Juvenile Scars and Future Inequality
Assessing the Impact of Youth Unemployment on Earnings Inequality in Germany using Structured Additive Distributional Regression

Alexander Sohn*[1]

[1]Chair of Statistics, University of Göttingen

Version 0.3
Last changes: 30th September 2016

Abstract

In this paper, we analyse the impact of unemployment spells during youth on conditional earnings distributions in Germany in 2013. For the estimation of the conditional earnings distributions, we propose the use of structured additive distributional regression. We consider full-time working, part-time working and unemployed individuals, both male and female.

We find that long-term unemployment during youth exhibits scarring that induces a negative effect on the wage distribution of the individual beyond its mean. Transcending the individual level, a counterfactual elevation of unemployment during youth is shown to have detrimental effects on that generation's earnings distribution, not only in terms of per capita earnings but also in terms of inequality and polarisation measures applied to the aggregate earnings distribution.

JEL-Classification: C13, C21, D31, J24, J31, J42, J45, N34

Keywords: Inequality; Income Distribution; Scarring; Unemployment

1 Introduction

Since the advent of the Great Recession, unemployment rates around the developed world have soared. Young people have been particularly hard hit by the worsening of the labour market situation according to the OECD (2010). Accordingly, the European Commission (2009, p.2) concedes that "life chances of many young people are blighted". In Europe, governments have thus scrambled to avert the creation of a scarred generation (see European Commission, 2009). Yet, while the recent economic literature has made considerable headway on the assessment of

*Corresponding author: asohn@uni-goettingen.de.

youth unemployment on expected labour market outcomes of the affected individuals later in their life, studies analysing the distributional impact in society are scarcely sown.

This paper contemplates the impact of youth unemployment on the earnings for later life. Building on the existing literature, it does not only focus on the conditional expected earnings as the outcome of interest but considers full conditional earnings distributions as well as the aggregate earnings distribution. Thus, we explicitly incorporate the dimensions of risk and inequality to the literature on unemployment scarring. More specifically, we apply structured additive distributional regression to data from the German Socio Economic Panel (SOEP) to estimate the effect of unemployment duration during youth for earning rates at a later point of time in life. The resultant estimates for the full conditional earning rate distributions provide not only an indication of risk/inequality faced by individuals experiencing youth unemployment but also allow for the construction of counterfactual aggregate earning rate distributions in the whole society. It thus allows for an assessment of the impact of youth unemployment on the level of inequality at large.

We find that the detrimental effects of long-term youth unemployment are not constrained to later effects on the expected earnings rate of the affected individual. Their expected utility is blighted even further if one assumes inequality aversion, as the conditional earnings rate distribution shows an adverse effect of long-term unemployment on its inequality, as measured by the Atkinson index. At the aggregate level, we find, using a counterfactual scenario, that increased unemployment (both short-term and long-term) not only leads to a lower expected earnings rate but to higher inequality and polarisation of the aggregate earnings distribution.

The remainder of this paper is structured as follows: In the next section, we briefly look at key contributions in the literature on unemployment scarring. In Section 3, we put forward the underlying economic rationale for considering the distributional impact of unemployment. In the subsequent section, we explain the employed estimation strategy, namely structured additive distributional regression. In Section 5, we assess single conditional earning distributions with respect to the unemployment level during the time of youth while conditioning on other covariates. Then, we use a counterfactual analysis to contemplate the societal impact of rising youth unemployment. In the last section, we conclude.

2 Existing Literature

Scores of publications analyse the effect of unemployment on economic variables of interest and there is a general consensus that unemployment blemishes the affected persons' prospective employment outcomes and thus the economic performance of society at large. Youth unemployment in particular is seen to have long lasting negative effects over the course of the affected person's employment career (see Bell and Blanchflower, 2011).

Concerning economic theory on the matter, the most straight forward explanation for these effects is that unemployment reduces a person's human capital stock, thus lowering that individual's potential productivity (see Pissarides, 1992; Becker, 1993). Another explanation brought forward is that of stigmatisation whereby employers interpret the unemployment as a negative signal (see Kübler and Weizsäcker, 2003; Biewen and Steffes, 2010). As a result, the demand for that person's labour is diminished, decreasing the likelihood of employment and the offered wages. On the supply side, negative effects are expected on the grounds of discouragement and negative habitual effects for those affected (see Clark et al., 2001). When assuming imperfections in the labour market such as imperfect information and search costs, additionally lowered labour supply may be found due to lower incentives to look for suitable jobs as affected individuals may reduce their search efforts in light of the smaller prospective gains that are to be associated with it (see Ljungqvist and Sargent, 1998). From the theoretical literature two possible effects may be deduced: an effect on the chance of finding employment and, given that employment is found, an effect on the wage rate earned.

Plentiful evidence exists on the empirical existence of the effect of unemployment on future employment outcomes. Little surprisingly, most studies are concerned with evidence from the United States (e.g. Kletzer and Fairlie, 2003; Kroft et al., 2013). For Europe, the evidence is most substantial for Britain (e.g. Arulampalam, 2001; Gregg and Tominey, 2005). For continental Europe Cockx and Picchio (2013), Möller and Umkehrer (2014) or Bell and Blanchflower (2015) are exemplary studies which analyse effects of youth unemployment for Belgium, Germany and Greece respectively.

This literature has shed considerable insights onto the workings and magnitude of the effect of youth unemployment. However, the literature focussed on the latter features a somewhat dichotomic streak. In general, the magnitude of the impact of youth unemployment is either considered with regard to its impact on prospects of unemployment or with regard to its impact on prospective wages. And while some papers analyse both effects, sometimes even modelling them jointly, the interpretation of each effect is normally distinctly separated.

For the analysis of unemployment, the analysis is either directed at the probability of unemployment in later life using simple logit/probit (e.g. Arulampalam, 2001)[1] or towards the duration of unemployment using a variety of methods, e.g. quantile regression (e.g. Schmillen and Möller, 2012) or survival analysis (e.g. Cockx and Picchio, 2013). For the analysis of wages, most papers focus on the impact of unemployment on the expected (log-)wage (e.g. Arulampalam, 2001; Gregg and Tominey, 2005). A few select papers have explored some distributional impacts of unemployment,

[1]This estimation is often only done for a Heckman-type correction of the wage estimation and not of interest in itself.

using quantile regression techniques (e.g. Möller and Umkehrer, 2014).

Yet, as far as we know, no paper exists to date which contemplates the effect of unemployment on the full conditional wage distribution as received by individuals. It is such a type of analysis which this paper pursues.

3 Theoretical Considerations concerning the Distributional Impact of Unemployment

In this paper, we generally conceive individuals to be in an imperfect labour market with various sources of friction such as imperfect information, geographic immobility, etc. (see Sohn, forthcoming). In light of these complexities, the labour market outcome cannot be modelled deterministically but only by models featuring which entail stochastic variation. Thus we conceive the observed labour market outcome, y, as a realisation of a random variable, Y.[2] Using observable characteristics, $\mathbf{X} = \mathbf{x}$, we are able to narrow down the stochastic element and explain parts of the variation. However, one must acknowledge that the capacity to explain the variation of the outcome is generally highly limited and that vast swathes of the variation cannot be explained by observable characteristics.

On the basis of this, we would argue that an analysis of labour market outcomes should not be restricted to an analysis focussed on the conditional expectation, denoted $\mathbb{E}(Y_{\mathbf{x}})$, but also ought to contemplate the variation therefrom, which we cannot explain and are thus forced to leave to the residual error term. Thus, rather than solely focussing on the conditional expectations and the variation among them across the covariate space, we propose to contemplate variations of the full conditional distribution $\mathcal{D}_{\mathbf{x}}$. Thus we explicitly contemplate differences between individuals with identical observable characteristics that often fall prey to the informational reduction from the full conditional distribution to a single conditional point measure, like the expectation of the distribution.

This distributional perspective has several ramifications for the analysis of the impact of unemployment, both at the individual/micro-economic level and at the societal/macro-economic level, which will be discussed in the following.

[2] For sake of simplicity, we will restrict ourselves to univariate labour market outcomes. Here we focus on the earnings rate as the only dependent variable. Other aspects like working hours, the hardship of the work, etc. are not considered. For possible methodological inroads to a multivariate analysis, see Klein et al. (forthcoming).

3.1 The micro perspective: The earnings rate distribution and utility

Following van Kerm (2013), let us assume that the work-related utility of individual i can be expressed as a functional of the earnings rate received by that person, y_i. This view is chosen on grounds that the hourly pay commanded by an individual's work effort cannot only be seen as the materialistic reward for the hardship of the work effort but also as a fairly good proxy for the societal esteem associated with it. The pay rate is thus critical for both the ability to draw utility from the purchase of commodities as well as drawing utility from the admiration brought forward from other people. Clearly, this view is a far-reaching oversimplification and deserves to be criticised from several fronts. However, given the elusive nature of utility, we believe it a fair simplification to make, as our focus is to contemplate the impact of youth unemployment which goes beyond the change in the expected earnings rate.

The individual's earnings rate, y_i, is thought of as a realisation of random variable Y_i. In light of lacking further information, this random variable is modelled by $Y_{\mathbf{x}_i}$, that is by a random variable defined for all observably equivalent individuals, with $\mathbf{X} = \mathbf{x}_i$. The view taken here is that the individual does not have full information and is consequently unaware of the exact outcome that is to be expected. Rather the individual's information is limited to the extent that assessment needs to be based on the distribution $\mathcal{D}_{\mathbf{x}_i}$ which governs the outcomes of the random variable $Y_{\mathbf{x}_i}$ from which the individual's labour market outcome is thought to realise. The assessment of the individual's labour market outcome should thus be based on the distribution of $Y_{\mathbf{x}_i}$ rather than the eventual outcome y_i.

