Schmollers Jahrbuch 128 (2008), 461 – 488
Duncker & Humblot, Berlin

On the Definition of Unemployment and its Implementation in Register Data – The Case of Germany

By Thomas Kruppe, Eva Müller, Laura Wichert, and Ralf A. Wilke*

Abstract

In this paper we suggest different implementation strategies for common international and German legal unemployment definitions for the Sample of the Integrated Employment Biographies (IEBS). The IEBS belongs to a new generation of German merged register data that is more comprehensive than previous data sets. Our descriptive figures show large differences in the number of spells and the unemployment duration across different definitions and implementations. This suggests that empirical results of labour market research are likely to depend on the underlying legal definition of unemployment and its implementation in this data.

Zusammenfassung

Der Beitrag stellt unterschiedliche Implementationen geläufiger internationaler und deutscher Definitionen von Arbeitslosigkeit in den Datensatz ‚Stichprobe der Integrierten Erwerbsbiografien' (IEBS) vor. Die IEBS gehört zu einer neuen Generation von zusammen geführten Registerdaten, die umfassender sind als bisherige Datenquellen. Die vorgelegten deskriptiven Analysen zeigen beträchtliche Unterschiede hinsichtlich der Anzahl der Arbeitslosigkeitsepisoden und der Arbeitslosigkeitsdauer. Hieraus wird geschlossen, dass empirische Ergebnisse der Arbeitsmarktforschung in hohem Maße von der zugrundeliegenden Definition von Arbeitslosigkeit und ihrer Implementation abhängig sind.

JEL Classification: C81, J64

Received: August 6, 2007
Accepted: June 13, 2008

* This work is a result of research cooperation with the IAB Nuremberg, the University of Nottingham and the ZEW Mannheim within the research project: "Contribution to Quality Monitoring of Integrated Employment Biographies: Development of Alternative Unemployment Classification Proposals". We thank two anonymous referees, Gert Wagner, Gesine Stephan, Dirk Oberschachtsiek, Hans Kiesl, and the participants at numerous seminars for helpful comments. Eva Mueller and Ralf Wilke thank the IAB for financial support during their employment at the ZEW. Laura Wichert thanks the ZEW for its hospitality.

1. Introduction

The determinants of unemployment duration are of high interest in social and applied economic research alike. A broad range of empirical and theoretical research in Germany is concerned with this topic. The empirical studies are based on different data such as surveys or administrative registers, which have both advantages and disadvantages. See for example Biewen / Wilke (2005) for a direct comparison of unemployment information in the German Socio-Economic Panel (SOEP) and the IAB Employment Sample. While the smaller sample size and errors in the reporting behaviour of the individuals are the main weaknesses of survey data. Important weaknesses of administrative individual data are a small number of observed variables and inconsistent administrative records. Inconsistencies occur due to the merging of different registers.

In administrative records missing interval information prevents the researcher from fully reconstructing individual employment trajectories using administrative registers. This often makes it impossible to compute the true length of the unemployment periods from this data, since unemployment is inherently a concept – or, moreover, a social category – that is closely related to other labour market concepts such as employment, inactivity, and being outside the labour market.[1] In applied research, the results are therefore likely to depend on the underlying implementation of unemployment duration in the data. Indeed, there is already empirical evidence for this. See Fitzenberger / Wilke (2004) or Lee / Wilke (2008) for an evaluation of a reform of unemployment compensation on unemployment duration in Germany. In addition to the problem of partial identification of the true length of unemployment periods, we focus in this paper on the question of how different legal definitions of unemployment can be implemented in register data.

In Section 2, we first direct our attention to international and national legal definitions of unemployment and their application. We take Germany as an example and show that the definition of unemployment is not stationary but a social category with varying characteristics. As reference we present international standardised unemployment rates, mainly based on the definitions of the International Labour Organisation (ILO), and the primary German regulations for the support of unemployed stated in the Third Volume of the Social Code (SGB III).

Arising from this discussion, Section 3 formulates six alternative concepts of unemployment. These concepts are then implemented in the Sample of Integrated Employment Biographies (IEBS). This data is individually merged

[1] On the genesis of unemployment as a social problem, from the first debates to the foundation of the first compulsory insurance against unemployment, and from individual status to the formation of a social institution, see i.e. Zimmermann, 2006.

register data containing periods of employment, claims of unemployment compensation, job-seeking periods, and information about participation in labour market programmes. Our empirical exercise is nontrivial given the complicated data structure and the massive amount of information in the data. Moreover, missing interval information and data inconsistencies make unique implementations of unemployment duration impossible. For this reason, we suggest several alternative approaches to the applied researcher. These implementations are freely available as Stata Do files (see http: // doku.iab.de / fdz / reporte / 2007 / MR_03-07_Do-files.zip). Some simple descriptive analysis shows strong empirical evidence for considerable differences in the length of unemployment periods depending on the underlying definition of unemployment and their implementations in the data.

In Section 4 we discuss further problems with these data that we do not address in this work. Section 5 discusses the main findings of the paper.

2. International and National Concepts for Measuring Unemployment in Germany

The international standardised unemployment rates, published by the *Statistical Office of the European Communities* (EUROSTAT) and the *Organisation for Economic Co-operation and Development* (OECD), are mainly based on the definitions of the *International Labour Organisation* (ILO) and calculated using cross-sectional random sample survey data sources, namely the *European Labour Force Survey* (LFS). Also, longitudinal data sources can be used to measure unemployment using different concepts, namely the *European Community Household Panel* (ECHP) or the *German Socio-Economic Household Panel* (SOEP). Last but not least, measurement can be based on register data of the German social security system.

The German national unemployment rate is officially announced by the *Federal Employment Service* (Bundesagentur für Arbeit (FES)) on a regular basis. This official measure for unemployment is based on a definition of unemployment that is codified in the *Third Volume of the Social Code* (Sozialgesetzbuch III (SGB III)) from January 1998. The *Second Volume of the Social Code* (Sozialgesetzbuch II (SGB II)), introduced in December 2004, broadens the definition of unemployment to all individuals capable of working, as well as the indigent, where the former is interpreted individually and the latter in a household context. This also refers to concepts of labour reserve, hidden unemployment, hidden labour force, and discouraged workers. When we compare the definitions of international standardised unemployment rates with the German national unemployment rate, we obtain evidence that the social category of unemployment does not have one unique definition.[2]

2.1 International Standardised Unemployment Rates

The *International Labour Office* (ILO) defined "unemployment" at the *Thirteenth Conference of Labour Statistics* as persons without work who are currently available for work and seeking work. The concept thus refers to the definition of employment: persons in paid work could be *"at work, [. . .] performing some work for wage or salary",* or could be *"with job but not at work, having already worked in their present job, were temporarily not at work [. . .] and had a formal attachment to their job".* Self-employed persons are defined as being *"at work, [. . .] performing some work for profit"* or being *„with an enterprise but not at work, [. . .] who are not at work [. . .] for any specific reason".* To comply with this definition, working for at least one hour is sufficient to qualify someone as "employed". The main criterion here is formal job attachment, not the main activity. Consequently, students with a part-time or marginal job are included, as well as persons on leave, laid-off persons, short-time workers, apprentices, and members of the armed forces (ILO 1983)[3]. In absence of a formal work attachment – but fulfilling the criteria mentioned above (available and searching for work) – persons are counted as unemployed irrespective of their primary activity status (i.e., student).

