FDZ Methodenreport

Methodische Aspekte zu Arbeitsmarktdaten

No. 06/2008

Bounds analysis of competing risks: a nonparametric evaluation of the effect of unemployment benefits on migration in Germany.

(Revised Version of the FDZ Methodenbericht No. 04/2007)

Melanie Arntz (ZEW Mannheim)
Simon M. S. Lo (University of Freiburg)
Ralf A. Wilke (University of Nottingham)

Bundesagentur für Arbeit

Data access

The dataset described in this document is available for use by professional researchers. For further information, see “Individual Data” on the website: http://fdz.iab.de/en/.

Bounds analysis of competing risks: a nonparametric evaluation of the effect of unemployment benefits on migration in Germany.*

Melanie Arntz[†]

Simon M. S. Lo[‡]

Ralf A. Wilke[§]

July 2008

Abstract

This paper suggests an approach for analyzing a dependent competing risks model in presence of partly identified interval data. We apply our nonparametric bounds framework to empirically evaluate the effect of unemployment benefits on the cumulative incidence of local job finding and inter-regional migration of unemployed workers in Germany. Our findings indicate that reducing the entitlement length for unemployment benefits has heterogenous effects depending on the household composition and the wage replacement ratio in absence of unemployment benefits.

Keywords: cumulative incidence curve, partially missing data, bounds analysis, difference-in-differences

JEL: C41, C14, J61

*We would like to thank Simon Lee for helpful discussions and the participants of numerous seminars for helpful comments. Arntz gratefully acknowledges financial support by the German Research Foundation (DFG) through the research project "Potentials for more flexibility of regional labor markets by means of interregional labor mobility". We use a sample drawn from the Employee and Benefit Recipient History (V6.0) of the Institute of Employment Research (IAB). The data preparation and analyses were made possible by the Research Data Centre (FDZ) of the German Federal Employment Agency at the IAB because the project provides new insights on data quality. We thank selected staff members of the Research Data Centre (FDZ) of the German Federal Employment Agency at the IAB and especially Stefan Seth for the help with the data.

[†]ZEW Mannheim, E–mail: arntz@zew.de.

[‡]University of Freiburg, E–mail: losimonms@yahoo.com.hk

[§]University of Nottingham, E–mail: ralf.wilke@nottingham.ac.uk

1 Introduction

In many cases, administratively merged interval data contain only incomplete individual employment records. As an example, German administrative data do not contain the exact length of unemployment periods because there are unobserved periods in an individual's employment trajectory. Parameters of interest are not point identified if the econometric framework accounts for this incomplete information. For example, Lee and Wilke (2008) suggest bounds for a difference-in-differences treatment effect on the marginal survival probability in unemployment. They bound the effect of a reform of unemployment compensation in Germany using a lower and an upper bound of the true unemployment duration. While their approach is restricted to a model with independent censoring, this paper considers a dependent competing risks model.

The contribution of this paper is twofold: First, we present a bounds framework for a competing risks model in presence of partially identified interval data by deriving bounds for the cumulative incidence curve (CIC, see also Kalbfleisch and Prentice, 1980). As we have access to a large data set that encompasses 50% of the male working population in Germany, we choose a nonparametric approach in order to impose only few assumptions that may be violated in the real world. Our work is relevant as partial identification of unemployment duration is a common problem in merged administrative individual data and potentially occurs in most national data sets. The literature using such data is growing in recent years and governments contract out research based on this data to analyze policy reforms. Second, we apply this framework to investigate empirically the effect of unemployment compensation on the cumulative incidence of local job finding and inter-regional migration. Our approach thus allows for analyzing a policy question that is highly debated in the public, namely the question whether unemployment benefits promote or inhibit migration of unemployed workers.

The findings of previous research on the effect of unemployment benefits on unemployment duration generally suggest a disincentive effect of unemployment compensation on the transitions from unemployment to employment (Katz and Meyer, 1990; Card and Levine, 2000; Lalive and Zweimüller, 2004; van Ours and Vodopivec, 2006). These findings are in line with the predictions from search theory that considers unemployment compensation to raise reservation wages (Atkinson and Micklewright, 1991). The effect of unemployment benefits on migration, however, is less clear. The negative effect of rising reservation wages and smaller geographical search horizons as a reaction to higher benefit levels (Hassler at al., 2005) contrasts a positive resource effect as higher unemployment benefits levels enable individuals to bear migration cost (Tatsiramos, 2008) and to increase expenditures that enhance the productiveness of job search (Barron and Mellow, 1979; Tannery, 1983). Most empirical studies, however, seem to suggest a mobility-reducing effect of unemployment benefits on migration (Goss and Paul, 1990, Antolin and Bover, 1997). Consistent with these findings, Arntz (2005) and Arntz and Wilke (2008) obtain some empirical evidence with data from Germany that unemployed with higher previous earning capacities who are entitled to receive unemployment benefits (UNB) for an extended period of more than 18 months are much less likely to leave unemployment via migrating to another region than individuals with a shorter period of UNB receipt. To some extent, the findings of these studies may be driven by an unobserved selection of immobile individuals into unemployment benefits or an extensive receipt of UNB. In a study with individual fixed effects that should mitigate such biases, Tatsiramos (2008) finds a positive effect of unemployment benefits on migration in Denmark and France, a result that he assigns to the mobility-enhancing resource effect of unemployment benefits. In contrast, he finds no effect of unemployment benefits on migration in the UK or Germany. As a drawback, however, this study does not take account of competing transitions to local employment in a duration model context. Our empirical analysis thus reexamines the effect of shorter unemployment benefit receipt on possibly dependent transitions to either local or non-local employment via migration. For this purpose, we exploit a natural experiment that generates some exogenous variation of entitlement length, namely the reform of unemployment benefit entitlements in Germany in 1997. This reform reduced the length of entitlements for certain age groups by up to 10 months. We obtain the following empirical findings:

- Our results confirm that missing interval information in German individual administrative data at first precludes any clear result as the bounds tend to be very wide. By introducing additional assumptions, bounds can be tightened.

- For high-skilled individuals, our results are indicative for the presence of effects of the entitlement length for unemployment benefits on the duration of unemployment. In contrast, there is no evidence that labour market outcomes for less skilled unemployed are affected by the entitlement length for unemployment benefits. This is explained by a wage replacement ratio that, for this group, is invariant to the receipt of unemployment benefits.

- The family or household background is an important determinant for the effect of a reduction in benefit entitlements in case of the high skilled: The cumulative incidence for local job finding increases for married unemployed, while it remains constant for singles. The cumulative incidence for non-local job finding via migration increases for singles while it is constant for married unemployed.

The paper is structured as follows. The following section presents the econometric framework. Section three contains the application and section four concludes.

2 Econometric Model

In this model we study state transition times with $k = 0, 1, \ldots, K$ different competing labor market states. We let T_{lk} be a random variable of the latent transition time from an original state $l = 0, 1, \ldots, K$ to a destination state $k = 0, 1, \ldots, K$. l denotes the original state and k denotes the destination state with $k \neq l$. We denote $k = 0$ as the state of unemployment whereas $k = 1, \ldots, K$ represents states such as employment in the local or a non-local area or being out of labor force. There are $i = 1, \ldots, n$ independent identically distributed realizations τ_{ilk} of T_{lk}. For simplicity, we suppress subscript i in the rest of this paper. In case of $k = 0$, the end date of the last employment spell is normalized as $T_{l0} = 0$, for $l = 1, \ldots, K$.

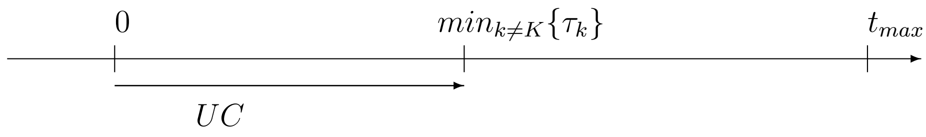
We are particularly interested in studying unemployment duration, i.e. the transition time T_{0k} from state $l = 0$ to state $k = 1, \ldots, K$, and its realization τ_{0k} with $k = 1, \ldots, K$. When there is no ambiguity, we suppress the subscript of the original state 0. The latent transition time is then $T_{0k} = T_k$ and its realization is $\tau_{0k} = \tau_k$. We assume that $\tau_{lk} \neq \tau_{lm}$ for all l and all $k \neq m$. X is a vector of exogenous individual characteristics and T_{lk} are some unknown functions of X. T_{max} is an exogenous random variable that refers to the maximum observation period and t_{max} is its realization. In other words, we have independent right censoring at the end of the observation period. In what follows, r is the exit state which has the shortest latent transition time among all τ_k with $k = 1, \ldots, K$, i.e.

$$r = \begin{cases} 1 & \text{if} \quad \tau_1 = min_k\{\tau_k\}; \\ & \quad\quad \vdots \\ K & \text{if} \quad \tau_K = min_k\{\tau_k\}. \end{cases}$$

Now assume that labor market state K pools all labor market states that are unobserved in the data. The model can be easily extended to more than one unobserved labor market state but it is not interesting to distinguish between different unobservable risks. In the German administrative data, unobservable labor market states that produce observational gaps are self-employment or being out of the labor force. Since an observational gap after the end of unemployment compensation transfers may also refer to continued unemployment without receiving unemployment compensation, it is not clear whether an individual remains unemployed or leaves to one of the unobserved exit states K after the end of transfer receipt. The data structure thus implies that the unemployment duration and the transition to one of the other exit states $k = 1, ..., K-1$ can unambiguously be identified only if the individual receives unemployment compensation during the entire time period of unemployment.

