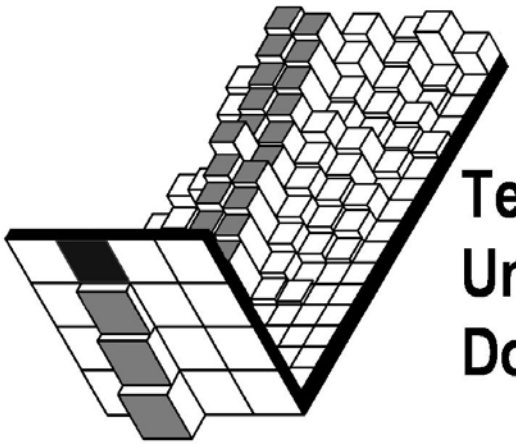
Technical Report

12/2009

Biases in the measurement of labour market dynamics

Ronald Bachmann, Sandra Schaffner



SFB 475
Komplexitätsreduktion in multivariaten Datenstrukturen

Sonderforschungsbereich 475
Universität Dortmund
44221 Dortmund
Germany

Biases in the Measurement of Labour Market Dynamics

Ronald Bachmann
RWI Essen

Sandra Schaffner *†
RWI Essen

June 29, 2009

Abstract

This paper analyses worker transitions on the German labour market derived from different data sources. These include the two German micro data sets which provide high-frequency observations on workers' employment and unemployment histories: the German Socioeconomic Panel (SOEP) and the IAB Employment Subsample (IABS). This exercise thus yields a comprehensive overview of German labour market dynamics. Furthermore, it highlights the differences between the results obtained from a retrospective survey, the SOEP, and a process-induced administrative data set, the IABS. In particular, our analysis shows which groups of the labour market are particularly affected by measurement error. We also show which role measurement issues play when establishing the stylised facts about the cyclicality of labour market dynamics.

Keywords: Gross worker flows; SOEP; IABS.

JEL classification: J63; J64; J62

*Address for correspondence: RWI Essen, Hohenzollernstr. 1-3, 45128 Essen, Germany. E-mail addresses: bachmann@rwi-essen.de, schaffner@rwi-essen.de.

†We are grateful to participants of the SOEP2008 User Conference, the EALE conference, and a seminar at RWI Essen for helpful comments, and to the staff at the IAB for hospitality and for help with the data. We gratefully acknowledge the support of the Collaborative Research Centre SFB 475 (DFG) and of the Leibniz Association. Sarah Mölders and Aneliya Petkova provided able research assistance.

1 Introduction

The analysis of labour market dynamics has been a very active field of research for the last two decades.[1] Recently, interest in this topic has been increased further by the debate on whether the cyclicality of unemployment in the U.S. is mainly due to variation in inflows into or to outflows from unemployment.[2] One of the methodological conclusions from this debate is the fact that the use of different data sources plays an important role for the results eventually obtained (cf. Yashiv, 2008). This has important implications for the analysis of German labour market dynamics as well.

In Germany, the two main sources of panel data that have been used to analyse labour market dynamics at a high frequency are the German Socio-Economic Panel (SOEP), and process-induced, administrative data sets provided by the Institute for Employment Research (IAB) of the German employment office (Bundesagentur für Arbeit), such as the IAB employment subsample (IABS). There are important differences between these two sources, which are due to the facts that the SOEP data derive from an annual household survey, while the IABS is constructed from administrative data. This has two major consequences. First, while the SOEP data are designed to be representative for the population as a whole, the IABS data are representative for people working in social security (dependent-status) employment, and only covers the unemployed if they receive benefits. The definition of labour market state in the latter data set is therefore non-standard. Second, the information on the

[1] Blanchard and Diamond (1989, 1990) are seminal articles on gross worker flows in the U.S.. For analyses of worker and job flows in European countries see Burda and Wyplosz (1994), and Contini and Rivelli (1997).

[2] This debate was initiated by Shimer (2007) who claimed that outflows are by far the most important factor in this respect. This finding has however been strongfully challenged by Elsby, Michaels, and Solon (2009) and Fujita and Ramey (forthcoming), stressing the importance of both inflows and outflows.

labour market state of an individual in the SOEP is collected retrospectively, which is potentially an important source of measurement error. The IABS, on the other hand, is likely to be more accurate because it is based on administrative data where individuals and firms are subject to penalties if the information is not supplied, or inaccurately so.

In order to quantify the differences between the results obtained from the SOEP and from the IABS, and to find the reasons for potential discrepancies, we proceed as follows. First, we transform the IABS to a monthly panel data set in order to obtain the same structure as the SOEP. We then make an unconditional comparison, i.e. we calculate unemployment rates, worker flows, and employment and unemployment spells, from the original samples. We repeat this exercise after restricting the SOEP to the IABS population (i.e. workers covered at least once by social security legislation). This is the conditional comparison, which also allows us to quantify the effect of not taking into account workers which were never covered by social security legislation in the IABS. In each case, we try to find the reasons for differences between the results from the two different data sets. Finally, we apply an external validation strategy by comparing the stocks of employment and unemployment implied by the computed flows to the official statistics provided by the German Employment Office. In order to obtain consistent time series on employment, unemployment, and labour market transitions, we conduct the entire analysis for East and West Germany separately.

Our study is also the first to provide a comprehensive set of stylized facts on high-frequency labour market dynamics for Germany. Correctly measuring labour market states and transitions between them is important for a number of reasons: It is crucial for understanding the dynamic nature of the labour market (cf. Schmidt, 2000), for analyses of the impact of institutions on labour

market dynamics as in Bauer, Bender, and Bonin (2007), and for the measurement of the effect of employment experience and tenure on job stability and wages (e.g. Burda and Mertens, 2001). Finally, results from studies of the effect of the length of unemployment spells on exit rates and wages (e.g. Tatsiramos, forthcoming) are likely to be affected as well.

The plan of the paper is as follows. The next section discusses the literature related to our analysis. In the third part of the paper, we describe the two data sets used, compare their design, and present our methodology for calculating stocks and flows. Section 4 presents the empirical evidence, emphasizing cross-sectional, time-series and cyclical features of the computed series. The last part of the paper summarizes and concludes the discussion.

2 The Measurement of Labour Market Dynamics in the Literature

As pointed out above, the analysis of labour market dynamics has been a particularly active field of research in the U.S. recently. However, empirical studies have come up with contradictory findings. Yashiv (2008) summarizes these studies. He argues that some data facts can be agreed upon, such as the fact that both accessions to employment and separations from it display important cyclicality and volatility. However, there are also areas of strong disagreement. These include the relative volatility of job finding and separation rates, where different data sets produce different results. Furthermore, the fit of the gross flows data with net employment growth data is found to differ across studies and not to be high.