From the rich literature on stochastic dominance in income distributions (see among others Moyes, 1999; Cowell, 2000; Chotikapanich and Griffiths, 2006), most reasonable assumptions lead to scenarios where first- and/or second-order distributional dominance critically depends not only on the expectation of $Y_{\mathbf{x}_i}$ but on its full distributional specification. For example, it is reasonable to assume that the relation between the earnings rate and the work-related utility follows a concave relationship as the utility drawn from income used for consumption is assumed to be concave (see Bourguignon, 1989; Ferrer-i Carbonell, 2005; Layard et al., 2008) while the relation between distress (i.e. negative utility) and pay as a key element of the social production function is generally thought to be convex (see Newman, 1999; Ormel et al., 1999; von Scheve et al., 2016). Supposing that an individual is ex-ante unaware of the position within the earnings distribution, it is fair to assume that the individual's assessment would be based on the expected utility derived from the full distribution and not the utility derived from one scalar number highlighting one aspect of the distribution.

More concretely, let us assume the utility function below which follows the framework suggested by Atkinson (1970) yielding constant (relative) risk/ inequality aversion:

$$U(Y_{\mathbf{x}_i}) = C_1 + C_2 \frac{Y_{\mathbf{x}_i}^{1-\epsilon}}{1-\epsilon}, \quad 0 \leq \epsilon < 1, \tag{1}$$

where $Y_{\mathbf{x}_i}$ denotes the earnings rate, conceived as a random variable, of individual i with characteristics $\mathbf{x}_i$ and where ϵ is a parameter representing the degree of risk/inequality aversion and where C_1 and C_2 are simply constants. Note that in terms of the risk/ inequality aversion parameter, we choose a rather strict constraint of $0 \leq \epsilon < 1$.[3] By virtue of this constraint we have an easily interpretable formulation whereby the work related utility is simply added to the utility captured in C_1 thought to represent utility from outside the realm of work related utility.

When assuming such a utility function, the change in the utility an individual can expect ($\Delta\mathbb{E}(U)$) given a change of covariates from $\mathbf{x}_i$ to $\mathbf{x}'_i$ can be represented as the change of the certainty equivalent earnings rate (CEER), i.e.

$$\Delta(\mathbb{E}(U)) \propto \mathcal{C}(Y_{\mathbf{x}'_i}) - \mathcal{C}(Y_{\mathbf{x}_i}) = \mathbb{E}(Y_{\mathbf{x}'_i})(1 - \mathcal{A}(Y_{\mathbf{x}'_i})) - \mathbb{E}(Y_{\mathbf{x}_i})(1 - \mathcal{A}(Y_{\mathbf{x}_i})), \tag{2}$$

where $\mathcal{C}$ denotes the certainty equivalent earnings rate of the earnings rate distribution for the individual, while $\mathbb{E}$ still denotes the expectation operator and $\mathcal{A}$ the Atkinson index for an earnings rate distribution for a given level of risk/ inequality aversion (see Atkinson, 1970).

At the individual level, the analysis of full conditional earnings rate distributions, via their certainty equivalents, allows to assess the change in the expected work-related utility when incorporating the assumption of a concave utility-earnings relationship with constant relative risk/ inequality aversion. Using information solely about the conditional means, this would not be possible. Contemplating full conditional earnings rate distributions thus stands in contrast to and goes beyond the implicitly made, unrealistic assumption of a linear relationship underlying mean-based assessments which are still predominant in the econometric literature on scarring effects.

3.2 The macro perspective: inequality and polarisation

In order to evaluate the effect of youth unemployment for society at large, we consider the impact of a change in the distribution of unemployment for the whole population on three measures of interest. To that end, we compare the observed empirical evidence of the earnings rate with a counterfactually constructed scenario for which we assume all variables to be identical with the exception of the distribution of unemployment. It must be stressed that this comparison does not contemplate general equilibrium effects. As such general equilibrium effects are likely and

[3] It is standard in the literature to consider at least $\epsilon = 1$ and frequently also $\epsilon > 1$.

can safely be assumed to trigger changes in the covariate distribution, the imposed counterfactual scenario of a change in only one covariate must be considered highly problematic. Unfortunately, it is to date infeasible to contemplate general equilibrium effects in the context of constructing counterfactuals (see DiNardo et al., 1996) so that there is little to be done to abate this major caveat in the analysis.
For the comparison between actual and counterfactual distribution, we focus on contrasting three measures of interest for the actual and counterfactual scenario.[4]

As a first measure, we consider the societal welfare out of a consequentialist perspective whereby the societal welfare can be assessed on the basis of the individuals' utility, as specified above.

$$\Delta\mathcal{W} \propto \int\limits_{\mathbf{x}\in\Xi} \mathcal{C}(Y_{\mathbf{x}})s^{cf}(\mathbf{x})d\mathbf{x} - \int\limits_{\mathbf{x}\in\Xi} \mathcal{C}(Y_{\mathbf{x}})s(\mathbf{x})d\mathbf{x}, \tag{3}$$

where $\mathcal{W}$ denotes the societal welfare defined as the expectation of individuals' utility as defined in Equation (1) while $s(\mathbf{x})$ and $s^{cf}(\mathbf{x})$ denote the shares for a given covariate combination $\mathbf{x}$ in the factual and counterfactual scenario respectively. Lastly, Ξ denotes the set of all relevant covariate combinations considered for the comparison. The resultant measure thus gives an indication for the overall change in the individuals' utility induced by a change in the distribution of youth unemployment in the population.

As a second measure, we consider the Atkinson index directly for the aggregate earnings rate distribution. Given the general concern both in the public and the academic sphere on the rise of inequality in Germany and beyond, this measure intends to emphasise on the impact that a rise in youth unemployment can have several years down the line.

The third measure is the level of polarisation. While connected to the concept of inequality, the concept of polarisation focusses more explicitly on the antagonisms within a society which can arise out of changes deemed equalising by measures of inequality which follow the Pigou-Dalton criterion (see Wolfson, 1994; Esteban and Ray, 1994). Thus the consideration of polarisation measures gives another, arguably more direct account of the impact that youth unemployment can have on the disintegration of social cohesion within a society.

[4]Note that we do not decompose the difference between the actual and the counterfactual distribution along the lines of Fortin et al. (2011).

4 Modelling the Distributional Impact of Unemployment

4.1 Estimation of conditional distributions

In order to allow for potential differences in the impact of the covariates on the earnings distributions for males and females, we will run two separate regressions for males and females, with each regression specified as follows.

4.1.1 A generic representation for the predictors

As a general framework, we consider structured additive distributional regression (SADR) (see Klein et al., 2015b) whereby the distribution of the earnings rate, $\mathcal{D}_{\mathbf{x}}$, is conditioned on a set of covariates, $\mathbf{X} = \mathbf{x}$. The conditional distribution is assumed to follow a parametric distribution. Thus, the conditional distribution can be written in the form $\mathcal{D}(\theta_1(\mathbf{x}), \ldots, \theta_K(\mathbf{x}))$, where $\theta_k(\mathbf{x})$ is the k-th parameter in the parametric distribution and is conditioned on the covariate combination of the specific stratum. For notational brevity, we will drop the suffix $(\mathbf{x})$ in the following. Additionally, we define $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_K)'$.

In this paper, we use the following generic representation for every parameter of the distribution: Each parameter θ_k, for $k \in 1, \ldots, K$, can be linked to a structured additive predictor, η^{θ_k}, via a suitably specified link function, g_k, mapping the predictor into the parameter space such that $\theta_k = g_k^{-1}(\eta^{\theta_k})$. The predictor η^{θ_k} can be specified in the following form:

$$\eta^{\theta_k} = \beta_0^{\theta_k} + f_1^{\theta_k}(\mathbf{x}) + \ldots + f_{J_k}^{\theta_k}(\mathbf{x}), \tag{4}$$

where $\beta_0^{\theta_k}$ represents the intercept of the predictor and the functions $f_j^{\theta_k}(\mathbf{x})$, $j = 1, \ldots, J_k$, can capture both linear and non-linear effects of single or multiple elements of the covariate vector $\mathbf{x}$. The latter is done by means of representing the function by a suitable linear combination of basis functions which are generally penalised such that the non-parametric estimate adheres to the required smoothness (see Fahrmeir et al., 2013).

In our application we will use the following regression set-up for all the predictors:

$$\begin{aligned} \eta^{\theta_k} = \beta_0^{\theta_k} + \beta_1^{\theta_k} stunem + \beta_2^{\theta_k} ltunem + \beta_3^{\theta_k} nat + \beta_4^{\theta_k} educ_2 + \beta_5^{\theta_k} educ_3 + \beta_6^{\theta_k} educ_4 \\ + f_1^{\theta_k}(age) + heduc \cdot f_2^{\theta_k}(age) + f_4^{\theta_k}(region), \end{aligned} \tag{5}$$

where *stunem* and *ltunem* are both binary, denoting short-term and long-term unemployment experience during youth, respectively. When the individual was out of employment for some time but less than one year during the time of youth, *stunem* is set to unity and is zero otherwise. If the individual spent one year or more in unemployment the variable, *ltunem* is unity and is zero otherwise. Next to these main effects of interest, we control for some standard variables commonly used in the labour market literature. We control for nationality by means of the binary variable *nat* which is unity if the person has German nationality and zero otherwise. Moreover, we use education-specific intercepts by three binary variables $educ_e$, where the subscript e denotes the education level, with the first education level taken as the base. Age, as a proxy for potential experience, is modelled by a flexible smooth function, based on P-splines (see Eilers and Marx, 1996; Brezger and Lang, 2006). Thereby, f generally consists of a number of basis functions allowing for a high degree of flexibility. To ensure smoothness, we impose a Gaussian prior on the splines' coefficients. To account for different developments over the lifespan depending on whether the person has enjoyed higher education and interaction with the binary variable *heduc* which is unity if the person has a degree in higher education, i.e. if the ISCED level is 6 according to the SOEP, and zero otherwise. In order to capture differences between the economic dynamism of different federal states in Germany, we include a hierarchical spatial effect, such that:

$$f_4^{\theta_k}(region) = \beta_7 east + \gamma_{region}, \tag{6}$$

where *east* is a binary variable that is unity if the federal state is situated in the East, thus capturing the difference between the former Federal Republic of Germany and the German Democratic Republic. The state-specific random effect is denoted γ_{region} and accounts for variations across individual states. For the random effect, we impose a Gaussian prior centred around zero, with the variance determined by a standard inverse gamma hyperprior.