Figure 1 illustrates the fully identified case where the observed transition time refers to the transition time from state 0 (unemployment) to state $k \neq K$ such that $\tau_r = \min_{k \neq K}\{\tau_k\}$. Thus, the transition time is point-identified with known exit state $r \neq K$.

Figure 1: A fully identified unemployment duration with $\tau_r = \min_{k \neq K}\{\tau_k\}$.



In contrast, if there is an observational gap in the data, the true unemployment duration is not point identified. Let us denote the beginning of the first unobserved period by the random variable C and its realization by ς. In our application, this is usually the end date of an unemployment compensation claim period. In the German context, unemployment compensation stops because unemployment benefits have been exhausted and the individual does not pass a means-test for unemployment assistance. Other reasons are benefit sanctions for unemployed who did not comply with the eligibility criteria. Both cases can be hardly predicted and are typically not random. Since the earliest exit to K occurs at the beginning of an unobserved time period, it holds that $\varsigma \leq \tau_K$ and thus C and T_K are not independent, i.e. $C = T_K - \xi(T_K)$ where $\xi(T_K) \in [0, T_K]$ is some positive random function. Given the data structure in our application, there may also be observations for which the receipt of unemployment compensation does not immediately start after the end of an employment period and thus there is an unobserved period starting at $T_{l0} = 0$ so that $\varsigma = 0$. Moreover, ς is observed only if $\varsigma \leq \tau_r$. In the case of our identified spell in Figure 1, we therefore have $\tau_r = \min_{k \neq K}\{\tau_k\} < \varsigma \leq \tau_K$ and ς is not observed.

Figure 2: A partially identified unemployment duration for which ς is observed and $\varsigma \leq s_{k \neq K} \leq t_{max}$.

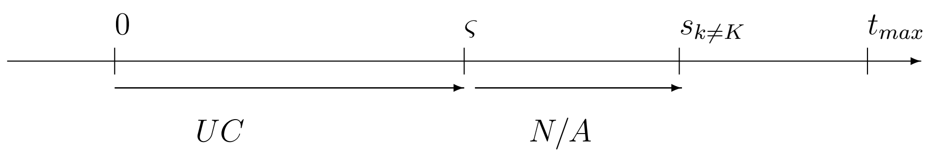


Figure 2 illustrates the case where we observe ς, i.e. $\varsigma \leq \tau_r$ due to an unobserved period after a period of unemployment compensation transfers. As discussed before, an unobserved period can also occur directly after an employment period ($\varsigma = 0$). In these cases, τ_r cannot be point identified from the data and has to be bounded. For this purpose we denote the first observed transition time after an unobserved period as $s_{k \neq K}$. It is possible to construct worst-case bounds for τ_r:

1. The upper bound of τ_r can be obtained by assuming that there is no exit to K during the unobserved period. Instead, unemployment continues until $s_{k \neq K}$ and so τ_r is equal to $s_{k \neq K} = min_{k \neq K}\{\tau_k\}$. In other words, by ignoring ς, τ_r would be identified as in Figure 1.

2. The lower bound of τ_r can be obtained by assuming that there is an exit to K during the unobserved period, i.e. $r = K$. Then the earliest transition to K can occur at ς so that the lowest value of τ_r is ς. In this case, $s_{k \neq K}$ equals to $\tau_K + \tau_{Kk}$ and can be ignored.

If we observe ς, the true value τ_r lies always in the interval $[\varsigma, s_{k \neq K})$. We now define a variable δ to formalize the identification of τ_r from the data as follows:

$$\delta = \begin{cases} 0 & \text{if } \varsigma \text{ is not observed;} \\ 1 & \text{if } \varsigma \text{ is observed} \end{cases}$$

Events $\delta = 0, 1$ are disjoint and can be distinguished in the data. If ς is not observed in the data, we have $\delta = 0$, and τ_r is fully identified. If we have $\delta = 1$, the unemployment duration is partially identified and τ_r is unknown. If we knew that $r \neq K$, then τ_r would be known. But if $r = K$, we only know that $\tau_K \in [\varsigma, s_{k \neq K})$. The difficulty in an application is that for $\delta = 1$ we do not know whether $r = K$ or $r \neq K$.

In addition to the identification problem that arises from the uncertainty of τ_r in the case of $\delta = 1$, our competing risk setting also implies another identification problem which is related to the general identification problem of competing risks. If risks are not independent, the marginal distribution for each competing risk cannot be identified without additional parametric assumptions (Cox, 1962; Tsiatis, 1975). In light of this additional identification problem, partial effects and changes in the latent distributions can also only be bounded. Non-parametric bounds on the marginal distribution as have been proposed by Peterson (1976) are typically too wide to infer some causal interpretation. As an alternative, parametric assumptions can be imposed to tighten bounds or to achieve full identification. Under rather restrictive assumptions, Heckman and Honoré (1989) and Abbring and van den Berg (2003) show identification of the semiparametric mixed proportional hazard model. Honoré and Lleras-Muney (2006) impose mild assumptions to obtain tight bounds for parameters within the accelerated failure time model. Our approach avoids such parametric assumptions which may not be met in our application and hence leaves the fundamental identification problem unresolved. As its main contribution, however, it tackles the identification problem that stems from partial identification of interval data. Moreover, by using bounds on the cumulative incidence curve, it provides a non-parametric tool which has a meaningful interpretation also in presence of dependent competing risks.

The CIC refers to the observed probability of experiencing a transition to a specific state prior to a certain time in the presence of all competing risks (see for example Moeschberger and Klein, 1995). It therefore does not recover the underlying risk-specific marginal distribution of latent durations. Instead, it refers to observed transition probabilities. Related literature also refers to this as the subdistribution (Kalbfleisch and Prentice, 1980). In the following, we derive bounds on the identification region of the CIC which reflect the partial identification of τ_r only. These bounds do not resolve the identification problem of competing risks models. Equivalent bounds can also be derived for the overall survivor curve while bounds for other functions such as the cause-specific hazard rate or the cause-specific cumulative hazard rate cannot be derived. In the following, we restrict our attention to the observable risks $k = 1, \ldots, K-1$.
The CIC for transition to state k at time t can be decomposed into a part that is due to fully identified observations ($\delta = 0$) and partially identified observations ($\delta = 1$):

$$\begin{aligned}
I_k(t|x) &= P(T_k \leq t, r = k|x) \\
&= P(T_k \leq t, r = k, \delta = 0|x) + P(T_k \leq t, r = k, \delta = 1|x) \\
&= P(T_k \leq t, r = k, \delta = 0|x) \\
&\quad + P(T_k \leq t, r = k, r \neq K, \delta = 1|x) + P(T_k \leq t, r = k, r = K, \delta = 1|x) \\
&= P(T_k \leq t, r = k, \delta = 0|x) + P(T_k \leq t, r = k, r \neq K, \delta = 1|x) \qquad (1)
\end{aligned}$$

for $k = 1, \ldots, K-1$. The second part of (1) is not identified since we cannot identify r in presence of $\delta = 1$. Therefore, as discussed before, the second part of (1) can only be bounded. Similar to the ideas outlined by Manski (2003), we bound the unknown probability $P(T_k \leq t, r = k, r \neq K, \delta = 1|x)$ by an interval which is identifiable with the available data structure. To see this, we rewrite (1) as:

$$\begin{aligned}
I_k(t|x) &= P(T_k \leq t, r = k, \delta = 0|x) \\
&\quad + P(T_k \leq t, r = k|r \neq K, \delta = 1, x)P(r \neq K|\delta = 1, x)P(\delta = 1|x). \qquad (2)
\end{aligned}$$

The worst-case lower bound relies on the assumption that unobserved periods correspond for sure to an unobserved labor market state $r = K$, i.e. $P(r \neq K|\delta = 1, x) = 0$. In this case, the second part of (2) is zero. The worst-case lower bound on the identification region of the CIC of risk $k \neq K$ is then given by:

$$I_k^{LB}(t|x) = P(T_k \leq t, r = k, \delta = 0|x). \qquad (3)$$

The worst-case upper bound assumes that there is for sure continued unemployment during an unobserved period, i.e. $P(r \neq K|\delta = 1, x) = 1$. In this case, τ_r is identified and $P(T_k \leq t, r =$

$k, r \neq K, \delta = 1|x)$ can be directly estimated from the data. The worst-case upper bound for the CIC is then:

$$\begin{aligned} I_k^{UB}(t|x) &= P(T_k \leq t, r = k, \delta = 0|x) \\ &+ P(T_k \leq t, r = k | r \neq K, \delta = 1, x) P(\delta = 1|x) \end{aligned} \tag{4}$$

Note that these worst case bounds assume the conditional probability $P(r \neq K|\delta = 1, x)$ to be either zero or one for all t. It directly follows that $I_k^{LB}(t|x) \leq I_k^{UB}(t|x)$ for all t.

The bounds given in (3) and (4) can be estimated nonparametrically by using Kaplan-Meier type estimators, as the censoring time T_{max} is independent (see Kalbfleisch and Prentice, 2002). Let $t_0 < \ldots < t_j < \ldots < t_J$ be the discrete times at which $\tau_{k \neq K}$, ς and t_{max} are observed. For the estimation of the lower bound (3) there are d_{kj}^{LB} observed exits to risk type $k \neq K$ at time t_j; d_{cj}^{LB} observed realizations of C at t_j; and d_{mj}^{LB} censored observations at $t_{max} = t_j$. For $k \neq K$, we have:

$$\begin{aligned} d_{kj}^{LB} &= \sum_{i=1}^{n} \mathbb{I}(\tau_{ik} = min\{\tau_{ik}, \varsigma_i, t_{i,max}\}) \\ d_{cj}^{LB} &= \sum_{i=1}^{n} \mathbb{I}(\varsigma_i = min\{\tau_{ik}, \varsigma_i, t_{i,max}\}) \\ d_{mj}^{LB} &= \sum_{i=1}^{n} \mathbb{I}(t_{i,max} = min\{\tau_{ik}, \varsigma_i, t_{i,max}\}) \end{aligned}$$

with $\mathbb{I}(Y)$ is the indicator function of the event Y.

