Labor mobility patterns over the life-course

Manzoni, A.

Publication date:
2010

Document Version
Publisher's PDF, also known as Version of record

Anna Manzoni

Labor mobility patterns over the life-course

A comparison of retrospective and prospective data in different labor markets

Labor mobility patterns over the life-course.

A comparison of retrospective and prospective data in different labor markets

ISBN 978-90-5335-251-9

Cover design by Simone Vinke

Printing: Ridderprint bv, Ridderkerk

Labor mobility patterns over the life-course.

A comparison of retrospective and prospective data in different labor markets

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit van Tilburg,
op gezag van de rector magnificus, prof. dr. Ph. Eijlander, in het openbaar te verdedigen
ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de
Universiteit op vrijdag 19 februari 2010 om 14.15 uur

door

Anna Manzoni
geboren op 12 mei 1981
te Lecco, Italië.

Promotor	Prof. R. J. A. (Ruud) Muffels
Copromotor	Dr. A.R.C.M. (Ruud) Luijkx
Promotiecommissie	Prof. Karl Ulrich Mayer Prof. Heike Solga Prof. Jeroen Vermunt Prof. Paul De Graaf Prof. Melinda Mills

Acknowledgements

Looking back over the past four years, I must say I have really enjoyed my time as a PhD student. I learned so much, and I am not just talking about 'formal' learning, which goes without saying. I greatly benefited from the environment I was in and I grew as a person through daily talks and interactions, and I believe this is a 'sine qua non' for a good researcher, especially in the social sciences.

Many people contributed to this and I would like to acknowledge my debt to them. They have to a different extent and in different ways accompanied and helped me in this journey.

My decision to start a PhD project was relatively unplanned. I was quite unaware of what I would face and of the way the academic world works. I am not sure I would have made the same choice if I knew what was in front of me in the years to come. But I am happy that I made this choice, which definitely I do not regret.
Although it was 100% my own decision, I believe it was highly influenced by my professor during my undergraduate years, Paolo Barbieri: I don't think I would be where I am if I hadn't met you on my way. I am grateful to you for this.

I am grateful to my co-promoter and daily supervisor, Ruud Luijkx, for guiding me: I enjoyed your structured way of thinking and working.
To Ruud Muffels, my promoter, go my thanks for the fruitful exchanges of ideas and for motivating me. I am certainly in debt to both 'Ruuds' for the insight into how to think like a scientist.
Thanks to Jeroen Vermunt for sharing his great methodological knowledge and giving me swift and helpful feedback, especially in the last year, and also for always smiling and showing interest in my research.
Thanks to Didier Fouarge for sharing many data problems and for always offering his advice: working with you was always very pleasant and efficient.

Thanks to Maike for sharing the office with me, which certainly isn't an easy task if you have seen how messy I can manage to make it. But also for always being up for a chat and a drink (or two). Without Maike's help, the Dutch summary of this thesis would have probably been quite difficult to read.
I also benefited, directly or indirectly, from the contacts with the rest of my colleagues, who certainly contributed to my development as a social scientist as well. I don't need to list them.
Thanks to Els Kuster, who helped me in dealing with some practical matters and made it easier for me to get through the dissertation approval process from a distance of thousands of miles.
Ton Heinen, the director of the PhD School of the Faculty of Social Sciences at Tilburg University, has been available at any time for any concerns I had, and Shirley Baert greatly supported me in the final stage.

Thanks to all the scholars who allowed me to email them with tons of questions and always offered very swift replies.

In terms of 'institutions', I am very grateful to NWO, which supported my PhD project, making my way possible and relatively easy. It gave me the opportunity to travel and

participate in high-level international conferences, providing me with the financial support in such a way that I could concentrate fully on my thesis.

Thanks to the Swedish Institute for Social Research (SOFI) at Stockholm University and to the German Institute of Economic Research (DIW Berlin Deutsches Institut für Wirtschaftsforschung) for hosting me in spring 2007 and summer 2008 respectively.

I definitely couldn't have coped with the many hours spent in front of the computer without the expectation of an evening or a weekend climbing. And my climbing buddies were just wonderful in stimulating me: My instructor Gijs, and in particular Kim and Lukas, with whom I share more than just the passion for climbing. But also the whole training group and all those always present in the hall: you might be unaware of it, but you definitely contributed to my positive mood, which is (highly) directly correlated with my work productivity. And Johannes, to whom I could always express my strongest emotions, for being always available to listen and support me, for understanding (and correcting) my German, and for boosting my love for the great outdoors, especially after a great time on the wall and over a well-deserved Belgian beer.

I was so lucky to live in a very nice house during this time. This wouldn't have been the case if I hadn't met Heejung. I also had the pleasure to live with super nice people. Owen, I was really happy you moved in. Certainly my time there wouldn't have been as good without you. Thanks for sharing my love for sports and languages and most of all for always being available for anything. And I really want to stress *always* and *anything*. Roberta, who is just a great person: I am afraid any word I would use to describe her would not capture what I mean. And I'll stop here or I will get emotional…You should just meet her.

Thanks to (in alphabetical order) Anne-Kathrin, Bea, Carmen, Gonul, Marco, Michele A. , Michele M., Miguel, for the nice talks during coffee breaks and dinners. And to many others in Tilburg with whom I shared pleasant moments.
Thanks to Olga & the all the Amsterdam crew for giving me a great getaway from Tilburg any time I needed it.

I have always felt good in The Netherlands. But home is home, and actually distance made my love for my hometown stronger than it was when I lived there. Certainly this is greatly due to the fact that any time I visited, all my friends there were totally available for me. Spok, thanks for always being available to share a great time on the rocks and on the ski slopes, and for being such a good friend. Ale and Rosti always made me feel like I had never left our trio, because you know I actually haven't. Elena and Silke, our reunions are always just great.

And of course my family. First of all Mum and Dad, to whom I know I gave a hard time leaving (especially considering that I am an only child, in an Italian family…) and were always just so proud of me.
But my 'extended' family as well (and I cannot mention everyone or the acknowledgements would end up being longer than the body of the thesis, given the numbers…).
To all of my friends and family, thanks for sharing your time with me, for making me laugh, for giving me a great time, for listening, supporting, for just being there, and for your unconditional love and faith in me.

TABLE OF CONTENTS

Chapter 1

Labor mobility patterns over the life-course: Introduction

1. Research questions and aims

The collection of empirical studies brought together in this thesis concerns the substantive and methodological analysis of labor market dynamics in various European labor markets using panel and life history data. The main substantive focus of the thesis resides in the analysis of patterns of labor market mobility over time and across birth cohorts to examine the alleged rising volatility on the labor market. The methodological analysis deals with concerns which are raised in the literature on the reliability of life history data associated with memory or recall bias. The basic idea underlying most of the studies is to examine the extent by which life history and panel data come to similar or different conclusions with respect to labor market mobility over time and to try, to the extent possible, to correct for memory bias in order to be able to properly address the substantive research questions.[1]

Substantive questions (SQ)

We aim at investigating the extent to which labor market patterns are changing over time and across cohorts addressing the current academic debate on the alleged increasing volatility on the labor market. While the general theoretical belief is that job stability is on decline, empirical evidence is scarce and ambiguous (Bergemann and Mertens, 2004; Biemann et al., 2009) and some authors claim that stability still prevails (Mayer, 2000, 2008; Winkelmann and Zimmermann, 1998).

We are also interested in corroborating claims that these long-term changes regard specific groups, characterized by gender, age, skill or educational level.

Furthermore, we want to unravel which dimensions affect labor market careers and determine the way in which people structure their life course, by examining the

[1] Here, we outline the main substantive and methodological research questions that we aim to address in this thesis, while we refer to paragraph 4 for the specific questions tackled in each chapter.

possible causes of the alleged changes, considering both micro and macro level determinants.

We also intend to look at the role played by events in one's life: unemployment, childbirth, divorce and life course stage. As it will become clearer in the explanation of our theoretical framework (Section 2), this stems from the basic idea of life course theory according to which people experience influential events in their lives and go through different phases which are linked to one another.

We want to study employment patterns both looking at specific events on the labor market (entry, exit and job to job shifts) as well as participation and at the effect of specific events on it. We not only look at levels of transitions and participation but we also try to unravel their determinants.

Eventually, we believe that different employment and gender regimes (Lewis, 1992; Lewis et al., 2008) might be differently affected by the globalization process allegedly induced by skill-biased technological change, changes in the demand for human capital or in the supply of education endowments, changes in women's labor market participation and in the diffusion of part-time and temporary jobs. Therefore, we want to investigate whether unregulated countries are more exposed to changing mobility patterns than regulated regimes.

Summarizing, our substantive research questions read as follows:

Substantive research questions:

SQ1: To what extent are labor market patterns changing over time and across cohorts?

SQ2: Are labor market careers becoming more volatile or not?

SQ3: Are there differences in the extent to which labor market patterns change across different groups in society?

SQ4: What are the causes for changing labor market patterns? And specifically, what is the role played by increasing demand for human capital, increasing supply of education endowments, increasing labor market participation of women?

SQ5: What is the role played by events (like unemployment, childbirth or divorce) and by the stage in the life course? How does childbirth specifically affect women's labor market participation?

SQ6: Are unregulated countries more exposed to changing mobility patterns than regulated regimes? How are women's participation patterns differently affected across welfare regimes?

Methodological questions (MQ)

Advancements in life course research highly depend on the availability of good data: Longitudinal data are deemed the most valuable type of data to investigate such issues, for they provide a dynamic perspective which allows us to analyze social changes (see Section. 3 in this chapter for a more throughout discussion on the advantages offered by longitudinal data). The availability of longitudinal data highly increased in the past years, but their reliability is an issue of concern. This is where our methodological questions come into play. We want to examine possible sources of bias in longitudinal surveys, directing specific attention to the reliability of retrospective surveys and to memory bias in particular. We do not want to limit the analyses to a descriptive assessment of bias: instead, we aim at correcting for it by examining the alleged causes and including variables addressing these causes in our models.

We are interested in differences between prospective panel and retrospective life-history data, and the extent to which different designs reveal dissimilar outcomes concerning both the level of transition rates and their causation. We want to model the effect of (increasing) recall time and account for several others sources of error in each type of data as well as for compositional differences, in the attempt to isolate memory bias. We want to investigate the role played by design differences and by memory bias specifically, in order to come to conclusions about how similar or different inferences made from life course data and from panel data are. In this way, we can assess the extent to which prospective and retrospective 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 leads to the explanatory question why differences may occur. Correcting for biases, we attempt to obtain comparable outcomes across the two designs.

While both life history and panel data have been extensively used for substantive analyses, very little work has been done, to our knowledge, comparing them. In this respect, this thesis offers a significant contribution to the literature.

Our methodological research questions can be summarized as follows:

Methodological research questions:
MQ1: Which are the sources of bias in prospective and retrospective longitudinal data?
MQ2: What is the role played by design characteristics and memory bias in particular in affecting the estimates of labor market dynamics?
MQ3: How can we correct for memory bias?
MQ4: Do prospective and retrospective data lead to different estimates of levels and determinants of labor market dynamics?
MQ5: What are the advantages and disadvantages of combining the two data sources to arrive at better estimates of the reasons for change?

2. Theoretical perspective

Life course perspective

Our interest in the changes of career patterns over time and across cohorts and the focus on the role played by prior events and experiences make the life course perspective the main theoretical paradigm to which this thesis refers. A longitudinal and life course perspective has become the current "gold standard" of quantitative social science (Mayer, 2008) and longitudinal studies have been celebrated as the most progressive and innovative research area of all social science (Butz and Torrey, 2006).

In the past 20 years the life course approach has turned into a broadly accepted perspective and research in many fields now routinely relies on it (Kohli, 2007). Life course sociology emerged and developed over several decades, and can be seen as a specific branch of a larger field which embraces several other disciplines, such as anthropology, demography, psychology and economics. Such link also shows up in this thesis, in that, moving within a sociological framework, we also borrow from economic, psychological and demographic theories. In that sense, we embrace an interdisciplinary approach.

At the basis of the life course perspective is the idea that over the life course people experience influential events and go through different stages: finishing full-time education, starting a new job, leaving the parental home, finding a partner, forming a

family, children leaving the family (empty nest) and retiring from the labor market. The concept of life course refers to the sequence of roles that connect the phases of life and specific interest goes into the ways in which individual lives are linked to social change and in the relation of earlier phases of life to later phases. The term 'life course' is used by sociologists to denote the sequence of states and events in various life domains spanning from birth to death.

In this thesis, the attention is specifically directed to employment biographies, looking at how previous career decisions and events as well as human capital investments but also marital histories affect the employment career, over time and across cohorts. The underlying assumption is that earlier experiences condition later behavior and events, while available resources and opportunities unfold over time.

The influence of prior experiences and events becomes apparent in the thesis through the focus on examining the impact of previous labor market experiences, while taking duration dependence (time into the state) and career complexity (i.e. number of previous spells) into account. Also, education reflects investments in human capital. Furthermore, looking at the life course stage, as indicated by the marital status and the number and age of children, we are able to link employment and demographic histories.

We are interested in time-related variation: here a note is necessary. Three types of time-related effects can be distinguished: Age effects refer to variation over time associated with increasing age brought about by physiological changes, the occurrence of events or the accumulation of experience, and/or role or status changes. Period effects regard variation across time periods impacting on all age groups simultaneously, such as a war, a natural catastrophe or a severe economic crisis as the current one. Cohort effects are associated with changes across groups of individuals who experience an initial event, such as entrance on the labor market, in the same year(s) (Yang and Land, 2006). While age variations reflect the developmental nature of true age changes, period and cohort effects reflect the influence of social forces.

As it will become clearer in the next Section, the use of longitudinal data and event history analysis techniques in the thesis signals a dynamic perspective using the life course theory as our overarching theoretical framework.

Increasing volatility or stability? The state of the art and our contribution

Due to individualization, ageing and globalization trends resulting from the rapidly changing social and economic context, labor markets are presumed to become less stable and more volatile. In the 1980s, an increasing amount of changes over the life-course started to take place.

Cyclical and structural developments (such as technological innovation, educational expansion, qualification upgrading, demographic changes, and increasing female labor force participation) affecting both the demand and the supply side, in parallel with the increasing availability of longitudinal data, have led to a flourishing research practice on labor market dynamics. The literature has concentrated on job (in)stability, occupational choices, impact of family cycle on employment, transitions out of employment and employment insecurity, labor force gender inequality, within-firm and across-firm mobility, class and job mobility.

A great deal of studies on employment instability focuses on social stratification and inequality issues by examining the influence of class versus education (Breen, 1997; Goldthorpe, 2002) and on the impact of the existence of different labor market segments across which mobility is assumed to be low (Mayer and Carroll, 1987). The life course theory has focused on life course events which may act as trigger events affecting later life as well as the future career. One such event is unemployment and there is growing evidence that it has a 'scarring' effect on future employment chances (Gangl, 2006). Furthermore, a great deal of interest went into the adverse role of the very strict or rigid labor market and social security institutions, allegedly affecting the mobility patterns of employed and unemployed workers as well as the outcomes in terms of segmentation and stratification (DiPrete et al., 2006).

Although much of the literature concentrates on trends in labor market mobility and stability, in the sociological debate already in the 1980s there was a growing interest in studying and explaining individual career opportunities, which are determined both by variables describing characteristics of individuals and variables capturing attributes of the labor market structure (Blossfeld, 1986).

Concepts such as *differentiation, de-institutionalization and de-standardization* have been introduced. Differentiation refers to an allegedly increasing number of distinct labor market states people go through over the life course. De-institutionalization points to the fact that states, stages and events are being reintegrated or fused, becoming less structured, which makes life courses less predictable. De-standardization indicates that states, events and their sequences are shared by a smaller part of a population or

occur at more dispersed ages and with more dispersed durations (Brückner and Mayer, 2005).

In the past decades, changes in the domains of longevity and population aging, processes of deregulation of the labor market, flexibility of employment patterns and discontinuities of family life as well as structural and cultural transformations have been accompanied by increasing heterogeneity across the population. The life course has allegedly become more complex, less predictable, less collectively determined, less stable, less orderly, more flexible, and more individualized and it seems to be co-constructed by individuals and institutions (Heinz and Marshall, 2003). Working lives in particular appear more unstable. One major aspect of life courses is their internal temporal ordering, that is, the relative duration in given states as well as the age distributions at various events or transitions.

The male normal work biography of continuous full-time employment that had become the rule by the 1960s in most of the advanced capitalist economies is fast giving way to various forms of discontinuous careers interspersed with spells of part-time work and non-work activities (Kohli, 2007). Cohorts or historical periods mark differences in life-course options and in passages in the social system of education, employment, family, social assistance and retirement. Recently, attention has been directed to examining the stability of careers and to investigating the causes and consequences of career instability.

Changes may reflect both the historical and the biographical context, and should therefore be investigated across cohorts and life domains. Individual lives are rooted into social structures and reflect their demands and opportunities. Variability occurs not only in the life course, but at the same time, also across historical and cross-national settings. Over time, shifts in the institutional setting, demographic transitions as well as specific developments in the educational and occupational system but in values and preferences as well, are likely to affect life course outcomes and labor market patterns specifically (Mayer, 2008).

From an historical perspective, employment participation patterns have been changing allegedly rapidly for both men and women in the last decades, encouraging a flourishing research practice on labor market dynamics and growing interest in individual career opportunities, shedding light on the importance of factors affecting both the demand and the supply side.

Therefore, attention needs to be paid to the influence of personal characteristics and individual behavior, but also to historical forces, cultural frames and social institutions. By paying attention to period and business cycle effects, we are able to

capture economic, educational and cultural changes, which are assumed to affect labor market behavior.
We want to stress that the way life course theory is treated in the thesis resembles an agency-structure perspective: employment positions and changes are the result of the interrelationship between choices and structural constraints as it evolves over people's life courses. We argue that changes in the labor supply occur during the life course and over time (across cohorts) and could reflect either changes in preferences or in structure or institutions. Preferences might also be subject to changes over the life course due to changes in life conditions or institutions (Crompton and Feuvre, 1996; Ginn et al., 1996; Healy, 1999).

It should also be stressed that a comparative perspective underlines the whole thesis, which becomes apparent from different stages in the research: comparison is carried out between different instances of the same survey (Chapter 2), between surveys (Chapter 3 and 4), and across countries (Chapter 5).

3. Longitudinal Data: Potential from different survey designs

Labor market research in a life course perspective needs adequate methodological tools for empirical study and causal analysis. However, obtaining accurate and complete information about the past, so as to show the evolving structure of the life course and to link earlier and later events within individual life, is not easy (Scott and Alwin, 1998); the life course is a complex social process, in which several spheres and life domains are interconnected (Feldhaus and Huinink, 2006).
Time plays an important role in causal inferences. Identifying the link between earlier events and subsequent processes is a necessary step to understand labor market dynamics. To explain causal relationships, reducing causal fallacies, the whole process evolving over a period of time needs to be examined, while it is not enough to measure characteristics of actors at one point in time and predict outcomes at a later time point. This requires dynamic modeling and high quality data (Feldhaus and Huinink, 2006).

In the social sciences, cross-sectional observations are the form of data most commonly used to assess the determinants of behavior (Davies, 1994; Blossfeld and Rohwer, 1995). However, cross-sectional surveys, being conducted at just one point

in time, are inappropriate for studying social change and are not able to resolve issues of causal order.
Longitudinal data, tracking the same subjects over time, provide social researchers with powerful instruments to get to the heart of many processes of social change, ensuring a more complete approach to social empirical research.
The availability of such data certainly represents a good starting point for valuable scientific analyses. However, data reliability has to be assured for findings to be meaningful and trustworthy.

Prospective and retrospective data: Concerns about data reliability
Two distinct methods of collecting data suitable for life course analysis are 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 past behavior. 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 labor market panel surveys rely to some degree on recalled data for collecting information on labor market dynamics by asking respondents to retrospectively fill in the gaps between interviews. These two research designs overcome limits of cross-sectional data and help tackling the causality problem, allowing to relate the change in future outcomes to conditions in the past at each time point.

On the one hand, the situation regarding data looks bright: Research has profited from better theory and from new strategies of data collection, especially panel and retrospective survey designs. Quantitative longitudinal studies have made impressive progress during the last two decades. The "maturing" of longitudinal surveys based on nationally representative samples has led to an upsurge of innovative and substantive research and the progress in terms of data access has been matched by innovations and new directions in the methods applied to such data. The invention and dissemination of statistical techniques for analyzing longitudinal data as well as considerable improvements with regard to easy-to-use software contributed to successfully analyze the data. With respect to data analysis, cohort studies, event history analysis, refined methods of panel and event history analysis allow connecting different levels of analysis and different interdependent

dimensions of the life course (Feldhaus and Huinink, 2006). Sequence-pattern analysis and optimal matching[2] are instead successful in capturing the time dynamics of entire life trajectories.

On the other hand, worries about data reliability, which especially concern retrospective data, mainly due to memory bias, represent a challenge to be overcome for our analyses to be valuable. Therefore, we make use of (prospective and retrospective) longitudinal data in our substantive analyses, but we first address concerns about their reliability.

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 the research purposes. We employ life-history and panel data, evaluate possible sources of error with specific attention to memory bias, correct for them whenever possible, and look at the extent to which they are able to expose labor market dynamics and the underlying causal mechanisms. While in both the designs individuals are followed over time, which allows an explanation of change, the belief is generally that life-history data are less reliable than panel data due to memory bias. Although retrospective data definitely have more practical advantages, due to lower money and time constraints as well as a longer time horizon, during the last few decades several concerns have been raised about the reliability of retrospective data (see Dex, 1991 for a literature review). Besides problems related to substantial selection bias due to prior mortality, since estimates are only representative for the survivors, a major problem refers to the fact that when individuals are asked to retrieve events, situations and attitudes which happened in the past, the possibility of inaccurate reporting may occur (Babbie, 1973; Powers et al., 1978). Inaccurate long term recall could make retrospective surveys a very poor substitute for panel data, from which the importance to assess the extent to which retrospective data are affected by memory bias ensues.

Some concerns may be raised for panel surveys too. Panel attrition and conditioning may undermine population representativeness, while the multi measurement occasions make it quite complicated to re-arrange the discontinuous information about statuses, events and transitions given in the several waves into a continuous flow of behavior, introducing problems related to changing measurement and seam effects (Solga, 2001). Furthermore, since they also rely on some

[2] These techniques are explained in Chapter 2.

retrospective information to provide monthly data, factors affecting recall might cast concerns.

4. Structure of the book

In this thesis, we assess the factors which affect labor market dynamics and the possible causes of bias in different longitudinal research designs. The thesis examines the substantive effects of such factors as well as the consequences of biases; it attempts to empirically correct for measurement error and memory bias specifically. The focus is on employment careers, but the link with other life domains is carefully accounted for. We examine both participation and transitions on the labor market. Participation refers to the employment status while transitions to events on the labor market, and specifically to movements in and out as well as between different labor market positions.

Besides this introduction, the thesis consists of four essays (Chapter 2 to 5) and one concluding chapter. Although each of the four essays represents a self-contained article, they are related to one another, and their presentation follows a logical order.
We first address our methodological questions (Chapter 2 and 3), so as to provide a sound basis on which we can then build to answer our substantive questions (Chapter 4 and 5). Tackling the methodological issues concerning data reliability and (retrospective data) memory bias, we can model careers dynamics in a more refined way.

We initially deal with concerns about the reliability of retrospective data, modeling measurement error and isolating memory bias specifically.
In Chapter 2, we focus on retrospective data in the attempt to investigate memory bias. Using Swedish data, we look into retrospective reports of employment biographies with different recall length. First, we give a descriptive picture of differences between reports, which offers a first insight into possible errors. Afterwards, we quantify such differences through an algorithmic approach, which also allows us to investigate how they relate to both respondents' and careers' characteristics. Eventually, using a latent class Markov model, we take measurement error (at each response separately) into account, isolating memory bias and arriving at a sound estimation of the (true) latent state.

In the third chapter instead, we address concerns about data reliability comparing life history and panel data. Emphasis is put on the comparability of the data and the extent to which they are able to measure labor and job mobility patterns over time.
We first discuss advantages and disadvantages of both types of data, outlining the main differences between prospective and retrospective designs, reviewing the literature on recall bias. Afterwards, we use the GLHS (German Life History Study) and the GSOEP (German Socio-Economic Panel) to exemplify retrospective and prospective data respectively and to explain labor market transitions while correcting for memory bias in each type of survey. Specifically, we address the question about the similarity of inferences made from retrospective and prospective data, assessing to which extent the two sources 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.
We assess the extent to which differences in survey design between life history and panel data might affect the results of the analyses on labor market mobility patterns, with respect to the level of the estimated transition probabilities. Our contribution to the existing literature concerns the empirical investigation of differences between the two designs, which we perform controlling for specific sources of bias in each survey design. We specifically account for the underreporting of short unemployment spells, the length of the recall period and seam effects, besides residual design differences.
Three main research questions are addressed in the third chapter: the first concerns the 'true' measurement of transitions and the extent to which in life-course and panel data memory bias and seam effects might have a leveling-down and leveling-up effect, respectively, on the transition probabilities between jobs and on exit and (re)-entry into employment. A second question deals with the extent to which the differences are due to memory bias effects associated with the underreporting of short spells and to the longer length of the recall period in the life-course data. Furthermore, a third more substantive question pertains to the degree to which substantive results are influenced by design effects and to what extent prospective and retrospective designs point to the same determinants of transitions or causal mechanisms on the labor market after correcting for design effects.

This last question is specifically treated in the fourth chapter where, using the same two datasets, we examine substantial parts of the labor market careers for several birth cohorts covering the period from the late 1930s up to the year 2005.

In this chapter, we address a substantive research question about long-term changes in labor market mobility across birth cohorts and demographic life-stages, examining the effect of career investments and employment history across birth cohorts in particular, but without discarding structural influences. We focus on differences in employment volatility by educational level, the effect of familial responsibilities, of career investments and human capital on labor market dynamics.

Having assessed patterns and differences across data, we combine the two data sources to arrive at better estimates of the reasons for change. Controlling for possible sources of error as well as for compositional differences, we can combine the data to arrive at better estimates of change and its determinants. Correcting for design differences, we want to exploit the advantages of both prospective and retrospective data since they, combined, enlarge our scope for examining life course and birth cohort changes and extend our time perspective: Life course data, due to the long period they cover, help unravel the influence of age, period and cohort on the observed changes in transition patterns over time and capture to what extent, in which periods and for whom these changes exert their effects on labor market mobility patterns.

A more methodological question tackled in this chapter pertains to the degree to which the substantive results are affected by the survey design and whether different kinds of research designs lead to the same or different estimated effects of career determinants. Do prospective and retrospective data show different results with respect to the significance and the strength of the causal effects even after correction for differences in recall bias and design differences and, if so, to what extent? Particular attention is paid to the length of the recall period and the alleged anchoring effect of major life events such as marriage or childbirth. The assumption underlying the anchoring hypothesis is that respondents might use major life events such as marriage or childbirth to recall their employment histories, temporally connecting events in their lives, especially when they have to remember events over a long time horizon.

In chapters two to four we follow a 'within country' approach and track changes over time. In the fifth chapter, instead, we move to a cross-national perspective and we direct specific attention to the link between employment and demographic careers, given the alleged strong influence that the stage in the life course exerts on the employment biography of people. We concentrate on childbirth events for women, as

one of the main factors determining the stage in the life course and affecting employment history.

The long term perspective offered by life-history data is exploited to analyze such patterns for successive birth cohorts in Germany, the Netherlands and Great Britain. A cross-national comparison might allow for a better understanding of the mechanisms generating varying patterns of life course outcomes and is likely to be better suited for disentangling the linkages between institutional features of societies and life-course outcomes.

The main research questions of this chapter refer to the extent to which the birth of a first child affects women's labor supply, and the effect of subsequent childbirths. Attention is paid to differences across birth cohorts as well as across countries and gender-regimes. We address such questions using retrospective life-history data, given their suitability to study changes in people's labor market behavior over the life-course across different cohorts.

Overall, this thesis attempts to add to the current substantive and methodological literature. Methodologically, because it investigates and proposes new approaches to deal empirically with memory bias. Substantively, because it tackles more robustly the major question about the extent to which labor markets are changing and how that differently affects the behavior and transition patterns of people belonging to different cohorts. By building on a much sounder empirical basis, it arrives at much more robust results than it has been attempted before; first, by combining and comparing different data sources and, second, by modeling and correcting for measurement error associated with memory bias.

References

Babbie, E. (1973). *Survey Research Methods*. Belmont, California: Wadsworth Publishing Company

Bergemann, A., Mertens, A. (2004). Job stability trends, layoffs and transitions to unemployment: an empirical analysis for West Germany. Centre for Economic Policy Research (CEPR), London

Biemann, T., Fasang, A., Grunow, D. (2009). Did Economic Globalization Destabilize Careers? Evidence from Germany. *CIQLE Working Paper* 2009-02

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, 208-25

Blossfeld H. P., Rohwer, G. (1997). Causal inference, time and observation plans in the social sciences. *Quality and quantity: European journal of methodology* 31: 361-84

Bradburn, N. M., Rips, L. J., Shevell, S. (1987). Answering autobiographical questions: the impact of memory and inference on surveys. *Science* 236, 157-61

Breen (1997). Risk, Recommodification and Stratification Sociology. *The journal of the British Sociological Association*, 31 (3)

Brückner, H. and Mayer, K. U. (2005). The De-Standardization of the Life Course: What It Might Mean and Whether It Actually Took Place. *Advances in Life Course Research* 9: 27-54

Butz, W. P. and Torrey, B. B. (2006.) Some Frontiers in Social Science, *Science* 312 (5782)

Crompton, R. and Le Feuvre, N. (2000). Gender, family and employment in comparative perspective: the realities and representations of equal opportunities in Britain and France. *Journal of European Social Policy* 10: 334-348

Davies (1994). Wives and Work: The Sex Role Revolution and Its Consequences, *Population and Development Review*, 10 (3)

Dex, S. (1991). The Reliability of Recall Data: A Literature Review. *ISER Working Paper*. Colchester, University of Essex

DiPrete, T., Goux, D., Maurin, E., Quesnel-Vallee, A. (2006). Work and pay in flexible and regulated labor markets: A generalized perspective on institutional evolution and inequality trends in Europe and the U.S. *Research in Social Stratification and Mobility* 24, 311-332

Elder, G. H. Jr. (1994). Time, human agency, and social change: Perspectives on the life course. *Social Psychology Quarterly,* 57(1), 4-15

Feldhaus M, Huinink J. (2006). Panel Anlysis of Intimate Relationships and Family Dynamics (PAIRFAM) Why we need it and what it is about. *Pairfam Working*

Paper No. 6, Universität Bremen, online at http://www.pairfam.unibremen.de/fileadmin/user_upload/redakteur/publis/arbeitspapier_6.pdf

Ginn et al. (1996). Feminist Fallacies: A Reply to Hakim on Women's Employment, *The British Journal of Sociology*, 47 (1)

Gangl, M., (2006). Scar Effects of Unemployment: An Assessment of Institutional Complementarities. *American Sociological Review* 71, 986-1013

Goldthorpe, J. (2002). Globalisation and Social Class, West European Politics, 25

Hauser, R. M. and Featherman, D.L. (1977) *The process of stratification: trends and analyses*. New York Academic Press

Heinz W.R. and Marshall V. W. (2003). *Social dynamics of the life course: transitions, institutions, and interrelations*. Aldine De Gruyter, New York.

Healy, G. (1999). Structuring Commitments in Interrupted Careers: Career Breaks, Commitment and the Life Cycle in Teaching. *Gender, Work and Organization*, 6 (4): 185-201

Kohli, M. (2007). The Institutionalization of the Life Course: Looking Back to Look Ahead. *Research in Human Development* 4: 253-71

Lewis, J. (1992). Gender and the Development of Welfare Regimes. *Journal of European Social Policy, 2*(3): 159-173.

Lewis, J., T. Knijn, et al. (2008). Patterns of Development in Work/Family Reconciliation Policies for Parents in France, Germany, the Netherlands and the UK in the 2000s. *Social Politics: International Studies in Gender, State and Society* 15(3): 261-286.

Mathiowetz, N. A. and Duncan, G. J. (1988). Out of work, out of mind: Response errors in retrospective reports of Unemployment. *Journal of Business and Economic Statistics* 6, 221-229

Mayer, K. U. and Carroll, G. R. (1987). Jobs and Classes: Structural Constraints on Career Mobility. *European Sociological Review* 3, 14-38

Mayer K. U. (2000). Promises fulfilled? A review of 20 years of life course research. *Archives Européennes de Sociologie* XL1: 259-82

Mayer K. U. (2004). Whose Lives? How History, Societies, and Institutions Define and Shape Life Courses. *Research in Human Development* 1: 161 - 87

Mayer, K. U. (2008) New directions in life course research. In *CIQLE Working Paper* 2008-07

Mortimer, J. T. and Shanahan, M. J. (2003). *Handbook of the life course*. New York Boston Dordrecht [etc.]: Kluwer Academic/Plenum Publishers

Paull, G. (2002). Biases in the reporting of labour market dynamics. The Institute for Fiscal Studies, London

Powers, E. A ., Goudy, W. J. and Keith, P. M. (1978). Congruence Between Panel and Recall Data in Longitudinal Research. *Public Opinion Quarterly* 42, 380-389

Reimer, M. (2005). Autobiografisches Gedächtnis und retrospektive Datenerhebung: die Rekonstruktion und Validität von Lebensverläufen. *Studien und Berichte* No. 70. Berlin, Max-Planck-Institut für Bildungsforschung

Scott, J. and Alwin, D. (1998). Retrospective Versus Prospective Measurement of Life Histories in Longitudinal Research. In Giele, G. H., Elder, J. Z. (eds.) Methods of Life Course Research: Qualitative and Quantitative Approaches. Thousand Oaks et al.: Sage Publications

Solga, H. (2001). Longitudinal Surveys and the Study of Occupational Mobility: Panel and Retrospective Design in Comparison. *Quality and Quantity* 35, 291 - 309

Yang, Y. and Land, K. (2006). A Mixed Models Approach To The Age-Period-Cohort Analysis Of Repeated Cross-Section Surveys, With An Application To Data On Trends In Verbal Test Scores, *Sociological Methodology* 35, 75-98

Winkelmann, R. and Zimmermann, K. F. (1998). Is job stability declining in Germany? Evidence from count data models. *Applied Economics* 30, 1413-1420.

Chapter 2

Memory bias in retrospectively collected employment careers: a model based approach to correct for measurement error *

Abstract. In this paper, we investigate bias in retrospective reports of employment biographies, presenting an original and innovative method for such purpose. The Swedish Level of Living Survey (LNU) represents a very informative source to study error in retrospective survey data, in that it collects work histories for the same respondents in the same period at two different interviews carried out about 10 years apart. We use an algorithmic approach to compare, pair-wise, the two reconstructions, which permits us to arrive at a metric measure of their dissimilarity. Applying regression analysis to such measure, we show that the occurrence of spells out of employment and career complexity particularly affect report, while recall accuracy doesn't seem to be influenced by respondents' socio-demographic characteristics. Using a (latent) Markov model, we introduce measurement error at both responses so to arrive at a sound definition of the (true) latent state, which also informs us about error at each response. The results are informative on the extent and the nature of errors in retrospective reports: Error is confirmed to lead to simplification and conventionalism. Career complexity and spells out of employment seem to make recall particularly problematic, especially with increasing recall time, but already at very short time distances.