Estimation is performed in a Bayesian framework using Markov Chain Monte Carlo (MCMC) techniques implemented in the software BayesX (Belitz et al., 2015). See Klein et al. (2015a) for details on the estimation procedure.

Concerning the coefficients' prior distributions, we employ non-informative flat priors for the linear effects. For the basis functions of the smooth effects, we use multivariate normal distributions for the basis function coefficient. The scale of their precision matrix determines the degree to which the desired property, i.e. smoothness of deviations from the linear effect, is enforced. This scale is in turn regulated by scale-dependent hyper priors as introduced in Klein et al. (forthcoming). This representation has the advantage over the conventional use of inverse gamma distributed hyper priors that it doesn't suffer from the theoretical problem that the inverse gamma prior puts zero density on the base model (see Simpson et al., 2015). For the regional random effect, we also use scale dependent priors which are now applied to the independent normal distributions applied to penalise the state-specific random effects deviating from the linear *east* effect. The priors are thus chosen in such a way that the deviation from a simple linear set-up is generally feasible but penalised where required. See Section A.2 in the appendix for further information on the prior specification.

Concerning the MCMC algorithm, we use 200,000 MCMC realisations for burn-in and thin out the following 800,000 MCMC realisations by a factor of 800. Thus, we use 1,000 MCMC realisations for each predictor η^{θ_k} to construct the posterior of the predictor, which is then transformed by the corresponding link function g_k into the posterior distribution of the parameter of interest. These distributions are proper under mild conditions (see Klein et al., 2015b). For our inferential purposes, we use the median from the posterior as point estimate for the parameters in order to provide estimates for the resultant full conditional distribution for the desired covariate combination. Using the MCMC realisations, we also provide point-wise credible intervals giving a notion of uncertainty attached to the estimators.

4.1.2 Estimating the conditional mixture distribution

For the conditional earnings rate distribution, $\mathcal{D}_{\mathbf{x}}$, we use a Type II Dagum mixture distribution taking the following form:

$$p(y \mid \pi, a, b, c) = \pi \mathbb{1}_{\{y=0\}} + (1-\pi) p_+(y \mid a, b, c), \tag{7}$$

where the probability of zero earnings rate/unemployment is given by π, with $\mathbb{1}_{\{y=0\}}$ denoting an indicator function which is unity for a zero earnings rate/unemployment. For earnings greater than zero, we obtain the probability density specified by a Type I Dagum distribution, denoted $p_+(y \mid a, b, c)$.

The model set-up is elaborated in Section A.3 the appendix. The parametric specification induced by this set-up is tested by means of an adaptation of a misspecification test proposed by Rothe and Wied (2013) for testing conditional distributions and found to perform adequately. See Section A.6 in the appendix for details.

4.2 Constructing counterfactual aggregate distributions

The construction of counterfactual distributions for analyses at the macro level is standard in the literature and numerous methodological approaches have been made.

Following the idea of the construction of the aggregate earning distribution from the conditional earning distributions estimated as in Equation (8), we estimate the actual aggregate earning distribution, $p_{agg}^{act}(y)$, as follows:

$$p_{agg}^{act}(y) = \int_{\mathbf{x}\in\Xi} p_{\mathbf{x}}(y) s^{act}(\mathbf{x}) d\mathbf{x}, \tag{8}$$

where $s^{act}(\mathbf{x})$ denotes the actual share of the covariate combination $\mathbf{x}$ in the population as observed. Note that Ξ not only encapsulates all existing combinations of $\mathbf{x}$ but can (and does) also entail covariate combinations that are not observed in the sample.[5] Consequently the counterfactual aggregate earning distribution, $p_{agg}^{cf}(y)$, can be defined analogously:

$$p_{agg}^{cf}(y) = \int_{\mathbf{x}\in\Xi} p_{\mathbf{x}}(y) s^{cf}(\mathbf{x}) d\mathbf{x}, \tag{9}$$

where $s^{cf}(\mathbf{x})$ is the share of the covariate combination $\mathbf{x}$ in the counterfactual scenario. Concretely, we consider a scenario of youth unemployment rising to levels similar to those in the hard hit economies in Southern Europe for a ten-year period.

Before we go on to look at results for Germany using this estimation strategy, it should be stressed that one needs to be very wary drawing causal inferences from the distributional models presented here. Rather, these models are intended to give an indication on the enlargement of the magnitude of the scarring effect, when considering the dimension of inequality and risk.

5 Results for Germany

5.1 Data

We use the German Socio-Economic Panel (SOEP) database (see SOEP, 2014) as our primary source of data. The SOEP is a representative panel study of private households in Germany. The data base contains a wide range of socio-economic information on its household members, including data on income and unemployment experiences.

In order to allow for a comprehensive assessment of the impact of youth unemployment and the construction of counterfactuals for the whole population, this paper goes beyond the restriction commonly found in other studies of focussing solely on unemployment effects for (mostly male) full-time workers. In this study, we contemplate the impact of unemployment for both men and

[5] Thus we are not forced to constrain the covariate space of the counterfactual scenario to that of the limits of the covariate space of the observed scenario. In other words, we don't necessarily require the overlap of the covariate spaces of the factual and the counterfactual scenario, as would effectively be the case of re-weighting approaches following DiNardo et al. (1996). This is made feasible by the possibility of additive regression models to interpolate and extrapolate in a highly dimensional covariate space. Naturally, this makes the highly contentious assumption of a adequately specified regression model.

women. Moreover, we do not restrict the sample to individuals working full-time but include all individuals, including part-time and unemployed individuals in our sample. More concretely, the sampling universe is all individuals sampled by the SOEP between 25 and 60 years of age for whom we could retrieve information on the dependent and independent variables (see below). This specification yields 8,344 observations (3,780 males and 4,564 females) in total.

As our dependent variable of interest we consider the personal hourly earnings rate which we define by dividing the gross market monthly labour earnings in the year 2013 by the number of actual hours worked. For individuals for whom we cannot observe any wage rate due to unemployment, we do not follow the bulk of the literature on scarring effects by attempting to estimate the job offers at disposal to the individual. Since our focus is first and foremost directed at the issue of inequality with respect to youth unemployment, an approach that considers the actual market labour outcomes rather than only its hypothetical demand side is deemed more appropriate. Therefore the unemployed are assumed to have an hourly wage rate of or close to zero Euros per hour. In terms of the welfare assessment we therefore assume that people in unemployment only have the baseline utility of C_1 from Equation (1) and do not get additional utility from pursuing an occupation and earning money. Additional utility from the freedom dispensed by the unemployment status is thus not considered. The underlying view is that in general the freedom of time afforded by unemployment does not yield utility (see Spencer, 2009).

The main covariate of interest we consider is the total duration of unemployment spells during the time of youth. Youth is defined as by the International Labour Office such that it covers all persons aged from 15 to 24 years (ILO, 2014). In addition, we control for age, education, nationality and the region of residence. Following the classical Mincer equation set-up, age is accounted for in a continuous manner and thought to control for potential labour market experience while education is the second component accounting for the human capital of an individual. The latter is not modelled continuously but by four discrete education levels based on grounds of the ISCED97 classification. Concerning the nationality, we use a binary categorisation of individuals with and without German nationality respectively. Lastly, we consider the federal state of residence of any individual to account for spatial differences in economic dynamism across Germany.

For more information on the variables see Section A.1 in the appendix.

5.2 The distributional impact of unemployment at the micro level

The conditional earning distributions which we estimate portray a structure whereby most measures of interest depend on various parameters. This set-up induces a relationship between the covariates and the measure of interest which does generally not allow for a straight forward ceterus paribus (c.p.) effect type of interpretation that is independent of the other covariate values. Even for very simple measures, like the unemployment rate which is directly related to the parameter π, we find that the logit-link leads to intertwining effects of the covariates.

For the interpretation, we thus use effect displays as suggested by Fox (1987) for generalised linear models. Thereby, the effect of one covariate is displayed keeping the other covariates fixed at "typical values". Here, we will concentrate on the effect of the three different unemployment experiences during the time of youth both for men and women. The person considered has German nationality, the age is fixed at the average age for the economically active population of 43 years, while we will consider the second education level (upper secondary and post-secondary but non-tertiary education) which is still the dominant education level in Germany. Concerning the region, we will consider the state of Lower Saxony, that can be seen as an average state in Germany with regard to per capita income, as well as population size. The wage distributions which we consider can thus be seen as those of a representative man and a representative woman with varying degrees of unemployment over youth. For other covariate combinations see Section A.5 in the appendix.

Figure 1 displays the changes of the estimated conditional earning distribution. The solid black line marks the conditional wage distribution for a person without any unemployment experience, while the pink dashed line denotes the distribution for a person with some short-term unemployment experience totalling less than one year. Lastly, the red dashed-dotted line yields the distribution of a person with long-term unemployment experiences during youth. The point with a bar on the left indicates the probability of a zero-income, with the corresponding scale on the left hand side of the graph. For incomes greater than zero, we provide a density with the corresponding scale on the right.
One immediately obvious observations is that long-term unemployment during the time of youth has severe negative implications for the chance of employment, with the probability of unemployment / zero earnings much higher for both men and women. Moreover, when considering the distribution above zero, one can see that the resultant distribution is also shifted to the left and thus features a lower expectation. Other aspects of the distribution above zero show a more promiscuous pattern with the male distribution featuring a lower coefficient of variation but a higher skewness, while for the women it is vice versa. All these higher order moment effects are insignificant at the usual levels, though.