In contrast, ς can be ignored for the estimation of the upper bound (4). Moreover, in this case we have $\tau_r = s_{r \neq K}$. Then we define d_{kj}^{UB} and d_{mj}^{UB} analogously as

$$\begin{aligned} d_{kj}^{UB} &= \sum_{i=1}^{n} \mathbb{I}(\tau_{ik} = min\{\tau_{ik}, t_{i,max}\}) \\ d_{mj}^{UB} &= \sum_{i=1}^{n} \mathbb{I}(t_{i,max} = min\{\tau_{ik}, t_{i,max}\}). \end{aligned}$$

Let $d_j^{LB} = \sum_{k=1}^{K-1} d_{kj}^{LB} + d_{cj}^{LB} + d_{mj}^{LB}$ and $d_j^{UB} = \sum_{r=1}^{K-1} d_{kj}^{UB} + d_{mj}^{UB}$. The number of observations at risk just before t_j is then given by

$$n_j^{LB} = d_j^{LB} + \ldots + d_J^{LB} \text{ and } n_j^{UB} = d_j^{UB} + \ldots + d_J^{UB}.$$

The Kaplan-Meier type estimators for the cause specific hazard rate and the overall survivor curve for the distribution of observed transition to state $k \neq K$ are

$$\hat{\lambda}_k^b(t_j|x) = d_{kj}^b / n_j^b \text{ with } b \in \{LB, UB\} \quad \text{and} \quad \hat{\lambda}_c^{LB}(t_j|x) = d_{cj}^{LB} / n_j^{LB};$$
$$\hat{S}^{LB}(t_j|x) = \prod_{u=1}^{j-1} \left(1 - \sum_{k=1}^{K-1} \hat{\lambda}_k^{LB}(t_u) - \hat{\lambda}_c^{LB}(t_u)\right) \quad \text{and} \quad \hat{S}^{UB}(t_j|x) = \prod_{u=1}^{j-1} \left(1 - \sum_{k=1}^{K-1} \hat{\lambda}_k^{UB}(t_u)\right). \tag{5}$$

Note that that these estimators are consistent as the right censoring is independent. A consistent estimator for the bounds given in (3) and (4) is then:

$$\hat{I}_k^b(t_j|x) = \sum_{u=1}^{j} \hat{\lambda}_k^b(t_u|x)\hat{S}^b(t_u|x) \quad \text{with} \quad b \in \{LB, UB\} \tag{6}$$

and $k \neq K$.

In analogy to Lee and Wilke (2008), we use the monotone relations given in (3) and (4) to bound a difference-in-differences estimator. Suppose there is a policy intervention in a natural experiment setting and we have $X = (G, P, Y)$. There are two groups, the control group ($G = g_0$) and the treatment group ($G = g_1$), and two time intervals, the pre-reform period ($P = p_{t0}$) and the post-reform period ($P = p_{t1}$). Y is a vector of other observable individual variables such as gender, age etc. The reform of interest is supposed to have an effect on the observed risk-specific transition distribution of the treatment group in the post-reform years. Under the assumption that the CIC of treatment and control group would have followed parallel paths without the reform, the effect of the reform can be estimated by a difference-in-differences estimator (DID) (see also Abadie, 2005 for a review of nonparametric identification of DID models) as

$$\Delta_{Ik}(t_j|y) = [I_k(t_j|g_1, p_{t1}, y) - I_k(t_j|g_0, p_{t1}, y)] - [I_k(t_j|g_1, p_{t0}, y) - I_k(t_j|g_0, p_{t0}, y)] \tag{7}$$

for $r = 1, \ldots, K-1$, where $I_k(t_j|g, p, y) = P(T_k \leq t_j, r = k|G = g, P = p, Y = y)$. Given that we can only identify intervals for the risk-specific cumulative incidence curve it is straightforward to bound Δ_{Ik} (Lee and Wilke, 2008):

$$\begin{aligned} l_{Ik}(t_j|y) = &\max[-1, \{I_k^{LB}(t_j|g_1, p_{t1}, y) - I_k^{UB}(t_j|g_0, p_{t1}, y)\} \\ &- \{I_k^{UB}(t_j|g_1, p_{t0}, y) - I_k^{LB}(t_j|g_0, p_{t0}, y)\}] \end{aligned} \tag{8}$$

and

$$\begin{aligned} u_{Ik}(t_j|y) = &\min[1, \{I_k^{UB}(t_j|g_1, p_{t1}, y) - I_k^{LB}(t_j|g_0, p_{t1}, y)\} \\ &- \{I_k^{LB}(t_j|g_1, p_{t0}, y) - I_k^{UB}(t_j|g_0, p_{t0}, y)\}] \end{aligned} \tag{9}$$

for $k = 1, \ldots, K-1$. Note that the lower and upper bound are restricted to be between -1 and 1. This is due to the fact that the maximum variation of probabilities cannot be larger than 1 in absolute values. The reform effect is estimated by replacing the upper and lower bounds by consistent estimators as defined in (6).

From (3)-(4) it can be seen that the width of the bounds of the DID changes in (8)-(9) depends on $P(r \neq K|\delta = 1, g, p, y)$ and $P(\delta = 1|g, p, y)$. As the worst-case bounds can be wide, there are several approaches to tighten them. In addition to monotonicity or independence assumptions as in Lee and Wilke (2008) one could use economic reasoning to tighten the feasible interval for

$P(r \neq K|\delta = 1, g, p, y)$ in (2). With this respect it is important to note that $P(r \neq K|\delta = 1, g, p, y)$ can be modelled as a function of time, while in (3) and (4) it was assumed to be constant. As an example, $P(r \neq K|\delta = 1, g, p, y)$ could be assumed to decrease with an increasing gap after ς. Another approach would be to increase the share of fully identified unemployment durations by determining an appropriate sample of unidentified spells that can be excluded from the estimations without changing the estimates for the CICs in (3) and (4). Formally, this approach reduces $P(\delta = 1|g, p, y)$.

Another possibility is to assume cross restrictions on the DID terms to preclude that some of them attain their lower and others their upper bound. As an example one can assume that the conditional probability $P(r \neq K|\delta = 1, g, p, y)$ is independent of G and P, i.e. $P(r \neq K|\delta = 1, g, p, y) = P(r \neq K|\delta = 1, y)$ which only depends on other individual characteristics Y. The DID changes of the CIC from (7) can then be decomposed as follows:

$$\Delta^c_{Ik}(t_j|y) \quad = \quad \Delta_{Ik}(t_j, \delta = 0|y) + P(r \neq K|\delta = 1, y)\Delta_{Ik}(t_j, \delta = 1|y) \tag{10}$$

with the effect of the reform on the CIC for different values of δ defined as :

$$\begin{aligned} \Delta_{Ik}(t_j, \delta = 0|y) &= I_k(t_j, \delta = 0|g_1, p_{t1}, y) - I_k(t_j, \delta = 0|g_0, p_{t1}, y) \\ &- I_k(t_j, \delta = 0|g_1, p_{t0}, y) + I_k(t_j, \delta = 0|g_0, p_{t0}, y), \text{ and} \qquad (11) \\ \Delta_{Ik}(t_j, \delta = 1|y) &= I_k(t_j, \delta = 1|r \neq K, g_1, p_{t1}, y) - I_k(t_j, \delta = 1|r \neq K, g_0, p_{t1}, y) \\ &- I_k(t_j, \delta = 1|r \neq K, g_1, p_{t0}, y) + I_k(t_j, \delta = 1|r \neq K, g_0, p_{t0}, y). \qquad (12) \end{aligned}$$

In order to determine the bounds we have to minimize and maximize (10) by assigning appropriate $P(r \neq K|\delta = 1, y)$ at each t_j: If $\Delta_{Ik}(t_j, \delta = 1|y) > 0$, set $P(r \neq K|\delta = 1, y) = 1$ and if $\Delta_{Ik}(t_j, \delta = 1|y) < 0$, set $P(r \neq K|\delta = 1, y) = 0$ to maximise (10). The minimum is attained in the reversed way. Thus the lower bound of $\Delta^c_{Ik}(t_j|y)$ is always smaller than the upper bound and the width of the bound is $|\Delta_{Ik}(t_j, \delta = 1|y)|$ which is tighter than the worst-case bound. Moreover, $P(r \neq K|\delta = 1, y)$ is now a function of t_j. Under the additional assumption, (11) and (12) are thus obtained by bounding the DID changes instead of bounding the CIC as in the case of the worst-case bounds in (8) and (9). A common difficulty of all approaches to tighten the bounds is to verify their validity in an application.

3 Empirical Application

We apply the above framework to bound the effect of reducing the maximum duration of receiving unemployment benefits on the observed transitions from unemployment to local and non-local employment via migration. We begin this section with a brief description of the German unemployment compensation system and discuss the 1997 reform of unemployment benefit entitlements. This discussion is based on the Employment Promotion Act (*Arbeitsförderungsgesetz*), the Social Welfare Act III (*Sozialgesetzbuch III*) and several secondary sources such as Plaßmann (2002), Oschmiansky et al. (2001) and Wolff (2003). We then introduce the data and discuss the selection of treatment and control group before we present our findings.