Measurement issues have been an important topic in labour market re-

search for some time. Abowd and Zellner (1985) propose an adjustment procedure for the Bureau of the Census and Bureau of Labor Statistics gross labor-force flows data that also takes into account classification errors in the CPS. They find that accounting for this source of error reduces estimated movements by 8%-49%, and increases estimated consecutive periods of unemployment by 18%. For the UK, Paull (2002) examines biases in the reporting of labour market dynamics in the British Household Panel Survey (BHPS). She finds that shorter spells of all kind (employment, self-employment, out of the labour force and education spells) are less likely to be recalled correctly than longer ones. She also finds recall bias to have important effects on the measurement of labour market dynamics: this bias significantly reduces transition rates between all types of state, and increases median spell lengths for employment, unemployment, and time out of the labour force.

For Germany, Jürges (2007) compares current and one-year retrospective data on unemployment in the SOEP for the time period 1985 to 2001. He finds that 17% of all unemployment spells are not reported one year later, and another 8 % are misreported. Furthermore, the ratio of retrospective to current unemployment is not stable over time.

There are several studies about German labour market dynamics. Erlinghagen (2005) uses the SOEP in order to analyse the evolution of lay-offs and job security for the time period 1985-2001. He finds that the business cycle is the most important determinant for the observed evolution, and that there is no discernible long-run trend. Schmidt (2000) and Kluve, Schaffner, and Schmidt (2006) also use the SOEP, stressing the heterogeneous experience of different demographic groups, especially with respect to their sensitivity to cyclical factors. Bachmann (2005) uses the IAB employment subsample (IABS), a large administrative data set for the time period 1975 to 2001, to

analyse the cyclical features of the German labour market. His most important findings are that direct job-to-job transitions are pervasive and strongly procyclical. Furthermore, hiring play an important role in determining labour market dynamics. Fitzenberger and Garloff (2005) use the same data set to calculate labour market transitions. However, they only consider year-on-year changes, which, as we show below, means that a lot of the actual dynamics are not recorded in their study. Biewen and Wilke (2005) analyze the exit rate out of unemployment to employment using both the IABS and the SOEP. They generate two samples from the IABS, which are comparable to the SOEP. In particular, they use the two definitions of unemployment introduced by Fitzenberger and Wilke (2004), unemployment between jobs, and spells of nonemployment containing at least one day of benefit recipiency. They find important differences. For example, the peak of the exit rate is at 12 months in the SOEP and at 20 months in the IABS.

3 The Data

The **IAB Employment Sample Regional File 1975-2004** (IABS-R04, IABS for short) is provided by the Institute for Employment Research (IAB) of the German Federal Employment Agency. The data base covers 2% of all the persons who, between the 1st January 1975 (for western German employees) or the 1st January 1992 (for eastern German employees) and the 31st December 2004, worked in an employment covered by social security. The data source consists of notifications made by employers to the social security agencies, which include health insurances, statutory pension schemes, and the unem-

ployment insurance agencies.[3] These notifications are made on the behalf of workers, employees and trainees who pay contributions to the social insurance system. This means that, for example, civil servants and the self-employed are not included. Overall, the subsample includes over 1.29 million people, of which 1.1 million are from western Germany. For 1995, the employment statistics, from which the IAB Regional File is drawn, cover nearly 79.4% of the employed persons in western Germany, and 86.2% of all employed persons in eastern Germany. As for the unemployed, only those entitled to unemployment benefits are covered. Workers becoming unemployed who achieved entitlement are not instantaneous eligible if they quit the job.

The **German Socioeconomic Panel** (SOEP) is a representative annual survey of private households in Germany that was started in 1984 in West Germany, and in 1992 in East Germany (see Wagner, Frick, and Schupp, 2007, for an overview). On average, the SOEP covers 4,500 households with 11,000 individuals (of which about 6,000 are employed) per year. Panel attrition arises if a person refuses to answer, dies or goes abroad. New households emerge if an individual separates from a previously interviewed household, e.g. by moving out, and forms or becomes part of a new household and by refreshment samples. In the questionnaire, among other things, respondents are asked to fill in a calendarium to report their major activity for each month of the preceding year. This calendarium can thus be used to calculate individual labour market transitions on a monthly basis. Employment in the SOEP refers to full-time work, part-time work, and vocational training; unemployment, on the other hand, refers to registered unemployment, and non-participation is the residual category, comprising among others schooling, military service, community

[3]For a complete description of a previous, and very similar version of this data set, see Bender, Haas, and Klose (2000); Drews (2007) provides a description of the current version.

service, maternity leave, and retirement. In the SOEP, parallel labour market episodes are recorded. Individuals doing so are not counted as employed if they declare to be both employed and, in the same month, to be in the military service, schooling, community service, maternity leave, or retirement. Students working during vacation or retired persons performing part-time jobs are not of key interest in the analysis of labour market transitions.

Comparing the two data sets, the main differences are as follows. First, while the SOEP is representative for the entire population, the IABS only covers workers who are subject to social security contributions, or who are unemployed and receive unemployment benefits. Second, the information on individual labour market histories in the SOEP is retrospective, self-reported, and available at a monthly frequency; the IABS data, on the other hand, are process-induced and exact to the day.

Stocks and flows are computed as follows from the two data sets. In the case of the IABS, we use the employment and unemployment spells in order to identify for each individual the labour market state (employment, unemployment, non-participation) which in terms of working days prevailed during a given month. As for the SOEP, the individual information on labour market status provided by the monthly calendarium is used. We construct a third data set by restricting the SOEP to the IABS population (SOEP_R), in order to be able to assess the importance of sample composition. The most important difference between the SOEP and the SOEP_R is that the latter does not include the self-employed and civil-servants ("Beamte"). For all three data sets, labour market transitions are defined as a change in labour market state from one month to the next.

Table A.1 displays the descriptive statistics for the three data sets used. It becomes apparent that the data sets compare well with respect to gender, age,

and geographical location. The skill categories, however, display significant differences, which we put down to different definitions in the two data sets. In the following analysis, we therefore do not distinguish between different skill groups, but concentrate on the former three characteristics.

