Employment mobility patterns and its determinants in Germany: Comparing the reliability of prospective and long-term retrospective data

-Draft version-

Anna Manzoni

Ruud Luijkx

Tilburg University, The Netherlands

To be presented at the RC28 Spring Meeting in Brno, Czech Republic, May 24-27, 2007

<u>Abstract</u>

To analyse employment mobility of people on the labour market, longitudinal studies are an ideal framework, but longitudinal surveys can be conducted using different research designs. This paper focuses on differences between prospective and retrospective data, to gauge the extent to which longitudinal surveys are able to unravel labour market dynamics and its underlying causal mechanisms.

We assess to which extent retrospective and prospective data lead to the same or different conclusions with respect to basic career moves. This descriptive question brings us to the explanatory question why differences between the designs may occur.

Using the German Life History Study (GLHS) and the German Socio-Economic Panel (GSOEP), we are paying attention on the one hand to the impact of recall on measured labour market dynamics, in order to assess to what degree biases in the reported data may affect the estimation of models of labour market dynamics, and on the other hand to "seam effects", which generate heaping in the transition probabilities between subsequent years (calendars).

When individuals are asked to recall their past, the possibility of inaccurate reporting may occur. Previous studies (Reimer and Künster, 2004; Reimer, Matthes and Pape, 2006) on recall bias in the GLHS have compared interviews obtained in different waves for the period for which they overlap and others have tried to link it with register data. Autobiographical remembering tends to simplify and conventionalise the life course and errors follow a memory pattern of simplification and conventionalisation, leading to an overall reduction in episodes and transitions. Moreover, they are not distributed evenly among all respondents.

We test whether and to what extent memory bias explains the differences in the estimates of labour market transitions obtained with the different data sources. Furthermore, we want to inspect whether determinants of careers are different when using the two designs.

After having tested the reliability of the calendar information in the GSOEP, we try to grasp the extent of the long term retrospective bias by comparing life-course data coming from the GLHS with the data gathered annually in the GSOEP. While, due to memory bias, we expect lower transition probabilities when using retrospective data, we also try to take into account possible seam effects (spurious transitions between calendars collected in subsequent years), which might overestimate transition probabilities in the panel data. By combining person-period records from prospective and retrospective data (for the period in time for which we have both) in one dataset and by constructing a dummy-variable indicating the type of data the record came from, we can assess the design effect. Furthermore, we can test various design implications by interacting this variable with other independent variables.

Multilevel (multistate) discrete-time models for event history data are used to model transitions between labour market states, including both time constant and time varying variables.

We expected events occurring further in the past to be less well recalled than more recent events and stable careers to be better recalled than volatile ones.

Our findings show that the two designs lead to the same patterns and reveal the same determinants of labour market transitions. Yet, the levels seem quite different. The results of the multivariate analyses we have carried out do not suggest a clear and significant effect played by the length of the recall period, but we might still think that other factors affecting the complexity of the reporting task play a major role in explaining lower transition chances in the retrospective survey.

Introduction

To analyse the mobility of people on the labour market, longitudinal studies are an ideal framework. As Menard (2002) points out, "longitudinal research serves two primary purposes: to describe patterns of change and to establish the direction [...] and magnitude [...] of causal relationships". But longitudinal surveys can be conducted with different research designs.

Correctly measuring individual dynamics in labour market behaviour has become increasingly important as research and policy attention has become more focussed on the relationship between current employment opportunities and past experience (Paull, 2002).

But obtaining accurate and complete information about the past, so to show the evolving structure of the life course and to link earlier and later events within individual life, is not easy (Scott & Alwin, 1998).

This paper focuses on two distinct methods of collecting data suitable for life course analysis: the prospective and the retrospective one.

Data can be collected prospectively, repeatedly asking individuals for their current situation using the same questionnaire, through what is known as panel study.

Alternatively, longitudinal data can be obtained through a retrospective study, where respondents are asked to recall their behaviour over a specified prior period. While there exist surveys designed entirely retrospectively, also panel studies often rely on retrospective information for gathering data related to the period before the first interview. Furthermore, most of the major labour market panel surveys rely to some degree on recalled data for collecting information on labour market dynamics by asking respondents to retrospectively fill-in the gaps between interviews.

One approach that embeds periodic retrospective measurement within the prospective panel design is the life history calendar (LHC). It provides a good example of how retrospective measurement methods can be embedded in what is otherwise a prospective panel study assessing the cumulation of experiences. Furthermore, the use of the LHC to link events and transitions over time improves the quality of data.

This paper focuses on the extent to which longitudinal surveys are able to expose labour market dynamics and the underlying causal mechanisms and it is specifically concerned about data quality. Our reliance or confidence in the data is affected by the issues of reliability and validity. The first concept refers to the repeatability of measurement, to its consistency over multiple attempts to measure the same things using the same methods. Validity, instead, is concerned with assessing whether what actually is being measured corresponds to some external reality; it refers to the degree to which a study accurately reflects or assesses the specific concept that the researcher is attempting to measure. While reliability is concerned with the accuracy of the actual measuring instrument or procedure, validity is concerned with the study's success at measuring what the researchers set out to measure, asking how far the data correspond to an external reference or records which are thought to be largely error free. It should not be taken for granted that independent records are always error free: there are a range of conceptual problems which can lead to mismatches between recall and independent records. It is much easier to measure reliability than validity. Some methods of measuring reliability make a closer approximation to measuring validity than others, but strictly speaking measures of reliability do not provide measures of validity.

Were both the research designs free from measurement errors, they should describe the same mechanisms. But different factors affect the reliability of both. In particular, we want to assess if and how these different factors make data collected through different research designs unreliable.

Using the GLHS (German Life History Study) and the GSOEP (German Socio-Economic Panel) to exemplify respectively retrospective and prospective data, we want to check whether different kinds of research designs lead to the same or different observed patterns, paying special attention to the

impact of recall on measured labour market behaviour and highlighting how and to what degree the biases in the reported data may affect the estimation of models of labour market dynamics (Paull, 2002). We want to verify whether the picture drawn using prospective and retrospective data is substantially different, paying attention both to the levels and to the determinants.

Specifically, we are trying to assess to which extent retrospective and prospective data lead to the same or different conclusions with respect to basic career moves, i.e., entry into and exit from employment as well as shifts between jobs, and this rather descriptive question brings us to the explanatory question why differences may occur. Emphasis will be on the comparability of data and the extent to which they are able to measure labour and job mobility patterns over time will be discussed, keeping data quality as our central concern.

Data

The GLHS and the GSOEP are both large-scale multipurpose surveys which, in a different way, collect data on a wide variety of subjects, paying special attention to the change in patterns of education and training, career transitions, and processes of family formation, making it possible to investigate many economic and social characteristics of households and their members.

The GLHS is a retrospective study on individual life courses of people belonging to specific birth cohorts (retrospective cohort study) in which respondents were asked, at a specific moment, questions about their past, and so they had to reconstruct their life histories. The GSOEP, on the other hand, is a panel study, in which the same set of participants is interviewed annually, using the same questionnaires.

The GLHS, carried out since 1983 at the Max Planck Institute Berlin under the direction of Prof. Karl Ulrich Mayer, collected retrospective data for about 8,500 men and women from 20 selected birth cohorts in West Germany and more than 2,900 men and women from 13 selected birth cohorts in East Germany. It consists of a set of singular retrospective standardized interviews (face to face or telephone) with persons belonging to certain birth cohorts and drawn from representative samples. The data covers a comparatively long time frame allowing for analyses that go back a period of time in history (Buchholz & Grunow, 2003). The GLHS is one of the major bases for comparative studies of social conditions before, during, and after the division of Germany. We are using the aggregate dataset (Gesamtdatenbank)¹, which contains data from four different surveys².

The GLHS allows to collect a complete retrospective career history of the respondents, who were asked the monthly beginning and ending dates of each job they had ever held and of each episode of self employment they had ever experienced. They were also asked to identify the occupation, the branch of industry, the size of the firm, and the beginning and ending wages for each job and self-employment episode. Finally, respondents were asked to indicate whether job shifts occurred across firms or within the same firm.

The GSOEP is a longitudinal survey of private households and persons in the Federal Republic of Germany which has been running since 1984 with an initial sample of 5,921 household containing a total of 12,290 individual respondents. After 21 waves 3,724 of these households with 6,811

¹ Prof. Karl Ulrich Mayer and his collaborators kindly made this last version of the data available for our analysis.

² The life histories of a first sample of 2,172 respondents representative for the Federal Republic and West Berlin from the cohorts born 1929-1931, 1939-1941 and 1949-1951 were collected from 1981 and 1983 (LV-West I). In the years 1985-1987 other 1,412 men and women belonging to the cohort 1919-1921 were interviewed (LV-West II). 2,008 respondents from the birth cohorts 1954-1956 and 1959-1961 took part in the 1988/1989 GLHS survey (LV-West III). Furthermore, a follow-up survey (LV-West 64/71), in which 2909 West German respondents belonging to the birth cohorts 1964 and 1971 were interviewed, was conducted in 1998-99.

respondents were still participating in 2004³. In the GSOEP a set of core questions on labour market and occupational dynamics is asked yearly. In order to measure change and stability across time, the SOEP questions are targeted at different dimensions of time using different measurements of time (information at a given point of time, periodical information, calendar information, life history information). Specifically, information on the labour market status is collected in different sources: current status information, short term (one year) retrospective information, and long term retrospective information. We are making use of the first two kinds of information. At each interview the individual is asked if s/he is currently engaged in paid employment and also specifically in which position/occupation⁴. Using the information on current activity status allows us to construct career patterns but based on annual information and therefore lacking detail on what happens within the year.

The GSOEP also collects data on transitions and events which occurred between two waves. Namely, it introduces a monthly calendar which helps reconstructing what happened in each month in the previous calendar year: Respondents are asked to report labour market states retrospectively for each month, from January to December, of the preceding year. Using the monthly calendar we are able to construct career patterns for a part of the life-course on a monthly basis, although they are also affected by some memory bias because they are based on information recalled for the previous year⁵. In addition, questions are asked concerning the date and the kind of job changes occurred from January 1st the year before the interview⁶, allowing us to split employment spells into job spells and to specify the kind of change intervened.

Although the individual information is available for all job spells, job specific information is only available if the respondent was in that job at the moment of interview. Therefore, specific description of moves on the labour market reflects changes in position held at each interview time, while short job spells tend to be underrepresented. Despite this, we do have information about changes in the working arrangements as far as it concerns the more general distinction between part time and full time and type of job change, so we can at least recognize whether a short job that do not coincide with any interview is in the same firm as the previous one or involves a change of employer (in case of multiple job changes from January the year before the interview only the last one is recorded).

³ We are using the 1984-2005 CD-ROM version of the GSOEP and, while when we are analyzing the reliability of the calendar data comparing them with current information we are using the whole sample, when testing long term recall, for reasons of comparison with the GLHS, we have selected only respondents living in West Germany.

⁴ The question for current status is "Are you currently engaged in paid employment? Which of the following applies best to your status?" providing (in the last wave) the options: full-time employed, part-time employed, in occupational / professional education or retraining, marginally ('geringfügig') or irregularly employed, approaching retirement part-time employment with zero working hours, doing your compulsory military service, doing community service as substitute for compulsory military service ('Zivildienst') / voluntary social year, not employed. Pensioners with a job contract are considered employed. Furthermore, there is a question asking specifically "What is your current position/occupation?".