Basic features of the unemployment compensation system. During the study period, the system of unemployment compensation in Germany consisted of two main components: unemployment benefits (UNB) and unemployment assistance (UNA). As an insurance benefit, UNB is limited in time depending on the length of socially insured employment during a period of seven years before the benefit claim. Moreover, the length of benefit entitlements positively depends on age with a maximum UNB receipt of 12 months for younger age groups and up to 32 months for older age groups in the years prior to the 1997 reform. After exhausting UNB, unemployed individuals receive the tax-funded unemployment assistance if they pass a means-test. Both UNB and UNA correspond to a percentage of former wage income. UNB replaces 63% (68%) of former wage income and UNA still reaches income replacement rates of 53% (57%) for individuals without (with) dependent children. For individuals with low pre-unemployment wages, income replacement rates irrespective of the type of unemployment compensation may even be close to 100% because of receiving complementary social benefits. This is the case if the unemployment compensation as a percentage of former wage income does not suffice to ensure the legally defined minimum standard of living.

Due to this design, the effect of shortening the length of entitlements to unemployment benefits is not homogeneous. In particular, recipients of complementary social benefits are not affected by a change in the length of UNB receipt. By contrast, unemployed individuals without additional social benefits but with eligibility for the means-tested UNA loose around 10% of their former wage income when switching from UNB to UNA. For this group, a shortening of UNB is likely to have a small effect only. Individuals who do not pass the means test for receiving UNA due to having other income sources or private savings even loose all unemployment compensation after exhausting UNB. The threat of entitlement loss should thus be strongest for this latter group of unemployed.

1997 Reform. In April 1997, a reform of the Employment Promotion Act came into force to shorten entitlements to UNB for some of the older age groups. In Germany, the potential UNB duration (PUNBD), i.e. the maximum duration of UNB receipt an individual is entitled to at the beginning of the unemployment period, positively depends on the period of socially insured employment within the seven years prior to the benefit claim. This so called extended claim period is restricted by previous benefit claims and thus may be shorter than seven years. In addition, the PUNBD positively depends on age. During the 1980s, the PUNBD had successively been expanded for older age groups. Thus, before the reform in 1997, entitlements to UNB lasted up to 32 months for individuals above the age of 42, while the PUNBD for individuals below this age range was only 12 months. A detailed description of these earlier reforms can be found in Hunt (1995). One well-documented result of these earlier reforms that demonstrates the disincentive effect of this system was the rapid increase of early retirees whose extremely long UNB receipt allowed for bridging the gap between employment and retirement age (Fitzenberger and Wilke, 2004).

In 1997, the PUNBD was reduced for some of the older age groups by lowering the age limits for certain maximum entitlement length (see Table 1). As a consequence, the PUNBD for individuals between 42 and 43 years of age was cut from 18 month before 1997 to 12 month after the 1997 reform. For individuals aged 44, UNB was even cut from a maximum receipt of 22 to a maximum receipt of 12 months. Individuals aged below 42 years were unaffected by the reform as they always received a maximum of 12 month of UNB. The 1997 reform thus provides a natural experiment with a credible source of variation in PUNBD that can be used to identify its causal effect. As a drawback, however, the implementation of the reform was partially cushioned. Until March 1999, new benefit claimants were treated according to the pre-reform regulations if there was a work history of more than one year during the three years prior to the benefit claim. Thus, the new regulations applied to all new benefit claims after March 1999 only. In addition, the introduction of stricter sanction rules for non-compliance with eligibility requirements may have accelerated transitions from unemployment to employment for all age groups after 1997 (Boone et al., 2002, 2004).

Table 1: Potential unemployment benefit duration (PUNBD) for UNB claimants up to age 47 by work history and age, IAB-R01

Soc. insured employment during claim period	**PUNBD (in month)** until 03/97	since 04/97
12 month	6	6
16 month	8	8
20 month	10	10
24 month	12	12
28 month	14 (age ≥42)	14 (age ≥45)
32 month	16 (age ≥42)	16 (age ≥45)
36 month	18 (age ≥42)	18 (age ≥45)
40 month	20 (age ≥44)	20 (age ≥47)
44 month	22 (age ≥44)	22 (age ≥47)

Source: Plaßmann (2002)

Two German studies already looked at the effect of the 1997 reform on transitions from unemployment to employment. Based on the German socio-economic panel (GSOEP), Wolff (2003) only finds very weak positive effects of shortening the PUNBD on the transitions to employment in eastern Germany. This finding may reflect the limited sample size of the GSOEP data as the study includes only a limited number of spells that were actually affected by the reform. Based on much more extensive administrative data, Müller et al. (2007) find strong evidence that the 1997 reform reduced the inflow into unemployment and the duration of unemployment among individuals above age 52. They reason that a shortening of the maximum UNB receipt mainly lowers the attractiveness of early retirement through the unemployment compensation system.

In the subsequent analysis, we use an administrative data set that is similar to the one used by Müller et al. (2007), but with an even larger 50% sample of the male working population. The data set thus allows for distinguishing between exits to local versus exits to non-local employment after migration. We can thus answer the question how job seekers changed their behavior in response to the reform by analyzing the effect of the PUNBD on transitions to local and non-local employment. We restrict the analysis to prime age individuals for whom early retirement should not be an issue.

Data. We use a sample drawn from the Employee and Benefit Recipient History (V6.0) of the Institute of Employment Research (IAB). This administrative data set contains information of employment periods that are subject to social insurance payments concerning spells of employment and spells for which the individual received unemployment compensation from the Federal

Employment Agency (*Bundesagentur für Arbeit*) such as unemployment benefits (UNB), unemployment assistance (UNA) and maintenance payments during training measures (MP). Our analysis is based on a 50% sample of the male working population. The data was prepared by the IAB to have the same structure as the IAB employment subsample 1975-2001 - regional file which is a 2% sample and available as a scientific use file (Hamann et al., 2004). As the access to the 50% sample is restricted, we did all the preliminary work with the 2% sample and switched to the 50% sample for the final and main estimations only.

The data consists of administrative records which are provided as spells on a daily basis. However, individual employment histories are only partly identified as there may be gaps in an individual's record. These gaps correspond to either a period of unemployment without receiving unemployment compensation or to other unobserved labor market states such as being self-employed, a civil servant or being out of labor force. This is why the true unemployment duration is often not observed (see also Lee and Wilke, 2008). An unemployment period is identified only if there is a permanent receipt of unemployment compensation between two employment spells. In this case, an unemployment spell belongs to the case $\delta = 0$. In the application, this requires an individual to receive unemployment compensation within one month after the end of an employment period. Moreover, intermediate gaps in the receipt of unemployment compensation or the gap between the end of transfer receipt and employment do not exceed one month. However, gaps in the record will often be longer and the true unemployment period is not identified. In this case, an observation belongs to the case $\delta = 1$ and we observe ς which is the beginning of an unobserved period.

As discussed in the previous section, we can derive an upper bound and a lower bound definition for the unemployment spell in the case of $\delta = 1$. The lower bound definition LB assumes that an exit to an unknown labor market state occurs at the beginning of the unobserved period, i.e. at ς. In contrast, the upper bound definition UB assumes that no such exit occurs in an unobserved period and thus interprets any intermediate unobserved period as a continued unemployment spell. It thus equates a period of non-employment with a period of unemployment. Note that in case of a fully identified unemployment spell ($\delta = 0$), both the upper bound and lower bound definition of unemployment yield the same spell length. Moreover, exits to an unobserved labour market state only occur in case of the lower bound definition of unemployment. In the data, right-censoring occurs at the end of the observation period, i.e. on 12/31/2005.

For all unemployed with a transition to employment, the IAB data allows for comparing the location of the old and the new workplace on the level of the 440 German counties. In the following analysis, we assume a job movement to involve a residential relocation if the distance between the county capitals of the old and the new workplace exceeds 100 km. Labor market regions that comprise daily commuting ranges typically do not exceed 75 km in Germany. Most job movements with a distance of 100 km between the old and new workplace region should thus necessitate residential mobility. For each spell of unemployment, the analysis thus distinguishes between exits to a local job, exits to a non-local job after migration and exits to other unknown destination states.

For our analysis, we use inflow samples for a pre- and a post-reform era. Due to the implementation of stricter sanction rules in 1994, extending the pre-reform era beyond 1995, might mix different reforms. We therefore consider an unemployment spell starting between 1995 and 1996 as a pre-reform spell. The post-reform era is predetermined by the fact that the implementation of new UB regulations did not start before 1999. The post-reform inflow sample thus consists of all unemployment spells starting in 1999 or 2000. Moreover, in order to keep the sample relatively homogeneous in terms of labor force attachment, we restrict our sample to men who have previously been full-time employed. In addition, we restrict the analysis to individuals born in western Germany because the working history for individuals from eastern Germany is not known before 1991 which aggravates the computation of entitlement lengths to unemployment benefits.

As discussed before, the reform effect is expected to be weaker for recipients of complementary social benefits. Unfortunately, the IAB data does not include enough information on the household context to actually identify recipients of complementary social benefits. In our analysis, we therefore use a lower educational degree as a proxy for lower earning capacities that increases the probability of receiving social benefits in addition to unemployment compensation. Indeed, we obtain similar result pattern when we repeat the empirical analysis for different wage levels similar to Lee and Wilke (2008). In order to take account of heterogeneous treatment effects, we therefore distinguish between high-skilled individuals who either have a tertiary education or are master craftsmen and less-skilled individuals. High-skilled individuals who are less likely to receive complementary social benefits are expected to react stronger to the shortening of the maximum receipt of unemployment benefits.