4 Empirical Evidence

4.1 Cross-Sectional Features

The unemployment rates from the IABS and the SOEP for different worker groups are displayed in Table A.2. As one can see, the unemployment rate in the IABS is generally higher than the unemployment rate in the SOEP. One potential explanation for this are differences in sample definitions. While only employees subject to social insurance contributions and unemployment benefit recipients are included in the IABS, the SOEP covers all workers and unemployed. The latter data therefore include more employed workers, which lowers the unemployment rate. The calculated unemployment rates in the restricted SOEP sample are closer to those from the IABS. This finding suggests that some differences can be explained by differences in sample composition. This result holds for both men and women. The table also shows that the difference in unemployment rates is particularly large for older workers, both in East and West Germany. These workers are more likely to be in early retirement schemes and similar policy measures. Participation in such policy measures are recorded differently in the SOEP compared to the administrative data of the IABS: While workers will tend to classify themselves as retired (non-participation) in the SOEP, in the IABS workers in early retirement schemes are counted as unemployed because they receive unemployment benefits. As

a result, unemployment rates of older workers are higher in the IABS than in the SOEP. For young workers in West Germany, the relationship is the reverse: Unemployment rates are higher in the SOEP than in the IABS. This can be explained by the fact that young workers, having a short labour market history, often do not meet the eligibility criteria for unemployment benefits. They are therefore not counted as unemployed in the IABS, but report being unemployed in the SOEP.

Unemployment rates are a result of the flows into and out of unemployment. We distinguish three core labour force states: employment (E), unemployment (U), and out-of-the-labour-force (N). The flows between the different states are denoted by F^{xy}, where x is the preceding state, while y is the current state. Therefore F^{UE} is the flow from unemployment to employment from $t-1$ to t. The first moments of the flows into employment are given in Panel (a) of Table A.3. The last row displays the range of values of U.S. studies as summarized by Yashiv (2008). The findings are quite similar for both data sets. The hiring rate from unemployment,$\frac{F^{UE}}{E}$, is about 0.6% per month. This is less than half of the lowest figure reported by Yashiv (2008) for the U.S. labour market. The job finding rate of the unemployed in Germany, $\frac{F^{UE}}{U}$, is even smaller in comparison to the U.S. figures. Our highest estimate (6.9%, from the IABS) does not even reach a third of the lowest value from Yashiv (2008). The picture is relatively similar for the separation flows from employment to unemployment and non-participation, which are presented in Table A.3(b), which are all considerably smaller in Germany than in the US. The job outflow rate to unemployment, $\frac{F^{EU}}{E}$, for example, reaches only 0.6% per month, which is half the lowest U.S. figure.

Given the differences between the construction of the IABS and the SOEP, one would expect large differences between the labour market dynamics calcu-

lated from the two data sets. Surprisingly however, the results are relatively similar. For example, the job finding rate of the unemployed is 6.5% per month according to the SOEP, and 6.9% according to the IABS. Looking at the flows between unemployment and employment for the different subgroups, one can see that there are hardly any differences between men and women; however, differences exist between East and West Germany (cf. Table A.4). In particular, worker flows are considerably higher in East Germany than in West Germany. The flow rates in East Germany are almost as high as the ones reported by Yashiv (2008) for the U.S. Note that the mean job loss rate is higher than the mean hiring rate in East Germany because employment is decreasing over the time period considered. The differences in the means of the different data sets are small for all four subgroups as well.

4.2 Time-Series Properties

Figure A.1 shows the evolution of the unemployment rates for West Germany for the time period 1983-2004 and for East Germany for the time period 1992-2004, for men and women separately. Each time series is computed using the micro data from the IABS and from the SOEP. While the general evolution of the unemployment rate is similar in the two data sets, there are differences between worker groups. In order to evaluate the impact of sample composition over time, the figure also displays the results from the SOEP data set restricted to the IABS population, SOEP_R. If the entire difference between the IABS and the SOEP was due to composition effects, the results from the IABS and SOEP_R would be identical. As Figure A.1 makes clear, this is generally not the case. In particular, sample composition seems to play an important role for men in West Germany. For East German men, sample composition also

seems to exert some influence, although less so than for West German men. For women in both East and West Germany (Panels b and d), however, the composition effect only plays a minor role, as witnessed by the fact that the results from the IABS and SOEP$_R$ are very similar. These results can be explained by the fact that West German men are more likely self-employed or civil servants than the other groups. The differences between the three curves remain relatively stable over time for men in West Germany and East Germany. By contrast the unemployment rates for women calculated from the three samples converge over time. One explanation for this is the fact that the coverage of social security is higher for women and that women are more likely to work part-time. Marginal part-time employment (so called Mini-Jobs) has only been covered by social security since 1999. Therefore the rates move more closely together since 1999.

To capture the comovement between the unemployment rates computed for the different data sets, Table A.5 shows the corresponding correlation coefficients. These figures are generally in line with the conclusions drawn from the cross-sectional analysis above. The correlations are strongest for West German workers, and especially for men. In the West, they are lower for younger workers (younger than 25 years), in the East they are lower for younger and older workers (older than 50 years). As the correlations between the IABS and the SOEP and the IABS and SOEP$_R$ are very similar, sample composition does not seem to play an important role in this context.

The correlations between the worker flows calculated from the different data sets are presented in Table A.6. Even though the mean flows presented in Table A.3 are quite similar, the correlations are relatively small, especially for the separation flows. While the average separation rates do not differ much between the IABS and the SOEP, the correlation is -0.148. The outflows rate

from employment, $\frac{F^{EN}+F^{EU}}{E}$, are more highly correlated, but the means differ from each other. For all separation flows, the correlations between the IABS flows and the SOEP$_R$ flows are higher than the correlations between the IABS flows and the SOEP flows. Therefore, different sample compositions have a role to play. However, the former correlations remain relatively low. Therefore, inaccuracies in the self-reported retrospective information of flows in the SOEP seem to be of some importance as well. The highest correlations can be found for the aggregate inflow into employment, $\frac{F^{UE}+F^{NE}}{E}$ while the correlation of the hiring flows from unemployment are small.

Looking at the correlations between the time series for different worker groups displayed in part (c) of Table A.6, it becomes obvious that the comovement for the hiring rates, $\frac{F_{UE}}{E}$, calculated from the IABS and the SOEP is relatively strong, with correlation coefficients of almost than 0.4 for men in both East and West Germany. The job loss rate into unemployment, however, is comparatively lower for every worker group, and even negative for women in East and West Germany. It is remarkable that all correlations are lower for women than for men. This result is consistent with the findings by Jürges (2007) that women are more likely to report wrong states in the retrospective data.