**This chapter was written together with Jeroen Vermunt, Ruud Luijkx and Ruud Muffels and will soon be submitted for publication*

1. Studying labor market dynamics using longitudinal data: on the reliability of retrospective surveys

Longitudinal data on careers provide much richer information than cross-sectional data since, by following the same respondents over time, they allow us to investigate career changes at the individual level. They allow us to study individual employment careers, their determinants, their differences among social groups, their changes over time, across cohorts as well as labor market and welfare contexts. Longitudinal data can be collected through panel studies, which repeatedly ask the same questions about people's current employment status, or through retrospective studies, which offer the possibility to ask to report at one occasion their complete (employment) history. Retrospective data provide a valuable source to study changes over time in that they trace, in a relatively fast and cheap way, entire individual careers for respondents belonging to different cohorts, covering a long period in historical time. Retrospective surveys usually collect information in the form of event histories, where respondents are asked to report the full sequence of episodes that they experienced up to the moment of interview, which constitute an indispensable form of longitudinal data for the analysis of employment biographies and certainly provide a valuable tool to observe respondents' life courses.

To generate valid and meaningful life-history data, respondents should correctly report their past employment histories, recalling and dating all the relevant episodes (and therefore transitions). However, retrospective surveys may suffer from memory bias, since information is based on the respondents' autobiographical remembering, in the course of which information might be forgotten or distorted, affecting reliability. Recall might turn out unreliable due to involuntary memory lapses but also to a voluntary modification of the reporting of events. Respondents' lack of memory accuracy may produce, voluntarily or involuntarily, an artificial sequence of episodes different from the real life course sequence. Memory's tendency towards reduction in complexity is usually expected to lead to transitions being more often forgotten than added.

Unemployment constitutes a special case, over which particular concern might be raised: a transition into unemployment is a profound and marked change in work-related social roles, activities and goals. However, findings from survey research show that episodes of unemployment have a particularly high risk of being forgotten (Elias, 1997; Paull, 1997; Dex, 1991; Reimer, 2004).

As a consequence, life courses might be misreported, which represents a threat to data accuracy and to the reliability of substantive findings. Completeness, accuracy and unbiasedness might be severely at risk when using memory-based reports, especially when the recall covers a longer period.
Recall bias endangers the reliability of such data and, for substantive research findings to be valuable, it is therefore important to rule out doubts on the reliability of results obtained with retrospective data.

In this chapter, we examine the quality of the reconstruction of employment biographies, as well as how such quality depends on the time lag between the events of interest and the interview. Life history data are excellently suited for studying changes over people's entire lives, and a deeper insight into their reliability may consequently be of great value for substantive analysis, since taking account of recall error will prevent researchers from drawing biased inferences. Moreover, once we know the type and amount of error, we may be able to correct for it.
The main research question concerns the occurrence of measurement error in the reconstruction of the sequences of episodes and the transitions between them in a retrospective survey due to memory bias. We wish to come to conclusions about whether retrospective reports on labor market experiences can be trusted or tend to bias analyses on labor market careers and, if so, in what way. We apply a variety of statistical methods, descriptive statistics on observed event sequences and matching techniques for comparing event sequences. In this way, we exploit the longitudinal nature of our data, improving over previous research in which cross-sectional data were used, often comparing survey data with administrative data or company records. The use of sequence analysis to study memory bias developing a measure of distances between different reports represents an original contribution to the previous literature.
Furthermore, we improve over the mainly descriptive ways in the current literature. We develop and apply a model-based approach to examine sequences of episodes, which certainly represents a very innovative approach to the study of memory bias. We make use of latent Markov models, in which - contrary to standard applications of this model- we relax the assumption that errors are independent across time points.

2. The Data

A major source for studying labor market dynamics in Sweden is the Swedish Level of Living Survey (LNU), in which retrospective information on work histories is collected.

The LNU, conducted by researchers at the Swedish Institute for Social Research (SOFI) at Stockholm University, is one of the longest running longitudinal social science surveys. It was first conducted in 1968 and then repeated in 1974, 1981, 1991 and 2000, using a panel approach. These surveys were carried out as face-to-face interviews with national probability samples of individuals between the ages of 18 and 75 (15 to 75 until 1981). The methodology of the survey has gradually developed and in 1991 a retrospective data section was added, where data on cohabitation, children, education and economic activities were recorded with a precision of one month duration (Jonsson and Mills, 2001). In 2000 the retrospective data from 1991 were updated, basically maintaining the same questionnaire.

In the working life biography section, individuals' economic activities are registered one by one in a temporal order, starting with the first real job, defined as a job that lasted six months or more. If respondents declare to have had a job that lasted for more than six months, they enter the economic activity history where, for each episode, they are asked information on the labor force status and end date of the episode,[3] with monthly precision in dates.[4]

Consequently, a respondent's employment history is a continuous sequence of employment episodes interrupted by episodes of unemployment and labor market inactivity: respondents move chronologically through mutually exclusive states.

When respondents are employed, questions on occupation, industry, firm size and sector are asked both in 1991 and 2000 while questions on part or full-time work, permanent or temporary employment, and whether the temporary employment is a part of labor market program are asked in 2000 only.

New respondents in 2000 as well as those who participated in 1991 but didn't answer the retrospective part have to report the whole employment history (from the first job

[3] The design assumes no gaps, since no information on start was collected, neither are parallel activities registered.

[4] In 1991 this information was collected only for respondents born between 1925 and 1965. Furthermore those who claimed to have had 15 jobs or more were diverted to a shorter section on occupations. We wanted to focus on employment (and not job) histories, and we consequently decided not to consider those who completed only the occupational history in our analysis, which means that we excluded respondents who declared to have had more than 15 jobs.

lasting at least 6 months), whereas for those who participated in the 1991 survey, the data collection in 2000 starts asking the activity in January 1990.
This allows us to reconstruct the whole career history of each respondent, until the last interview, offering a great source to analyze the entire career patterns of several birth cohorts over a long time period. Yet, the survey relies on retrospective reports, and errors in such reports represent a threat to the reliability of analyses based on such data.

As we explain in the next paragraph, the specific features of the data also allow a deeper insight into the importance of memory bias on reports of life courses and therefore on the reliability of life course data.
Understanding how reliable life course data are and what causes memory bias allows researchers to correct for it in the explanatory modeling phase to avoid misspecification and biased estimates.
This represents the starting point for a substantive analysis, where concerns on data reliability are dealt with and, hence, ruled out.

The overlap in the retrospective work histories

The Swedish Level of Living Survey provides a unique opportunity to study memory bias in retrospective reports. As previously mentioned, respondents who participated in the 1991 survey are re-interviewed in 2000 (when possible) and asked to report their employment biography starting from the activity they held in January 1990.
As a side effect, respondents report a second time on the period from January 1990 to the date of the first interview in 1991.[5] For respondents who participated in both the waves indeed, the economic activity data from the 1991 survey end at the interview date in that survey (between January and December 1991) and the economic activity data from the 2000 begin in January 1990, which generates an overlap between the information given in 1991 and in 2000, for a period in the years 1990 and 1991.
Therefore, the same respondents provide information about the same time period at two different interviews, which are carried out at different moments in time: in the first interview (in 1991) respondents recall what happened in their work life

[5] Although the period reconstructed is relatively short, the features of the Swedish labor market, characterized by an intermediate level of job mobility, that is predominantly upward in direction (DiPrete et al., 1997), weak occupational boundaries and organization-based labor market boundaries such that jobs are held to be less stable than in other countries and firms resort more quickly to layoffs—even of experienced workers—as a method of adjustment (Björklund and Holmlund, 1987; Büchtemann, 1993; Grubb and Wells, 1993; OECD, 1994) offer the analytical advantage of a sufficiently large number of difficult cases in which memory error can be observed.

approximately one year before, in the second one (in 2000) the reconstruction concerns approximately the 10 years period before the interview, meaning that the reference period (1990 and 1991) is the same as in the first interview. Figure 1 illustrates the data structure.

Figure 1. Representation of the data structure

year	…	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
restrospective distance																	
interview 1 (1991)	…	6	5	4	3	2	1	0									
interview 2 (2000)							10	9	8	7	6	5	4	3	2	1	0

While this could generate a problem if we want to obtain complete employment histories for respondents who participated in both the interviews (due to the need to match information from multiple sources), it also offers a great source to study the reliability of such reports. The data allow us to investigate whether and how information gathered at different time points is affected by error: besides being able to identify inconsistencies in the two reports, we have a greater amount of information to identify true employment patterns. Inconsistencies between reports for the reference period can be attributed to inconsistent autobiographical reconstructions and to the greater recall difficulties in the second interview since about 10 years have elapsed in 2000 for the recall of events in 1990/1991.

As explained, the data are collected in the form of event histories, where respondents have to recall and date the full sequence of subsequent episodes in their life course.
The state space contains four main states, which define respondents as: employed (E), self-employed (S), not employed (N) or unemployed (U). During the recall process respondents have to label episodes, describing their main activity, and reconstruct transitions between episodes by providing the month and year in which each episode ends. This occurs through the subjective reconstruction by the respondent, which might be affected by (and adapted to) normative considerations about the logic of a life course and the individual's self perception of that at the time of recall.
Errors may arise when respondents recall their status at a given point in the past and when they reconstruct a transition by dating the end of an episode. They may completely forget episodes but also erroneously add new ones and consequently artificially increase the number of transitions. Since the reference period is limited in time, misdating the end of an episode could mean that the whole episode (and the transition between episodes) is moved to a date before or after the reference period.

Incorrectly dating the occurrence of an event may lead it to being reported as having occurred before or after our observation window (i.e. before January 1990 or after the interview in 1991), which would lead to the loss of episodes within our reference period.

3. Analysis and Results

Respondents may report their employment biographies differently when asked at different moments in their life and at different time distance from the occurrence of the events.

The availability of reports of employment biographies of the same respondents given at two different (retrospective) interviews for the same reference period offers a possibility to study the reliability of retrospective reports, and allow us to get a grip on differences in recall patterns for different length of the recall period.

If the same respondents are re-interviewed and no changes occurred in the survey format, which is the case in our data, differences in the sample as well as in the survey design and questionnaire format are controlled for. The two sources of data differ instead in the length of the recall period (about 1 year against about 10 years).

We could reasonably assume that the more similar the reports are, the more reliable they are, even though they might both contain errors. When differences are recorded instead, in (at least) one case there were mistakes in the reconstruction, which decreases the reliability.

Inconsistencies in respondents' careers may be attributed to inconsistent autobiographical reconstruction and to the greater recall difficulties at the second interview, since about 10 years elapsed between the interviews. Investigating the patterns of inconsistencies should inform us on the quality of the recall over time. Since memory is presumed to decay over time, we contend that recall error increases with longer recall periods. For that reason, it seems reasonable to put more trust in the reports that are closer to the interview (i.e. the 1991 report).

This means that we see differences between the two interviews as entirely caused by memory bias due to the length of the recall period and assume that other sources of bias not related to memory have no impact, which may not be the case. Other external factors generating differences (like the conditions in which the interview took place or data quality/coding) cannot be fully ruled out, and measurement error might (differently) affect both the reports to some extent. Error might indeed occur at

each interview and reasons (related for example to the interview context) might also exist which lead to greater bias in the 1991 response (closer to the event), which puts into question the assumption that the 1991 report contains no error. Both assumptions, either that the 1991 report is perfect or that it also contains error, will be included in our model-based approach.

We start with an overall description of the reports at the two interviews based on the number of reported episodes and transitions at each interview.

Table 1. Number of reported episodes and transitions, by interview

	1991	**2000**	% reduction in 2000
Episodes			
Employment	1732	1652	4.62
Self-Employment	155	127	18.06
Non employment	534	474	11.24
Unemployment	96	55	42.71
Total	2517	2308	8.30
Transitions	544	335	38.42

As we can see in Table 1, it appears that, at the aggregate level, the number of episodes reported drops at the second interview (2000). The reduction is particularly sharp for unemployment episodes.

After presenting this aggregate picture, we move to a sequence perspective (par. 3.1), where we exploit the longitudinal nature of our data. At this point, we compare each reported individual career sequence as reported by the respondent at the two interviews and develop a metric measure (par. 3.2) quantifying their (pairwise) distance, using the report in 1991 as reference.
Eventually, we use a latent class Markov (LM) model to control for measurement error and investigate memory bias (par. 3.3). The latent aspect of the model, which refers to the true state, that cannot be observed, is at the individual level, which explains the association in time points for one person. This is a model for each respondent, which looks at all the information (sequence) for each respondent (i.e. looks at the vector of all the states). Having two indicators for each observation we can, under the Markov assumption, make inferences on memory bias allowing

dependency across time in the classification error and define a 'true' (latent) state which might serve as a reference point to which to compare each observed report.

3.1 Comparing sequences of episodes

Starting from this simple descriptive account of differences in the two reports, we want to have a deeper insight into the nature of the (mis)reports in the career histories exploiting the longitudinal nature of our data. We now consider as unit of analysis the sequence of episodes in the biography of each respondent and the interest is in the sequential character of all events together, which means on the career trajectory.

Sequence data are present in the social sciences to investigate life courses, marital histories or employment biographies. Generally spoken a sequence is an ordered list of elements, whereas an element is a certain status (e.g. labor market or marital status) and a set of consecutive elements of the same type defines an episode.

Here, we use sequence-based methods to compare the reports given nearly 10 years apart. We look at the whole trajectory trying to identify differences between the two reports, aiming at unraveling in which cases and how recall affects the career sequence reconstructed. Further digging into the pattern of differences in the employment history reconstruction and into its association with respondents' and careers' characteristics offers an insight into the occurrence and strength of memory bias.

The analyses are based on the reports of 1973 respondents, of which 47% are men and 53% women. For these respondents, the two measurements in 1991 and 2000 allowed us to reconstruct a period between 3 and 22 months, on average 15 months, covering the period January 1990 (the earliest possible date) and November 1991 (the latest recorded interview date in 1991).[6] The average length of the recall distance is 8 months when using the 1991 survey and nearly 10 years when using the 2000 survey. Respondents in 2000 therefore reconstructs events which happened much farther away in the past.

[6] A few respondents reported their activity only from a date later than January 1990, although this was the date from which they were asked to reconstruct their work biography in the questionnaire. Consequently, a few respondents reconstructed less than 13 months –i.e. from January 1990 to January 1991, being January 1991 the earliest possible interview date for the 1991 survey-.

Table 2. Descriptive sequences characteristics, by interview

	1991		**2000**		% reduction in 2000
	mean	St. dev	mean	St. dev	
Number of different elements	1.21	0.44	1.15	0.377	4.96
Number of episodes	1.28	0.64	1.17	0.45	8.59
Number of employment episodes	0.88	0.44	0.84	0.41	4.55
Number of unemployment episodes	0.05	0.24	0.03	0.18	40.00
N	1973		1973		

In Table 2 we provide a descriptive comparison of the sequences. Employment careers appear less heterogeneous according to the report in the second interview (2000) than in the first (1991): Fewer different elements (i.e. states) and fewer episodes seem to be reported on average in a career sequence when the reconstruction takes place longer after the occurrence of the events (in 2000).

Table 3 instead shows how exactly each type of sequence is reported at the two occasions.

Table 3. (Ten most common) Sequences as reported in 1991 and 2000

sequence in 1991	***sequence in 2000***											
	E	**N**	**S**	**EN**	**NE**	**EU**	**ES**	**ENE**	**EUE**	**NEN**	***else***	**N**
E	**95.5**	0.8	0.5	1.2	0.6	0.3	0.2	0.3	0.2	-	*0.4*	1282
N	5.3	**85.6**	-	5.3	2.3	-	2.7	-	-	-	*1.5*	188
S	5.4	9.0	**79.3**	-	-	-	-	-	-	-	*3.6*	111
EN	14.2	18.1	-	**59.1**	1.6	-	-	1.6	-	3.1	*2.3*	127
NE	29.3	7.3	1.2	1.2	**48.8**	-	-	4.9	-	3.7	*3.6*	82
EU	52.4	-	-	9.5	-	**19.1**	-	4.8	-	-	*14.2*	21
ES	52.9	5.9	-	-	-	5.9	**35.9**	-	-	-	-	17
ENE	54.6	-	3.0	3.0	6.1	3.0	-	**24.2**	3.0	-	*3.1*	33
EUE	64.3	-	-	-	14.3	-	-	-	**7.1**	-	*14.3*	14
NEN	-	35.7	-	35.7	14.3	-	-	-	-	**14.3**	-	14
else	*5.3*	*85.64*	-	*5.3*	*2.3*	-	-	-	-	-	*1.5*	84
N	1356	230	103	111	69	11	11	21	6	9	46	1973

Continuous sequences (one episode), and hence less complex sequences, seem the ones less often mistakenly reported, and employment seems the better reported state. Being the sequences ordered from the most simple to the most complex,[7] higher numbers under the main diagonal indicate that in the second report a simplification of sequences occurs. Later reports (i.e. from the interview in 2000) seem to involve less change and describe simpler careers.

[7] The order in which sequences are presented is based on the number of different episodes in the trajectories and their frequency.

These findings seem particularly relevant given the current debate about the increasing (in)stability[8] on the labor market, since they suggest that retrospective reports might simplify the earlier part of career patterns (reporting fewer transitions in earlier life than there were in reality) and thus overestimate[9] this trend. To further address such concerns, we employ the recently developed concept of turbulence (Elzinga, 2007a; Elzinga, 2007b; Elzinga and Liefbroer, 2007), which refers to the number of employment changes, the predictability of the order of jobs, and the variability of the durations spent in different employment states. Elzinga's measure takes into account the order of career states as well as their duration variation in different career states.[10]

We have therefore computed (and confronted) this measure of turbulence for the two responses, which should indicate whether the estimated turbulence of a career might be affected by the retrospective nature of the survey data and specifically might be biased downwards by a longer recall time.

Table 4. Sequence turbulence, by interview

	turbulence	
	mean	bootstrap C.I.
1991	1.698	(1.632 - 1.763)
2000	1.481	(1.417 - 1.545)

The results of our analyses, presented in Table 4, suggest an underestimation of turbulence with increasing recall time. Turbulence at about 10 years distance from the occurrence of the events turns out significantly lower than at 1 year (or less) distance, as proven by the 'not overlapping' bootstrap confidence interval (calculated

[8] See the introductory chapter in the thesis for a more extensive discussion on the current debate on the increasing instability of employment careers..

[9] If increasing length of the recall period leads to more simplified careers, career volatility is underestimated especially for the very past and less for the time closer to the interview, which would overestimate the trend towards increasing instability over time.

[10] As a measure of turbulence taking durations into account, Elzinga (2007b: 33) proposed using

$$0 \le T(x,t) = \log_2\left(\phi(x)\frac{s^2_{t,\max}+1}{s^2_t+1}\right) \qquad [2.1]$$

with $\phi(x)$ denoting the number of distinct subsequences, s^2_t denoting the variance of state durations and $s^2_{t,\max}$ the maximum of that variance given the total duration of the sequence.

T(x, t) is therefore a sequence property quantified such that the sequencing of the states and the variance of their durations is taken into account; it is not sensitive to the specific time scale used and it increases with decreasing variance of the durations of the states (Elzinga and Liefbroer, 2007).

with the bias corrected and accelerated method, BC*a*) for the mean turbulence (Efron and Tibshirani, 1993). This is relevant in that the length of the recall period might affect the findings of substantive studies looking at employment dynamics over time.

3.2 Sequence similarity: Algorithmic approach

A further step consists in not simply looking at differences and their patterns but also in trying to quantify them. An algorithmic approach may be used to define a measure for the differences between pairs of sequences. In a following step, it is possible to use regression analysis to study what affects such differences in the reports, paying attention to both respondents' attributes as well as career characteristics.

In comparing sequences, the first question concerns how the difference between (pairs of) sequences should be measured. Different strategies and algorithms have been developed to do that.

The most known method of sequence analysis in the social sciences is optimal matching (OM) (Abbott, 1995; Abbott and Forrest, 1986). It compares sequences using the so called Levenshtein distance, a measure based on the minimal cost (i. e. the most efficient set of operations) to transform one string into another using a series of "elementary operations" of insertion, deletion and substitution. The Needleman-Wunsch algorithm finds this alignment (for a more detailed explanation of OM see Abbott, 1995; Abbott and Forrest, 1986; Abbott and Tsay, 2000)

In OM, the costs of transformation operations have to be defined, implying assumptions about the relative cost of substituting, inserting and deleting. Substitution costs between elements (which in our application are represented by different labor market states) have to be explicitly parameterized, which involves in general serious theoretical considerations. [11] We employ optimal matching to compute the distances between pairs of sequences reported at the two interviews.

[11] This step in OM is always arbitrary to a certain extent and in the social sciences a theory is not always sufficiently precise to inform us about appropriate cost specification, which may cast worries that there is no theoretical basis to determine the substitution costs matrix and therefore no way to justify the analysis. Theoretical considerations about the similarity between elements may guide the definition and differentiation of substitution costs. Alternatively, it has been proposed that they are fully derived from the data themselves based on the frequency of transitions from one state to another, with lower costs for high transition frequency, based on the intuition that two statuses are less different when there are more transitions between them. We didn't use such specification of the subcost matrix in our analyses since we believe that the underlying assumption that more frequent transitions between states indicate greater similarity doesn't hold in our application: we consider for

Such distance measure provides an indication of how much the report at the second interview differs from the report at the first interview. The measure itself is not very informative. However, it can be used as a dependent variable in a regression model to assess whether the extent of differences between the sequences reported at different points varies across respondents according to their characteristics or to the attributes of their career sequences.

In Table 5, we present the results of a regression analysis on the distance between the two observed reports. The first column refers to the OM distance, computed using a subcost matrix in which all the states are considered equally different and substitution costs between all pairs of states are assigned a cost equal to the double of the 'indel' cost (i.e. the cost of inserting or deleting elements), which is set to one (this is the default subcost matrix in Stata).

Table 5. OLS regression on 1991-2000 sequence (standardized) distance (OM)

	latent-observed	(a) latent-1991	(b) latent-2000	sig. of (a)-(b)
Female	-0.011	-0.015*	0.007	
Birth year	0.0001	0.0005*	-0.0002	
Episodes out of employment	0.105**	0.049**	0.162**	**
Career complexity	0.118**	0.03**	0.21**	**
Later report (2000)	0.145**			
Constant	-0.271**	-0.998*	0.593	
N	3946	1973	1973	
R squared	0.14	0.09	0.14	

Our results indicate that gender and age [12] do not affect sequence dissimilarity. The occurrence of episodes out of employment and career complexity,[13] instead, have a strong positive effect on sequence distance, indicating that recall accuracy is lower when respondents experience more episodes and transitions in the course of their employment history and when it involves episodes out of employment in particular.

example employment more similar to self-employment than to not employment, although we observe fewer transitions in the first case.

[12] The time of the interview being fixed, birth cohort also indicates the age at interview, where later cohorts –i.e. increasing birth year- are younger at the time of interview.

[13] Career complexity was operationalized as the number of episodes according to the 1991 report and standardized to the number of months that are observed calculating the degree to which a respondent's number of episodes deviates from the average number of episodes in his/her 'length of observation group,' where three groups were separated according to the time for which the respondent was observed.

Trusting the first report (1991) more, these results also suggest that in such circumstances recall is more problematic when the recall period is longer.[14]

However, we acknowledge that OM has been subjected to many critiques, especially for the way it (doesn't) handle duration and because it is not able to account for the order of events and the direction of time. Some aspects in OM might be particularly problematic. The subcost matrix in OM has to be symmetric which, in our case, means that reporting of being employed when actually unemployed is implicitly considered be equal to reporting of being unemployed when actually employed. Moreover, such approach doesn't handle right censoring, neglects nonlinear time dependencies and lacks concern with covariates, for which there is no obvious place (Halpin, 2003). However, as it will become clearer later, these shortcomings constitute more of a problem for substantive analysis, and less for our methodological application.

Several alternative techniques [15] to the standard OM have been proposed. Yet, our results turn out robust not only to different specifications of the subcost matrix [16] but also to alternative techniques. In the second column of Table 5, we prove this showing the results of a regression analysis on a distance measure computed using one of the latest developed techniques for sequence comparison, i.e. the Elzinga's measure (Elzinga, 2007b), a non-alignment technique that accounts for the order of events within sequences avoiding arbitrary specifications of transformation costs, since it is not based on transformation operations as in OM. [17]

[14] We also tried to include education in the model (the results of which are available from the first author upon request), but the effect was found not significant. Due to a slight loss of power resulting from a few cases with missing information on education, we preferred to present the result of the model not including education. However, the effects of the considered variables turned out to be very similar.

[15] Some proposed alternative techniques to standard OM are: *OMAV*, an optimal matching technique adjusting for duration; *Degenne*, a method focusing on comparing vectors of cumulated duration at each time point; *Hamming* distance, which compares sequences element by element such that the inter-sequence distance is the sum of the element-wise distances; *Elzinga*'s technique, which calculates the similarity between the sequences based on identical subsequences.

[16] The results of such analyses are not shown here but are available from the first author upon request.

[17] This technique is based on a different metric and handles durations in a more consistent way: it calculates the similarity between the sequences based on identical subsequences. Distance is defined as the number of non-common subsequences, weighted by the frequency with which they occur in either sequence and their duration. It has advantages concerning duration and the importance of the order of non-adjacent states. The algorithm consists in an efficient way of counting the number of common n-tuples in pairs of sequences.

Such an algorithmic approach doesn't make assumptions about the processes that generate the data, keeping it in this way at the exploratory level. It might develop in time into a stochastic technique with which we can formally test hypotheses and make inferences, although a lot of research should be still done on that and this is therefore not likely to happen soon (Biemann et al., 2009).

Furthermore, most sequences of interest to social scientists [18] do develop stochastically through time, where earlier events determine the probabilities of occurrence of later events (path dependency). That is, for most sequences we will rightly be more interested in the processes that created them than their actual occurrence through time (Halpin, 2003: 9). These drawbacks led us to look into the issue with a different approach, which may hopefully help us to overcome them.

3.3 Model-based approach

To overcome several of the shortcomings of the algorithmic approach presented above, we now investigate the recall bias problem using a model-based approach. The two key requirements of the stochastic model for modeling the two sets of employment status reports are that it should be suited for (1) modeling event sequences or discrete-time longitudinal data on transitions across a finite number of states and (2) dealing with measurement error in the recorded states, which in our application also involves combining the possibly inconsistent information in the 1991 and the 2000 measurements of the month-specific labor status.

A suitable stochastic model is the Markov model, which yields estimates for the probability of making a transition between each pair of states of the outcome variable of interest. While a simple Markov model solves the problem associated with the OM approach that adjacent time points are treated as somewhat disconnected, it does not tackle the measurement error issue discussed above. However, there exists a more extended variant of the Markov model called the latent (or hidden) Markov model that can be used for this purpose. Here, "latent" or "hidden" refers to the fact that a person's true state at time point *t* is unknown, and it should thus be treated as a latent variable. In a latent Markov model, the true states are the categories (classes) of a dynamic latent (unobservable) variable. The latent

[18] The evolution of patterns though time (which can refer to several concepts, like age, period cohort or life stage) are especially a matter of concern within life course sociology, which is the basic framework in which this thesis is embedded.

states are connected to the observed states using a model similar to a latent class model, which defines the measurement error component of the latent Markov model.

There exist various examples of applications of latent Markov models for the analysis of discrete-time longitudinal data subject to measurement error (Collins and Wugalter, 1992; Hagenaars, 1990; Poulsen, 1982; Van de Pol and Langeheine, 1990; Vermunt, Magidson and Tran, 2008; Magidson, Vermunt and Tran, 2009). In our application, we use the latent Markov model to estimate the error in the recorded employment status at the two interviews, which we believe will yield more interesting information on the pattern and amount of recall bias in the two retrospective reports than was obtained with the descriptive and algorithmic approaches. It should be noted that the latent Markov model not only yields estimates of transition and measurement error probabilities, but that it can also be used to obtain (posterior mode) estimates of the true state occupied at each time point. Hence, whereas in OM we could only compare the distance of the 2000 report to the 1991 report, which was in fact used as a kind of perfect measure, the latent Markov approach allows us to compare each of these two reports with the estimated true sequences, and thus to relax the assumption that the information from the 1991 interview is free from errors.

Typically, latent or hidden Markov models are applied with a single report per respondent per time point. However, our data consist of two (imperfect) reports of a respondent's employment status per time point (per person-month). As it is explained in more detail below, this makes it possible to relax certain key assumptions of the basic latent Markov model which are unrealistic in our application.

We denote the 1991 and 2000 report of the employment status of person i at month t by y_{it}^{1991} and y_{it}^{2000}, respectively. Note that y_{it}^{1991} and y_{it}^{2000} are categorical variables with M=4 categories;[19] that is, $1 \leq y_{it} \leq M$. The total number of time points (months) for which person i provides information is $T_i + 1$, with $0 \leq t \leq T_i$.[20] The observed data for a particular respondent consist of two response vectors of length $T_i + 1$ denoted by $\mathbf{y}_i^{1991}$ and $\mathbf{y}_i^{2000}$, respectively. The probability of having the observed set of responses is denoted by $P(\mathbf{y}_i^{1991}, \mathbf{y}_i^{2000})$. Note that $P(\mathbf{y}_i^{1991}, \mathbf{y}_i^{2000})$ is the probability for which we are going to define a statistical model.

[19] The four categories represent the four states: 1=employed (E), 2=self-employed (S), 3=not employed (N), 4=unemployed (U).

[20] The number of months for which we have employment information varies across respondents, which is why we use the index i.

Besides the observed responses, the latent Markov model for $P(\mathbf{y}_i^{1991}, \mathbf{y}_i^{2000})$ contains $T_i + 1$ discrete latent variables, each having K categories or latent states. We denote the true (latent) state at time point t by x_t, where $1 \leq x_t \leq K$. In the current study, K=M=4; that is, the number of latent labor force states is assumed to be equal to the number of observed labor force states.

The latent Markov model has the following form:

$$P\left(\mathbf{y}_i^{1991}, \mathbf{y}_i^{2000}\right) = \sum_{x_0=1}^{K} \sum_{x_1=1}^{K} \ldots \sum_{x_T=1}^{K} P(x_0) \left[\prod_{t=1}^{T_i} P\left(x_t \mid x_{t-1}\right)\right]\left[\prod_{t=1}^{T_i} P\left(y_{it}^{1991} \mid x_t\right) \prod_{t=1}^{T_i} P\left(y_{it}^{2000} \mid x_t\right)\right] \qquad [2.2]$$

The unknown model probabilities to be estimated are the initial latent state probabilities $P(x_0=s)$, the latent transition probabilities $P(x_t = r \mid x_{t-1}=s)$, and the classification error probabilities $P(y_{it}^{1991}=\ell \mid x_t=s)$ and $P(y_{it}^{2000}=\ell \mid x_t=s)$.[21] In the Latent GOLD software (Vermunt and Magidson, 2008), that we used to estimate our models, these model probabilities are parameterized using multinomial logistic equations:

$$\mathrm{P}\left(x_0 = s\right) = \frac{\exp\left(\alpha_s\right)}{\sum_{k=1}^{K} \exp\left(\alpha_k\right)} \qquad [2.3]$$

$$\mathrm{P}\left(x_t = r \mid x_{t-1} = s\right) = \frac{\exp\left(\gamma_{rs}\right)}{\sum_{k=1}^{K} \exp\left(\gamma_{ks}\right)} \qquad [2.4]$$

$$\mathrm{P}\left(y_{it}^{1991} = \ell \mid x_t = s\right) = \frac{\exp\left(\beta_{\ell s}^{1991}\right)}{\sum_{m=1}^{M} \exp\left(\beta_{ms}^{1991}\right)} \qquad [2.5]$$

$$\mathrm{P}\left(y_{it}^{2000} = \ell \mid x_t = s\right) = \frac{\exp\left(\beta_{\ell s}^{2000}\right)}{\sum_{m=1}^{M} \exp\left(\beta_{ms}^{2000}\right)} \qquad [2.6]$$

Under the identifying restrictions[22] $\alpha_1 = 0$ and $\gamma_{11} = \gamma_{22} = \gamma_{33} = \gamma_{44} = 0$ and $\beta_{11} = \beta_{22} = \beta_{33} = \beta_{44} = 0$, we get the following rather easy interpretation of the logit parameters:

[21] Note that with s and r we refer to a particular latent state and with ℓ to a particular observed state.

[22] Because these are multinominal logit models, we need one constraint on α, four on γ (one for each value of x_{t-1}) and four on β (one for each value of x). We imposed constraints in such a way that one category of the dependent variable in the equation concerned is treated as reference category, which gives the parameters an interpretation that makes very much sense in latent Markov models. In the transition and error codings this reference category changes with the value of the independent variable in the equation.

$$\alpha_s = \log \frac{\mathrm{P}\left(x_0 = s\right)}{\mathrm{P}\left(x_0 = 1\right)} \qquad [2.7]$$

$$\gamma_{rs} = \log \frac{\mathrm{P}\left(x_t = r \mid x_{t-1} = s\right)}{\mathrm{P}\left(x_t = s \mid x_{t-1} = s\right)} \qquad [2.8]$$

$$\beta_{\ell s}^{1991} = \log \frac{\mathrm{P}\left(y_{it}^{1991} = \ell \mid x_t = s\right)}{\mathrm{P}\left(y_{it}^{1991} = s \mid x_t = s\right)} \qquad [2.9]$$

$$\beta_{\ell s}^{2000} = \log \frac{\mathrm{P}\left(y_{it}^{2000} = \ell \mid x_t = s\right)}{\mathrm{P}\left(y_{it}^{2000} = s \mid x_t = s\right)} \qquad [2.10]$$

that is, as

- the log odds of having initial state s rather than 1 [2.7]
- the log odds of making a transition from state s to state r rather than staying in state s ('transition coding') [2.8]
- the log odds of misclassifying someone with latent state s by assigning observed state ℓ rather than the correct state s. We estimate a logit for error for each response ('error coding') [2.9 for 1991 response and 2.10 for 2000 response].

Note that 2.8 and 2.9/2.10 follow from the fact that the "no transition" and the "no error" categories, respectively, serve as the *reference* categories; that is, from the fact that the reference category changes with the origin state (true state).

The model described in equation [2.2] is the basic latent class Markov model. It makes certain assumptions about 1) the way people move between latent states (the transitions) and 2) the way the observed states are reported given the true states (the measurement error). The main assumption about the (latent) transitions is that these are in agreement with a first-order Markov process, which implies that the state occupied at time point t depends only on the state occupied at time point t-1 (the previous state), and that it is thus independent of the states occupied at all other time points. In addition, it is assumed that transition probabilities are time homogeneous and that there are no observed or unobserved factors affecting the transition probabilities. Although it would be possible to relax each of these sets of assumptions, we keep the model for the transitions as simple as possible since our main interest is not in the parameters of a transition model but in the measurement error parameters, as well as in posterior estimates of the respondents' true states. It turns out that this information of main interest is not (or very weakly) affected by the model we use for the transitions. This is because we have two measurements of a

person's true state (and not just one) and because we have a rather long time series. We obtain very good (almost perfect) estimates for the true event sequences, which cannot be improved by a more sophisticated transition model. Relaxing the first-order Markov assumption, adding covariates (including time) in the model, and modeling unobserved heterogeneity using a time constant latent variable will improve the prior estimates of the transition probabilities for a specific respondent, but will not affect the estimates of the measurement error nor of the posterior estimates of a person's true state. Therefore, we decided to keep our models as simple as possible as far as the transition part of the model is concerned.