The unemployment figures published annually by the ILO include registered unemployment, also published by the FES, as well as unemployment based on tables from the Federal Statistical Office (Destatis), based on the ILO-Definitions and measured by the German Microcensus, (ILO 1997). Between January 2005 and October 2007 there was an interim solution for monthly unemployment based on a telephone survey. The comparison in the following section refers to this interim solution.[4]

2.2 Registered Unemployment in Germany

In Germany, unemployment aid is regulated by the *Third Volume of the Social Code (SGB III).* It contains, among other things, legislation on the eligibility for unemployment benefits as part of the social security system. It distinguishes between persons in or not in employment, jobseekers, unemployed persons, disabled persons, and re-entrants, while these definitions are only valid in the context of the SGB III:

---

2 On the statistical differences in unemployment in European countries, see Werner (1984). On the U.S. concept and an adjustment of the Canadian and European unemployment rate with respect to this concept, see Sorenntino (2000).

3 The proper statistical handling of persons on extended leave and seasonal workers was under discussion (ILO 1996, 1998).

4 For background information on the new time series, the differences in results involved in the changeover to the new data source, the method applied in estimating the seasonally adjusted results, and the method of calculating unemployment rates see Federal Statistical Office of Germany (2007).

Persons are counted as **not in employment,** if they are not employed or are employed for less than 15 hours per week, and are searching for employment of at least 15 hours with compulsory social insurance contributions. This means that the person has to and is willing to try all possibilities to end the period of having no job, including being at the disposal for placement through the Federal Employment Service. This availability to work is defined more precisely as being capable of work and willing to take a reasonable job under usual or standard conditions.

Jobseekers are defined as persons looking for dependent employment[5] with a duration of more than seven calendar days at home or abroad. They are required to register with the Federal Employment Service for placement, must be capable of and allowed to pursue the occupation in which they are to be placed, and must reside in Germany.

Unemployed persons are *jobseekers,* who are temporarily *not in employment* but searching for employment with compulsory social insurance contributions and have registered at the Federal Employment Service personally. This definition also includes persons who are not receiving benefits. On the other hand, the criteria of receiving unemployment benefits is not sufficient to be counted as unemployed, which is the case for persons who cannot take up work due to periods of illness lasting up to six weeks.

From this definition we can conclude that persons who are not counted as unemployed are:

- more than marginally employed (more than 15 hours a week),
- not capable of or allowed to work (such as individuals younger than 16 years),
- not available without a convincing reason,
- not registered personally at the Federal Employment Service, do not show up again for longer than three months, or repeatedly failed to keep an appointment without cogent reason,
- taking part in active labour market policy measures (exception: short training measures called "Trainigsmaßnahmen" before 2004),
- resident only in a foreign country,
- unable to work because of an illness lasting longer than six weeks,
- performing military or compulsory service or under arrest,
- secondary school students, college students, or school-leavers looking only for professional training,

[5] "Dependent employment" as used here is equivalent to "wage and salaried employment" as used by the International Labour Organisation.

- foreigners without a compulsory work permit (and persons seeking or having been granted asylum) including members of their family (receiving or not receiving benefits), if the labour market is closed to them,
- on leave (i.e., parental leave),
- doing short-time term work (with as little as zero hours of working time),
- aged 58 years or older and thus do not fulfil the definition of unemployment since they are justifiably unwilling to pursue all possibilities to end the period of having no job or are unable to take up work immediately.

Any unemployed person receiving benefits has to personally renew their registration every three months. Not doing so automatically leads to a removal from the register. Exemptions are only given to older (over 55 years) and hard-to-place workers. One side-effect of this regulation is the reduction of long-term unemployment due (mainly) to short interruptions that occur in the case of late renewal.

Any work is considered as "reasonable". This leads to a downward displacement in filling vacancies from more to less qualification-specific occupations and, therefore, in the composition of the qualifications for registered unemployment, where the lower the qualification, the higher the risk of becoming and remaining unemployed.

Any registered unemployed individual is obliged to search for a job on his or her own. The Federal Employment Service is allowed to ask for evidence of active job search such as job applications, reading newspaper want ads, or being interviewed by firms. If this evidence cannot be provided, the individual could lose his or her status as registered unemployed and therefore become ineligible for unemployment benefits. In this context, it is worth noting that training measures can be used explicitly to check for the capability, readiness, and willingness to take up work.

Two groups are excluded from being counted as unemployed: those who were unable to work because of illness, and persons aged 58 or older who would be eligible for early retirement after the unemployment period.

Within the context of the *Second Volume of the Social Code* (Sozialgesetzbuch II (SGB II)), introduced in December 2004, the definition of unemployment is broadened with respect to the SGB III to include anyone who is needy and capable of work, where both are interpreted in a household context. This refers also to concepts of labour reserve, hidden unemployment, hidden labour force, and discouraged workers. In contrast to the former legislation, any person who is in need of and wants to receive social benefits is presumed to be unemployed if he or she is able to work for at least three hours a day. This is also applied to any resident of the household who is liable for alimony or palimony. This change in legislation has had a clear statistical impact on official unemployment numbers, which have been significantly higher since 2005.

The survey results displayed in Table 1 present the primary differences with respect to operationalisation between the survey of the international standardised unemployment definition based on the ILO criteria and the national unemployment definition arising from SGB III.

Table 1

Survey method of ILO- and FES statistic in Germany[6]

	ILO-Statistic (Federal Statistical Office)[7]	Registered Unemployment (FES)
Source	– Computer Assisted Telephone Interview (CATI) – Random sample 30,000 interviews – Six-month panel with monthly interview – Monthly expansion-based estimatefor total population	– Individual registration at FEA checked by case worker – Complete count – Information may be out of date – Monthly reference day
Definition of "active job search"	– Search for employment of at least one hour / week and – Specific search during the last four weeks	– Search on employment of at least 15 hours / week – The case worker concludes that the person is using all possibilities to find a job
Availability	– Taking up work is possible within two weeks	– Person is willing and able to take up work immediately
Without employment	– No employment or employment less than one hour / week	– No employment or employment less than 15 hours / week

The consequence of these different definitions and survey methods can be illustrated in the ongoing reporting of unemployment: While the FES reports a reduction of 869,000 to a total number of 4,108,000 unemployed persons from March 2006 to March 2007[8], the Federal Statistical Office reported a reduction of about 690,000 to a total of 3.03 million unemployed individuals. But the Federal Statistical Office also states:

"As the telephone survey is based on a random sample of about 30,000 persons per month, the random sampling error has to be taken into account when interpreting the results. For the number of unemployed observed for March 2007, that error may amount to a maximum of +/–190,000. This means that with an observed result of 3.03 million, the actual number of unemployed per-

6 Mainly based on Hartmann / Riede (2005).

7 2005 – 2007.

8 Source: Pressrelease 023 29. 03. 2007 http: // www.arbeitsagentur.de / zentraler-Content / Pressemeldungen / 2007 / Presse-07-023 (30. 5. 2007).

sons was very probably within the range of 2.84 and 3.22 million in that month."[9]

2.3 The Concept of the Potential Labour Force

The composition of the unemployed is directly connected to the concept of labour force. In the past 20 years, the use of the concept of the *"Stille Reserve"*[10] to estimate the size of the potential labour force was always a matter of discussion, whereas today, critique is pointed mainly at the statistical handling of persons in active labour market policy measures. The high number of participants in job creation schemes, further training, and retraining measures leads to a distortion of the unemployment rate through hidden unemployment. Aside from this, early retirement and underemployment, partially due to labour market policy measures like short-time work and partially due to those working at an unwanted part-time job while seeking full-time work, are a major subject of discussion in the context of extended unemployment rates.

