.— Job mobility rates

Notes: Mobility rates are defined as the ratio of individuals who experience at least one change of employer to the total number of individuals who are employed for at least one day in any particular year. *Direct* changes of employer are defined as changes with an interruption of employment of less than three weeks. If the interruption lasts longer and the worker is not recalled by his/her former employer, then it is counted as an *indirect* change. Dark-shaded areas denote years with negative GDP growth and grey-shaded areas those with positive GDP growth not exceeding 2 %.

a high amount of prime-age unemployment).[15], [16]

To sum up our interpretation of the descriptive evidence: unemployment tends to be very persistent over the professional career. At the same time, there is at least some evidence in favor of the view that periods of unemployment during the first years on the labor market are part of an adjustment process (judging from

[15]A rank correlation of 0.4 between early-career and prime-age unemployment makes it clear that the long-term mobility in the distributions of unemployment is higher than the short-run mobilities reported above. This results is partly mechanical. But it could also be viewed as evidence in favor of the hypothesis that early-career unemployment is — at least to a certain extent — an expression of early job-mobility and does not necessarily have to be damaging in the long-run.

[16]Table IX in Appendix C shows that the incidence of unemployment falls over the course of the professional career. At the same time, the table demonstrates that the mean of total unemployment generated within each year increases with early-career unemployment as well as over time. So with the proportion of people experiencing unemployment declining, a shrinking group of individuals seems to experience more and/or longer spells of unemployment. This is evidence against (time-invariant) heterogeneity as the only link between early and subsequent unemployment but perfectly in line with true state dependence.

TABLE III

Transition probabilities between certain positions in the distributions of early-career and prime-age unemployment (in percent).

	early-career unemployment										
prime-age unemployment	*p51 to p55*	*p56 to p60*	*p61 to p65*	*p66 to p70*	*p71 to p75*	*p76 to p80*	*p81 to p85*	*p86 to p90*	*p91 to p95*	*p96 to 1*	*0 to p50*
p61 to p65	5.8***	6.4***	6.2***	6.2***	6.4***	6.4***	6.4***	5.8***	5.0	3.2	42.2
p66 to p70	5.2	6.0***	6.2***	6.6***	6.8***	6.8***	7.0***	6.8***	6.2***	4.4	38.0
p71 to p75	5.0	5.4	6.0***	6.8***	6.8***	6.6***	7.0***	8.6***	6.4***	5.2***	36.2
p76 to p80	5.0	5.4	5.6***	6.2***	6.4***	6.8***	7.4***	7.8***	7.0***	5.8***	36.6
p81 to p85	4.6	4.8	5.8***	6.6***	6.8***	7.4***	8.2***	8.8***	8.2***	8.2***	30.6
p86 to p90	4.2***	4.6	5.6	5.8***	6.2***	7.4***	8.0***	9.6***	9.4***	10.4***	28.8
p91 to p95	3.4***	3.8***	4.8	5.4	6.0***	7.2***	8.6***	11.0***	11.8***	15.8***	22.2
p96 to 1	2.2***	2.6***	3.2***	4.0***	4.8	5.8***	7.2***	10.0***	13.4***	32.2***	14.6
0 to p60	64.6	61.0	56.6	52.4	49.8	45.6	40.2	31.6	32.6	14.8	37.1

Notes: For the definitions of early-career and prime age unemployment, cf. the notes to Table I. *** indicates significance at the 1 % level as indicated by Pearson's chi-squared tests with the null hypothesis of independence between early-career and prime-age unemployment. The hypothesis that all rows and columns in the table are independent is rejected with an overall $\chi^2(361) = 2.7exp^6$.

Figure 2 this adjustment process takes approximately eight years which is the reason behind our cut-off of the early career and prime age). As forcefully argued by Heckman and Borjas (1980), in the end only a multivariate analysis that takes into account the potential endogeneity of youth unemployment can tell whether the observed unemployment persistence is due to true state dependence. This is the aim of the next sections.

5. RESULTS

5.1. *Baseline Estimates*

Table IV summarizes the outputs of nine different estimates of the conditional expectation function of prime-age unemployment. Even though the focus is on the question whether unemployment exhibits true state dependence, coefficients for the most interesting control variables are also displayed. Besides, for all IV regressions the table contains the instrument's coefficients and first stage F-statistics. Throughout, standard errors are clustered at the district level.

As a starting point, in column (1) prime-age unemployment is regressed on early-career unemployment and a constant. The resulting regression suggests that on average every additional day of early-career unemployment is associated with 0.93 more days of prime-age unemployment and that this relationship is statistically significant. The picture remains practically unchanged if one controls for the full set of observable characteristics described in Section 3, as is done in column (2). The same is true if location-specific fixed effects are also included

TABLE IV

Different Estimates of Prime-age Unemployment — Baseline Regressions.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Model	OLS	OLS	OLS	Tobit	Tobit	IV	IV	Tobit-IV	Tobit-IV
Regressions of prime-age unemployment									
Early-career unemployment	0.93***	0.89***	0.89***	0.57***	0.57***	1.91***	2.62***	1.29***	1.98***
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.26)	(0.33)	(0.15)	(0.20)
Age	—	-1.64*	-5.40***	-5.12***	-7.78***	-3.61**	-6.55***	-6.46***	-8.79***
		(0.97)	(0.75)	(0.78)	(0.53)	(1.47)	(0.98)	(1.02)	(0.70)
Remuneration	—	-2.70***	-2.20***	-2.47***	-2.02***	-0.18	1.27*	-0.63	0.79*
		(0.23)	(0.21)	(0.20)	(0.17)	(0.71)	(0.71)	(0.50)	(0.48)
Size of training firm	—	-0.42	-1.44**	-3.70***	-4.36***	1.89**	1.87**	-2.02***	-1.69**
		(0.34)	(0.35)	(0.45)	(0.60)	(0.98)	(0.94)	(0.67)	(0.70)
Median wage of training firm	—	-0.25	-1.38***	-1.26***	-1.95***	1.49***	0.65	0.01	-0.32
		(0.21)	(0.19)	(0.17)	(0.14)	(1.49)	(0.42)	(0.27)	(0.30)
Occupation (reference category: agricultural occupations)									
Unskilled manual occ.	—	49.09**	43.89**	10.63	4.51	50.03***	58.43***	11.33	16.37
		(21.81)	(21.72)	(14.38)	(14.14)	(17.94)	(19.70)	(11.82)	(14.02)
Skilled manual occ.	—	-78.77***	-75.14***	-84.24***	-83.71***	-21.62	35.14	-42.12***	5.36
		(20.22)	(20.13)	(13.05)	(13.00)	(22.95)	(19.91)	(16.12)	(21.27)
Technicians and engineers	—	-122.48***	-118.08***	-132.46***	-131.79***	-27.01	59.67	-62.34***	11.84
		(21.59)	(20.71)	(14.42)	(13.73)	(32.38)	(40.89)	(23.56)	(28.87)
Unskilled services	—	71.04***	53.24**	27.73*	11.69	39.04*	14.12	4.49	-20.03
		(23.12)	(22.64)	(15.11)	(14.51)	(21.23)	(20.46)	(13.83)	(14.38)
Skilled services	—	-60.17**	-68.05***	-78.10***	-86.48***	-12.59	23.79	-43.01***	-12.45
		(26.11)	(26.15)	(15.96)	(15.95)	(23.81)	(30.13)	(15.73)	(21.07)
Semiprofessions	—	-122.30***	-116.45***	-148.77***	-147.58***	-8.12	89.54*	-64.97**	18.88
and professions		(24.22)	(25.11)	(15.36)	(16.60)	(39.42)	(48.15)	(26.38)	(33.61)
Unskilled commercial occ.	—	11.67	-4.53	-25.45*	-39.04***	107.24***	166.31***	43.33**	99.68***
		(20.16)	(20.27)	(13.45)	(13.17)	(30.57)	(39.78)	(20.57)	(26.56)
Skilled commercial occ.	—	-93.51***	-87.04***	-130.93***	-128.97***	27.48	133.80***	-42.69*	49.72
and managers		(20.66)	(20.25)	(13.68)	(13.38)	(36.77)	(48.24)	(25.75)	(33.48)
Regressions of early-career unemployment									
Unemployment at graduation	—	—	—	—	—	18.27***	27.20***	18.27***	27.20***
						(2.57)	(5.59)	(2.57)	(5.59)
Other variables included in regressions									
District dummies	No	No	Yes	No	Yes	No	Yes	No	Yes
Cohort dummies	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector dummies	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Unemployment at transition	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First stage F-statistics	—	—	—	—	—	50.41***	23.96***	—	—
Number of observations	827,114	739,432	739,432	739,432	739,432	739,432	739,432	739,432	739,432

Notes: Standard errors clustered at the district level in parentheses. * , (**), [***] indicates significance at the 10, (5), [1] % level. IV regressions are performed with Hansen, Heaton and Yaron's (1996) continuously-updated GMM estimator implemented in the Stata command *ivreg2* by Baum, Schaffer and Stillman (2003, 2007); Tobit-IV regressions are calculated with Smith and Blundell's (1986) conditional maximum likelihood estimator. In both cases the instrument is the local unemployment rate at graduation. Tobit and Tobit-IV models report the average marginal effects on the observed amount of prime-age unemployment; for all factor variables the discrete first differences from the base categories are calculated. The delta method is used to compute standard errors. Apart from the instrument, variables included in the regressions of early-career unemployment are the same as in the estimates of prime-age unemployment. For variable definitions see Section 3.

[cf. column (3)]. This implies that initial sorting of individuals into their training firms' labor market districts hardly seems to bias those estimates that do not account for it.

As discussed above, nearly 60 percent of the individuals in our sample are not unemployed for a single day during their prime age. Thus, we are faced with the typical case of a corner solution outcome. To address this issue, the OLS regressions are supplemented by simple Tobit models in the tradition of Tobin (1958). In columns (4) and (5), results are yet again shown both with and without dummy variables for the training firms' labor market districts. Importantly, these columns do not directly contain the Tobit models' coefficients. These coefficients measure how the latent amount of prime-age unemployment m^*_{t2} changes with respect to changes in the regressors. But in the context of a corner solution model, we do not really care about the latent dependent variable. Instead, the marginal effects on the observed amount of prime-age unemployment m_{t2} appear much more relevant. They are therefore displayed in table IV [cf. Wooldridge (2002)]. Since these marginal effects depend on the explanatory variables' values, one must decide at what values to report them. As is commonly done in the literature, the table shows the average marginal effects. For factor variables discrete first differences from the base categories are calculated; the delta method is used to compute standard errors.[17]

A comparison of column (2) and column (4) — neither of which incorporates dummy variables for the training firms' labor market districts — shows that the Tobit specification exhibits a somewhat lower marginal effect of early-career unemployment than the OLS regression. This result is practically unchanged by the inclusion of location-specific fixed effects [cf. columns (3) and (5)]. In both Tobit regressions, the average marginal effect is around 0.57 days.