We observe high transition rates in the SOEP data between December and January which seems to be due to the retrospective data. Furthermore, additional seasonal volatilities seem to be less pronounced in the SOEP than in the IABS. We therefore seasonally adjust all time series and calculate the correlations of these series. The results are presented in Table A.7. All correlations are higher than without seasonal adjustment. The correlations are highest for the transition rates between employment and unemployment, the hiring and job loss rate.

Having found that there exist some differences between the flows computed from the IABS and from the SOEP, especially at a high frequency, we are now interested in which data set yields flows that are consistent with official labour market figures. In order to do so, we use data on employment and unemployment stocks provided by the German Statistical Office and the German Employment Office. We then use our flow measures calculated from the two micro data sets together with the stock-flow identities of the labour market to compute implied employment and unemployment stocks. The stock-flow used read as follows:

$$E_{t+1} = E_t + F^{UE} + F^{NE} - F^{EU} - F^{EN} \tag{1}$$

$$U_{t+1} = U_t + F^{EU} + F^{NU} - F^{UE} - F^{UN} \tag{2}$$

where F^{xy} is a flow from state x to state y between time period t and time period $t+1$. The starting values for E_0 and U_0 are set equal to the officially published figures. The following values for employment and unemployment are calculated from the computed flows and the stock-flow identity. We thus obtain simulated series for the two labour market states, in each case one series simulated with the IABS flow measures, the other one with the SOEP flow measures. A comparison of these implied stocks with the official labour market statistics shows how well the IABS and the SOEP are able to mirror the evolution of aggregate labour market dynamics.

The results of our simulation for the implied employment stocks are displayed in Figure A.2. For West Germany, it becomes apparent that the SOEP flows yield an implied employment stock which is too low compared to the official figures. The general trend is captured well by the IABS for West Germany and the SOEP for East Germany. However, the unemployment rate

in West Germany is better captured by the SOEP. The correlation with the official statistic is the highest for the $SOEP_R$ data. All correlations except for employment in West Germany are high, ranging between 0.74 and 0.89. Migration between the two parts of Germany can possibly explain some differences between the two data sets and the official data. While a respondent's region in the SOEP is his area of residence, in the IABS it the place of the firm for employment spells, and his residence for unemployment spells only. Furthermore, in the SOEP changes in the region can only observed at the interview dates.

The implied unemployment stocks are displayed in Figure A.3. Again, the stock derived from the IABS flows captures dynamics well; however, the unemployment stock is underestimated in the West, and exaggerated in the East. The unemployment stocks implied by the SOEP and SOEP$_R$ do not capture short-run dynamics well and are generally too low compared to the official figures.

4.3 Spells

In order to give a first impression of the spells computed from the IABS and the SOEP, Table A.8 reports averages and percentiles for the duration of employment and unemployment spells computed from the two data sets for different worker groups. East and West Germany differ in their mean duration of employment and unemployment spells: Employment spells are longer in West Germany, which is partly due to the fact that West German workers are observed for a longer time period (i.e. before 1990) than their East German counterparts. On the other hand, unemployment spells in East Germany are longer than in West Germany, which can be attributed to the fact that un-

employment is higher and exit rates out of unemployment are lower in East Germany.

When looking at the differences between the two data sets, several features are noteworthy. First, both, employment and unemployment spells are on average longer in the SOEP than in the IABS, with employed women in West Germany being the only exception. For example, the mean employment spell lasts 36 months in the IABS, but more than 38 months in the SOEP. For unemployment spells, the corresponding figures are 9.8 and 11.3 months in the IABS and the SOEP, respectively. Second, the difference between the two data sets with respect to spell duration is larger for men than for women. While the difference in employment spells is larger in West Germany than in East Germany, it is the other way round for unemployment spells. Third, an inspection of the percentiles shows that the differences between the employment spells computed from the two data sets can mainly be found in the lower and middle parts of the distribution; for unemployment spells, the differences are larger in the upper parts of the distribution.

For a better illustration of the differences in spell lengths, the distribution of employment and unemployment durations is depicted in Figures A.4 and A.5, respectively. It becomes apparent that that there is a general tendency in the SOEP to underreport short employment and unemployment spells (with the exception of durations of 1 month, which are more prevalent in the SOEP), which is in all likelihood due to the retrospective nature of the SOEP. This means that there is some recollection error involved when workers report their labour market history of the preceding year. The second striking difference between the distributions is a much larger year effect in the SOEP than in the IABS for both employment and unemployment spells, i.e. there is a strong tendency to report durations of 12 months and multiples thereof in the SOEP.

Again this seems to be caused by retrospection, which leads to recall bias: when filling in the employment calendarium of the SOEP, survey respondents tend to systematically adjust the beginning/end of their employment or unemployment spells to January/December of the year, although the true starting date was February (or later) and the true ending date was November (or earlier). Both effects, the underreporting of short durations and the 12 months peak are more pronounced for the employment spells than for the unemployment spells. Furthermore, the differences seem to be smallest for East German women.

5 Conclusion

There exist two large micro data sets yielding information on gross worker flows in Germany: the IABS and the SOEP. While the IABS is a process-induced administrative data set, the SOEP is a survey with retrospective information on labour market states. On the one hand, the SOEP suffers from recall error of the respondents; the IABS, on the other hand, does not cover the entire labour force. Therefore labour market transitions are subject to measurement error in both data sets. By comparing the flows computed from the two data sets, we quantify the differences in the gross flows.

Comparing the averages of the unemployment rates and the flows between unemployment and employment, the differences are not substantial. The differences in unemployment rates are larger for old workers which could be due to early retirement. Generally, the unemployment rate is higher in the IABS than in the SOEP, with young workers in West Germany being the only exception.

The differences between the data sets remain small also when the subgroups

are investigated. The levels of the gross flows therefore seem to be relatively accurately measured. However, the correlations between the time series computed from the two data sets are relatively small, especially for women. The differences in the specifications of the labour market states as well as the recall error in the SOEP seem to be most important for women. It is remarkable that the correlations of hiring rates are higher than those of the job loss rates. This holds true for all groups.

The time series of the unemployment rate suggest that composition effects seem to play a bigger role for men than for women. They are more likely not to be captured by the IABS data, whereas the unemployment rates of women seem to converge over time. This could be an indication that the coverage of marginal employment by social security accounts leads to a good coverage of the whole female labour force by administrative data.

Although the differences are relatively small between unemployment and gross flows of the two data sets, there are some shortcomings for both data sources in the extrapolation of employment and unemployment of official data. While dynamics are captured by the IABS, the main trend is better captured by the SOEP.

Finally, the distribution of the spell lengths shows that short spells are underrepresented in the SOEP while there are peaks at twelve months which seem to be a result of misreporting. These findings are less pronounced for unemployment spells than for employment spells.