⁵ Respondents are shown a grid having as columns all the months in the previous calendar year and as rows the following states (in 2005): full-time employed, part-time employed, I held a mini-job, in first-time company raining/apprenticeship, in further training/retraining, further professional education, registered as unemployed, in retirement or early retirement 'Vorruhestand', on maternity leave/ child rearing leave/ 'Elternzeit', in school/ at university or 'Fachschule', military service,/ Reserve Duty Training Exercise/ Community Service/ Voluntary Social Year, housewife/houseman, other. They are asked to check all the months in which they had each of the mentioned states.

⁶ The questions we are using are, in the last wave (2005): "Did you change your job or start a new job after December 31, 2003? When did you start your current position (month in 2004; month in 2005) What type of employment change was that?". These questions were asked only to old respondents before 1994, so that it is not possible to reconstruct job changes in the first year for which we have calendar information for respondents who had their first interview before that year.

Factors affecting prospective and retrospective data quality

Much in the literature has been written comparing the prospective and the retrospective survey designs, tracing advantages and disadvantages of both. As Scott and Alwin (1998) point out, the implied polarization between retrospective and prospective methods is to be rejected. "It is important not to make too rigid a division between retrospective and prospective designs because, in practice, nearly all social surveys elicit some information about the distant past {...}. Moreover, prospective studies are almost inevitably in the business of collecting retrospective data" (Scott & Alwin, 1998:106). Several aspects have to be taken into account when evaluating prospective and retrospective data: practical considerations of time and money, substantive concerns, as well as issues of measurement quality have to be weighted to judge which data better suit our research purposes.

The main differences between these two measurement strategies involve elements that have implications for the quality of data (Scott & Alwin, 1998) and this is what we are especially interested in.

A prospective design allows data to be collected concurrently with the events in question but major costs as well as concerns raised by panel attrition and conditioning have also to be considered.

Selective survival, refusal, loss of subjects through mobility, may make the reinterviewed subjects quite different from the original sample, undermining population representativeness and leading to severe problems for making reliable population estimates over time, since non response does not occur at random.

Another issue concerns whether interviewing the same people over time affects the essence and quality of respondents' answers (known as "Hawthorne effect"). Panel conditioning may affect the quality of data reported. Yet, on the positive side, panel participation may improve the quality of data, because, over time, respondents can be "trained" to keep records and documents that will help them report past events and changes with more accuracy.

Moreover, the multi measurement occasions in panel studies introduce problems related to changing measurement and seam effects, such that people memories in terms of dating of transitions seem to be influenced by the reference period of the measurement instrument.

The multi-measurement occasion information makes it quite complicated to re-arrange the discontinuous information about statuses, events and transitions given in the several waves into a continuous flow of behaviour (Solga, 2001). Since time elapses between two interviews, there is no guarantee that the meaning of a measure will remain unchanged as well as the quality of answers, and this might generate mobility in place of stability. This leads to the risk of artificial mobility, problem which does not arise in a single measurement design and which shows itself in relatively small changes in responses to a question for adjacent months within a reference period and more abrupt changes for adjacent months across reference periods ("seam effects").

Retrospective data surely have more practical advantages: lower money and time constraints and a longer time horizon. Despite the absence of multiple measurement occasions eliminates some kinds of bias ensuing from repeated contact and reinterviews, they suffer from measurement errors, which come from two main sources. Recall data face a major problem of substantial selection bias due to prior mortality, since estimates are only representative for the survivors. Furthermore, when individuals are asked to retrieve events, situations and attitudes which occurred in the past, the possibility of inaccurate reporting may occur, either from unintentional memory lapse or a conscious misrepresentation of the past (Babbie, 1973; Powers et al., 1978). Individuals may omit an event altogether or they may misclassify it or when it occurred.

It is very important to assess the extent to which retrospective data are affected by memory bias, since inaccurate long term recall could make retrospective surveys a very poor substitute for panel data but the quality of retrospective data is not easy to assess, especially considering that recent experiences and events may bias the recollections people make about their earlier experiences and

people may revise their recollections of the past in the light of subsequent outcomes and present experiences.

During the last few decades there have been serious questions raised about survey methodology. There has been a general concern to know whether retrospective data suffer from increased bias or measurement error compared with questions asked about current status (Dex, 1991) and many studies have researched the reliability of retrospective data.

If using the GLHS we rely on long term retrospective information, where the time distance between the event reconstructed and the interview ranges from the same year to 52 years before, the use of information coming from a LHC in the GSOEP also leads us to rely on retrospective information, although the length of the recall period is much shorter, since respondents are asked to recall only the previous calendar year. In general, the longer the recall period, the greater the concerns about the reliability of retrospective data, although there are important differences by subject matter and salience. (Scott & Alwin, 1998). Thus, prospective designs seem likely to yield more reliable dating of life transitions and changes because the length of recall between waves is usually relatively short. Yet, biases in the recall might occur, and it is possible to analyse retrospective short-term errors as well (Jürges, 2005 for GSOEP; Paull, 2002 for BHPS).

Several factors have been shown to influence the degree of recall error: the length of recall period, the complexity of the reporting task, the salience, pleasantness and social desiderability of the events being recalled. Studies in cognitive psychology have identified three major factors which are responsible for the accuracy of recall: interference, salience of the event to be reported and length of recall period (Eisenhower et al., 1991). Interference refers to the fact that it is easier to remember rare events ceteris paribus than frequent events, since the occurrence of many similar or related events reduce the memorability of each individual event. Salience instead refers to how important an event is to the respondent. As memory decays over time, the probability of accurate recall generally decreases with the length of the recall period. However, shortening recall periods does not always reduce recall bias since if the recall period is too short, over-reporting may arise from the tendency of respondents to "telescope" rare events into that period.

Different studies have researched the importance of the length of the recall period for recall accuracy (Powers at al., 1978; Mathiowetz & Duncan, 1988; Paull, 2002; Auriat, 1993) finding differences among different groups and according to the subject matter to be recalled. Different studies show sometimes contrasting results. Although some studies (Bradburn, Rips and Shewell, 1987) argue that the longer ago an event is, the less reliable the answers concerning events are, others, in contrast with the previous believe that the retention curve of recall and autobiographical memory over time follows an exponential decay function, have shown that very recent events are recalled best but then the decay in memory does not continue over time but stays stable for a long period (Hauser and Featherman, 1977 ??). In accordance with these findings, Matthiowetz and Duncan (1988) found the recall period to have an insignificant effect on unemployment recall error once the salience and difficulty of the task had been controlled. The question whether recall errors are a function of time distance between the event and the interview is a very important one, since if this is the case, it would suggest that retrospective measurement for small intervals like panel interviewing times are reliable but long recall periods might not be. But there is no easy answer regarding the length of time over which events can be recalled with some reliability (Scott & Alwin, 1998).

Mayer (2006) argues that "time distance to the event is related to recall reliability in a curvilinear fashion. Recall is best for very short time distances, than worsens, but stays very stable as time distance goes on. This result could support retrospective studies with a long time frame into the past. In contrast, errors in dating events seem to increase with time distance in a linear fashion. Experimental data shows that early childhood events are badly recalled, events for the age between 15 and 25 are well-remembered and little is recalled for middle age. This suggests that recall is not only a function of time, but also of the probability of occurrence of given events". These

conclusions rely heavily on Reimer's (2005) findings, according to which the likelihood of error does not seem to be related to time: the consistent lack of a time to interview distance factor shows that no simple memory laws apply.

Analysing data of a German 1987 three wave panel life history study in order to investigate the reliability and consistency of individual occupational career data, De Graaf et al. (1989) show that retrospective variables are no less reliable than other survey data (income being a noticeable exception). Few sources are found to have a significant impact on the unreliability and inconsistencies encountered. The amount of job spells is found to have a clear positive effect on the inconsistency of the number of spells recalled while the time variable indicating how long ago a certain event took place is found not to have an effect.

Hauser and Feathermann (1977) conclude that reports on jobs that respondents had five years ago do not exhibit greater unreliability than reports on jobs of few months ago, and "reject the implication that occupation information is subject to unusual distortion and decay as a function of time".

Furthermore, the psychology of autobiographical memory has rapidly advanced and the emergence of cognitive survey psychology has made major progress, so to give more insights in the mechanisms at play, even though findings from the psychology of autobiographical memory might not be easily applied to standardized retrospective surveys (Mayer, 2007).

In this paper, we are studying retrospective bias regarding labour market dynamics, looking at both the short term and the long term recall bias. Namely, we are first using the German Socio-Economic Panel to test the reliability of (calendar) one-year retrospective information comparing it with current status information, and then we try to gauge the extent of the long term retrospective error, comparing the biographical long term retrospective information gathered in the German Life History Study with the calendar information from the GSOEP.

Analyses

Given the information available in the datasets we are using, we try to answer our research questions carrying out different kinds of comparisons: first, we test the reliability of short term retrospective information comparing the calendar (one-year retrospective) data with the information gathered at the moment of each interview in the GSOEP.

Then, we proceed comparing the information collected through the GSOEP yearly survey with that gathered through the life history study.

Testing short-term retrospective information

The SOEP monthly calendars are mostly used in labour market studies to generate spell data for event history analysis. Due to the rather short recall period (one year on average) and the implementation of the LHC as a way to improve memory, there should not be much scope for retrospective bias, also considered that respondents are asked for factual information (Juerges, 2005). Yet, their reliability has to be assessed, since measurement errors may potentially bias the results.

Reliability tests comparing calendar and current data

To test the reliability of short term retrospective information in the GSOEP it is possible to compare current data and calendar data collected one year later. At each interview, respondents are asked to report, besides their current employment status, their employment condition for each month from January to December of the previous calendar year. Thus, an overlap is generated between the information given at the moment of interview and that given retrospectively in the calendar one year later with reference to the month of interview in the previous year.

Previous studies (Jürges, 2005) have compared current and one-year retrospective data on unemployment in the GSOEP, finding that calendars suffer from a sizeable amount of underreporting. Jürges (2005) shows that about 13% of all unemployment spells are not reported one year later, and another 7% are misreported. Furthermore, individuals with weak labour force attachment have been proved to have the greatest propensity to underreport unemployment retrospectively. Thus, research based on such data, should always try to account for the shortcomings they can present, since retrospective errors are unlikely to be random. Recent comparisons (Biewen and Wilke, 2005) of transition rates based on SOEP calendars suggest that they underestimate the length of unemployment by a considerable margin, especially for women in western Germany.

We extend the analysis including all the labour market states. We try to look at the consistency of current information and one-year retrospective information more in general, looking at the declaration of the employment status, instead of focussing exclusively on unemployment. For doing this, we compare the information provided in the calendar (the monthly employment status) with the answer to the question explicitly asking for the current employment status at each interview⁷.

We find that in nearly 97% of the cases in which we have data coming from both the sources, respondents give consistent information, if we consider the general distinction between employment and non employment status.

Mismatches in both the directions (failing to report having been employed during a particular month in the previous year despite reporting being employed when interviewed in the month itself, and the other way around, reporting having been employed during a particular month despite reporting of being not employed when interviewed in the month itself) seem to be slightly more common among women, who specifically tend to underestimate employment when asked retrospectively.

	<i>v</i>		
	Men	Women	Total
No error	96,87	96,36	96,61
False negative ¹	1,64	1,77	1,71
False positive ²	1,48	1,87	1,68

Tab. 1 Probability of incorrect report in the calendar information by sex (column percentages)

¹ The respondent states at the moment of interview that s/he is employed while in the calendar s/he declares s/he was not.