The Institute for Employment Research (IAB) brings these aspects together in defining the overall *"Stille Reserve"* as persons not in employment who are looking for a job without being registered as unemployed. One part is comprised of jobless individuals in labour market policy programs, especially in full-time measures for further vocational training (including vocational rehabilitation measures and German language courses) and in early retirement. Taking out this part of the group reveals the *"Stille Reserve"* in its traditional meaning, which cannot be quantified exactly (Fuchs 2003; Brinkmann / Klauder / Reyher / Thon 1987).

The following section discusses alternative implementations of these concepts using German register data, based on individual status information: starting from the duration of unemployment as calculated by the FES, six alternative concepts are developed. Comparing these against two additional benchmark concepts, empirical evidence shows that there is a high impact on the results dependent upon on the underlying definition of unemployment.

3. The Duration of Unemployment as an Empirical Concept

The individual unemployment status as discussed in Section 2 is time-dependent. Different underlying definitions of unemployment can lead to different interpretations of individual employment histories in the data. We now focus on the duration of unemployment. A benchmark figure for this aspect of

[9] Press release No. 181 / 02. 05. 2007: ILO labour market statistics March 2007 Source: http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/EN/press/test/Neu_oHS_d_PE07_181_133.psml (04. 02. 2008).

[10] Literal translation: hidden reserve or hidden labour force.

the data is the number of long-term unemployed persons, which is generally seen as an indicator of the persistence of unemployment. Up to 1985, all persons who had not worked more than 13 weeks consecutively after their unemployment registration were counted as long-term unemployed after 12 months of unemployment. Since 1985, any interruption (employment, illness of more than six weeks, further training etc.) leads to a restart of the time counted for the 12 month period. In practice, this change had a clear statistical impact. A recalculation of registered long-term unemployment using the criteria from 1985 for the period before this change (1977 to 1984) shows a statistical reduction of between 12 and 23 percentage points through the new definition (Ministry of Work, Health and Social Services for North Rhine-Westphalia 1998). This must be seen as an additional, cumulative factor alongside a general underestimation of long-term unemployment when using register data (Auer, 1984; Karr, 1997). While any interruption is still counted as an outflow from and – at the end of the interruption – as an anew inflow into unemployment, the introduction of the SGB III changed the definition of long-term unemployment once again. The duration is now a retrospectively calculated sum of periods rather than a fixed status at a single point in time. This definition is used to provide more access to benefits and active labour market policy measures.[11] Not deducted from the duration of unemployment is any period of

- participating in active labour market policy measures,
- illness or maternity protection,
- childcare or long-term care of a family member in need,
- employment or self-employment up to six months,
- periods without the legal right to take up employment,
- short interruptions without supporting documentation.

The following section deals with several concepts of which periods should be included in the calculation and which periods should be counted as interruptions.

3.1 Concepts for the Implementation of Unemployment Duration

We now establish general links between the different definitions of unemployment and the actual calculation of unemployment duration. Based on the

[11] Recently, the BA developed a concept of unemployment including periods of participation in labour market programs, which is also a retrospective sum of periods rather than a status at one point in time and is calculated in two versions as the time spent unemployed, including periods of participating in active labour market policy measures. While one version also includes periods of subsidized work, this is not the case in the second version.

general information available in German register data, unemployment duration can be measured according to one of the following concepts:

1. **Concept 1:** Each uninterrupted unemployment period shown by the administrative record.
2. **Concept 2:** Concept 1, corrected for periods of dependent employment of over 15 hours a week.
3. **Concept 3:** Concept 2, corrected also for periods of dependent employment of less than 16 hours a week.
4. **Concept 4:** Concept 2, with added periods of participating in (any) active labour market policy measures.
5. **Concept 5:** Concept 4, with added periods of illness, identified by a variable on the reason for leaving and entering registered status of unemployment.
6. **Concept 6:** Concept 5, with added period(s) without information on the employment status of individuals, presuming that most people have to search for employment and are willing to work under good conditions, even if they are not registered.[12]

In the following we illustrate these rather theoretical concepts by using a fictitious individual employment history that could occur in real data. In particular, we illustrate how the foregoing concepts are implemented to determine the status of unemployment and to calculate its duration. Figure 1 illustrates the problem of – partially parallel – spells of different labour market states over seven time periods. In real data these spells would in fact be merged from different administrative registers and would have different lengths. The figure shows the resulting unemployment spells derived from the above six concepts. It is evident that the resulting number of unemployment spells and their duration differ across the concepts as we obtain two to three unemployment spells and the cumulative unemployment duration ranges from two (Concept 3) to six (Concept 6) periods. The length of the last unemployment duration varies between one (Concept 1 to Concept 5) and five (Concept 6) periods.

3.2 Data Sources

Empirical work on unemployment duration can be carried out with several competing data sources for Germany. Until the 1990s, such analyses were mainly based on survey data. Since then administrative data from the FES emerged as an attractive alternative. Which data set is most appropriate depends

[12] This concept therefore also defines periods of employment which are possibly not observed in the data as unemployment. This problem is relevant if unobserved periods may correspond to employment and unemployment.

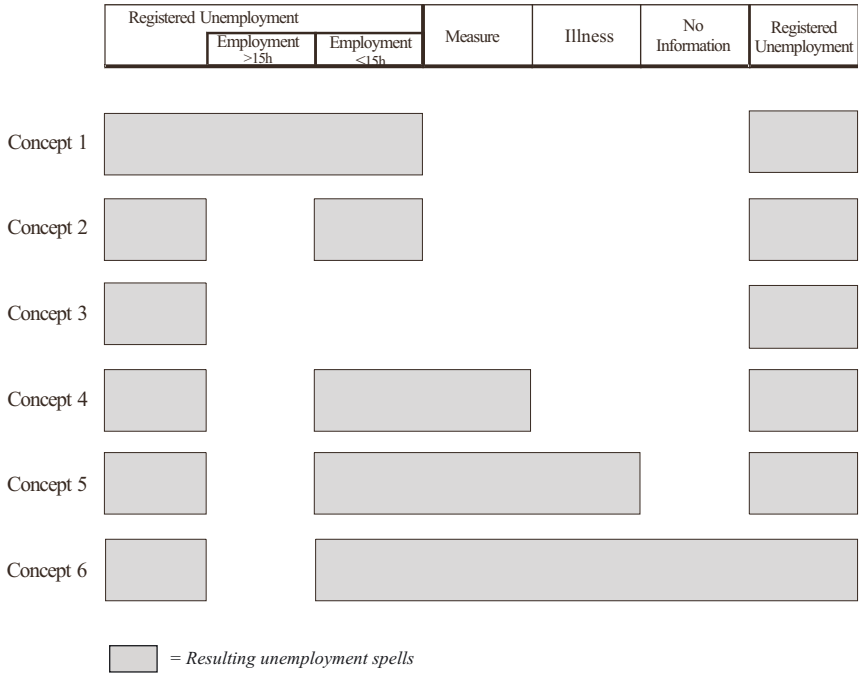

Figure 1: Illustration of six unemployment concepts using a fictitious employment history

on the specific research question at hand, as every data set has its advantages and disadvantages. In the following we briefly compare major survey data sets (German Socio-Economic Panel, [SOEP], and Microcensus, [MC]) with register data from the FES. More information can be found in Biewen / Wilke (2005).