Reform effects are also likely to differ depending on the household context. In particular, being married and having dependent children have typically been found to raise migration costs (see Ghatak et al., 1996). For these households, a shortening of PUNBD is thus less likely to provoke a higher level of migration than for single households. Distinguishing between different household contexts may thus be an additional important distinction in order to identify heterogeneous reform effects. Since the data does not include information on dependent children, but only on the marital status, the following analysis distinguishes between four groups: high- or less skilled singles and high- or less-skilled married individuals. High-skilled individuals are expected to react stronger to the cut in PUNBD. In addition, married individuals are less likely to respond to the reform by higher migration rates than their single counterparts.

Choosing the treatment and control group. Due to the reform in 1997, eligibility to an extended UNB duration of more than 12 month was cut for individuals aged 42-44 years, while the PUNBD of individuals below this age was unaffected by the reform. Thus individuals aged 36-41 years serve as the group to control for changing labor market conditions when comparing transitions to local and non-local employment before and after the reform. However, since only individuals with long UNB entitlements are affected by the reform, the exact choice of treatment and control group has to be conditioned not only on age, but also on the entitlement length at the beginning of the unemployment period. This is because choosing the treatment and control group based on their actual UNB entitlements results in a non-comparability of individuals in the control and treatment group with regard to their working history because the criterium to reach maximum entitlements is less strict for the younger cohort (see Table 1).

In order to ensure that treatment and control group are comparable with regard to their working history, a suitable selection rule should thus be the same for both groups. For this purpose, we compute counterfactual UNB entitlements, i.e. hypothetical UNB entitlements in the absence of the 1997 reform had the individual been aged 42-44 at the time of benefit claim (see Appendix A for details). As can be seen in Table 2, the resulting counterfactual UNB entitlements are quite comparable for both age groups. For the subsequent analysis, we choose all unemployment spells that begin with a receipt of unemployment benefits and whose counterfactual UNB duration exceeds 12 month. This selection rule ensures the comparability between the treatment and the control group and the existence of some minimum treatment for the treatment group. For characteristics that are observable in the IAB data, Appendix B confirms that treatment and control group are mostly quite comparable. Small but notable differences can be found in the share of married individuals and the educational degree. Since these are important characteristics for the expected strength and type of reaction to the reform, however, we separately estimate reform effects for high- and less-skilled singles and married individuals anyway. Our DiD approach assumes that both treatment and control group experience similar changes in labor market conditions in the post- compared with the pre-reform period. We cannot test this assumption, but as the average age between the two groups differs by five years only, we believe that the extend of non parallel changes is small.

Table 2: Estimated counterfactual UNB entitlement length for unemployment spells in the pre- and post-reform era by age group[a], IAB data

	Age 36-41		Age 42-44	
UNB duration	# spells	%	# spells	%
$\leq$ 2 months	24,469	7.7	10,131	8.2
3-4 months	18,340	5.7	7,289	5.9
5-6 months	19,133	6.0	7,706	6.2
7-8 months	19,659	6.2	7,783	6.3
9-10 months	20,556	6.4	8,079	6.5
11-12 months	18,354	5.8	6,969	5.6
13-14 months	18,748	5.9	7,091	5.7
15-16 months	18,824	5.9	6,920	5.6
17-18 months	161,045	50.4	61,592	49.9
Total	319,128	100.0	123,560	100.0

[a] Includes all previously full-time employed men born in West Germany whose unemployment spell starts with the receipt of unemployment benefits.

For the resulting control group and the post-reform treatment group, Table 3 shows that the estimated actual entitlement length that is subject to the 1997 reform and the true age of the individual is up to 12 months only. Note that for some individuals who do not fulfill the criterium for the maximum entitlement length, but still pass the selection criterium, the true UNB duration may be lower than 12 months. In the pre-reform era, the treatment group is entitled to 18.5 month of UNB receipt on average, while in the post-reform era this average UNB duration falls to 11.8 month. This latter UNB receipt corresponds to the UNB duration for the control group in the pre- and post-reform era. The average treatment thus is a reduction of UNB entitlements of 6.7 month with the treatment ranging from a reduction of one to a reduction of ten month for individuals aged 44 with maximum UNB entitlements.

Table 3: Estimated actual UNB entitlement length for unemployment spells with counterfactual UNB of >12 months in the pre- and post-reform era by treatment and control group, IAB data

	Control group		Treatment group	
UNB duration	pre-1997	post-1997	pre-1997	post-1997
6-8 months	2.2%	1.6%	0.00%	1.6%
9-11 months	7.2%	5.6%	0.00%	5.3%
12 months	90.6%	92.8%	0.00%	93.1%
13-14 months	0.0%	0.0%	6.9%	0.0%
15-16 months	0.0%	0.0%	7.5%	0.0%
17-18 months	0.0%	0.0%	58.6%	0.0%
19-20 months	0.0%	0.0%	2.6%	0.0%
21-22 months	0.0%	0.0%	23.3%	0.0%
Average months	11.8	11.8	18.5	11.8
Total spells	104,069	94,309	39,434	36,104

Table 4 shows exit types and median unemployment duration for both the upper and lower bound definition of unemployment. Exits to unknown destination states only occur in case of the lower bound definition. Moreover, note that in the lower bound definition, around 40% of all unemployment spells exit to an unknown labor market state. In other words, 40% of the observations are only partially identified and belong to the case $\delta = 1$. According to the lower bound definition, around 50% (10%) of all unemployment spells exit to a local (non-local) job. According to the upper bound definition that rules out any exits to an unknown state, around 70% (15%) of all unemployment spells end in a local (non-local) job, while the rest is right-censored. Moreover, the degree of right-censoring is more pronounced in the post-reform year. Both unemployment definitions point towards a somewhat higher probability of finding non-local employment via migration in the post-reform years.

Note also that median unemployment durations are shorter for all groups in the post-reform years. As has been discussed previously, this may reflect a combination of better labor market conditions compared to the pre-reform years as well as the stricter sanction rules that applied to both the control and the treatment group. Moreover, the descriptive statistics for both unemployment definitions suggest that the treatment group has a somewhat longer unemployment duration and that the gap between treatment and control group becomes somewhat smaller for the upper bound definition of unemployment only. There is thus no clear descriptive evidence from Table 4 in favor of a reform effect. However, this may be due to the fact that the treatment effect is unlikely to be homogeneous. Appendix B thus additionally displays the median unemployment durations for the four sub-samples of less- and high-skilled single and married men and suggests that the positive gap in the median unemployment duration between the treatment and control group in the pre-reform year becomes much smaller in the post reform years for high-skilled singles, while only small reductions can be found for married men. The following analysis thus distinguishes between these four groups of unemployed.

Table 4: Descriptive summary of full sample, IAB data

	Control group		Treatment group	
	pre-1997	post-1997	pre-1997	post-1997
LB spells				
median duration (days)	185	152	205	172
exit to local job	50.1%	50.8%	48.1%	49.4%
exit to non-local job	8.7%	10.4%	8.9%	10.2%
exit to other destination	41.2%	38.8%	43.0%	40.4%
right-censored	0.0%	0.0%	0.0%	0.0%
total exits	100.0%	100.0%	100.0%	100.0%
UB spells				
median duration (days)	258	204	307	232
exit to local job	75.2%	71.1%	72.5%	69.6%
exit to non-local job	14.1%	15.2%	14.2%	14.7%
exit to other destination	0.0%	0.0%	0.0%	0.0%
right-censored	10.7%	13.7%	13.3%	15.7%
total exits	100.0%	100.0%	100.0%	100.0%
Total spells	104,069	94,309	39,434	36,104

Bounds Analysis We estimate bounds for the effect of the reform on the cumulative incidence of non-local and local re-employment for different skill groups by applying formulas (8)-(9). Moreover, we also add the asymptotically valid 90% joint confidence intervals for upper and lower bounds, which are computed following the bootstrap procedure of Horowitz and Manski (2000), also applied in an earlier version of Lee and Wilke (2008). The bootstrap repetitions are 500. Figure 3 shows that when applying the extreme bounds the partial identification problem of our interval data precludes any clear result pattern as none of the bounds cross the zero line during the treatment period.

As another interesting observation, we find that the resulting bounds do not coincide with the point estimates for the lower and upper bound of the latent variable. As shown in Appendix C, point estimates for the different definitions of the unemployment duration data do not span the full width of our estimated bounds. This suggests that a sensitivity analysis based on different transition time definitions alone may be misleading. Moreover, note that we generally observe a smooth variation of the bounds with the duration of unemployment. This does not suggest any remarkable jumps in the hazard rate or survivor function at the begin of the treatment. Our results therefore support the theoretical results of non-stationary job-search (van den Berg, 1990).

Approaches to tighten the bounds As indicated in the theoretical part there are several approaches to tighten the bounds. In a first attempt, we assume monotonicity and independence as done by Lee and Wilke (2008). Bounds become tighter but in many cases they are still wide. Moreover, by using the 2% sample we found some indication that the independence assumption may not be valid as bounds cross for some short intervals in few cases. Another natural attempt to tighten the bounds is to restrict the interval for $P(r \neq K|\delta = 1)$. One could for example assume a shorter interval such as $[0.1, 0.9]$. Alternatively, one could also assume that $P(r \neq K|\delta = 1)$ decreases with the distance between ς and $s_{k \neq K}$, i.e. the length of the unobserved period. In other words, if the unobserved period is short, it is less likely that someone transits to an unobserved labor market state such as self-employment. As an example, one could assume that $P(r \neq K|\delta = 1)$ is an exponential density. We have estimated the bounds under several such scenarios using the 2% sample of the IAB data. Resulting bounds are tighter and suggest that the change in the CICs is positive for the high-skilled group. Results are, however, not presented as it is difficult to justify these assumptions. These approaches may, however, be justifiable in other applications. As an example, a tightening of bounds could be achieved by estimating $P(r \neq K|\delta = 1)$ with additional information from survey data.