Results from both the OLS and the Tobit models discussed so far should probably be interpreted as a confirmation of the descriptive evidence presented in Section 4. They demonstrate that unemployment is quite persistent over the professional career but say little about whether true state dependence exists between (involuntary) early-career and prime-age unemployment. That is the purpose of the regressions summarized in columns (6) to (9). These regressions instrument (involuntary) early-career unemployment with the local unemployment rate prevailing at the training firm's location right before graduation.

For all instrumental variable specifications, F-statistics against the null that the excluded instrument is irrelevant are statistically significant. More importantly, they are quite a bit higher than ten. Following Staiger and Stock (1997)

[17]Additionally, Table X in Appendix C reports the marginal effects on the latent amount of prime-age unemployment (i.e. the model's coefficients). Table X also summarizes the marginal effects on the observed amount of prime-age unemployment if all explanatory variables take on their average value and — as recommended by Wooldridge (2002) — the average marginal effects on the observed amount of prime-age unemployment among the subpopulation for which prime-age unemployment is not at a boundary. Qualitatively, the different marginal effects are all very similar.

and Stock, Wright and Yogo (2002), in the presence of one endogenous variable this value is commonly used as rule of thumb in order to decide if an instrument appears strong. So we feel confident that we do not have to worry about weak instrument problems. In any case, we will return to the question whether the local unemployment rate at graduation is a strong instrument in section 5.2. Then we will also test whether early-career unemployment should actually be treated as endogenous.

The specifications reported in columns (6) and (7) of table IV are very similar to the ones shown in columns (2) and (3) but for the instrumentation of (involuntary) early-career unemployment. Instead of using the canonical two-stage least-square estimator, the IV regressions rely on Hansen, Heaton and Yaron's (1996) continuously-updated GMM procedure. This is generalization of the limited-information maximum likelihood estimator to the case of possibly heteroskedastic and autocorrelated disturbances. It has the advantage that all statistics are not only robust to heteroscedasticity and clustering at the district level but also efficient.

If one compares the output summarized in column (6) with that of column (2), one notices that the coefficient associated with early-career unemployment remains statistically significant. In fact, it is higher in the IV than in the OLS regression. Consistent with findings by Gregg (2001), Neumark (2002) and Gregg and Tominey (2005), a simple OLS regression apparently understates the scarring effect of early-career unemployment. At first glance, this might seem surprising. One might intuitively assume that omitting information on unobserved individual characteristics — e.g. an individual's ability or motivation — would upwardly bias simple OLS estimates. However, as discussed above, there might be good reasons for why they are in fact downward-biased. In particular, the presence of voluntary unemployment might contribute to a downward bias in OLS estimates of the scarring effects of involuntary unemployment. Moreover, it appears reasonable to assume that the local unemployment rate at graduation used in the IV regressions' first stages has a stronger effect on involuntary than on voluntary unemployment. The returns to job hopping and/or the returns to search might also be considerably higher for individuals with high ability [cf. Neumark (2002)].

Columns (7), (8) and (9) again add controls for initial sorting of individuals by including fixed effects for the training firms' districts and/or use Smith and Blundell's (1986) conditional maximum likelihood estimator for a Tobit model with continuous endogenous regressors to take account of the corner solution outcome. Again, the Tobit models report the average marginal effects on the observed amount of prime-age unemployment and again, for all factor variables the discrete first differences from the base categories are calculated and standard errors are computed with the help of the delta method.

For all IV specifications, the estimated average amount of prime-age unemployment that is induced by an additional day of early-career unemployment rises as compared to the regressions that regard early-career unemployment as exogenous. The marginal effects associated with this variable are 2.62 days when we include district dummies in the IV regressions, 1.29 days when we take account of the corner solution outcome and 1.98 days when we do both.

Ultimately, the regression reported in column (9) of Table IV takes all the various sources of bias discussed in Section 2 into consideration. Thus, it represents our preferred specification and we conclude that early-career unemployment in fact causes future unemployment. With on average one day of early-career unemployment leading to two days of joblessness during prime age, this scarring effect is not only statistically significant but also economically important. Besides, because prime age is by our definition twice as long as the early phase of the professional career, a marginal effect of two hints at an elasticity of prime-age unemployment with regard to early-career unemployment of almost exactly one.

Before discussing the scarring effect of unemployment in greater detail, we will now briefly shift the attention to some of the more interesting control variables. Generally speaking, many of them exhibit statistically and economically significant coefficients (of course these should not be interpreted as a revelation of causal relationships). This confirms the existence of strong correlation between initial conditions and later labor market outcomes. Moreover, while for many control variables the size of their coefficients and sometimes also their levels of statistical significance vary quite a bit between the different specifications summarized in Table IV, most signs consistently stay the same.

Focusing on column (9) of Table IV, we see that having a higher graduation age is associated with less prime-age unemployment, *ceteris paribus*. The variable measuring the size of the training firm also has the expected (negative) sign while the firm's average wage level is not significantly related to prime-age unemployment. The coefficient associated with the remuneration earned at graduation is not statistically significant either, at least not on a level that appears appropriate for the large data set we use. Lastly, even though most of the specifications summarized in Table IV document a strong link between the occupation pursued early in the professional career and the amount of unemployment that an individual experiences later, this is not really the case in column (9). Here, many occupation dummies are in fact not statistically significant.

5.2. *Sensitivity and Specification Tests*

We will now report the outcomes of sensitivity checks that evaluate whether our central result — namely the long-run scarring effect of early-career unemployment — is robust to variations of the empirical setup. Results for a number of such checks are reported in Table V which focuses exclusively on the main variables of interest. Reference point is the regression reported in column (9) of Table IV, that is, the conditional maximum likelihood Tobit-IV estimation that includes district fixed effects and instruments (involuntary) early-career unemployment with the local unemployment rate right before graduation. Reported

TABLE V

Different Estimates of Prime-age Unemployment — Robustness Regressions.

	(1)	(2)	(3)	(4)	(5)	(6)
Specification	Unempl. in origin at transition as control	Minimum unempl. in prime age as control	At least one observation during last four years	Less than six years of seasonal empl.	Nonempl. I instead of unempl.	Nonempl. II instead of unempl.
Model	Tobit-IV	Tobit-IV	Tobit-IV	Tobit-IV	Tobit-IV	Tobit-IV
Regressions of prime-age unemployment [prime-age non-employment in (5) and (6)]						
Early-career unemployment	2.00***	1.91***	2.15***	1.90***	—	—
	(0.20)	(0.17)	(0.22)	(0.21)		
Early-career non-employment	—	—	—	—	1.61***	1.20***
					(0.13)	(0.11)
Regressions of early-career unemployment [early-career non-employment in (5) and (6)]						
Unemployment at graduation	27.54***	29.06***	27.09***	23.37***	60.09***	41.04***
	(5.55)	(5.20)	(5.53)	(4.86)	(10.85)	(7.94)
Other variables included in regressions						
District dummies	Yes	Yes	Yes	Yes	Yes	Yes
Unempl. at transition (current)	No	Yes	Yes	Yes	Yes	Yes
Unempl. at transition (origin)	Yes	No	No	No	No	No
Minimum unempl. in early career	No	Yes	No	No	No	No
Number of observations	740,394	739,432	648,644	652,206	739,432	739,432

Notes: Standard errors clustered at the district level in parentheses. * , (**), [***] indicates significance at the 10, (5), [1] % level. All regressions are calculated with Smith and Blundell's (1986) conditional maximum likelihood Tobit-IV estimator and report the average marginal effects on the observed amount of prime-age unemployment [prime-age non-employment in (5) and (6)]. The delta method is used to compute standard errors. The instrument is the local unemployment rate at graduation. Unless otherwise noted, covariates are the same as in column (9) of Table IV. In (1) the local unemployment at the transition from youth to prime-age for the district of the last apprenticeship spell is used as a control variable; in (2) the minimum local unemployment rate during the early career is used as a control variable; in (3) individuals who are not observed during the last four years of their prime age are excluded; in (4) individuals with more than five years of seasonal employment as defined by Del Bono and Weber (2008) are excluded; in (5) early-career and prime-age non-employment modeled on the definition by Fitzenberger and Wilke (2010) are used instead of early-career and prime-age unemployment; in (6) early-career and prime-age non-employment are modeled on an alternative definition by Schmieder, von Wachter and Bender (2012). Apart from the instrument, variables included in the regressions of early-career unemployment [early-career non-employment in (5)] are the same as in the estimates of prime-age unemployment [prime-age non-employment in (5) and (6)]. For variable definitions see Section 3.

are the average marginal effects on the observed amount of prime-age unemployment.[18]

So far, we have used county-specific unemployment rates to capture local labor demand at the transition from youth to prime age where the appropriate county has been determined by the location of the last pre-transition employment spell. However, one might wonder whether individuals' geographical mobility during the first years of the professional career should not be viewed as endogenous. In particular, one might expect individuals with (unobserved) beneficial characteristics to be more likely to end up in a labor market district with a comparatively

[18]Table XI in Appendix C summarizes the corresponding Tobit regressions without the use of instrumental variables.

low unemployment rate eight years after their labor market entry. That is why in column (1) of Table V we continue to control for county-specific unemployment rates but do not use the individuals' location at the transition to prime age. Instead, we use the unemployment rate that prevailed at that point in time in their county of origin, i.e. the county where their last apprenticeship spell was recorded. As discussed above, conditional on the district fixed effects we regard this location as exogenous. Table IV shows that controlling for the unemployment rate at the district of origin does not significantly change the coefficient associated with early-career unemployment.

Column (2) of Table V controls for yet another unemployment rate faced by the individuals in our sample. Beaudry and diNardo (1991) show that in a model with implicit labor market contracts and moderately costly mobility, the lowest unemployment rate since the beginning of a job influences the current wage, even if one controls for the current unemployment rate. Moreover, they present empirical evidence that confirms their model's prediction: once they include the lowest unemployment rate since the beginning of a job, neither the unemployment rate at the start of the job nor the current rate is significantly associated with wages.[19] Based on Beaudry and diNardo's (1991) work and loosely following Neumark (2002), column (2) of Table IV includes the minimum unemployment rate an individual faces during the first eight years on the labor market. Again, this does not alter the coefficient associated with early-career unemployment by much.

Next, we shift the attention to the issue of sample attrition. As argued above, one of the many advantages of not relying on survey but administrative data is that one need not worry too much about panel mortality or non-responses. In fact, Figure 5 in Appendix C shows that the annual sample attrition rate — i.e. the rate of individuals that disappear from the observable part of the German labor market — is almost constant over time and consistently lower than two percent. Still, it might be the case that our estimates of the scarring effects of early-career unemployment are biased because individuals with a high amount of youth unemployment are more or less likely to exit the part of the German labor market covered by our data set (potentially in order to become civil servants, self employed or inactive). In column (3) of Table V all individuals who are not observed during the last four years of their prime age are excluded from the regression. Though this reduces the estimation sample by about 15 percent, the coefficient measuring the scarring effect of early-career unemployment is not significantly altered.