Survey data is generally characterised by a smaller sample size and a richer set of interesting household related variables, especially related to households where respondents are living in. The SOEP is a representative German panel survey that was started in 1984 in West Germany and extended to East Germany after reunification in 1990. See Haisken-DeNew / Frick (2005) and Wagner et al. (2007) for more details about the SOEP. The SOEP contains a monthly retrospective employment calendar that records which months (if any) of the year prior to the interview the individual was registered unemployed. As it is produced for academic purposes, it is easy to access and available in an easily usable format.

The MC is a representative 1 % random sample of all households. Response to most of the questions is compulsory. It contains information about job seek-

ing status, registered unemployment and the receipt of unemployment compensation at the date of the interview. Information about the length of a job seeking period is also provided. People living in a certain flat are followed up to four years only. Further information about the data can be obtained from the Federal Statistical Office and its data research centre.

It is important to note that retrospective information in survey data may be subject to measurement error if individuals do not (or do not want to) remember when or even whether they were unemployed. For example, Kraus / Steiner (1998) found clear heaping effects in the SOEP unemployment data at the end of each year. Furthermore, by comparing retrospective with contemporaneous unemployment information, Jürges (2007) concludes that 7 % of all retrospectively reported unemployment spells in the SOEP may be subject to error and 13 % of all unemployment spells are not reported one year later. The amount of error is related to other observed characteristics. We are not aware of similar analysis for the MC.

In contrast to the survey data, information in the FES data seems to be less subject to measurement error (Bernhard et al., 2006). It contains information that is drawn from administrative registers rather than that recorded by interviews. Misreporting and missing recalls are therefore not relevant. Data inconsistencies due to the administrative recoding process are, however, possible (Waller, 2008). Important information about the household background is often not available, although the FES contains detailed information about policy interventions such as ALMP on individual level. For this reason, it is currently an important database for evaluating the effect of the German labour market reforms conducted during the period 2003 – 2006. Administrative data has another difficulty, namely that it does not cover the entire employment trajectories of individuals. This leaves unobserved periods to the researcher and it may cause the true length of an unemployment period to be unidentifiable in this data (see also Fitzenberger / Wilke (2004) or Lee / Wilke (2008)). Moreover, it can be difficult to identify the labour market state if different merged administrative registers contain parallel records with potentially conflicting information. These limitations in administrative data motivate the following empirical analysis. Due to data protection laws it can take some time until access to the data is granted under strict data security regulations.

3.3 Empirical Evidence

In our empirical exercise we implement the six concepts from section 3.1 using merged administrative data from Germany.[13] In addition, we compare our results to two benchmark proxies for unemployment which had been used

[13] A direct comparison between SOEP and MC is not possible because both data sets cannot be linked with the register data.

in the former literature on unemployment using this data. In our empirical analysis we use the Integrated Employment Biographies (IEBS), which are made available for scientific use by the research data center of the FES (FDZ).[14] They include information on dependent employment (Source: BeH), registered unemployment, job search (BewA), unemployment benefits (LeH) and participation in active labour market policies (MTG) as a representative 2.2 percent sample covering about 80 percent of the labour force (Jacobebbinghaus / Seth 2007).[15] Whereas information on employment is included from 1990 until 2003 and unemployment benefits from 1990 until 2004, information on participation in active labour market policy measures is only reliable in the period 2000 – 2004. We use data between 2000 and 2003 exclusively in order to focus on a period for which all registers are available. The key properties of our sample are described in Table 2.

Table 2

Number of episodes and parallel episodes in the IEBS for the period 2000 – 2003

	BeH	LeH	MTG	BewA
Number of episodes	4,937,224	1,059,985	313,786	1,516,699
Parallel episodes				
BeH	275,630	199,901	107,918	443,013
LeH	199,901	1,877	136,531	776,205
MTG	107,918	136,531	19,068	170,728
BewA	443,013	776,205	170,728	4,992

While implementing the six concepts using this data set, we faced certain problems such as data inconsistencies and missing interval information. Data inconsistencies due to unfeasible overlapping of register information have already been extensively analysed by Jaenichen et al. (2005) and Bernhard et al. (2006). For this reason we do not directly address them in our implementations but strongly recommend performing the data corrections suggested by Bernhard et al. (2006) before applying our implementations to the data. The main focus of this work is to address the problem of missing interval information. We do, however, take into account contradictory information about the length of training measures which are funded by the FES.

Since there are several ways to deal with the present data problems, we discuss different approaches to implementing the above six concepts in the Ger-

[14] For Information on the IEBS and how to access it see: http: // fdz.iab.de.

[15] For a detailed description see Hummel et al. (2005) or Jacobebbinghaus / Seth (2007).

man register data. In a first step the researcher has to decide which administrative information is used as the core information for an unemployment period, since BewA and LeH provide sets of unemployment information which are related to the receipt of unemployment compensation or to job seeking activities. For this reason, the source from which the unemployment information is taken has to be defined: BewA, LeH or from both (BewA+LeH). The choice should be made according to the specific research question at hand. When we are interested in some economic effects of unemployment benefits, the LeH may be sufficient. Using only BewA may be useful for analysing registered unemployment. To obtain the most comparable results on unemployment, using information from both BewA and LeH may be the right decision. With this paper we provide program code for both approaches, but we decided to focus the following analysis to unemployment information taken from the BewA only. This is only done for reasons of simplification and does not mean that using the BewA information is more suitable for a specific empirical problem at hand.

In order to implement the above six concepts, the researcher has to make further decisions:

1. **How to deal with parallel employment information?**

   In case of parallel full-time or part-time employment and unemployment information, the researcher has to decide if the information is assessed as employment or unemployment. While there are some regulations that allow registration as unemployed parallel to dependent employment, this, for example, could be interpreted as underemployment (see also Chapter 2.3).

2. **How to deal with training measures, illness, or information on those out of the labour force?**

   In a second step the researcher has to make a decision about the unemployment status of periods of training, illness, or absence from the labour force. Here we face similar questions as when analysing employment. In general (short) illness is not shown by register data on employment spells and therefore is not counted as an interruption. While some "training measures" in Germany are used to check for people's readiness to work, others are used to train unemployed persons to write letters of application or to give them practical advice in direct connection to a subsequent job. Further vocational training can range from short modules of several weeks to long-term measures lasting two years or more, providing a recognized vocational qualification. If periods of active labour market measures count as unemployment, the researcher has to make another decision about their length. The duration of active labour market periods is also not uniquely defined in the data since there are different sources of possibly contradictory information in the MTG. For this reason Waller (2008) suggests two concepts to deal with the data inconsistencies:

a) The "naïve concept" relies on the information concerning the participation in labour market policy measures, i.e., we use the end date of the measure given in the data as the real end measure, even if there are parallel employment spells.

b) The "standard concept" relies on the information from employment spells, i.e., if an employment spell starts before the end date of the labour market measure given in the data, we consider the observation as an employment spell.

In the following we present the results for implementation b only. The results for implementation a are available in an Internet appendix (http: // doku.iab. de / fdz / reporte / 2007 / MR_03-07_appendix.doc).