Additional assumptions concerning $P(r \neq K|\delta = 1)$ are thus able to tighten the bounds but appear arbitrary in the context of our application. Since the share of partly identified spells ($\delta = 1$) in our sample is about 40%, another approach is to reduce it. As an attempt, we exclude observations that do not start with the receipt of unemployment compensation within one month after the end of employment, i.e. we exclude spells with $\delta = 1$ and $\varsigma = 0$. This is done because these spells are least informative in the sense that the interval $[\varsigma, s_{k \neq K}]$ is often large. The resulting bounds are again tighter and suggest some weak increase in the CICs for the high skilled while there are no apparent changes for the less skilled. As discussed above, this approach is only valid if the exclusion of spells is a random sample in the sense that the sample composition does not change and the shape of the conditional CICs in (3) and (4) does not change. As the latter condition is not well supported by the 2% data, we decided not to proceed in this way. Although these attempts seem unappropriate for our application, we decided to mention them as they may be more suitable for other research.

Our final attempt to tighten the bounds is therefore to impose the additional independence assumption $P(r \neq K|\delta = 1, g, p, y) = P(r \neq K|\delta = 1, y)$. The resulting bounds in Figure 4 are much tighter. For less skilled individuals, however, we do not find much of an effect irrespective of the marital status and thus only display the result patterns for all less-skilled males together. In contrast, Figure 4 suggests larger changes in observed exit probabilities for high-skilled job seekers for whom the threat of entitlement loss after exhausting UNB is likely to be larger. Moreover, we find heterogeneous result patterns for high-skilled men depending on the marital status. While the bounds for single males - although only scratching the significance level - weakly suggest a higher probability of migration as a main reaction to a cut in PUNBD, we find a strong and significant positive effect of the cut in PUNBD on the probability of finding local employment among married men.

If we assumed independent risks, these findings would have a causal interpretation in the sense that extensive unemployment benefits among singles mainly allow for avoiding or postponing migration such that the reduction of UNB entitlements primarily fosters the willingness to migrate. From a theoretical perspective, this finding is quite plausible in light of the institutional design in Germany because the counteracting resource effect suggested by Tatsiramos (2008) is likely to be small. This is because unemployed individuals irrespective of whether receiving UNB or UNA get financial support for search costs and moving costs. The positive effect of lower reservation wages in case of a cut in PUNBD should thus likely exceed the negative resource effect. However, in the case of married men, high migration cost seem to dominate the response to a cut in PUNBD. For married men, extensive unemployment benefits rather seem to allow for extending unemployment before re-entering local employment. As a consequence, a cut in UNB entitlements is unlikely to foster migration among married unemployed.

We thus find some interesting evidence that the strength and the type of reaction to reforms of the unemployment compensation system critically hinge on the household context. This may also explain why Tatsiramos (2008) could not establish any average effect of unemployment benefits on migration in Germany. However, our findings are only suggestive for some reform effects on leaving unemployment locally or non-locally as the effects on the marginal distributions for the risk of finding employment locally or non-locally is not identified in our econometric model without imposing additional assumption on the dependence structure between risks. In a follow up paper, Lo and Wilke (2008) check the robustness of our result pattern with respect to the assumed dependence structure using the 2% sample of our data. They find that the sign of the estimated treatment effect is indeed quite robust. For this reason we believe that the sign of changes in the CIC is also a natural candidate for the sign of the true treatment effect.

4 Conclusion

This paper has presented a nonparametric approach that allows for analyzing a competing risk model with partially identified interval data. Our bounds analysis is a highly relevant approach for applied researchers who face this data limitation. It extends the nonparametric bounds analysis by Lee and Wilke (2008) to a dependent competing risk setting and derives bounds for the risk-specific cumulative incidence curve. Although our approach does not resolve the non-identifiability of competing risks and thus precludes a direct causal inference, it provides a flexible descriptive tool for the observed risk-specific transition distribution. In particular, our approach is fully nonparametric and we avoid strong assumptions on our duration model that may be violated in the real world. Moreover, we suggest several approaches to tighten the bounds in an application.

In our empirical application with German data, we have explored the effect of reducing the receipt of unemployment benefits on the observed transitions to either local or non-local employment via migration. Our results show that partial identification is a big problem in merged administrative individual data that may preclude any causal inference. Moreover, point estimates for the lower and upper bound of the latent variable do not span the full width of our estimated bounds. Therefore, a sensitivity analysis based on the two point estimates alone may be misleading. We obtain considerably tighter bounds by imposing additional assumptions. Similar to Lee and Wilke (2008) the resulting estimates are suggestive for a reform effect on observed exit probabilities for high-skilled individuals for whom the threat of entitlement loss after exhausting UNB is likely to be largest. Our results indicate that the effect of extensive unemployment benefits strongly depends on the family background. The cumulative incidence for migration increases for singles while it remains unchanged for married men. In contrast, we do not observe changes in the cumulative incidence for local jobs for singles while it increases significantly for married men. Unfortunately, administrative data does contain many household background variables. For this reason, we are not able to capture the decision process about the location of a future job in more detail.

Other limitations of our approach point towards some interesting extensions. With regard to data limitations, data with more information on individual and household characteristics would be desirable to reexamine our empirical results. Such additional information would also allow to distinguish groups for whom a shorter receipt of unemployment benefits implies different entitlement losses. In addition, the causal inference from our empirical results is limited because of the unresolved identification problem of the competing risks. A promising route for future research thus is to combine our bounds framework for partially missing data with attempts to derive tighter identification bounds such as Honoré and Lleras-Muney (2006) or to assume a dependence structure as in Lo and Wilke (2008). However, as a disadvantage to our current bounds framework for cumulative incidence curves, such attempts necessitate additional assumptions. Moreover, a tightening of bounds could be achieved by estimating $P(r \neq K|\delta = 1)$ with additional information. This could be done for example if a sample of the administrative data would be merged with survey data that fully identifies the employment trajectories.

References

[1] Bundesanstalt für Arbeit (2001), IAB-Kurzbericht Nr. 1/2001, Nürnberg: Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit.

[2] Abbring, J.H. and G.J. van den Berg (2003), The identifiability of the mixed proportional hazards competing risks model, *Journal of the Royal Statistical Society B* Vol. 65, 701–710.

[3] Abadie, A. (2005), Semiparametric Difference-in-Differences Estimators, *Review of Economic Studies* Vol. 72, 1–19.

[4] Antolin, P. and O. Bover (1997), Regional Migration in Spain: The Effect of Personal Characteristics and of Unemployment, Wage and House Price Differentials Using Pooled Cross-Sections, *Oxford Bulletin of Economics and Statistics* Vol. 59, 215–35.

[5] Arntz, M. (2005), *The Geographical Mobility of Unemployed Workers*, ZEW Discussion Paper No. 05-34, Mannheim.

[6] Arntz, M. and R.A. Wilke (2008), Unemployment Duration in Germany: Individual and Regional Determinants of Local Job Finding, Migration and Subsidized Employment, *Regional Studies*, forthcoming.

[7] Atkinson, A.B. and J. Micklewright (1991), Unemployment compensation and labor market transitions: A critical review, *Journal of Economic Literature* Vol. 29, 1679-1727.

[8] Barron, J.M. and W. Mellow (1979), Search Effort in the Labor Market, *The Journal of Human Resources* Vol. 14, 389–404.

[9] Boone, J., P. Fredriksson, B. Holmlund, and J.C. van Ours (2002) *Optimal Unemployment Insurance with Monitoring and Sanctions*, IFAU Working Paper 21, Uppsala.

[10] Boone, J., A. Sadrieh, and J.C. van Ours (2004), *Experiments on Unemployment Benefit Sanctions and Job Search Behavior*, IZA Discussion Paper No. 1000, Bonn.

[11] Card, D.E. and P.B. Levine (2000), Extended benefits and the duration of UI spells: Evidence from the New Jersey Extended Benefit Program, *Journal of Public Economics* Vol. 78, 107-138.

[12] Cox, D.R. (1962), *Renewal Theory*, London.

[13] Fitzenberger, B. and R.A. Wilke (2004), *Unemployment Durations in West-Germany Before and After the Reform of the Unemployment Compensation System During the 1980s*, ZEW Discussion Paper No. 04-24, Mannheim.

[14] Fitzenberger, B. and R.A. Wilke (2007), *New Insights on Unemployment Duration and Post Unemployment Earnings in Germany: Censored Box-Cox Quantile Regression at Work*, ZEW Discussion Paper No. 07-007, Mannheim.

[15] Ghatak, S., Levine, P., and S.W. Price (1996), Migration Theories and Evidence: An Assessment, *Journal of Economic Surveys* Vol. 10, 159–198.

[16] Goss, E. and C. Paul (1990), The Impact of Unemployment Insurance Benefits on the Probability of Migration of the Unemployed, *Journal of Regional Science* Vol. 30, 349–358.

[17] Hamann, S., G. Krug, M. Köhler, W. Ludwig-Mayerhofer, and A. Hacket (2004), Die IAB-Regionalstichprobe 1975-2001: IABS-R01, *ZA-Information* Vol. 55, 36–42.

[18] Hassler, J., J.V. Rodriguez Mora, K. Storesletten, and F. Zilibotti (2005), A positive theory of geographic mobility and social insurance, *International Economic Review* Vol. 46, 263-303.