² The respondent states at the moment of interview that s/he is not employed while in the calendar s/he declares s/he was.

To check whether errors are evenly distributed among respondents, we use logistic regression analysis, estimating the probability of inconsistency between the two sources of information. A set of variables related both to socio-demographic characteristics of the respondents as well as to the survey design are included as independent variables.

⁷ The exact question we have used to retrieve this information was "Are you currently employed? Which of the following applies to you?". The possible options for an answer have been recoded into two categories: employed (full-time employed, full-time short work hours, (regularly) part-time employed, part-time short working hours, minimally or irregularly employed, disabled employment) and non employment (all the other options).

	Dela
Female	0.176**
Age (at interview)	-0.005**
Age squared	0.000**
Number of job spells	0.098**
Education (CASMIN)	
labc (ref.)	
2 <i>ab</i>	-0.132**
2c	0.285**
3ab	-0.328**
work experience	-0.007**
spell length	-0.025**
time distance between data	0.049**
Number of interviews	0.074**
Year	-0.047**
Constant	91.814**
R2	0.161
Chi2	10884.8
df	14
Ν	232255
* = <0.05 ** = <0.01	

Tab. 2 Logistic regression estimates of incorrectly recalling employment status

Date

* p<0.05, ** p<0.01

Women's retrospective reports seem to be less accurate than men's. We might think this to be related to the effect of salience of employment, since salience has been proved to be inversely related to retrospectively misreporting (Dex, 1991; Jürges, 2005; Matthiowez and Duncan, 1988; Linton 1982 – in Neisser, Loftus 1979).

We might think that there is a positive relation between age and difficulty in recalling, so that older respondents are more likely to underreport, due to deteriorating memory. To test this hypothesis, we include age at interview in our regression model but, as Matthiowez and Duncan (1988), we do not find support for this relation, since age does not show a positive effect on the odds of misreport.

Education does not seem to play a significant role in one direction.

Trying to account for salience, we introduce the number of months the respondent has been employed, as an indicator for labour force attachment⁸. As expected, this variable appears inversely related to misreport, being respondents more attached to employment also more accurate in reporting employment spells.

We may think that the length of the spell also plays a role in the facility to recall: If on average the spells length is greater, than it should be easier to recall them correctly. Longer spells could be reported more precisely then shorter because they are more salient, and because the longer the spell, the smaller the probability that respondent report the wrong month. But the available information on spell duration is in some sense endogenous because it can only be derived from the employment calendars.

The difficulty of the reporting task may also be considered an additional correlate of response error: interference theory (Crowder, 1976) suggests that the probability of reporting a particular event is inversely related to the number of related events an individual experiences. In accordance with our expectations, we do find a positive effect of the number of jobs held⁹ on the probability to misreport:

⁸ Since this variable is computed only on the months the respondent is observed prospectively, and not on the whole career, it is not highly correlated with age.

⁹ This variable indicates the number of jobs a respondent held in his/her career and it is standardized for age by calculating the degree to which a respondent's number of job spells deviates from the average in his/her five year age group, so to correct for the high correlation with age.

The more spells a respondent has, the more likely it seems that s/he misreports his/her employment status when asked one year later.

Furthermore, we have included a variable indicating the distance¹⁰ to the following interview, which tells us the time distance between the information which is compared. It appears that, the farther away the recalled information is from the current information, the greater the difficulty in recalling/the probability to misreport.

The "calendar year" variable, introduced to assess the development of misreports during time, shows a decrease in their number from 1983 to 2005, suggesting that data quality increases in the course of a panel survey (Jürges, 2005).

We might also think that the mismatch is not due to the fact that a respondent forgets a spell altogether, but simply that s/he misclassifies it, declaring it to start in a different month, leading to a contrast with the current data. Thus, we can identify different degrees of failures: allowing for a less rigorous definition of error, we can identify cases in which respondents fail to report the same employment status in the month of interview, but within the previous and the following month the same status can be recognised. We have defined errors which are still occurring under this more permissive definition as "hard errors": in this way, mismatches are reduced by approximately 30%¹¹.

Tab. 3 Probability of incorrect report in the calendar information, all errors and hard errors (column percentages).

	Error	Hard error
No error	96,61	97,63
False negative	1,71	1,16
False positive	1,68	1,22

The reduction in the number of errors may be read as an indication that the calendar information leads to problem in misdating, while it occurs less often that spells are completely forgotten. Hence, we may argue that the use of calendar information should not affect much the levels of the transition probabilities, although errors may occur in the exact dating of the spells. Yet, we can never be sure that what is recalled (under the more permissive definition) is indeed the same spell which has been misdated or is instead a different spell, while one was completely forgotten.

Seam effects

Another source of measurement error in panel data, as previously explained, is related to the multimeasurement occasions, which lead to the so-called "seam effects".

In the GSOEP, on the one hand, the use of a LHC (Life History Calendar) to collect monthly information related to the previous year yields exhaustive information to generate spell data for event history analyses but, on the other hand, we still have a multi-measurement design, and, besides the well-known problem of panel attrition, we do face problems related to "seam effects", which lead to spurious transitions between calendars collected in subsequent years. In the GSOEP calendar information, a very high proportion of all employment spells apparently begins in January or ends in December of each year, leading to much higher transition probabilities¹² in December¹³.

¹⁰ In case of temporary drop out it turns out to be problematic to compute this variable. We think these cases should be dropped from the analysis, since they are based on a different kind of reconstruction. Up to now, we have simply dropped cases in which the distance between interviews exceeds 23 months (which is the maximum distance if two interviews are carried out in two consecutive years). Yet, in this way we still have some cases based on reconstruction of temporary drop out (because for example if the drop out was only for one year the distance between the interview before and the one after the drop out may be less than 23 months).

¹¹ Errors are reduced from 3,39% to 2,38%.

¹² In this instance, with transition probabilities, we are referring to exit from and entry into employment, which are the events which define the beginning and the end of an employment spell as resulting from the declaration of the

	EXIL	Entry
January	6,12	6,39
February	6,42	8,53
March	8,97	9,44
April	6,34	7,04
May	5,91	6,49
June	8,46	8,56
July	7,44	8,85
August	8,19	9,34
September	8,82	7,99
October	7,08	5,69
November	7,33	4,15
December	18,94	17,54

Tab. 4 Exit and entry probabilities by calendar month (column percentages)

Comparisons with data from the Federal Labour Office show that this concentration cannot be explained by cyclical factors alone (Kraus and Steiner, 1998) and it is instead due to seam effects: In the GSOEP calendar, information for December and January of the following year comes from two different interviews, so that the likelihood of a difference between these two months is higher, due to a multi-measurement problem.

If, on the one hand, seam effects tend to overestimate transitions, we might also think that, on the other hand, prospective generated longitudinal data may underestimate change, because they are not able to report changes between interviews. The use of a LHC in the GSOEP allows to record monthly changes in the employment status, partly solving the issue (but, as we have shown, generating other problems). The problem still persists for job to job changes, since these are not reported in the LHC, given that they do not imply a change in the employment status (but only in the job), and since at most one change a year can be reported.

Double changes

Besides some tests concerning the reliability of the short term retrospective data, we carried out additional analyses to check how consistent retrospective information are among themselves, using questions concerning job changes, which often cover the same period of time twice ¹⁴.

Failing to drop changes reported twice would lead to an overestimation of job changes of about 9%; in 15,6% of the cases we find a different type of change declaration, which can be corrected choosing the report closer to the event. However, there might still be an overestimation of job changes due to the fact that while when exactly the same month but a different kind of change is declared we can choose for the information closer to the interview date, we can still have an overestimation of events in case there is a misdating, and a different job change date is declared in two interviews for the same change. And we may think that this is more likely to occur the later in the year the interview occurs.

employment status given at different interview time points. Job to job mobility instead does not imply a change in the employment status, but simply of the job, and it is not affected by this measurement problem, since the information concerning job to job shifts does not come from the comparison of statuses declared in different interviews, but instead from a question explicitly asking the date of a job change.

¹³ This is due to the fact that we have defined an event as occurring in the last month of the "old" employment status, so that exit is reported in the last month of employment and re-entry in the last month of non-employment, before reentering. Defining an event in the first month of a new spell, would lead to an overestimation of the transition probabilities in January. However, the cause of the problem is always the difference in the status declaration between December and January of the following year.

¹⁴ See appendix for a detailed explanation concerning how this might occurr and how we have dealt with these cases.

Tab 5 Percentage	of double reports, equa	l and	contrasting	information
Single report	91.00%			

Double report	9,00%
equal report	15,60%
contrasting report	84,40%

Testing the reliability of long term retrospective information

After having explored the reliability of the one year retrospective questions implemented in the GSOEP yearly questionnaire, we come to our main interest. Are prospective and retrospective designs revealing the same mechanisms in the labour market? To which extent are measurement errors in both the research designs affecting the description of the true underlying mechanisms in the labour market? Are there differences in the aggregate descriptive statistics? And in the determinants?

To answer these research questions, in each of the studies we have generated a person-month dataset in which, for each respondent, we have reconstructed multiple job spells.

We define an employment spell as a period in which a person is employed and employment spells are separated by non employment periods. A job spell, instead, is defined as any period in which a person was employed in a specific position with specific characteristics; a job spell can also follow another job spell, where the characteristics of the job were different.

All the types of job changes are recorded: entry and re-entry into the labour market, exit from the labour market, job to job mobility. We have then combined the two person-month files in one dataset, generating a dummy variable indicating from which survey the data came from (see Appendix for details concerning the procedure we have implemented to generate our person-month file).

Descriptive analysis

In the first part of our analysis, we look at the various transition chances in a bivariate fashion.



Fig. 1a Exit probabilities by age, GSOEP and GLHS, males and females

Employment exit chances for men follow a U-shaped pattern, being higher at the lowest and the highest ages.

While men are less likely to exit the labour market in their central ages, women have less exit chances than men at the beginning and at the end of their career, so that their curve appears to be slightly flatter. The same pattern emerges from both the datasets, even though from the analysis conducted on the life history data transition probabilities appear much lower for both sexes.

We might assume differences in the transition levels between prospective and retrospective studies to be different by time distance to the interview, so that the further ago the period reconstructed through the retrospective survey is, the greater the difference with the prospective one, while the closer the period is to the interview date, the narrower the differences with the prospective data, due to the assumption that the length of the recall period is important for recall accuracy and events occurring further in the past are less well recalled than more recent events. Thus, we have plotted the transition probabilities by time to the interview for the GLHS.

Fig. 1b Exit probabilities by age, GSOEP and GLHS by time distance to interview, males



Fig. 1c Exit probabilities by age, GSOEP and GLHS by time distance to interview, females



From the graphs (Fig. 1b and 1c), it does not seem that time distance to the interview plays an effect. We might expect curves for the period closer to the interview to be closer to the GSOEP one, whereas the ones referring to periods further away from the interview to be more distant. From the graphs, instead, we do not see much difference by distance to the interview. In the following multivariate analysis, we are trying to clarify whether there is indeed no effect played by time to interview or there is instead a compositional effect, which averages it out.



Fig. 2 Re-entry probabilities by age, GSOEP and GLHS, males and females

Re-entry chances gradually decline with age: the curve is steeper for males, being their re-entry chances higher at lower ages, while in the central ages re-entry probabilities are higher for females. Both the datasets give the same results but again using the retrospective survey transition probability is lower, although the path is the same.