In Table 1 we display some distributional measures of the portrayed distributions, where we consider the whole observed distribution again, i.e. we include zero earnings in the computation for all measures.

The most noticeable difference is change in the probability of unemployment. For the men considered, the probability increases from 9% for those without any unemployment during youth to

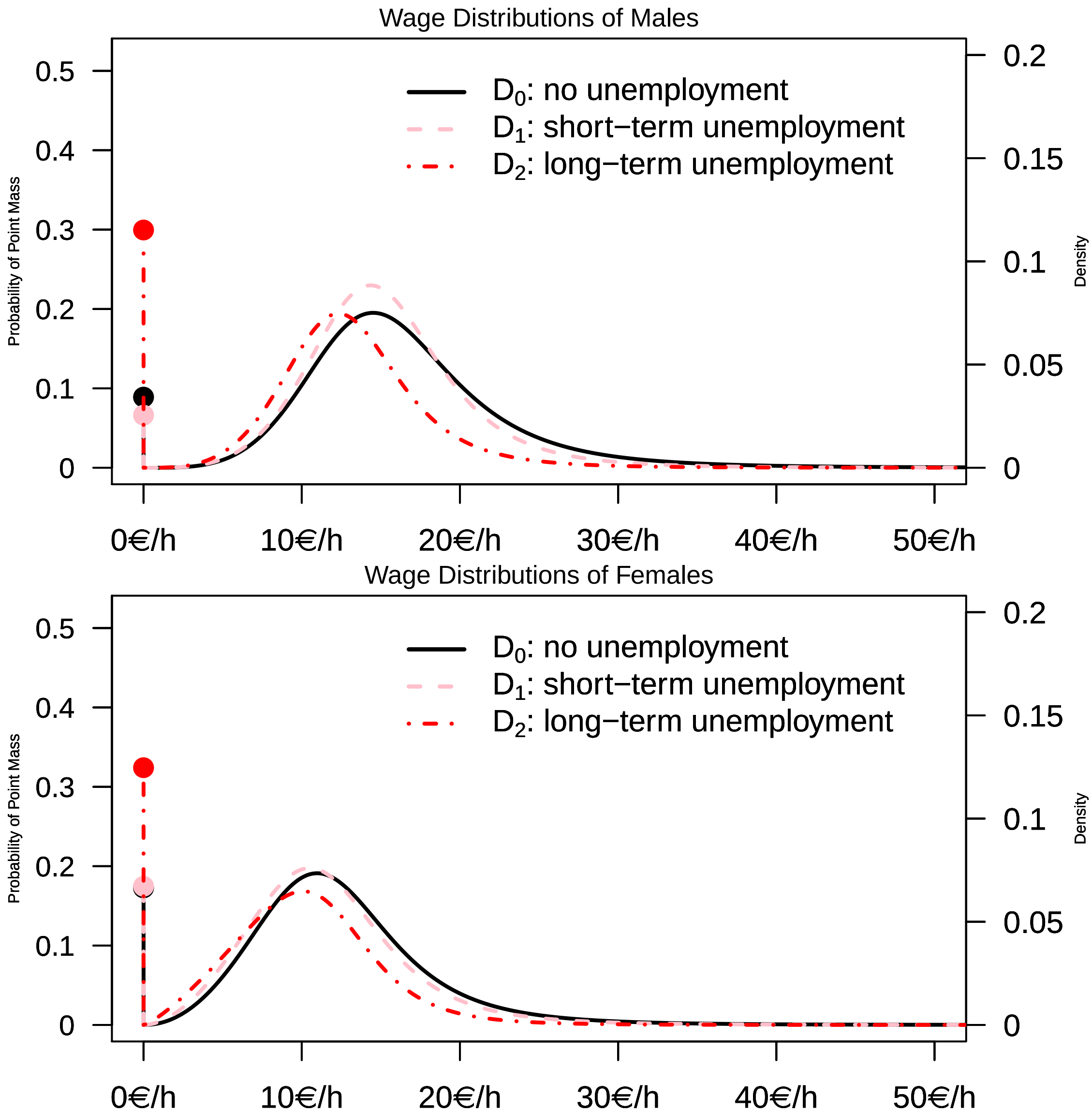

Figure 1: Distribution estimates of German citizens with medium education living in Lower Saxony.

30% for those with long-term unemployment. For the women, the increase is from 17% to 32%. Interestingly, the men with short-term unemployment feature the same or even lower unemployment probabilities later in life than those without any unemployment during youth, although these results are not significant at the 5% level. The results thus concur with the literature which indicates that short-term unemployment, especially among young individuals, is not necessarily scarring but can be seen first and foremost as frictional unemployment as young individuals look for the "true calling" (see Gervais et al., 2014). By contrast, long-term unemployment seems to exert heavy scarring, as the literature would leave us to expect (see among others Arulampalam et al., 2000; Schmillen and Möller, 2012; Eriksson and Rooth, 2014).

Concerning the expected wage rate, we can observe that both for men and women, long-term

	D_{0m}	D_{1m}	D_{2m}	D_{0f}	D_{1f}	D_{2f}
$\pi_{\mathbf{x}}$	0.09[0.06; 0.12]	0.07[0.05; 0.09]	0.30[0.23;0.38]	0.17[0.13; 0.23]	0.17[0.13; 0.23]	0.32[0.26;0.40]
$\mu_{\mathbf{x}}$	13.27[11.71;14.95]	13.09[11.61;14.62]	6.41[5.12;7.84]	10.00[8.97;11.02]	9.32[8.23;10.46]	6.59[5.51;7.82]
$\mathcal{A}_{\mathbf{x}}(\epsilon{=}0.1)$	0.03[0.02; 0.04]	0.02[0.02; 0.03]	0.08[0.07;0.10]	0.03[0.03; 0.04]	0.03[0.03; 0.04]	0.06[0.04;0.07]
$\mathcal{A}_{\mathbf{x}}(\epsilon{=}0.5)$	0.22[0.17; 0.27]	0.17[0.13; 0.22]	0.50[0.42;0.59]	0.22[0.18; 0.27]	0.23[0.18; 0.28]	0.37[0.30;0.44]
$\mathcal{A}_{\mathbf{x}}(\epsilon{=}0.9)$	0.61[0.51; 0.71]	0.51[0.41; 0.61]	0.91[0.86;0.94]	0.59[0.50; 0.69]	0.60[0.50; 0.69]	0.80[0.72;0.86]
$\mathcal{C}_{\mathbf{x}}(\epsilon{=}0.1)$	12.87[11.29;14.50]	12.77[11.31;14.28]	5.88[4.60;7.27]	9.67[8.63;10.67]	9.01[7.94;10.12]	6.21[5.13;7.43]
$\mathcal{C}_{\mathbf{x}}(\epsilon{=}0.5)$	10.39[8.64;12.07]	10.88[9.28;12.36]	3.20[2.14;4.39]	7.78[6.60; 8.83]	7.23[5.98; 8.37]	4.16[3.16;5.24]
$\mathcal{C}_{\mathbf{x}}(\epsilon{=}0.9)$	5.24[3.58; 7.05]	6.47[4.71; 8.20]	0.59[0.31;1.05]	4.05[2.88; 5.26]	3.75[2.63; 4.88]	1.33[0.81;2.04]

Table 1: Some distribution measures for the conditional wage distributions

youth unemployment lowers the expected wage rate for the set of individuals considered here, while the effect for short-term unemployment is rather timid and not statistically significant at the 5% level. For men, the mean wage falls by 52% from 13.27€ to 6.41€ when comparing no youth unemployment with long-term unemployment experiences. For women, we see a decrease by 34% from 10.00€ to 6.59€. Again evidence of this is plentiful in the literature already.

Looking at additional attributes of the distribution, we can observe that the effect of unemployment is not restricted to the expected earnings but that the inequality of the conditional earnings distribution, as measured by the Atkinson index, is significantly greater for those individuals with a long-term unemployment experience. Depending on the level of aversion, the inequality increases between 49% ($\epsilon = 0.9$) and 160% ($\epsilon = 0.1$) for men and between 35% ($\epsilon = 0.9$) and 100% ($\epsilon = 0.1$) for women.

Concerning the effects for the certainty equivalent earnings rate, we see a decrease between 54% ($\epsilon = 0.1$) and 88% ($\epsilon = 0.9$) for men and 36% ($\epsilon = 0.1$) and 67% ($\epsilon = 0.9$) for women when comparing no unemployment with long-term unemployment experience. The effects on the expected utility of long-term unemployment are thus exacerbated if considering inequality aversion in the assessment.

For the assessment of the adequacy of our parametric specification by structured additive distributional regression we use the misspecification test specified in Section A.6. Using 1,000 bootstrap repetitions, we consider the adequacy both for males and for females and obtain the following results. For men, we obtain a p-value 0.34. For women, the p-value is 0.13. As we do not reject the hypothesis of a parametric misspecification at all usual levels, we consider the approximation by our parametric fit of the conditional earning distributions to be adequate.

5.3 The distributional impact of unemployment at the macro level

In this section, we look at a coarsely constructed counterfactual scenario in which Germany's fortunes are assumed to have gone south after the years of unification similar to the economic problems encountered by many South-European countries in the past years. Specifically, we consider the effects on the aggregate distribution in 2013 given hypothetical change in characteristics for one cohort, i.e. those in their youth in 1990, assuming that all other attributes of the covariate distribution stay the same. For this counterfactual scenario, we simply assume that individuals in the SOEP born between 1967 and 1976 faced a youth unemployment situation similar to those in their youth in Greece in 2013, with a 28% risk of short-term unemployment and a 30% risk of long-term unemployment. See Section A.1.2 in the appendix for details.