In the first step – how to deal with parallel information – the researcher has to choose from three possibilities:

1. **Implementation A** (based upon Concept 1):

   Each uninterrupted unemployment period shown by the administrative record, i.e., each uninterrupted BewA spell, including those with parallel BeH spells, are considered as unemployment spells.

2. **Implementation B** (based upon Concept 2):

   It includes each uninterrupted unemployment period shown by the administrative record, corrected for periods of dependent full-time employment, i.e., all uninterrupted BewA spells except for those with parallel BeH-spells coming from a full-time job. If a parallel BeH spell comes from a full-time job, the spell is considered as an employment spell; if it comes from a part-time job, it is considered as an unemployment spell (the variable used to identify full-time or part-time employment is *"Erwerbsstatus"*).

3. **Implementation C** (based upon Concept 3):

   Each uninterrupted unemployment period shown by the administrative record, corrected for all periods of employment. Thus, only BewA spells without any kind of parallel BeH spell are considered to be unemployment spells.[16]

The following figure presents the differences in the median unemployment lengths depending on which implementation is chosen:

[16] This implementation may be considered as somewhat crude because in many cases such overlapping periods correspond to employment subsidies with a focus on the second labour market (e.g., ABM, SAM). It would also be plausible to define these cases as unemployment, but we decided not to distinguish at all between different types of labour market measures in order to keep the number of implementations manageable. As with any other analysis, this decision is in the responsibility of the researcher.

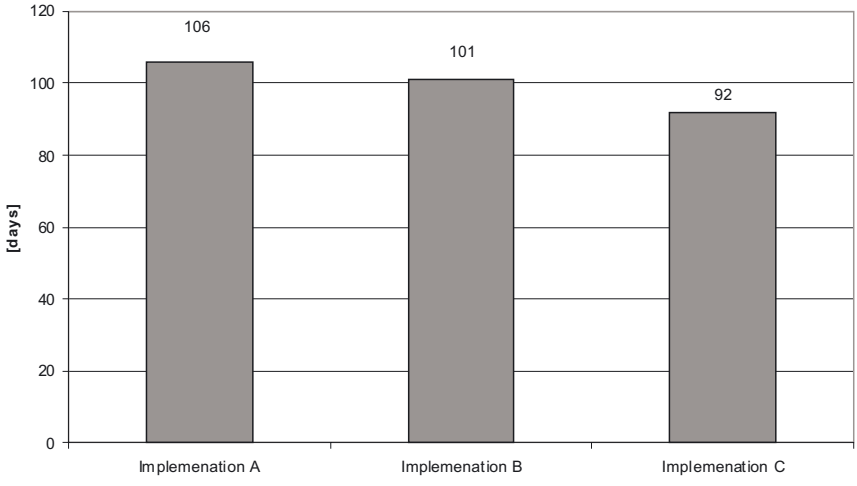

Figure 2: Median unemployment length for males for different treatments of parallel employment spells (Source: BewA)

Additional information about the participation in active labour market policy measures can be found in the *"Maßnahme-Teilnehmer-Grunddatei"* (MTG). This information allows for the further specification of requirements for unemployment (please note that the MTG only contains information about primary policy measures funded by the Federal Employment Agency). In the second step there are again three alternatives:

1. **Implementation I** (based upon Concept 4):

   Spells with information about the participation in active labour market policy measures are also considered to be unemployment spells.

2. **Implementation II** (based upon Concept 5):

   Spells with information about illness are also considered to be unemployment spells.[17] These periods can be directly identified from the BewA or they are identified using the variable on the reason for leaving and entering the registered status of unemployment.

3. **Implementation III** (based upon Concept 6):

   Periods without any information about the employment status of an individual are considered to be unemployment periods. We have to keep in mind, though, that people who cannot even be found in the data set could be unemployed according to this concept.

The different implementation strategies are summarized in Table 3.

[17] Sick leave during an employment period does not interrupt the employment spell.

Table 3

Summary of Implementation strategies for unemployment in the IEBS

Decision	Alternative	Brief explanation
1	A	Unemployment = each uninterrupted BewA or LeH period (also including any parallel employment information)
	B	Unemployment = each uninterrupted BewA or LeH period except for parallel periods of dependent full-time employment.
	C	Unemployment = each uninterrupted BewA or LeH period without any kind of parallel employment information.
2	I	Unemployment = each uninterrupted BewA, LeH or MTG period except for parallel periods of dependent full-time employment.
	II	Unemployment = each uninterrupted BewA, LeH, MTG or "illness" period except for parallel periods of dependent full-time employment.
	III	Unemployment = each uninterrupted BewA, LeH, MTG or "illness" period and all periods without any information about the employment status (except for parallel periods of dependent full time employment).

Depending on the question, the researcher has to choose which concept is most appropriate for the purpose of the analyses. This is especially the case for the second decision when the researcher decides how to handle illness or missing information (alternative I, II or III). Therefore, we can not give any general advice about which concept out of the 63 possibilities to use. However, when comparing our implementations to the different existing unemployment concepts discussed in section 2.2, we can deduce some basic guidelines concerning the first decision step (alternative A, B or C): we expect that implementation A is not likely to be appropriate in a real-world application since the requirements are very weak and therefore the delimitation of unemployment is rather blurred. Alternative C offers a comparatively strict definition of unemployment which can be used when the researcher wants to use the unemployment definition given in the ILO-Statistic. Alternative B is not only a compromise between the two extreme positions of implementation A and C but also sensible if the criteria for registered unemployment according to the FES (SGB III) should be fulfilled in the analysis.

In order to illustrate the importance of each of the decisions, we present exemplarily the different outcomes of the median length of unemployment for males.[18] For this purpose we hold constant all the remaining decisions, except the one in question:

[18] Moreover, we use seven days as the maximum tolerated interruption length, because there is some evidence that short interruptions could be caused by minor failures

- Information is taken from BewA;
- We consider unemployment spells with any kind of parallel employment information as unemployment (Implementation A);
- Training measures are also treated as unemployment periods (Implementation I);
- The duration of the length of a labour market policy measure is determined using the "standard concept" (Implementation b).

This makes the implementation AI (BewA) the default implementation in what follows when we report descriptive figures. To this end we also hold the implementation of the length of training measures constant and choose the "standard" implementation b as default. By doing this we can easily illustrate the changes in unemployment duration when we deviate from implementation AIb.

The following figure shows the median length of unemployment for the possibilities AIb, AIIb and AIIIb:

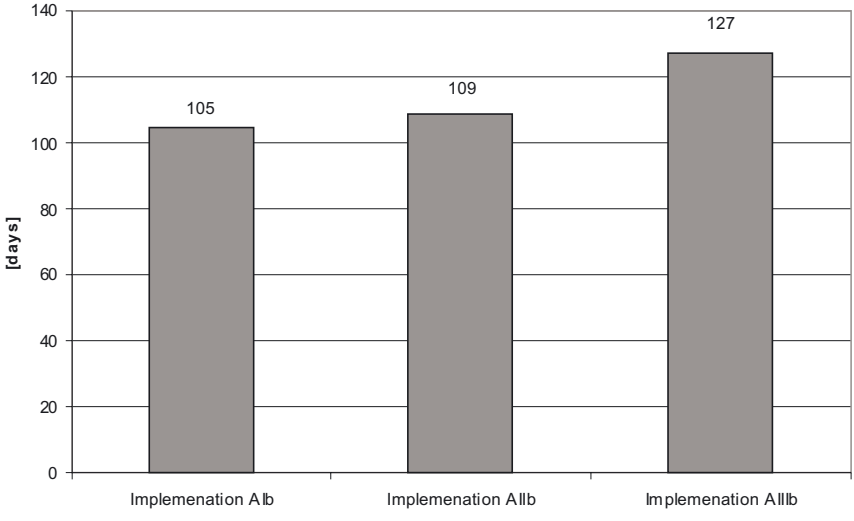

Figure 3: Median unemployment length for males for different treatments of training and illness periods (Source: BewA)

The distinction (A / B / C) refers to three different implementations of unemployment. Based upon these three definitions and using information from the

in the data generating process. Also, official calculation of unemployment duration tolerates interruptions caused by short employment. Since results are likely to depend on the chosen value, it can be set as a parameter in our STATA code.