This is not the case either if those individuals that have more than five years of seasonal employment during the first 24 years of their professional career are excluded [cf. column (4) of Table V]. This exclusion of seasonal workers is meant to make sure that our results are not purely driven by men who "only" have a

[19]Recently, Hagedorn and Manovskii (forthcoming) have argued that variables summarizing past aggregate labor market conditions lose any predictive power for current wages once match qualities are accounted for.

very elevated amount of unemployment because they are seasonally employed during a large portion of their professional career. In order to identify seasonal employment we draw on Del Bono and Weber (2008) and label two or more employment spells that last for at least two but less then eleven months and end at about the same time in consecutive calender years a seasonal job. Also following Del Bono and Weber (2008), we allow for a three-month window at the end dates of a spell.[20]

Additionally, we evaluate if altering the measure for unemployment durations changes our results. In particular, we make use of two alternative definitions that use the length of nonemployment spells as a measure for unemployment durations. The first definition (*non-employment I*) relies on Fitzenberger and Wilke (2010). Here, all time periods not recorded as employment that follow an employment spell and contain at least one spell of receiving unemployment benefits are counted as non-employment. The second definition of non-employment (*non-employment II*) is based on Schmieder, von Wachter and Bender (2012). Schmieder, von Wachter and Bender (2012) measure nonemployment as the time between the start of receiving unemployment benefits and the date of the next registered employment spell, where all nonemployment durations are capped at 36 months. Modeled on early-career and prime-age unemployment, early-career and prime-age non-employment are given by the total length in days of all non-employment spells of an individual in the eight years after finishing the first apprenticeship and the subsequent 16 years, respectively. As columns (5) and (6) of Table V demonstrate, the scarring effect of youth non-employment [and especially of non-employment as defined by Schmieder, von Wachter and Bender (2012)] appears somewhat smaller than that of youth unemployment. Qualitatively, however, results stay the same.

As a further check as to whether our finding of a long-run scarring effect of early-career unemployment is robust to variations of the empirical setup, we consider a second instrument, a dummy variable for whether an individual's training firm closes in the year of his graduation from the dual education system. This dummy variable not only represents a second source of exogenous variation. It also allows us to exploit a different form of such variation, namely establishment-level variation instead of variation on the level of the local labor market.

We consider the dummy variable for whether an individual's training firm closes in the year of his graduation to be a relevant instrument because it determines whether a graduate is forced to search for a job outside his training firm upon labor market entry (recall that nearly 60 percent of individuals in our sample stay with their training firm after graduating from the dual education system). Besides, it is ignorably assigned: Hamermesh (1987) demonstrates

[20] Table XIII in Appendix C shows that with this definition of seasonal employment about half of the individuals in our sample are seasonally employed for at least two years during their early career or their prime age. A much smaller proportion of individuals — around 15 percent of the sample — is seasonally employed for at least five out of twenty-four years while less than three percent of sampled men are seasonally employed for ten years or more.

that plant closures tend to surprise the workers who are affected. As compared to those already in employment, individuals who start their apprenticeship are even less likely to have the necessary information to correctly forecast the likelihood of their training firm closing down a few years later. Changing the training firm during the course of an apprenticeship is also rather difficult (both for practical reasons and because of the rather restrictive paragraph 22 of Germany's Vocational Training Act of 1969). Lastly, the instrument is excluded because — as with the local unemployment rate at graduation — economic conditions that change over time, the accumulation of human capital and matching processes early in the professional career should prevent it from influencing prime-age unemployment through channels other than youth unemployment.[21]

Results for regressions that use the closure of a graduate's training firm as instrument for (involuntary) early-career unemployment are reported in Table VI. The table summarizes the outputs of seven regressions. These differ along the following four dimensions: First, while in columns (1) and (5) a dummy variably for establishment death is the only instrument, in the other columns both this variable and the local unemployment rate at graduation are jointly used. The resulting models are overidentified, which allows us to perform a number of specification tests. Second, because many of the most common tests are only available for linear IV but not for Tobit-IV, the table contains estimates obtained with the help of both models. Columns (1) to (4) relate to IV regression; Tobit-IV is used in columns (5) to (7). Once again, the table only displays results for the main variables of interest; for the Tobit-IV estimates (which continue to represent our preferred specification) the average marginal effects on the observed amount of prime-age unemployment are reported.

Third, all regressions but the one reported in column (4) of Table VI control for district dummies. In column (4) a linear IV regression instead contains dummy variables for individuals' training firms. The aim is to capture initial sorting into these firms. Estimating Tobit-IV regressions with training firm fixed effects appears computationally unfeasible. Fourth, establishment fixed effects only make sense if an establishment's size surpasses a certain threshold. We follow von Wachter and Bender (2006) and only consider individuals graduating from training firms with at least 50 employees subject to social security contributions and five graduating apprentices in a given year in the respective regression [column (4)]. In order to ensure that the resulting outcomes are not driven by the selection of this sub-sample, columns (3) and (7) contain regressions for the smaller sample that include the usual district fixed effects.

[21] Hethey and Schmieder (2010) note that restructuring and relabeling of firms is often poorly measured in administrative data sets and that this can potentially create large biases. Using worker flows between German establishments they credibly identify establishment births and deaths in the BHP. We rely on their classification of establishment closures and assign a value of one to all establishments that experienced either a "small death", an "atomized death" or a "chunky death" [according to Hethey and Schmieder's (2010) classification] in an individual's year of labor market entry.

TABLE VI

DIFFERENT ESTIMATES OF PRIME-AGE UNEMPLOYMENT — IV REGRESSIONS.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Model	IV	IV	IV	IV	Tobit-IV	Tobit-IV	Tobit-IV
Regressions of prime-age unemployment							
Early-career unemployment	1.29***	2.17***	1.98***	1.91***	0.87***	1.67***	1.59***
	(0.15)	(0.19)	(0.20)	(0.19)	(0.10)	(0.13)	(0.14)
Regressions of early-career unemployment							
Unemployment at graduation	—	26.94***	29.47***	25.85***	—	29.40***	30.07***
		(5.54)	(5.07)	(2.23)		(5.03)	(4.94)
Establishment closure	65.73***	56.81***	81.93***	46.36*	55.73***	42.93***	70.24***
	(6.03)	(5.86)	(23.81)	(28.09)	(6.03)	(5.82)	(5.82)
Other variables included in regressions							
District dummies	Yes	Yes	Yes	No	Yes	Yes	Yes
Establishment dummies	No	No	No	Yes	No	No	No
Number of observations	809,498	739,158	298,471	298,471	809,498	739,158	298,471
Difference-in-Sargan exogeneity test	7.40***	28.61***	24.35***	—	—	—	—
Smith-Blundell exogeneity test	—	—	—	—	10.08***	10.08***	40.76***
First stage F-statistics	104.66***	70.23***	23.23***	—	—	—	—
Hansen J statistic	—	11.65***	0.47	—	—	—	—
Anderson-Rubin test	78.39***	111.15***	40.89***	—	67.03***	622.05***	202.78***
Conditional likelihood-ratio test	—	108.94***	39.52***	—	—	582.42***	201.14***
Lagrange multiplier test	—	105.97***	35.72***	—	—	502.93***	198.36***
J overidentification test	—	5.19**	5.17**	—	—	119.12***	4.42**
Sample							
All Establishments	√	√			√	√	
Large establishments only			√	√			√

Notes: Standard errors clustered at the district level in parentheses. * , (**), [***] indicates significance at the 10, (5), [1] % level. "Large establishments only" means that the sample only contains individuals graduating from training firms with at least 50 employees subject to social security contributions and five graduating apprentices in a given year. IV regressions are performed with Hansen, Heaton and Yaron's (1996) continuously-updated GMM estimator implemented in the Stata command *ivreg2* by Baum, Schaffer and Stillman (2003, 2007). Tobit-IV regressions are calculated with Smith and Blundell's (1986) conditional maximum likelihood estimator; they report the average marginal effects on the observed amount of prime-age unemployment. The delta method is used to compute standard errors. Unless otherwise noted, covariates are the same as in column (9) of Table IV. In (1) and (5) a dummy variable for establishment closure is used as instrument; in (2), (3), (4), (6) and (7) the same dummy variable and the local unemployment rate at graduation are both used as instruments. Apart from the instrument(s), variables included in the regressions of early-career unemployment are the same as in the estimates of prime-age unemployment. The Hansen J statistic is an overidentification test for all instruments. The Anderson-Rubin test [cf. Anderson and Rubin (1949)], the conditional likelihood-ratio test, the Lagrange multiplier test by Moreira (2003) and Kleibergen (2007) and the J overidentification test are all tests of weak IV robust inference For variable definitions see Section 3.

Table VI shows that instrumenting (involuntary) early-career unemployment with a dummy variable for establishment closure at graduation leaves our main result unchanged: early-career unemployment does exhibit long-run scarring effects. And these scarring effects are both statistically and economically significant. For the whole sample, an additional day of youth unemployment on average leads to an increase of prime-age unemployment by 1.67 days, *ceteris paribus*. For larger establishments, the marginal effect of an additional day of early-career unemployment amounts to 1.59 days.

Qualitatively, these results do not change no matter if we use only a dummy variable for establishment closure or both our excluded variables as instruments. They also stay the same irrespective of whether we rely on an IV or a Tobit-IV model and are robust to controlling for establishment dummies and for restricting the estimation sample to larger establishments.

Moreover, one might want to compare Table VI with the results reported in Table IV that do not take account of the likely endogeneity of early-career unemployment. Such a comparison reveals that for all the seven IV/Tobit-IV specifications of Table VI the link between early-career and prime-age unemployment is stronger than in a simple OLS/Tobit regression. In other words, OLS and Tobit estimates are again shown to be downward biased.

Even if early-career unemployment was in fact exogenous, point estimates from Hansen, Heaton and Yaron's (1996) continuously-updated GMM estimator and Tobit-IV would be consistent. In this case, however, OLS or Tobit would be more efficient. That is one reason why for both the IV and the Tobit-IV models we test for the endogeneity of early-career unemployment. In the linear model this is done with the help of a heteroskedasticity-robust form of the Difference-in-Sargan exogeneity test while for the Tobit-IV model Smith and Blundell's (1986) conditional maximum likelihood estimator can directly be used as a test of exogeneity. For both tests the null hypothesis is that early-career unemployment can be treated as exogenous. As Table VI shows, all tests reject this hypothesis on the one percent level.