The two key assumptions in the measurement part of the basic latent Markov model are that, conditional on a person's true states, the 1991 and 2000 responses at time point *t* are 1) independent of one another and 2) independent of the responses at the other time points. In latent class analysis, the former assumption - which is connected to the fact that we have two measurements - is referred to as "local independence". The second assumption is sometimes referred to as "independent classification errors" (ICE). It is especially the ICE assumption that turns out to be problematic in our application, which is why we propose various modifications of the basic latent Markov model for relaxing this assumption.

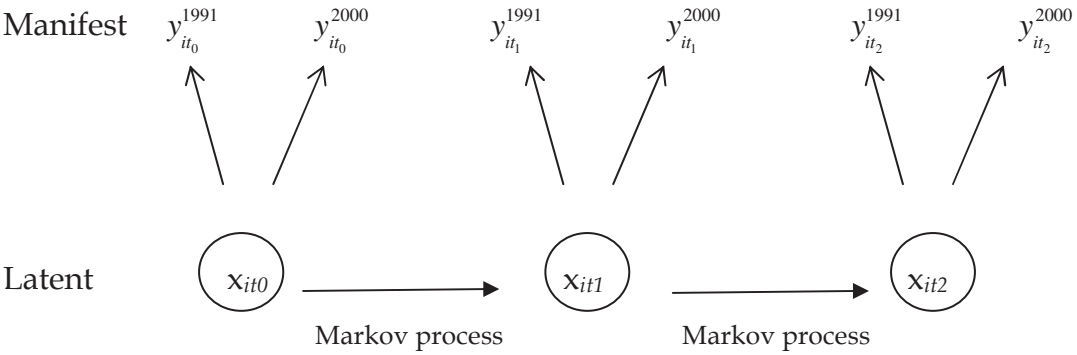
In our application, we used three kinds of specifications for the measurement error component of the latent Markov model. Figure 2 provides a graphical representation of the models.

Model 1, which serves as the baseline model, uses the rather strong ICE assumption (we label it 'ICE' model). In addition, we assume the response at the 1991 interview to be perfect, which is similar to what was done in the OM approach, which quantified how well the 2000 sequences match the 1991 sequences. Note that the probability of making a classification error can be fixed to zero by equating each of the logit parameters $\beta_{\ell s}^{1991}$ appearing in equations 2.5 and 2.9 to a large negative value (e.g. -100) for $\ell \neq s$.

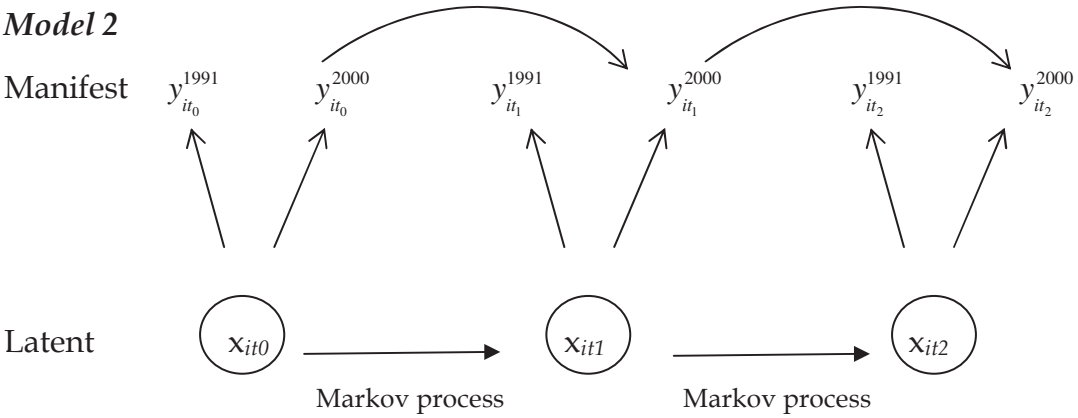
Figure 2. Models Graphical Representation

For t=0, 1, 2, ...T

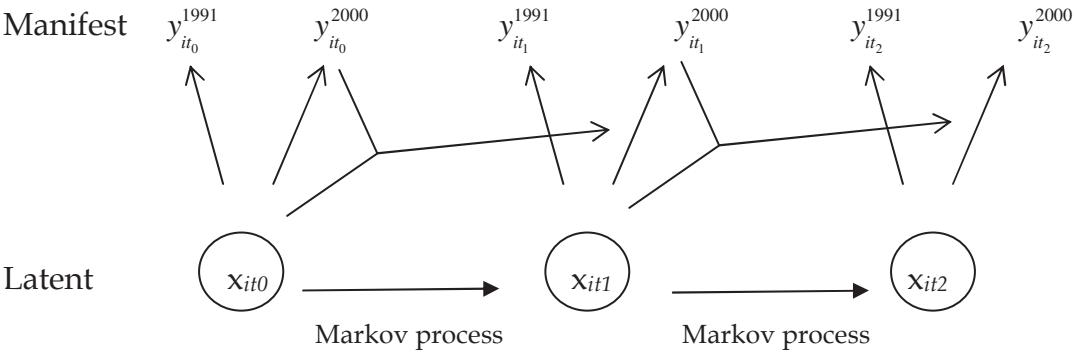
Model 1



Model 2



Model 3



We also estimated an ICE model that does not assume that the 1991 response is perfect. This, however, leads to the very unrealistic result that the 2000 report is much better than the 1991 report. This is due to an artifact of the basic latent Markov

model; that is, that it will see the larger stability of the 2000 series as an indication that there is less measurement error. This problem, which is caused by the ICE assumption, becomes visible when both the 1991 and 2000 measurement error probabilities are estimated freely. It can be solved by relaxing the ICE assumption, which is feasible in our application because we have two measurements of a person's state at each occasion.

The reason why the ICE assumption is unrealistic in our application relates to the fact that in a retrospective survey, during the recall process, a respondent could forget or misplace episodes, which may be more likely with increasing recall time. The consequence of these kinds of mistakes is that the same mismatch between true and observed state (the same type of error) occurs at consecutive months. As a result, errors at consecutive time points are not independent of one another, and the ICE assumption is thus clearly invalid. Below, we describe various ways to relax the ICE assumption for the 2000 report. Although it is unlikely that such correlated errors are fully absent in the 1991 report, it is reasonable to assume that these are much weaker when reporting over the previous year than when reporting about approximately ten years ago. If we were to allow dependent classification errors for both responses, we would run into the risk of fully distorting the structure of the latent Markov model, as well as of running into identification problems.

Model 2 (which we label 'state dependent classification error' model) differs from the ICE Model 1 in that it treats the classification error probabilities for 1991 as free parameters (rather than fixing them to zero) and that it relaxes the ICE assumption for the 2000 interview. The latter is achieved by making the 2000 reported labor status for month *t* depend on the observed status in the previous month (y_{it-1}^{2000}). This non-ICE latent Markov model is obtained by replacing the term $P(y_{it}^{2000} = \ell \mid x_t = s)$ in equation [2.2] with $P(y_{it}^{2000} = \ell \mid x_t = s, y_{i,t-1}^{2000} = k)$. This means that there are four sets of free error matrices to be estimated, one for each value of *k*, which can take on the values 1 to 4. Without additional constraints this would yield 48 (4 times 12) free measurement error parameters. A more parsimonious specification is obtained using a logistic model for y_{it}^{2000} with additive effects of x_t and $y_{i,t-1}^{2000}$; that is,

$$\log \frac{P(y_{it}^{2000} = \ell \mid x_t = s, y_{i,t-1}^{2000} = k)}{P(y_{it}^{2000} = s \mid x_t = s, y_{i,t-1}^{2000} = k)} = \beta_{\ell s}^{2000} + \lambda_{\ell k}. \qquad [2.11]$$

In other words, the logit equation for the 2000 response also contains the lagged 2000 observed state (y_{it-1}^{2000}) as an additional predictor. Compared to the ICE model, we have 9 additional free $\gamma_{\ell k}$ parameters.[23]

Model 3, labeled 'error dependent classification error' model, contains a more specific structure for the dependency of errors across occasions. Rather than allowing the observed states at adjacent months to be related as in Model 2, we allow the probability of a particular type of error at month *t* to depend on the type of error made at month *t*-1. This requires that y_{it}^{2000} depends not only on the observed state at *t*-1 but also on the latent state at *t*-1. This is achieved by replacing the term $P(y_{it}^{2000} = \ell \mid x_t = s)$ appearing in equation [2.2] with the term $P(y_{it}^{2000} = \ell \mid x_t = s, x_{t-1} = r, y_{i,t-1}^{2000} = k)$.

Given that *K*=*M*=4, introducing both the lagged latent state and the lagged 2000 reported state (x_{t-1} and y_{it-1}^{2000}) into the measurement error part of the latent Markov model implies that there will be 16 (4 times 4) matrices of error probabilities, one for each possible combination of lagged observed and latent state. On the one hand, this shows that this model allows great flexibility. On the other hand, while it is possible to estimate all 16 sets of error probabilities freely, this may yield many parameters possibly difficult to interpret. We therefore opt for imposing more structure on the measurement error parameters for the 2000 report. More specifically, we assume that there are two sets of error probabilities (instead of 16), one for the situation in which the *t*-1 report was correct and one for the situation in which the *t*-1 report was incorrect. Such a more restricted structure can be easily defined using a logit parameterization of the error probabilities; that is,

$$\log \frac{P(y_{it}^{2000} = \ell \mid x_t = s, x_{t-1} = r, y_{i,t-1}^{2000} = k)}{P(y_{it}^{2000} = s \mid x_t = s, x_{t-1} = r, y_{i,t-1}^{2000} = k)} = \beta_{\ell s}^{2000}\delta_{rk} + \lambda_{\ell s}(1-\delta_{rk}). \qquad [2.12]$$

Here, δ_{rk} is an indicator variable (a dummy) taking on the value 1 if *k*=*r* and 0 otherwise. Hence, compared to the ICE model, we have two sets of parameters, one for the situation in which the report at *t*-1 was correct ($\delta_{rk}=1$) and one for the

[23] This because y_{it-1}^{2000} and y_{it}^{2000} are a nominal predictor with 4 categories. Since one serves as reference, we have 3*3=9 free parameters in the relationship between them. Model 2 has, in total, 12+9 extra parameters compared to Model 1 because also the $\beta_{\ell s}^{1991}$ parameters are freely estimated instead of fixed to -100.

situation in which it was incorrect ($\delta_{rk} = 0$). As in the ICE model, for identification, we set $\beta_{\ell s}^{2000} = \lambda_{\ell s} = 0$ for $\ell = s$. This yields 2 sets of 12 measurement error coefficients.

The constraints we impose represent one of the many possible restrictions that could be imposed. However, it turned out that the results of main interest are robust to the exact specification that is used. One of the other specification we tried, rather than distinguishing reports on and off the main diagonal at *t*-1 (no error versus any error), looks at the type of mistake and in particular at whether at the previous time point the respondent made the same mistake. The underlying idea is that a particular type of mistake is more likely to occur if the same mistake occurred at the previous time point: the log odds of a mistake at *t* depends on having had the same mistake at *t*-1. This specification is obtained with a logit model of the form $\beta_{\ell s}^{2000} + \lambda_{\ell s}\delta_{\ell srk}$, where $\delta_{\ell srk} = 1$ if $\ell \neq s \,\&\, \ell = k \,\&\, s = r$, and otherwise 0.

In Table 6 we present the estimated classification error probabilities obtained from the three models described above. Note that these sum to 1 within columns. The numbers appearing in the first line represent the estimated probability of being in a particular latent state averaged across time points.

ICE Model (Model 1): Perfect response at the first (1991) interview and independent measurement error at the second (2000).

The results obtained with the baseline ICE Model 1 show that the estimated probability of being truly employed is .75. For this state the measurement error in the 2000 report is rather small (less than 5%). The self-employed and non-employed states (containing 7 and 17% of the person periods) have much larger measurement errors (22 and 19%). For the unemployment state (estimated to contain 1.5% of the person months), the measurement error is extremely large with a misclassification rate of 69%.

Table 6. Markov models results

Independent Classification.Error(I.C.E.) Model (Model 1)					*State dependent classification error Model (Model 2)*						*Error dependent classification error Model (Model 3)*					
BIC: 29922.2284	**X**				BIC: 12496.1606	**X**					BIC: 19010.0015	**X**				
LL: -14813.1616; df: 39					LL: -6065.9848; df: 48						LL: -9311.5243; df: 51					
	E	**S**	**N**	**U**			**E**	**S**	**N**	**U**			**E**	**S**	**N**	**U**
Size	0.7497	0.0664	0.1686	0.0152	Size		0.7492	0.0664	0.169	0.0154	Size		0.7773	0.0524	0.1489	0.0214
y1991					**y1991**						**y1991**					
E	1	0	0	0	E		1	0.0001	0.0019	0.0007	E		0.9988	0.0001	0.0163	0.2122
S	0	1	0	0	S		0	0.9999	0	0	S		0.0001	0.999	0.0001	0.0168
N	0	0	1	0	N		0	0	0.9981	0.0041	N		0.0011	0.0009	0.9834	0.1024
U	0	0	0	1	U		0	0	0	0.9951	U		0	0	0.0002	0.6686
y2000						**X**						**X**				
E	0.9559	0.1137	0.168	0.4735	**lag2000**	**y2000**	**E**	**S**	**N**	**U**	**lag2000**	**y2000 Xlag**	**E**	**S**	**N**	**U**
S	0.0073	0.7779	0.0127	0.0576	E	E	0.9962	0.9501	0.9237	0.9615	E/S/N/U	E match	0.9971	0.0088	0.0185	0.1258
N	0.0323	0.1033	0.8083	0.1578	E	S	0.0002	0.0244	0.0012	0.0008	E/S/N/U	S match	0.0003	0.9874	0.0005	0.0035
U	0.0045	0.0052	0.011	0.3111	E	N	0.0029	0.0198	0.0699	0.0098	E/S/N/U	N match	0.0026	0.0038	0.981	0.0299
					E	U	0.0006	0.0057	0.0052	0.0279	E/S/N/U	U match	0	0	0	0.8408
					S	E	0.0178	0.0001	0.0026	0.0041						
					S	S	0.9696	0.9991	0.9578	0.9533	E/S/N/U	E mismatch	0.1747	0.4346	0.7544	0.6611
					S	N	0.0093	0.0005	0.0352	0.0074	E/S/N/U	S mismatch	0.1692	0.1559	0.0597	0.0802
					S	U	0.0033	0.0003	0.0044	0.0352	E/S/N/U	N mismatch	0.6561	0.4095	0.1859	0.201
					N	E	0.103	0.0159	0.0045	0.0319	E/S/N/U	U mismatch	0	0	0	0.0577
					N	S	0.0005	0.0093	0.0001	0.0006						
					N	N	0.8957	0.9735	0.9951	0.9553						
					N	U	0.0008	0.0013	0.0003	0.0122						
					U	E	0.1723	0.0211	0.0215	0.0046						
					U	S	0	0.0002	0	0						
					U	N	0.0228	0.0196	0.0725	0.0021						
					U	U	0.8049	0.9591	0.906	0.9934						

State dependent classification error model (Model 2): Relaxing the ICE assumption in the later response (2000) and allowing (independent) measurement error in the first response (1991).

Model 2 relaxes the assumption that the 1991 response is perfect and makes the 2000 response dependent on the previous observed state. The estimated measurement error probabilities for the first response (1991) indicate that hardly any classification errors occur in each of the four states. The picture for the second response (2000) is completely different in the sense that the observed state seems to be strongly related to the observed state at the previous time point and, although the probabilities are slightly higher for the true responses, that it is only weakly related to the true state. Therefore, in our opinion, this specification does not yield very informative results about the measurement error in the 2000 report, since we should not simply consider the observed state at the previous time point, but rather whether it was reported correctly or not.

Error dependent classification error model (Model 3): Allowing dependency on previous errors.

Model 3 contains a more precise specification of the dependence of errors across months in the 2000 sequence by assuming that the likelihood of an error does depend not only on the lagged observed state (as in Model 2) but also on the type of lagged error. Specifically, we comment on the results obtained from a model in which we distinguish two situations, namely any error against no error. The assumption is that respondents have a stronger tendency to incorrectly report their current state if they made a(ny) mistake at the previous time point.

The results for the 1991 response show that there is basically no error for the employment and self-employment states, very small error for the not employment state, and much larger error (33%) for the unemployment state. The much larger measurement error in the unemployment state is well documented in the literature (Magidson, Vermunt and Tran, 2009).

The estimated measurement error probabilities show that the error at the second interview is very small if a respondent correctly reported the state at the previous time point. For unemployment though, the misreport is larger than for the other states (about .16), also when no error in the lag occurred. When any type of classification error occurred at the previous time point instead, much more errors occur at the current time point. This applies to all four true states. In this situation,

self-employment and unemployment are mainly confused with employment and not employment, while employment and not employment are often reversed.

An interesting result is that when no error occurred at *t*-1, the quality of the response at the second, more distant, interview (2000) is similar to the one of the first interview (1991), and sometimes even better (error for unemployment is smaller). On the other hand, when error occurred at a previous time point,[24] the quality of the response greatly deteriorates. These results confirm that this model specification makes sense: after filtering out the bad cases, the quality of the 1991 and 2000 reports are similar, which means for the shorter and longer recall periods.

Overall, assuming independent measurement error seems to give a distorted picture due to the fact that people make consistent error. However, we are able to relax the independent classification error (ICE) assumption in a way which makes sense given the way the data were collected. By specifying two different sets of error probabilities in the later report (2000) (Model 3), we are able to distinguish one situation which is similar to (or even more favorable than) the first response (1991) and another which is clearly less favorable.

Comparing latent and observed career sequence

After estimating a particular latent class Markov model, based on the parameter estimates and the observed response sequences, one can obtain posterior estimates for the true state at each time point *t*. The quality of the prediction of the true states is extremely high in our application: we obtain a correct prediction in 99.8% of the person-month combinations.

We use this classification output for a further analysis in which the estimated sequence of true responses is compared with the observed sequences. Note that this represents an improvement compared to the OM approach in which we quantified how well the 2000 sequence matches the 1991 sequence, which means that we implicitly assumed that the 1991 sequence is free of errors. Tables 7 to 9 provide the relevant results obtained by comparing the estimated true response according to Model 3 with both the 1991 and 2000 sequences.

[24] This is the case for less than 8% of the person-time points.

Table 7. (Standardized) distances (OM)

	OM distance	
	mean	bootstrap C.I.
1991-latent	0.020	(0.016 - 0.023)
2000 - latent	0.165	(0.147 - 0.182)
1991-2000	0.195	(0.173-0.216)

In Table 7,[25] we present a measure of the distance between the observed and the latent sequences, computed in the same way as we did for 2000 in Table 5, column 1 (par. 3.2). While in the latter case the observed response in 1991 was used as a reference, we now take the true response as reference, and we compare both the 1991 and 2000 responses to it. As it can be seen in the table, the average distance to the true sequence is higher in the later (2000) than in the first (1991) report and, as indicated by the bootstrap confidence interval, this difference is statistically significant.

Comparison with the results of the distance between the two observed sequences in 1991 and 2000 shows that the mismatch of 2000 sequences with the true sequence is smaller than its mismatch with the 1991 sequence. This illustrates that the amount of recall bias in the 2000 interview is overestimated when it is assessed based on a comparison with another better but still imperfect observed sequence.

Table 8. OLS regression on the distance between latent and observed response, by interview

	latent-observed	(a) latent-1991	(b) latent-2000	sig. of (a)-(b)
Female	-0.011	-0.015*	0.007	
Birth year	0.0001	0.0005*	-0.0002	
Episodes out of employment	0.105**	0.049**	0.162**	**
Career complexity	0.118**	0.03**	0.21**	**
Later report (2000)	0.145**			
Constant	-0.271**	-0.998*	0.593	
N	3946	1973	1973	
R squared	0.14	0.09	0.14	

[25] The distance measures presented in Table 7 are based on the 'default' subcost matrix as defined in par. 3.2.

Similarly to what we did in paragraph 3.2 (see also Table 5), we use regression analysis to quantify the effect of respondents' and career characteristics on the quality of the reports. Table 8 shows the results of such analysis. In the first column, we show the results of a regression analysis on the full set of sequence distances, where we look at the overall effect of the covariates of interest, controlling for the source of observed response, i.e. whether the distance to the latent sequence refers to the report gathered in 1991 or in 2000. We find that the complexity of the career and the presence of episodes out of employment have a significant effect, increasing the amount of errors both at shorter (1991) and at longer (2000) recall times, and we confirm that the later reports (2000) are more different from the true sequence. In the second and the third column of Table 8 instead we present the results of a regression of the distance between observed and latent response separately for the two interviews. The last column also shows whether the effects of the covariates are significantly different in the two cases.[26] The strength of the effects of episodes out of employment and career complexity appears much larger in the later report (2000), which may suggest that increasing recall time makes reports more unreliable when careers are more complex and the respondent was out of employment for some time.

Table 9, where the observed sequence patterns at each interview are compared to the latent one, shows that sequences reported at the first interview (1991) appear closer to the latent sequences than sequences in 2000 do. However this seems to be especially the case for careers involving (more) transitions. Again, comparing the two observed responses with one another instead of with the true sequence seems to overestimate the amount of bias.

[26] The significance of the difference is obtained through a model including the interactions between the interview report (1991 vs 2000) which is compared to the latent sequence and each of the covariates. We use the cluster option in Stata to adjust the standard error to take into account the dependence of the 1991-latent and 2000-latent distance within respondents.

Table 9. Observed and true (most common) sequences

	sequence in 2000										
true sequence	**E**	**N**	**S**	**U**	**EN**	**NE**	**EU**	**ES**	**SE**	**UE**	*else*
E	**96.2**	0.8	0.5	-	1.4	0.7	-	0.2	0.2	-	-
N	4.8	**84.1**	0.5	-	7.3	2.9	-	-	-	-	*0.4*
S	5.4	9.0	**80.2**	-	-	-	-	2.7	-	-	*2.7*
U	11.1	11.1	-	**55.6**	-	-	-	-	-	11.1	*11.1*
EN	16.8	13.7	0.8	-	**63.4**	5.3	-	-	-	-	*0.0*
NE	28.4	6.3	1.0	-	10.5	**52.6**	-	-	-	-	*1.2*
EU	51.1	v	-	-	6.4	4.3	**34.0**	-	-	-	*4.2*
ES	56.3	0.0	6.2	-	-	-	-	**37.5**	-	-	-
SE	85.7	-	14.3	-	-	-	-	-	-	-	-
UE	23.5	-	-	-	-	-	-	-	-	**70.6**	*5.9*
else	*13.2*	*22.6*	*5.7*	-	-	*9.4*	-	-	-	*3.8*	***45.3***
N	1356	230	103	5	111	69	11	11	3	9	65

	sequence in 1991										
true sequence	**E**	**N**	**S**	**U**	**EN**	**NE**	**EU**	**ES**	**SE**	**UE**	*else*
E	**98.6**	-	-	-	0.8	0.4	-	0.1	0.1	-	*0.0*
N	-	**89.9**	-	-	4.8	4.8	-	-	-	-	*0.5*
S	-	-	**99.1**	-	-	-	-	-	-	-	*0.9*
U	11.1	-	-	**11.1**	-	-	11.1	-	-	22.2	*44.5*
EN	0.0	-	-	-	**100.0**	-	0.0	-	-	-	*0.0*
NE	-	-	-	-	8.4	**90.5**	-	-	-	-	*1.1*
EU	12.8	-	-	-	2.1	-	**72.3**	2.1	-	-	*10.7*
ES	-	-	-	-	-	-	-	**100.0**	-	-	*0.0*
SE	-	-	-	-	-	-	-	-	**100.0**	-	*0.0*
UE	23.5	-	-	-	5.9	5.9	11.8	-	-	**52.9**	*0.0*
else	*1.9*	*3.8*	*1.9*	*1.9*	*7.6*	*1.9*	*1.9*	-	-	*3.8*	*75.3*
N	1282	188	111	1	166	103	37	18	8	11	48

4. Discussion and conclusions

In this chapter, we investigate measurement error in retrospective reports of employment biographies. We use retrospective data collected at two different time points using the same questionnaire and interviewing the same respondents about their employment histories in the same reference period to study measurement error and memory bias in particular. We present a new model-based approach to investigate the importance of memory bias, which represents an improvement over purely descriptive techniques.

Focusing on retrospectively collected employment histories, we first apply an explorative algorithmic sequence technique to investigate differences between retrospective reports provided at different time points, i.e. for reports differing in the length of the recall period. The use of Optimal Matching allows us to define a measure of the differences between the observed career sequences, and using

regression analysis we can also investigate the extent to which this relates to respondents' and career characteristics.
However, applying an algorithmic approach such as OM we have to use one of the two responses as reference for our comparison, and it seems reasonable to trust the report with shorter recall period (1991) more. Although we may think that longer time distance between the period to be reconstructed and the interview might increase bias in the later reports (2000), the occurrence of error at the first interview cannot be ruled out and we do not know what the true state is.

Using latent class Markov models, we can correct for measurement error in each response and introduce a latent variable for the true state, which we don't know. We can introduce measurement error in each response, which means that we do not assume that the response at the first interview (1991), which is the one with shorter recall period, is perfect, as it is instead implicitly done when comparing the two observed reports. Furthermore, we propose various modifications of the latent Markov model that allow us to relax the assumption of independent classification errors.

The results of the analyses performed in this chapter are informative on the extent and the nature of errors in retrospective reports and warn scholars interested in carrying out research using such data against possible shortcomings.
We confirm that recall errors tend to lead to a simplification of careers especially of the earlier career parts –having them occurred farther away in time- resulting in more continuous careers, characterized by fewer episodes and transitions than they actually have. This may consequently lead to an overestimation of careers' stability in retrospective data. Career complexity further increases the difficulty of the recall process and unemployment appears to be more often misreported, even shortly after the occurrence of the event. However, the recall of a more complex career and of episodes out of employment seem to be more problematic when the recall period is longer.
We also find that, while for some cases error is persistent over time, for the great majority the quality of the reports is not affected by the length of the recall period.

Overall, it seems that the quality of retrospective data is not as bad as some people tend to believe. Some particular sequences seem better suited to be recalled through retrospective questioning than others (unemployment) and recall problems seem to

be limited to very specific situations. In most cases the reporting is of high quality justifying no doubt on the inferences derived from retrospective surveys. Furthermore, once researchers are aware of the possible errors involved, it is possible to quantify them and to correct for them.

The chapter deals with a very complex issue and, although many questions still remain open, our approach might be an important step forward in our knowledge of the importance of memory bias in life history data. The information obtained appears very useful for subsequent studies addressing substantive research questions using retrospective life-history data (on careers).
The findings may also be empirically exploited in subsequent substantive studies to correct for measurement error. Without re-estimating measurement error, the parameters obtained from our Latent Markov model might be used as fixed coefficients for measurement error in which the correction could be integrated in the likelihood function. Furthermore, in more substantive research, we might include predictors to model the dynamic process of interest (not measurement error).

References

Abbott, A. (1988). Transcending general linear reality, *Sociological Theory*, 6, 169–186

Abbott, A. (1995). A comment on "Measuring the agreement between sequences", *Sociological Methods and Research*, 24(2), 232–243

Abbott, A. and Forrest, J. (1986) Optimal matching methods for historical sequences, *Journal of Interdisciplinary History*, XVI(3), 471–494

Abbott, A. and Tsay, A. (2000). Sequence analysis and optional matching methods in sociology, *Sociological Methods and Research*, 29(1), 3–33

Björklund, A. and Holmlund, B. (1987). Worker Displacement in Sweden: Facts and Policies. Paper presented at the American Economic Association meeting, Stockholm, December

Blomskog, S. (1997). *Essays on the functioning of the Swedish labour market* (Vol. 27). Stockholm: Swedish Institute for Social Research

Brückner, H. and Mayer, K. U. (2005). The De-Standardization of the Life Course: What It Might Mean and Whether It Actually Took Place. *Advances in Life Course Research* 9:27-54

Büchtemann, C.F., (1993). Employment security and labor markets. In: Buechtemann, C.F. (Ed.), Employment Security and Labor Market Behavior. International Industrial and Labor Relations Report 23. ILR Press, Ithaca

Collins, L. M., and Wugalter, S. E. (1992). Latent class models for stage-sequential dynamic latent variables. Multivariate Behavioral Research, 27, 131–157

Dex, S. (1991). The Reliability of Recall Data: A Literature Review. *ISER Working Paper*. Colchester, University of Essex

DiPrete, T. (2002). Life course risks, mobility regimes, and mobility consequences: A comparison of Sweden, Germany, and the United States. *American Journal of Sociology,* 108 No.2.

DiPrete, T. A., de Graaf, P. M., Luijkx, R., Tahlin, M., and Blossfeld, H.-P. (1997). Collectivist versus individualist mobility regimes? Structural change and job mobility in four countries. *American Journal of Sociology,* 103(2), 318-358

DiPrete, T. A., Goux, D., Maurin, E., and Tåhlin, M. (2001). Institutional determinants of employment chances. The structure of unemployment in France and Sweden. *European Sociological Review,* 17 No.3, 233-254

Efron, B., and R. J. Tibshirani (1993). *An introduction to the bootstrap.* New York: Chapman & Hall

Elias, P. (1997). Who forgot they were unemployed? *ESRC Working Paper*

Elzinga, C.H. (2007a). Turbulence in categorical time series, *Sociological Methods and Research* (submitted for publication)

Elzinga, C. H. (2007b) CHES*A 2.1 User manual*. Amsterdam, Vrije Universiteit Amsterdam

Elzinga, C. H. and Liefbroer, A. C. (2007). De-standardization and differentiation of family life trajectories of young adults: A cross-national comparison'. *European Journal of Population* 23: 225-250.

Grubb, D. and Wells, W. (1993). Employment Regulation and Patterns of Work in EC Countries, OECD Economic Studies, No. 21, Winter, pp. 7-58

Gustafsson, S. S., Wetzels, C. M. M. P., Vlasblom, J. D., and Dex, S. (1996). Women's labor force transitions in connection with childbirth: A panel data comparison between Germany, Sweden and Great Britain. *Journal of Population Economics,* 9(3), 223

Hagenaars, J. A. (1990). Categorical Longitudinal Data - Loglinear Analysis of Panel, Trend and Cohort Data. Newbury Park: Sage.

Jonsson, J. O., and Mills, C. (2001). *Cradle to grave. Life-course change in modern Sweden*: Sociologypress

leGrand, C., and Tahlin, M. (2002). Job mobility and earnings growth. *European Sociological Review,* 18 No. 4, 381-400

Magidson, J., Vermunt, J.K., and Tran, B. (2009). Using a mixture latent Markov model to analyze longitudinal U.S. employment data involving measurement error. In: K. Shigemasu, A. Okada, T. Imaizumi, and T. Hoshino (eds.), *New Trends in Psychometrics*. Universal Academy Press

OECD (1994). The OECD Jobs Study: Evidence and Explanations. 2 vols, Paris .

Paull, G (2002). Biases in the reporting of labour market dynamics. London, The Institute for Fiscal Studies

Poulsen, C. S. (1982). Latent structure analysis with choice modeling applications. Arhus: The Arhus School of Business Administration and Economics

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, Berlin, Max-Planck-Institut für Bildungsforschung

Van de Pol, F., and Langeheine, R. (1990) Mixed Markov latent class models. *Sociological Methodology*, 20, 213–247

Vermunt, J.K. (2003). Multilevel latent class models. *Sociological Methodology*, 33, 213–239

Vermunt, J.K. and Magidson, J. (2007). LG-Syntax User's Guide: Manual for Latent GOLD 4.5 Syntax Module. Belmont, MA.: Statistical Innovations

Vermunt, J. K., Tran, B., and Magidson, J. (2008). Latent class models in longitudinal research. In S. Menard (Ed.), Handbook of Longitudinal Research: Design, Measurement, and Analysis. Elsevier. In press

Vermunt, J. K., and Magidson, J. (2004). Latent class analysis. In M. S. Lewis-Beck, A. E. Bryman, and T. F. Liao (Eds.), *The Sage Encyclopedia of Social Science Research Methods* (pp. 549-553). Thousand Oaks, CA: Sage.

Chapter 3

Explaining differences in labor market transitions between panel and life-course data in West-Germany *

Abstract. Panel and life-course data are ideally suited to unraveling labor market dynamics, but their designs differ, with potential consequences for the estimated relationships. To gauge the extent to which these two data designs produce dissimilar transition rates and the causation thereof, we use the German Life History Study and the German Socio-Economic Panel. Life-course data in particular suffer from recall effects due to memory bias causing understated transition probabilities. Panel data suffer from seam effects due to spurious transitions between statuses recalled in activity calendars that generate heaps at particular time points and cause overstated transition probabilities. We combine the two datasets and estimate multilevel (multistate) discrete-time models for event history data to model transitions between labor market states taking these factors into account. Though we find much lower transition rates in the life-course study, confirming the results of Solga (2001) in this Journal for East-Germany, part of the difference can be explained by short spells recall bias. The estimated models on exit, re-entry and job mobility on the combined datasets show indeed a negative retrospective design effect. Another specification that includes the length of the recall period shows no significant decrease in the transition probabilities with increasing length, suggesting that the negative design effect is due to other design differences.

** This chapter is a very similar version the publication Manzoni, A., Luijkx R. and Muffels, R.(2009) Labour Market Transitions in Panel and Life-course Data: Examining Differences Between The Two Designs in Germany. Quality and Quantity. International Journal of Methodology.*

1. Introduction

Obtaining accurate and complete information about people's life course is not easy (Scott and Alwin, 1998). During the last few decades, serious questions were raised about the survey methodology and data reliability of retrospective surveys, with a general concern for increased memory bias or measurement error compared to a panel design in which questions are asked about people's current status (Dex, 1991).

In this paper, we want to check whether different designs lead to dissimilar conclusions with respect to the transition rates and the determinants of career moves, i.e. entry into and exit from employment as well as moves between jobs. Also to assure that the results are not affected by contextual differences, we restrict ourselves to data in one country, i.e. (West) Germany, using two carefully designed socio-economic datasets aimed at measuring labor market statuses and changes therein over time: The German Life History Study (GLHS) and the German Socio-Economic Panel (GSOEP).

Retrospective surveys like the GLHS ask respondents to recall their behavior over a specified prior period whereas prospective panel surveys repeatedly ask individuals about their current situation using the same questionnaire. However, they may also gather retrospective information, both for the period before the first interview and to fill in the gaps between consecutive yearly interviews, as it is in the GSOEP. One approach that embeds periodic retrospective measurement within the prospective panel design is the life history calendar (LHC), made to link events and transitions over time and therefore improving the quality of the data (Freedman et al., 1988; Taris, 2000).

Earlier in this Journal, Solga (2001) compared the same two datasets to examine occupational mobility. She found strong differences across the two designs in the transition rates for East-Germany, which she tried to explain by comparing the methodological features of the two designs. We build on her approach but shift the focus to labor market and job mobility patterns in West-Germany. We add to her approach through the elaboration and testing of an explanatory model on job mobility as well as on exit and re-entry mobility, combining the two datasets. Moreover, we use this empirical model to explain the dissimilar outcomes in the transition rates across the two designs, but after correcting for design differences. Finally, we model the sources of bias in the two designs and in particular memory bias, from which especially the German life-history study suffers.

2. Prospective and retrospective designs: A review of the differences

A prospective panel follows members forward in time, while a retrospective study follows members backward in time (Tuma and Hannan, 1984). This basic feature leads to many differences between the two types of longitudinal surveys. Much of what has been written in the literature on the comparison between prospective and retrospective surveys deals with tracing the advantages and disadvantages of both designs (Blossfeld and Rohwer, 1995). As Scott and Alwin (1998) point out, the implied polarization between retrospective and prospective methods has to be rejected because "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 and Alwin, 1998: 106). Several aspects have to be weighted when evaluating prospective and retrospective data: practical considerations of time and money, substantive concerns, as well as issues of measurement quality. We deal in particular with the latter.