Fig. 3 Job to job mobility by age, GSOEP and GLHS, males and females



Job moves are more likely at the beginning of the career, but for women they tend to decline earlier (right after 20s) while for men they reach the peak around 30s. While in the GLHS men show higher chances of job to job moves until the very late stage of the career, in the GSOEP transition chances are higher for women already from their 40s.

From this graph (Fig. 3), we can see that in the GSOEP differences between men and women are less pronounced than in the GLHS and differences between the two surveys are more marked for women than for men. We may think that this is strictly related to the definition of job spell in the two studies: In the GLHS the definition seems to be much wider, while the GSOEP classifies as job changes only specific kinds of changes (namely a change of employer or a change of position within firm). Although we have tried to make the definitions as comparable as possible, redefining a job change so to include also changes in working hours in the GSOEP, still the definition in the GLHS remains much broader. This could be the main cause of the big differences we encounter,

and of the sometimes strange results, especially for males, which probably do face more often the kinds of job changes which are not considered in the GSOEP, while it is possible that instead for females changes in working hours capture great part of job to job moves.

Differences in the levels of transition probabilities may be explained by memory bias in the retrospective data, as people might recall less well events happened in the further past. But actually our first analyses do not show a clear effect of the length of the recall period. Still, the mechanisms of reconstruction at play in the life history study are quite different from those in the panel one, a main difference lying in the length of the period to be recalled¹⁵.

Due both to the different research designs and to the time when the surveys were carried out, the birth cohorts, age and period composition of the two datasets differs quite widely¹⁶. We do believe that age, period and cohort effects could play quite some role in the comparison of our findings, and in the following multivariate analysis we are trying to take it into account, although the linear relation among these three factors make it difficult to assess the role played by each of them.

From the more substantive point of view, the trend observed seems to confirm the high protection of the core of employment in Germany.

Multivariate analysis

We have then proceeded carrying out some multivariate analyses, estimating regression models in which transition rates are the dependent variable. Event history analysis is used to analyse transition rates between different employment statuses.

Separate models for men and women (aged 18-64) including time constant and time varying covariates are estimated, in which the dependent variable is the instantaneous rate of change from one state, j, to another state, k, defined as

$$r_{jk}(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} P_{jk}(t \le T < t + \Delta t \mid t \le T) , j \ne k$$

where r_{ij} is the instantaneous probability of having an event in a time interval $(t, t+\Delta t)$, conditional on having no event until *t*.

Logistic regression is used to analyse transition rates, the dependent variable resulting as the odds of an event occurring given that it has not occurred in the month before.

The models estimated from the data at hand cover the following types of transitions and changes:

I. transitions out-of-employment (exit);

II. transitions into employment (re-entry)¹⁷;

III. job to job changes (i.e. any job move which implies the start of a new job spell, given that the respondent was employed before).

¹⁵ We might think that this could be a factor explaining differences but we cannot test for the effect of the length of the period to be recalled because in the GLHS the reconstruction is a minimum of 11 years. The shorter reconstructed period is too long to check whether in case of a shorter period to be recalled the differences between the two designs are smaller.

¹⁶ For example, old age people in the GLHS exclusively belong to the earlier birth cohorts, while old age respondents in the GSOEP belong also to much younger birth cohorts.

¹⁷ Entry into first job is not considered.

After having selected the period for which we have both prospective and retrospective data and having dropped left censored cases¹⁸ we apply discrete-time models for multilevel¹⁹ event history data.

We first estimate a joint model²⁰ for transitions out of employment and into employment²¹. In this way we allow for correlation across transitions in unobserved individual risk factor. By jointly modelling transitions from employment and non-employment, it is possible to test explicitly for state-dependent covariate effects. A joint model also allows for residual correlation in individual transition rates across states, which might arise because of unobserved factors that affect more than one type of transition.

The model is a generalised multilevel discrete-time event history model in which the response variable is the binary indicator of event occurrence.

The model estimates the chances of leaving employment, given that a respondent is employed, and the chances of re-entering employment, given that a respondent is not employed but has already experienced an employment spell earlier in his/her life (Steele et al. 2004).

Job to job mobility is instead modelled separately in a multilevel event history model estimating the chances of any job changes, given that the respondent is employed. Job to job mobility is defined as any job move which implies the start of a new job spell, given that the respondent was employed before.

Dependent variables

In the models we are estimating, the dependent variables are represented by events occurring on the labour market. The transitions we are analysing are defined as follows:

-exit from employment arises when a respondent is employed in the current month [*t*=employed] but not in the following [t+l=not employed]

-re-entry into employment is defined for subjects who have already experienced an employment spell and are currently not employed [t=not employed] but are employed in the following month [t+l=employed]

-job to job moves originate when a new job spell starts, given that the respondent was employed before. Hence, the switch from one job to the other occurs without any non-employment spells in between.

Independent variables

Since our main interest lays in the comparison of the two different research designs, an important independent variable is the one indicating the data source. When pooling the two datasets, we have generated a variable indicating whether the information comes from the GSOEP or from the GLHS. Moreover, we have created a variable indicating both the survey design and the time distance to the

¹⁸ We are using a more conservative approach, i.e. we limit the data to the period for which we have both retrospective data and prospective panel data. Left censored cases have to be dropped when carrying out event history analysis, because time in the spell is unknown and might consequently bias the duration analysis.

¹⁹ A multilevel model is used to allow for the hierarchical structure that arises from having repeated episodes (of employment or non-employment) nested within individuals. When an event may occur more than once over an individual's lifetime, the durations between events may be correlated due to the presence of unobserved individual-level factors. Repeated events are usually handled by including individual-specific random effects in an event history model, leading to a multilevel model. Longitudinal data can indeed be viewed as multilevel data, with repeated measurements nested within individuals. This leads to a two-level model, with the series of repeated measures at the lowest level, and individual persons at the highest level (Hox, 2002). The importance of the use of a hierarchical structure is related to the fact that many types of events can occur more than once to an individual, and we cannot assume that the durations of episodes from the same individual are independent. There may be unobserved individual-specific factors which affect the hazard of an event for all episodes. Some individuals might be more at risk of experiencing an event than others, and it is unlikely that the reasons for this variability in the hazard will be captured by covariates. The presence of unobserved individual-specific risk factors leads to unobserved heterogeneity in the hazard.

²¹ We are modelling only re-entry, and not first entry into employment.

interview for the GLHS. In fact, we are comparing data coming from the GSOEP and data coming from the GLHS, but while in the first case the reconstruction always concerns the previous calendar year, in the second we have large variation in the distance between the period reconstructed and the time of the interview. Thus, we have created a categorical variable indicating whether the information comes from the GSOEP or from the GLHS, distinguishing this last in 4 categories according to the time distance to the interview. Hence, we distinguish five categories (one for GSOEP and four for GLHS according to the time distance to the interview) and code them with cumulative contrasts, so that the effect of each category has to be compared with the previous one (instead of with the reference category).

As we have previously explained, there are strong a priori reasons for the believe that microeconometric duration models estimated on the GSOEP data may yield severely biased estimates due to heaping effects. In our models, we try to take heaping into account by including a time-varying dummy variable for this month (in our case we introduce a "December" dummy, since events are defined in the last month of a spell) in the set of regressors (Huijer and Schneider 1989, Huijer Loewenbein and Schneider 1990 and Hunt 1995 have done the same). Kraus and Steiner (1998) proved that accounting for the disproportionate number of spells beginning in January or ending in December in the GSOEP by including dummy variables for these months seems to be a reasonable way to deal with the problem. Furthermore, the way the heaping mechanism is modelled has little effect on the estimated parameters in duration models.

Since this type of measurement error occurs specifically in the GSOEP, the dummy variable introduced to control for seam effects indicates only records coming from the GSOEP.

While we control for seam effects when modelling exit and re-entry, we do not need to account for seam effects in the analysis of job to job moves instead, since in this case the information is not reconstructed merging monthly calendar data coming from different interviews, but through an explicit question asking the date of a job change from January of the previous year up until the moment of interview, so that no heapings between December and January should be generated by a multimeasurement design.

We also calculate and include into the models time into employment (when estimating exit chances) and non employment duration (when estimating re-entry chances), coded in a categorical variable with four categories (<13 months –reference-, 13-24 months, 25-48 months, >48 months). When estimating job to job mobility instead, we include a categorical variable indicating job tenure, which is defined as the number of months in the current job spell²².

We include the number of jobs a respondent held in his/her career as a time varying covariate, standardizing it for age by calculating the degree to which a respondent's number of job spells deviates from the average in his/her five year age group, so to correct for the high correlation with age (Kalmijn & Luijkx, 2006).

Education is measured by the CASMIN classification (Brauns et al., 2003), which distinguishes between the level of education and general vocational tracks within different levels. We have recoded the CASMIN merging categories 1a, 1b and 1c into primary education, 2a and 2b into lower secondary, 2c voc and 2c gen into upper secondary, 3a and 3b into tertiary. The resulting categorical variable is included as a time varying covariate, indicating the highest level of education at each point in time.

 $^{^{22}}$ Both the variables indicating the duration of the current spell (either employment or non employment) and job tenure are computed taking into account the information given in the biographical questionnaire, so to obtain the correct duration variable also fur the spell/job hold at the time of the first interview, and avoiding problems related to left-censoring.

In the GLHS, we reconstructed this variable combining the information coming from a variable indicating the highest level of schooling reached at the time of interview and another variable indicating any vocational training degree (Berufausbildungabschluss) reached²³.

In the GSOEP, it is derived from a question asking the highest level of education at the moment of interview. Information from the different waves is then merged to the person month file and extended to all the months following an interview up until a different educational level is recorded and to the months before the interview if the subject was not in education.

Information on marital status and children is also included in time varying form.

The life-course survey collects retrospective information on marital histories during the entire life course. For each marriage that respondents experienced in their life, they have reported the beginning and end dates, as well as additional information on the partner. Information related to children is contained in a special file, which allows to link any respondent to his/her children. The date of birth of each child is reported as well, allowing to reconstruct their age in each month. Starting from this information we have generated a variable indicating the number of children²⁴ for each respondent in each month, distinguishing between children aged 0-5 and 6-16.

In the GSOEP the monthly marital status is recorded in a monthly spell file. Information is available for each respondent since January 1st of the calendar year preceding a person's first SOEP interview²⁵. Starting from this spell file we construct a person-month file²⁶, indicating the marital status of each respondent in each month and we then merge this information with the person month file created starting from the job spells file²⁷. Also in the GSOEP we generate a variable indicating the number of children aged 0-5 and 6-16, thanks to the presence of yearly files linking each child to his mother and her partner²⁸.

Following Kalmijn and Luijkx (2006), we combine the information coming from the generated variables indicating marital status and presence of children to indicate the stage in the life course; we distinguish the combination of marital status and child status in five²⁹ mutually exclusive groups and code these with cumulative contrasts:

- 1. single
- 2. married without children
- 3. young children (any children aged 0 to 5, independent of the marital status)
- 4. older children (no children under 6 but children under 16, independent of the marital status)

²³While the variable indicating the vocational training degree varies during time for each individuals, the variable related to the school level indicates the highest school concluded at the time of interview, so that is is not time-varying. In the aggregate data we do not have the time varying information related to the school, but only information related to the school level that respondents are attending at the moment of interview, which doesn't necessarily imply that that level was concluded successfully getting a degree. Nevertheless, we think that the information related to the highest level of school is precise enough, as the very great majority of the respondents reaches the highest level of education before entering the labour market.