	Factual	Counterfactual	Difference
μ_{agg}	13.241[12.620;13.851]	12.764[12.170;13.381]	−0.478[−0.571;−0.373]
$\mathcal{C}_{agg}(\epsilon{=}0.1)$	12.893[12.270;13.500]	12.405[11.781;13.014]	−0.486[−0.575;−0.387]
$\mathcal{C}_{agg}(\epsilon{=}0.5)$	10.899[10.182;11.548]	10.359[9.651;11.015]	−0.529[−0.602;−0.454]
$\mathcal{C}_{agg}(\epsilon{=}0.9)$	4.804[4.145; 5.576]	4.342[3.728; 5.090]	−0.457[−0.501;−0.412]
$\mathcal{A}_{agg}(\epsilon{=}0.1)$	0.036[0.034; 0.040]	0.039[0.036; 0.042]	0.002[0.002; 0.003]
$\mathcal{A}_{agg}(\epsilon{=}0.5)$	0.239[0.223; 0.262]	0.254[0.236; 0.278]	0.015[0.013; 0.017]
$\mathcal{A}_{agg}(\epsilon{=}0.9)$	0.851[0.822; 0.887]	0.875[0.848; 0.908]	0.024[0.020; 0.026]
$\mathcal{P}_{agg}(\alpha{=}1, \beta{=}1)$	0.085[0.065; 0.106]	0.098[0.079; 0.119]	0.013[0.010; 0.016]

Table 2: Some measures for the macro effects

For the aggregate expected earnings rate, we see a 3.6%[2.8%;4.3%] decrease, when contrasting the factual with the counterfactual scenario. The observed effect is more pronounced when considering the effect on the certainty equivalents, with a decrease ranging from 3.8% [3.0%;4.4%] to 9.6% [8.6%;10.3%], for $\epsilon = 0.1$ and $\epsilon = 0.9$ respectively. For $\epsilon = 0.9$, we thus observe a relative difference of more than twice the size than the difference observed when only considering the expected earnings rate. As one can see from the credible intervals, this difference can be considered significant at the 5% level. In other words, the magnitude of the effect of a counterfactually change is affected significantly by the consideration of risk/ inequality aversion at the disaggregated level in the assessment.

Shifting the perspective to inequality measures for the aggregate earnings distribution, we see an increase of 6.2% [5.7%;6.9%] for $\epsilon = 0.5$. For $\epsilon = 0.1$ the increase stands at 5.8% [4.9%;6.8%] while for $\epsilon = 0.9$ the increase is much smaller at 2.9% [2.3%;3.2%]. We thus find that unemployment during youth does not only affect the earnings outcomes at the individual level but also increases inequality at the aggregate level. Additionally considering a measure of polarisation, which is found to provide another important indicator for the degree to "potential (if not open) social conflict" (Esteban and Ray, 1994, p.821). This measure increases much more dramatically in relative terms with the polarisation from Esteban et al. (2007). Taking $\alpha = \beta = 1$ (based on Gradín, 2000), we find a relative increase of 15.5% [15.5%;21.5%].
Hence, we find that youth unemployment seems to increase inequality and polarisation in a society at later points in time. Given the increased political hazards that are associated with higher levels

of inequality and polarisation[6], this paper argues that a comprehensive assessment of the effects of youth unemployment should give equal weight to the relation with future inequality as to the assessment of the welfare measures from above, despite the greater difficulties involved.

6 Conclusion

In this paper, we have analysed the relation between the earnings rate in later life with respect to unemployment during youth, using structured additive distributional regression to control for other characteristics like age, education, etc.
At the microeconomic level, we find that an effect of long-term unemployment not only blights the expected earnings rate of affected individuals. Going beyond a mean-based analysis, we see that the effect of long-term unemployment during youth is aggravated. When considering the inequality in the conditional earnings distributions, as measured by the Atkinson index, the impact of unemployment increases. Put differently, the relative change in expected utility is significantly above the relative change in expected earnings if we assume risk/inequality aversion.

In order to assess the magnitude of the effect of youth unemployment at the macro economic level, we consider a counterfactual scenario of heightened youth unemployment levels in Germany, akin to those currently observed in Greece. We find that such a shock is likely to dimish the expected earnings rate at the aggregate level and the expected welfare measures used based on inequality aversion. Additionally, we find that the counterfactual scenario of youth unemployment also increases aggregate inequality, again measured by the Atkinson index, and polarisation, as measured by the polarisation concept put forward by Esteban et al. (2007).

In 1973, John Richard Hicks remarked that contemporary wage structures are “the fossilised remains of historical shortages”(Hicks, 1973, p.320). This paper indicates that the unemployment experiences of current European youth may well feature as fossilised remnants in the inequalities of earnings rates in the future. While much work remains to be done on the issue, the findings of this paper indicate that an analysis that solely judges the magnitude of the scarring of unemployment on the basis of conditional mean estimates potentially neglects important detrimental effects occurring both at the micro- and the macro level. A more thorough analysis of the effects of youth unemployment, that also considers the distributional effects is thus warranted.

[6] As Amartya Sen (1973, p.1) notes: “the relation between inequality and rebellion is a close one”.

References

D. W. K. Andrews (1997): A Conditional Kolmogorov Test, in: Econometrica, 65(5), pp. 1097–1128.

W. Arulampalam (2001): Is Unemployment Really Scarring? Effects of Unemployment Experiences on Wages, in: Economic Journal, 111(475), pp. 585–606.

W. Arulampalam, A. L. Booth and M. Taylor (2000): Unemployment persistence, in: Oxford Economic Papers, 52(1), pp. 24–50.

A. B. Atkinson (1970): On the Measurement of Economic Inequality, in: Journal of Economic Theory, 2(3), pp. 244–263.

G. S. Becker (1993): Human capital: A theoretical and empirical analysis, with special reference to education, University of Chicago Press, Chicago, 3rd ed.

C. Belitz, A. Brezger, N. Klein, T. Kneib, S. Lang and N. Umlauf (2015): BayesX - Software for Bayesian Inference in Structured Additive Regression Models: Version 3.0, URL `http://www.bayesx.org`.

D. N. F. Bell and D. G. Blanchflower (2011): Young People and the Great Recession, in: Oxford Review of Economic Policy, 27(2), pp. 241–267.

— (2015): Youth unemployment in Greece: Measuring the challenge, in: IZA Journal of European Labor Studies, 4(1), p. 1.

M. Biewen and S. P. Jenkins (2005): A Framework for the Decomposition of Poverty Differences with an Application to Poverty Differences Between Countries, in: Empirical Economics, 30(2), pp. 331–358.

M. Biewen and S. Steffes (2010): Unemployment Persistence: Is the Evidence for Stigma Effects?, in: Economics Letters, 106(3), pp. 188–190.

F. Bourguignon (1989): Family Size and Social Utility, in: Journal of Econometrics, 42(1), pp. 67–80.

A. Brezger and S. Lang (2006): Generalized Structured Additive Regression Based on Bayesian P-splines, in: Computational Statistics & Data Analysis, 50(4), pp. 967–991.

D. Chotikapanich (ed.) (2008): Modeling Income Distributions and Lorenz Curves, Springer, New York.

D. Chotikapanich and W. E. Griffiths (2006): Bayesian Assessment of Lorenz and Stochastic

Dominance in Income Distributions, in: J. Creedy and G. Kalb (eds.), Dynamics of Inequality and Poverty, pp. 297–321, Elsevier, Amsterdam.

A. E. Clark, Y. Georgellis and P. Sanfrey (2001): Scarring: The Psychological Impact of Past Unemployment, in: Economica, 68(270), pp. 221–241.

B. Cockx and M. Picchio (2013): Scarring Effects of Remaining Unemployed for Long-term Unemployed School-leavers, in: Journal of the Royal Statistical Society: Series A (General), 176(4), pp. 951–980.

F. A. Cowell (2000): Measurement of Inequality, in: A. B. Atkinson and F. Bourguignon (eds.), Handbook of Income Distribution, pp. 87–166, Elsevier, Amsterdam.

J. DiNardo, N. M. Fortin and T. Lemieux (1996): Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach, in: Econometrica, 64(5), pp. 1001–1044.

P. H. C. Eilers and B. D. Marx (1996): Flexible Smoothing with B-splines and Penalties, in: Statistical Science, 11(2), pp. 89–102.

S. Eriksson and D.-O. Rooth (2014): Do Employers Use Unemployment as a Sorting Criterion When Hiring? Evidence from a Field Experiment, in: American Economic Review, 104(3), pp. 1014–1039.

J. Esteban, C. Gradín and D. Ray (2007): Extensions of a Measure of Polarization, with an Application to the Income Distribution of five OECD Countries, in: Journal of Economic Inequality, 5(1), pp. 1–19.

J. Esteban and D. Ray (1994): On the Measurement of Polarization, in: Econometrica, 62(4), pp. 819–851.

European Commission (2009): European Union's Strategy for Youth – Investing and Empowering, URL http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52009DC0200&from=EN.

Eurostat (2016): Unemployment by sex and age - annual average, URL http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=une_rt_a&lang=en.

L. Fahrmeir, T. Kneib, S. Lang and B. D. Marx (2013): Regression: Models, Methods and Applications, Springer, Berlin.

A. Ferrer-i Carbonell (2005): Income and Well-Being: An Empirical Analysis of the Comparison Income Effect, in: Journal of Public Economics, 89, pp. 997–1019.

L. Flabbi (2010): Gender Discrimination Estimation in a Search Model with Matching and Bargaining, in: International Economic Review, 51(3), pp. 745–783.

N. M. Fortin, T. Lemieux and S. Firpo (2011): Decomposition Methods in Economics, in: O. Ashenfelter and D. E. Card (eds.), Handbook of Labor Economics, Vol. 4A, pp. 1–102, North-Holland, Amsterdam.

J. Fox (1987): Effect Displays for Generalized Linear Models, in: C. C. Clogg (ed.), Sociological Methodology, vol. 17, pp. 347–361, American Sociological Association, Washington D.C.