2.1 Selection and memory bias in retrospective data

Retrospective data suffer from two main measurement errors: substantial selection bias due to prior mortality, since estimates are only representative for the survivors, and memory bias. When individuals are asked to retrieve events, situations and attitudes which occurred in the past, the possibility of inaccurate reporting (omission or misclassification) may occur, either from unintentional memory lapse or a conscious misrepresentation of the past (Powers et al., 1978). Inaccurate long-term recall could make retrospective surveys a very poor substitute for panel data, but their quality is not easy to assess, especially considering that earlier experiences may be revised in the light of subsequent ones and that both willingness and ability to recall play a role.

2.2 Length of the recall period

Studies in cognitive psychology have identified three major factors responsible for recall accuracy: interference (it is easier to remember rare events than frequent ones since the occurrence of many similar or related events reduces the memorability of each individual event), salience of the event to be reported (meaning how important it is to the respondent), and length of the recall period (Eisenhower et al., 1991). As

memory decays over time, the probability of accurate recall is generally thought to decrease with the length of the recall period. However, shortening the recall period 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 (Sudman and Bradburn, 1973). Assuming that interference and salience are the same in the two designs,[27] we are left with the length of the recall period. The issue as to whether recall errors are a function of time distance between the event and the interview date is a very important one. If this is the case, it would suggest that retrospective measurement of 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 and Alwin, 1998). Different studies have researched the importance of the length of the recall period for recall accuracy (Mathiowetz and Duncan, 1988; Paull, 2002; Powers et al., 1978) finding differences among different groups and depending on the subject matter. Results are sometimes contrasting: Although some studies (Bradburn et al., 1987) argue that the longer ago an event is, the less reliable the answers concerning events are, suggesting a linear decay with time (Linton, 1982; Thompson et al., 1996; Wagenaar, 1986), others have shown that very recent events are recalled best but then the decay in memory does not continue over time and stays stable for a long period (Rubin and Baddeley, 1989). 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 have been controlled for. Heavily relying on Reimer's (2005), findings, Mayer (2007) 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". According to Reimer (2005) the likelihood of error does not seem to be related to time: the consistent lack of support for the relevance of a 'time to interview' distance factor shows that no simple memory laws apply. De Graaf et al., in an unpublished work on data of a German 1987 three wave panel life history study, investigate the reliability and consistency of individual occupational career data, showing that retrospective variables are no less reliable than other survey data. The amount of job spells seems 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 to have no effect. Hauser and Featherman (1977) conclude that reports on jobs that respondents had

[27] This hypothesis seems realistic since we don't have any reasons to believe that work histories are more complex in one of the two samples nor that events have different importance.

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

2.3 Short spells memory bias

Previous findings in the literature corroborate the recall risk of non-standard episodes (like unemployment), which are markedly less well remembered (Dex, 1991; Duncan and Hill 1985; Elias, 1997; Jacobs, 2002; Mathiowetz and Duncan, 1988; Paull, 2002; Reimer, 2004). Reimer and Künster (2004), comparing the GLHS with register data, show that short episodes are omitted or merged, when transitions are not very marked, leading to simpler, more conventional and stable careers. Therefore, we may assume that, retrospectively, respondents underreport short unemployment spells.

2.4 Short-term recall bias

The longer the recall period, the more the concerns about reliability, although there are important differences by subject matter and salience (Scott and Alwin, 1998). Prospective designs seem to suffer much less from recall bias than retrospective surveys, due to shorter length of recall between. Yet, biases in recall might occur in the short-term as well (Jürges, 2005; Paull, 2002).

2.5 Seam effects

The use of multi measurement occasions in panel studies 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 behavior (Solga, 2001). Due to the time elapsed between two interviews, the meaning of a measure and the quality of the answers may change, leading to the risk of artificial mobility: relatively small changes in response to a question for adjacent months within a reference period and more abrupt changes for adjacent months across reference periods might appear (Callegaro, 2006; Lynn et al., 2005).

2.6 Panel attrition

Due to panel mortality and attrition,[28] Panel data face considerable problems with population representativeness (Solga, 2001): The sample size decreases wave by wave, which might lead to severe problems for making reliable population inferences over time, especially because non-response doesn't occur at random, with the consequent risk of biased estimates (Hagenaars, 1990; Rendtel, 1989). The longer the 'life' of a panel survey is, the higher this risk is. On the other hand, the assumption that right censoring is independent of spell duration cannot be rejected: attrition is neither related to the elapsed or the remaining spell length (Pyy-Martikainen and Rendtel, 2003).[29] Furthermore, since most of the variables explaining attrition are included in our regression models, the possible effect on the transition rates and on the differences between the datasets is negligible.

3. Data

The GLHS and the GSOEP are both large-scale multipurpose surveys collecting data in a different way on a variety of subjects. 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. The GSOEP is a panel study, in which the same sample of participants is interviewed annually, using the same questionnaires.

The GLHS,[30] carried out since 1983 at the Max Planck Institute for Human Development in Berlin under the direction of Professor Karl Ulrich Mayer, permits to construct a complete retrospective career history of the respondents, who were asked

[28] See Haisken-DeNew and Frick (2005) and Kroh and Spieß (2006) for details about attrition in the GSOEP.

[29] A popular strategy for coping with panel attrition is weighting. Yet, it is not clear how effective it is in reducing attrition bias (Vandecasteele and Debels, 2007). DuMouchel and 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 and 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 (longitudinal) weights in panel models, we have not used weights in the presented analyses.

[30] We are using the aggregate dataset (Gesamtdatenbank), which contains data from four different surveys in which different birth cohorts are interviewed. Prof. K. U. Mayer and his collaborators kindly made this last version available for our analysis. For further details on the GLHS see Mayer (2007).

the monthly beginning and ending dates of each job or self-employment spell they had ever experienced.[31] They were also asked to identify the occupation, the branch of industry, the size of the firm, and the wages at the start and the end of each job and self-employment episode. Finally, respondents were asked to indicate whether occurring job shifts referred to in-firm or between-firm job moves.

The GSOEP [32] is a longitudinal survey of private households and persons in Germany that has been running since 1984. While some long-term retrospective information is collected, to reconstruct background information on work history[33] (but not only), a set of core questions on labor market and occupational changes is asked yearly to assess the current employment status and employment changes. At each interview, respondents should report whether they are currently engaged in paid employment and if so, in which occupation, allowing us to construct career patterns though based on annual information. The GSOEP also collects monthly information on transitions and events occurring between two waves, using a monthly calendar asking respondents to reconstruct the employment situation for each month in the previous calendar year.[34] Therefore, the monthly calendar information is also affected by some memory bias.[35] Respondents are shown a grid[36]

[31] See http://www.mpib-berlin.mpg.de/en/forschung/bag/projekte/lebensverlaufsstudie/index.htm.

[32] See Haisken-DeNew and Frick (2005); find the questionnaire on http://panel.gsoep.de/soepinfo2006/).

[33] We are only using this information to reconstruct job duration of spells which started before the first interview.

[34] Since the monthly calendar allows to report multiple states in one months, some problems originate when this indeed happens. Contrary to previous studies (Bergemann and Mertens, 2004; Ernicke, 1997), which have dealt with the issue imposing a hierarchy of the spells, we consider the type of spells reported in the same month as well as the length of the multiple spells. We assume that employment status cannot coexist only with "unemployment", "housewife" or "military service" and that in case of multiple spells for one or two consecutive months, independent of the kind of spells, multiple states are not possible, so we observe one (or two) event(s). In case of three or more consecutive months, we distinguish two different situations: if the respondent declared to have been employed and, in the same month either housewife, unemployed or involved in the military or civil service, we exclude all the records from that moment onwards for that person from our analysis., while we assume the respondent hold more states in case of other multiple states declaration. This decision is due to the fact that in the first case, 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.

[35] We have checked the reliability of short term retrospective information in the GSOEP comparing 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. On our selected sample, we find that in nearly 96% of the cases respondents give consistent information on the employment status, and

having as columns all the months in the previous year and as rows a large number of states. In addition, questions are asked concerning the date and the kind of job changes occurred from January 1st the year before the interview,[37] allowing us to split employment spells into job spells and to assess the kind of change intervened.[38] We have information about within-year changes as far as it concerns shifts between part-time and full-time work (from the calendar) and we can assess whether a short job spell that falls between two interviews is in the same firm as the previous one or whether it involves a change of employer. More detailed information on the job is only available if the respondent was in that job at the moment of interview.[39]

Job changes are defined differently in the two datasets: in the GLHS job spells are recalled chronologically and are separated by changes of firms, professional position, income, activity or working time. In the GSOEP, they are separated by changes of firm, professional position within a firm or switches into self-employment. Income or working time changes are not sufficient conditions to classify them as job changes.

Altogether, the two surveys differ in several respects, mainly deriving from the research design. One of the most relevant differences is the different time frame observed. The GLHS reconstructs, at a single point in time, events occurring also long before the interview date. It is only representative for specific birth cohorts and,

---

discrepancy halves allowing for a less rigorous definition of error (more details available from the authors upon request).

[36] Using the 2005 questionnaire these were: full-time employment, part-time job, mini-job, first-time company training/apprenticeship, further training/retraining, further professional education, unemployed, ('Vorruhestand' or early) retirement, 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 and other. See http://panel.gsoep.de/soepinfo2006/, question 99 in 2005.

[37] 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?"

[38] 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, we made sure that no double counting occurred, choosing for the type of change reported in the first interview, which is closer to the event, in case a change was reported twice but a different type of change was reported at the two interviews. Yet, an error could occur also in dating, leading to an overrepresentation of changes, in case a single change is reported as occurring at different time points. (We could recognise mistakes only when two changes were reported with less than one month distance in the months preceding the interview). While the inability to correct for some double reported changes might overestimate the number of changes, the need for agreement in the employment status between the date of change and calendar and the possibility to report at most one change a year lead to a loss of declared changes.

[39] This may lead to missing information for spells not hold at any interview but also to multiple contrasting information for spells hold at more interviews. This might happen, for example, when reporting job characteristics, which might contrast across waves even if no change is reported (for example the respondent might report different occupational status across waves without reporting a change in occupational status).

dependent on the moment in life at which respondents are interviewed, it reconstructs a longer or shorter period of their life. The panel survey is instead a representative sample of the residential population of the FGR in 1984[40] and monitors changes wave by wave. It is representative for the whole population, but it captures information only starting from the moment of the first interview, because of which left-censoring can easily occur.[41] As a consequence, older cohorts only report information on relatively older ages compared to younger cohorts.

We do not believe that selecting only the cohorts and period for which the data overlap, in the attempt to make the two datasets more similar would be a good solution, because it would leave us with a very short period in which all the cohorts are interviewed (and each cohort would be observed at different ages), while more recent years would entail only very young people born in more recent birth cohorts. Nothing could be said anymore nor on the period effect (because of the relatively short period under observation and because of the different age-cohort composition across periods) nor on the cohort effects (since each cohort is observed at different ages). Furthermore, we would be selecting only respondents' reports for years quite close to the interview for most of the cohorts, so that reports for periods more distant from the interview would only come from the younger birth cohorts reporting about very young ages, ruling out the possibility to study the effect of the length of the recall period. Therefore, we rely on the results of the analyses on the combined data.[42]

The difficulty of the reporting task might also be affected by the specific way the information is retrieved. In the life history study, respondents have to recall their labor market situation on a spell basis, being asked the beginning and end date of each spell. Based on the calendar instead, respondents in the GSOEP simply have to recall their monthly status, while spells and their timing are reconstructed by the researcher.[43] Even though in principle this doesn't necessarily need to be the case, retrospective surveys are usually event-based, while the 'calendar' format is more common for shorter recall periods.

---

[40] Additional samples were added later on (Haisken-DeNew and Frick, 2005).

[41] Matching cross-sectional information on the time spent in the current spell, left-censored spells become left truncated and can be used in the duration analysis.

[42] To strengthen our results and to fully rule out any APC composition effects, we have replicated our analyses on the selected overlapping samples: The results (available upon request) confirm our previous findings and show that the design effect is not due to the different APC composition of the two samples.

[43] This holds at least for exit and re-entry events, while the dates of job-to-job changes are reconstructed through a question asking explicitly when the event occurred.

Issues related to data quality, both in the collection (interviewing method, interviewers' training, wording and ordering of questions etc.) and in the editing and coding process (due to time/money constraint and personnel expertise) as well as unspecified background variables and processes of selectivity, cannot be completely ruled out and might also account for differences affecting the comparison.

4. Research Questions

Our first research question deals with the 'true' measurement of transitions and the extent to which in life-course and panel data memory bias and seam effects might have a leveling-down and leveling-up effect, respectively, on the transition probabilities between jobs and on exit and (re)-entry into employment.

A second question concerns the extent to which the differences are due to memory bias effects associated with the underreporting of short spells and to the longer length of the recall period in the life-course data.

A third more substantive question pertains to the degree to which our substantive results are influenced by design effects and to what extent prospective and retrospective designs point to the same determinants of transitions or causal mechanisms on the labor market after correcting for design effects. In this article, we focus on the first two methodological research questions, while the third will be treated in a separate paper.

5. Method

With respect to the first research question, we report in Table 1 on the various sources of bias that we have to take into account for comparing the outcomes between the two designs and on the way we try to control for them.

Table 1. Overview of biases and how we control for them

Design	Origin of bias	Data to be used	Models
Retrospective life-course	Overall design effect	Merged GLHS/GSOEP	Model I, dummy for survey
	Length of recall period	Merged GLHS/GSOEP	Model II-III, splines indicating time distance to the interview in GLHS
	Underreporting of short spells out of employment	Merged GLHS/GSOEP	Model Ia, dummy for survey excluding spells out of employment shorter than 3 months
Prospective panel data	Short term recall bias	GSOEP	Assessment of its significance comparing current and calendar information (not presented here)
	Seam effects	Merged GLHS/GSOEP	Model I-II-III, dummy for December in the GSOEP

The basic idea is to merge the datasets and to estimate models for exit, re-entry and job mobility. By including a design dummy indicating the information source, we try to gauge whether and to what extent the found differences are due to survey design differences. The design dummy should, after correcting for heaps at the seam in the panel data,[44] mainly represent memory bias in the retrospective data, which is acknowledged as the main source of difference between the two datasets (Model I). Although the impact of unobserved characteristics cannot be ruled out entirely there are no a priori reasons to assume that, if any, they will exert a dissimilar effect in the two surveys. Furthermore, we might assume that their effects are small, since the richness of the datasets enabled us to capture most of the theoretically presumed relevant covariates in our models.

To examine the effect of the length of the recall period, we estimate a separate model (Model II), in which we add splines (Marsh and Cormier, 2002) specifying time distance from the interview in the life-course data. Seam effects are accounted for by including a dummy variable for December in the GSOEP, since events are defined in the last month of a spell and calendars run from January to December (Huijer and Schneider, 1989; Kraus and Steiner, 1998). To correct for the underreporting of short unemployment spells, we estimate our first model excluding spells lasting less than three months (Model Ia).

[44] Spurious transitions arise between calendars collected in subsequent years, which generate heaping in the transition probabilities. In the GSOEP, specifically, we observe heaps between December and January, which cannot be explained by cyclical factors alone (Kraus and Steiner, 1998).

In an additional step (Model III), we also include a number of controls and explanatory variables derived from the literature on labor market mobility (Allmendinger, 1989; Blossfeld, 1986; Buchholz, 2004; Buchholz and Grunow, 2003; DiPrete, 2002; DiPrete et al. 1997; Mayer and Carroll, 1987; Mertens, 1999; Muffels and Luijkx, 2008; Rosenfeld, 1992; Sicherman and Galor, 1990; Uunk et al., 2005).

5.1 The empirical model

We elaborate regression models with the transition rates between different employment states as dependent variables and we apply discrete-time models for multilevel[45] event history data, covering exit, re-entry and job-to-job changes.[46] To assess the risk of event occurrence, we use the discrete-time hazard function $h(t)$, which is defined as the conditional probability of having an event during interval *t*, given no occurrence in *t*-1 (Singer and Willett, 2003).[47]

$$h_j(t) = \Pr\left(y_j(t) = 1 \mid y_j(t-1) = 0\right) \qquad [3.1]$$

Logistic regression[48] is used to analyse the transition rates, the dependent variable measuring the odds of an event occurring given that it has not occurred in the month before. Transitions out of employment and into employment are modelled jointly: The model is a generalised multilevel (multistate) discrete-time event history model (Steele et al., 2004) estimating 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. The model can be written as:

[45] A multilevel model is used to allow for the hierarchical structure arising from repeated episodes 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-specific factors, which leads to unobserved heterogeneity in the hazard. Repeated events are usually handled by including individual-level random effects in an event history model.

[46] Exit is defined when the respondent is employed in the current month [*t*=employed] but not in the following [*t+1*=not employed]; Re-entry if subjects who have already experienced an employment spell are currently not employed [*t*=not employed] but are employed in the following month [*t+1*=employed]; Job-to-job changes are any job moves implying the start of a new job spell, given that the respondent was employed before.

[47] Each respondents might have multiple spells.

[48] We use the xtlogit command (with random effects) in Stata 10.

$$\text{logit}\left[h_{ij}^{E}(t)\right]=\alpha^{E}(t)+\beta^{E}x_{ij}^{E}(t)+u_{j}^{E} \text{ for transitions out of employment} \quad [3.2]$$

$$\text{logit}\left[h_{ij}^{N}(t)\right]=\alpha^{N}(t)+\beta^{N}x_{ij}^{N}(t)+u_{j}^{N} \text{ for transition into employment} \quad [3.3]$$

where $h_{ij}^{E}(t)$ and $h_{ij}^{N}(t)$ are, respectively, the hazard of exit and of re-entry in time interval *t* during episode *i* of individual *j;* $x_{ij}^{E}(t)$ and $x_{ij}^{N}(t)$ are covariates included in the exit and in the re-entry model respectively; u_{j}^{E} and u_{j}^{N} are random effects representing unobserved characteristics of individual *j*. We assume u_{j}^{E} and $u_{j}^{N} \sim N\left(0,\sigma_{u}^{2}\right)$ and we allow the $u_{j}{}^{E}$ and $u_{j}{}^{N}$ to be correlated.

Job-to-job mobility is instead modeled separately in a multilevel event history model estimating the chances of any job change, given that the respondent is employed. The model can be written as follows:

$$\text{logit}\left[h_{ij}(t)\right]=\alpha(t)+\beta x_{ij}(t)+u_{j} \quad [3.4]$$

where h_{ij} is the hazard of a job-to-job change in time interval *t* during episode *i* of individual *j*.

The hazard rates are estimated separately for each gender, as the employment behavior of men and women shows clear differences.

5.2 Independent variables

Education is measured by the CASMIN classification (Müller et al., 1989), which distinguishes between levels of education and general vocational training. We differentiate between primary (1a, 1b, 1c), lower secondary (2a and 2b), upper secondary (2c voc and 2c gen) and tertiary education (3a and 3b) and we include the resulting categorical variable as a time varying covariate, indicating the highest level of education at each point in time. Information on marital status and children is combined and included in a time varying form. In the life-course survey, for each marriage, respondents have to report the beginning and the end dates, as well as additional information on the partner and on children, if any. In the GSOEP, the monthly marital status is reported and information on children is gathered.[49]

[49] The definition of child identifies parenthood in a social, not necessarily in a strict biological way.

Following Kalmijn and Luijkx (2006), we combine the information about marital status and children to indicate the stage in the life course, distinguishing five mutually exclusive groups and coding them with cumulative contrasts: single, married without children, young children (any children aged 0 to 5, independent of the marital status), older children (no children under 6 but children under 16, independent of the marital status), empty nest (no children under age 16 anymore in the household).

We include age in linear and, to examine the non-monotonic relationship between an individual's age and employment, also in quadratic form, as control variables. We apply our analyses only to the age group 18-64. We consider all the periods covered by each survey, which range from late 1930s to 1999 in the GLHS and from 1983 to 2005 in the GSOEP.

We control for period and age composition differences, but not for cohort effects, due to the linear relation among age, period and cohort, which makes it impossible to include all the three effects in one regression model.

We include the ISEI score[50] of the current job (for exit and job to job changes) or the most recent (for re-entry) occupation as a measure of the occupational status, which has been shown to be an important determinant of labor mobility (Kalmijn and Luijkx, 2006).

In a further step, despite some concerns for an endogeneity problem, we include time in employment and non-employment as categorical variables in the exit and in the re-entry model respectively. In the job-to-job mobility model, we include job tenure, defined as the number of months in the current job spell.[51] The number of jobs a respondent held in his/her career is included as a time varying covariate, 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 avoid high correlation with age (Kalmijn and Luijkx, 2006).

To capture differences between the two surveys, we include a design dummy indicating the data source, with the GSOEP being the reference category. Due to

[50] In the panel survey, the occupational code is only available if a job was held at the time of interview so that we do not have any information for (relatively short) jobs held in between two interviews. This leads in our case to missing ISEI information for about 15% of the person-months in employment, plus nearly 8% in which contrasting information across waves (for the same spell) couldn't be solved attributing the modal value (in case of contrasting information for one spell the modal value was attributed if more than half of the report agreed with the modal value).

[51] Both duration into employment and job tenure take into account the information given in the biographical questionnaire, so to avoid left-censoring problems.

possible recall errors, reports might change dependent on the length of the recall period. A spline model is preferred over a dummy approach, since it captures smoothly the changes in the effects of the time distance, avoiding therewith abrupt drops. We have defined three knots,[52] namely at 2, 5 and 10 years distance from the interview, which generate four time ranges: less than 2 years, between 2 and 5, between 5 and 10 and more than 10. Spline functions are used to impose continuity in the effect of the length of the recall period, allowing the effects to change gradually within each category indicating time distance to the interview, by estimating different linear slopes for each range, so to avoid inappropriate jumps which would be generated by a simple dummy variable model. The equation for our Model II reads:

$$\begin{aligned}\text{logit}\left[h(t)\right] = {} & \alpha + \beta_1 Age + \beta_2 Age^2 + \beta_3 P1900-49 + \beta_4 P1950-82 + \beta_5 P1990-94 + \beta_6 P1995-99 + \beta_7 P2000-05 \\ & + \beta_8 Married + \beta_9 SmallChildren + \beta_{10} OldChildren + \beta_{11} EmptyNest \\ & + \beta_{12} EduLowerSecondary + \beta_{13} EduUpperSecondary + \beta_{14} EduTertiary + \beta_{15} SeamEffect + \beta_{16} Isei \\ & + \beta_{17} Glhs + \beta_{18} Spline2yrs + \beta_{19} Spline5yrs + \beta_{20} Spline10yrs + \beta_{21} Splinemore10yrs + u_j \end{aligned} \quad [3.5]$$

where the β coefficient for each spline indicates the change in the log odds of the considered transition associated with a month increase in the length of the recall period, in the range considered by each spline, where Spline2yrs indicates distance from the interview shorter than 2 years, Spline5yrs between 2 and 5 years, Spline10yrs between 5 and 10 years and Splinemore10 longer than 10 years.

We assume that after correcting for all the theoretically expected systematic possible sources of bias, our survey indicator summarises the remaining differences between the surveys. Since our explanatory models not only contain the theoretically most relevant variables but also account for a number of personal and demographic variables, we may safely assume that we capture also most of the random differences which might occur due to the fact that we are dealing with two different samples.

[52] The term 'knot', in spline regression models, indicates the endpoints of the ranges in the independent variable.

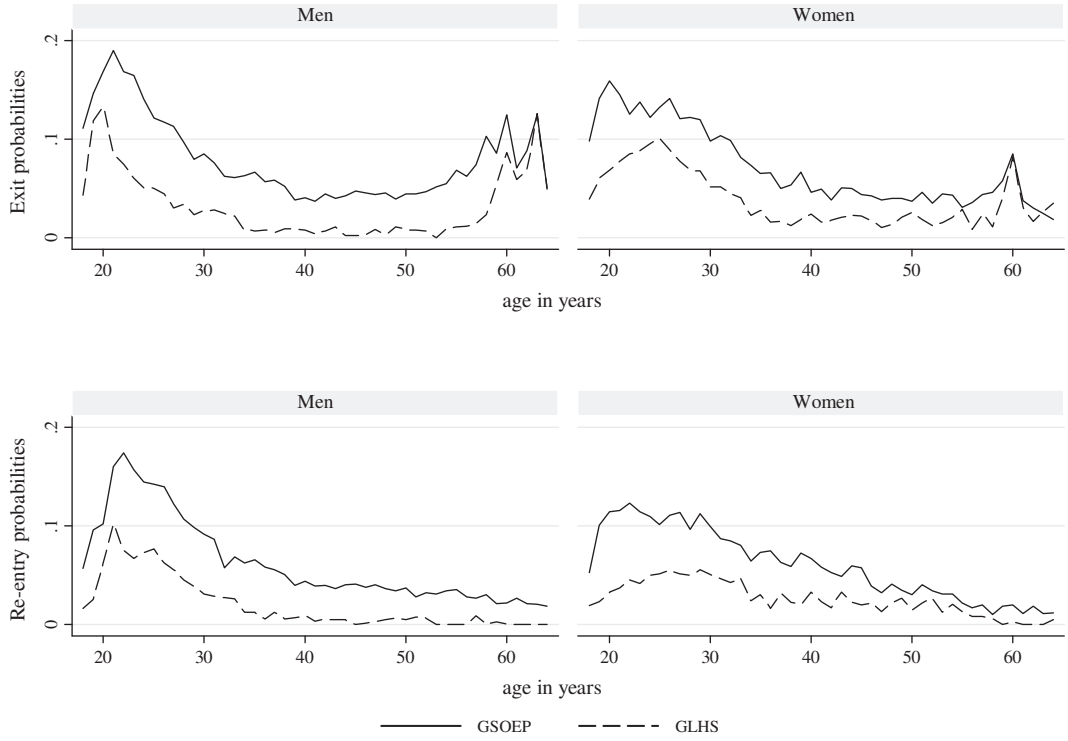
6. Results

First, we present some descriptive information on the differences across designs in the observed transition rates with respect to exit, re-entry and job-to-job moves. In a next step, we estimate explanatory models in which we aim to control for memory bias and seam effects.

6.1 Descriptive analyses

To acquire a first insight into the differences in the level of transitions across designs, we depict the exit and re-entry transition rates by age and sex.

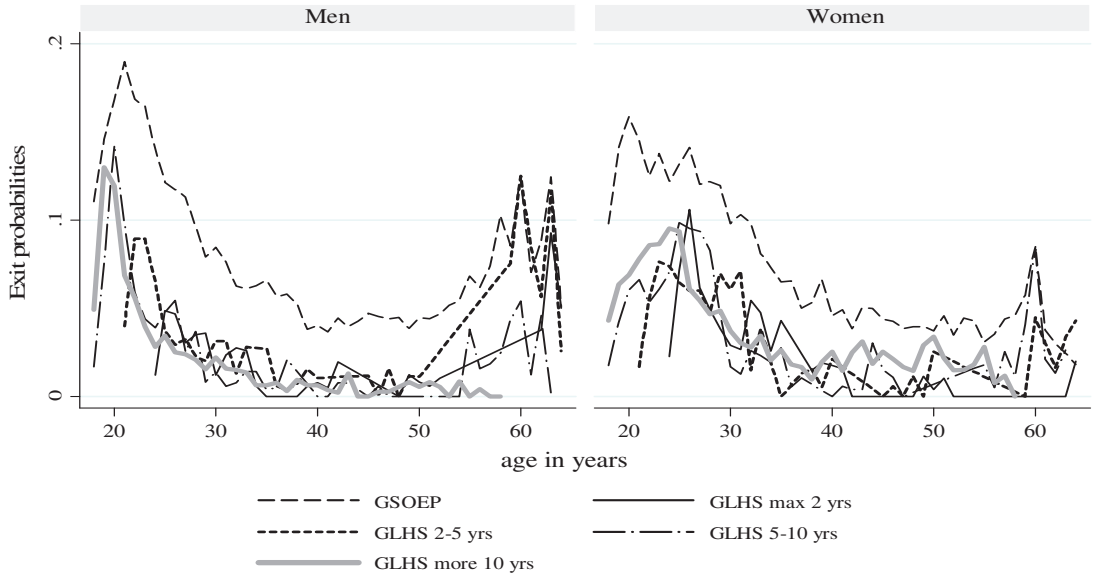
Figure 1. Exit and Re-entry probabilities by age and sex, GSOEP and GLHS



Both datasets show the same U-shaped pattern for employment exit chances for men but slightly flatter exit curves for women. Women seem to have higher exit chances at the beginning of their career. Transition probabilities appear lower for both sexes in the life history study though.

We contended that, due to recall bias, the level of transition probabilities in the life-course data is lower and the more so the larger the time distance to the interview. Here, we plot the transition probabilities by time to the interview for the GLHS.

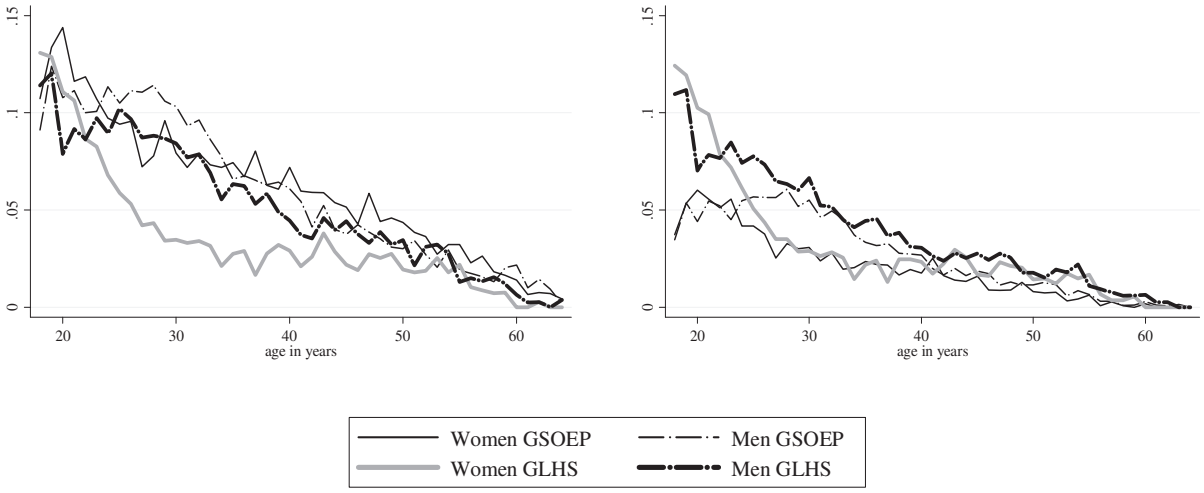
Figure 2. Exit probabilities by age and sex, GSOEP and GLHS by time distance to interview



The length of the recall period doesn't seem to play any significant effect (Fig. 2). We might expect the curves for the periods closer to the interview to be closer to the GSOEP one, whereas those referring to periods further away from the interview to be more distant. On the contrary, we do not observe much difference. Also for re-entry rates, both datasets reveal the same path, with the re-entry chances declining with age; the retrospective survey shows lower transition probabilities.

Figure 3. Job-to-job mobility by age.

Figure 3a. Employer change probabilities by age.



In the analysis of job mobility (Fig. 3), we try to make the job spell definitions in the two designs closer to one another by taking working time changes into account.[53] We might contend that our attempt works better for women, for whom a sizeable share of job-to-job moves involves a change of working time. This might not hold for men instead, whose kinds of job changes may be less likely to be captured by the definition adopted in the GSOEP (see 'Data' section). This may cause an underestimation of job moves for males in the GSOEP compared to the GLHS and could be responsible for the more similar transition probabilities in the two surveys. Moreover, job transition rates in the GLHS might be overstated due to the underreporting of short non-employment spells. The reason is that a reported job change in the GLHS should instead have been counted as an exit into non-employment and a re-entry into employment, especially when the worker changes employer, because in that case it is more likely that a short non-employment spell occurs in between. This might also explain the lower exit and re-entry transition rates in the GLHS. To check this last hypothesis further, we focus on employers' changes only, expecting them to be more likely to occur with a non-employment spell in between, making the transition rates particularly affected by the failure to report short non-employment spells.

Figure 3a clearly shows that employer change probabilities are indeed higher in the retrospective data and might therefore be responsible for our results on job-to-job mobility.

6.2 Multivariate analyses

Table 2 shows the coefficients associated with the dummy variable for survey, estimated with a model including (Model I) and one excluding (Model Ia) short spells out of employment.

In the life history study the odds of exit are about 27% lower for women and about 53% lower for men. Re-entry odds are instead 59% and 52% lower respectively. The design effects seem therefore rather strong. We believe that these differences may be to a great extent due to memory bias. In the following, we unravel it first by viewing the impact of short spells recall bias and the impact of the length of the recall period.

[53] We split job spells in the panel data in the case of a change from a part-time to a full-time job or vice versa, even when the respondent did not mention explicitly a job change.

The observed lower exit and re-entry rates in the GLHS seem partly due to the underreporting of short spells out of employment. Leaving them out, the design effect becomes much weaker, the odds of exit now being only 7% lower for women, though still 41% lower for men in the life-course study. Re-entry odds are now 36% lower for women and 17% lower for men. Although the differences have not disappeared, they are much smaller, showing that memory bias with respect to short spells is indeed affecting the transition rates in the GLHS.

Table 2. Selected parameter estimates for survey effect: Full model (I) versus no short spells model (Ia).

	Women		**Men**	
	Exit	***Reentry***	***Exit***	***Reentry***
Survey effect (ref: Gsoep)				
GLHS: Full model (Model I)	-0.254**	-0.900**	-0.749**	-0.724**
GLHS: Without short spells (Model Ia)	-0.076	-0.537**	-0.453**	-0.186**

*p<0.05; **p<0.01

With respect to job-to-job mobility, we make a distinction between employer and within-firm mobility. We contend that workers who change employer are more likely to experience a short unemployment spell in between than workers who change jobs within the firm. Hence, we expect the retrospective data to understate within-firm job moves but not employer's changes. Our results indeed support our expectations: within-firm mobility shows a strong significant negative effect for the retrospective design dummy, whereas employer changes exhibit opposite results, the retrospective design effect being insignificant and positive. Again, this witnesses the tendency to forget short (unemployment) spells in retrospective data designs.

Table 3. Estimated coefficients from multilevel multistate model of transitions out of (exit) and into (re-entry) employment and job-to-job changes.