²⁴ We could at least correct this variable using the information related to children death, available in the aggregate dataset. Information about other reasons leading to leave the parents household are not directly available in the data.

²⁵ Since the precise question allowing to reconstruct the marital status was introduced in 1985, for some respondents this data are available only from January 1984, although they were first interviewed in 1984. For other it was possible to retrieve this information in other ways.

²⁶ Furthermore, we have corrected for some missing information using the data stored in a yearly spell file containing yearly information coming from the biography questionnaire.

²⁷ Some corrections were necessary, as the monthly spells overlap themselves in each case by one month. In the spell file this is due to the fact that a person can have two different family statuses during a month if a family status change occurred. Generating the person-month file, in order to avoid double person-month records, we considered the changing month as part of the new status, making the previous one end one month earlier. We didn't implement this attribution only in few cases of spells lasting just one month: in these cases we needed increment the beginning date of the following spell by one month.

²⁸ This implies that the definition of child identifies parenthood in a social, not necessarily in a strict biological way

²⁹ A sixth groups including respondents with no children but neither married nor single, together with other cases for which information on the life course stage is missing.

5. empty nest (no children under age 16 but any children under that age in any previous month)

Since we assume that employment behaviour depends on age, we include age in linear and, to examine the nonmonotonic relationship between an individual's age and employment, also in quadratic form³⁰, as control variables. We apply our analyses only to the age group 18-64.

A categorical variable indicating period is also included as control in our models. As mentioned, we have selected only the period for which we have information coming from both the surveys. This ranges from 1983 to 1999.

We broke down the period variable in 3 categories: 1983-1989, 1990-1994, 1995-1999 and the first is used as reference category.

In this way, we are controlling for differences in period and age composition. While period effects indicate the exceptional impact of certain time periods on life chances, and are the same for all respondents irrespective of their age, age effects assume that chronological age determines the outcomes studied. Still, we are not controlling for cohort effects, which would imply that life conditions and opportunity structures changed over time so that members of the same (birth) cohort are faced with the same conditions, but those conditions are different from those of other cohorts (Solga 2001). This is due to the well know APC problem, originating from the linear relation among age, period and cohort, which makes it not possible to include all the three effects in one regression model. At the operational level the three concepts of Age, Period and Cohort are linearly dependent. Once a person's score on two of these variables is known, the score on the third variable follows automatically. This implies that statistically it is impossible to identify the effects of all three variables in the same analysis.

Since we may assume that differences between GSOEP and GLHS are, at least partly, due to recall accuracy, and may think that respondents with different characteristics have different ability to recall, it is possible to introduce interactions between the survey dummy and factors which are thought to be related to memory bias. If an effect turns out to be stronger/weaker in the retrospective survey, we might think that it is because people with some characteristics recall worse/better, so that the effect becomes stronger/weaker due to the influence of the accuracy in recalling.

The literature on recall bias has proved that errors are not evenly distributed among respondents but affect mostly those respondents who have complex careers: the less conventional and more complex a life, the more likely there will be recall error (Reimer, 2004). Thus, we may think that the number of jobs held in the career affects recall accuracy, and estimates in the GLHS might be more affected. Hence, we introduce an interaction term between the variable indicating the number of spells held and the survey dummy.

We carry out our analysis in different steps. First, we want to gauge the whole "survey effect", which estimates the total difference between the two data sources. Then, we specify different categories in the life history study, according to the time distance to the interview, to assess the effect played by the length of the recall period.

Furthermore, since a common finding in the literature on recall bias (Dex, 1991; Paull, 2002; Duncan and Hill, 1985; Jacobs, 2002; Elias, 1997; Mathiowetz & Duncan, 1988; Reimer, 2004) is that shorter spells as well as unemployment spells are more likely to fail to be recalled retrospectively, we introduce some additional analyses to test whether difference between the two survey design are largely explained by this.

³⁰ The quadratic term for age is introduced in the model divided by ten.

Results

To analyse the survey effect, we proceed by steps. First, we estimate a model using a simple dummy variable indicating whether the data come from the GLHS or from the GSOEP. From table 6 (total rows) we can see that, for both men and women, the odds of exit and re-entry is lower in the life history study. For women, the odds of exit is about 25.7% lower (exp(-0.2857)-1), and for men about 44% lower. The odds of re-entry is instead approximately 29% lower for females and 35% lower for males.

Looking at job-to-job mobility instead, while we see a negative effect of the GLHS for women, we do not see it for males. We do not have a straightforward explanation for this, but we think that the different definition of job spells in the two studies might play a big role, perhaps making the data hardly comparable.

Previous findings in the literature corroborate the recall risk of non-standard episodes (like unemployment), which are markedly less well remembered; Reimer and Kuenster (2004: 22), comparing the GLHS with register data, show that short episodes are omitted or merged, when transitions are not very marked or short times of employment interruptions are ironed out, leading to simpler, more conventional and stable careers. Given these findings, we may assume that, retrospectively, respondents underreport short unemployment spells. Hence, we want to check whether the difference between the retrospective and the panel findings is indeed mostly caused by the underreport of short spells out of employment, or instead there is some error which goes beyond this. To test this hypothesis, we have tried to estimate the models explained above not considering short spells (less than 3 months) out of employment. Table 6 shows the results of a multilevel model estimated considering all the spells (total rows), and not considering short unemployment spells (no short rows) respectively. It is pretty evident that the second option produces coefficients which are much less strong and significant, indicating that the difference between the two surveys is in large part due to the difference in reporting short spells out of employment. Moreover, this may also explain differences in job to job mobility. Our original models on job to job mobility didn't show a clear negative effect of the retrospective survey, differently from what was happening in the estimation of exit and re-entry models. This may be seen as an overestimation of job to job moves in the retrospective study or as an underestimation in the prospective one. If, one the one hand, reasons for the last might be advanced (such as the fact that at most one change a year may be reported, respondents were not asked this question in their first year of interview before 1994, the information stems from a question looking retrospectively at what happened in the previous year, and the information on the date has to match with the employment status declared in the LHC), on the other hand, a very important factor of underestimation in the life history study might be that, retrospectively, respondents tend not to report short spells of unemployment. This might lead to an overestimation of job to job moves, because what is actually an exit followed by a re-entry is seen as a direct job to job change. If this is the case, employer changes should be mainly affected, since it is more likely that a non employment spell occurs in between two jobs with a different employer, while this is usually not the case for within firm changes. Indeed, the analysis of within firm job changes shows significant negative effect of the survey design for both men and women while the analysis of employer changes shows opposite results: while the survey design (indicating retrospective) is positive but not significant for women, the effect is significant for men. Differences between within firm and between firm mobility may be seen as a results of the overestimation of employer change in a retrospective design due to the underreport of short non employment spells.

Tab. 6 Logit coefficients for GLHS effect³¹.

		Women	Men
Exit	total	-0.2857**	-0.587**
	no short	-0.106*	-0.321**
Re-entry	total	-0.34**	-0.4373**
	no short	-0.224**	-0.117*
Job to job	total	-0.1741**	0.0706
	within	-0.998**	-0.415**
	employer	1.47	0.199**
	* 0.01		

* p<0.05, ** p<0.01

As we have extensively explained before, the two surveys have two quite different approaches to reconstruct social change, and a main difference is related to the length of the time period reconstructed. While the GSOEP asks year by year to reconstruct what happened the year before, the GLHS asks in one shot to reconstruct the whole career. Consequently, respondents have to recall events very close to the interview time but also events which occurred many years before. Thus, we want to control for these differences, paying attention to the effect of distance between the reconstructed event and the interview, trying to understand if it plays a major role in determining the differences between the two studies.

Hence, we substitute the dummy for survey design with a categorical variable indicating, besides the data source, the time distance to the interview for the GLHS records.

This variable doesn't show a significant effect of the time distance to the interview, nor on exit nor on re-entry nor on job to job mobility, and this holds for both men and women. Although we do see a survey effect, we cannot simply attribute it to the fact that in the life history study the length of the recall period is greater. We do not find a significant difference in the transition probabilities by time distance to the interview looking within the GLHS either (tables not shown).

³¹ These are the parameters resulting from a joint multilevel event history model for exit and re-entry and a separate model for job to job moves. All the independent variables mentioned are included in the model, but no interactions. These coefficient represent log odds ratio, and GSOEP is the reference category.

	W	omen	Men		
	Exit	Re-entry	Exit	Re-entr	
Age	-0.016**	0.014**	-0.022**	0.022**	
Age2	0.000**	-0.000**	0.000**	-0.000**	
Period					
1983-89 (ref.)					
1990-94	0.004	-0.102**	0.070*	-0.123**	
1995-99	0.004	-0.205**	-0.009	-0.213**	
Time into employment/out of employment					
less than 13 months (ref.)					
13-24 months	-0.428**	-0.647**	-0.668**	-0.355**	
25-48 months	-0.612**	-0.810**	-0.964**	-0.883**	
more than 48 monts	-0.643**	-1.682**	-1.464**	-1.161**	
Lifecourse					
single (ref.)					
married not children	0.881**	-0.390**	-0.257**	0.316**	
small children	-0.156**	-0.734**	0.076	-0.099	
only old children	-0.241**	0.612**	-0.020	0.029	
empty nest	-0.522	0.338	-0.014	0.130	
Education (CASMIN)					
labc(ref.)					
2ab	-0.100**	0.210**	0.011	0.018	
2c	0.279**	-0.022	0.452**	-0.500**	
3ab	0.118*	0.293**	-0.437**	0.058	
Survey/time distance to interview					
GSOEP (ref.)					
glhs max2 yrs	-0.057	-0.268**	-0.479**	-0.391**	
glhs 3-5 yrs	-0.153*	-0.104	0.012	-0.053	
glhs 6-10 yrs	-0.039	0.015	-0.144*	-0.019	
glhs more than 10 yrs	-0.523**	0.228*	-0.064	0.016	
Seam effect (GSOEP)	1.014**	0.944**	1.077**	0.806**	
Number of job spells	0.050**	0.113**	0.040**	0.013	
Number of spells * GLHS	-0.061**	-0.190**	-0.067**	-0.015	
Constant	-0.518*	0.355**	0.909**	-6.386**	
variance component (lev. 2)	-1.	466**	-1.	458**	
st. dev.	0.4	480**	0.4	482**	
rho	0.0	656**	0.0)66**	
-2 LL	100	667,63	126	335,08	
df		47		47	
N obs	67	75501	90)6997	
11 003					

 Tab.
 7 Estimated coefficients from multilevel multistate model of transitions out (exit) and into employment (re-entry), women and men