In line with the approach summarized in Table IV, F-statistics against the null that the excluded instrument is irrelevant are computed for the GMM instrumental variable specifications [cf. columns (1), (2) and (3) of Table VI]. Again, these F-statistics are statistically significant and higher than ten. So, once more, we feel confident that we do not have to worry about weak instrument problems.[22]

[22] Yet, to be on the safe side, we draw on the literature on how to deal with inference in IV models when instruments are weak [which implies point estimates are biased and Wald tests unreliable, cf. Stock, Wright and Yogo (2002)]. In particular, we make use of Finlay and Magnusson's (2009) tests of weak IV robust inference. These have the correct size even when instruments are weak. In Table VI test results are shown separately for the linear and the Tobit instrumental variable models. For the former, the tests allow for estimations that are robust to arbitrary heteroskedasticity or intracluster dependence. For the latter they assume an i.i.d error term. More specifically, the Anderson-Rubin test [cf. Anderson and Rubin (1949)], a conditional likelihood-ratio test, the Lagrange multiplier test by Moreira (2003) and Kleibergen (2007) and a J overidentification test are performed. For all these tests, the null hypothesis

A third set of tests we make use of are Hansen J overidentification tests. Here, the null hypothesis is that both instruments are in fact exogenous. While for the sample that only encompasses larger establishments this null cannot be rejected [cf. column (3) of Table VI], column (2) shows that it is in fact rejected on the one percent level for the whole sample. Yet, the table also shows that the scarring effects found for the two samples do not differ significantly. So any endogeneity of one or both instruments is not strong enough to actually be behind our main result.

5.3. *Heterogeneity in Scarring Effects*

Pure location-shift models confined to the mean of the dependent variable's distribution assume marginal effects to be constant over this distribution. In contrast, quantile regression models — pioneered by Koenker and Bassett (1978) — not only allow the regressors to alter the location of the dependent variable's distribution but also to impact its shape or scale. In the context of the scarring effects of early-career unemployment, this allows an emphasis on the right tail of the (conditional) distribution of prime-age unemployment and a test of whether scarring varies over this distribution.

That is why we will now report results obtained with the help of Chernozhukov and Hong's (2002) 3-step procedure for censored quantile (CQ) regressions. These results do not account for the possible endogeneity of early-career unemployment but will serve as a useful benchmark. Additionally, we will use the 4-step censored quantile instrumental variable estimator developed by Chernozhukov, Fernández-Val and Kowalski (2011). The authors' estimator not only allows an emphasis on the right tail of the (conditional) distribution of prime-age unemployment but also takes care of the corner solution of prime-age unemployment and the possible endogeneity of early-career unemployment. More technically, it combines two approaches. The first is Powell's (1986) idea to deal with censoring semiparametrically through the conditional quantile. The second is a control function approach in the tradition of Hausman (1978). For computation, Chernozhukov and Hong's (2002) algorithm for CQ regressions is augmented with the estimation of the control variable. The estimator's advantages include that it does not require the error term to be homoscedastic. Estimates are consistent and asymptotically normal independent of the distribution of the error term as long as the conditional quantile of the error term is zero.[23]

is that the coefficient associated with early-career unemployment is zero. And as Table VI shows, the hypothesis is rejected by all tests on the one percent level (the one exception is the J overidentification test for the IV model which rejects the null hypothesis "only" on the five percent level).

[23]Cf. Appendix B for a detailed description of Chernozhukov, Fernández-Val and Kowalski's (2011) estimator and Kowalski (2009) for an application in the context of estimating the price elasticity of expenditures on medical care. An alternative CQIV estimator was developed by Blundell and Powell (2007). Because both the CQ and the CQIV procedure are computationally rather demanding, results are reported for a 25 percent sample of our original data set.

TABLE VII

DIFFERENT ESTIMATES OF PRIME-AGE UNEMPLOYMENT — CENSORED QUANTILE (INSTRUMENTAL VARIABLE) REGRESSIONS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Percentile	p50	p55	p60	p65	p70	p75	p80	p85	p90	p95
Censored Quantile Regressions of prime-age unemployment(step 3)										
Early-career unemployment	0.65***	0.69***	0.75***	0.82***	0.93***	1.09***	1.28***	1.49***	1.70***	1.98***
Lower bound	(0.63)	(0.66)	(0.73)	(0.78)	(0.89)	(1.04)	(1.23)	(1.43)	(1.64)	(1.80)
Upper bound	[0.68]	[0.72]	[0.78]	[0.85]	[0.97]	[1.15]	[1.32]	[1.54]	[1.76]	[1.98]
Marginal effect	0.18	0.23	0.32	0.45	0.65	0.92	1.18	1.45	1.67	1.98
Censored Quantile Instrumental Variable Regressions of prime-age unemployment (step 4)										
Early-career unemployment	3.56***	3.66***	3.62***	3.09***	2.91***	3.18***	4.09***	5.09***	6.32***	6.47***
Lower bound	(2.82)	(3.18)	(3.29)	(3.09)	(2.84)	(2.96)	(3.76)	(4.17)	(3.22)	(x.xx)
Upper bound	[4.33]	[4.26]	[4.13]	[4.05]	[3.12]	[3.47]	[4.92]	[5.80]	[8.44]	[x.xx]
Marginal effect	0.96	1.24	1.56	1.70	2.04	2.67	3.76	4.94	6.20	6.47
Control term	-2.69***	-2.71***	-2.60***	-1.97***	-1.65***	-1.71***	-2.39***	-3.16***	-4.11***	-4.01***
Lower bound	(-3.43)	(-3.74)	(-3.49)	(-2.01)	(-1.85)	(-1.99)	(-3.20)	(-3.88)	(-5.84)	(-x.xx)
Upper bound	[-1.92]	[-2.17]	[-2.27]	[-1.94]	[-1.55]	[-1.52]	[-2.09]	[-2.26]	[-1.12]	[-x.xx]
Marginal effect	-0.73	-0.92	-1.12	-1.08	-1.17	-1.44	-2.20	-3.07	-4.03	-4.01
District dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Censored quantile regressions use Chernozhukov and Hong's (2002) 3-step procedure and report lower bounds of 99% confidence intervals in parentheses and upper bounds in brackets. Also reported are the average marginal effects on the observed amount of prime-age unemployment. Censored quantile instrumental variable regressions rely on the estimator developed by Chernozhukov, Fernández-Val and Kowalski (2011). Here, the whole 4-step procedure is bootstrapped and lower bounds of bias-corrected 99% confidence intervals are in parentheses and upper bounds in brackets. *** indicates that the 99% confidence interval does not include zero. All quantile regressions are calculated using Stata's *qreg* command with 50 replications. The instrument is the local unemployment rate at graduation. Covariates are the same as in column (9) of Table IV. For variable definitions see Section 3.

Outputs of the final (and crucial) steps of ten CQ as well as ten CQIV regressions of prime-age unemployment on early-career unemployment are summarized in Table VII. They are visualized in Figure 3, too. For all regressions, the control variables introduced in Section 3 as well as dummy variables for the training firms' districts are again included (to save space neither are displayed in the table). Throughout, results are presented for selected quantiles of the conditional distribution of prime-age unemployment. A large proportion of sampled individuals exhibit no or little prime-age unemployment. Besides, we are most interested in those individuals that conditional on observables suffer from a very elevated amount of unemployment. Therefore, our regressions start at the median and proceed in steps of five percentiles all the way to the 95^{th} percentile.

As in the Tobit model, the CQ and CQIV regressions' coefficients measure how the latent amount of prime-age unemployment m_{t2}^* changes with respect to changes in the regressors. That is why the average marginal effects on the observed amount of prime-age unemployment m_{t2} are also displayed in table VII [cf. Kowalski (2009) and Chernozhukov, Fernández-Val and Kowalski (2011)].

The table's upper panel displays the censored quantile regressions' outputs. In line with the OLS regression results discussed above, a significant and positive relationship between early-career unemployment and prime-age unemployment exists even if all our control variables are taken into account. What is illuminating is that this relationship is especially pronounced in the right tail of the (conditional) distribution of prime-age unemployment: at the 95$^{\text{th}}$ percentile an additional day of early-career unemployment goes hand in hand with an increase of prime-age unemployment by 1.98 days.

Outputs from CQIV regressions are presented in the lower panel of Table VII. Here, explanatory variables not only include early-career unemployment but also a control term generated in the CQIV regressions' first stage. This control term's coefficient directly gives the direction and magnitude of the bias that results if one ignores the endogeneity of early-career unemployment (cf. Appendix B). Qualitatively, the table confirms the CQ regressions' main result, namely the existence of a significant and positive relationship between early-career and prime-age unemployment. Moreover, because of the control variable approach we can now interpret this relationship as causal: unemployment early in the professional career has a long-term scarring effect. And this scarring effect is present not only at the mean or median but at all the estimated quantiles. Moreover, it is statistically significant at all these quantiles.

Confirming the results of the mean estimates, the CQIV regressions' coefficients are larger than those found with the help of censored quantile regressions for all quantiles studied here. By implication, the estimates produced with the help of CQ regressions are downward-biased. This conclusion is also mirrored by the consistently negative coefficients associated with the control terms in the CQIV regressions' fourth steps. A closer look at these different coefficients reveals that the downward bias is most pronounced in the right tail of the distribution of prime-age unemployment.

The scarring effect of early-career unemployment varies considerably across the quantiles studied here. In fact, confidence intervals from the Tobit-IV model and the CQIV procedure overlap only between the 55$^{\text{th}}$ and the 75$^{\text{th}}$ percentile. For all other percentiles, estimates are inconsistent with the premise that early-career unemployment exerts a pure location shift.

Even more importantly from an economic point of view, Table VII and Figure 3 show that scarring is strongest in the right tail of distribution of prime-age unemployment. Thus, individuals who experience more unemployment during their prime age than others with comparable observable characteristics are particularly affected by early-career unemployment. This might be due to unobservables exogenous to the scarring effect of early-career unemployment that alter the signal sent by and/or the degree of human capital lost during early-career unemployment and thereby influence the position in the conditional distribution of prime-age unemployment.