[19] Heckman, J.J. and B.E. Honoré (1989), The identifiability of the competing risks model, *Biometrika* Vol. 76, 325–330.

[20] Honoré, B.E. and A. Lleras-Muney (2006), Bounds in Competing Risks Models and the War on Cancer, *Econometrica* Vol. 74, 1675–1698.

[21] Horowitz, J.L. and C.F. Manski (2000), Nonparametric analysis of randomized experiments with missing covariate and outcome data (with discussion), *J. Amer. Statist. Assoc.* Vol. 95, 77–88.

[22] Hunt, J. (1995), The Effect of Unemployment Compensation on Unemployment Duration in Germany, *Journal of Labor Economics* Vol. 13, 88-120.

[23] Kalbfleisch, J.D. and R.L. Prentice (1980), The Statistical Analysis of Failure Time Data, *Wiley Series in Probability and Statistics.*

[24] Katz, L.F. and B.D. Meyer (1990), The impact of the potential duration of unemployment benefits on the duration of unemployment, *Journal of Public Economics* Vol. 41, 45-72.

[25] Lalive, R. and J. Zweimüller (2004), Benefit Entitlement and Unemployment Duration: Accounting for Policy Endogeneity, *Journal of Public Economics* Vol. 88, 2587–2616.

[26] Lee, S. and R.A. Wilke (2008), Reform of Unemployment in Germany: A nonparametric Bounds Analysis using Register Data, *Journal of Business and Economic Statistics*, forthcoming.

[27] Lo, S.M.S. and R.A. Wilke (2008), A copula model for dependent competing risks, mimeo.

[28] Manski, C.F. (2003), *Partial Identification of Probability Distributions*, New York.

[29] Moeschberger, M.L. and J.P. Klein (1995), Statistical Methods for Dependent Competing Risks, *Lifetime Data Analysis* Vol. 1, 195–204.

[30] Müller, E., R.A. Wilke, and P. Zahn (2007), Beschäftigung und Arbeitslosigkeit älterer Arbeitnehmer: Eine mikroökonometrische Evaluation der Arbeitslosengeldreform von 1997, *Jahrbücher für Nationalökonomie und Statistik* Vol. 227, 65-86.

[31] Oschmiansky, O., S. Kull, and G. Schmid (2001), *Faule Arbeitslose? Politische Konjunkturen einer Debatte*, Discussion Paper FS 01-206, Wissenschaftszentrum Berlin für Sozialforschung.

[32] Peterson, A.V. (1976), Bounds for a Joint Distribution With Fixed Sub-Distribution Functions: Application to Competing Risks, *Proceedings of the National Academy of Science* Vol. 73, 11–13.

[33] Plaßmann, G. (2002), *Der Einfluss der Arbeitslosenversicherung auf die Arbeitslosigkeit in Deutschland. Beiträge zur Arbeitsmarkt- und Berufsforschung*, No. 255, Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit (IAB), Nürnberg.

[34] Tannery, F.J. (1983), Search Effort and Unemployment Insurance Reconsidered, *The Journal of Human Resources* Vol. 18 432–440.

[35] Tatsiramos, K. (2008), Geographic Labour Mobility and Unemployment Insurance in Europe, *Journal of Population Economics*, forthcoming.

[36] Tsiatis, A. (1975), A Nonidentifiability Aspect of the Problem of Competing Risks, *Proceedings of the National Academy of Sciences* Vol. 72, 20–22.

[37] Van den Berg, G.H. (1990), Nonstationarity in Job Search Theory, *Review of Economic Studies* Vol. 57, 255–277.

[38] Van Ours, J.C. and M. Vodopivec (2006), How Shortening the Potential Duration of Unemployment Benefits Affects the Duration of Unemployment: Evidence from a Natural Experiment, *Journal of Labor Economics* Vol. 24, 351-378.

[39] Wolff, J. (2003), *Unemployment Benefits and the Duration of Unemployment in East Germany*, Sonderforschungsbereich 386 Discussion Paper, No. 344.

Appendix A - Computation of actual and counterfactual UNB entitlements

The entitlement length at the beginning of the unemployment spell is not included in the data and has to be computed based on the known employment history, age and the known regulations and changes across time. For this purpose, we compute the claim period which encompasses a maximum of three years prior to making the UNB claim, but ends with a previous UNB claim within this three years period. In the same token, we calculate the employment duration within the relevant extended claim period of up to seven years prior to making the claim. As previously mentioned, UNB entitlements depend on the duration of socially insured employment within the relevant claim and the relevant extended claim period. Unemployment benefits exceeding 6 month necessitate at least 12 month socially ensured employment within the claim period. Thus, an individual with at least 12 month socially ensured employment within the claim period and 24 month within the extended claim period gets 12 month of UNB. If there is a shortened claim period due to a previous UNB claim, the new UNB claim based on the employment periods after this last unemployment period may be extended up to the age-specific PUNBD by remaining entitlements at the end of the previous unemployment period if the beginning of the last UNB claim lies within the last seven years.

For the estimation of actual UNB entitlements all changing regulations throughout the 1980s and 1990s have been applied. For the counterfactual UNB entitlements, we apply the pre-reform conditions to the post-reform period and compute the UNB entitlements as if all individuals had been 42 by the time of the benefit claim. More precisely, we adjust the whole age history of an individual as if, for example, an individual aged 38 at the beginning of the unemployment period had always been four years older. This adjustment alone does not ensure the comparability of the resulting counterfactual entitlements for the pre- and post-reform period because entitlements depend on the entire work history which is subject to all previous changes in regulations. We therefore compute the counterfactual entitlements for the post-reform period had all changes in regulations been shifted by five years, the difference between the pre- and post-reform period. This procedure ensures a twofold: (i) the comparability of counterfactual UNB entitlements for all age groups irrespective of whether the unemployment period starts prior or after the reform and (ii) the equivalence of counterfactual and actual UNB entitlements for the treatment group in the pre-reform era. As a consequence, the treatment group in the pre-reform period with counterfactual UNB entitlements of more than 12 month actually has entitlements of more than 12 month while all others who fulfil this criterium actually receive UNB for a maximum of 12 month only, but are comparable to the former group in terms of their employment history.

Appendix B

Table 5: Descriptive summary of sample characteristics, IAB data

	Control group		Treatment group	
Age (years)	38.3	38.3	43.0	43.0
Married	64.7	60.0	71.0	67.4
High school degree	19.9	17.7	18.5	18.2
Vocational training	72.9	75.1	74.8	75.1
Tertiary education	7.2	7.2	6.8	6.7
High-skilled single	2.7	3.2	1.9	2.3
Less-skilled single	32.6	36.9	27.2	30.3
High-skilled married	5.3	4.8	6.0	5.0
Less-skilled married	59.4	55.1	65.0	62.3
Skilled blue-collar	42.1	41.5	42.9	41.0
Unskilled blue-collar	32.3	31.8	30.3	32.0
White-collar	25.6	26.7	26.8	27.0
1st wage quintile	33.9	34.6	34.0	35.4
2nd wage quintile	23.0	24.1	21.8	22.9
3rd wage quintile	15.7	16.1	15.2	15.4
4th wage quintile	14.1	13.2	14.1	12.8
5th wage quintile	13.4	12.0	14.9	13.5
Tenure prev. job (days)	1253	1287	1489	1481
Previously unemployed	63.1	71.5	58.2	69.6
Number of prev. unempl. spells	2.8	3.0	2.5	3.0
Total spells	104,069	94,309	39,434	36,104

Table 6: Median unemployment duration by sub-sample and definition of unempoyment

	Control group		Treatment group	
	pre-1997	post-1997	pre-1997	post-1997
LB spells				
high-skilled singles	274	178	364	214
less-skilled singles	243	183	303	214
high-skilled married men	184	123	233	165
less-skilled married men	159	134	183	154
UB spells				
high-skilled singles	488	278	648	369
less-skilled singles	336	243	449	291
high-skilled married men	367	246	434	305
less-skilled married men	211	177	245	200

Appendix C

The point estimates for the treatment effect in Figure 5 are obtained by using the following formulas:

$$\begin{aligned} l_{Ik}(t_j|p_{t0},p_{t1},x) &= \{I_k^{LB}(t_j|1,p_{t1},x) - I_k^{LB}(t_j|0,p_{t1},x)\} \\ &\quad - \{I_k^{LB}(t_j|1,p_{t0},x) - I_k^{LB}(t_j|0,p_{t0},x)\} \\ \text{and} & \\ u_{Ik}(t_j|p_{t0},p_{t1},x) &= \{I_k^{UB}(t_j|1,p_{t1},x) - I_k^{UB}(t_j|0,p_{t1},x)\} \\ &\quad - \{I_k^{UB}(t_j|1,p_{t0},x) - I_k^{UB}(t_j|0,p_{t0},x)\} \end{aligned}$$

for $k = 1, \ldots, K-1$.

Figure 3: Lower and upper bound of the DiD changes of the cumulative incidence of local (left) and non-local (right) exits to employment among selected groups.