MTG, we can combine different permutations of the three refinements (A / B / C, I / II / III, and a / b) in order to get, e.g., A-I-a or C-II-b. This leaves altogether 21 different definitions of unemployment based upon the Concepts 1 to 6 from the previous chapter. These 21 definitions can be computed using LEH, BewA or both as the core information of the unemployment period. Our implementations therefore yield 63 definitions of unemployment in the IEBS.

Although the differences in the median length in the examples presented above are already remarkable, they are still relatively small compared to those if more than one of the decisions is changed. Figure 4 presents the maximum difference in median duration for males over all implementations discussed in this section. The median ranges from 92 to 172 days. If we also alternate over a variety of stratifications such as sex, age, and East / West, we identify the group of those aged > 49 as the one with the largest difference across implementations. For this group the smallest median is 127 days for implementations CIIIb and BIIIb using the BewA, while the greatest median (335 days) is obtained for AIIIa using LeH and BewA.

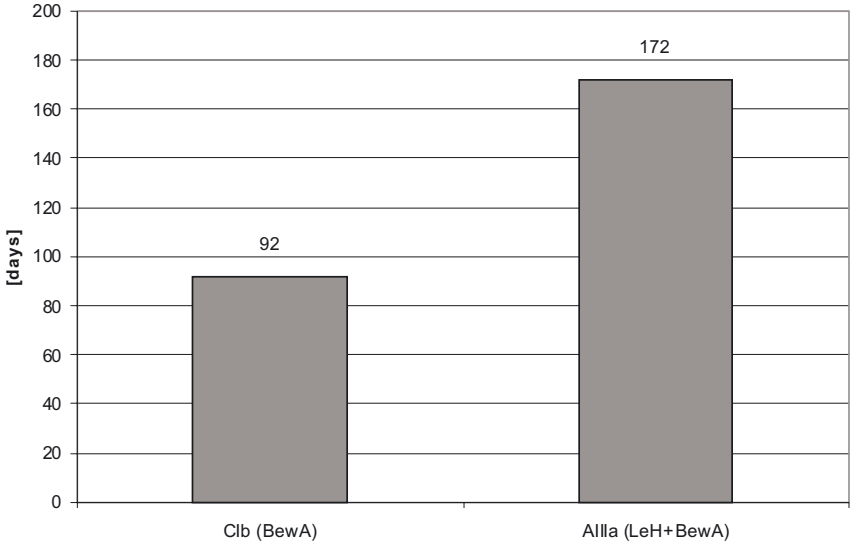

Figure 4: Minimum and maximum median length for males of all the discussed implementations

Before we discuss the descriptive statistics for the different implementations, two additional proxies for unemployment in the IAB data are introduced. These proxies are based on the information content in the IAB employment sample (IABS) and they have already been used in the econometric literature (see, e.g., Fitzenberger / Wilke, 2004 or Lee / Wilke, 2008). For this

reason we refer to these concepts as benchmarks in what follows. Moreover, necessary assumptions for the programming and differences in the definition of censoring for each of the concepts mentioned above are presented.

1. **Nonemployment (Fitzenberger / Wilke) NE:**

   All periods of nonemployment that follow an employment period and contain at least one period of receiving income transfers from the German Federal Employment Service. A period of nonemployment is right-censored if the last nonemployment spell is not followed by an employment spell. (If an income transfer spell (LeH) is parallel to an employment spell (BeH), the period is treated as an employment spell.)

2. **Unemployment with permanent income transfer (Lee / Wilke) UPIT:**

   All periods of nonemployment after an employment period with a continuous flow of unemployment compensation from the German Federal Employment Service. The maximum interruption in compensation transfers is one month (30 days). The limit of one month is chosen because unemployed who do not receive unemployment compensation for more than one month lose their social insurance protection. This implies that there is a strong financial incentive for not having long gaps and if they are observed, they may be related to periods out of the labour market. An observation is marked as right-censored if the interruption in transfer payments is longer than 30 days or if there is no other observation after a compensation payment spell. (If an income transfer spell (LeH) is parallel to an employment spell (BeH), the period is treated as an employment spell.)

The following figure presents the median length of unemployment for males according to the example discussed above (males, AIb (BewA)) and the two benchmark proxies NE and UPIT.

Both proxies for unemployment are conditional on a preceding employment period. This is an important difference compared to the concepts described earlier in this section. Conditioning on a foregoing employment period has two main implications: first, it restricts the set of unemployment periods to those who were just transitioning into unemployment, while it excludes individuals with a loose labour market connection and without any employment spell. Second, by conditioning on an observed employment period in the data, the sample of unemployment spells becomes more representative with respect to transitions from socially insured employment into unemployment. Thus, problems of sample selection do not apply.

A long list of descriptive statistics for all different definitions based on the IEBS from 1 January 2000 to 31 December 2003 can be found in Tables 5 and 6 in the Appendix. Additional stratifications with respect to age groups and East / West Germany are available as an Internet appendix http: // doku.iab.de / fdz / reporte / 2007 / MR_03-07_appendix.doc. This appendix also presents de-

tailed results for implementations a and b and which underlying register (BewA, LeH or both) was chosen.

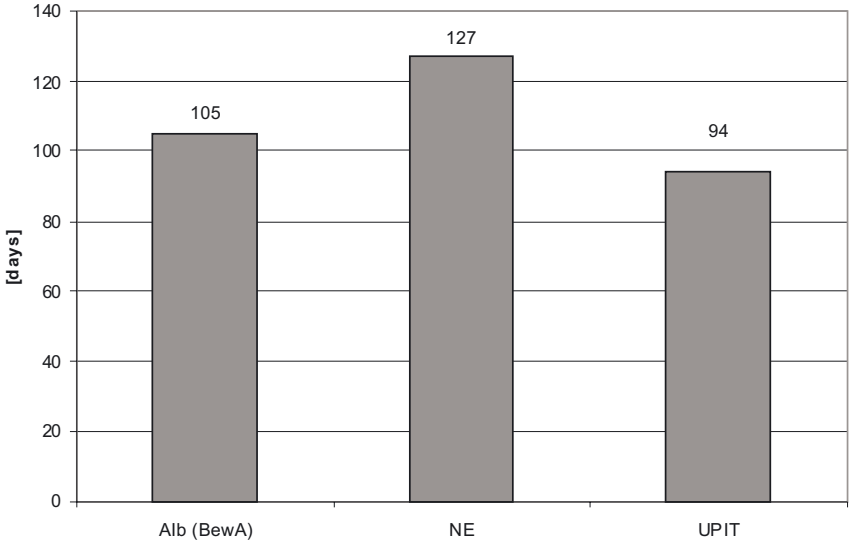

Figure 5: Median unemployment length for males compared to the two proxies NE and UPIT

In the following we briefly discuss how the different definitions of unemployment are related in terms of unemployment duration, since the different implementations imply a certain ordering. Table 4 shows the resulting ordering for a given unemployment period.