M. Gervais, N. Jaimovich, Siu, Henry, E and Y. Yedid-Levi (2014): What Should I Be When I Grow Up? Occupations and Unemployment over the Life Cycle, in: NBER Working Paper, No. 20628.

C. Gradín (2000): Polarization by Sub-populations in Spain, 1973-91, in: Review of Income and Wealth, 46(4), pp. 457–474.

P. Gregg and E. Tominey (2005): The Wage scar from male youth unemployment, in: Labour Economics, 12(4), pp. 487–509.

Hellenic Statistical Authority (2016): Unemployed persons of 15 years and over by age groups, duration of unemployment and sex, URL http://www.statistics.gr/.

J. R. Hicks (1973): The Theory of Wages, Macmillan, London, 2nd ed.

ILO (2014): Key Indicators of the Labour Market, International Labour Organisation, Geneva, URL http://www.ilo.org/empelm/what/WCMS_114240/lang--en/index.htm.

C. Kleiber and S. Kotz (2003): Statistical Size Distributions in Economics and Actuarial Sciences, Wiley, Hoboken.

N. Klein, T. Kneib, S. Klasen and S. Lang (forthcoming): Bayesian Structured Additive Distributional Regression for Multivariate Responses, in: Journal of the Royal Statistical Society: Series C (Applied Statistics).

N. Klein, T. Kneib and S. Lang (2015a): Bayesian Generalized Additive Models for Location, Scale and Shape for Zero-Inflated and Over-Dispersed Count Data, in: Journal of the American Statistical Association, 110(509), pp. 405–419.

N. Klein, T. Kneib, S. Lang and A. Sohn (2015b): Bayesian Structured Additive Distributional Regression with an Application to Regional Income Inequality in Germany, in: Annals of Applied Statistics, 9(2), pp. 1024–1052.

L. G. Kletzer and R. W. Fairlie (2003): The Long-Term Costs of Job Displacement for Young Adult Workers, in: Industrial and Labor Relations Review, 54(4), pp. 682–698.

K. Kroft, F. Lange and M. J. Notowidigdo (2013): Duration Dependence and Labor Market

Conditions: Evidence from a Field Experiment, in: The Quarterly Journal of Economics, 128(3), pp. 1123–1167.

D. Kübler and G. Weizsäcker (2003): Information Cascades and the Labor Market, in: Journal of Economics, 80(3), pp. 211–229.

R. Layard, G. Mayraz and S. Nickell (2008): The Marginal Utility of Income, in: Journal of Public Economics, 92(8), pp. 1846–1857.

L. Ljungqvist and T. J. Sargent (1998): The European Unemployment Dilemma, in: Journal of Political Economy, 106(3), pp. 514–550.

J. Möller and M. Umkehrer (2014): Are there Long-Term Earnings Scars from Youth Unemployment in Germany?, in: ZEW Discussion Papers, No. 14-089.

P. Moyes (1999): Stochastic Dominance and the Lorenz Curve, in: J. Silber (ed.), Handbook of Income Inequality Measurement, Kluwer Academic, Boston.

K. S. Newman (1999): Falling from Grace: Downward Mobility in teh Age of Affluence, University of California Press, Berkely.

OECD (2010): Off to a Good Start? Jobs for Youth, OECD Publishing, Paris.

J. Ormel, S. Lindenberg, N. Steverink, Verbrugge Lois M, J. Ormel and L. M. Verbrugge (1999): Subjective Well-Being and Sociat Production Function Theory, in: Social Indicators Research, 46(1), pp. 61–90.

C. Pissarides (1992): Loss of Skill during Unemployment and the Persistence of Unemployment Shocks, in: Quarterly Journal of Economics, 107, pp. 1371–1391.

C. Rothe and D. Wied (2013): Misspecification Testing in a Class of Conditional Distributional Models, in: Journal of the American Statistical Association, 108(501), pp. 314–324.

A. Schmillen and J. Möller (2012): Distribution and determinants of lifetime unemployment, in: Labour Economics, 19(1), pp. 33–47.

A. K. Sen (1973): On Economic Inequality, Oxford University Press, Oxford.

D. P. Simpson, H. Rue, T. G. Martins, A. Riebler and S. H. Sørbye (2015): Penalising model component complexity: A principled, practical approach to constructing priors, in: arXiv: 1403.4630, v4, URL `http://arxiv.org/abs/1403.4630`.

SOEP (2014): Socio-oeconomic Panel Study (SOEP), data of the years 1984-2013, version 30, doi: 10.5684/soep.v30.

A. Sohn (2015): Beyond Conventional Wage Discrimination Analysis. Assessing Comprehensive Wage Distributions of Males and Females using Structured Additive Distributional Regression, in: SOEPpapers on Multidisciplinary Panel Data Research, No. 802.

— (forthcoming): The Regression of an Equitable Market Economy: Analysing Income Inequality in Germany From the Bottom Up Using Structured Additive Distributional Regression, Ph.D. thesis.

A. Sohn, N. Klein and T. Kneib (2015): A Semiparametric Analysis of Conditional Income Distributions, in: Schmollers Jahrbuch, 135(1), pp. 13–22.

D. A. Spencer (2009): The Political Economy of Work, Routledge, London.

P. van Kerm (2013): Generalized Measures of Wage Differentials, in: Empirical Economics, 45(1), pp. 465–482.

C. von Scheve, F. Esche and J. Schupp (2016): The Emotional Timeline of Unemployment: Anticipation, Reaction, and Adaptation, in: Journal of Happiness Studies.

G. G. Wagner, J. R. Frick and J. Schupp (2007): The German Socio-Economic Panel Study (SOEP) - Scope Evolution and Enhancements, in: Schmollers Jahrbuch, 127(1), pp. 139–169.

M. C. Wolfson (1994): When Inequalities Diverge, in: American Economic Review, 84(2), pp. 353–358.

A Appendix

A.1 Data

As primary source for our data, we use the SOEP database (Wagner et al., 2007). We use samples A to J, i.e. all available samples excluding only the last refreshment sample 2012 and the migration sample 2013. Concerning the waves, we only use information only from wave BD, i.e. the wave for 2013. Only taking those values for which we have the full set of variables, as described below, this yields 8,344 observations (3780 males and 4564 females).

A.1.1 Variables used

In order to obtain our dependent variable earnings rate, we use the gross-market labour income obtained in the previous months (`bdp7701`). This number is divided by the hours worked per week (`bdp1011`) times 4.2 (for weeks per month, following van Kerm, 2013).

For the explanatory variable age, we simply use the year of birth (`gebjahr`) and subtract it from 2013, while the sex is determined by the variable `bdsex`.

The education level is taken on grounds of the variable `ISCED13` from the person-related status and generated variables (`PGEN`) available in the SOEP. All observations equal or lower than 5 (higher vocational training) are put in the category no higher education with only those persons with a value of 6 (higher education) considered for the category higher education.

The nationality is obtained directly from the SOEP based on the person's contemporary status (`BDP143`).

For the spatial effect we use the variable `bdbula` with the variable *east* set to unity for all federal states formerly belonging to the German Democratic Republic, including the whole of Berlin. West Berlin (as defined prior to 1990) is not accounted for in our sample and treated like a state from the former East.

For the length of the personal unemployment spell, we use the available information from the spell data (`PBIOSPE` and `ARTKALEN`) for each individual to construct the duration of the unemployment spells prior to the 25$^{\text{th}}$ birthday. Subsequently, the variable *stunem* is defined to be unity if the resultant length of all the unemployment spells is positive but not above one year and zero otherwise. Additionally, *ltunem* is unity if the sum of all unemployment spells is longer than one year.

All observations are weighted using the variable `bdphrf`. This weighting accounts for different selection probabilities due to the stratified in the sampling and differences in response rates by the contacted individuals.

A.1.2 The factual and counterfactual covariate distributions

For the construction of the factual shares $s^{act}(\mathbf{x})$, we simply weight each covariate combination according to the weighted number of individuals recorded in the SOEP.
For the construction of the counterfactual shares $s^{act}(\mathbf{x})$, we simply take the SOEP and randomly assign new unemployment characteristics to the affected cohort (i.e. those between 37 and 46 in 2013) using the random numbers generator in `R`, with the probabilities specified as follows:

$$P(\text{no Unemployment}) = 0.417$$
$$P(\text{short-term Unemployment}) = 0.280$$
$$P(\text{long-term Unemployment}) = 0.303$$

These estimates are based on estimates on unemployment experience for people between 15 and 24 in 2013 provided by the Hellenic Statistical Authority as well as estimates from Eurostat on the economically active population in their youth. From the former we got the following estimates: 88,200 Greek youth had experienced short-term unemployment, while 95,400 young Greeks head experienced long-term unemployment in 2013 (Hellenic Statistical Authority, 2016). The overall active population is estimated at 315,600 (Eurostat, 2016), yielding the estimates shown above.
Note that given this data we can only give unconditional probabilities, such that we do not condition the probability on variables like education, which are likely to influence the chance of both short-term and long-term unemployment experiences. Consequently, the provided estimates for the counterfactual scenario ought to be considered very rough at best. However, since the example only serves exemplary purposes, we believe that such a rough-and-ready approach is justifiable for the purpose at hand.

A.2 Priors and Hyperparamters

The priors for the effects in Equation 5 are as follows:

- For the linear effect (β_0, β_1, β_2, β_3, β_4 and β_5), we simply use flat priors, so that no penalisation is incurred.
- For the smooth effect for age (f_1 and f_2), we use scale dependent priors with the scale parameter (denoted by θ in Klein et al. (forthcoming)) set to 1×10^{-2}. The age effect is

thus only loosely smoothed. The reasoning is that we want to put little restraint on the age as non-linearities are likely and are indeed standard in the literature in form of a quadratic polynomial. Moreover, the relatively even distribution of the covariate allows for a robust estimation over the space of the age covariate.