	Women					Men				
Model I	*Exit*	*Reentry*	*Job-to-job*	*Employer*	*Within firm*	*Exit*	*Reentry*	*Job-to-job*	*Employer*	*Within firm*
Age	-0.017**	0.009**	-0.006**	-0.003**	0.007*	-0.034**	0.019**	-0.004**	-0.002	0.009**
Age squared (/10)	0.014**	-0.015**	0.026**	0.021	-0.100**	0.033**	-0.025**	0.063	-0.051**	-0.128**
Period (ref: 1983-1989)										
1900-1949	-0.302**	-0.037	0.468**	0.447**	0.012	0.274**	-0.991**	0.330**	0.333**	0.179
1950-1982	-0.420**	-0.500**	0.092	0.212**	-0.012	-0.486**	-0.007	0.046	0.139**	-0.392**
1990-1994	0.204**	-0.151**	0.123**	0.207**	0.046	0.175**	-0.105**	0.06	0.104*	-0.175*
1995-1999	0.149**	-0.125**	0.184**	0.171**	-0.314*	0.055	-0.127**	0.112**	0.184**	-0.620**
2000-2005	0.104*	0.047	0.212**	0.122	-0.748**	-0.035	-0.085	0.111**	0.196**	-0.963**
Lifecourse stage (ref: Single)										
Married	0.936**	-0.458**	-0.240**	-0.449**	-0.164	-0.286**	0.360**	0.006	-0.003	0.121
Small children	-0.013	-0.987**	-0.062	-0.273**	-0.568**	-0.061	0.039	-0.108**	-0.150**	0.048
Children more 6 yrs	-0.593**	0.552**	0.259**	0.435**	0.311	0.007	-0.006	-0.046	0.026	-0.108
Empty nest	0.662	0.435	-0.447	0.634	-18.922	-0.086	0.529	-0.322	-1.073	-0.448
Education (ref: Primary)										
Lower secondary	-0.086*	0.239**	0.034	0.057	0.215*	0.130**	-0.063	0.037	-0.064	0.446**
Upper secondary	0.488**	0.043	0.231**	0.149*	0.404**	0.816**	-0.554**	0.301**	0.029	0.685**
Tertiary	0.266**	0.542**	0.477**	0.398**	0.396*	0.151**	-0.219**	0.449**	0.404**	0.639**
Seam effect	0.869**	0.950**				1.044**	0.853**			
ISEI	-0.008**	0.005**	-0.005**	-0.008**	0.016**	-0.016**	0.008**	-0.006**	-0.009**	0.002
Survey (ref: Gsoep)										
GLHS	-0.254**	-0.900**	-0.628**	0.057	-1.180**	-0.749**	-0.724**	-0.231**	0.155**	-0.460**
Constant	-0.755**	-3.503**	-2.820**	-3.700**	-8.353**	9.414**	-6.089**	-3.281**	-3.982**	-8.641**
variance component (lev. 2)	-0.631		-0.734	-0.683	-0.234	-0.513		-0.721	-0.545	0.09
st. dev.	0.729		0.693	0.711	0.89	0.774		0.697	0.761	1.046
rho	0.139		0.127	0.133	0.194	0.154		0.129	0.15	0.249
Log Likelihood	-75178.8		-366660.9	-23750.0	-5238.7	-80846.9		-51643.1	-33369.6	-11101.6
df	41			19		41			19	
N obs	1260010			746871		1482910			1257695	
N groups	11312			9248		12744			11919	

*p<0.05; **p<0.01

In Model II, we examine the effect of the length of the recall period indicated by time distance to the interview (Table 4). In none of the models we find a significant effect of the length of the recall period. The coefficients associated with the spline parameters are mainly insignificant and certainly do not show a gradual decay over time.

Table 4. Selected parameters estimates for length of recall effects from multilevel multistate model of exit, re-entry and job-to-job changes. Model II and III.

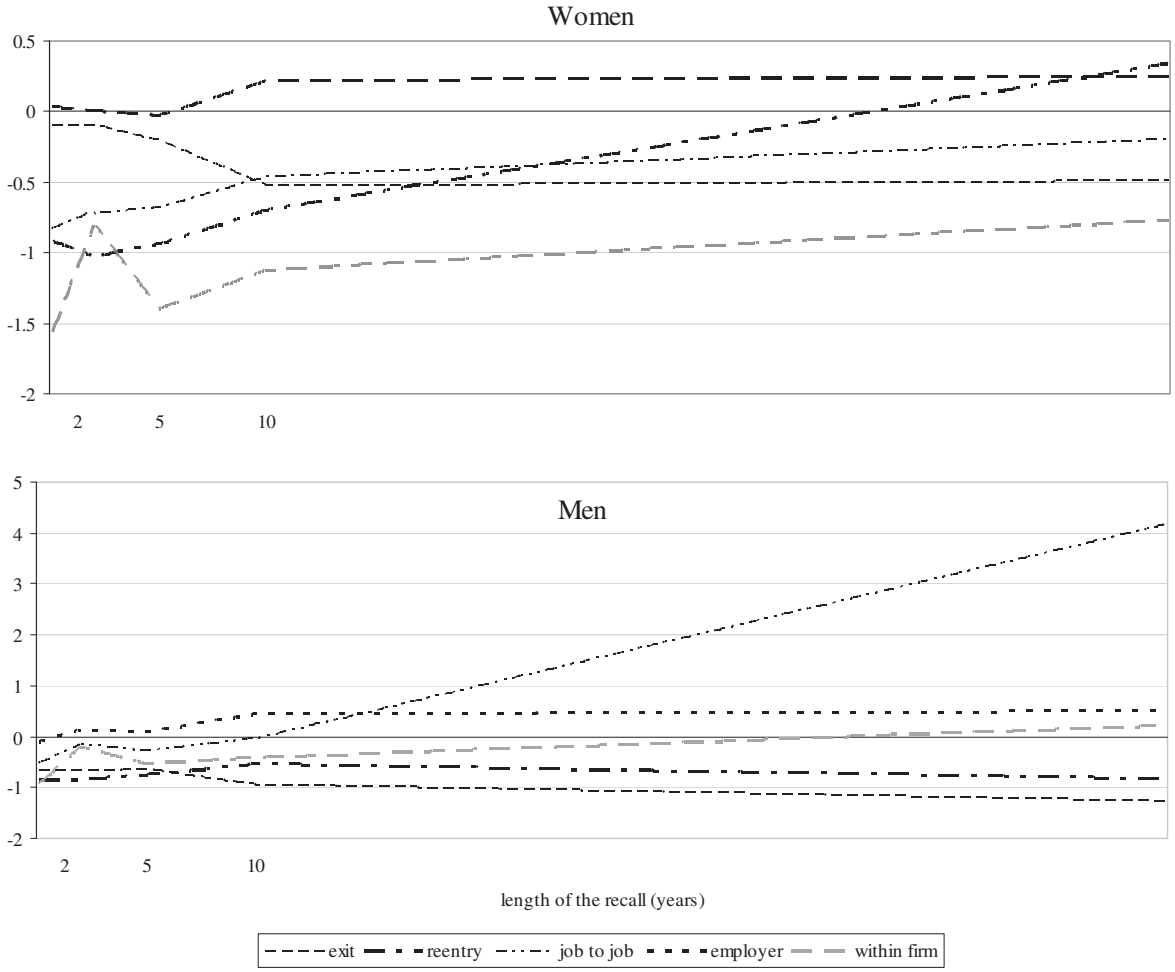
	Women									
	Exit		***Reentry***		***Job-to-job***		***Employer***		***Within firm***	
	Model II	*Model III*	*Model II*	*Model III*	*Model II*	*Model III*	*Model II*	*Model III*	*Model II*	*Model III*
Survey/Recall time (ref: Gsoep)										
GLHS	-0.1020	-0.0892	-0.9124**	-0.5887**	-0.83878**	-0.8168**	0.0383	0.0902	-1.6022**	-1.5561**
up to 2 years	-0.0001	0.0014	-0.0047	-0.038	0.0054	0.0061	-0.0015	-0.0004	0.0329	0.0326
2-5 years	-0.0031	-.0.0028	0.0025	0.0007	0.0014	0.0015	-0.0008	-0.0011	-0.0167	-0.0170
5-10 years	-0.0054**	-0.0053**	0.0041**	0.0027	0.0037**	0.0028*	0.004**	0.0030*	0.00466	0.0043
more than 10 years	0.0001	0.0001	0.0002	0.0003	0.0005	0.0005	0.0001	0.0001	0.0007	0.0006

	Men									
	Exit		***Reentry***		***Job-to-job***		***Employer***		***Within firm***	
	Model II	*Model III*	*Model II*	*Model III*	*Model II*	*Model III*	*Model II*	*Model III*	*Model II*	*Model III*
Survey/Recall time (ref: Gsoep)										
GLHS	-0.6753**	-0.5271**	-0.8521**	-0.6835**	-0.5408**	-0.5250**	-0.1097	-0.0923	-0.9736**	-0.9693**
up to 2 years	0.0001	0.0017	0.0005	-0.0011	0.0156**	0.0157*	0.0105	0.0105	0.0318*	0.0321*
2-5 years	0.0007	0.0001	0.0026	0.0026	-0.0033	-0.0034	-0.0013	-0.0013	-0.0093	-0.0098
5-10 years	-0.0051**	-0.0069**	0.0037*	0.0030*	0.0040**	0.0034*	0.0059**	0.0053**	0.0022	0.0023
more than 10 years	-0.0006	.-0.0003	-0.0059	0.0006	0.0008**	0.0008*	0.0001	0.0001	0.0013*	0.0013*

*p<0.05; **p<0.01

Figure 4 graphs the coefficients associated with the spline parameters from Model II. Differences in the levels of transition probabilities between the two datasets remain, but greater time distance to the interview doesn't seem to increase the gap. The shape of the memory error seems to fluctuate more closer to the interview date, while with longer recall time, recall errors do not increase. This tempts us to reject the assumption that recall of employment status information decays as a function of time (Hauser and Featherman, 1977). The conclusion must be that other design differences between the two surveys lead to lower transition rates in the retrospective study, which are not captured by the effect of the different time distance to the events that have to be recalled. The lengthening of the recall time indeed appears not to raise the gap in the estimated transition probabilities between the two surveys. The results also show that adding controls for time dependence and career complexity does not alter the results (Model III, Table 4).

Figure 4 Spline coefficients for length of the recall bias from multilevel multistate model of transitions out (exit) and into employment (re-entry) and job-to-job changes. Model II.



7. Discussion

The paper aimed at assessing the extent to which differences in survey design between life course and panel data affect the results of the analyses of labor market patterns in West-Germany. We theoretically discussed the differences between the two designs with respect to their sample design and composition as well as the various sources of measurement error and estimated multilevel multi-state event-history models on job-to-job, exit and re-entry mobility to gauge differences while controlling for the acknowledged sources of errors.

While the overall picture is fairly similar across the two designs, the estimated transition probabilities are quite different, being significantly lower in the life-course data. The survey design effect, examined by including a dummy for survey in the

models estimated on the two pooled datasets, turns out to be rather strong and significant. We contended that the underreport of short unemployment spells would account for it and that the longer the recall period the more the transition probabilities would be understated due to memory bias and, therefore, the longer the recall period, the larger the difference in the observed transition rates between the two designs. After controlling for short spells recall bias, the estimated differences in the transition probabilities between the two designs were reduced, while the lengthening of the recall period appeared not to increase the gap between the two surveys.

Our findings suggest that dissimilarities across the two designs cannot be simply accounted for by the factors we tried to correct for, such as length of the recall period and short unemployment spells, nor by compositional differences, which are ruled out by including the variables in which the two samples differ and by testing our results on the overlapping period as far as APC composition is concerned. They seem instead associated with other aspects which are not yet sufficiently discussed in the literature. One of these might be the way spells are derived or constructed from the two designs. In the GSOEP, where respondents have to recall a short period of one year, they are simply asked to report their monthly employment status, and spells are constructed by the researcher. In the GLHS instead, respondents are asked to identify all employment and non-employment spells in their life, and to provide the beginning and end dates of each spell, which may affect the difficulty of the reporting task, particularly for more complex careers.

We might also presume that age at interview plays a role: the older, the longer the period to recall and the higher the difficulties to recall. Unfortunately, the high correlation with the actual age does not allow us to control for that.

The methodological issues which we try to resolve in this paper and which are expected to explain the observed differences in the transition probabilities across the two longitudinal survey designs show that a deep insight into the survey design and into the way the information has to be recalled and reported is needed, especially in comparative research, to make plausible inferences about the prevalence of transition patterns. We also show that there are ways to control for some differences in the design of panel and life-course data, such as dealing with the length of the recall period, 'seam effects', and with the occurrence of short non-employment spells, to arrive at closer estimates of the transition probabilities, although they do not account for all the difference.

The paper discussed issues of data quality and data reliability using two very important designs in current sociological research dealing with longitudinal data. The findings warns against too easily formulated claims about the caveats of life-course studies compared to panel studies because the differences across the two designs seem more associated with rather general design differences than with the alleged larger recall or memory bias in retrospective designs.

The paper contributes to the existing literature for it not only sheds light on the sources of differences between the two designs by giving a detailed account of them, but also by trying to model the design effects and to disentangle memory bias effects from more general design differences. Furthermore, the paper reminds researchers of the importance of being aware of the specific design features of their data sources as well as their quality and reliability and the possible impact they might have on the inferences researchers generally want to make.

References

Allmendinger, J. (1989). Career Mobility Dynamics. A Comparative Analysis of the United States, Norway, and West Germany. Berlin: Max-Planck-Institut für Bildungsforschung

Bergemann, A. and Mertens, A. (2004). Job Stability Trends, Layoffs and Transitions to Unemployment: An Empirical Analysis for West Germany. London: *Centre for Economic Policy Research (CEPR), Discussion Paper* No. 4792

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, 208-225

Blossfeld, H. P., Rohwer, G. (1995). Techniques of event history modeling: new approaches to causal analysis. Mahwah,.Erlbaum

Bradburn, N. M. et al. (1987). Answering autobiographical questions: the impact of memory and inference on surveys. *Science* 236,157-161

Buchholz, S. (2004). Men's Late Careers and Career Exits in West Germany. *Globalife Workng Paper* No. 63

Buchholz, S., Grunow, D. (2003). Globalisation and Women's employment in West Germany. *Globalife Workng Paper* No. 47

Callegaro, M. (2006). Seam Effects in Panel Data Collected With Event Hsitory Calendar Methodology: The Case of Labor Force Transitions in the PSID

Dex, S. (1991). The Reliability of Recall Data: A Literature Review. *ISER Working Paper*. Colchester, University of Essex

DiPrete, T. (2002). Life Course Risks, Mobility Regimes, and Mobility Consequences: A Comparison of Sweden, Germany, and the United States. *American Journal of Sociology* 108 (2)

DiPrete, T., de Graaf, P., Luijkx, R., Tahlin, M., Blossfeld, H. P. (1997). Collectivist Versus Individualist Mobility Regimes? Structural Change and Job Mobility in Four Countries. *American Journal of Sociology* 103, 318-358

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: 508-532

Eisenhower, D. Mathiowetz, N. A. and Morganstein, D. (1991). Sources and Bias Reduction Techniques. In Measurement errors in surveys, pp. 127-144 New York, John Wiley and Sons

Elias, P. (1997). Who forgot they were unemployed? *ESRC Working Paper*

Ernicke, B. (1997). Datenaufbereitung aus dem Sozio-oeconomischen Panel: Erwerbsverlauf und Familiendaten. Bremen.

Freedman, D. et al. (1988). The Life History Calendar: A Technique for Collecting Retrospective Data. *Sociological Methodology* 18, 37-68

Hagenaars, J. A. (1990). Categorical longitudinal data: log-linear panel, trend, and cohort analysis. Newbury Park, CA. Sage Publications

Haisken-DeNew, J. P. Frick, J. R. (2005). DTC Desktop Companion to the German Socio-Economic Panel (SOEP)

Hauser, R. M., Featherman, D. L. (1977). The process of stratification: trends and analyses. New York Academic Press

Hoem, J. M. (1989). The issue of weights in panel surveys of individual behavior. Panel Surveys. Wiley

Horowitz, J. L., Manski, C. F. (1995). Censoring of outcomes and regressors due to survey nonresponse: identification and estimation using weights and imputations. *Econometrics. EconWPA*

Hujer, R., Schneider, H. (1989). The analysis of labor market mobility using panel data. *European Economic Review* 33, 530

Jacobs, S. C.: Reliability and recall of unemployment events using retrospective data. *Work, Employment and Society*, vol. 16, 537–548 (2002)

Jürges, H. (2005). Retrospective error in SOEP calendar data: the case of unemployment. *Schmollers Jahrbuch* 125, 157-165

Kalmijn, M., Luijkx, R.: Changes in Women's Employment and Occupational Mobility in the Netherlands between 1955 and 2000. In Blossfeld, H. P., Hofmeister, H. (eds.) (2006). Globalization, Uncertainty, and Women's Careers: An International Comparison, pp. 84-112. Cheltenham, UK, and Northampton (MA-USA), Edward Elgar

König, W., Lüttinger, P., Müller, W. (1988). A Comparative Analysis of the Development and Structure of Educational Systems. Methodological Foundations and the Construction of a Comparative Educational Scale. *CASMIN Working Paper* No. 12

Kraus, F., Steiner, V. (1998). Modelling Heaping Effects in Unemployment Duration Models -With an Application to Retrospective Event Data in the German Socio-Economic Panel. *Jahrbücher für Nationalökonomie und Statistiek*. Stuttgart

Kroh, M., Spieß, M. (2006). Documentation of Sample Sizes and Panel Attrition in the German Socio Economic Panel (SOEP) (1984 until 2005). DIW Data Documentation 15

Linton, M. (1982). Transformations of Memory in Everyday Life, 77-91 in Neisser, U. (ed.) Memory Observed: Remembering in Natural Contexts, San Francisco: Freeman

Lynn, P. et al. (2005). A Review Of Methodological Research Pertinent To Longitudinal Survey Design And Data Collection. *ISER Working Paper*. Colchester, University of Essex

Marsh, L. C., Cormier, D. R. (2002). Spline regression models. Thousand Oaks, CA. Sage

Mathiowetz, N. A., Duncan, G. J. (1988). Out of work, out of mind: Response errors in retrospective reports of Unemployment. *Journal of Business and Economic Statistics* 6, 221-229

Mayer, K. U. (2007). Retrospective Longitudinal Research: The German Life History Study. In Menard, S. (ed.) Handbook of Longitudinal Research: Design, Measurement and Analysis, Elsevier , San Diego

Mayer, K. U., Carroll, G. R. (1987). Jobs and Classes: Structural Constraints on Career Mobility. *European Sociological Review* 3, 14-38

Mertens, A. (1999). Job Stability Trends and Labor Market (Re-)Entry in West Germany 1984-1997

Muffels, R., Luijkx, R. (2008). Labour Market Mobility and Employment Security in Europe: 'Trade-off' or 'Flexicurity'? *Work, Employment and Society* 22 (2), 221-42

Müller, W. Lüttinger, P. and Wolfgang, K. (1989). Class and Education in Industrial Nations. *International Journal of Sociology* 19, 3-39

Paull, G. (2002). Biases in the reporting of labour market dynamics. London, The Institute for Fiscal Studies

Powers, E. A. et al. (1978). Congruence Between Panel and Recall Data in Longitudinal Research. *Public Opinion Quarterly* 42, 380-389

Pyy-Martikainen, M., Rendtel, U. (2003). The Effects of Panel Attrition on the Analysis of Unemployment Spells. CHINTEX

Reimer, M. (2005). Autobiografisches Gedächtnis und retrospektive Datenerhebung: die Rekonstruktion und Validität von Lebensverläufen. Studien und Berichte No. 70. Berlin, Max-Planck-Institut für Bildungsforschung

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, Berlin, Max-Planck-Institut für Bildungsforschung

Reimer, M., Künster, 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. Berlin, Max-Planck-Institut für Bildungsforschung

Rendtel, U. (1989). Über den Einfluß der Panelselektivität auf Längsschnittanalysen. Vierteljahreshefte zur Wirtschaftsforschung 58, 45–61

Rosenfeld, R. A. (1992). Job Mobility and Career Processes. *Annual Review of Sociology* 18, 39-61

Rubin, D. C., Baddeley, A. D. (1989). Telescoping Is Not Time Compression - A Model Of The Dating Of Autobiographical Events. *Memory and Cognition* 17, 653-661

Scott, J., Alwin, D. (1998). Retrospective Versus Prospective Measurement of Life Histories in Longitudinal Research. In Giele, G. H., Elder, J. Z. (eds.) Methods of Life Course Research: Qualitative and Quantitative Approaches. Thousand Oaks et al.: Sage Publications

Sicherman, N., Galor, O. (1990). A Theory of Career Mobility. *Journal of Political Economy* 98, 169-192

Singer, J. D., Willett, J. B. (2003). Applied longitudinal data analysis: modeling change and event occurrence. New York, Oxford University Press

Solga, H. (2001). Longitudinal Surveys and the Study of Occupational Mobility: Panel and Retrospective Design in Comparison. *Quality and Quantity* 35, 291 - 309

Steele, F. Goldstein, H. and Browne, W. (2004). A general multilevel multistate competing risks model for event history data, with an application to a study of contraceptive use dynamics. *Statistical Modeling* 4, 145-159

Sudman, S., Bradburn, N. M. (1973). Effects of Time and Memory Factors on Response in Surveys. *Journal of the American Statistical Association* 68, 805-815

Taris, T. (1996). A primer in longitudinal data analysis. London, Sage (2000)

Thompson, C.P. et al (1996). Autobiographical Memory, Mahwah, New York: Erlbaum

Tuma, N. B., Hannan, M. T. (1984). Social dynamics: models and methods. San Diego, Academic Press

Uunk, W., Mach, B. W. and Mayer, K. U. (2005). Job Mobility in the Former East and West Germany: The Effects of State-Socialism and Labor Market Composition. *European Sociological Review* 21, 393-408

Vandecasteele, L., Debels, A. (2007). Attrition in Panel Data: The Effectiveness of Weighting. *European Sociological Review* 23, 81-97

Wagenaar, W.A. (1986). My Memory: A Study of Autobiographical Memory over Six Years. *Cognitive Psychology* 18, 225-252

Chapter 4

Explaining labor market mobility patterns of West German men and women over the life course: A cohort study combining life-history and panel data *

Abstract . We examine labor market mobility patterns of West German men and women born between the 1920s and the 1980s. We combine the German Life History Study (GLHS) and the German Socio-Economic Panel (SOEP) to study the entire employment histories of these cohorts up to 2005, correcting for design biases and differences. Random effects models are estimated on labor market transitions using a multilevel discrete-time event history framework. Substantial effects of career investments and life course phase are found in both designs, although the strength of the effects turns out to be rather different. Correcting for recall bias and implicit anchoring does not remove the differences in strength, showing that they are associated with overall design differences. The findings confirm the alleged increased labor market volatility of the younger birth cohorts, while a high education appears less of a guarantee for a successful career for these cohorts than it used to be for the older.

**This chapter has been submitted for publication to Social Science Research, as Manzoni, A., Luijkx, R. and Muffels, R. Explaining labor market mobility patterns of West German men and women over the life course: A cohort study combining life-history and panel data.*

1. Introduction

The study concerns the employment histories of five birth cohorts of West German men and women born between the 1920s and the late 1980s correcting for design differences and possible sources of bias. Combining the two data sets allows us to extend the time period analyzed and to broaden the cohort composition of our sample. The use of the life-course data only would mean having information on the entire life courses of the pre-war birth cohorts but lacking information on the later careers of younger generations. The use of only panel data, instead, would not allow us to study the earlier careers of the older generations, because the data are censored at the first year of observation in 1984. By combining, we are able to create entire employment histories and life-course biographies for West German men and women born between 1918 (65 years old in 1983) and 1987 (18 years old in 2005).

The paper provides evidence on whether and to what extent survey design differences affect the significance, sign and strength of the major substantive effects on mobility. The main issue addressed concerns the impact of career investments indicated by human capital formation, employment history and life course phase on the labor market mobility patterns of various birth cohorts. The study shows how the combination of both data sources can tell us more about long-term trends in labor market mobility in Germany and the occurring shifts in mobility patterns across birth cohorts over time.

2. Theoretical and empirical perspective

Rising volatility and labor market institutions

Labor market mobility patterns are believed to change substantially over time due to changes in the institutional and socio-economic context as well as in people's values, norms and preferences. Life course research (Blossfeld et al., 2007) suggests that younger birth cohorts face more unstable careers than older cohorts also due to the functioning of internal labor markets protecting insiders (middle aged and older workers) and creating barriers for outsiders (younger workers) to enter steady employment (Lindbeck and Snower, 2001).

DiPrete (2002) elaborates the concept of *life course mobility regimes* to explain why particular labor markets exhibit lower or higher mobility rates than others. He especially assigns a large role to the welfare state (generosity and duration of

benefits), the labor market (employment protection and the wage bargain), as well as to training (apprenticeship) and the education institutions (preparing for specialized occupations) in society to explain cross-country differences. Historically, the West German institutional setting is featured by a strongly skill-based institutional labor market structure, highly protected employment statuses and long-term employment relationships causing stable returns to investments in skill formation well into the mid and later career stages (see, e.g., Gangl, 2004). The German educational system is highly standardized and stratified (Allmendinger, 1989) and with its well-known dual system it combines training on-the-job with vocational training. Standardization of occupational titles and certificates allows for flexibility between firms and within the same occupation, while due to the close link between educational certificates and occupational position (Shavit and Müller, 1998), flexibility between occupations is hindered (Kurz et al., 2002). This holds true even though the ideal of the stable uninterrupted occupational career has always been a norm restricted to the core worker in the German labor market (Grunow and Mayer, 2007). Despite some policy changes and reforms, there is still a high level of institutionalized regulation of the rights of core workers on the internal labor market in Germany.

Globalization, skill formation and 'scarring'

The alleged rising volatility of employment careers is assumed to mirror a shift from the standard 'life-time employment' into a more volatile and diverse career pattern, the so-called 'patchwork career' (Blossfeld, 2007) or the 'boundaryless career'(Arthur and Rousseau, 1996; Stone, 2006). This has been paralleled with the prolongation of time spent in education and training, the shift from industrial to service employment, the upgrading of the occupational structure and the higher prevalence of part-time and temporary work. Younger birth cohorts seem to pass through longer phases of vocational training but also unemployment during their life course than the older cohorts.

The larger volatility might also be closely associated with the stronger economic cycles due to globalization trends. During the last decades, because of globalization forces and the process of skill-biased technological change (Acemoglu, 2002; Buchholz et al., 2009; DiPrete, 2002; Mills, 2009), the demand for low-skilled labor declined, and firms were urged to react faster and in a more flexible way to market demand changes. At the same time, the erosion of the internal labor market (Gazier, 2000; Grimshaw et al., 2001) and the wider use of fixed-term contracts since the mid 1980s have led to more unstable employment careers, with more volatility and higher

risks of intermittent unemployment allegedly endangering the position of the low skilled in particular. Some authors point to the 'scarring' thesis suggesting that the longer the unemployment spells last (for the low skilled) the longer it takes to re-enter and the more it inhibits the chances to move into steady employment also due to human capital depreciation (Gangl, 2004, 2006; Barbieri, 2009). For the employed the reverse might also be true, since the longer they have stayed in the labor market or the longer tenure they have in their current job, the more opportunities they have had to build up firm-specific human capital through 'on-the-job' learning and training and the less likely they are to exit. DiPrete (2005) reviewing the situation in France and Germany argues that the low skilled are more likely to be assigned to the insecure temporary jobs instead of becoming unemployed, as the unified theory of Blau and Kahn (2001) predicts for the US, therewith arguing for a trade-off in Europe between wages or job rewards and employment security (DiPrete, 2005; Maurin and Postel-Vinay, 2005). This trade-off might also be relevant for labor market mobility over time due to increasing volatility among the least protected workers in temporary, small or part-time jobs. Furthermore, overall wage mobility tends to be lower and employment stability (indicated by tenure) larger in Europe compared to the US (DiPrete, 2002). This might especially be true for countries with strong employment protection and internal labor markets such as Germany.

A changing normative context

Due to changes in gender role patterns and social norm behavior, household formation processes altered and changed the life course trajectories of younger cohorts of women which, though still interrupt their work career at childbirth, exit for shorter periods (Fouarge et al., 2009). Household formation and labor market events are however closely interrelated. There is abundant evidence that these changes in fertility and marital behavior affect especially female labor market participation over the life course. Previous studies found for women strong positive effects of marriage on exit (Luijkx et al., 2006; Stier and Yaish, 2008; Tienda and Stier, 1996), especially for those having more traditional orientations towards work and care (Drobnič et al., 1999). The birth of the first child and of following children seem to act as 'trigger' events affecting labor participation and enrolment in part-time or full-time jobs (Fouarge et al., 2009). Marriage and childbirth might affect exit and re-entry of men and women in opposite direction. Men tend to increase participation after marriage or childbirth while women reduce it at least temporarily. Familial responsibilities tend to make men more but women less 'employable'. Germany has

witnessed a significant shift departing from a breadwinner welfare state from the 1940s on up to the early 1980s into a one and a half earner welfare state with women working part-time, with lower withdrawal rates and less reduction in their working hours at childbirth in the 1980s up to the mid 2000s (Statistiches Bundesamt, 1999).

Evidence from existing studies

Although the literature suggests that due to the aforementioned structural and cultural changes labor market mobility has increased over time, empirical evidence has lagged behind the theoretical proliferation on the argument (Biemann et al., 2009) and several studies on occupational mobility in Germany show limited evidence of increasing instability (Mayer, 2000, 2008; Winkelmann and Zimmermann, 1998). Also a recent study (Grunow and Mayer, 2007) renders little support for the alleged rising instability of occupational careers, although some empirical support is found for Germany concerning the increased prevalence of interrupted careers due to intermittent spells of unemployment. Furthermore, empirical evidence suggests that employment careers are much more stable over time than in theory sometimes is assumed and globalization cannot "be identified as a driving force behind the dynamics of individuals' careers" (Biemann et al., 2009).

Longitudinal data, especially panel but also life-history data, have been widely used to study labor market dynamics in Germany, but never in combination. Although in some studies survey design differences are examined (Buchholz, 2004; Drobnič et al., 1999), very little work has been done, to our knowledge, comparing the German life-history and the socio-economic panel study. An exception is the work of Solga (2001) for East Germany, though she only looked into the transition rates and not into their causes, and Manzoni et al. (forthcoming) for West Germany. Panel data have been exploited in Germany to study, for example, trends in job stability (Bergemann and Mertens, 2004; Winkelmann and Zimmermann, 1998), occupational choices (Constant and Zimmermann 2004), and the impact of the family situation on exit and re-entry (Drobnič et al., 1999). Life-course data have been widely used to study for example the (in)stability of German employment careers (Kurz et al., 2002), the employment insecurity and labor market attachment of different birth cohorts (Buchholz and Grunow, 2003), the effects of firm or organizational size, social class and industrial segmentation on job-shift patterns (Carroll and Mayer, 1986), gender inequality in the labor force across birth cohort (Trappe and Rosenfeld, 1998), job-shift patterns within and across firms (Uunk et al., 2005), the relation between class or job mobility (Mayer

and Carroll, 1987) and the life-course, and cohort and period effects in the status attainment process (Blossfeld, 1986). Although the results of these studies are often contradictory, they facilitated the selection of the relevant covariates to be included in our exit and re-entry models.

Accounting for design differences across the two surveys

One of our main concerns is how design differences across the two surveys might affect the substantive findings on the changes in transition patterns over time. In an earlier paper (Manzoni et al., forthcoming) we tested the reliability of life-history data for making inferences about the size of the transitions. We showed that transition rates are lower when using retrospective data compared to panel data, even after controlling for compositional differences as well as for possible sources of recall bias (seam effects, underreporting of short spells and length of the recall period). Correcting for these design differences appears needed and strongly reduces the dissimilarity across the two designs, though not entirely. The accuracy of the recall might also depend on general design differences between the two types of studies such as the amount of information that has to be recalled, the interviewing method and the way information is asked.. Finally, there are (random) differences due to sampling and attrition rates. Whilst panel data suffer from attrition, it plays hardly any role in retrospective surveys. Design differences might affect the rate of transitions as well as the sign and strength of the parameter estimates, for which reason we need to correct for these differences by including them as covariates in our explanatory models.

Life-history data are presumed to be less reliable than panel data due to initial selection and memory bias (Blossfeld and Rohwer, 1995; Powers et al., 1978; Reimer, 2005; Scott and Alwin, 1998): The longer recall period, due to having to reconstruct more events occurring further in the past, may affect recall accuracy. According to the literature, memory bias is supposed to be associated with career complexity as indicated by the number of transitions and short spells (Reimer and Künster, 2004). Since the transition probability itself might depend on career complexity, due to endogeneity, a weaker effect might be found than in reality exists. Recall accuracy is also affected by the salience of the event (Mathiowetz and Duncan, 1988). If the length of a spell is an indicator of its importance, short spells might be considered less important, and might therefore more often fail to be reported. Finally, recall bias might result from the anchoring of particular transitions close to major life events such as marriage or the birth of a child.

However, concerns for recall bias are raised for panel surveys too, since they also rely on retrospective questions asked on monthly employment statuses (Biewen and Wilke, 2005; Jürges, 2005, Kraus and Steiner, 1998; Paull, 2002).[54] The recall period is however much shorter, referring to the preceding calendar year only. Furthermore, panel surveys based on repeated-interviewing might suffer from artificial mobility, which shows itself in relatively small changes in response to a question for adjacent months within a reference period and more abrupt changes for adjacent months across reference periods.[55] These biases are usually referred to as 'seam effects', since heaps tend indeed to occur at the 'joint point' between waves (Solga, 2001).

2. Research questions and hypotheses

The substantive research question of the paper concerns the changing labor market mobility patterns and employment histories across successive birth cohorts. We are particularly interested in the long-term changes in labor market mobility across birth cohorts and different demographic life-stages indicated by fertility and marital behavior. We further examine the effects of career investments and employment history. Since Germany faced, after the reunification, a long period of slowed economic growth that affected labor market exit and re-entry, and knowing that the size of the transitions fluctuates with the business cycle, we also control for the effect of the GDP growth rate.

Hypothesis 1.

- Globalization, increased international competition and increased institutional support to working women have changed the labor market mobility patterns of younger birth cohorts, which tend to become more volatile, meaning that they experience more transitions over their life-course than older cohorts (1a).
- Younger cohorts exhibit different preferences for work and tend to combine work and care more than older cohorts (Hakim, 2001; Fouarge et al., 2009). We suspect that especially higher educated women want to combine work and care and therefore exit

[54] Information on the reliability checks on the GSOEP activity calendar are available upon request.

[55] Spurious transitions arise between calendars collected in subsequent years, which generate heaping in the transition probabilities between subsequent years (calendars). In the GSOEP, specifically, we observe heaps between December and January, which cannot be explained by cyclical factors alone (Kraus and Steiner, 1998).

less and re-enter more after childbirth or taking up caring duties than lower educated women, in particular those belonging to the younger cohorts (1b).
- Familial responsibilities connected to marriage and childbirth are expected to have opposite effects on exit and re-entry of men and women during their life-course. Women with (young) children, because of caring responsibilities, are more likely to exit from the labor market and have lower chances to re-enter. Men, due to the higher economic needs implied by having a family and children, are less likely to exit and more likely to re-enter after marriage or childbirth (1c).

Hypothesis 2.
- According to the thesis of skill-biased technical change and human capital theory, we expect that the more men and women have invested in their career through work experience or education and vocational training, as especially the younger cohorts do, the less likely they will exit the labor market and the more likely they will re-enter into employment over their career. For the same reason the low educated, especially when they belong to the younger cohorts, are expected to exit more and re-enter less (2a).
- Following the 'scarring' thesis, we suspect that due to human capital losses during the time out of work, non-employment inhibits re-entry into employment (Gangl, 2006; Luijkx and Wolbers, 2009). We therefore expect that the longer the total time spent in non-employment, the more the future career is 'scarred' by a lower probability to re-enter (2b).

Hypothesis 3.
- The highly institutionalized and stratified German labor market protects insiders and creates barriers for outsiders because of which the rising volatility on the labor market is concentrated among workers with lower human capital (low educated) experiencing higher exit and re-entry chances though in more unstable jobs (3a).
- During the 1990s, due to the reunification, Germany faced a long period of slowed economic growth with relatively high levels of unemployment because of which in particular the exit chances of the older cohorts will be high and their re-entry chances low (3b).