Table 4

Ordering of the different implementations with respect to the length of an unemployment spell

Length of unemployment				
A	≥	B	≥	C
III	≥	II	≥	I
a	≥	B		
NE	≥	UPIT		

1. Implementation A results in a longer duration compared to implementations B and C. The latter yields the shortest duration. This is due to the fact that the implementation A includes all unemployment periods shown in the ad-

ministrative data, regardless of any kind of parallel employment information. Following implementation B or implementation C, a parallel full-time job or any kind of parallel employment, respectively, leads to an interruption and thus to a shorter unemployment spell.

2. Implementation I is shorter than II because it excludes periods of illness. Implementation III yields the longest unemployment duration because it also includes out of the labour force periods. In many cases implementation III merges two or more durations according to implementations I and II.
3. Implementation a treats the end of a policy measure as the end of unemployment while according to implementation b, unemployment ends when a job starts. Therefore, the spell length in the latter case is shorter.
4. Since the NE proxy relies on weaker requirements than UPIT in terms of tolerated interruption length, the observed unemployment spell is longer in the first than in the latter case.

The marginal distributions of unemployment duration for the implementations may not possess this ordering because of differing numbers of observations. Also, it is less clear in ranking implementations according to their number of observations. It is, however, evident from tables 5 and 6 in the appendix that there are in any case several hundred thousand observations, which is more than enough for an empirical application. Since NE and UPIT require foregoing employment periods, they have the lowest number of observations.

Finally we remark on the notion of censoring in the computed unemployment periods. Contrary to the two benchmark proxies NE and UPIT, censoring takes place only at the start and the end of the observation period (1 January 2000 and 31 December 2003). Note that this notion of censoring differs from those of NE and UPIT: There, left censoring is impossible because it requires the observed transition from employment to unemployment. Right-censoring is not only due to the end of the data; it is also at the end of income transfers if there is no observed transition to employment. This reflects the fact that unobserved periods impose some uncertainty in the length of the true unemployment period which is not observed. Since in a variety of applications estimators capable at handling left censoring in an appropriate manner do not yet exist, left censoring may cause difficulties in applied empirical work.

4. Further Topics

When an applied researcher is analyzing the registers of the German social insurance system missing or conflicting spell information are not the only problems. In addition, the researcher faces the problem that the registers are typically not representative of the entire population. In order to appear in an administrative record, one needs to be in contact with one of the relevant data

generating administrations. In Germany, this is unlikely to be the case for the self-employed, civil servants and other individuals who have never contributed to the national insurance system. This implies that during a period of unemployment these individuals are likely to be ineligible for income transfers from the FES. While these unemployed persons consider themselves unemployed, they may not register as job seekers if they do not consider this to be beneficial for them. For this reason, the BewA is not representative because appearance in this register depends heavily on a subjective choice. Such sample selection issues should also to be taken into account when working with this data.

A non-representative sample of unemployment spells cannot be considered to be a random sample and the results of statistical analysis may be biased. This is especially the case if selection is not independent of the variable to be explained. This may prevent the researcher from consistently estimating the causal effect of policy interventions using this data. Since the degree and the kind of selectivity are unknown, we cannot suggest a general solution to this problem.

We think, however, that for some specific empirical questions one may create samples which are (almost) representative. This is the case, for example, if one conditions the sample of unemployment spells on those who have foregoing employment durations of a minimum length before unemployment[19] and to relate unemployment periods to the receipt of income transfers. The two benchmark definitions NE and UPIT proceed in this way. Research questions related to the receipt of unemployment benefits can be analysed quite well with this data. Their fields of application are, however, limited to quite specific empirical problems and in many problems it may not make sense to use them in applied analysis.

5. Discussion

In this paper we discuss several theoretical and legal concepts of unemployment. As a consequence, the labour market state and the duration of an unemployment period of an observation depend on the nature of the concepts. This is likely to cause difficulties for applied research on unemployment, since results may depend heavily on the definition of unemployment.

[19] To be eligible for the receipt of unemployment compensation from the FES, one needs to be employed at least once for a minimum duration during a claim period, which can be more than five years depending on the age of the unemployed and the specific legal conditions in force. This condition may not be satisfied for young unemployed persons just starting their career. For this reason, the sample of NE and UPIT is also not representative with respect to all transitions from employment to unemployment.

Given the theoretical notions of unemployment, we focus on the question of how the labour market state, unemployment, and the duration of unemployment can be defined in real-world data. In our empirical work we use the Sample of Integrated Employment Biographies (IEBS), which is the German merged administrative individual data. In addition to two well-known benchmarks, we develop more than 60 different implementations of unemployment in this data. A short descriptive analysis shows considerable differences in the number of unemployment spells and in the length of unemployment periods, which provides evidence for the relevance of our work. Our implementations can be provided to users of this data by the research data centre of the FES (fdz.iab.de). They are provided as ready-for-use Stata do files.

Given the theoretical concepts and the data at hand, we face several difficulties: in some cases the legal definitions of unemployment cannot be implemented directly because not all necessary information is available in the data, e.g., the true unemployment duration according to a theoretical concept cannot be identified exactly given the data. For this reason we use several variants of the implementation that take this difficulty into account. Our implementations therefore provide alternative unemployment classifications in the data.

Unobserved periods in the employment trajectories are an important weakness of this data, which may cause a variety of problems in applied analysis. This requires by its very own careful judgement of whether an empirical question can really be solved by analyzing the register data at hand. Indeed, our descriptive figures reveal remarkable differences in the number of observations and in the length of unemployment across the implementations. This provides additional motivation for the relevance of this work. Future work may address the question of whether missing information or differences in the length of unemployment periods across our implementations are random or whether they are correlated with observable variables.

Our work delivers more than 60 different implementations of unemployment in the data. This is already quite comprehensive and we think that we have addressed several important topics. There is, however, further scope for interesting extensions. For example, one could define finer classifications which distinguish between the types of labour market measures. This is important because labour market programs are heterogeneous in their purposes and intents.

References

Auer, P. (1984): Maßnahmen zur Reintegration von Langzeitarbeitslosen, Discussion Paper IIM / LMP 84 – 20, Social Science Research Center, Berlin.

BA – Bundesanstalt für Arbeit (1999): Arbeitsmarkt 1998. Amtliche Nachrichten der Bundesanstalt für Arbeit, Jg. 47, Sondernummer vom 12. 06. 1999, Nürnberg.

Bernhard, S. / *Dressel,* C. / *Fitzenberger,* B. / *Schnitzlein,* D. / *Stephan,* G. (2006): Überschneidungen in der IEBS: Deskriptive Auswertung und Interpretation, FDZ Methodenreport, 04 / 2006, Nürnberg.

Biewen, M. / *Wilke,* R. A. (2005): Unemployment Duration and the Length of Entitlement Periods for Unemployment Benefits: Do the IAB Employment Subsample and the German Socio-Economic Panel Yield the Same Results? Allgemeines Statistisches Archiv 89 (2), 209 – 236.