WomenMenAge 0.004^* -0.000 Age2 -0.000^{**} -0.000^* Period $1983-89 (ref.)$ -0.005 $1990-94$ 0.101^* -0.005 $1995-99$ 0.191^{**} 0.068 Tenureless than 13 months (ref.) $13-24$ months -0.146^{**} -0.056 $25-48$ months -0.177^{**} -0.185^{**} more than 48 monts -0.492^{**} -0.462^{**} Lifecoursesingle (ref.) $married$ not childrenmarried not children -0.278^{**} -0.124^{**} only old children 0.347^{**} -0.013 empty nest -0.853 0.411
Age Age2 0.004^* -0.000^{**} -0.000 -0.000^{*} Period 1983-89 (ref.) 1990-94 0.101^* 0.101^* $-0.0051995-99-0.0050.191^{**}Tenureless than 13 months (ref.)13-24 months-0.146^{**}-0.146^{**}-0.05625-48 months-0.177^{**}-0.185^{**}-0.462^{**}Lifecoursesingle (ref.)married not children-0.159^{**}-0.278^{**}-0.124^{**}-0.013empty nest-0.853Education (CASMIN)-0.004^*$
Age 0.004^* -0.000 Age2 -0.000^{**} -0.000^* Period $1983-89 (ref.)$ -0.005 $1990-94$ 0.101^* -0.005 $1995-99$ 0.191^{**} 0.068 Tenureless than 13 months (ref.) $13-24$ months -0.146^{**} -0.056 $25-48$ months -0.177^{**} -0.185^{**} more than 48 monts -0.492^{**} -0.462^{**} Lifecoursesingle (ref.) 0.085 small children -0.278^{**} -0.124^{**} only old children 0.347^{**} -0.013 empty nest -0.853 0.411
Age2 -0.000^{**} -0.000^{*} Period1983-89 (ref.)1990-94 0.101^{*} -0.005 1995-99 0.191^{**} 0.068 Tenureless than 13 months (ref.)13-24 months -0.146^{**} 0.13^{-24} months -0.146^{**} 0.177^{**} 0.185^{**} more than 48 monts -0.492^{**} Lifecoursesingle (ref.)married not children -0.159^{**} 0.085 small children 0.0278^{**} -0.124^{**} $only old children$ 0.347^{**} -0.013 empty nest -0.853 0.411
Period $1983-89 (ref.)$ $1990-94$ 0.101^* $0.0051995-99-0.0050.068Tenureless than 13 months (ref.)13-24 months-0.146^{**}0.146^{**}-0.05625-48 months25-48 months-0.177^{**}0.462^{**}-0.462^{**}Lifecoursesingle (ref.)married not children0.0278^{**}0.0850.085small children-0.278^{**}0.013empty nest-0.8530.411$
Period 1983-89 (ref.) 1990-94 0.101* -0.005 1995-99 0.191** 0.068 Tenure less than 13 months (ref.) 13-24 months -0.146** -0.056 25-48 months -0.177** -0.185** more than 48 monts -0.492** -0.462** Lifecourse single (ref.) 0.085 married not children -0.159** 0.085 small children -0.278** -0.124** only old children 0.347** -0.013 empty nest -0.853 0.411
$\begin{array}{cccccccccccccccccccccccccccccccccccc$
$\begin{array}{cccccccccccccccccccccccccccccccccccc$
1995-99 0.191** 0.068 Tenure less than 13 months (ref.) 13-24 months -0.146** -0.056 25-48 months -0.177** -0.185** more than 48 monts -0.492** -0.462** Lifecourse single (ref.) 0.085 married not children -0.159** 0.085 small children -0.278** -0.124** only old children 0.347** -0.013 empty nest -0.853 0.411
Tenure less than 13 months (ref.) 13-24 months -0.146** 25-48 months -0.177** more than 48 monts -0.492** Lifecourse -0.462** Lifecourse single (ref.) married not children -0.159** 0.085 small children -0.278** -0.124** only old children 0.347** -0.013 empty nest -0.853 0.411
less than 13 months (ref.) 13-24 months -0.146** -0.056 25-48 months -0.177** -0.185** more than 48 monts -0.492** -0.462** Lifecourse single (ref.) 0.085 married not children -0.278** -0.124** only old children 0.347** -0.013 empty nest -0.853 0.411
13-24 months -0.146** -0.056 25-48 months -0.177** -0.185** more than 48 monts -0.492** -0.462** Lifecourse single (ref.) 0.085 married not children -0.159** 0.085 small children -0.278** -0.124** only old children 0.347** -0.013 empty nest -0.853 0.411
25-48 months -0.177** -0.185** more than 48 monts -0.492** -0.462** Lifecourse single (ref.) married not children -0.159** 0.085 small children -0.278** -0.124** only old children 0.347** -0.013 empty nest -0.853 0.411
more than 48 monts-0.492**-0.462**Lifecourse single (ref.)-0.159**0.085married not children-0.159**0.085small children-0.278**-0.124**only old children0.347**-0.013empty nest-0.8530.411
Lifecourse single (ref.) married not children -0.159** 0.085 small children -0.278** -0.124** only old children 0.347** -0.013 empty nest -0.853 0.411 Education (CASMIN)
Lifecourse single (ref.) married not children -0.159** 0.085 small children -0.278** -0.124** only old children 0.347** -0.013 empty nest -0.853 0.411 Education (CASMIN)
single (ref.) married not children -0.159** 0.085 small children -0.278** -0.124** only old children 0.347** -0.013 empty nest -0.853 0.411
married not children-0.159**0.085small children-0.278**-0.124**only old children0.347**-0.013empty nest-0.8530.411
small children -0.278** -0.124** only old children 0.347** -0.013 empty nest -0.853 0.411 Education (CASMIN)
only old children 0.347** -0.013 empty nest -0.853 0.411 Education (CASMIN)
<i>empty nest</i> -0.853 0.411 Education (CASMIN)
Education (CASMIN)
Labe(ref)
$\frac{1}{2}ab = 0.052 \qquad 0.070*$
$2ab 0.032 0.079^{\circ}$
$2c = 0.221^{\text{AAA}} = 0.120^{\text{AAA}}$
5 <i>ab</i> 0.271*** 0.251***
Survey/time distance to interview
GSOEP (ref.)
glhs max2 yrs -0.315** 0.035
glhs 3-5 yrs 0.126 0.166*
glhs 6-10 vrs 0.137 0.050
glhs more than 10 yrs 0.097 -0.213*
Number of job spells 0.121** 0.160**
Number of spells * GLHS -0.051 -0.103**
•
Constant -5.272 -4.593
variance component (lev. 2) -1.483*** -1.341***
st. dev. 0.476 0.512
-21 J26171 72 -2988/ 81
df 22 22
N 390310 747404
N groups 7915 11789

Tab. 8 Estimated coefficients from multilevel model of transitions from job to job, women and men

* p<0.05, ** p<0.01

Age: for both men and for women it has a negative effect on exit and a positive effect on re-entry and on job to job moves. Figures 4 to 6 plot the age effect resulting from the two variables indicating age. Its effect on exit chances seems for women more stable during the life course than for men. Although for both the effect is weaker (less negative) at the end of the career (clearly related to exit due to retirement –early retirement-), for men the negative effect is much stronger in the middle ages, leading to much lower exit probabilities for men, and thus to much more discrepancy by gender. From the more substantive point of view, this vouches for the strong protection of the core of employment (men in the central ages) in Germany.



Fig. 4 Age effect on exit probabilities, males and females

The effect of age on re-entry appears very different between men and women, showing for the firsts a U-shaped form and for the second an inverse U-shape. Re-entry probabilities are for men lower in the middle ages, while for women the opposite is the case. We may link this to the lower exit of men in the central ages, which consequently lead also to less re-entries.

Fig. 5 Age effect on re-entry probabilities, males and females



More similar is instead the effect on job to job moves, which show in both the cases a decreasing effect of age, which is constantly stronger for males, showing that men's careers are overall more volatile than women's.

Fig. 6 Age effect on job to job mobility, males and females



Period: doesn't seem to have a significant effect on exit, while it seems that re-entry probabilities decreased during the 90s. Job to job mobility appears to increase during the 90s.

Time into employment/tenure: both for men and women, the longer people stay in employment, the less likely they are to exit (in agreement with the human capital theory) and to change job (considering job tenure). The effect is strong and highly significant

Time out of employment: both for men and women, the longer people stay out of employment, the fewer are the chances to re-enter the labour market.

Lifecourse: Stage in the life course seems to have a significant effect on behaviour on the labour market especially for women. Marriage implies for them greater chances to leave employment and lower chances to re-enter the labour market or to change job. Having small children, instead, leads to a decrease in the chances to exit but to change job and to re-enter as well. Decreased exit probabilities may be due to an anticipation effect, such that women who are going to have a child exit the labour market before giving birth, or to more economical needs connected with having a small child, but when woman are already out of the labour market having small children make it more difficult to enter again, and the same holds for changing job, so that women with children are less volatile.

When children grow older, exit chances become lower (probably due to selection of women who remain employed, while those who wanted to exit have already done it before), while re-entry chances increase, and women seem more volatile.

Life course effects instead seem for men quite different and less significant. Marriage seems to decrease exit, in accordance with the greater economical needs connected with supporting a family and to increase re-entry chances. Further stages in the labour market seem less significant instead.

Education: difficult to interpret. many n.s.

Highly educated women seem to have higher chances of re-entry, though they also exit more than the lower educated. High educated men seem instead to re-enter more often, but also to exit less. This pattern might be reasonable thinking that high educated women tend to temporarily go out of the labour market when having a child, since then they do not have difficulties in re-entering. Low educated women instead, might avoid going out of the labour market because of the greater difficulties they have to face to re-enter. Men instead have less need to go out, and when it happens it is maybe less of a choice, so that the high educated experience it less often. High educated people show more volatile careers (they switch more often job, without non employment spells in between), probably due to the lower difficulties in finding a new job. The more jobs a person had, the higher the transition probabilities, but since parameter estimates are less strong in the GLHS (which is more affected by recall accuracy) compared to the GSOEP³², we might argue that people with more jobs also have more difficulties in recalling (they forget some spells).

The interaction term between the variable indicating the number of job spells and the survey design, indicates that the effect is less strong in the life history study. We might argue that people with less stable careers also have more difficulties in recalling, so that the effect of the number of job spells appears less strong in the GLHS. Thus, while on the one hand the more job spells in the past the more likely is that a respondent experience a transition, on the other hand, a more volatile career might act in the opposite direction, because of difficulties in recall, and therefore less accuracy.

The dummy variable accounting for seam effects in the GSOEP confirms indeed that heapings are taking place in December, and it is controlling for it.

Tub. 7 Logii coefficienis for genuer effeci					
	Exit	Re-entry	Job to job		
Female	0.179**	-0.683**	0.149**		
Female * GLHS	0.163**	0.160**	-0.318**		
GLHS	-0.445**	-0.501**	0.170**		

Tab. 9 Logit coefficients for gender effect

* p<0.05, ** p<0.01

From the models estimated separately on men and women we have noticed that the survey effect has a different strength by gender, being much more negative for the firsts. Hence, we want to test whether there is a significant difference between men and women in the survey effect. This is done estimating one model for men and women and introducing an interaction term between the survey dummy and a dummy for sex (with male as reference category).

Table 9 clearly shows that there is indeed a significant difference in the survey effect by gender, being for example the odds of exit 19,6% ([exp(0.179)-1]*100) higher for female than for males according to the GSOEP and 41% ([exp(0.179 + 0.163) -1]*100) higher for females according to the GLHS. Women exit more than men, but according to the GLHS the difference is even bigger: we may think that is due to the fact that women recall more/better.

The odds of re-entry about 49% lower for females according to the GSOEP estimates and only 41% lower for female according to the GLHS. Women re-entry less than men, but, according to the GLHS the difference is smaller, maybe due to the fact that they recall better, and this decreases the differences.

The overall survey effect appears stronger for men: the odds of exit are for males 36% lower with the GLHS, while for females are only 25% lower.

Unexpected results when analysing job to job mobility may again be probably due to the very different definition of job spells in the two surveys.

Discussion

In this paper, we have presented the results of some analyses on labour and job mobility in Germany using two different data sources.