- For the smooth effect for unemployment (f_3), we use scale dependent priors with the scale parameter set to 1×10^{-5}, implying a much more stringent penalisation. This more heavy regularisation is required as the unemployment is much more unevenly distributed leaving only about 85% of the observations not featuring any unemployment leaving only 1300 observations above zero which are in turn heavily concentrated towards the lower spectrum. With this uneven distribution, unpenalised deviations are likely to be driven by very few observations yielding a possibly erratic predictor.
- For the random effects for the federal states included in f_4, we use a weakly penalising scaling parameter of 1×10^{-2} again. The reasoning is analogue to age.

A.3 Two-Stage Model Set-up

The model set-up for the estimation of the conditional wage distributions is as follows:

First, we estimate the conditional probability of unemployment, π, which is equivalent to the person receiving a zero-wage, such that $\pi = P(y = 0)$:

$$g_1(\pi) = \eta^{\pi}, \tag{10}$$

where g_1 is a logit-link and η^{π} is the predictor as specified in Equation (5) for a given covariate combination.

For the positive wages, we use the Type I Dagum distribution which has a track record of performing very well for modelling positive earnings at the aggregate level (e.g. Kleiber and Kotz, 2003; Chotikapanich, 2008).[7] A natural alternative which is the work-horse distribution in the literature

[7] This distribution is very similar to the Singh-Maddala distribution which has been used by Biewen and Jenkins (2005) for conditional earning distributions and by van Kerm (2013) for modelling wage rates. Yet, Kleiber and Kotz (2003) remark that the Dagum distribution generally performs slightly better than the Singh-Maddala. It should be noted that other more complex distributions, like the Generalised Beta of Second Kind, have been suggested in the literature for aggregate income distributions which outperform the Dagum distribution. However, these more complex distributions have so far proven to be too complex for stable estimation. A comparative study on the performance of the fit of these various parametric alternatives would be needed though to provide a more profound assessment on this issue.

is the log-normal distribution (see Flabbi, 2010), yet previous studies have found this distribution to have problems in modelling conditional earning distributions (Sohn et al., 2015; Sohn, 2015).

The density of the Type I Dagum distribution is given by

$$p_{+}(y \mid a, b, c) = \frac{acy^{ac-1}}{b^{ac}(1+(y/b)^{a})^{p+1}}, \quad a \in \mathbb{R}_{>0},\ b \in \mathbb{R}_{>0},\ c \in \mathbb{R}_{>0}. \tag{11}$$

The estimation of the three parameters a, b and c is done using the following generic predictor set-up as discussed above:

$$g_2(a) = \eta^{a}, \tag{12}$$

$$g_3(b) = \eta^{b}, \tag{13}$$

$$g_4(c) = \eta^{c}, \tag{14}$$

where all three link functions are log-link functions ensuring a positive support for the parameters and inducing a multiplicative connection between the covariates.

Overall, the estimation procedure gives us four parameters to estimate the wage distribution over the covariates space. Using this parametrisation, the density of the conditional wage distribution can be expressed as a mixture of a point-mass at zero and a continuous distribution thereafter:

$$p(y \mid \pi, a, b, c) = \pi \mathbb{1}_{\{y=0\}} + (1-\pi)p_{+}(y \mid a, b, c), \tag{15}$$

where p denotes the probability mass or probability density for a given wage y. For the point mass of zero wages $\mathbb{1}_{\{y=0\}}$ denotes an indicator function which is unity for a wage of zero and thus gives a probability mass of π. For earnings greater than zero, we obtain the density as specified by the Type I Dagum distribution.

It should be noted that currently a lot of work is being done on various other statistical approaches allowing for the estimation of conditional distributions. For a discussion of other estimation strategies see Sohn (2015).

A.4 Effects on the parameters

Due to the set-up of SADR, we have a specification of the impact of each covariate on the parameters of the distribution. While some parameters have a straight forward interpretation others are less easily interpreted. Here we focus on the effect of unemployment at the time of youth as captured by the variable *stunem* and *ltunem* on the parameters. The results are displayed in Table 3.

	short-term unemployment		long-term unemployment	
	males	females	males	females
π	−0.321[−0.361;−0.275]	0.015[−0.013;0.046]	1.474[1.439; 1.512]	0.832[0.797;0.867]
a	0.191[−0.011; 0.431]	0.053[−0.148;0.258]	0.188[−0.126; 0.517]	0.241[−0.122;0.620]
b	0.002[−0.103; 0.099]	−0.022[−0.112;0.069]	−0.136[−0.271;−0.019]	−0.020[−0.167;0.131]
c	−0.219[−0.584; 0.151]	−0.125[−0.453;0.174]	−0.251[−0.786; 0.235]	−0.532[−1.116;0.004]

Table 3: Effects for *stunem* and *ltunem* on parameter predictors

As can be observed in the table, the effect of long-term unemployment during the time of youth has a positive and significant effect on the contemporary chance of unemployment. This finding is in line with what would be expected following the literature on scarring. A little surprisingly an opposite albeit much smaller effect is found for short-term unemployment for males.

For the parameters of the Dagum distribution, we can observe varying directions and levels of significance with most parameters not significant at the 5% level. Due to the difficulty of interpreting the parameters independently, we will only point to the fact that the scaling parameter b generally points in the expected downward direction for long-term unemployment, while for short-term unemployment the picture is more mixed again. The main aspect to emphasise here is the large levels of uncertainty attached to the parameter estimates, which is hardly surprising given the complex nature of distributional regression (see Klein et al., 2015a).

A.5 Unemployment effect for different covariate combinations

A.5.1 Unemployment effect for other ages

Tables 4 and 5 show some distribution measures of the earning distribution two other ages, namely the first and the third quartile of the age range considered, 34 years and 51 years, respectively. All other covariates are kept the same as in the main example in the text.

As can be seen from the tables, the general structure persists, whereby the probability of unemployment (π) rises as individuals experience unemployment during their youth, while the expected income (μ) falls. The inequality associated with the distributions as expressed by the Atkinson index $\mathcal{A}$ generally increases leading to a more drastic decrease in the welfare measure $\mathcal{C}$ for all three levels of inequality aversion considered.

	D_{0m}	D_{1m}	D_{2m}	D_{0f}	D_{1f}	D_{2f}
$\pi_{\mathbf{x}}$	0.07[0.05; 0.09]	0.05[0.03; 0.07]	0.24[0.18;0.31]	0.20[0.15;0.25]	0.20[0.15;0.26]	0.36[0.29;0.44]
$\mu_{\mathbf{x}}$	12.63[11.21;14.07]	12.46[11.10;13.74]	7.32[6.10;8.65]	8.69[7.65;9.64]	8.13[7.10;9.10]	5.52[4.52;6.60]
$\mathcal{A}_{\mathbf{x}}(\epsilon{=}0.1)$	0.02[0.02; 0.03]	0.02[0.01; 0.02]	0.06[0.04;0.07]	0.04[0.03;0.05]	0.04[0.03;0.05]	0.07[0.06;0.09]
$\mathcal{A}_{\mathbf{x}}(\epsilon{=}0.5)$	0.14[0.11; 0.19]	0.11[0.09; 0.14]	0.37[0.30;0.45]	0.27[0.22;0.33]	0.27[0.22;0.33]	0.44[0.37;0.51]
$\mathcal{A}_{\mathbf{x}}(\epsilon{=}0.9)$	0.44[0.35; 0.54]	0.35[0.28; 0.44]	0.82[0.74;0.88]	0.67[0.59;0.76]	0.68[0.59;0.76]	0.85[0.79;0.90]
$\mathcal{C}_{\mathbf{x}}(\epsilon{=}0.1)$	12.37[10.95;13.78]	12.26[10.87;13.55]	6.91[5.69;8.25]	8.34[7.28;9.31]	7.80[6.75;8.79]	5.14[4.15;6.24]
$\mathcal{C}_{\mathbf{x}}(\epsilon{=}0.5)$	10.80[9.39;12.19]	11.06[9.64;12.37]	4.60[3.39;5.90]	6.36[5.21;7.43]	5.90[4.80;6.98]	3.12[2.27;4.11]
$\mathcal{C}_{\mathbf{x}}(\epsilon{=}0.9)$	7.12[5.48; 8.68]	8.05[6.51; 9.49]	1.34[0.76;2.18]	2.84[1.89;3.86]	2.63[1.75;3.61]	0.80[0.47;1.31]

Table 4: Some distribution measures for the conditional wage distributions for 34 years (all other covariates the same)

	D_{0m}	D_{1m}	D_{2m}	D_{0f}	D_{1f}	D_{2f}
$\pi_{\mathbf{x}}$	0.13[0.09; 0.17]	0.10[0.07; 0.13]	0.39[0.30;0.48]	0.15[0.12; 0.21]	0.16[0.12;0.21]	0.30[0.23;0.38]
$\mu_{\mathbf{x}}$	14.39[12.87;16.10]	14.02[12.50;15.60]	7.99[6.58;9.54]	9.31[8.15;10.45]	8.65[7.43;9.88]	5.63[4.56;6.85]
$\mathcal{A}_{\mathbf{x}}(\epsilon{=}0.1)$	0.02[0.02; 0.03]	0.02[0.02; 0.02]	0.06[0.05;0.08]	0.04[0.04; 0.06]	0.05[0.04;0.06]	0.08[0.06;0.09]
$\mathcal{A}_{\mathbf{x}}(\epsilon{=}0.5)$	0.17[0.13; 0.21]	0.13[0.10; 0.17]	0.40[0.33;0.49]	0.30[0.24; 0.36]	0.30[0.24;0.37]	0.47[0.39;0.54]
$\mathcal{A}_{\mathbf{x}}(\epsilon{=}0.9)$	0.48[0.39; 0.58]	0.39[0.31; 0.48]	0.84[0.77;0.90]	0.72[0.63; 0.80]	0.72[0.63;0.80]	0.88[0.82;0.92]
$\mathcal{C}_{\mathbf{x}}(\epsilon{=}0.1)$	14.04[12.48;15.74]	13.75[12.24;15.32]	7.48[6.08;9.04]	8.89[7.73;10.02]	8.25[7.03;9.50]	5.21[4.15;6.43]
$\mathcal{C}_{\mathbf{x}}(\epsilon{=}0.5)$	12.01[10.37;13.60]	12.20[10.68;13.74]	4.76[3.48;6.19]	6.56[5.26; 7.77]	6.04[4.81;7.27]	3.01[2.15;4.03]
$\mathcal{C}_{\mathbf{x}}(\epsilon{=}0.9)$	7.44[5.71; 9.22]	8.51[6.81;10.17]	1.26[0.70;2.10]	2.63[1.68; 3.71]	2.42[1.56;3.48]	0.69[0.39;1.16]

Table 5: Some distribution measures for the conditional wage distributions for 51 years (all other covariates the same)

A.5.2 Unemployment effect for another education level

Table 6 shows the distribution measures for an individual with higher education (i.e. education level 4 rather than 2). All other covariates are kept the same as in the main example in the text.