Summing up, it can be stated that there is no clear sign which of the two data sets fulfil best the needs for an analysis of labour market transitions. Therefore the possible measurement errors and the shortcomings of the data sets have to be taken into account. If the main interest is in short spells the retrospective data of the SOEP seem to be insufficient. A part of the labor

force is not covered by the IABS which has influences on the unemployment rates especially for men. Potential biases emanating from these measurement errors should be taken into account by researchers using either of the two data sets for analyses pertaining to employment and unemployment dynamics.

References

Abowd, J. M., and A. Zellner (1985): "Estimating gross labor-force flows," *Journal of Business & Economic Statistics*, 3(3), 245–283.

Bachmann, R. (2005): "Labour market dynamics in Germany: Hirings, separations, and job-to-job transitions over the business cycle," Collaborative Research Center 649 Discussion Paper 2005-45, Humboldt-Universität zu Berlin.

Bauer, T. K., S. Bender, and H. Bonin (2007): "Dismissal protection and worker flows in small establishments," *Economica*, 74(296), 804–821.

Bender, S., A. Haas, and C. Klose (2000): "IAB employment subsample 1975-1995. Opportunities for analysis provided by the anonymised subsample," IZA Discussion Paper 117, Institute for the Study of Labor (IZA).

Biewen, M., and R. A. Wilke (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?," ZEW Discussion Papers 05-05, ZEW - Zentrum für Europäische Wirtschaftsforschung / Center for European Economic Research.

Blanchard, O., and P. A. Diamond (1989): "The Beveridge Curve," *Brookings Papers on Economic Activity*, 1989(1), 1–76.

——— (1990): "The cyclical behaviour of gross flows of workers in the US," *Brookings Papers on Economic Activity*, 2, 85–11.

Burda, M. C., and A. Mertens (2001): "Estimating wage losses of displaced workers in Germany," *Labour Economics*, 8(1), 15–41.

Burda, M. C., and C. Wyplosz (1994): "Gross worker flows and job flows in Europe," *European Economic Review*, 38(6), 1287–1315.

Contini, B., and R. Rivelli (1997): "Gross flows vs. net flows in the labor market: What is there to be learned?," *Labour Economics*, 4(3), 245–263.

Drews, N. (2007): "Variablen der schwach anonymisierten Version der IAB-Beschäftigtenstichprobe 1975-2004," FDZ Datenreport 3/2007, Institut für Arbeitsmarkt- und Berufsforschung (IAB).

Elsby, M. W., R. Michaels, and G. Solon (2009): "The Ins and Outs of Cyclical Unemployment," *American Economic Journal: Macroeconomics*, 1(1), 84–100.

Erlinghagen, M. (2005): "Entlassungen und Beschäftigungssicherheit im Zeitverlauf - Zur Entwicklung unfreiwilliger Arbeitsmarktmobilität in Deutschland," *Zeitschrift für Soziologie*, 34(2), 147–168.

Fitzenberger, B., and A. Garloff (2005): "Labor market transitions and the wage structure in Germany," *Jahrbücher für Nationalökonomie und Statistik*, 227(2), 115–152.

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 Papers 04-24, ZEW - Zentrum für Europäische Wirtschaftsforschung / Center for European Economic Research.

Fujita, S., and G. Ramey (forthcoming): "The cyclicality of separation and job finding rates," *International Economic Review*.

Jürges, H. (2007): "Unemployment, life satisfaction and retrospective error," *Journal of the Royal Statistical Society*, 170(1), 43–61.

Kluve, J., S. Schaffner, and C. Schmidt (2006): "Labor Force Status Dynamics in the German Labor Market: Individual Heterogeneity and Cyclical Sensitivity," Technical Report 20/06, SFB 475.

Paull, G. (2002): "Biases in the reporting of labour market dynamics," IFS Working Papers W02/10, Institute for Fiscal Studies.

Schmidt, C. M. (2000): "Persistence and the German unemployment problem: Empirical evidences on German labor market flows," *Economie et Statistique*, 332-333(2000-2/3), 83–95.

Shimer, R. (2007): "Reassessing the Ins and Outs of Unemployment," NBER Working Papers 13421, National Bureau of Economic Research, Inc.

Tatsiramos, K. (forthcoming): "Unemployment insurance in Europe: Unemployment duration and subsequent employment stability," *Journal of the European Economic Association*.

Wagner, G. G., J. R. Frick, and J. Schupp (2007): "The German Socioeconomic Panel Study (SOEP) - Scope, Evolution and Enhancements," *Schmollers Jahrbuch (Journal of Applied Social Science Studies)*, 127(1), 139–169.

Yashiv, E. (2008): "U.S. labor market dynamics revisited," *Scandinavian Journal of Economics*, 109(4), 779–806.

Appendix A Appendix

A.1 Data

The IAB Employment subsample is a dataset with daily information on employment and unemployment benefits. In order to obtain a data set with monthly frequency, on a monthly basis for each individual, long spells are split into monthly intervals. When the status changes during a particular month, the status featuring the highest number of working days within the month is chosen. Individuals are taken as not participating every time they are not in the data and aged between 16 and 64 in the observation period.

The SOEP is a yearly survey. In addition, monthly information on the labour market status is collected retrospectively for the months of the preceding year. We use this "calendarium" to calculate monthly transition rates. For weighting, the cross sample weights of the survey year are used. Because of panel attrition and other reasons, there exist several refreshment samples during our observation period. All these refreshment samples are also included. While this is not problematic for calculating unemployment rates, some issues emerge when calculating transition rates because the status before entering the sample is not known. Therefore the first transition is taken between January and February for the refreshment samples or other people who enter the sample because of entering a household. The weights for the transitions between December and January are those of the preceding year and are representative without the refreshment sample. However, some people always leave the dataset. Because of the retrospective nature of the data, individuals leave the dataset only between December and January. Their last transition is measured between November and December. Therefore, only those who stay in the sample define the transition rates between December and January. Because of the retrospective data, there is no relationship between the time of leaving the data which is always December and the reasons for not taking part in the next survey which is more than one year later.

In both data sets, some individuals change from East to West Germany or vice versa. These transition rates are not included in our data except for the extrapolation. In all other cases, these transitions are treated as entries to and exits of the data set.