The second research question is more methodological and pertains to the degree to which the substantive results are affected by the survey design. Do prospective and retrospective data show different results with respect to the significance and strength

of the causal effects, after correction for differences in recall bias and design differences and, if so, to what extent? If the results for the main effects are similar across the two designs, we can safely combine the two datasets to estimate our event history models.

Previous findings in the literature point to the need to account for recall bias that shows itself in the anchoring of events around major life events or in a reduced recall accuracy with increasing length of the recall period in retrospective surveys (Manzoni et al., forthcoming) as well as in the occurrence of seam effects in panel data (Callegaro, 2007).

Hypothesis 4.

- We contend that major life events such as marriage or childbirth might act as implicit anchors[56] for respondents to temporally connect events affecting their recall accuracy. Our hypothesis is that people in the life history study have to remember events over a longer time horizon and are therefore more likely to anchor events around childbirth or marriage. We correct for this anchoring of employment transitions by including a time dummy indicating their occurrence in the period close to these life events. If anchoring exists the time dummies should have a positive and significant effect on exit and re-entry in the life history study compared to the panel. When the effects of the time dummies in the GLHS are insignificant and/or negative, we assume that there is no anchoring. .

The main line of reasoning in the paper is that correcting for design differences leads to more reliable estimates and reduces spurious differences in outcomes between the two surveys with respect to labor market mobility, making researchers more confident about using life-course data for studying long-term labor market changes. While correcting for design differences, we want to exploit the advantages of both datasets since they, combined, enlarge our scope for examining life course and birth cohort changes and extend our time perspective.

[56] By 'implicit' anchors we mean that such events are not explicitly mentioned in the questionnaire as landmark events to help people recall their work history, but respondents might use them as such in their mind.

3. Data

The GLHS and the GSOEP are both longitudinal multipurpose surveys, which collect data in a rather different way. The GLHS (Mayer, 2007) is a retrospective study on individual life courses of people belonging to specific birth cohorts (retrospective cohort study) in which respondents are asked, at a specific date, questions about their past career. It consists of a set of standardized interviews (face-to-face or telephone) with persons belonging to certain birth cohorts and drawn from representative samples, who were asked the monthly beginning and ending dates of each job or self-employment spell they had ever experienced. This allows the retrospective reconstruction of their complete career history.[57] We use the aggregate dataset (Gesamtdatenbank),[58] which contains data from four different surveys.

The GSOEP (Frick, 2005) is a panel study that started in 1984, in which the same respondents are interviewed (mostly face-to-face) annually. Information about employment history is only collected on a yearly basis. Questions are asked yearly about labor market and occupational changes based on the retrospective information. At each interview, respondents are asked whether they are currently engaged in paid employment and if so, in which occupation. 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. Using the monthly information gathered through a one year's retrospective calendar [59] instead, we are able to construct career patterns for at least a part of the life-course on a monthly basis. Although the individual information is available for all job spells, job specific information is only available if the respondent was actually in the job at the time of the interview. Further to this, information is available about the changes in working time as far as it concerns the distinction between part time and full time work.

---

[57] Questions were further asked about the occupation, the branch of industry, the size of the firm, and the wages at the start and the end of each job episode.

[58] Specifically, 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 to 1983 (LV-West I). In the years 1985-1987 another 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.Karl Ulrich Mayer and his collaborators kindly made this last version of the data available for our analysis.

[59] Respondents are shown a grid having as columns all the months in the previous year and as rows a large number of states. Respondents were asked to check for every month whether they were in each of the mentioned states.

For reasons of comparison with the GLHS dataset, we selected only respondents living in West Germany (including West Berlin), and we do not include foreigners, immigrants nor respondents from the East-German sample. Because we want to compare life history data with panel data, we are not making use of the biographical information.[60]

4. Method

To answer our research questions, we reconstruct employment spells for each respondent, defined as periods in which a person is employed and separated by non-employment periods. After pooling the two datasets, we estimate generalized discrete time multilevel event history models,[61] in which exit and re-entry[62] are the dependent variables indicating transitions during each month in the entire career. We use the discrete-time hazard function $h_j(t)$, which is defined as the conditional probability of having an event *j* during interval *t*, given that the respondent is at risk (Singer and Willett, 2003).

$$h_j(t) = \Pr(y_j(t) = 1 \mid y_j(t\text{-}1) = 0) \qquad [4.1]$$

The event history models[63] measure the odds of the event occurring, given that it has not occurred in the month before. 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.

[60] This information is only used in some cases to reconstruct initial conditions. Specifically, the time the respondent was in the initial (employment or not employment) state, to correct for left-censoring.
[61] A multilevel model is applied to allow for the hierarchical structure arising from repeated episodes 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-specific factors, which leads to unobserved heterogeneity in the hazard. Repeated events are usually handled by including individual-level random effects in an event history model.
[62] Exit from employment is defined in the last month of employment before a non –employment state while re-entry in the last month out of employment before re-entering it. We are aware that our dependent variable is highly aggregated and includes several age groups and very different reasons for a transition. Exit to retirement may for example be considered very different from other exit destinations. However, controlling for cohort and age should pick up a great share of the differences. Excluding exit to retirement doesn't indeed make much difference for our results. Also, controlling for the business cycle we partly correct for the involuntary nature of some transitions.
[63] We use the xtlogit command (with random effects) in Stata 10.

The model is estimated as a random effects logit panel regression model for entry into and exit out of employment conditional on stay out of or in employment.[64] The hazard is defined formally as:

$$h_{ij}^{E,N}(t) = \frac{1}{1+e^{-(\alpha^{E,N}(t)+\beta^{E,N}X_{ij}^{E,N}(t)+u_j^{E,N})}} \qquad [4.2]$$

where $h_{ij}^{E,N}(t)$ represent the hazards of exit (E) and of re-entry (N) in time interval *t* during episode *i* of individual *j;* the $X_{ij}^{E,N}(t)$ are covariates included in the exit and in the re-entry model and $u_j^{E,N}$ are random effects representing unobserved characteristics of individual *j*.

To estimate this model, we pool the two data sets and generate a person-month data file where a survey dummy indicates whether the data come from the panel or the life-history study. The dummy is expected to pick up the main sample and design differences across the two surveys. Since our main interest goes to changes in mobility by birth cohort and since we want to correct for sample composition differences related to birth cohort, we include an interaction term between the survey dummy and birth cohort. To further check whether the changes in mobility patterns differ by education level, we include an interaction between cohort and education level.[65] To check whether the substantive effects of the covariates ($X_{ij}^{E,N}(t)$) differ between the two surveys, we include their interactions with the survey dummy (Model 1). To test our hypothesis about the anchoring of labor market transitions to a marriage (Model 2) or a childbirth (Model 3) event, we include a time dummy indicating the occurrence of a transition in the 6 months period around the event that is interacted with the survey dummy to examine whether as we contend anchoring occurs more often in the life history study because people have to recall events over a longer time horizon.[66]

[64] Since our focus is on birth cohort differences and cohort is a constant that would be removed from the model we did not apply a fixed effects specification. A random effects specification is considered appropriate because the unobserved heterogeneity (effort, motivation, personality traits or ability) is likely not to be associated with birth cohort.

[65] Such interactions are not included in the models for women given to not significant results.

[66] Model 1 equals Model 2 (and 3) but doesn't include marriage (or childbirth). No substantive differences were found between these models in the effects of the covariates. Consequently, results from Model 1 are not presented, although available from the authors upon request.

The random effects logit model is defined as:

$$\text{logit}\left(\frac{h_{ij}^{E,N}(t)}{1-h_{ij}^{E,N}(t)}\right) = \alpha^{E,N}(t) + \beta_1^{E,N} SURVEY_{ij}(t) + \beta_2^{E,N} COHORT_{ij} + \beta_3^{E,N} ANCHOR_{ij}^{mar/chb}$$
$$+ \beta_4^{E,N} COHORT_{ij} * SURVEY_{ij}(t) + \beta_5^{E,N} COHORT_{ij} * EDUC_{ij}(t)$$
$$+ \beta_6^{E,N} X_{ij}^{E,N}(t) + \beta_7^{E,N} X_{ij}^{E,N} * SURVEY_{ij}(t) + \beta_8^{E,N} SURVEY_{ij} * ANCHOR_{ij}^{mar/chb} + u_j^{E,N} \qquad [4.3]$$

where SURVEY is the survey dummy (GLHS=1); COHORT*SURVEY the interaction between cohort and survey, COHORT*EDUC the interaction between cohort and education level, $X_{ij}^{E,N}(t)$ the main effects of the (substantive) covariates and $X_{ij}^{E,N}$ * SURVEY the interaction term between these covariates and the survey dummy. The SURVEY*ANCHOR term is the two-way interaction term between the survey and the anchoring around marriage (mar) or childbirth (chb), and $u_j^{E,N}$ the random effects. The main effects show the results for the SOEP panel and the interactions provide evidence on the extent to which the results for the life history data (GLHS) are different from the panel estimates.

Eventually, in Model 4, we leave out the interaction terms between the survey dummy and the substantive covariates.

Since mobility patterns and employment histories are very different for men and women, and the effects of the covariates may differ across gender, all our models are estimated separately for men and women.[67]

5. Measures included in the empirical model

Our models include a set of substantive measures referring to career investment, measured by human capital and employment history, and life course stage.

- Career investments partly deal with investments in human capital through formal education. Education level is measured by the CASMIN classification (Müller et al., 1989) recoded into 5 categories.[68] The resulting categorical variable is included as a time varying covariate, indicating the highest level of education at each point in time (Manzoni et al., forthcoming).

[67] Models have also been estimated to test whether there are significant differences across gender; results are available from the authors upon request.

[68] Categories 1a, 1b and 1c are merged into primary education, 2a and 2b into lower secondary, 2c gen is upper secondary, 2c voc is upper secondary vocational, 3a and 3b into tertiary.

- Also age might be seen as an indicator for the current level of human capital. We suspect an inverse U-shaped relationship, because of which we also included age squared. Worker's human capital increases with age up to a certain threshold after which it declines due to the obsolescence of human capital, therewith raising exit (reinforced by generous early-retirement options) and reducing re-entry. Age also impacts the timing of life course decisions associated with e.g. marriage and childbirth and the occurrence of life events. Finally, it controls for compositional differences associated with age across the two samples.
- Duration dependence is measured through the effect of time in a particular state (spell duration) and coded as a categorical variable with four categories (<13 months, 13-24 months, 25-48 months, >48 months). This variable indicates another aspect of career investment or disinvestment, i.e. the amount of human capital built up through on-the-job learning and the depreciation of human capital when out of employment.
- Other indicators for employment history, referring to the current (exit) or previous job spell (re-entry), are: occupational status, measured by the ISEI score of the current (exit) or the most recent (re-entry) occupation;[69] sector (public versus private) and part-time employment.[70]
- Employment history is further indicated by the number of job spells people had, acting as an indicator for the volatility of the career as well as its complexity. We include it as a time–varying covariate, 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, to correct for the high correlation with age (Luijkx et al., 2006).
- Information on life course stage is measured by marital status and number and age of children. These measures are time varying. We distinguish single, married/cohabiting and divorced/separated people whereas a continuous variable specify the number of children. A dummy indicates the presence of any children aged 0-5 and another one indicates whether the household contains only older children (6-16).

[69] In the GSOEP, the ISEI code, as all the specific characteristics related to the job, is only available if a job was held at the time of interview so that we do not have any information for (relatively short) jobs held in between two interviews. For jobs for which no ISEI information was available we have attributed the ISEI of the most recent occupation.

[70] In the GLHS we use the number of working hours to define part-time employment (using 35 hours as cut-off point); in the GSOEP instead the respondent's self reported status is used, since working hours are not available on a monthly basis.

- One of our main hypotheses deals with differences in mobility patterns across cohorts. We distinguish 5 cohorts (<1932; 1932-1945; 1946-1953; 1954-1963; >1963).
- From a substantive point of view, we are interested in age, cohort and period effects. Due to the well-known linearity in the relationship among them, we cannot simply include them simultaneously in our models. Given the importance of age and our main interest in cohort differences, we focus on these two. However, age and birth cohort partly capture the effect of period and also the two surveys refer to different periods. The panel survey covers the period from 1983 up to 2005 whereas the Life History Study reconstruct the time from the 1930s to 1999.
  Our supply data lack information at the demand side. However, the over time varying opportunities and constraints workers face due to the business cycle or structural shifts in demand are important (Rosenfeld, 1992). Therefore, we add a time-varying variable controlling for business cycle effects indicated by the annual GDP growth rate.[71]

We further include various measures to correct for recall bias in both datasets and a dummy for survey, which accounts for all other design differences across the datasets.

- To control for seam effects a dummy indicating the month of December in the panel is included.
- To control for the length of the recall period in the life-course study, spline functions with knots at 24, 60 and 240 months are used, which allow for continuous (monthly) changes in the effect of the recall time, avoiding therewith abrupt drops (Manzoni et al., forthcoming). Presuming that increasing length of the recall period might increase recall problems affecting the likelihood of events (Bradburn et al., 1987; De Graaf et al; 1989, Hauser and Featherman, 1977; Mathiowetz and Duncan, 1988; Mayer, 2007; Paull, 2002; Powers et al., 1978; Reimer, 2005; Rubin and Baddeley, 1989; Sudman and Bradburn, 1974) it is important to control for it. Another reason is that controlling for it assures that the effect of other characteristic on the hazard is not due to the different distribution of the various groups by the length of the period recalled.
- To control for the anchoring of transitions around major life events we also include time dummies indicating the 6 months time period around marriage or childbirth, that is the three months before and the three months after it.

[71] A dummy variable is introduced for the period before 1950, since no data is available, and for year 1970, due to discontinuity in the time series.

To make the interpretation of the coefficients more meaningful, all the continuous variables are mean centered, so that the presented effects refer to the average respondent.

6. Results

6.1. Descriptive findings: sample characteristics

Table 1 gives an overall picture of the observed characteristics in our sample and their differences in the two datasets, separately for men and women.

Table 1. Sample composition: descriptive characteristics

	GLHS		GSOEP	
	Men	Women	Men	Women
Person-time level				
Spell duration (months)				
employment	120	93	194	121
non-employment	29	97	37	115
Employed (%)	87.68	61.77	60.71	82.54
Exit rate	0.46	0.57	0.88	0.75
Re-entry rate	0.41	0.36	0.84	0.64
Education				
Primary	63.42	60.29	50.22	47.98
Lower secondary	19.04	27.32	19.76	30.09
Upper secondary	4.38	3.11	4.19	3.36
Upper secondary voc	4.47	4.82	5.61	6.16
Tertiary	8.67	4.43	19.28	11.58
Life-course (%)				
Single	39.52	31.42	18.11	12.15
Married	57.68	60.57	77.61	80.08
Divorce/widow	2.80	8.01	3.83	7.35
Nr. Children	0.69	0.78	0.66	0.71
Having small children (%)	23.67	27.47	13.78	14.84
Having only older children (%)				
age (in years)	32.0	32.0	41.9	41.6
isei	40.84	43.23	45.77	43.03
public sector (%)	21.59	24.93	24.45	30.20
part-time (%)	1.30	9.95	3.04	27.69
Complexity/Volatility	0.074	0.194	0.002	0.154
Person level				
Birth Cohort (%)				
1964-…	35.82	33.55	34.20	35.50
1954-63	27.86	26.03	23.01	23.40
1946-53	9.79	9.95	14.46	13.94
1932-45	13.44	12.47	22.20	20.90
…-1932	13.09	17.99	6.13	6.26
Time (year) level				
GDP	6.56 (st.dev 3.87)		3.91 (st. dev 2.29)	
N (person-months)	670700	695847	846994	858334
N (respondents)	3431	3407	9537	9653
N (years)	51		23	

Some characteristics are time-constant and reported at the person level (e.g. birth cohort). Others are instead time-varying (e.g. age, marital status) and are reported at the person-time (i.e. person-month) level. Moreover, we also report a 'context' variable (i.e. GDP growth), that only changes over time (year), in the same way for all the individuals.

Due to the mentioned design differences, the composition of the sample by age, period and cohort is quite different. Table 1 shows much lower transition rates in the retrospective study (ranging from 40 to 60% compared to 60 to 90% in the panel) where men experience fewer exits and re-entries. Furthermore, we observe more singles, fewer married people and people with small children in the life course study. For the other characteristics differences are much smaller.

6.2. Multivariate analyses: testing the differences across designs

In Table 2, we show the results from Model 2, testing the differences in the effects of our covariates across surveys.[72]

Our substantive focus is on the effect of human capital investments and employment history (age, education level, time in employment, sector, working time and occupational status) and of changes in familial responsibilities associated with changes in the life course stage (marital status, number and age of children).

Although the hypothesis that the effects in the two surveys are equal must be rejected both for men and women and for exit and re-entry, we found few significant differences in the sign of the substantive effects. After correction for design differences and recall bias indicated by recall length, seam effects, the anchoring of events and career complexity, the differences in the substantive effects are strongly reduced and refer mainly to differences in the strength of the effects.

We find strong evidence for negative duration dependence in both surveys, meaning that the longer the time in employment or out of employment, the lower the exit and the re-entry chances respectively. The effects are much weaker in the life-history study though (the interaction effects are significant), This is likely due to the overestimation of the spell duration as a consequence of the underreporting of short spells out of employment (Manzoni et al., forthcoming).

[72] The main effects of each parameter in Model 2a and 2b, presented in the left columns, represent the effects in the SOEP panel while the interaction effects in the right column show whether the life history estimates differ (significantly) from the panel estimates.

Table 2. Multilevel discrete time event history models (random effects logit) for exit and re-entry from the labor market. Effects for GSOEP and differences of GLHS with GSOEP.

	EXIT (Mod 2a)				*RE-ENTRY (Mod 2b)*			
	Women		***Men***		***Women***		***Men***	
	GSOEP	**GLHS interaction**	**GSOEP**	**GLHS interaction**	**GSOEP**	**GLHS interaction**	**GSOEP**	**GLHS interaction**
Spell duration (ref: <13 months)								
13-24 months	-0.417**	0.307**	-0.684**	0.541**	-0.606**	0.163*	-0.504**	1.345**
25-48 months	-0.708**	0.544**	-1.024**	0.766**	-0.786**	0.444**	-0.793**	1.135**
>48 months	-0.942**	0.864**	-1.658**	1.024**	-1.357**	0.583**	-1.303**	2.207**
Education (ref: Low)								
Lower secondary	-0.068	0.102	-0.072	0.554**	0.145**	0.006	-0.197**	-0.145
Upper secondary	0.554**	0.691**	0.493**	0.873**	-0.009	-0.827**	-0.167*	-1.671**
Upper secondary vocational	0.048	0.146	0.065	0.867**	0.244**	-0.330*	-0.036	-0.679**
Tertiary	0.257**	0.23	0.042	0.426**	0.568**	-1.032**	0.508**	-2.034**
Marital status (ref: married)								
divorced/widow	-0.107	-0.436**	0.235**	-0.079	0.220**	0.677**	-0.291**	-0.104
single	-0.530**	-0.595**	0.138**	0.102	0.350**	0.607**	-0.259**	0.02
Nr. children	-0.004	-0.271**	-0.002	-0.053	0.049*	-0.230**	-0.055	0.05
Any small children (<6)	0.263**	0.316**	0.038	-0.29	-0.643**	-0.067	0.104	0.021
Only older children (6-16)	-0.081	-0.154	-0.086	0.094	-0.096	0.019	0.145	-0.461*
age	-0.002**	-0.001	-0.001	-0.002**	-0.002**	0.002**	-0.004**	-0.003**
age squared (/100)	0.002**	-0.001**	0.002**	0.002**	-0.002**	0.0004	-0.003**	-0.002**
GDP growth	0.016*	-0.013	-0.002	-0.011	0.009	-0.004	0.015	-0.008
Birth Cohort (ref: 1964-..)								
1954-1963	-0.031	-0.09	0.031	0.014	0.036	-0.085	0.246**	-0.218*
1946-1953	-0.113	-0.214	-0.103	-0.054	0.114	-0.662**	0.402**	-0.708**
1932-1945	0.204*	-0.453**	-0.044	-1.047**	0.028	-1.425**	0.690**	-0.412*
<1932	0.586**	-0.843**	0.521**	-1.023**	-0.115	-1.610**	-0.114	0.525
Current/last job attributes								
isei	-0.013**	-0.00002	-0.019**	0.002	0.003*	0.010**	0.0002	0.009**
public sector	-0.215**	0.018	-0.088*	-0.254**	-0.046	0.09	0.023	-0.283**
part-time	0.165**	-0.207**	0.464**	-0.295	0.187**	-0.045	0.062	0.054
seam effect	0.773**		1.111**		1.158**		0.943**	
Complexity/Volatility	0.054**	-0.015	0.076**	-0.048**	0.074**	-0.040*	0.036**	-0.005
Survey: Glhs	-0.471**		-1.607**		-0.815**		-0.645**	
Recall time(Glhs)								
max 24 months		0.005		-0.003		-0.001		0.002
25-60 months		0.001		0.001		0.003		0.003
61-120 months		-0.006**		-0.011**		0.003		0.002
more than 120 months		-0.0002		0.003**		0.004**		-0.004**
*Education*Cohort*								
Lower secondary*1954-1963			-0.158				0.15	
Upper secondary*1954-1963			-1.021**				1.105*	
Upper sec voc*1954-1963			-0.621*				0.957*	
Tertiary*1954-1963			-0.316				1.045**	
Lower secondary*1946-1953			0.220*				0.115	
Upper secondary*1946-1953			-0.447				0.067	
Upper seco voc*1946-1953			0.296				0.835**	
Tertiary*1946-1953			-0.043				-0.111	
Lower secondary*1932-1945			-0.237				0.235	
Upper secondary*1932-1945			-0.052				-0.179	
Upper sec vocl*1932-1945			-0.273				0.1	
Tertiary*1932-1945			-0.29				0.107	
Lower secondary*<1932			-0.193*				-0.009	
Upper secondary*<1932			-0.001				-0.186	
Upper sec voc*<1932			-0.089				-0.278	
Tertiary*<1932			-0.054				-0.223	
Constant	-4.243**		-4.514**		-2.674**		-2.264**	
Variance lev. 2 (person)	-1.318**		-1.953**		-0.983**		-0.840**	
St. dev. random intercept	0.517		0.377		0.612		0.657	
Rho	0.075		0.041		0.102		0.116	
N	835288		1132039		598800		232555	

*p<0.05; **p<0.01

The coefficients come from a model also controlling for the anchoring effects around marriage and its interaction with cohort (coefficients presented in Table 3)

For men and women with upper secondary general education the odds of exit and re-entry are in both surveys significant and pointing in the same direction but with much stronger positive effects for exit[73] and much stronger negative effects for re-entry[74] in the life-history study. However, for the higher educated the results are rather different across surveys: While in the panel study higher educated women and men with tertiary education show larger odds of re-entry, the results for the life-history study are opposite. But also for men and women with lower secondary education, the odds of exit are different; negative in the panel and positive in the retrospective survey, although not always significant.

The effects of the occupational status (ISEI) point in the same direction in both surveys, although the odds for re-entry are stronger positive in the retrospective study. The effect of age, which is considered a human capital indicator, points in the same direction in the two designs. However, the birth cohort effects are quite different across the designs in size as well as in strength, signaling sample design differences. Older cohorts seem to have higher odds of exit and re-entry in the panel but much lower odds of exit and re-entry in the life history study, both for men and women.

Employees in the shielded public sector are less likely to exit in both surveys. Also the positive effect of part-time work on exit turns into a negative effect in the retrospective study at least for women. The odds for re-entry are positive for part-timers and similar in both designs.

Career complexity, another aspect of the employment history, has a positive effect on exit and re-entry in both surveys. The effects though are weaker in the life history study, particularly for men's exit and women's re-entry behavior. It appears that, although career complexity leads to a higher exit and re-entry and hence to a higher volatility, it also makes recall more difficult leading to the underreporting of job spells in the retrospective study. Consequently, the effect on exit and re-entry is less positive in the life history data.

The effects of life course stage or familial responsibilities on labor market behavior point in the same direction, although with different strength, at least for women. For single compared to married women, the odds of exit are lower and those of re-entry

[73] For men, this holds at least born after the WWII, while for those born before the panel survey doesn't show positive effects on exit.

[74] In the panel survey only the oldest cohort of men shows positive effects on re-entry.

higher with stronger negative and positive effects respectively in the retrospective study.
The number of children shows a significantly stronger negative effect on exit and re-entry for females in the life-history study. The presence of small children below 6 exerting a positive effect on women's exit appears much stronger in the retrospective data. For men the effect of the presence of older children seems to hinder re-entry especially according to the retrospective study.

Finally, with respect to our methodological concerns about recall bias in the two surveys, the seaming effects in the panel indeed show a strongly significant and positive effect on exit and re-entry. The effect of length of the recall period indicates that for longer recall periods in the life history study the odds of exit and re-entry are only significantly lower for women for the '5-10 years' and for men for the 'more than 10 years' period. Correcting for seam effects and the length of the recall period is therefore needed to improve the fit of the model but also to control for the rather different sources of recall bias in both surveys.

6.3. Anchoring of transitions around major life events

We hypothesized (hypothesis 4) that the differences are also related to the anchoring of transitions to marriage and childbirth. The effects are presented in Table 3, where the effects of the time dummies for marriage and child birth for exit and re-entry, computed from the estimation results of Models 2 and 3, are shown. The coefficients show the main effect of the timing variable for marriage or childbirth in the panel study and the difference with that effect in the life history study. The coefficients point to anchoring when the effects of the time dummies for marriage or birth showing the difference of the effect in the life history study compared to the panel study are significantly positive indicating that in the period around the occurrence of major life events differences in transition rates across datasets are reduced or cancelled out.
The results reported in Table 3 show little evidence for the downsizing effect of anchoring on the design differences. More specifically, childbirth doesn't seem to act as an anchoring event in any of our models, while some evidence is found for anchoring in the case of marriage in the exit model for women (survey differences are cancelled out) and in the re-entry model for men (survey differences are reduced). A possible explanation for the relative unimportance of anchoring to explain the design

differences might be that salient events in a person's life improve the respondent's ability to date transitions but do not influence the estimation of their occurrence.

Table 3. Anchoring effects by birth cohort: effects of the time dummy for marriage and childbirth.

	EXIT				*RE-ENTRY*			
	Women		***Men***		***Women***		***Men***	
	GSOEP	**GLHS interaction**	**GSOEP**	**GLHS interaction**	**GSOEP**	**GLHS interaction**	**GSOEP**	**GLHS interaction**
	Mod 2a				*Mod 2b*			
Marriage	0.518**	0.573**	-0.174	0.268	0.03	-0.25	-0.069	0.564**
Childbirth	3.916**	-0.849**	0.208	-0.342	-1.005**	-0.157	0.006	0.137

*p<0.05; **p<0.01

Coefficient are from Model 2 (for marriage) and Model 3(for childbirth) as specified in the method section. The effects of the other variables of Model 2 are shown in Table 2. The estimates of Model 3 are largely similar to those in Model 2 and are therefore not shown

The general lesson to be learned from these findings is that the substantive results across the two datasets after correcting for seam effects, recall bias and anchoring are generally quite similar even though the strength of the effects is rather different, which seems attributable to the different design and composition of the samples. The former analysis is important to understand and explain the observed differences across the two longitudinal studies and to correct for their design differences. More important though is that it justifies combining the datasets for examining the career patterns of several birth cohorts over their entire life course.

6.4 Examining career patterns combining panel and life course data

To arrive at more robust conclusions on career mobility across birth cohorts, we subsequently viewed the effects of our substantive covariates estimating random effects models on the combined datasets.[75] This allows us to observe the entire life-course for several birth cohorts, correcting for biases due to each specific survey design. In Table 4, the estimation results of Model 4 are reported.

[75] Effects are assumed not to differ across datasets and therefore the interactions with the survey dummy are excluded from these models. We only include an interaction between cohort and the survey dummy and, in the model on men, an interaction between education and cohort.

Table 4. Multilevel discrete time event history models for exit and re-entry from the labor market. Pooled GSOEP and GLHS

	EXIT (Mod 4a)				*RE-ENTRY (Mod 4b)*			
	Women		***Men***		***Women***		***Men***	
Spell duration (ref: <13 months)								
13-24 months	-0.291**		-0.458**		-0.571**		0.002	
25-48 months	-0.465**		-0.720**		-0.670**		-0.484**	
>48 months	-0.547**		-1.366**		-1.223**		-0.424**	
Education (ref: Low)								
Lower secondary	-0.080*		0.115*		0.175**		-0.200**	
Upper secondary	0.846**		0.699**		-0.222**		-0.466**	
Upper secondary voc	0.043		0.392**		0.210**		-0.244**	
Tertiary	0.209**		0.135		0.413**		-0.027	
Marital status (ref: married)								
divorced/widow	-0.313**		0.200**		0.437**		-0.361**	
single	-1.026**		0.210**		0.562**		-0.278**	
Nr. children	-0.047*		0.002		-0.004		-0.069*	
Any small children (<6)	0.260**		-0.103		-0.710**		0.162**	
Only older children (6-16)	-0.213**		-0.073		-0.094*		0.102	
age	-0.003**		-0.001**		-0.002**		-0.005**	
age squared (/100)	0.001**		0.003**		-0.001**		-0.002**	
GDP growth	0.01		-0.007		0.01		0.01	
Birth Cohort (ref: 1964-..)	Gsoep	Cohort*Glhs	Gsoep	Cohort*Glhs	Gsoep	Cohort*Glhs	Gsoep	Cohort*Glhs
1954-1963	-0.069	-0.056	0.176**	-0.197*	0.035	-0.073	0.321**	-0.286**
1946-1953	-0.107	-0.368**	0.08	-0.373**	0.058	-0.600**	0.481**	-0.702**
1932-1945	0.313**	-0.727**	0.084	-1.307**	-0.125	-0.990**	0.864**	-0.570**
<1932	0.754**	-1.228**	0.482**	-1.273**	-0.323*	-0.621**	0.384	-0.043
Current/last job attributes								
isei	-0.009**		-0.015**		0.004**		0.003*	
public sector	-0.214**		-0.180**		-0.027		-0.058	
part-time	0.109**		0.438**		0.190**		0.106*	
Complexity/Volatility	0.063**		0.072**		0.065**		0.042**	
seam effect	0.759**		1.107**		1.157**		0.944**	
Survey: Glhs	-0.023		-0.179		-0.634**		-0.319*	
Recall time(Glhs)								
max 24 months	0.007		0		0		-0.001	
25-60 months	-0.001		0.002		-0.001		0	
61-120 months	-0.008**		-0.010**		0.004*		0.003*	
more than 120 months	0.001		0.003**		0.001**		-0.004**	
*Education*Cohort*								
Lower secondary*1954-1963			-0.173				0.039	
Upper secondary*1954-1963			0.274*				-0.563**	
Upper sec voc*1954-1963			-0.621*				-0.258	
Tertiary*1954-1963			-0.137				-0.361**	
Lower secondary*1946-1953			-0.316*				0.322	
Upper secondary*1946-1953			0.042				-0.673*	
Upper seco voc*1946-1953			-0.558*				0.094	
Tertiary*1946-1953			-0.453**				0.041	
Lower secondary*1932-1945			0.034				0.136	
Upper secondary*1932-1945			-0.605				-0.459	
Upper sec vocl*1932-1945			0.003				0.962**	
Tertiary*1932-1945			-0.260*				0.058	
Lower secondary*<1932			-0.085				0.066	
Upper secondary*<1932			-0.416				-0.116	
Upper sec voc*<1932			-0.622*				0.737	
Tertiary*<1932			-0.302				0.488	
Constant	-4.312**		-4.922**		-2.775**		-2.329**	
Variance lev. 2 (person)	-1.259**		-1.790**		-0.937**		-0.738**	
St. dev. random intercept	0.533		0.409		0.626		0.692	
Rho	0.079		0.048		0.106		0.127	
N	835288		1132039		598800		232555	

*p<0.05; **p<0.01

Career investments, human capital and scarring effects

The results largely confirm our earlier findings (Models 1, 2 and 3) for which reason we only briefly discuss them. We find strong support for the existence of duration dependence showing that prior career investments reduce exit for male workers stronger than for female workers. Similarly, a longer stay in unemployment inhibits women's careers much stronger than men's, confirming that women endure stronger scarring effects.[76] Prior career investments seem therefore to protect insiders but in particular men and much less so women, confirming hypotheses 2a and 2b about the impact of prior career investments and 3a about the insider-oriented German labor market. Contrary to our expectations, higher educated men and women do not always exit less and re-enter more. Higher educated women compared to lower educated women participate more in the labor market, exit more and re-enter more, showing that they are more mobile, possibly due to the larger volatility among part-time workers. The interaction between education and cohort shows that higher educated men belonging to the younger cohorts saw their relative position impaired. A high education appears less of a guarantee to sustain employment or to have better options to re-enter. Men and women with upper secondary general education show the highest exit and the lowest re-entry chances, which is confirmed by the effects of the two-way interactions with cohort at least for men.[77] But for men at intermediate education level, enrolment in vocational training seems to prevent employment exits to some extent. and to increase the re-entry chances compared to general education, especially for the older cohorts. This might signal the rising difficulties of outsiders belonging to the youngest birth cohorts to enter the ranks of the core worker. In addition, the attainment of tertiary education appears to become less a guarantee for the career of the younger cohorts than it used to be for the oldest generations contrary to hypothesis 2a and 3a about the improved chances of the better educated younger cohorts.

Employment history and familial responsibilities

The employment history variables also acting as indicators for prior career investments exhibit strong effects on mobility, for both men and women. A higher occupational status protects workers from exit and facilitates re-entry confirming hypothesis 3a about the insider-protecting German labor market. Working in a part-

[76] Gender differences turn out to be significant in a formal test.

[77] The only exception is for cohort 1932-45, where we find highest exits for upper secondary vocational education.

time job not only increases exit but also facilitates re-entry, leading therefore to higher overall volatility. Employees in the shielded public sector show a lower risk of exit. Career complexity, as indicated by the number of previous job spells, clearly raises the odds of both exit and re-entry, and hence volatility, suggesting that an unstable work history tends to act as a barrier to enter stable employment.

Familial responsibilities show significant effects: single women exhibit much lower odds of exit and much higher odds of re-entry compared to married women, while for men the opposite holds. The presence of more children reduces women's exit, confirming that also women's behavior is affected by the financial needs incurred by large families. Due to increasing familial duties and consequently more difficulties in combining work and care, women with young children (pre-school age) show higher odds of exit. The presence of young children also reduces the re-entry chances compared to households without young children, although the inactivity might be voluntary. On the other hand, women with older children face lower odds of exit, probably due to the selection of mothers into employment at this stage of their lives, although re-entry is lower than for women with younger children. The results for males are opposite, confirming hypothesis 1c; Men exit less and re-enter more possibly due to the financial needs implied with having children.