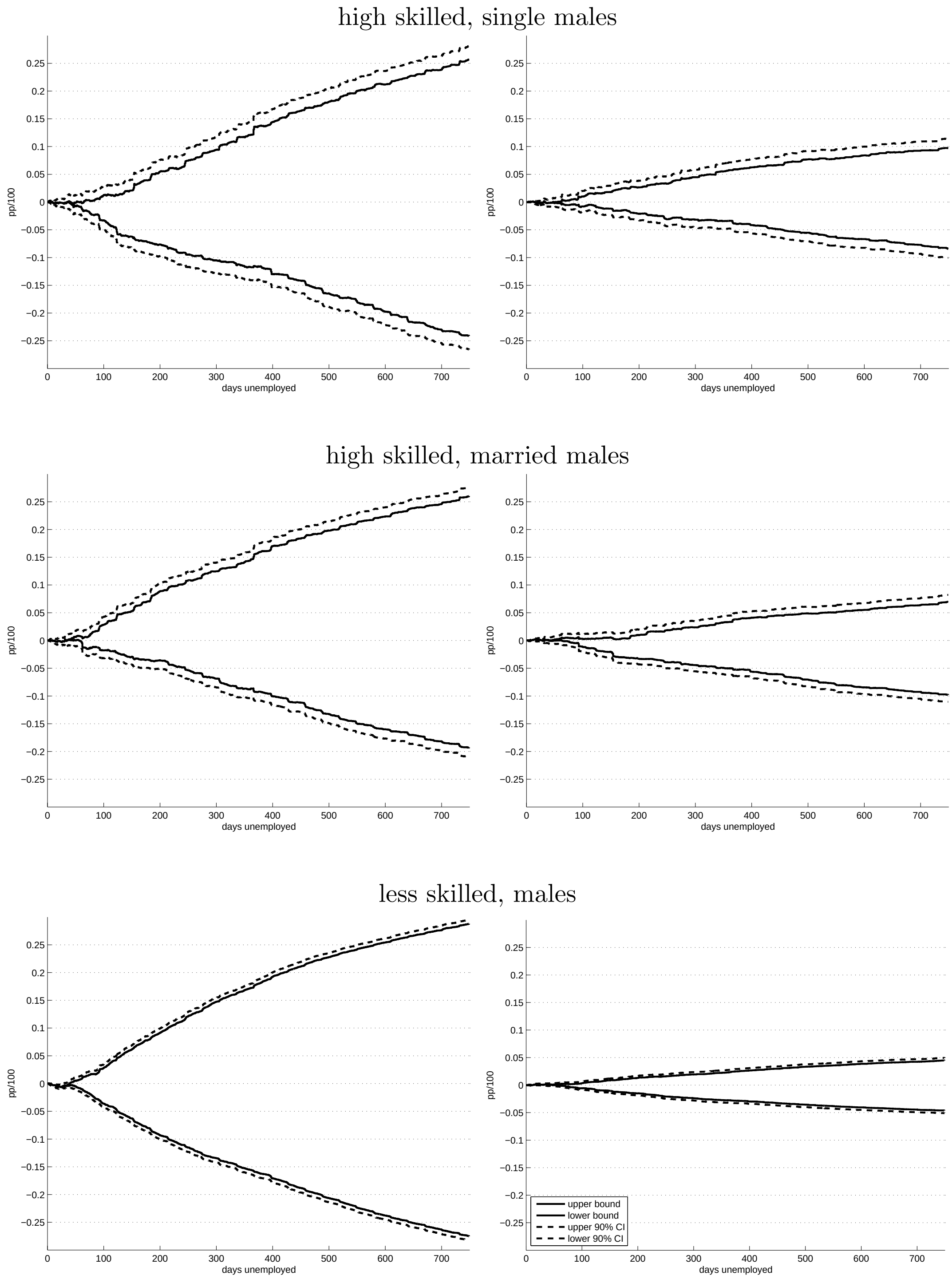


Figure 4: Lower and upper bound of the DiD changes of the cumulative incidence of local (left) and non-local (right) exits to employment among selected groups, additional assumption.

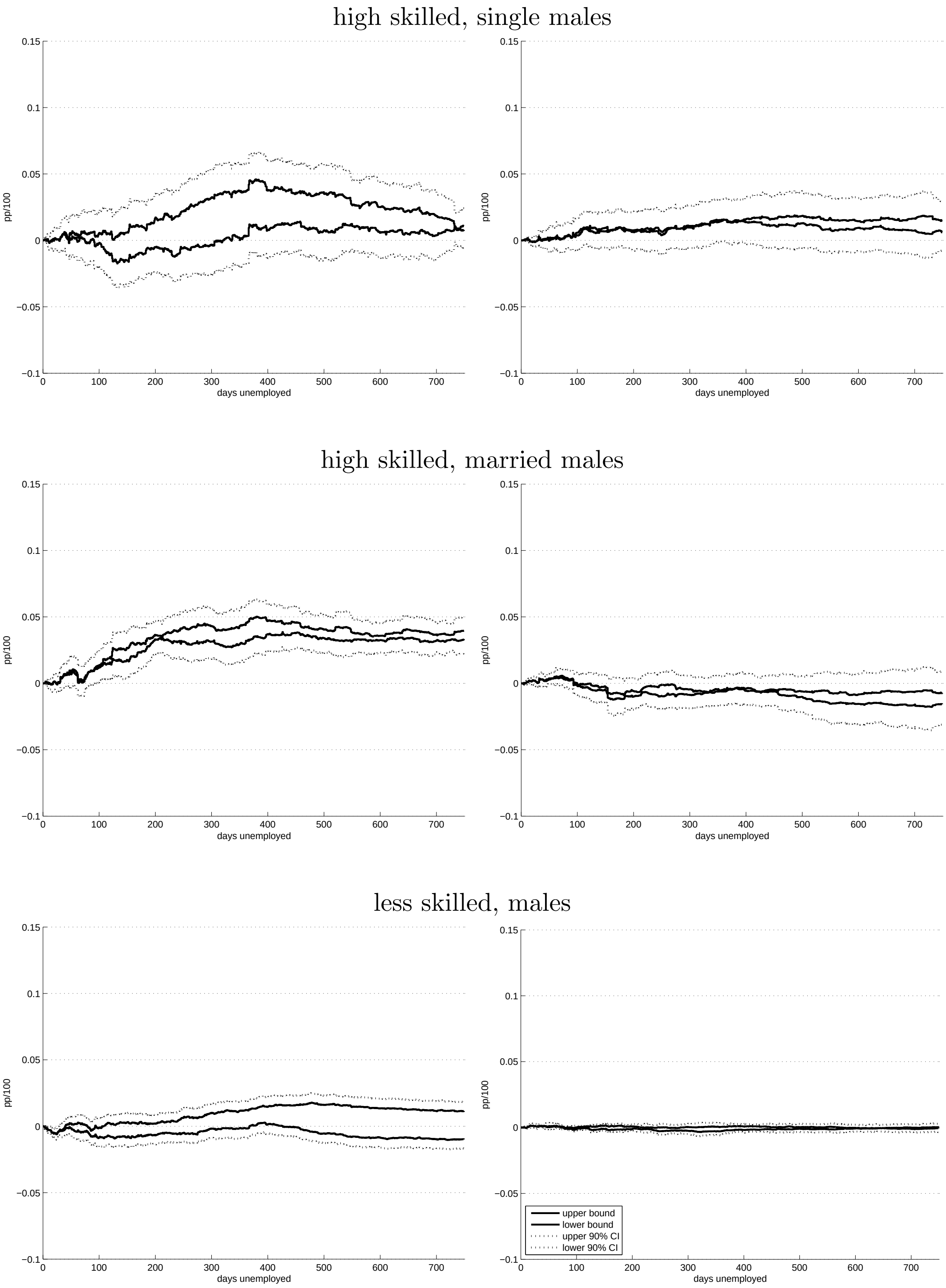
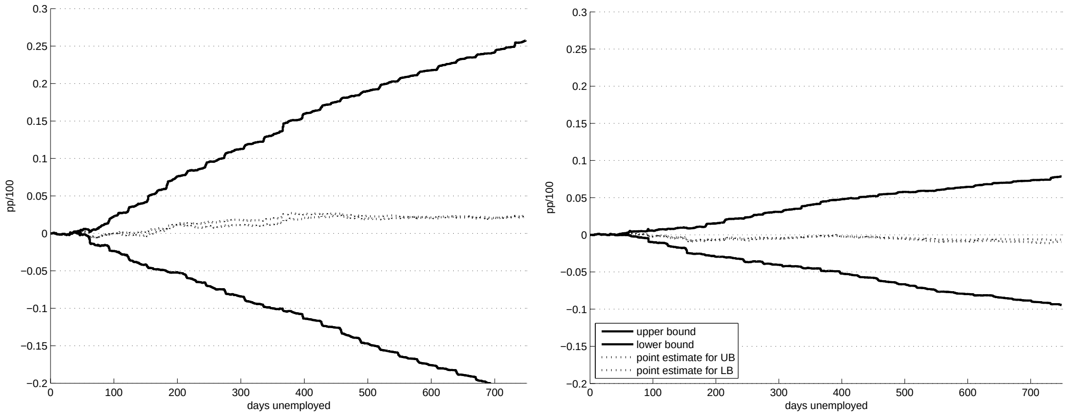


Figure 5: Point estimates for lower and upper bound of treatment effect on the cumulative incidence of local(left) and non-local(right) exits to employment among high-skilled unemployed, married males.



Imprint

FDZ Methodenreport
No. 06/2008

Publisher
The Research Data Centre (FDZ)
of the Federal Employment Service
in the Institute for Employment Research
Regensburger Str. 104
D-90478 Nuremberg

Editorial staff
Stefan Bender, Dagmar Herrlinger

Technical production
Dagmar Herrlinger

Download
http://doku.iab.de/fdz/reporte/2008/MR_06-08.pdf

Internet
http://fdz.iab.de/

Corresponding author
Melanie Arntz
Centre for European Economic Research
(ZEW Mannheim)
L 7, 1 · 68161 Mannheim · Germany
Email: Arntz@zew.de