A.2 Graphs and Tables

Table A.1: Descriptive statistics

	SOEP/SOEP$_R$		IABS
Individual characteristics			
Male	49.2%		53.4%
Age	39.97		36.15
West-Germany	85.0%		85.4%
Low-skilled	20.8%		28.9%
Medium-skilled	58.9%		63.4%
High-skilled	20.3%		7.7%
Labour market statistics	SOEP	SOEP$_R$	
Employment rate	0.5875	0.4949	0.4816
Unemployment rate	0.0478	0.0478	0.0502
Non-participation rate	0.3647	0.4573	0.4682
Duration of spell (months)	31.51	30.72	35.03
No. of persons	36,208		2,226,485
No. of person-month observations	3,652,392		274,112,915

Data source: IABS-R04 and SOEP, own calculations. Time period under consideration: 1983-2004.

Table A.2: Unemployment rates

(a) By sex in East and West Germany

	Total		Men		Women	
	West	East	West	East	West	East
SOEP	6.2	16.5	6.2	14.0	6.3	19.3
	[1.8]	[3.3]	[2.0]	[4.0]	[1.8]	[2.9]
$SOEP_R$	7.4	18.2	7.6	16.0	7.0	20.5
	[2.2]	[3.9]	[2.4]	[4.9]	[2.0]	[3.2]
IABS	8.0	19.0	8.4	17.7	7.4	21.2
	[1.7]	[5.2]	[2.2]	[6.1]	[1.2]	[2.7]

(b) By age class

	West			East		
	< 25	$25 - 50$	> 50	< 25	$25 - 50$	> 50
SOEP	7.6	5.1	8.7	12.7	15.1	23.4
	[3.5]	[1.7]	[2.9]	[4.6]	[4.4]	[6.9]
$SOEP_R$	7.9	6.1	10.7	12.9	16.7	25.6
	[3.6]	[1.9]	[3.5]	[4.7]	[4.7]	[7.5]
IABS	6.1	6.9	12.3	13.6	17.0	29.7
	[2.0]	[1.5]	[3.4]	[3.5]	[5.2]	[6.5]

Data source: IABS-R04 and SOEP, own calculations. $SOEP_R$: SOEP restricted to those worker groups that are also observed in the IABS. Time period under consideration: 1983-2004 for West Germany, 1992-2004 for East Germany.

Table A.3: Hiring and separation flows

(a) Hiring Flows to employment

	$\frac{F^{UE}}{E}$	$\frac{F^{UE}}{U}$	$\frac{F^{NE}}{E}$	$\frac{F^{UE+NE}}{E}$	$\frac{F^{NE}}{N}$
SOEP	0.006	0.065	0.013	0.019	0.014
	[0.002]	[0.029]	[0.018]	[0.018]	[0.0018]
SOEP_R	0.006	0.065	0.013	0.019	0.014
	[0.002]	[0.029]	[0.018]	[0.018]	[0.018]
IABS	0.007	0.069	0.009	0.016	0.011
	[0.003]	[0.029]	[0.003]	[0.005]	[0.030]
Yashiv (2008)	0.015-0.020	0.247-0.321	0.013-0.029	0.030-0.046	

(b) Separation flows from employment

	$\frac{F^{EU}}{E}$	$\frac{F^{EN}}{E}$	$\frac{F^{EN+EU}}{E}$
SOEP	0.006	0.0200	0.027
	[0.003]	[0.036]	[0.038]
SOEP_R	0.006	0.013	0.020
	[0.003]	[0.014]	[0.015]
IABS	0.006	0.011	0.017
	[0.002]	[0.014]	[0.014]
Yashiv (2008)	0.013-0.020	0.015-0.032	0.029-0.050

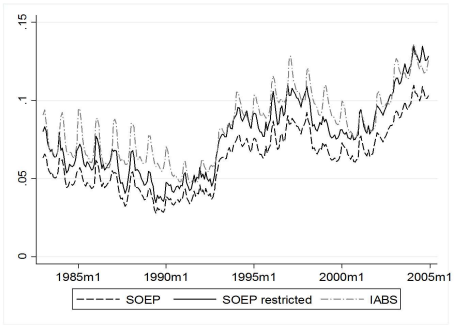
Data source: See notes to Table A.2. Yashiv (2008) summarizes results of different studies for the U.S.

Table A.4: The flows between employment and unemployment for different worker groups, and their correlations between different data sets

	Men West-Germany		Men East-Germany		Women West-Germany		Women East-Germany	
	$\frac{F^{UE}}{E}$	$\frac{F^{EU}}{E}$	$\frac{F^{UE}}{E}$	$\frac{F^{EU}}{E}$	$\frac{F^{UE}}{E}$	$\frac{F^{EU}}{E}$	$\frac{F^{UE}}{E}$	$\frac{F^{EU}}{E}$
SOEP	0.004	0.005	0.011	0.012	0.004	0.005	0.011	0.013
	[0.002]	[0.003]	[0.005]	[0.006]	[0.003]	[0.003]	[0.006]	[0.007]
SOEP_R	0.005	0.006	0.013	0.014	0.005	0.005	0.012	0.013
	[0.003]	[0.003]	[0.006]	[0.007]	[0.003]	[0.003]	[0.006]	[0.007]
IABS	0.006	0.006	0.013	0.017	0.004	0.006	0.010	0.014
	[0.003]	[0.004]	[0.007]	[0.019]	[0.001]	[0.002]	[0.003]	[0.009]

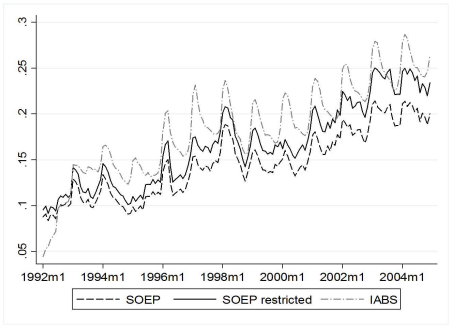
Data source: See notes to Table A.2.

Figure A.1: Unemployment rates of men and women in West and East Germany



(a) Men - West Germany

(b) Women - West Germany



(c) Men - East Germany

(d) Women - East Germany

Data source: IABS-R04 and SOEP, own calculations

Table A.5: Correlations between the unemployment rates from the IABS and the SOEP for different worker groups

(a) By sex in East and West Germany

	Total		Men		Women	
	West	East	West	East	West	East
corr(IABS, SOEP)	0.934	0.845	0.949	0.925	0.832	0.618
corr(IABS, $SOEP_R$)	0.927	0.850	0.943	0.922	0.827	0.652

(b) By age class

	West			East		
	< 25	$25 - 50$	> 50	< 25	$25 - 50$	> 50
IABS, SOEP	0.661	0.766	0.768	0.491	0.830	0.644
IABS, $SOEP_R$	0.673	0.783	0.810	0.477	0.885	0.632

Data source: See notes to Table A.2.