What is striking is the magnitude of the heterogeneity in scarring effects: while at the median an additional day of youth unemployment increases prime-

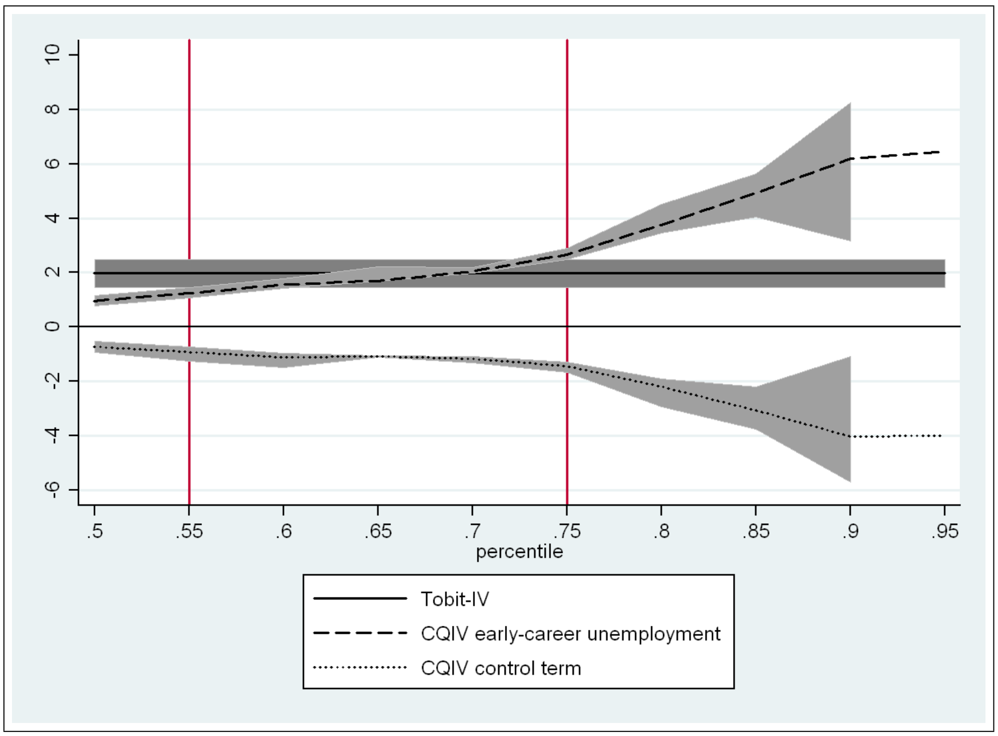

FIGURE 3.— Different Estimates of Prime-age Unemployment – Censored Quantile (Instrumental Variable) Regressions

Notes: Average marginal effects of early-career unemployment and a control term on the observed amount of prime-age unemployment and 99% confidence intervals. Censored quantile regressions use Chernozhukov and Hong's (2002) 3-step procedure. The Tobit-IV regression is calculated with Smith and Blundell's (1986) conditional maximum likelihood estimator; it reports the . The delta method is used to compute standard errors. All quantile regressions are calculated using Stata's *qreg* command with 50 replications. The instrument is the local unemployment rate at graduation. Covariates are the same as in column (9) of Table IV. For variable definitions see Section 3.

age unemployment by 0.96 days, scarring is more than six times stronger at the 95$^{\text{th}}$ percentile. Here, another day of early-career unemployment induces 6.47 days of prime-age unemployment.[24]

6. CONCLUSIONS

In an influential paper, Heckman and Borjas (1980, p. 247) asked: "Does unemployment cause future unemployment?" In this study, we attempted to answer their question with German administrative matched employer-employee data that allowed us to follow more than 800,000 individuals over 24 years. Using the innovative censored quantile instrumental variable estimator introduced by Chernozhukov, Fernández-Val and Kowalski (2011) and instrumenting (involuntary) early-career unemployment with local labor market conditions at labor market entry, we showed that unemployment is very persistent over the professional career and that youth unemployment has significant and long-term scarring effects. These effects are especially pronounced in the right tail of the (conditional) distribution of prime-age unemployment. While at the median an additional day of youth unemployment leads to an increase of prime-age unemployment by less than one day, at the 95$^{\text{th}}$ percentile another day of early-career unemployment induces almost six and a half days of prime-age unemployment.

These findings have several important implications: First, they imply that early-career joblessness contributes to the inequality of unemployment experience over the professional career documented by Schmillen and Möller (2012). Second, they lend support to theoretical models of state dependence like those by Vishwanath (1989), Lockwood (1991) or Pissarides (1992) and are in line with the findings by Raaum and Røed (2006), von Wachter and Bender (2006) and others that having good or bad luck early in the professional career can have significant and long-lasting consequences. Third, concerning labor market policies they suggest that these should emphasize the (re-)integration of youths into the labor market, the furthering of efficient and transparent early-career matching processes and, above all, the prevention of early-career unemployment.

While this study has focused on graduates from Germany's dual education system, we would argue that it allows to draw lessons for other economies, too. First of all, this is because — as already mentioned — dual education systems

[24] A different form of heterogeneity in scarring effects is the subject of Table XII in Appendix C. The table summaries nine Tobit-IV regressions where the dependent variables are the total amounts of unemployment over overlapping eight-year-long subperiods of prime age. In other words, in the first estimation, unemployment during years nine to 16 on the labor market is regressed on early-career unemployment, in the second regression the dependent variable is unemployment during years ten to 17 etc. Following a similar exercise by Gregg and Tominey (2005) all regressions control for the amount of unemployment experienced between the early years of the professional career and the period on the left-hand side of the estimation equation. As is evident from Table XII, early-career unemployment has a scarring effect during all phases of the professional career considered here. Unsurprisingly and in accordance with what is found by Gregg and Tominey (2005), this effect generally weakens as the professional career progresses.

play a prominent role not only in Germany but also in many other countries (e.g. in Austria, Switzerland or on the Balkans). In yet another group of countries, including in the United States and the United Kingdom, there has long been a discussion about whether to strengthen the importance of education programs that combine vocational training in a company and learning at a school [cf. for instance Heckman (1993) or Neumark (2002)].

Even more importantly, von Wachter and Bender (2006) point to the basic similarities in the labor markets for young workers between Germany and the United States while according to Ryan (2001) state dependence is unlikely to be specific to any one economy. These observations make our results conceptually relevant for developed economies more generally. Going even further, Ryan (2001, p. 49) asserts that "[a]ny adverse effects [of youth unemployment] on subsequent outcomes should be weakest in tight labor markets, where jobs are easy to find." Compared to the high rates of youth unemployment in many OECD economies at the moment, the individuals in our sample were faced with relatively low rates of joblessness during their first years on the German labor market. Thus, our results might represent a lower bound for the scarring effects youths currently unemployed in the United States, Italy or Spain will have to cope with in the future.

In closing, we would like to stress that in our view much more research on the scarring effects of youth unemployment is needed. In particular, this study has not attempted to investigate through which transmission channels scarring actually operates. Besides, an instrumental variable technique like the one used here can never be beyond doubt. We can for instance not completely rule out the possibility that widespread early-career unemployment influences a region's work norms and thus has scarring effects beyond the ones found here. We would see it as desirable if our study was complimented by other investigations that make use of a different set of instruments or even natural experiments (difficult as that may be to achieve). Lastly, our focus has solely been on the consequences of early-career joblessness for future unemployment. The resulting scarring effect might represent only one aspect of the actual extent of state dependence. In fact, Bell and Blanchflower (2011) use British data to show that even at age 50 individuals who suffered from youth unemployment report worse physical and mental health and lower job satisfaction than observationally similar individuals with no youth unemployment experience. While Bell and Blanchflower's (2011) findings should probably not be interpreted as causal, it might be worthwhile to investigate whether early-career unemployment has a long-term impact on the quality of employment, health or even mortality.

ACKNOWLEDGEMENTS

We thank Joachim Möller, Joshua Angrist, Stefan Bender, Philipp vom Berge and Mariana Viollaz as well as conference and seminar participants in Bonn, Göttingen, Kallmünz, Malaga and Nuremberg for helpful comments and sugges-

tions. The usual disclaimer applies.

REFERENCES

Acemoglu, Daron (1995). Public Policy in a Model of Long-Term Unemployment. *Economica* **62** 161–178.

Acemoglu, Daron and Pischke, Jörn-Steffen (1998). Why Do Firms Train? Theory and Evidence. *Quarterly Journal of Economics* **113** 79–119.

Anderson, T.W. and Rubin, Herman (1949). Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations. *Annals of Mathematical Statistics* **20** 46–63.

Arulampalam, Wiji (2001). Is Unemployment Really Scarring? Effects of Unemployment Experiences on Wages. *The Economic Journal* **111** 585–606.

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

Baum, Christopher; Schaffer, Mark and Stillman, Steven (2003). Instrumental Variables and GMM: Estimation and Testing. *Stata Journal* **3** 1–31.

——— (2007). Enhanced Routines for Instrumental Variables/GMM Estimation and Testing. *Stata Journal* **7** 465–506.

Beaudry, Paul and DiNardo, John (1991). The Effect of Implicit Contracts on the Movement of Wages Over the Business Cycle: Evidence from Micro Data. *Journal of Political Economy* **99** 665–688.

Bell, David and Blanchflower, David (2011). Young People and the Great Recession. *Oxford Review of Economic Policy* **27** 241–267.

Ben-Porath, Yoram (1967). The Production of Human Capital and the Life Cycle of Earnings. *Journal of Political Economy* **75** 352–365.

Blossfeld, Hans-Peter (1987). Labor-Market Entry and the Sexual Segregation of Careers in the Federal Republic of Germany. *American Journal of Sociology* **93** 89–118.

Blundell, Richard and Powell, James (2007). Censored Regression Quantiles with Endogenous Regressors. *Journal of Econometrics* **141** 65–83.

Bönke, Timm; Corneo, Giacomo and Lüthen, Holger (2011). Lifetime Earnings Inequality in Germany. *IZA Discussion Paper 6020*.

Bowlus, Audra (1995). Matching Workers and Jobs: Cyclical Fluctuations in Match Quality. *Journal of Labor Economics* **13** 335–350.

Burgess, Simon; Propper, Carol; Rees, Hedley and Shearer, Arran (2003). The Class of 1981: The Effects of Early Career Unemployment on Subsequent Unemployment Experiences. *Labour Economics* **10** 291–309.

Chernozhukov, Victor and Hansen, Christian (2006). Instrumental Quantile Regression Inference for Structural and Treatment Effect Models. *Journal of Econometrics* **132** 491–525.

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

Chernozhukov, Victor, Fernández-Val, Iván and Kowalski, Amanda (2011). Quantile Regression with Censoring and Endogeneity. *NBER Working Paper 16997*.

Clark, Andrew; Georgellis, Yannis and Sanfey, Peter (2001). Scarring: The Psychological Impact of Past Unemployment. *Economica* **68** 221–241.

Corcoran, Mary and Hill, Martha (1985). Reoccurrence of Unemployment among Adult Men. *The Journal of Human Resources* **20** 165–183.

Del Bono, Emilia and Weber, Andrea (2008). Do Wages Compensate for Anticipated Working Time Restrictions? Evidence from Seasonal Employment in Austria. *Journal of Labor Economics* **26** 181–221.

The Economist (2011) Left Behind, September 10th, 60–62.

Edin, Per-Anders and Gustavsson, Magnus (2008). Time out of Work and Skill Depreciation. *Industrial and Labor Relations Review* **61** 163–180.

Ellwood, David (1982). Teenage Unemployment: Permanent Scars or Temporary Blemishes. In Freeman, R. and Wise, D. (Eds.). *The Youth Labor Market Problem: Its Nature, Causes, and Consequences*, NBER Books, Cambridge.