The two datasets we are using exemplify two very different research designs. We have clearly explained the main differences between the two. Basically, they have different time frames to observe (social) changes. The Panel survey monitors social changes wave by wave starting in 1984, so that information is only captured starting from the time of the first interview (and one year

³² We found this in an analysis (not shown here) done separately on the two datasets.

retrospectively) and left censoring problems can easily occur. In contrast, the Life History Study reconstructs in a single time point social changes which occurred also before the interview. Moreover, in the panel data respondents start from the present and recall events backwards over the period whereas in the employment life-histories respondents are asked about their employment status from age 15 up to the interview date, forwards. This discrepancy may be crucial to recall. We are trying to assess the extent to which differences in the research design may affect the results of estimations concerning transitions on the labour market.

The descriptive analyses regarding transition probabilities based on both the datasets have revealed the same pattern with the panel data and the life course data. Although the path seems to be the same, the level of the transition chances is quite different. According to our analyses, life course data appear to underestimate transition probabilities. We want to assess the role played by memory bias in generating different estimates in the two datasets, but this is not an easy task. Despite we do see a survey effect, which shows itself in lower transition probabilities when using retrospective data, we do not find a clear significant effect of time distance to the interview, which would lead us to conclude that the length of the recall period has a significant effect.

Concerning the analysis of job to job changes, it has to be stressed that in the GLHS a job spell is defined as a by the respondent reported episode of occupational activity in which s/he is engaged in one main, paid occupation for at least 20 hours a week. Job spells are recalled chronologically by respondents from completion of primary education, and spells are separated by times of unemployment or labour market inactivity, and changes of firms, professional position, income, activity or work time arrangements (Uunk, Mach and Mayer, 2005). In the GSOEP, on the other hand, employment spells are detached, according to our reconstruction, by unemployment or labour market inactivity, while job spells are separated by changes of firms, professional position within a firm, or switch to self employment status. Income or working time arrangements changes are not sufficient conditions to identify a different job spell. Furthermore, the respondent has to declare explicitly that s/he had a job change. Because of the different criteria for job spells definition, many more employment episodes might be split into job spells in the GLHS generating more job changes. Hence, we have also tried to estimate job to job moves taking into account changes in working time arrangements splitting a spell in case of change from part time to full time or vice versa. This is not sufficient to homogenise job spells definition in the two datasets, but at least brings the GSOEP definition³³ closer to the GLHS one. We might think that our attempt to adjust the definition of job spell so to make it closer in the two studies works better for women, for whom a big share of job to job moves involves a change in working time arrangements. This does not hold for men instead, whose kinds of job changes are more likely not to be captured by the definition adopted in the GSOEP. This may lead to an underestimation of job moves for males in the GSOEP compared to the GLHS and could be responsible for the closer transition probabilities for men in the two surveys.

To study job to job changes with less interference coming from problems in the job definition we may try to focus exclusively on employer change, which define only a specific kind of job to job change, whose definition should be more homogeneous between the two surveys.

To try to assess the true effect played by the research design in the estimation of transition probabilities, we have carried out some multivariate analyses, trying to control for all the differences between the two surveys in their sample composition as well as for sources of measurement errors.

³³ Furthermore the underestimation of job spells in the GSOEP could be due to the fact that there needs to be consistency between the information concerning the job change and that concerning employment status, coming from the calendar. A mismatch between them could cause the loss of a job change.

The first analyses on the reliability of the calendar information in the panel survey have revealed some sources of errors which may bias the estimation of transition probabilities, which we are trying to keep in mind and for which we are controlling, whenever possible.

Although we have some information concerning characteristics which are more often related to panel attrition³⁴, we do not have a clear overview of the possible effects of panel attrition on the estimation of labour market transitions. We are not aware of specific studies on the GSOEP but researches based on other data have shown that the assumption that right censoring is independent of spell duration cannot be rejected: attrition is not related either to the elapsed or to the remaining spell length (Pyy-Martikainen & Rendtel, 2003)³⁵. The use of a LHC providing short term retrospective information seems to lead to some incongruence with the current information but still our analysis leads us to think that this is mostly due to misdating, therefore not affecting much the number of transitions. Seam effects seem instead to be responsible for a possible overestimation of transitions. Job changes might on the one hand be overestimated due to double report but, on the other hand, the possibility to report at most one change per year may negatively affect the number of events, as well as the need of consistency with the calendar information concerning employment status.

The research design appears to have a role in determining the levels of transition probabilities, although differences do not appear to simply depend on the length of the recall period.

The two designs seem instead to reveal the same effects played by different covariates, although differences in recall accuracy among respondents with certain characteristics may increase or decrease the strength of the effects³⁶.

We may think that differences are related more in general to the whole way how the recall takes place: in the GSOEP, although we are using data which are to a certain extent retrospective, people are recalling a shorter period, while in the GLHS they are asked about all their life, so that the length of the period to be reconstructed differs quite a lot. This is one of the main differences between the two surveys and we may think that respondents have more difficulties in recalling events when asked to reconstruct a longer period. Unfortunately, we are not able to test whether estimates in the life history study are closer to those in the panel study when the period to be reconstructed is shorter, since the birth cohort selection in the GLHS is such that respondents have always to reconstruct at least ten years.

We have tried to take into account the different APC composition of the two surveys adopting a more conservative approach, i.e. selecting only the periods in which we have information coming from both. Furthermore, we are controlling for period and for age as well. Still, we cannot control for birth cohort due to the linear relation among the three factors³⁷.

³⁴ Specifically, important characteristics affecting attrition have been documented to be: household moved, split-off, large city, single household for unsuccessful follow-up and resident of East Berlin, age of head of household, female head, household moved, split-off, separation/divorce of partner, change of interviewer, number of interviews with the same interviewer, low household income, item non-response on income, expected loss of job, migration from East to West Germany for refusal of respondents (Haisken-DeNew and Frick , 2005)

³⁵ A popular strategy for copying with panel attrition is weighting. Yet, it is not clear how effective it is in reducing attrition bias (Vandecasteele & Debels, 2007).(DuMouchel & Duncan, 1981) argue that while weighting may be appropriate for estimating population means and totals, relying on weighted estimates may be dangerous in regression problems. Hoem (1989) argues that weighting can be superfluous and (Horowitz & Manski, 1995) think that the use of weights may lead to estimates of regression outside the set of logically possible values. Due both to the uncertainty in the effectiveness of weights in regression models and to further difficulties we encountered in using weights in panel models, up to now we have not used weights to account for panel attrition. Yet, we might do it in a further step, using the longitudinal weights available in the GSOEP.

³⁶ This can be proved estimating separate models on the GLHS and GSOEP (tables not shown).

³⁷ However, we have tried some analyses selecting only the overlapping cohorts and including a variable controlling for cohorts, and we have essentially obtained the same results.

When we started our analyses, we had some assumptions concerning the effect of age at interview: we might indeed think that the older, the higher the difficulties to recall. But we cannot introduce age at interview in the model, because it is highly correlated with the actual age. The main question which arises in this context concerns whether the effect associated with the age variable is a true age effect indicating the differences in the transition probabilities by age or is instead a memory effect associated with age, generated by the different ability/accuracy in recalling by age (at interview). Although due to the high correlation with age we cannot find a way to test for the effect of age at interview on the estimation of transition probabilities, we could study the effect of age at interview in the short term retrospective calendar. In that case we are looking at different mechanisms, since we are not reconstructing events on the labour market but we are just checking for congruence in different reports related to the same period, but still it is important to notice that we do not find a decline in the ability to recall by age.

The problems which arose in these analyses stress the importance of a deep insight in the survey design and in the way a specific information is recalled, especially in comparative research.

Appendix Detailed explanation concerning the construction of the dataset

We are using the CD-ROM Version 2005 for the GSOEP and the Gesamtdatenbank der Studien LV1, LV2, LV2T, LV3T und LVW6471 for the GLHS. Stata version 9 is used for data analysis.

Calendar information from the GSOEP is available in the ARTKAL file, which contains monthly spells starting in January 1983 *(for detailed info see DTC p. 107)*. This file contains the information stemming from the calendar presented to each respondent during the yearly interview and allows us to reconstruct employment spells, distinguishing between part time and full time work.

Starting from this file³⁸ we have added information related to job changes occurred from January of the year before, so to split employment spells in job spells. The information concerning job changes could be recorded twice, namely in the current year if it occurred before the interview date and in the following year as a change occurred in the previous year. In constructing our data set we made sure that no double counting occurred. Both the kind and the date of change were reported³⁹. In case a change was reported twice but a different type of change was reported at the two interviews, we have chosen for the type of change reported in the first one, which is closer to the event. Yet, an error could occur not only in stating the type of changes, in case, in two different interviews, a single change is reported as occurring at different time points. We have partially controlled for this not allowing two changes to be closer than 3 months. In this case, we assume there was actually only one change, but the second time the question was asked, in relation to the same reference period, a misdate occurred, leading to a second fake job change. Thus, we have relied on the job change date given at the first interview⁴⁰.

Starting from this information, we have proceeded transforming the data in a person-month format. A major problem arises in case of multiple spells in one month. Previous analyses (Ernicke, 1997; Bergemann and Mertens 2004) have dealt with the issue imposing a hierarchy of the spells. We think this way of dealing with the problem is not fully exhaustive, since different situations should be distinguished. Specifically we have to pay attention to the type of spells which are *co present* (?) in one month and to the length of the *copresence* (?).

Concerning the first point, we assume that employment status cannot coexist with "unemployment", "housewife" or "military service", while it could coexist with the other statuses⁴¹; concerning the second argument, instead, we assume that, in case the subject has reported two spells in one (or two) consecutive months, it could actually be that there was a change of status in that month(s), and the subject hold one status for part of the month (s) and another status for the remaining part. Thus, we will see one event (entry or exit) in case of one single month with multiple spells and two events (first entry, then exit, or vice versa) in case of two consecutive multiple spell months. The assumption is that in these cases, independent of the kind of spells, there was no coexistence, so we observe one (or two) event(s).

³⁸Furthermore, we have recoded the spelltyp variable in the following way: [......]

³⁹ In case of missing date of change, the modal month of change in that interview year was attributed.

⁴⁰ Furthermore, we have carried out also some reliability checks on the job change information. We have noticed that the later the interview was carried out, the higher the likelihood a change was reported twice, for obvious reasons. Thus, we expect an overestimation of job changes in the earlier months of the year compared to the latest, since for these months the likelihood of being investigated twice is higher. While cases in which only the declaration of the type of spell was different could be detected and corrected, misdating could not. Thus, we expect an overestimation of job changes stem from two different questions, is possible to have inconsistencies between the two. Namely, it is possible that respondents declare to have a job change in one month but in the calendar they declare not to be employed in that month. This mismatch in the dating, which is in most of the case just one month, leads to the loss of some job changes.

⁴¹ When this s the case, we would also like to add a weight to employment, although we have not done it yet.

In case of three or more consecutive months in which two or more spells coexist, we have distinguished two different situations: while in case of spells which might coexist with employment we have assumed *co presence*⁴², when the respondent has declared to have been employed in one month and to have been at the same time either housewife, unemployed or involved in the military or civil service (since we assume that coexistence of employment with unemployment, housewife or military service is not possible), we have flagged all the records from that moment onwards for that person, and we have then excluded them from our analysis. This decision is due to the fact that in these cases, were we to rely on the information given, we would see continuous movements in and out of employment (because each status seems to have been hold for part of a month) and this would mean that the employment spells are very short ones.