Age and cohort effects

The relationship with age is decreasing for both exit and re-entry signaling lower mobility among older people. However, after some age threshold exit increases, likely to be due to early retirement, and re-entry further decreases both for men and women, reflecting the reduced labor market opportunities at higher age. Due to the unbalanced composition of the sample in the two surveys, we allow cohort effects to vary by survey. The way the life-history survey is designed makes it better suited to capture cohort effects, while in the panel study, which asks about the current situation, respondents belonging to different cohorts are observed in different life stages (younger cohorts at younger ages and older cohorts at older ages), which makes it difficult to disentangle cohort from life-stage effects. Exit as well as re-entry chances appear lower for the older cohorts of men and women, showing increased volatility among the younger cohorts confirming hypothesis 1a and 3b.[78] Only for men, we allow different cohort effects by education, since the effects for women were insignificant: The picture varies within cohort across educational levels, but shows no

[78] These results are not shown but are available from the authors upon request.

decreasing trend for exit nor an increasing trend for re-entry for the higher educated within all the cohorts providing little evidence for the occurrence of skill-biased technological change worsening the position of the low educated as formulated in hypothesis 2a (see also Luijkx et al., 2006).

7. Conclusions

In this paper, we investigate the labor mobility patterns of several cohorts of West German men and women covering the period from the late 1930s to 2005 by combining retrospective life history and prospective panel data. We use the GLHS life-history and the GSOEP panel data which allow us to analyze a longer time period, to broaden the cohort and age composition and to overcome particular shortcomings associated with each single dataset.

In the first part of the study we carefully look into the shortcomings in each design and whether they can be resolved. Most concerns deal with memory bias, since in the literature serious questions are raised about the quality and reliability of retrospective data, although there is little empirical evidence to date to support these concerns. We therefore try to dig further into this issue by examining the differences in findings across the two designs after correcting for memory bias.

Firstly, we explain the differences in duration dependency effects found in the two designs referring to earlier work showing that correcting for the bias in the recall of shorter or less salient spells indeed reduces the differences across the two designs, although it does not remove them entirely and therefore only partly explains the weaker state dependence in the retrospective study (Manzoni et al., forthcoming). This study indeed shows robust findings with respect to the existence of strong negative duration dependence effects in both designs, but with weaker strength in the life-history study.

Secondly, we investigate the impact of seam effects in the panel and the length of the recall period in the life history study by including a dummy for the heaping of events in December in the panel and spline functions for the time distance to the interview in the GLHS. We show that correcting for these sources of recall bias is needed to obtain more robust estimates. We finally investigate the role of implicit anchors such as marriage or childbirth events for explaining the differences between the surveys, but find no clear evidence for that (hypothesis 4). The effects of the implicit anchors are shown to be mostly insignificant providing little evidence for the impact of

anchoring. This is quite reassuring, in the sense that it provides little evidence that labor market mobility patterns over time are strongly affected by the use of retrospective data of allegedly lower quality. On the contrary, the paper shows that combining data from surveys with different designs and correcting for design differences partly resolve the shortcomings of each design. The sample composition is broadened by combining the two studies and the substantive results are therefore more robust with respect to the observed changes in mobility patterns by birth cohort, education level and prior career investments than with each separate design.

In the second part of the paper we focus on the substantive results combining the two data sets. Firstly, we find that due to familial responsibilities, women face higher exit and reduced re-entry chances than men, supporting the idea that particularly women face a more 'patchwork career'. Men instead react differently and increase their participation on the labor market due to the financial needs implied with having a family (hypothesis 1c). The findings also clearly show that prior career investments through work experience or education pay off for the career (hypothesis 2a), which is likely more in favor of the insiders on the internal labor market than the outsiders such as the low educated (hypothesis 3a). Our hypothesis on the occurrence of 'scarring' effects in the German labor market is also confirmed: the longer non-employed people stay out of the labor market the lower their chances to re-enter in a stable job later in the career (hypothesis 2b).

Hence, the main results confirm and mirror the main features of the German labor market as an insider-oriented, institutionalized and highly stratified occupational system coupled to relatively low levels of mobility.

The study also confirms the alleged rising volatility on the labor market: younger birth cohorts indeed display more mobility than older cohorts (hypothesis 1a). These changing mobility patterns seem to affect all education levels and not the low educated or low skilled in particular (hypothesis 2a, 3a), as the notion of the skill-biased technological change suggests. Tertiary education serves less as a guarantee for a successful career of the younger cohorts than it used to be for the older generations. The results therefore support on the one hand the larger career volatility among younger birth cohorts but provide on the other hand little support for the skill biased technical change thesis since a high education appears less of a guarantee for a successful career than it used to be for the older cohorts.

References

Acemoglu, D. (2002). Technical Change, Inequality, and the Labor Market. *Journal of Economic Literature* 40, 7-7

Allmendinger, J. (1989). Career Mobility Dynamics. A Comparative Analysis of the United States, Norway, and West Germany. Studien und Berichte 49. Max-Planck-Institut für Bildungsforschung, Berlin

Arthur, M. B., Rousseau, D. M., (1996). The boundaryless career: a new employment principle for a new organizational era. Oxford University Press, New York

Barbieri, P. (2009). Flexible Employment and Inequality in Europe. *European Sociological Review*. Advance access published May 4, 2009, doi:10.1093/esr/jcp020.

Bergemann, A., Mertens, A., (2004). Job stability trends, layoffs and transitions to unemployment: an empirical analysis for West Germany. Centre for Economic Policy Research (CEPR), London

Bertola, G., Blau, F. D., Kahn, L. M., (2001). Comparative Analysis of Labor Market Outcomes: Lessons for the US from International Long-Run Evidence. *NBER Working Papers* 8526

Biemann, T., Fasang, A., Grunow, D. (2009). Did Economic Globalization Destabilize Careers? Evidence from Germany. *CIQLE Working Paper* 2009-02

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

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, 208-25

Blossfeld, H. P., Rohwer, G. (1995). Techniques of event history modeling: new approaches to causal analysis. Erlbaum, Mahwah

Blossfeld, H. P., Mills, M., Bernardi, F. (Eds) (2007). Globalization and uncertainty of men's careers. An international comparison. Edward Elgar, Cheltenham

Bradburn, N. M., Rips, L. J., Shevell, S. (1987). Answering autobiographical questions: the impact of memory and inference on surveys. *Science* 236, 157-61

Buchholz, S., (2004). Men's Late Careers and Career Exits in West Germany. *Globalife Working Paper* No. 63.

Buchholz , S., Grunow, D. (2003). Globalisation and Women's employment in West Germany. *Globalife Working Paper* No. 47.

Buchholz, S., Hofacker, D., Mills, M., Blossfeld, H. P., Kurz, K., Hofmeister, H. (2009). Life Courses in the Globalization Process: The Development of Social Inequalities in Modern Societies. *European Sociological Review* 25, 53-71

Callegaro, M. (2007). Seam Effects in Panel Data Collected With Event History Calendar Methodology: The Case of Labor Force Transitions in the PSID. Survey Research and Methodology program (SRAM) Dissertations

Carroll, G. R., Mayer, K. U. (1986). Job-Shift Patterns in the Federal Republic of Germany: The Effects of Social Class, Industrial Sector, and Organizational Size. *American Sociological Review* 51, 323-41

Constant, A., Zimmermann, K. F., (2004). Occupational Choice Across Generations. *DIW discussion paper* 395.

De Graaf, N. D., Wegener, B., Liebig, S. (1989). Reliability of retrospective survey data: a comparison of job career and social resource variables. Unpublished paper

Dex, S., Laurie, H. (1995). The Reliability of Recall Data: A Literature Review. *Bulletin de Methodologie Sociologique* 49, 58-89

DiPrete, T. (2002). Life Course Risks, Mobility Regimes, and Mobility Consequences: A Comparison of Sweden, Germany, and the United States. *American Journal of Sociology* 108 (2), 267-309

DiPrete, T. (2005). Labor Markets, Inequality, and Change: A European Perspective. *Work and Occupations* 32, 119-39

Drobnič , 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, 133-46

Fouarge, D., Manzoni, A., Muffels, R., Luijkx, R. (2009). Childbirth and cohort effects on mother's labour supply: A comparative study using life-history data for Germany, the Netherlands and Great-Britain, *Work, Employment and Society*, forthcoming

Frick, J. (2005). A General Introduction to the German Socio-Economic Panel Study (SOEP). Design, Contents and Data Structure [waves A-U, 1984-2004]

Gangl, M. (2004). Welfare States and the Scar Effects of Unemployment: A Comparative Analysis of the United States and West Germany. *American Journal of Sociology* 109, 1319-64

Gangl, M. (2006). Scar Effects of Unemployment: An Assessment of Institutional Complementarities. *American Sociological Review* 71, 986-1013

Gazier, B. (2000). The connection between local and global justice: The case of transitional labor markets. *Revue Economique* 51, 571-81

Grimshaw, D., Ward, K.G., Rubery, J., Beynon, H. (2001. Organisations and the Transformation of the Internal Labor Market. *Work Employment Society* 15, 25-54

Grunow, D., Mayer, K. U. (2007). How stable are working lives? Occupational Stability and Mobility in West Germany 1940s-2005. *CIQLE Working Paper* 2007-03

Hauser, R. M., Featherman, D.L. (1977). The process of stratification: trends and analyses. New York Academic Press

Jürges, H. (2005). Retrospective error in SOEP calendar data: the case of unemployment. *Schmollers Jahrbuch* 125, 157-65.

Kraus, F., Steiner, V. (1998). Modelling Heaping Effects in Unemployment Duration Models -With an Application to Retrospective Event Data in the German Socio-Economic Panel. Jahrbücher für Nationalökonomie und Statistik. Stuttgart

Kurz, K., Hillmert, S., Grunow, 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

Lindbeck, A., Snower, D. J. (2001. Insiders versus Outsiders. *The Journal of Economic Perspectives* 15, 165-88

Luijkx, R., Kalmijn, M., Muffels, R. (2006). The impact of globalisation on job and career mobility of Dutch men: life-history data from the mid 1950s to the year 2000. In: Blossfeld, H. P., Mills, M., Bernardi, F., (Ed.) Globalisation, Uncertainty and Men's Careers. Cheltenham, UK/Northampton, MA, USA, Edward Elgar: 117-144

Luijkx, R., Wolbers, M. (2009). The Effects of Non-Employment in Early Work-Life on Subsequent Employment Chances of Individuals in The Netherlands. *European Sociological Review*. Advance access published February 3, 2009, doi:10.1093/esr/jcp002

Manzoni, A., Luijkx, R., Muffels, R., Explaining Differences in Labour Market Transitions Between Panel and Life-course Data in West Germany, *Quality and Quantity*, forthcoming.

Mathiowetz, N. A., Duncan, G. J. (1988). Out of work, out of mind: Response errors in retrospective reports of Unemployment. *Journal of Business and Economic Statistics* 6, 221-9

Maurin, E., Postel-Vinay, F. (2005). The European Job Security Gap. *Work and Occupations* 32, 229-52

Mayer, K. U. (2007). Retrospective Longitudinal Research: The German Life History Study. In: Menard, S. (Ed.) Handbook of Longitudinal Research: Design, Measurement and Analysis. Elsevier, San Diego

Mayer, K. U., Carroll, G. R. (1987). Jobs and Classes: Structural Constraints on Career Mobility. *European Sociological Review* 3, 14-38

Mertens, A. (1999). Job Stability Trends and Labor Market (Re-)Entry in West Germany 1984-1997. *Sonderforschungsbereich 373* 1999-60.

Mills, M. (2009). Globalization and Inequality. *European Sociological Review* 25, 1-8

Müller, W., Lüttinger, P., Wolfgang and Wolfgang, K. (1989. Class and Education in Industrial Nations. *International Journal of Sociology* 19, 3-39.

Paull, G. (2002. Biases in the reporting of labor market dynamics. *IFS Working Paper* W02/10.

Peters, E. (1988). Retrospective Versus Panel Data in Analyzing Lifecycle Events. *Journal of Human Resources* 23, 488-513

Powers, E. A., Goudy, W. J., Keith, P. M. (1978). Congruence Between Panel and Recall Data in Longitudinal Research. *Public Opinion Quarterly* 42, 380-9

Reimer, M. (2005. Autobiografisches Gedächtnis und retrospektive Datenerhebung: die Rekonstruktion und Validität von Lebensverläufen. Studien und Berichte 70. Max-Planck-Institut für Bildungsforschung, Berlin

Reimer, M., Künster, 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. Arbeitsberichte 8-2004. Max-Planck-Institut für Bildungsforschung. Berlin

Rosenfeld, R. A. (1992). Job Mobility and Career Processes. *Annual Review of Sociology* 18, 39-61

Rubin, D. C., Baddeley, A.D. (1989). Telescoping Is Not Time Compression - A Model Of The Dating Of Autobiographical Events. *Memory and Cognition* 17, 653-61

Scott, J., Alwin, D. (1998). Retrospective Versus Prospective Measurement of Life Histories in Longitudinal Research. In Giele, J. (Ed.), Methods of Life Course Research: Qualitative and Quantitative Approaches., Thousand Oaks et al. Sage Publications

Shavit, Y., Müller, W. (1998). From school to work: a comparative study of educational qualifications and occupational destinations. Clarendon Press Oxford

Singer, J. D., Willett, J. B. (2003). Applied longitudinal data analysis: modeling change and event occurrence. Oxford University Press, New York

Solga, H. (2001). Longitudinal Surveys and the Study of Occupational Mobility: Panel and Retrospective Design in Comparison. *Quality and Quantity* 35, 291 – 309

Statistiches Bundesamt (1999). Die Erwerbsbeteiligung im Familienzusammenhang und ausgewaehlte Einflussfaktoren - Ergebnisse de Mikrozensus 1997. *Wirtschaft und Statistik* 1, 28-38

Stier, H., Yaish, M. (2008). The Determinants of Women's Employment Dynamics: The Case of Israeli Women. *European Sociological Review* 24(3), 363-377

Stone, K. V. W. (2006). Thinking and Doing--the regulation of workers' human capital in the United States. *Socioeconomic Review* 4, 121-38

Sudman, S., Bradburn, N. (1974). Response Effects in Surveys. Aldine Publishing Company, Chicago

Tienda, M., Stier, H. (1996). Generating Labor Market Inequality: Employment Opportunities and the Accumulation of Disadvantage. *Social Problems* 43, 147-65

Trappe, H., Rosenfeld, R. A. (1998). A Comparison of Job-Shifting Patterns in the Former East Germany and the Former West Germany. *European Sociological Review* 14, 343-68

Uunk, W., Mach, B. W., Mayer, K. U. (2005). Job Mobility in the Former East and West Germany: The Effects of State-Socialism and Labor Market Composition. *European Sociological Review* 21, 393-408

Winkelmann, R., Zimmermann, K. F. (1998). Is job stability declining in Germany? Evidence from count data models. *Applied Economics* 30, 1413-1420

Chapter 5

Childbirth and cohort effects on mothers' labor supply: A comparative study using life-history data for Germany, the Netherlands and Great-Britain *

Abstract. The negative effect of childbirth on mothers' labor supply is well documented though most studies examine the short-term effects only. We use retrospective life history data for Germany, the Netherlands and Great Britain to investigate the long-term effects of childbirth on mothers' labor supply for successive birth cohorts. We estimate probit regression models correcting for selection (Tobit) into motherhood and the number of births. We find strong drops in participation before first childbirths, and strong recovery after the birth of the last child especially in Great Britain. Younger cohorts display a less sharp decline in participation around childbirth and a faster increase in participation in the 20 years after childbirth, especially in the Netherlands. However, mothers' participation rate does not return to its pre-birth level in any of the countries. Labor market conditions and institutional public support seem to contribute to explain the cross-country variation in participation behavior after childbirth.

**A very similar version of this chapter in under revision (revise and resubmit) in Work, Employment and Society. The article is cohautored with Didier Fouarge, Ruud Luijkx and Ruud Muffels.*

1. Introduction

Studies in the realm of economics and sociology have shown that childbirth has a negative effect on women's labor market participation. The occurrence of a childbirth event (Drobnič , 2000), the number of subsequently born children (birth sequence) and the timing of these births (Miller and Xiao, 1999), exert a substantial effect on mothers' labor supply over the life-course. However, the evidence to date is mostly based on a comparison of (repeated) cross-sectional data. Exceptions are Kenjoh (2005) and Vlasblom and Schippers (2006) in this Journal who used panel data, but with relatively short time horizons. An alternative is to use life history data covering the entire life-course of people (Manzoni et al., forthcoming).

Our contribution to the literature is threefold. First, we use unique retrospective life-history data to provide comparative evidence on the effect of childbirth on the labor supply of mothers from subsequent birth cohorts in three different labor markets and gender regimes (Lewis, 1992), i.e. Germany, the Netherlands and Great Britain. Second, our empirical model of mothers' labor supply examines the entire fertility history of women by incorporating the birth sequence and timing of subsequent childbirths. Third, we model the anticipation and recovery effects of subsequent childbirths on mothers' labor supply from two years before up to 20 years after childbirth. The main idea is that changes in women's labor market behavior across birth cohorts and over the life-course reflect the increasing labor market dynamics and cohort-related changing females' working time preferences due to individualization or emancipation trends, changes in educational attainment, shifts in occupational structure, and changing policies to support women's participation. The main research questions are as follows:

1. To what extent does the birth of a first child affect women's labor supply, and what is the effect of subsequent childbirths?
2. To what extent does the birth of a child affect the labor market participation of subsequent birth cohorts in a different way?
3. To what extent are the findings different across countries and gender-regimes, and which factors might be responsible for the observed differences?

We discuss the theoretical and empirical background of the study (Section 2), present the data and explain the research design (Section 3), report on the descriptive findings (Section 4), and discuss the results from the model estimations (Section 5). In Section 6, we draw conclusions and discuss further implications for research and policy.

2. Theoretical and empirical background

Basically, our approach departs from an 'agency-structure' perspective according to which labor market decisions are believed to be affected by 'rational choice' (Hakim, 2002) and structural constraints (Ginn et al., 1996). In our view it is not only agency or choice as Hakim's preference theory suggests but also contextual and structural constraints that matter for explaining women's decisions to participate in the labor market (Pfau-Effinger, 1993). There has been ample debate and supporting evidence that in particular contexts, family related time, money constraints and the lack of labor market opportunities determine women's employment and occupational choices (Healy, 1999; Crompton and Le Feuvre, 2000; Gash, 2008). Secondly, we disagree with Hakim's idea that women have advanced and stable career preferences since labor market preferences can change over the life course, and might be affected by the institutional and labor market context. For that reason we compare females' choices across three different institutional contexts while in our empirical model we explicitly allow female labor market choices to change at any point in their career. Changes in the labor supply of mothers across cohorts could therefore result from a number of changes: shifts in preferences or norm behavior, institutional changes associated with equal opportunities and child care policies, changes in educational attainment, or shifts in occupational and employment structures and mobility patterns (Crompton and Le Feuvre, 2000).[79] Our main interest, however, is with changes in participation patterns around childbirth by cohort.

The interdependence of labor supply within the family has received a great deal of attention in the literature (Becker, 1981, Drobnič et al., 1999). Becker (1981) expects that

[79] Vlasblom and Schippers (2004) suggest that changes in behavior are the most likely driving force behind the increase in female employment in Europe.

partners specialize in either labor market participation or home production, depending on their respective competitive advantage. Generally speaking, men are assumed to specialize in labor market activities, and women in home production, especially after childbirth. However, Becker's theory ignores the heterogeneity implied in male and female working time patterns (Dex et al., 1998). Evidence shows that single parents behave differently from married or cohabiting couples (Kalmijn and Luijkx, 2006).
Childbirth exerts a negative effect on women's labor force participation (Dekker et al., 2000; Drobnič , 2000; Gornick, 1994; Rosenfeld and Birkelund, 1995; Stier et al., 2001; Van der Lippe, 2001). However, women's participation behavior is significantly dissimilar across the three regimes. They share historically the 'male breadwinner' type of welfare regime that evolved over time into a 'dual earner' gender regime though with rather different gender labor supply patterns (Lewis, 1992; Lewis et al., 2008). The labor market context, and the education or occupational structure might also affect these supply patterns. The 'child effect' is, therefore, not equally strong across European welfare states (Gornick, 1994). Furthermore, younger generations of women are more likely to continue work after childbirth in longer-hours jobs than older generations.

The main question is to what extent the birth of a first child and subsequent children lead to a dissimilar change in women's labor supply in the three employment regimes. The British regime is characterized by a lower level of family support as indicated by the availability of childcare services (De Hanau et al., 2006), and by leaving work/family reconciliation predominantly a private responsibility (Lewis et al., 2008). Public support for mothers' employment is much stronger in Germany and the Netherlands where parental leave schemes are more generous. Though women's participation rates are not very different (62% in Germany, 66% in the UK and 68% in the Netherlands in 2007), they are very dissimilar for mothers with children. Vlasblom and Schippers (2006, p. 335) show that in the 1980s and 1990s the participation of women with one and two children at first childbirth is respectively 66% and 54% in the Netherlands, 31% for both in the UK and 11% and 12% in Germany.

The availability of part-time jobs is essential for the participation of women and Delsen et al. (2007) showed that demand-side factors such as the growth of the service sector, the need to extend operating hours and the greater flexibility it offers to employers can

explain the rise of part-time work. The share of part-timers varies from 42% of British, 46% of German to 75% of Dutch women in 2007, though women working much shorter hours in Germany and the Netherlands than in Great-Britain. The Netherlands is therefore characterized as the 'one-and-a-half earner' gender regime. The birth of a child is therefore likely to exert a larger adverse effect on the labor participation of mothers in Great Britain than in the other countries. Due to the greater availability of good quality part-time jobs in the Dutch labor market, we expect the birth of a child to have a smaller negative effect on the labor market participation of Dutch women (Fouarge and Muffels, 2009). Due to the lower public support in the unregulated British labor market, we expect British mothers to re-enter more quickly after birth to earn a living than German or Dutch working mothers with more extended and generous institutional support.

The labor market participation of women has increased across subsequent birth cohorts (Goldin, 2006). Institutions have been adapted through the supply of extended parental leave and childcare services, and to better accommodate to the needs to combine care and work (Uunk et al., 2005; Vlasblom and Schippers, 2006; Lewis et al., 2008). The occupational structure has also changed in favor of female employment due to a shift to service employment. However, these changes in institutions and the labor market context are relatively recent, and they affect the younger cohorts of women relatively more than older cohorts. Therefore, we expect the birth of a child to have a smaller effect on the labor supply of younger compared to older cohorts, especially in the Netherlands. Dutch women belonging to the younger cohorts have stronger preferences for combining work and care, have more opportunities to work part-time, and are higher educated, which improves their career opportunities compared to older cohorts. Furthermore, we expect the pace of recovery after childbirth to be steeper for younger cohorts. Eventually, we expect the drop in participation before childbirth to be more pronounced for first than for subsequent childbirths as was shown by Vlasblom and Schippers (2006). They argue that women anticipate the arrival of a second child by reducing their participation already before first childbirth indicating their different work orientation. Likewise, the pace of recovery after first childbirth is supposedly more pronounced for women with one child than for women with more children.

3. Data and methods

3.1. Life-history data

We use life-history data for Germany, the Netherlands and Great Britain, in which individuals are asked retrospectively about their school, work and demographic careers.[80] For Germany, data come from the German Life History Survey (GLHS) (Mayer, 2007), a study of eight cohorts born between 1919 and 1971 which allows the construction of a complete retrospective career and marital history.

The data for the Netherlands are from four retrospective life history surveys based on (stratified) random samples of the Dutch population which gathered retrospective monthly information on work histories, changes in the family structure and marital situation.[81]

For Great Britain, we use the British Household Panel Survey (BHPS). The labor market, marital and demographic histories are constructed from the retrospective information gathered in the second wave (1992), supplemented with information from the panel waves (Maré, 2006).

We selected women and reconstructed their employment spells as well as demographic events on a monthly basis starting from the first labor market entry – after leaving full-time education – up to the current situation at the time of interview. Due to differences in the legal retirement age, our analyses include Dutch and German women of 18-65 and British women of 18-60 years.

3.2 Empirical model

For assessing the impact of first and following childbirths on women's employment, we estimate a probit regression model explaining the probability to be employed in each month preceding and following childbirth. We perform event history analysis of the labor market participation of various cohorts of women dependent on the sequence of

[80] Gutierrez-Domenech (2005) uses similar data but with shorter time horizons. The literature provides little evidence for concerns about the reliability of retrospective data due to recall error being different across cohorts.

[81] The four surveys are: the Netherlands Family Survey 1992-1993 – FNB1992; the Survey Households in the Netherlands 1995 – HIN1995; the Family Survey Dutch Population 1998 – FNB1998 and the Family Survey Dutch Population 2000 – FNB2000 (see e.g. De Graaf *et al.*, 2000).

childbirths, while controlling for other characteristics. To capture possible anticipation and recoup effects of childbirth on mothers' labor supply, we include two time variables measuring the time before (t^b) and after (t^a) birth. However, the relationship with time is non-linear, as it is clearly shown in Figure 1, showing mothers' participation rates from 24 months before up to 240 months after childbirth. The effect of time before and after birth on labor supply apparently follows a 'saddle-shaped' pattern, meaning that it decreases prior to birth (anticipation) whereas it increases after birth (recovery) in a curve-linear way. This pattern is captured including a linear and quadratic time trend.

We estimate a probit model instead of a panel regression model. The main reason is that, due to the interdependency of mothers' subsequent childbirth decisions, we had to re-design our data into a mother-birth dataset for each of the childbirth sequences.[82] The probit model accounts for the differential career effect of first childbirth as well as of each following child by birth cohort of the mother. We model up to three birth events and control for additional births.

Given the unobserved index variable y_i^*, we observe the labor market outcome $y_i = 1$ if $y_i^* > 0$, and $y_i = 0$ if $y_i^* \le 0$, with $\Pr(y_{it} = 1) = \Phi(\beta X_{it})$, where y indicates the labor supply of female i in month t (with $y_{it} = 0$ when out of employment), $\Phi(.)$ is the standard normal cumulative distribution, X_{it} are the time constant and time varying covariates, and β the set of associated coefficients. The model can be written as follows:

$$\Pr(y_{it} = 1) = \alpha + \beta_{1b}\left(t_{it}^b * C_i * S_{it}\right) + \beta_{2b}\left(\left(t_{it}^b\right)^2 * C_i * S_{it}\right)$$
$$+\beta_{1a}\left(t_{it}^a * C_i * S_{it}\right) + \beta_{2a}\left(\left(t_{it}^a\right)^2 * C_i * S_{it}\right) + \gamma V_{it} + \varepsilon_{it} \qquad [5.1]$$

where $t^b{}_{it}$ and $t^a{}_{it}$ are time trends for the 24 months before and 240 months after childbirth, which are included linearly and quadratically, C_i represent the dummies for each birth cohort j and S_{it} the birth sequence of children. α is a constant term, β_{1b}, β_{2b}, β_{1a},

[82] Another reason is that since we observe women for long periods of time, it seems untenable to assume as in panel regression that unobserved individual effects are fixed or stable over time, particularly when a major life event such as childbirth is being considered. For estimation of the probit model on the pooled data we corrected the standard error of the parameters to account for the clustered nature of the data.

β_{2a} and γ are vectors of coefficients to be estimated, and ε_{it} is a normally distributed error term.[83]

The birth sequence variable (S_{it}) distinguishes among the 1st child in a one-child family, the 1st child in families with more children, the 2nd child in a family with two children, and the 2nd and 3rd child in families with more children. The birth cohort dummies (C_i) allow us to capture the major shifts in female labor supply that took place in the last century. We distinguish the 1900-1930, 1931-1940, 1941-1950, 1951-1960, and 1961-1970 cohorts in the Netherlands and Great Britain, and the 1919-21, 1929-31, 1939-41, 1949-51, 1954-56, 1959-61, 1964 and 1971 cohorts in Germany. The model allows for different estimates of the child effect depending on the birth sequence of subsequent children, and the birth cohort to which the mother belongs. The vector of parameters β_{1b}, β_{2b}, β_{1a} and β_{2a} captures the anticipation and recovery effects of subsequent children born in the various cohorts.[84]

The matrix of covariates (V_{it}) includes controls for educational level,[85] age at first childbirth, marital status, and a dummy for four or more children. Since we have no information on career orientations or preferences in our data we follow Vlasblom and Schippers (2006) assuming that mother's age at first childbirth can be viewed as a proxy. We suspect that the lower the age at first childbirth, the more likely mothers to be family-oriented or 'home-centred', and to reduce labor supply around childbirth. Likewise, the higher the age is, the more 'work-centred' women are and the less likely they are to interrupt their career. Education level is included to capture the rising level of education across cohorts, and because we assume that higher educated women are more career-oriented and therefore have reduced withdrawal rates around childbirth. We further control for differences in marital status since previous studies looking at labor supply effects of marriage for women found strong negative effects (Kalmijn and Luijkx, 2006; Stier et al., 2001). More traditionally oriented women are more likely to

[83] The equation represents Model III.

[84] In the three-way interactions, cohort is entered as a continuous variable to limit the number of coefficients to be estimated.

[85] Education is measured according to the CASMIN classification scheme in Germany (Müller *et al.* 1989) and the ISCED classification in the Netherlands and Great Britain.

marry and to interrupt the career (Drobnič et al., 1999), whereas career oriented mothers are more likely to stay single or cohabiting (Vlasblom and Schippers, 2006).[86]

Self-selection into motherhood

Depending on the birth cohort, between 10 and 20% of the sample never become mothers. Mothers could differ systematically from non-mothers in characteristics unobserved by the scientist, such as career intentions or occupational choices (Del Boca and Sauer, 2006). Would such factors also influence employment decisions, ignoring the selection process underlying the childbirth decision would bias the estimates of the impact of certain variables on the employment decision.

In a first step, we estimate a Tobit selection model to account for the probability to become a mother and the number of children. The model takes the form: $y_i^* = Z_i\beta + u_i$, with $u_i \sim N(0, \sigma_u^2)$, where y_i^* is an index variable of which the outcome $y_i = 0$ if $y_i^* \leq 0$ and $y_i = y_i^*$ if $y_i^* > 0$, with $\Pr(y_i > 0) = \Phi(-Z_i\beta/\sigma_u)$, $\Phi(.)$ being the standard normal cumulative distribution. In the model, y_i equals zero when a woman never becomes a mother, Z_i are covariates and β represents the coefficients to be estimated. We control for marital status, educational level and birth cohort. Identification is ensured by the inclusion of interaction terms between birth cohort and the two other variables in the selection model, which are omitted in the outcome equation. From the selection model, we retrieve the Tobit residuals, which are included as an additional explanatory variable in the outcome model [1].[87] The residuals turn out to be significant in Germany and Great Britain, but not in the Netherlands.

We estimate three alternative specifications of the probit model:

- *Model I (baseline)* includes main effects for a) all the control variables, i.e. education, marital status, age at first birth and Tobit residuals b) birth cohort, c) the dummies for birth sequence and time before and after childbirth.

[86] Unfortunately, our data do not allow making a proper distinction between married and cohabiting women. However, this is not a major issue since this distinction is mostly relevant for younger cohorts.
[87] Due to space constraints, the results from the Tobit regressions are not reported here, but are available from the first author upon request.

- *Model II* adds to Model I the two-way interaction between the dummies for the time before and after childbirth, and the birth sequence, since we contend the anticipation and recoup effects to differ by birth sequence.
- *Model III* adds to Model II the three-way interactions between birth sequence, the time dummies and birth cohort, since we assume the fertility history and its effects on mothers' participation to be different across birth cohorts.

4. Descriptive findings

Table 1. Mothers' participation rates and background statistics by birth cohort, birth sequence and country

	First birth of one				First of more births				Second birth		
	% Married	Education level	% Employed	Age at first birth	% Married	Education level	% Employed	Age at first birth	Age at second birth	Nchild	% Employed
Germany											
1919-1921	79.3	1.5	36.6	25.9	82.9	1.3	39.3	24.6	28.6	2.8	30.7
1929-1931	90.3	1.3	30.6	26.6	93	1.2	29.6	24.9	28.7	2.9	22.3
1939-1941	88.3	1.4	48.1	25.4	95.1	1.4	31.6	23.7	27.2	2.6	18.4
1949-1951	94.2	1.6	41.7	25.1	92.8	1.3	34.1	22.7	26.1	2.3	14.6
1954-1956	88.4	1.8	28.1	27	93.2	1.8	34.4	24	27.2	2.3	18.4
1959-1961	87.2	1.8	26.2	25.5	91.4	1.8	22.1	22.5	25.5	2.3	11.7
1964	89.9	2.2	14.3	28.7	87.9	2	7.6	25.4	28.5	2.3	9.2
1971	88.2	2	8.9	24.9	88.3	1.6	2.8	22.1	24.8	2.2	2.8
Total	87.3	1.7	27.7	26.2	89.6	1.5	27	24.2	27.7	2.5	18.4
Netherlands											
1900-1930	100	1.7	0	29.1	90	2	15	27.1	29.4	3.4	8.2
1931-1940	92.5	1.9	17.5	29.1	94.4	2	19.1	25.8	28.4	2.9	12.7
1941-1950	91.5	2.5	37.2	27.4	96.1	2.4	28.7	25.1	27.7	2.4	15.2
1951-1960	95.5	2.8	60.2	30.1	96.4	2.7	51.6	26.4	29.1	2.5	30.6
1961-	96.5	2.8	75.1	27.9	97.7	2.7	60.1	25.8	28.2	2.3	41.1
Total	95.3	2.7	61	28.4	96.2	2.5	49.3	25.8	28.4	2.5	26
Great Britain											
1900-1930	94.8	1.6	24.9	29.5	94.8	1.6	17.1	25.8	29.6	2.9	13.5
1931-1940	91.4	2	24.3	27	97.8	2	14.2	24.4	27.3	3	16.1
1941-1950	85.7	2.4	27.6	27.8	92	2.5	21.8	24.3	27.1	2.6	20.8
1951-1960	86.7	2.8	49.2	28.5	91	2.8	28.7	24.7	27.3	2.6	23.3
1961-	67.9	2.9	48.4	24.8	72.2	2.6	29	22	24.3	2.4	20.3
Total	85.2	2.3	35.2	27.7	91.2	2.3	21.7	24.5	27.5	2.7	18.8

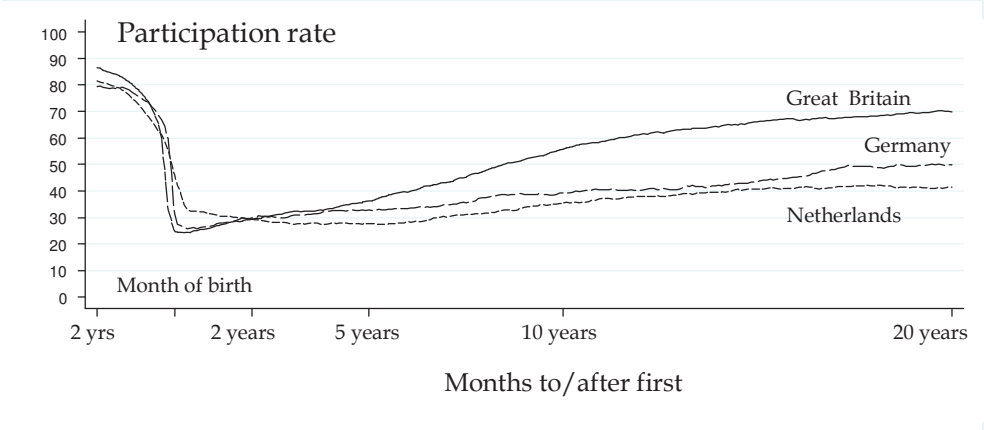
Source: Germany: GLHS, 1981-1999, Netherlands: FNB 1992, 1998, 2000; HIN 1995, Great-Britain: BHPS 1992; BHPS 1991-2003

In Table 1, we present descriptive findings on mothers' participation rates at childbirth and some background characteristics by birth cohort, birth sequence and country.