Finlay, Keith and Magnusson, Leandro (2009). Implementing Weak-instrument Robust Tests for a General Class of Instrumental-variables Models. *Stata Journal* **9** 398–421.

Fitzenberger, Bernd and Wilke, Ralf (2010). Unemployment Durations in West-Germany Before and After the Reform of the Unemployment Compensation System During the 1980s. *German Economic Review* **11** 336–366.

Freeman, Richard and Wise, David (1982). The Youth Labor Market Problem: Its Nature Causes and Consequences. In Freeman, R. and Wise, D. (Eds.). *The Youth Labor Market Problem: Its Nature, Causes, and Consequences*, NBER Books, Cambridge.

Gibbons, Robert and Katz, Lawrence (1991). Layoffs and Lemons. *Journal of Labor Economics* **9** 351–380.

Gregg, Paul (2001). The Impact of Youth Unemployment on Adult Unemployment in the NCDS. *Economic Journal* **109** 626–653.

Gregg, Paul and Tominey, Emma (2005). The Wage Scar from Male Youth Unemployment. *Labour Economics* **12** 487-509.

Hagedorn, Marcus and Manovskii, Iourii (forthcoming). Job Selection and Wages over the Business Cycle. *American Economic Review*.

Hamermesh, Daniel (1987). The Costs of Worker Displacement. *The Quarterly Journal of Economics* **102** 51–75.

Hansen, Lars Peter; Heaton, John and Yaron, Amir (1996). Finite-Sample Properties of Some Alternative GMM Estimators. *Journal of Business & Economic Statistics* **14** 262–280.

Hausman, Jerry (1978). Specification Tests in Econometrics. *Econometrica* **46** 1251–1271.

Heckman, James (1993). Assessing Clinton's Program on Job Training, Workfare, and Education in the Workplace. *NBER Working Paper 4428*.

Heckman, James and Borjas, George (1980). Does Unemployment Cause Future Unemployment? Definitions, Questions and Answers from a Continuous Time Model of Heterogeneity and State Dependence. *Economica* **47** 247–283.

Hethey, Tanja and Schmieder, Johannes (2010). Using Worker Flows in the Analysis of Establishment Turnover – Evidence from German Administrative Data. *FDZ Datenreport 06/2010*.

Hethey-Maier, Tanja and Seth, Stefan (2010). The Establishment History Panel (BHP) 1975-2008. *FDZ Datenreport 04/2010*.

Hippach-Schneider, Ute; Krause, Martina and Woll, Christian (2007). *Vocational Education and Training in Germany: Short Description*, European Centre for the Development of Vocational Training, Thessaloniki.

Hoffmann, Florian (2010). An Empirical Model of Life-Cycle Earnings and Mobility Dynamics. *Mimeo, University of British Columbia*.

Keane, Michael and Wolpin, Kenneth (1997). The Career Decisions of Young Men. *Journal of Political Economy* **105** 473–522.

Kleibergen, Frank (2007). Generalizing Weak Instrument Robust IV Statistics towards Multiple Parameters, Unrestricted Covariance Matrices and Identification Statistics. *Journal of Econometrics* **139** 181–216.

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

Kowalski, Amanda (2009). Censored Quantile Instrumental Variable Estimates of the Price Elasticity of Expenditure on Medical Care. *NBER Working Paper 15085*.

Kroft, Kory; Lange, Fabian and Notowidigdo, Matthew (2012). Duration Dependence and Labor Market Conditions: Theory and Evidence from a Field Experiment. *Mimeo, University of Toronto*.

Lockwood, Ben (1991). Information Externalities in the Labour Market and the Duration of Unemployment. *Review of Economic Studies* **58** 733–753.

Margolis, David; Simonnet, Véronique and Vilhuber, Lars (2001). Using Early Career Experiences and Later Career Outcomes to Distinguish between Models of Labor Market Behavior under Institutional Constraints. *Centre National de la Recherche Scientifique Working Paper 35/2001.*

Moreira, Marcelo (2003). A Conditional Likelihood Ratio Test for Structural Models. *Econometrica* **71** 1027–1048.

Mortensen, Dale (1986). Job Search and Labor Market Analysis. In Ashenfelter, O. and Layard, R. (Eds.). *Handbook of Labor Economics, Vol. 2*, Elsevier, Amsterdam.

Mroz, Thomas and Savage, Timothy (2006). The Long-Term Effects of Youth Unemployment. *Journal of Human Resources.* **41** 259–293.

Mühleisen, Martin and Zimmermann, Joachim (1994). New Patterns of Labour Mobility: A Panel Analysis of job Changes and Unemployment. *European Economic Review.* **38** 793–801.

Neumark, David (2002). Youth Labor Markets In The United States: Shopping Around Vs. Staying Put. *The Review of Economics and Statistics* **84** 462–482.

Niedergesäss, Markus (2012). Duration Dependence, Lagged Duration Dependence, and Occurrence Dependence in Individual Employment Histories. *University of TübingenWorking Papers in Economics and Finance, No. 26.*

Nilsen, Øivind and Holm, Katrine (2011). Scarring Effects of Unemployment. *CESifo Working Paper 3675*

Nordström Skans, Oskar (2011). Scarring Effects of the First Labor Market Experience. *IZA Discussion Paper 5565*

Oberschachtsiek, Dirk; Scioch, Patrycja; Seysen, Christian and Heining, Jörg (2009). Integrated Employment Biographies Sample IEBS: Handbook for the IEBS in the 2008 Version. *FDZ Datenreport 03/2009.*

OECD (2012). *Labor Force Statistics by Sex and Age*, retrieved in June 2012, OECD, Paris.

Oreopoulos, Philip; von Wachter, Till and Heisz, Andrew (2012). The Short- and Long-Term Career Effects of Graduating in a Recession. *American Economic Journal: Applied Economics* **4** 1–29.

Pissarides, Christopher (1992). Loss of Skill during Unemployment and the Persistence of Employment Shocks. *Quarterly Journal of Economics* **107** 1371–1391.

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

Raaum, Oddbjørn and Røed, Knut (2006). Do Business Cycle Conditions at the Time of Labor Market Entry Affect Future Employment Prospects? *The Review of Economics and Statistics* **88** 193–210.

Ryan, Paul (2001). The School-to-Work Transition: A Cross-National Perspective. *Journal of Economic Literature* **39** 34–92.

Schmelzer, Paul (2010). The Consequences of Job Mobility for Future Earnings in Early Working Life in Germany: Placing Indirect and Direct Job Mobility into Institutional Context. *European Sociological Review* **28** 85–95.

Schmieder, Johannes; von Wachter, Till and Bender, Stefan (2012). The Effects of Extended Unemployment Insurance Over the Business Cycle: Evidence from Regression Discontinuity Estimates Over 20 Years. *The Quarterly Journal of Economics* **127** 701–752.

Schmillen, Achim and Möller, Joachim (2012). Distribution and Determinants of Lifetime Unemployment. *Labour Economics* **19** 33–47.

Shapiro, Carl and Stiglitz, Joseph (1984). Equilibrium Unemployment as a Worker Discipline Device. *American Economic Review* **74** 433–444.

Smith, Richard and Blundell, Richard (1986). An Exogeneity Test for a Simultaneous Equation Tobit Model with an Application to Labor Supply. *Econometrica* **54** 679–685.

Staiger, Douglas and Stock, James (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica* **65** 557–586.

Stevens, Katrien (2008). Adverse Economic Conditions at Labour Market Entry: Permanent Scars or Rapid Catch-up? *Mimeo, University of Sidney.*

Stock, James; Wright, Jonathan and Yogo, Motohiro (2002). A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments. *Journal of Business & Economic Statistics* **20** 518–529.

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

Topel, Robert and Ward, Michael (1992). Job Mobility and the Careers of Young Men. *The Quarterly Journal of Economics* **107** 439–479.

Uhlig, Harald and Xu, Yexiao (1996). Effort and the Cycle: Cyclical Implications of Efficiency Wages. *Tilburg University, Center for Economic Research Discussion Paper 1996-49*.

Vishwanath, Tara (1989). Job Search, Stigma Effect, and Escape Rate from Unemployment. *Journal of Labor Economics* **7** 487–502.

von Wachter, Till 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.

Wesselbaum, Dennis (2010). Reciprocity and Matching Frictions. *Kiel Institute Working Paper 1616*.

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

APPENDIX A: DATA SELECTION AND CLEANSING

As mentioned in Section 3, our analysis focuses on all those individuals that graduated from Germany's dual education system between 1978 and 1980. Moreover, in order to ensure valid and undistorted results and to limit the impact of non-standard professional careers, it excludes a number of groups. Maybe most importantly, women are excluded because of data problems. In particular, these are related to the wakweak female labor market attachment (especially in the cohorts studied here) and the comparatively large number of women who do not qualify for unemployment benefits. Another group that is not considered are East Germans because their employment history has only been recorded in our data since the early 1990s. We label as "East German" all those individuals whose first employment or unemployment spell registered by the social security system takes place in East Germany.

Furthermore, our analysis does not cover foreign nationals, i.e. individuals that at no point during their professional career possessed a German passport. Not included either are individuals that held a high school diploma ("Abitur") when they graduated from their first apprenticeship. For the labor market entry cohorts considered here this was the case for only around five percent of individuals and we conjecture that in terms of unobserved characteristics they might hardly be comparable to the rest of our estimation sample. For similar reasons, we also exclude individuals who finished their first apprenticeship either at age 14 or earlier or at age 27 or later. Finally, we leave out all individuals for whom there are no IEB records at all in the eight years after they finished their first apprenticeship and/or the subsequent 16 years.

While in general the information contained in our administrative matched employer-employee data set can be considered highly reliable, it is not totally free from questionable information. That is why we went through all our main and control variables and replaced implausible data points with missing values. For example, the IEB contains a small number of occupational codes that have been documented to be erroneous and some figures listed for the remuneration prior to graduation from the dual education system appeared unrealistically low or high.