	D_{0m}	D_{1m}	D_{2m}	D_{0f}	D_{1f}	D_{2f}
$\pi_{\mathbf{x}}$	0.02[0.01; 0.03]	0.01[0.01; 0.02]	0.08[0.06; 0.12]	0.17[0.13; 0.23]	0.18[0.13; 0.23]	0.33[0.26; 0.41]
$\mu_{\mathbf{x}}$	26.76[23.73;29.94]	24.63[22.02;27.71]	20.28[17.52;23.43]	17.64[15.89;19.40]	16.41[14.68;18.40]	13.07[11.32;15.24]
$\mathcal{A}_{\mathbf{x}}(\epsilon{=}0.1)$	0.01[0.01; 0.02]	0.01[0.01; 0.01]	0.01[0.01; 0.02]	0.02[0.02; 0.03]	0.02[0.02; 0.03]	0.03[0.03; 0.04]
$\mathcal{A}_{\mathbf{x}}(\epsilon{=}0.5)$	0.06[0.05; 0.08]	0.05[0.04; 0.07]	0.09[0.07; 0.12]	0.13[0.11; 0.16]	0.14[0.11; 0.17]	0.22[0.17; 0.27]
$\mathcal{A}_{\mathbf{x}}(\epsilon{=}0.9)$	0.14[0.11; 0.18]	0.11[0.08; 0.15]	0.26[0.19; 0.34]	0.38[0.31; 0.46]	0.38[0.31; 0.46]	0.59[0.49; 0.68]
$\mathcal{C}_{\mathbf{x}}(\epsilon{=}0.1)$	26.42[23.45;29.53]	24.38[21.81;27.40]	19.98[17.25;23.04]	17.28[15.55;18.99]	16.08[14.34;17.99]	12.63[10.91;14.83]
$\mathcal{C}_{\mathbf{x}}(\epsilon{=}0.5)$	25.07[22.24;27.81]	23.37[20.93;26.07]	18.45[15.85;21.23]	15.31[13.64;17.00]	14.20[12.48;16.05]	10.15[8.41;12.24]
$\mathcal{C}_{\mathbf{x}}(\epsilon{=}0.9)$	23.13[20.48;25.60]	21.92[19.57;24.48]	15.00[12.40;17.70]	10.92[9.00;12.88]	10.14[8.25;12.09]	5.31[3.78; 7.22]

Table 6: Some distribution measures for the conditional wage distributions for educlvl 4 (all other covariates the same)

Again one can observe that the basic structure persists.

A.5.3 Unemployment effect for another federal state

Table 7 shows the distribution measures for an individual who lives in Mecklenburg-West Pomerania, an economically depressed state in the former East of Germany. All other covariates are kept the same as in the main example in the text.

	D_{0m}	D_{1m}	D_{2m}	D_{0f}	D_{1f}	D_{2f}
$\pi_{\mathbf{x}}$	0.16[0.11; 0.21]	0.12[0.09; 0.16]	0.45[0.36;0.54]	0.26[0.20; 0.33]	0.27[0.20; 0.33]	0.45[0.36;0.53]
$\mu_{\mathbf{x}}$	12.82[10.93;14.87]	12.76[10.93;14.70]	6.35[4.83;8.07]	9.35[7.99;10.78]	8.67[7.36;10.18]	5.66[4.50;7.13]
$\mathcal{A}_{\mathbf{x}}(\epsilon$=0.1)	0.03[0.02; 0.04]	0.02[0.02; 0.03]	0.08[0.06;0.10]	0.04[0.04; 0.06]	0.05[0.04; 0.06]	0.08[0.06;0.10]
$\mathcal{A}_{\mathbf{x}}(\epsilon$=0.5)	0.21[0.16; 0.27]	0.16[0.12; 0.21]	0.49[0.40;0.58]	0.30[0.24; 0.37]	0.30[0.24; 0.38]	0.47[0.39;0.55]
$\mathcal{A}_{\mathbf{x}}(\epsilon$=0.9)	0.59[0.48; 0.70]	0.49[0.38; 0.61]	0.90[0.84;0.94]	0.72[0.62; 0.81]	0.72[0.62; 0.81]	0.88[0.82;0.92]
$\mathcal{C}_{\mathbf{x}}(\epsilon$=0.1)	12.45[10.57;14.51]	12.48[10.65;14.41]	5.84[4.34;7.57]	8.92[7.58;10.36]	8.28[6.92; 9.77]	5.23[4.09;6.67]
$\mathcal{C}_{\mathbf{x}}(\epsilon$=0.5)	10.16[8.14;12.32]	10.70[8.86;12.64]	3.22[2.04;4.74]	6.56[5.09; 8.01]	6.06[4.66; 7.46]	3.00[2.03;4.26]
$\mathcal{C}_{\mathbf{x}}(\epsilon$=0.9)	5.26[3.39; 7.52]	6.48[4.58; 8.70]	0.62[0.30;1.21]	2.58[1.54; 3.91]	2.38[1.43; 3.65]	0.68[0.35;1.24]

Table 7: Some distribution measures for the conditional wage distributions for Mecklenburg-West Pomerania (all other covariates the same)

Again one can observe that the basic structure persists.

A.5.4 Unemployment effect for foreign nationality

Table 8 shows the distribution measures for an individual who does not have German nationality. All other covariates are kept the same as in the main example in the text.

	D_{0m}	D_{1m}	D_{2m}	D_{0f}	D_{1f}	D_{2f}
$\pi_{\mathbf{x}}$	0.07[0.05; 0.10]	0.05[0.04; 0.07]	0.25[0.19;0.33]	0.23[0.17;0.29]	0.23[0.18;0.30]	0.40[0.33;0.49]
$\mu_{\mathbf{x}}$	12.09[10.62;13.77]	12.02[10.68;13.63]	6.44[5.24;7.77]	7.50[6.54;8.48]	7.03[6.11;8.02]	4.79[3.97;5.74]
$\mathcal{A}_{\mathbf{x}}(\epsilon$=0.1)	0.02[0.02; 0.03]	0.02[0.01; 0.02]	0.07[0.05;0.09]	0.04[0.03;0.05]	0.04[0.03;0.05]	0.07[0.05;0.09]
$\mathcal{A}_{\mathbf{x}}(\epsilon$=0.5)	0.17[0.13; 0.23]	0.13[0.10; 0.18]	0.44[0.36;0.52]	0.26[0.21;0.33]	0.27[0.21;0.33]	0.43[0.36;0.51]
$\mathcal{A}_{\mathbf{x}}(\epsilon$=0.9)	0.52[0.42; 0.63]	0.43[0.33; 0.53]	0.87[0.81;0.92]	0.67[0.58;0.76]	0.67[0.58;0.76]	0.85[0.78;0.90]
$\mathcal{C}_{\mathbf{x}}(\epsilon$=0.1)	11.80[10.32;13.43]	11.80[10.47;13.37]	5.99[4.81;7.34]	7.22[6.22;8.19]	6.76[5.79;7.75]	4.47[3.64;5.41]
$\mathcal{C}_{\mathbf{x}}(\epsilon$=0.5)	9.99[8.44;11.58]	10.40[9.04;11.94]	3.60[2.56;4.81]	5.52[4.46;6.54]	5.16[4.08;6.18]	2.75[1.98;3.60]
$\mathcal{C}_{\mathbf{x}}(\epsilon$=0.9)	5.79[4.19; 7.45]	6.90[5.29; 8.54]	0.83[0.45;1.44]	2.49[1.66;3.44]	2.32[1.53;3.23]	0.72[0.43;1.18]

Table 8: Some distribution measures for the conditional wage distributions for foreign nationality (all other covariates the same)

Again one can observe that the basic structure persists.

A.6 Misspecification Testing

In order to assess the adequacy of the parametric assumptions made for SADR, we test the following null-hypothesis $\mathcal{H}_0$ against the alternative hypothesis $\mathcal{H}_1$:

$\mathcal{H}_0$: The conditional wage distributions can be modelled by our parametric form, $p(y \mid \boldsymbol{\theta})$, for all observed values of y and some values of $\boldsymbol{\theta}$ derived for the corresponding covariates, $\mathbf{x}$.

vs.

$\mathcal{H}_1$: The conditional wage distributions cannot be modelled by our parametric form, $p(y \mid \boldsymbol{\theta})$, for all observed values of y and any values of $\boldsymbol{\theta}$ derived for the corresponding covariates, $\mathbf{x}$.

The test used employed is an adaptation of the Kolmogorov-Smirnov test and derivations thereof for the distributional regression case as proposed by Andrews (1997) and Rothe and Wied (2013) for the frequentist framework. The underlying idea is to transform the conditional moment restrictions imposed by the parametric specification of the model into unconditional ones (see Rothe and Wied, 2013). This transformation allows for the computation of one test statistic whose distribution under the null can be derived by Monte Carlo techniques. The details of our implementation in the Bayesian setting are elaborated in Sohn (2015).