Table A.6: Correlations between the hiring and separation flows from the IABS and the SOEP

(a) Hiring flows to employment

	$\frac{F^{UE}}{E}$	$\frac{F^{UE}}{U}$	$\frac{F^{NE}}{E}$	$\frac{F^{UE+NE}}{E}$
IABS, SOEP	0.116	0.069	0.578	0.630
IABS, $SOEP_R$	0.115	0.069	0.578	0.630

(b) Separation flows from employment

	$\frac{F^{EU}}{E}$	$\frac{F^{EN}}{E}$	$\frac{F^{EN+EU}}{E}$
IABS, SOEP	-0.145	0.281	0.240
IABS, $SOEP_R$	-0.145	0.299	0.257

(c) Hiring and separation for different worker groups

	Men West-Germany		Men East-Germany		Women West-Germany		Women East-Germany	
	$\frac{F^{UE}}{E}$	$\frac{F^{EU}}{E}$	$\frac{F^{UE}}{E}$	$\frac{F^{EU}}{E}$	$\frac{F^{UE}}{E}$	$\frac{F^{EU}}{E}$	$\frac{F^{UE}}{E}$	$\frac{F^{EU}}{E}$
IABS, SOEP	0.395	0.170	0.349	0.086	0.049	-0.139	0.226	-0.104
IABS, $SOEP_R$	0.404	0.165	0.372	0.099	0.060	-0.141	0.252	-0.109

Data source: See notes to Table A.2.

Table A.7: Correlations between the hiring and separation flows from the IABS and the SOEP - seasonal adjusted data

(a) Hiring flows to employment

	$\frac{F^{UE}}{E}$	$\frac{F^{UE}}{U}$	$\frac{F^{NE}}{E}$	$\frac{F^{UE+NE}}{E}$
IABS, SOEP	0.650	0.562	0.080	0.201
IABS, SOEP_R	0.650	0.562	0.080	0.201

(b) Separation flows from employment

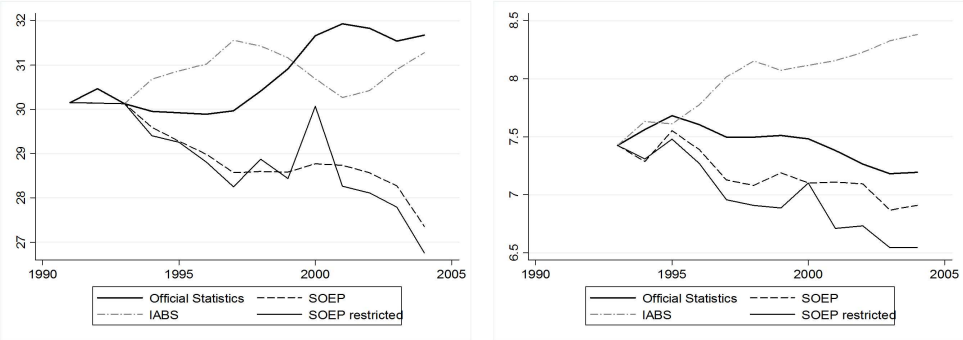
	$\frac{F^{EU}}{E}$	$\frac{F^{EN}}{E}$	$\frac{F^{EN+EU}}{E}$
IABS, SOEP	0.551	-0.046	-0.016
IABS, SOEP_R	0.551	0.096	0.116

(c) Hiring and separation for different worker groups

	Men West-Germany		Men East-Germany		Women West-Germany		Women East-Germany	
	$\frac{F^{UE}}{E}$	$\frac{F^{EU}}{E}$	$\frac{F^{UE}}{E}$	$\frac{F^{EU}}{E}$	$\frac{F^{UE}}{E}$	$\frac{F^{EU}}{E}$	$\frac{F^{UE}}{E}$	$\frac{F^{EU}}{E}$
IABS, SOEP	0.398	0.169	0.351	0.091	0.094	-0.149	0.204	-0.103
IABS, SOEP_R	0.409	0.160	0.376	0.105	0.106	-0.163	0.230	-0.100

Data source: See notes to Table A.2.

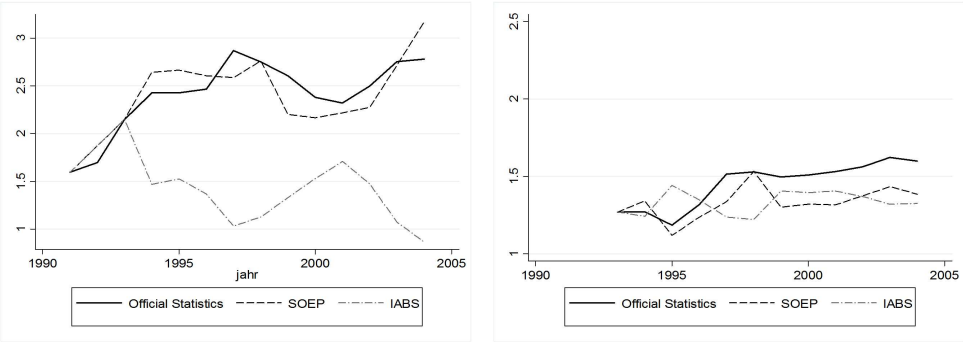
Figure A.2: Extrapolation of employment for West and East Germany



(a) West Germany (b) East Germany

Data source: IABS-R04 and SOEP, own calculations; BA Statistics are officially published figures from the German Employment Office (*Bundesagentur für Arbeit*).