APPENDIX B: CENSORED QUANTILE INSTRUMENTAL VARIABLE REGRESSION

Assume linearity in parameters and a conditional quantile function of the dependent variable $m^*_{t2} = Q_{m^*_{t2}}(\tau|d_{t1}, w, o_{t1}, u_{t2})$ (prime-age unemployment) at quantile τ that depends on the regressor of interest d_{t1} (early-career unemployment), a vector of exogenous covariates w (including a constant and possibly the censoring variable), a latent and unobserved variable o_{t1} which is correlated with m^*_{t2} as well as with d_{t1} and the error term u_{t2} with a conditional quantile of zero, $\mathrm{Q}_{u_{t2}}(\tau|d_{t1}, w, o_{t1}) = 0$.[25] Then, with $\tau \in [0, 1]$ indexing the quantile and $\{i = 1, ..., N\}$ indicating the individual, we arrive at the following system of equations:

$$m^*_{i,t2} = d_{i,t1}\alpha(\tau) + w'_i\beta(\tau) + o_{i,t1}\gamma(\tau) + u_{i,t2}, \tag{B.1}$$

$$d_{i,t1} = w'_i\dot{\beta} + z_{i,t0}\pi + o_{i,t1}, \tag{B.2}$$

where $\alpha(\tau)$, $\beta(\tau)$ and $\gamma(\tau)$ are parameters to be estimated. Further assume conditional independence of u_{t2} and o_{t1}, $u_{t2} \sim U(0, 1)|d_{t1}, w, z_{t0}, o_{t1}$ and $o_{t1} \sim U(0, 1)|w, z_{t0}$. As long as we cannot control for o_{t1}, estimates of $\alpha(\tau)$ would be biased and inconsistent because o_{t1} would be absorbed by the new error term "inducing endogeneity or selection bias, so that the conditional quantile of selected $[m^*_{t2}]$ given the selected $[d_{t1}]$, is generally not equal to the quantile of potential outcome" [Chernozhukov and Hansen (2006, p.494)]. While we cannot observe o_{t1} directly, we can estimate it from the residuals of Equation B.2. To accomplish this, we need to use the "instrumental variable" z_{t0} that is excluded from Equation B.1 but influences d_{t1} through π in Equation B.2. This instrumental variable enables us to control for any endogenous

[25] Chernozhukov and Hansen (2006) note that neither the hypothetical values of m^*_{t2} which would evolve under random assignment of treatment nor its corresponding quantiles are actually observable if endogeneity is present. However, CQIV still allows to recover the structural parameters of $Q_{m^*_{t2}}(\tau|.)$.

variation of d_{t1} in Equation B.2 and thus to recover the parameters of interest. This is why o_{t1} is known as the *control term* and Equation B.2 as the *control function*.

Our study uses labor market conditions at the time of graduation as instruments. Therefore, o_{t1} could be interpreted as the marginal propensity to experience early-career unemployment evaluated at the respective position in the distribution of prime-age unemployment conditional on the quality of initial matching of apprentices to firms and further exogenous characteristics.

Additionally, we face a corner solution with positive probability mass at zero. That is why we interpret m^*_{t2} as the latent amount of prime-age unemployment as opposed to the actually observed amount of prime-age unemployment, m_{t2}. It holds that

$$m_{i,t2} = \begin{cases} m^*_{i,t2} & \text{if } m^*_{i,t2} \geq 0 \text{ and} \\ 0 & \text{if } m^*_{i,t2} < 0. \end{cases} \tag{B.3}$$

The conditional quantile function of m_{t2} is

$$\mathrm{Q}_{m_{t2}}(\tau|X) = \max(X'\psi(\tau), 0), \tag{B.4}$$

where $X \equiv [d_{t1}, w, o_{t1}]$ and $\psi(\tau) \equiv [\alpha(\tau), \beta(\tau), \gamma(\tau)]$. Equation B.4 holds because quantiles are equivariant against monotone transformations, such as censoring. In the presence of exogenous regressors, the model presented so far could be consistently estimated with Powell's (1986) estimator. Better applicability is achieved by the semi-parametric estimator developed by Chernozhukov and Hong (2002) which is asymptotically as efficient as Powell's (1986) estimator but far less computationally demanding.

Chernozhukov, Fernández-Val and Kowalski (2011) combine Chernozhukov and Hong's (2002) estimator with a control function approach. The authors show that under mild regularity assumptions, $\sqrt{n}$-consistent and asymptotically normal estimates for $\psi(\tau)$ at every quantile τ can be obtained by

$$\hat{\psi}(\tau) = arg \min_{\psi \in \mathbb{R}^{dim(X)}} \frac{1}{N} \sum_{i=1}^{N} I(\hat{S}_i'\hat{\delta} > k)\rho_\tau(m_{i,t2} - \hat{X}_i'\psi). \tag{B.5}$$

Here $I(.)$ is an indicator function taking on unity when the expression holds and zero otherwise, $\rho_\tau(u_{t2})$ is Koenker and Bassett's (1978) absolute asymmetric loss function, $\hat{X} = x(d_{t1}, w, \hat{o}_{t1})$, $\hat{S} = s(\hat{X}, 0)$ and $x(.)$ as well as $s(.)$ are vectors of transformations of (d_{t1}, w, o_{t1}) or $(X, 0)$, respectively. $I(\hat{S}'\hat{\delta} > k)$ is called "selector" by Chernozhukov, Fernández-Val and Kowalski (2011) because - by identifying uncensored observations with censored predictions - it selects the subset of observations for which a linear form of the conditional quantile function can be assumed. Unfortunately, linear programming cannot be used to solve Equation B.5. Instead, one may rely on an algorithm proposed by Chernozhukov, Fernández-Val and Kowalski (2011) which augments the 3-step procedure of Chernozhukov and Hong (2002) by an additional step. The resulting four steps are as follows:

Step 1. Run an OLS regression of d_{t1} on the instrument z_{t0} and exogenous regressors w and obtain a prediction for the control term $\hat{o}_{t1} = \hat{F}_d(d_{t1}|w, z_{t0})$ from the residuals. This allows the construction of $\hat{X} = x(d_{t1}, w, \hat{o}_{t1})$.

Step 2. Identify the linear part of the conditional quantile function $X'\psi_0(\tau)$. To do so, choose a subset of observations for which the conditional quantile line is "sufficiently" above zero, $\{i : X_i'\psi_0(\tau) > 0\}$. Estimating a logit model for the conditional probability of non-censoring $P(m_{t2} = 1|S)$,

$$P(m_{i,t2} = 1|\hat{S}_i) = \Lambda(\hat{S}_i'\delta_0), \tag{B.6}$$

allows to choose a sample $J_0(c)$ that contains those observations which satisfy

$$J_0(c) = \{i : \Lambda(\hat{S}_i'\hat{\delta}_0) > 1 - \tau + c\}, \tag{B.7}$$

with $0 < c < \tau$. Chernozhukov and Hong (2002) suggest to choose c such that $\#J_0(c)/\#J_0(0) = 0.9$.

Step 3. Run an ordinary quantile regression on subsample $J_0(c)$. This gives

$$\hat{\psi}_0(\tau) = arg \min_{\psi \in \mathbb{R}^{dim(X)}} \sum_{i \in J_0(c)} \rho_\tau(m_{i,t2} - \hat{X}_i' \psi), \tag{B.8}$$

a consistent but inefficient estimate. To gain efficiency, the subset of observations used in Step 2 is updated by choosing $J_1(k)$ according to:

$$J_1(k) = \{i : \hat{X}_i' \hat{\psi}_0(\tau) > k\}, \tag{B.9}$$

where the fitted values from Equation B.8 are used and the cut-off k plays a similar role as c did in Step 2.

Step 4. Finally, repeat Step 3 but this time on subsample $J_1(k)$.

APPENDIX C: SUPPLEMENTARY TABLES AND FIGURES

TABLE VIII

Summary statistics on explanatory variables

variable	mean	standard deviation	minimum	maximum
local unemployment rate at graduation	3.64	1.28	0.9	8.2
local unemployment rate at transition	8.98	3.54	0.9	19.8
class of 1978	0.29	—	0	1
class of 1979	0.36	—	0	1
class of 1980	0.35	—	0	1
age at graduation	18.69	1.67	15	26
remuneration at graduation	10.88	5.84	0.01	176.60
agriculture	0.03	—	0	1
energy and mining	0.02	—	0	1
manufacturing	0.50	—	0	1
construction	0.18	—	0	1
trade	0.14	—	0	1
transport and communications	0.03	—	0	1
financial intermediation	0.02	—	0	1
other services	0.08	—	0	1
non-profits and households	0.003	—	0	1
public administration	0.02	—	0	1
size of the establishment	984.46	4482.37	1	57236
median wage of the establishment	38.06	9.04	1.15	82.44
agricultural occupations	0.02	—	0	1
unskilled manual occupations	0.08	—	0	1
skilled manual occupations	0.67	—	0	1
technicians and engineers	0.04	—	0	1
unskilled services occupations	0.02	—	0	1
skilled services occupations	0.01	—	0	1
semiprofessions and professions	0.02	—	0	1
unskilled commercial occupations	0.03	—	0	1
skilled commercial occupations and managers	0.13	—	0	1

Notes: For variable definitions see Section 3.

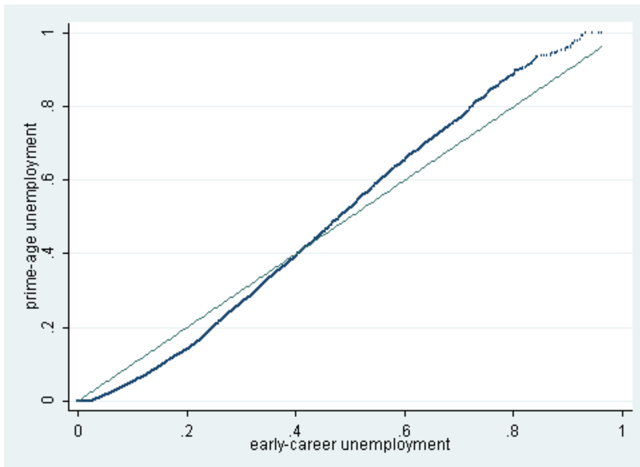

FIGURE 4.— Quantile-quantile plot of early-career vs. prime-age unemployment, measured as proportion of potential time on the labor market.

TABLE IX

RELATION BETWEEN EARLY-CAREER UNEMPLOYMENT AND LATER UNEMPLOYMENT - COHORTS ENTERING THE LABOR MARKET IN 1978, 1979, 1980 AND 1981 POOLED.

early-career unemployment	obs.	later unemployment	employment years 9 to 12	13 to 16	17 to 20	21 to 24
0 to p50	643,606	occurrence	0.07	0.09	0.08	0.09
		mean amount	18.63	31.05	34.45	37.67
p51 to p60	129,760	occurrence	0.15	0.14	0.13	0.13
		mean amount	35.60	49.8	53.81	57.39
p61 to p70	127,099	occurrence	0.21	0.17	0.17	0.16
		mean amount	56.46	64.07	71.03	73.78
p71 to p80	128,823	occurrence	0.27	0.22	0.21	0.2
		mean amount	71.01	87.09	96.17	98.56
p81 to p90	128,205	occurrence	0.37	0.29	0.27	0.25
		mean amount	110.34	124.07	137.12	138.24
p91 to p95	64,299	occurrence	0.45	0.34	0.31	0.29
		mean amount	156.13	165.12	180.66	175.91
p96 to 1	64,229	occurrence	0.67	0.5	0.45	0.41
		mean amount	359.07	339.89	351.64	326.55

Notes: *Occurrence* is measured as the proportion of individuals registered as unemployed for at least one day within each time-frame. *Mean amount* denotes the mean total unemployment generated within each time-frame.