Brinkmann, Ch. / *Klauder,* W. / *Reyher,* L. / *Thon,* M. (1987): Methodische und inhaltliche Aspekte der Stillen Reserve, in: Mitteilungen aus der Arbeitsmarkt und Berufsforschung (MittAB), Jg. 20. Nürnberg.

Federal Statistical Office of Germany (2007): Notes on the ILO labour market statistics from reference month September 2007, http: // www.destatis.de / jetspeed / portal / cms / Sites / destatis / Internet / EN / Content / Statistics / Arbeitsmarkt / ILOArbeitsmarktstatistik / Hinweise_ILO,templateId=renderPrint.psml (04. 02. 2008).

Fitzenberger, B. / *Wilke,* R. A. (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. ZEW, Mannheim.

Haisken-DeNew, J. P. / *Frick,* J. (eds.) (2005): Desk Top Companion to the German Socio-Economic Panel Study (GSOEP) – Version 8.0, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin.

Hartmann, M. / *Riede,* T. (2005): Erwerbslosigkeit nach dem Labour-Force-Konzept – Arbeitslosigkeit nach dem Sozialgesetzbuch: Gemeinsamkeiten und Unterschiede, in: Wirtschaft und Statistik, 4 / 2005, Wiesbaden.

Hummel, E. / *Jacobebbinghaus,* P. / *Kohlmann,* A. / *Oertel,* M. / *Wübbeke,* C. / *Ziegerer,* M. (2005): Stichprobe der Integrierten Erwerbsbiografien, IEBS 1.0., Handbuch-Version 1.0.0 FDZ Datenreport 6 / 2005. Nürnberg.

ILO – International Labour Office, Bureau of Statistics (1983): Resolution concerning economically active population, employment, unemployment and underemployment, in: Bulletin of Labour Statistics, 1983 – 3, Geneva.

ILO – International Labour Office, Bureau of Statistics (1996): Resolution of the Prag Meeting, in: Bulletin of Labour Statistics, 1996 – 1, Geneva.

ILO – International Labour Office (1997): Yearbook of labour statistics, Geneva.

ILO – International Labour Office, Bureau of Statistics (1998): General Report on the Sixteenth International Conference of Labour Statistics, http: // www.ilo.org / english / 120stat / techmeet / 16thicls / report4.htm (25. 08. 98), Geneva.

Jacobebbinghaus, P. / *Seth,* S. (2007): The German Integrated Employment Biographies Sample IEBS, in: Schmollers Jahrbuch 127, 335 – 342.

Jaenichen, U. / *Kruppe,* T. / *Stephan,* G. / *Ullrich,* B. / *Wießner,* F. (2005): You can split it if you really want – Korrekturvorschläge für ausgewählte Inkonsistenzen in IEB und MTG, FDZ Datenreport Nr. 04 / 2005, Nürnberg.

Jürges, H. (2007): Unemployment, life satisfaction and retrospective error, Journal of the Royal Statistical Society: Series A, 170, 43 – 61.

Karr, W. (1997): Die Erfassung der Langzeitarbeitslosigkeit – Ein kaum beachtetes Meßproblem, in: IAB Kurzbericht, Nr. 5, Nürnberg.

Kraus, F. / *Steiner,* V. (1998): Modelling heaping effects in unemployment duration models -with an application to retrospective event data in the German Socio-Economic Panel, Jahrbücher für Nationalökonomie und Statistik 217, 550 – 573.

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

Ministry of Work, Health and Social Services for North Rhine-Westphalia (Ed.) (1998): Arbeitslose, Langzeitarbeitslose und ihre Familien, Landessozialbericht, Vol. 8, Düsseldorf.

OECD – Organisation For Economic Co-Operation And Development (1999): Quarterly Labour Force Statistics, No. 1 / 1999, Paris.

Sorenntino, C. (2000): International unemployment rates: how comparable are they? Monthly Labor Review 123 (6), Washington.

Wagner, G. G. / *Frick,* J. R. / *Schupp,* J. (2007): The German Socio-Economic Panel Study (SOEP) – Scope, Evolution and Enhancements, Schmollers Jahrbuch 127 (1), 139 – 171.

Waller, M. (2008): On the Importance of Correcting Reported End Dates of Labor Market Programs, Schmollers Jahrbuch 128 (2), 213 – 236.

Werner, H. (1984): Unterschiede in der Erfassung der Arbeitslosigkeit, Mitteilungen aus der Arbeitsmarkt und Berufsforschung Nr. 3, Nürnberg.

Zimmermann, B. (2006): Arbeitslosigkeit in Deutschland – Zur Entstehung einer sozialen Kategorie. Campus, Frankfurt am Main / New York.

Appendix

Table 5

Descriptive statistics for marginal unemployment duration distribution using the BewA stratified by gender

Implementation	num_obs	mean	median	0.25-Quantile	0.75-Quantile
A	587278	215.8124	115	51	268
males	341310	201.2612	106	47	244
Females	245968	236.0039	125	56	298
B	589072	211.0129	109	46	261
males	342735	196.1846	101	43	238
Females	246337	231.644	122	52	291
C	593740	185.156	94	37	223
males	348339	177.1649	92	36	212
Females	245401	196.4991	100	40	244

Implementation	num_obs	mean	median	0.25-Quantile	0.75-Quantile
AIb	635683	216.0664	113	46	273
males	366743	203.122	105	44	251
Females	268902	233.7174	122	51	301
Missings	38	239.3684	184	92	365
BIb	636680	212.0321	108	43	268
males	367559	198.8578	100	40	245
Females	269083	230.0239	121	47	295
Missings	38	239.3684	184	92	365
CIb	634145	190.9311	94	34	237
males	370163	182.976	92	33	222
Females	263944	202.0805	99	36	254
Missings	38	239.3684	184	92	365
AIIb	625774	228.1369	117	48	284
males	357501	216.1743	109	45	269
Females	268235	244.0791	123	51	311
Missings	38	239.3684	184	92	365
BIIb	628301	223.8717	112	44	277
males	359302	211.5826	105	42	261
Females	268961	240.2863	122	48	306
Missings	38	239.3684	184	92	365
CIIa	609526	231.9031	108	41	290
males	350611	223.2859	103	39	275
Females	258877	243.5728	117	43	311
Missings	38	239.3684	184	92	365
CIIb	632355	204.0694	97	35	250
males	364231	196.3675	94	34	240
Females	268086	214.5285	103	37	273
Missings	38	239.3684	184	92	365
AIIIb	489461	281.7411	136	53	365
males	278833	270.5537	127	50	350
Females	210590	296.5614	151	57	385
Missings	38	239.3684	184	92	365
BIIIb	204074	217.6278	109	31	275
males	106654	218.6805	108	29	278
Females	97382	216.4663	111	33	275
Missings	38	239.3684	184	92	365
CIIIb	204074	217.6278	109	31	275
males	106654	218.6805	108	29	278
Females	97382	216.4663	111	33	275
Missings	38	239.3684	184	92	365

Table 6

Descriptive statistics for marginal unemployment duration distribution for the benchmark concepts using the LEH stratified by gender

Implementation	num_obs	mean	median	0.25-Quantile	0.75-Quantile
NE	318596	246.2681	136	52	352
males	195928	240.1096	127	51	337
females	122668	256.1047	151	55	365
UPIT	284369	187.9828	100	40	243
males	174360	180.419	94	39	226
females	110009	199.9711	110	42	273