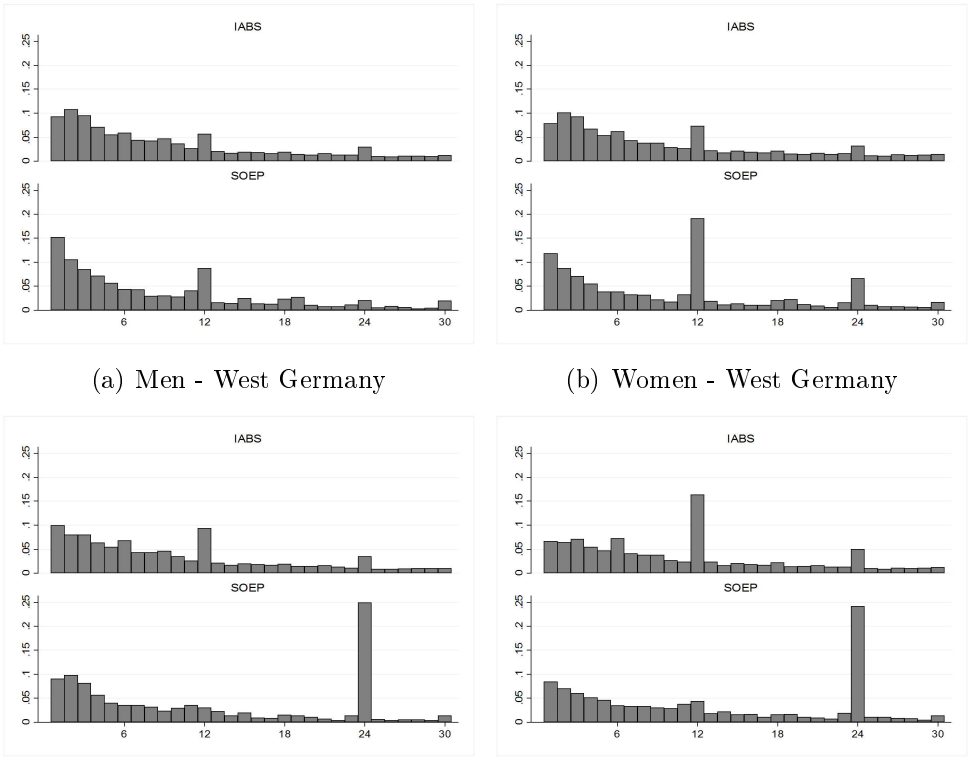
Figure A.3: Extrapolation of unemployment for West and East Germany



(a) West Germany (b) East Germany

Data source: See notes to Figure A.2.

Figure A.4: Duration distribution of employment spells for men and women in West and East Germany



(a) Men - West Germany

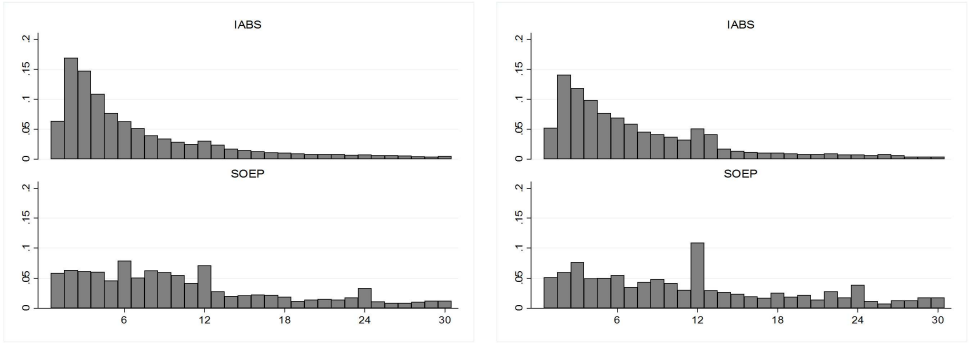
(b) Women - West Germany

(c) Men - East Germany

(d) Women - East Germany

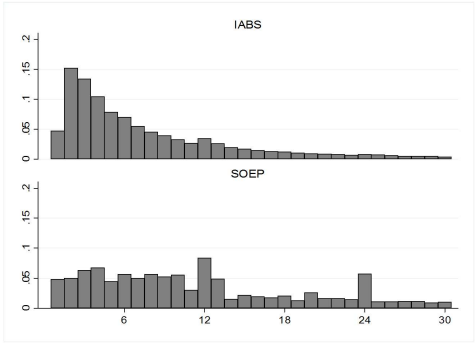
Data source: IABS-R04 and SOEP, own calculations

Figure A.5: Duration distribution of unemployment spells for men and women in West and East Germany

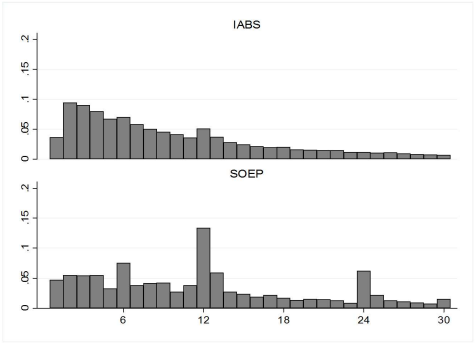


(a) Men - West Germany

(b) Women - West Germany



(c) Men - East Germany



(d) Women - East Germany

Data source: IABS-R04 and SOEP, own calculations

Table A.8: Statistics of the spell length from the IABS and the SOEP for different worker groups in months

			Percentiles					
Employment Spells								
	Mean	S.D.	10	25	50	75	90	100
	Total							
IABS	34.333	48.694	2	5	13	41	96	264
SOEP	37.248	45.0	4	10	21	48	92	252
	West Germany							
IABS	35.591	50.8	2	5	14	42	100	264
SOEP	38.38	46.9	4	11	23	48	96	252
	East Germany							
IABS	26.672	34.1	2	5	12	34	72	264
SOEP	37.682	38.7	3	8	17	46	84	156
	Men West Germany							
IABS	35.955	52.5	2	5	13	42	105	264
SOEP	44.89	53.3	4	12	24	60	120	252
	Women West Germany							
IABS	35.099	47.9	2	5	15	43	96	264
SOEP	30.92	36.7	4	10	17	38	72	252
	Men East Germany							
IABS	24.248	31.9	2	5	12	30	63	260
SOEP	34.27	37.9	3	9	19	48	84	156
	Women East Germany							
IABS	30.002	36.6	3	6	13	36	84	264
SOEP	32.68	37.8	3	8	16	42	84	156
Unemployment Spells								
	Total							
IABS	10.15	12.98	2	3	6	12	23	264
SOEP	9.85	12.21	1	3	6	12	23	165
	West Germany							
IABS	9.75	12.93	2	3	5	12	22	264
SOEP	10.43	12.56	2	3	6	12	24	165
	East Germany							
IABS	11.48	13.06	2	3	7	14	27	157
SOEP	10.43	12.15	1	3	6	13	24	143
	Men West Germany							
IABS	9.80	13.65	2	3	5	11	23	264
SOEP	9.88	13.53	1	2	5	12	24	165
	Women West Germany							
IABS	9.68	11.71	2	3	6	12	22	264
SOEP	8.96	10.04	1	3	6	12	20	111
	Men East Germany							
IABS	10.00	11.89	2	3	6	12	23	157
SOEP	8.74	10.42	1	2	5	11	21	89
	Women East Germany							
IABS	13.29	14.15	2	4	9	17	30	156
SOEP	12.10	13.44	2	3	8	15	27	143

Data source: IABS-R04 and SOEP, own calculations.