Furthermore we have used the retrospective life history information provided in the biography questionnaire (stored in the PBIOSPE file⁴³), to correct for left censored cases, adding, whenever possible, time duration in the left censored spell.

Regarding the GLHS< starting from the spell files, we have also reconstructed a person-month file, which we have then combined with the GSOEP one, creating a big dataset containing information from both the surveys.

⁴² We will also add a weight in this case.

⁴³ The file PBIOSPE contains yearly information on the job status in a spell form, updated with year-level aggregated monthly information from the P files of the current wave for the past calendar year.

References

- Auriat, N. (1993). "my wife knows best": A comparison of event dating accuracy between the wife, the husband, the couple, and the belgium population register. *Public Opinion Quarterly*, *57*(2), 165-190.
- Bergemann, A., & Mertens, A. (2004). Job stability trends, layoffs and transitions to unemployment: An empirical analysis for west germany (Vol. 4792). London: Centre for Economic Policy Research (CEPR).
- Bielby, W. T., Hauser, R. M., & Featherman, D. L. (1977). Response errors of black and nonblack males in models of the intergenerational transmission of socioeconomic status. *The American Journal of Sociology*, 82(6), 1242-1288.
- Blossfeld, H.-P. (1986). Career opportunities in the federal republic of germany: A dynamic approach to the study of life-course, cohort, and period effects. *European Sociological Review*, 2(3), 208-225.
- Bradburn, Rips Shevell (1987)
- Brückner, E. (1993). Lebensverläufe und gesellschaftlicher wandel: Konzeption, design und methodik der erhebung von lebensverläufen der geburtsjahrgänge 1919 – 1921. Berlin: Max-Planck-Institut für Bildungsforschung. –Cologne: Zentral Archiv –ZA2646, ZA2647 [distributor].
- Brückner, E. a. K. U. M. (1998). Collecting life history data. Experiences from the german life history study. In J. Z. E. J. G. H. Giele (Ed.), *Methods of life course research: Qualitative and quantitative approaches*.: Thousand Oaks et al.: Sage Publications.
- Brückner, H. (1995). Surveys don't lie, people do? An analysis of data quality in a retrospective life course study (Vol. 50). Berlin: Max-Planck-Institut für Bildungsforschung.
- Callegaro, M. (2006). Seam effects in panel data collected with event hsitory calendar methodology: The case of labor force transitions in the psid.
- Cleves, M. A., Gould, W. W., & Gutierrez, R. G. (2004). *An introduction to survival analysis using stata*. College Station, Texas: Stata Press.
- Crowder, R. G. (1976). *Principles of learning and memory*. Hillsdale New York: Erlbaum. Distributed by Wiley.
- De Graaf, N. D., Wegener, B., & Liebig, S. (1989). Reliability of retrospective survey data: A comparison of job career and social resource variables.
- Dex, S. (1991). The reliability of recall data: A literature review. from http://ideas.repec.org/p/ese/iserwp/1991-11.html
- Drobnic, S. (2000). The effects of children on married and lone mothers' employment in the united states and (west) germany. *Eur Sociol Rev, 16*(2), 137-157.
- Drobnic, S., Blossfeld, H.-P., & Rohwer, G. (1999). Dynamics of women's employment patterns over the family life course: A comparison of the united states and germany. *Journal of Marriage and the Family*, *61*(1), 133-146.
- DuMouchel, W., & Duncan, G. J. (1981). Using sample survey weights in multiple regression analyses of stratified samples.
- Duncan, G. J., & Hill, D. H. (1985). An investigation of the extent and consequences of measurement error in labor-economic survey data. *Journal of Labor Economics*, *3*(4), 508-532.
- Featherman, D. (1980). Retrospective longitudinal research: Methodological considerations. Journal of Economics and Business, 32(2), 152-169.
- Freedman, D., Thornton, A., Camburn, D., Alwin, D., & Young-DeMarco, L. (1988). The life history calendar: A technique for collecting retrospective data. *Sociological Methodology*, *18*, 37-68.
- Frick, J. (2005). A general introduction to the german socio-economic panel study (soep). Design, contents and data structure [waves a-u, 1984-2004].

- Gibson, J., & Kim, B. (2005, July 29th). Can we trust retrospective recall to replace panel data? from <u>http://www.caer.unsw.edu.au/Laour_Econometrics_05/Gibson_Kim_LEW05.pdf</u>
- Haisken-DeNew, J. P., & Frick, J. R. (2005). *Dtc desktop companion to the german socio-economic panel (soep)* (Vol. Vers. 8).
- Hill, D. H. (1987). Response errors around the seam: Analysis of change in a panel with overlapping reference periods.
- Hoem, J. M. (1989). The issue of weights in panel surveys of individual behavior. In Wiley (Ed.), *Panel surveys*.
- Horowitz, J. L., & Manski, C. F. (1995). Censoring of outcomes and regressors due to survey nonresponse: Identification and estimation using weights and imputations. *Econometrics*
- Hujer, R., & Schneider, H. (1989). The analysis of labor market mobility using panel data. *European Economic Review*, 33(2-3), 530.
- Jacobs, S. C. (2002). Reliability and recall of unemployment events using retrospective data. *Work, employment and society, Volume 16*(3), 537–548.
- Jürges, H. (2005). Retrospective error in soep calendar data: The case of unemployment. *Schmollers Jahrbuch*, 125, 157-165.
- Kalmijn, M., & Luijkx, R. (2006). Changes in women's employment and occupational mobility in the netherlands between 1955 and 2000". In H.-P. Blossfeld & H. Hofmeister (Eds.), *Globalization, uncertainty, and women's careers: An international comparison* (pp. 84-112). Cheltenham, UK, and Northampton (MA-USA): Edward Elgar.
- 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. In *Jahrbuecher fuer nationaloekonomie und statistiek* (Vol. 217/5). Stuttgart.
- Kurz, K., Hillmert, S., & Grunov, D. (2002). Increasing instability in employment careers? Men's job mobility and unemployment in west germany: A comparison of the birth cohorts 1940, 1955 and 1964. *Globalife Working Paper No. 34*, from <u>http://web.unibamberg.de/sowi/soziologie-i/globalife/workingpapers/index.html</u>
- Loftus, E. F. (1979). Eyewitness testimony. Cambridge, Mass. [etc.]: Harvard University Press.
- Loftus Marburger (1983)
- Mayer (2007) Achilles hee or problem solved? Autobiographical memory and retrospective measurement, Spring Meeting of the Methodological Section of the American Sociological Association
- Mayer, K. U. (forthcoming) Retrospective longitudinal research: The German Life History Study, in Menard, S. (ed.) *Handbook of longitudinal research: design, measurement and analysis*. San Diego: Elsevier
- Mayer, K. U., & Carroll, G. R. (1987). Jobs and classes: Structural constraints on career mobility. *European Sociological Review*, *3*(1), 14-38.
- Mayer, K. U., & Tuma, N. B. e. (1987). *Applications of event history analysis in life course research* (Vol. Nr. 30). Berlin: Max-Planck-Inst. für Bildungsforschung.
- Mathiowetz and Duncan (1988)
- McCulloch, S. D. a. A. (1997). The reliability of retrospective unemployment history data. from <u>http://www.reformmonitor.org/httpd-cache/doc_abstracts-1246.html</u>
- Mertens, A. (1999). Job stability trends and labor market (re-)entry in west germany 1984-1997.
- Moen, P. (1991). Transitions in mid-life: Women's work and family roles in the 1970s. *Journal of Marriage and the Family, 53*(1), 135-150.
- Morgenstern, R. D., & Barrett, N. S. (1974). The retrospective bias in unemployment reporting by sex, race and age. *Journal of the American Statistical Association*, 69(346), 355-357.
- Moss, L., & Goldstein, H. E. (1979). The recall method in social surveys.
- Neisser, U. (1982). Memory observed: Remembering in natural contexts. San Francisco: W.H. Freeman.
- Paull, G. (2002). Biases in the reporting of labour market dynamics. from http://www.ifs.org.uk/wps/wp0210.pdf

- Peracchi, F. (2002). The european community household panel: A review. *Empirical Economics*, 27, 63-90.
- Peters, H. E. (1988). Retrospective versus panel data in analyzing lifecycle events. *Journal of Human Resources*, 23(4), 488-513.
- Pierret, C. R. (2001). Event history data and survey recall: An analysis of the national longitudinal survey of youth 1979 recall experiment. *Journal of Human Resources*, *36*(3), 439-466.
- Poterba, J. M., & Summers, L. H. (1986). Reporting errors and labor market dynamics. *Econometrica: Journal of the Econometric Society*, 54(6), 1319-1338.
- Powers, E. A., Goudy, W. J., & Keith, P. M. (1978). Congruence between panel and recall data in longitudinal research. *Public Opinion Quarterly*, 42(3), 380-389.
- Pyy-Martikainen, M., & Rendtel, U. (2003). The effects of panel attrition on the analysis of unemployment spells.
- Reimer, M., and Matthes, B. (2006). Collecting Event Histories with True Tales: Techniques to Improve Autobiographical Recall Problems in Standardized Interviews, Quality and Quantity
- Reimer, M. (2005a). Autobiografisches Gedächtnis und retrospektive Datenerhebung. Die Rekonstruktion und Validität von Lebensverläufen, Max-Planck-Institut für Bildungsforschung, Studien und Berichte Nr. 70, Berlin
- Reimer, M. (2004). Collecting event history data about work careers retrospectively: Mistakes that occur and ways to prevent them. *Kognition und Kommunikation dei der Laengsschnittdatenerhebung*, from <u>http://www.mpib-berlin.mpg.de/dok/full/reimer/collecti/Arbeitsber1.pdf</u>
- Reimer, M., & Kuenster, R. (2004). Linking job episodes from retrospective surveys and social security data: Specific challenges, feasibility and quality of outcome. *Ausbildungs- und Berufsverlaeufe der Geburtskohorten 1964 und 1971 in Westdeutschland*, from <u>http://www.mpib-berlin.mpg.de/dok/full/reimer/linkingj/Arbeitsber8.pdf</u>
- Reimer, M. (2001). Die Zuverlässigkeit des autobiographischen Gedächtnisses und die Validität retrospektiv erhobener Lebensverlaufsdaten. Kognitive und erhebungspragmatische Aspekte. Max-Planck-Institut für Bildungsforschung, Berlin (Materialien aus der Bildungsforschung 71).
- Rubin, D.C. (Ed.) (1986). Autobiographical Memory. Cambridge University Press, Cambridge
- Rubin, D.C. (Ed.), (1996). Remembering Our Past: Studies in Autobiographical Memory. Cambridge University Press, Cambridge.
- Schwarz, N., & Sudman, S. (1993). Autobiographical memory and the validity of retrospective reports: Springer-Verlag.
- Scott, J., & Alwin, D. (1998). Retrospective versus prospective measurement of life histories in longitudinal research. In J. Z. E. J. G. H. Giele (Ed.), *Methods of life course research: Qualitative and quantitative approaches*.: Thousand Oaks et al.: Sage Publications.
- Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY [etc.]: Oxford University Press.
- Solga, H. (2001). Longitudinal surveys and the study of occupational mobility: Panel and retrospective design in comparison. *Quality and Quantity*, 35(3), 291 309.
- Vandecasteele, L., & Debels, A. (2007). Attrition in panel data: The effectiveness of weighting. *Eur Sociol Rev, 23*(1), 81-97.