TABLE X

Different Estimates of Prime-age Unemployment – Tobit-IV Regressions.

	(1)	(2)	(3)	(4)
Model	Tobit-IV	Tobit-IV	Tobit-IV	Tobit-IV
Marginal effect	Average marginal effects on latent variable	Average marginal effects on observed variable	Marginal effects on observed variable at the average	Average marginal effects on positive observations
Regressions of prime-age unemployment				
Early-career unemployment	5.14***	1.98***	2.14***	1.74***
	(0.60)	(0.20)	(0.25)	(0.21)
Age	-22.87***	-8.79***	-9.53***	-7.77***
	(1.86)	(0.70)	(0.77)	(0.63)
Remuneration	2.06	0.79*	0.86	0.70
	(1.23)	(0.48)	(0.53)	(0.44)
Size of training firm	-4.41**	-1.69**	-1.83**	-1.45**
	(1.78)	(0.71)	(0.74)	(0.60)
Median wage of training firm	-0.82	-0.32	-0.34	-0.28
	(0.77)	(0.30)	(0.32)	(0.26)
Unskilled manual occ.	42.56	16.37	17.74	14.47
	(36.74)	(14.02)	(15.30)	(12.50)
Skilled manual occ.	13.94	5.36	5.81	4.73
	(55.50)	(21.27)	(23.14)	(18.87)
Technicians and engineers	30.79	11.84	12.84	10.46
	(75.52)	(28.87)	(31.48)	(25.69)
Unskilled services	-52.07	-20.03	-21.71	-17.70
	(37.48)	(14.38)	(15.62)	(12.74)
Skilled services	-32.36	-12.45	-13.49	-11.00
	(54.37)	(21.07)	(22.67)	(18.46)
Semiprofessions	49.09	18.88	20.46	16.69
and professions	(88.11)	(33.61)	(36.74)	(29.99)
Unskilled commercial occ.	259.18***	99.68***	108.04***	88.08***
	(72.96)	(26.56)	(30.37)	(24.96)
Skilled commercial occ.	129.27	49.71	53.89	43.94
and managers	(89.09)	(33.48)	(37.13)	(30.37)
Other variables included in regressions				
District dummies	Yes	Yes	Yes	Yes
Cohort dummies	Yes	Yes	Yes	Yes
Sector dummies	Yes	Yes	Yes	Yes
Unemployment at transition	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes
Number of observations	739,432	739,432	739,432	739,432

Notes: Standard errors clustered at the district level in parentheses. * , (**), [***] indicates significance at the 10, (5), [1] % level. Tobit-IV regressions are calculated with Smith and Blundell's (1986) conditional maximum likelihood estimator. The instrument is the local unemployment rate at graduation. In (1) the marginal effects on the latent amount of prime-age unemployment (i.e. the model's coefficients) are reported; in (2) the average marginal effects on the observed amount of prime-age unemployment are reported; in (3) the marginal effects on the observed amount of prime-age unemployment are reported if all explanatory variables take on their average value; in (2) the average marginal effects on the observed amount of prime-age unemployment are reported among the subpopulation for which prime-age unemployment is not at a boundary. For all factor variables the discrete first differences from the base categories are calculated. Apart from the instrument, variables included in the regressions of early-career unemployment are the same as in the estimates of prime-age unemployment. For variable definitions see Section 3.

TABLE XI

Different Estimates of Prime-age Unemployment — Tobit Robustness Regressions.

	(1)	(2)	(3)	(4)	(5)	(6)
Specification	Unempl. in origin at transition as control	Minimum unempl. in prime age as control	At least one observation during last four years	Less than six years of seasonal empl.	Nonempl. I instead of unempl.	Nonempl. II instead of unempl.
Model	Tobit	Tobit	Tobit	Tobit	Tobit	Tobit
Regressions of prime-age unemployment [prime-age non-employment in (5) and (6)]						
Early-career unemployment	0.57***	0.57***	0.61***	0.50***	—	—
	(0.01)	(0.01)	(0.01)	(0.01)		
Early-career non-employment	—	—	—	—	0.58***	0.30***
					(0.01)	(0.01)
Other variables included in regressions						
District dummies	Yes	Yes	Yes	Yes	Yes	Yes
Unempl. at transition (current)	No	Yes	Yes	Yes	Yes	Yes
Unempl. at transition (origin)	Yes	No	No	No	No	No
Minimum unempl. in early career	No	Yes	No	No	No	No
Number of observations	740,394	739,432	648,644	652,206	739,432	739,432

Notes: Standard errors clustered at the district level in parentheses. * , (**), [***] indicates significance at the 10, (5), [1] % level. All regressions are performed with Hansen, Heaton and Yaron's (1996) continuously-updated GMM estimator implemented in the Stata command *ivreg2* by Baum, Schaffer and Stillman (2003, 2007) and report the average marginal effects on the observed amount of prime-age unemployment [prime-age non-employment in (5) and (6)]. The delta method is used to compute standard errors. Unless otherwise noted, covariates are the same as in column (5) of Table IV. In (1) the local unemployment at the transition from youth to prime-age for the district of the last apprenticeship spell is used as a control variable; in (2) the minimum local unemployment rate during the early career is used as a control variable; in (3) individuals who are not observed during the last four years of their prime age are excluded; in (4) individuals with more than five years of seasonal employment as defined by Del Bono and Weber (2008) are excluded; in (5) early-career and prime-age non-employment modeled on the definition by Fitzenberger and Wilke (2010) are used instead of early-career and prime-age unemployment; in (6) early-career and prime-age non-employment are modeled on an alternative definition by Schmieder, von Wachter and Bender (2012). Variables included in the regressions of early-career unemployment [early-career non-employment in (5) and (6)] are the same as in the estimates of prime-age unemployment [prime-age non-employment in (5) and (6)]. For variable definitions see Section 3.

TABLE XII

Different Estimates of Sub-periods of Prime-age Unemployment – Mean Regressions.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Model	Tobit-IV	Tobit-IV	Tobit-IV	Tobit-IV	Tobit-IV	Tobit-IV	Tobit-IV	Tobit-IV	Tobit-IV
Years on the labor market	9–16	10–17	11–18	12–19	13–20	14–21	15–22	16–23	17–24
Regressions of sub-periods of prime-age unemployment									
Early-career unemployment	1.07***	1.37***	1.40***	1.48***	1.44***	1.35***	1.29***	1.20***	1.00***
	(0.12)	(0.28)	(0.34)	(0.42)	(0.41)	(0.36)	(0.33)	(0.28)	(0.22)
Unemployment in year 9	—	-1.47**	—	—	—	—	—	—	—
		(0.64)							
Unemployment in years 9–10	—	—	-0.85*	—	—	—	—	—	—
			(0.45)						
Unemployment in years 9–11	—	—	—	-0.67	—	—	—	—	—
				(0.41)					
Unemployment in years 9–12	—	—	—	—	-0.50	—	—	—	—
					(0.34)				
Unemployment in years 9–13	—	—	—	—	—	-0.33	—	—	—
						(0.25)			
Unemployment in years 9–14	—	—	—	—	—	—	-0.22	—	—
							(0.20)		
Unemployment in years 9–15	—	—	—	—	—	—	—	-0.12	—
								(0.15)	
Unemployment in years 9–16	—	—	—	—	—	—	—	—	-0.01
									(0.10)
Regressions of early-career unemployment									
Unemployment at graduation	27.20***	15.09***	12.70***	11.06***	10.60**	10.75**	11.15**	11.54***	12.13***
	(5.56)	(4.21)	(4.17)	(4.20)	(4.19)	(4.22)	(4.41)	(4.43)	(4.38)
Number of observations	739,432	731,178	731,611	732,243	733,130	734,517	735,982	737,597	739,432

Notes: Standard errors clustered at the district level in parentheses. * , (**), [***] indicates significance at the 10, (5), [1] % level. All regressions are calculated with Smith and Blundell's (1986) conditional maximum likelihood Tobit-IV estimator and report the average marginal effects on the observed amount of prime-age unemployment. In all cases the instrument is the local unemployment rate at graduation. Unless otherwise noted, covariates are the same as in column (9) of Table IV. Apart from the instrument, variables included in the regressions of early-career unemployment are the same as in the estimates of prime-age unemployment. For variable definitions see Section 3.

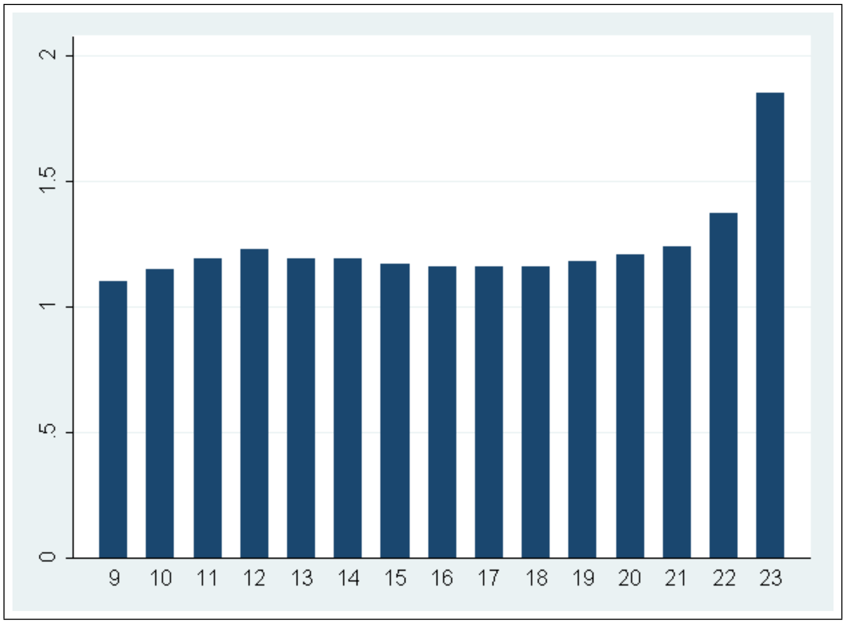

FIGURE 5.— Annual Sample Attrition Rates (in %)

Notes: Annual rates of individuals that disappear from the observable part of the German labor market(in %) by year on the labor market.

TABLE XIII

NUMBER OF YEARS WITH SEASONAL EMPLOYMENT SPELLS

number of years with seasonal employment spells	observations	share of sample in %
0	430,655	47.74
2	211,571	23.45
3	50,697	5.62
4	74,522	8.26
5	36,829	4.08
6	30,043	3.33
7	19,544	2.17
8	13,928	1.54
9	9790	1.09
10 or more	24,551	2.72
total	902,130	100

Notes: *Seasonal employment* denotes two or more employment spells that last for at least two but less then eleven months and end at about the same time in consecutive calender years. For variable definitions see Section 3.