Broadly speaking, in Great Britain and the Netherlands, mothers' participation rates as well as their education level have increased across cohorts. The percentage of employed mothers at first childbirth is larger if there is only one child than if there are more children. For the younger cohorts, employment rates seem on average highest in the Netherlands and Great Britain. Dutch mothers postpone first childbirth longer and might therefore be more 'career oriented'.

Figure 1 shows the labor market participation rates before, at and after first childbirth. The pattern confirms the results of Vlasblom and Schippers (2006): mothers tend to reduce labor supply already before first childbirth.

Figure 1. Mothers' participation rates by time to and after first childbirth by country



Source: Germany: GLHS, 1981-1999, Netherlands: FNB 1992, 1998, 2000; HIN 1995, Great-Britain: BHPS 1992; BHPS 1991-2003

Although the general pattern is similar in all three countries, there are noticeable differences. As expected, pre-birth participation rates are highest and their decrease around childbirth largest in Great Britain though mothers seem to recover faster after birth. In Great Britain, mothers' participation rates drop sharply to about 25% at childbirth, but increase to more than 35% already after 5 years, 55% after 10 years and 70% after 20 years. In the Netherlands they decline to 34% and keep declining up to 5 years after first childbirth but then increase slowly to 40% after 20 years. In Germany,

the drop is nearly as dramatic as in Great Britain but the recovery is much slower while they increase gradually to 50% after 20 years. However, in none of the countries do the participation rates return to their initial level. Even in Great Britain, 20 years after childbirth it is still 15 percentage points lower than the pre-birth level. The model-based estimates in Section 5 show identical results.

5. Results from the model estimations

Table 2 presents the marginal effects from the probit model. They are calculated at the mean of the continuous variables (time, age), and as the difference with the reference category for the categorical variables (education, birth sequence, marital status, birth cohort). To ease the interpretation, the effects of the interaction terms of birth sequence, cohort and time before and after childbirth are not reported in the table but are shown graphically in Figures 2 and 3.[88]

Higher educated mothers tend to participate more in all three countries but especially in Great Britain and the Netherlands. Compared to married women, singles are more likely to participate in Germany and Great-Britain than in the Netherlands. The higher the age at first childbirth is, the higher the participation in the Netherlands, but the lower the participation in Great Britain.[89] These findings suggest that career-oriented women are less likely to interrupt their career in the Netherlands due to a stronger employment support for working mothers compared to Great Britain. The time variables confirm the 'saddle-shaped' form of the participation curve before and after childbirth.

Birth cohort effects confirm the descriptive information in Table 1. In the Netherlands and Great-Britain, participation rates increase steadily for younger birth cohorts (Model

[88] In Figure 3, the main effects of the cohorts pertain to upward or downward shifts of the curves across graphs. The main effects of the birth sequences pertain to upward or downward shifts of the curves within a graph. The time interactions are reflected in the curvature of the various curves within and across graphs.

[89] Including an interaction term between age at first birth and education level shows that all interaction effects are insignificant for the Netherlands. In the UK, the interaction effect for 'medium high' education is negative, but small and only significant at the 10%-level. In Germany, we find a small negative interaction effect for the 'medium low' educational level.

I). This suggests that the availability of part-time jobs tends to raise mothers' labor market participation. For Germany all cohorts participate less than the oldest cohort born around 1920, though the differences become smaller for subsequent cohorts. [90]

Table 2. Marginal effects of childbirth and birth cohort on mother's labor market participation, probit regression with Tobit selection (main effects only)

	Germany			Netherlands			Great Britain		
Models	I	II	III	I	II	III	I	II	III
Birth sequence									
First birth of one	0.119**	0.041	-0.001	0.056*	0.096**	0.089**	0.110**	0.163**	0.163**
Second birth of two	-0.002	-0.101**	-0.123**	-0.002	-0.025**	-0.027**	0.059**	0.005	0.005
Second of more	-0.043**	-0.047*	-0.062**	-0.019**	-0.038**	-0.036**	-0.015	-0.006	-0.005
Third of more	-0.019	-0.073**	-0.093**	0.003	-0.023*	-0.024*	0.060**	0.049**	0.050**
Time dummies									
Time to birth	-0.041**	-0.070**	-0.026**	-0.017**	-0.036**	-0.025**	-0.042**	-0.083**	-0.066**
Time after birth	0.002**	0.0003	-0.001**	0.001**	0.000**	0.002**	0.004**	0.003**	0.003**
Time before squared	-0.138**	-0.234**	-0.046*	-0.046**	-0.102**	-0.030*	-0.123**	-0.236**	-0.164**
Time after squared	-0.002*	0.003**	0.006**	-0.0001	0.002**	-0.005**	-0.008**	-0.003**	-0.001
Time after 4th child	-0.081*	-0.089*	-0.085*	-0.055**	-0.068**	-0.066**	-0.076**	-0.096**	-0.098**
Education level (ref.:low)									
Medium low	0.040	0.038	0.040	0.027	0.028	0.028	0.076**	0.075**	0.077**
Medium high	-0.048	-0.055	-0.054	0.071**	0.077**	0.076**	0.065*	0.064*	0.065*
High (University)	0.103*	0.101*	0.108**	0.239**	0.256**	0.255**	0.177**	0.179**	0.182**
Marital status (ref.:married/cohabited.)									
Single	0.337**	0.263**	0.261**	0.093**	0.043	0.042	0.158**	0.085**	0.088**
Divorced/separ.	0.179**	0.172**	0.173**	0.053	0.052	0.053	-0.014	-0.020	-0.021
Age at first birth	0.001	0.001	0.001	0.004**	0.004*	0.004*	-0.004**	-0.005**	-0.005**
Birth cohort (ref.:1900-30, De: 1919-1921)									
1931-40 (De:1929-31)	-0.113**	-0.121**	-0.147**	0.037	0.038	0.037	0.138**	0.143**	0.134**
1941-50 (De: 1939-41)	-0.116**	-0.122**	-0.168**	0.100**	0.102**	0.099**	0.149**	0.155**	0.137**
1951-60 (De: 1949-51)	-0.059*	-0.061*	-0.132**	0.219**	0.232**	0.240**	0.175**	0.183**	0.155**
1961-... (De: 1954-56)	-0.047	-0.046	-0.142**	0.313**	0.327**	0.355**	0.112**	0.119**	0.079*
(De: 1959-61)	-0.062*	-0.063	-0.177**						
(De: 1964)	-0.046*	-0.042	-0.173**						
(De: 1971)	-0.126**	-0.141**	-0.292**						
Interactions									
Two-way:	No	Yes	No	No	Yes	No	No	Yes	No
Three-way:	No	No	Yes	No	No	Yes	No	No	Yes
Selection into motherhood									
Tobit residuals	-0.021*	-0.019	-0.023*	-0.004	-0.002	-0.003	-0.041**	-0.037**	-0.039**
Pseudo-R-square	0.086	0.098	0.103	0.113	0.132	0.134	0.128	0.148	0.148
N-individuals		2,878			2,841			3,022	

Note: De=Germany
Source: Germany: GLHS, 1981-1999, Netherlands: FNB 1992, 1998, 2000; HIN 1995, Great-Britain: BHPS 1992; BHPS 1991-2003
*p<0.05; **p<0.01

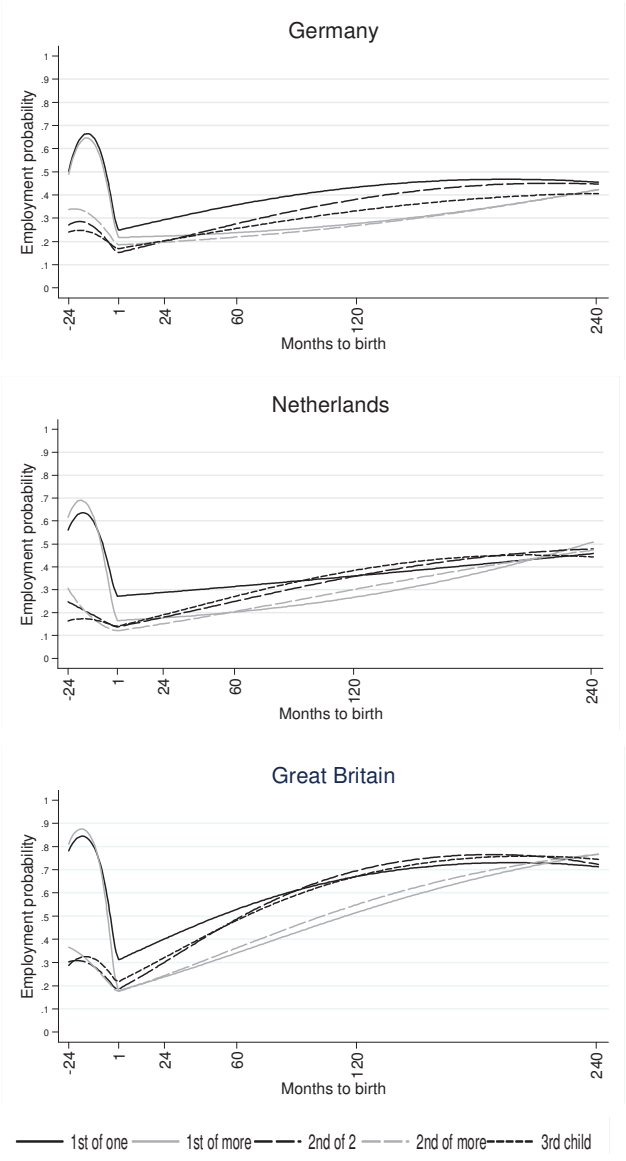
[90] The difference for the youngest cohort is larger, which could be due to the fact that women are on average younger at the time of the survey.

Figure 2 depicts the participation rates as calculated from the parameter estimates of Model II. For all countries, it shows smaller withdrawal rates at first childbirth when only one child is born compared to when more children are born. This suggests that career oriented women are less likely to interrupt their career than family oriented women. The drop in participation around first childbirth – both for mothers with only one or more children – is, as expected, especially large in Great Britain. This is likely explained by fewer working-time opportunities in firms allowing mothers to combine work and care or lack of public support to mothers in Great Britain.

Examining the anticipation effects, the curves appear more negatively sloped for first childbirths than for second or third childbirths in all three countries. Mothers with two or three children tend to participate less since they withdrew already at first childbirth, suggesting also that they are more family-oriented.

The recovery effects appear rather different across countries. In the Netherlands, mothers with only one child seem to recoup at a lower pace than in Germany or Great-Britain. This might reflect the stronger preferences for non-working time in the Netherlands or the insufficient childcare facilities. As expected, British mothers exhibit the strongest pace of recovery reflecting the stronger work norms in a double earner employment regime. In all countries, mothers seem to recover less quickly when the childbirth is not the last and more children follow (first and second births in families with more children). This might signal their stronger family-orientation compared to women with only one child. When the child is the last one born (1st child in 1 child family, 2nd child in 2 children family, or 3rd child), mothers seem to recover more quickly, which might be associated with reduced time constraints associated with the other children being older and needing less care.

Figure 2. Estimated mothers' labor supply (Model II): probit estimates of two-way interactions between birth sequence and time to/after childbirth by country

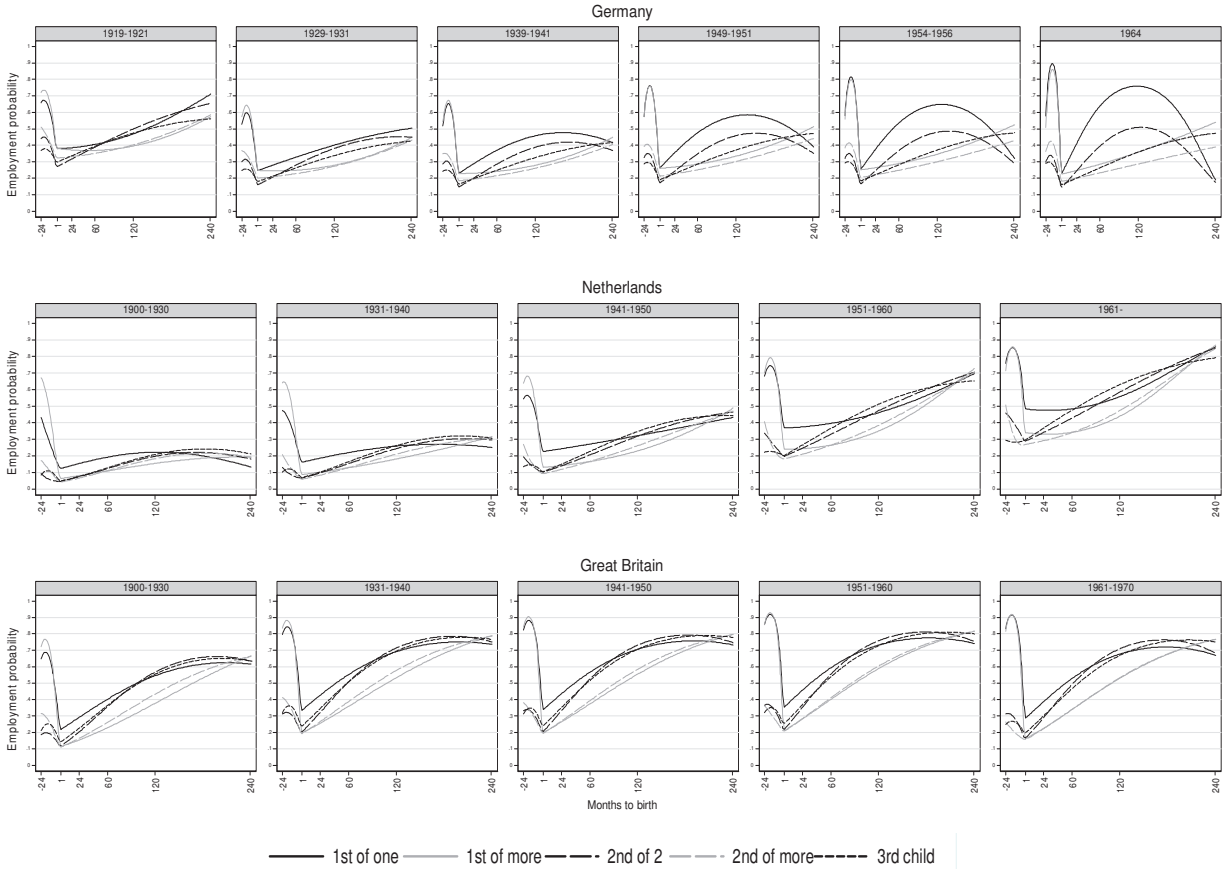


Source: Germany: GLHS, 1981-1999, Netherlands: FNB 1992, 1998, 2000; HIN 1995, Great-Britain: BHPS 1992; BHPS 1991-2003

Figure 3 depicts the participation rates as calculated from the parameter estimates of Model III. Younger cohorts in the Netherlands and Britain tend to participate more than older cohorts even after controlling for education level, marital status and age at first birth. Secondly, in all countries the steepness of the decline in labor participation in the two years preceding the first childbirth is significantly larger for the younger cohorts,

either in a one child family or in a family with more children. This means that subsequent cohorts tend to stay employed longer but when they stop working they do that closer to the time of birth. Thirdly, the pace of recovery is significantly steeper for subsequent cohorts, showing that mothers from younger cohorts tend to return to the labor market more rapidly than their predecessors. Fourthly, we find in all countries but especially in the Netherlands and Germany across cohort a larger variation in the pace of recovery after childbirth between the various birth sequences, signaling an increasing diversity of the life course.

Figure 3. Estimated mothers' labor supply (Model III): probit estimates of three-way interactions between birth sequence, time to/after birth and birth cohort by country



Source: Germany: GLHS, 1981-1999, Netherlands: FNB 1992, 1998, 2000; HIN 1995, Great-Britain: BHPS 1992; BHPS 1991-2003

For German and British mothers belonging to the younger cohorts, the curves for the first child in one-child families and the second child in two-child families are more strongly saddle-shaped than for otherwise similar Dutch mothers. This suggests that German and British mothers belonging to the younger cohorts increase their participation strongly in the first ten years after the birth of the last child, whereas Dutch women increase their participation at a slower but steadier pace. In Great Britain cohort differences in participation after childbirth are much smaller than in the other countries, suggesting that older cohorts already recovered more quickly after childbirth. The Netherlands display the most dramatic increase in participation rates across cohorts: younger cohorts appear to differ more in terms of their career patterns than the older ones, but also in Germany, where they participate less. More and more women belonging to the younger cohorts combine work and care after the third, which is often the last, child is born. This seems to signal the increasing financial needs associated with having children, the greater availability of working-time options and the shifting preferences of the younger generations of mothers to combine work and care.

6. Conclusions and discussion

The main contribution of the paper is to document the long-term effects of childbirth on the labor supply of various cohorts of mothers in Germany, the Netherlands and Great Britain. We use life-history data for three clearly distinct labor markets and gender regimes. We elaborate an empirical model in which we control for the selectivity into motherhood using a Tobit selection model in the first, and a probit model in the second step to explain mothers' labor supply. The model accounts for the birth sequence and the timing of births over the life-course. Our main findings are that 1) more mothers withdraw from the labor market around childbirth in Great Britain, but mothers also return faster into paid employment already in the 10 years following childbirth compared to the other countries; 2) successive cohorts display lower withdrawal rates before and at childbirth, and higher participation rates 20 years following childbirth, especially in the Netherlands; 3) younger birth cohorts show a larger diversity in participation patterns over the life course dependent on the birth sequence and timing of

births, particularly in Germany and the Netherlands. In all countries, when the child is not the last one, women recover less quickly. This is especially the case for Dutch women belonging to the younger birth cohorts. The results might signal the stronger family orientations of mothers with more children though younger cohorts seem to recoup more quickly than older cohorts.

The three countries share historically the same 'male breadwinner' type of gender regime and they all – in the last decade – seem to strive for enhancing mothers' labor supply, though with different levels of success. The results show increasing participation rates for younger cohorts in all countries also when women get more children, suggesting either that they exhibit stronger preferences for combining work and care or that they are exposed to more favorable institutions that help them combine work and care compared to older cohorts. The findings across country suggest that younger cohorts of women due to their improved education coupled to a larger availability of working-time options and stronger institutional support, especially in Germany and the Netherlands, have stronger preferences for combining work and care, and are better able to realize these preferences.

A policy lesson that might be drawn from the larger diversity in life-courses among the younger birth cohorts pertains to the existing income and employment support systems to become more life-course tuned or targeted to take account of the different career preferences of younger cohorts. The findings also suggest that regimes providing a wider portfolio of income and employment support options – such as the Netherlands – might help to create more and better part-time jobs because of shifting social norms in favor of part-time employment. Further evidence on a larger number of countries is needed to be able to disentangle the effects of choices and structural constraints indicated by changes in fertility behavior, career orientations, labor market conditions and policy and gender regimes.

References

Becker, G. S. (1981). *A Treatise on the Family*. Cambridge, MA: Harvard University Press

Crompton, R. and Le Feuvre, N. (2000) 'Gender, family and employment in comparative perspective: the realities and representations of equal opportunities in Britain and France'. *Journal of European Social Policy* 10: 334-348

De Graaf, N. D., De Graaf, P. M., Kraaykamp, G., and Ultee, W. C. (2000). Family Survey Dutch Population 2000, Nijmegen: Nijmegen University.

De Henau, J., Meulders, D. and S. O' Dorchai, 2006. 'The Childcare Triad? Indicators Assessing Three Fields of Child Policies for Working Mothers in the EU-15'. *Journal of Comparative Policy Analysis: Research and Practice,* 8(2):129-148

Dekker, R., Muffels, R. J. A., and Stancanelli, E. (2000). A Longitudinal Analysis of Part-time Work by Women and Men in the Netherlands. In D. E. Meulders (Ed.), *Gender and the Labor Market: Econometric Evidence of Obstacles to Achieving Gender Equality* (pp. 260-287). New York: St. Martin's Press

Del Boca, D., and Sauer, R. (2006). Life Cycle Employment and Fertility across Institutional Environments. *CHILD Working Papers*, n. 14/2006

Delsen, L., Bosworth, D., Gross, H. and Munoz de Bustillo y Llorente, R. (2007). *Operating Hours and Working Times: A Survey of Capacity Utilisation and Employment in the European Union*. Heidelberg, Physica

Dex, S., Joshi, H., Macran, S., and McCulloch, A. (1998). Women's Employment Transitions around Childbearing. *Oxford Bulletin of Economics and Statistics, 60*(1): 79-98

Drobnič , S. (2000). 'The Effects of Children on Married and Lone Mothers' Employment in the United States and (West) Germany.' *European Sociological Review, 16*(2): 137-157.

Drobnič , S., Blossfeld, H.-P., and 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.

Fouarge, D. and Muffels, R. (2009). Working part-time in The British, German and Dutch Labour Market: Scarring for the Wage Career? , *Schmollers Jahrbuch, 129:* 1-10

Gash, V. (2008). 'Preference or Constraint? Part-time Workers Transitions in Denmark, France and the United Kingdom.' *Work, Employment and Society 22*(4): 655-674.

Ginn et al. (1996). Feminist Fallacies: A Reply to Hakim on Women's Employment, *The British Journal of Sociology*, 47 (1)

Goldin, C. (2006). 'The Quiet Revolution That Transformed Women's Employment, Education, And Family.' *American Economic Review*, *96*(2): 1-21.

Gornick, J. C. (1994). *Women, Employment, and Part-time Work: A Comparative Study of the United States, the United Kingdom, Canada, and Australia.*, Harvard University, Cambridge, MA.

Gutierrez-Domenech, M. (2005). 'Employment after Motherhood: A European Comparison.' *Labour Economics, 12*(1): 99-123.

Hakim, C. (2002). 'Lifestyle Preferences as Determinants of Women's Differentiated Labor Market Careers.' *Work and Occupations, 29*(4): 428-459.

Healy, G. (1999). 'Structuring Commitments in Interrupted Careers: Career Breaks, Commitment and the Life Cycle in Teaching'. *Gender, Work and Organization*, 6 (4): 185-201.

Kalmijn, M., and Luijkx, R. (2006). Changes in Women's Employment and Occupational Mobility in the Netherlands between 1955 and 2000. In H.-P. Blossfeld and H. Hofmeister (Eds.), *Globalization, Uncertainty, and Women's Careers: An International Comparison* (pp. 84-112). Cheltenham, UK, and Northampton (MA-USA): Edward Elgar.

Kenjoh, E. (2005). 'New Mother's Employment and Public Policy in the UK, Germany, the Netherlands, Sweden, and Japan. *Labour 19*: 5-49.

Lewis, J. (1992). 'Gender and the Development of Welfare Regimes.' *Journal of European Social Policy, 2*(3): 159-173.

Lewis, J., T. Knijn, et al. (2008). "Patterns of Development in Work/Family Reconciliation Policies for Parents in France, Germany, the Netherlands and the UK in the 2000s." *Social Politics: International Studies in Gender, State and Society 15*(3): 261-286.

Manzoni, A., R. Luijkx and R. Muffels (2009). Labour market transitions in panel and life-course data: examining differences between the two designs in Germany. *Quality and Quantity*, forthcoming.

Maré, D. (2006). Constructing Consistent Work-life Histories: A Guide for users of the British Household Panel Survey. *ISER Working paper, 39*.

Mayer, K. U. (2007). Retrospective Longitudinal Research: The German Life History Study. In S. Menard (Ed.), *Handbook of Longitudinal Research: Design , Measurement and Analysis*. San Diego: Elsevier.

Miller, C., and Xiao, J. (1999). 'Effects of Birth Spacing and Timing on Mothers' Labor Force Participation.' *Atlantic Economic Journal, 27*(4): 410-421.

Müller, Walter, Lüttinger, Paul, König, Wolfgang and Wolfgang Karle (1989). 'Class and Education in Industrial Nations.' *International Journal of Sociology* 19: 3-39.

O'Connor, J., Orloff, A. S., and Shaver, S. (1999). *States, Markets and Families: Gender, Liberalism and Social Policies in Australia, Canada, Great-Britain and the U.S.*. Cambridge: Cambridge University Press.

Pfau-Effinger, B. (1993). 'Modernisation, Culture and Part-time Employment: The Example of Finland and West Germany.' *Work, Employment and Society, 7*(3): 383-410.

Rosenfeld, R. A., and Birkelund, G. E. (1995). 'Women's Part-time Work: A Cross-national Comparison.' *European Sociological Review, 11*(2): 111-134.

Stier, H., Lewin-Epstein, N., and Braun, M. (2001). 'Welfare Regimes, Family-supportive Policies, and Women's Employment along the Life-course.' *American Journal of Sociology, 106*(6): 1731-1760.

Uunk, W., Kalmijn, M., and Muffels, R. (2005). 'The Impact of Young Children on Women's Labour Supply: A Reassessment of Institutional Effects in Europe.' *Acta Sociologica, 48*(1): 41-62.

Van der Lippe, T. (2001). The Effect of Individual and Institutional Constraints on Hours of Paid Work of Women; an International Comparison. In L. Van Dijk (Ed.), *Women's Employment in a Comparative Perspective,* pp. 221-243. New York: Aldine de Gruyter.

Vlasblom, J.D. and Schippers, J. (2004). 'Increases in Female Labour Force Participation in Europe: Similarities and Differences.' *European Journal of Population 20*: 375–392

Vlasblom, J. D., and Schippers, J. (2006). 'Changing Dynamics in Female Employment around Childbirth: Evidence from Germany, the Netherlands and the UK.' *Work, Employment and Society, 20*(2): 329-347.

Chapter 6

Labor mobility patterns over the life-course: Conclusions

1. The project, its scientific relevance and contribution

In the introduction, we already highlighted the substantive and methodological research questions that the thesis attempts to answer.[91] The substantive questions concern the changing labor market mobility patterns over time and across cohorts using panel and life history data. We tried to explain these mobility patterns by examining their main determinants, as derived from the rich literature on the subject. The main focus was on the effects of career investments like the building up of human capital on exit and re-entry but we also examined the impact of life course events and changes in labor market conditions over time. From a comparative perspective, we explored the influence of country and institutional differences.

Since it is rather risky to derive substantive conclusions from the analysis of allegedly less reliable retrospective data, the substantive questions were supplemented with relevant methodological questions concerning the quality of retrospective data. We were therefore interested in examining the possible sources of bias in retrospective surveys, and our concern was above all with the reliability issue and in particular with the impact of memory bias. The aim was to investigate differences in estimates derived from prospective and retrospective longitudinal survey designs with respect to career patterns and their determinants using the two alternative designs. From a methodological viewpoint, the focus was on the possible implications of the design effect and whether individual determinants of careers are different when the two designs are compared, looking both at micro- and macro-level determinants.

[91] For a more throughout overview of the research questions see the introduction.

We started off the thesis addressing our methodological questions, which constituted the basis on which to build substantive analyses. Although there has been a lot of theoretical discussion on the data reliability of retrospective surveys (see the introduction), the issue has received too little attention in substantive research on labor market dynamics, which seldom takes such concerns into account. Also, while both prospective and retrospective data sources have been extensively used to analyze employment careers, there is hardly any work done comparing and combining such different sources.

We first accounted for differences across survey designs in levels and determinants of career dynamics, and then developed a more substantive interest in the influence of factors derived mainly from human capital and life course theories on career dynamics and in changes over the life course and across cohorts. We also looked into the possible impact that the institutional context might have on these transition patterns. With a view to the alleged structural changes in the life courses of individuals across birth cohorts, we exploited the advantages of both the prospective and the retrospective survey designs to investigate our substantive questions. In particular, given the recent developments in the evolution of the life course (see Section 2 in this chapter for a more extensive discussion), a cohort perspective was taken, for that we believe that it provides a particularly interesting insight into this matter.

While the literature has widely concentrated on employment dynamics and job (in)stability, findings are not always consistent and the debate over increasing instability or volatility on the labor market is still open (Biemann et al., 2009). Although we have no final answer to this debate yet, through our analyses we certainly added to it, in that we arrived at new evidence about these alleged changing patterns across cohorts.

Within such framework, a dynamic analytical approach appeared the proper setting, which had become dominant in the literature in the last decades. Longitudinal surveys have become highly valued due to their potential to provide a long-term view on what is going on in society. Longitudinal data are deemed the most valuable instrument to analyze the life courses of people. Through the introduction of time, they permit the transition from a static to a dynamic view, which allows the analysis of social change. Combining observations from multiple time points, they provide measures of stability or instability, also shedding light on causality.

However, while the successful implementation of longitudinal surveys can be extremely rewarding, there are considerable complexities involved in designing and carrying out a longitudinal survey over and above those that apply to other surveys (cross sections) (Lynn, 2009). This is relevant because low reliability might make longitudinal data a poor source, despite their great potential. Moreover, different types of longitudinal data might have different quality and reveal different patterns.

In order to study changing life-course patterns across cohorts, retrospective data provide a valuable source, in that they collect information on the entire life course of individuals covering a long historical time frame for respondents belonging to different birth cohorts. In sociological research, retrospective data are widely employed to analyze life course patterns. However, they are often criticized due to concerns raised by data contamination due to memory bias (Dex, 1991; Elias, 1997; Paull, 2002; Reimer, 2004; Reimer and Künster, 2004; Scott and Alwin, 1998).

Gauging the reliability of retrospective data and assessing whether and to what extent different dynamic designs diverge in their findings is the first step in our analysis to test the hypotheses about changes in career patterns.

We first examined the quality of retrospective data on employment biographies modeling measurement error looking at retrospective reports with different recall times (Chapter 2) and then comparing the levels (Chapter 3) and the determinants (Chapter 4) of labor market transitions across retrospective and prospective survey designs.

Building on this and correcting for recall bias associated with the length of the recall period, the underreporting of short spells out of employment and seam effects, we then moved the focus to our substantive research questions. Combining the two designs and therefore exploiting the advantages of both, we investigated the effects of substantive factors affecting the career patterns (Chapter 4). The substantive focus was on individual employment trajectories, as the key element of people's life courses and how they change over time, across cohorts and social group. Further to this, we explored the causes of such changes and the way they are affected by particular events in respondents' lives ,as well as their interrelatedness with and dependence on events occurring in other life domains. Eventually, concentrating on retrospective life history data, we were able to investigate more thoroughly the cohort differences in the careers of women at particular instances in their life course, i.e. around childbirth. The chapter further came to grip with how these labor market patterns are affected by the

institutional context by way of making a comparison across a limited set of countries (Chapter 5).

Longitudinal (comparative) data allowed us to analyze patterns of labor mobility over time and across cohorts. Time refers to several distinct dimensions. The interest has been on changing patterns across cohorts and period in particular but also across the various life course stages. We wanted to account for the major determinants of labor market dynamics as derived from the rich literature. The theoretical framework was the life course theory, which departs from the idea that individuals' lives evolve over time over different life course stages and that important events in one's life might affect the subsequent evolution of people's life (and career) patterns. The choice for the lifecourse perspective implied that we examined the impact of so-called 'trigger events' like childbirth or unemployment on people's careers. This appeared the ideal theoretical starting point to capture the impact of career investments indicated by human capital, work history and life course phase on employment mobility patterns across birth cohorts. It also permit to gain more insight not only into the transition patterns over time, across cohorts and socio-economic contexts but also on how transition patterns are affected by events in different life domains such as in family life (divorce) and in employment (job loss).

Overall, this thesis contributed to an empirical assessment of the extent to which different survey methods are likely to generate biased results through recall or measurement error. The information presented allows analysts to judge whether and to what extent substantive conclusions concerning labor market dynamics are likely to be affected by reporting bias. It therefore added to the existing knowledge on the nature of bias in the recall of labor market spells. We concentrated on prospective and retrospective survey designs, assessed their respective advantages and disadvantages (Chapter 3), compared them, and exploited their potential to get a more in depth perspective on the life course of people (Chapter 4). In this way, the thesis contributes to improve the methodological and substantive treatment of labor market mobility patterns over time. On the one hand, we came to more robust conclusions, since the findings derived from these retrospective and prospective data were controlled for bias in the two surveys and corrected for the underreporting of short unemployment spells,

the length of the recall period and seam effects. In addition, we controlled for the residual design difference which were left after taking account of these memory bias sources. The approach adopted here to deal empirically with memory bias also extended the analysis of memory bias, which had so far been confined to psychological research (e.g., Johnson et al., 1998; Sudman and Bradburn, 1973), to life course research . On the other hand, we gained substantive knowledge since, by combining data from panel and life history data, we were able to extend our time horizon and we arrived at more robust substantive findings by correcting for various sources of memory bias. The use of life-course data in a comparative perspective as in Chapter 5 allowed us to employ a cross-cohort perspective over a relatively long time horizon that hitherto, to our knowledge, had not been performed.

2. Background of the study: recent developments on the labor market

The twentieth century has witnessed rapid social change in several life domains: a plurality of new social forms has appeared and insecurity has become pervasive in contemporary life-courses. This concerns not only the labor market but the family sphere as well, with increasing risks such as unemployment or divorce (Vandecasteele, 2007). People's lives have become less predictable and more volatile, and show a larger a variety of life course patterns or trajectories than the standard one observed about 50 years ago. Some authors (Grunow and Mayer, 2007) point to recent structural changes, like the increase in time spent in education and training, the shift from industrial to service employment, the upgrading of the occupational structure, the higher prevalence of part-time work and the wider use of fixed-term contracts. Such changes may contribute to a shift from the standard 'life-time employment' into more diverse career patterns, characterized by more unstable careers, with more volatility and higher risks of intermittent unemployment (allegedly endangering the position of the low skilled in particular).

Consequently, increasing attention has been lately placed on changes in career patterns, which are claimed to be associated with the globalization of the world economy, increasing labor market flexibility and non-standardized work.

However, behavior on the labor market does not only depend on structural constraints; preferences and values are also important driving forces. Labor market preferences can change over the life course, and might be affected by the institutional context. Changes in preferences and values have started to occur, altering the way work (but not only) is viewed and leading to the development of new career concepts. The notion of lifetime employment has been slightly replaced by multiple employment episodes for men, while for women the combination of work and care represents an alternative to withdraw from work and interrupt the career (if any) after marriage or childbirth.
At the same time, the development of family arrangements other than the standard couple, mainly due to the postponement of marriage and the diffusion of divorce, is likely to influence the behavior on the labor market, in particular of women, who more and more often cannot rely on the support of a partner.

Labor market mobility patterns over time are affected by both changes in the institutional and socio-economic context and at the individual level. The latter pertains to changes in people's values, norms and preferences concerning work as well as the family, but also to structural changes in individual characteristics (like education, for example).
On the one hand, internal labor markets, as they still exists in some sectors and countries, tend to protect insiders (middle aged and older workers) and create barriers for outsiders (younger workers) to enter steady employment (Lindbeck and Snower, 2001). They are therefore partly responsible for the increased career instability of the younger cohorts having less chances to enter a steady job (e.g. within the public sector). On the other hand, globalization forces and the process of skill-biased technological change (Acemoglu, 2002; Buchholz et al., 2009; DiPrete, 2002; Mills, 2009) have led to a decline in the demand for low-skilled and have urged firms to react faster and more flexibly to market demand changes. This induced the alleged erosion of internal labor markets in some particular sectors and countries and harmed the stability of the jobs on the internal labor market.
Due to changes in gender role patterns and social norm behavior, also household formation processes altered and changed the life course trajectories of younger cohorts. And household formation is closely interrelated with labor market events. However, differences may occur among different social groups.

The current debate stresses the increasing instability of employment careers, as well as the increasing career heterogeneity across the population (see introduction). Such changes might also translate into a deterioration of workers' labor market position, given that more flexibility may occur also on the employers' side. While gender differences in mobility patterns supposedly tend to become smaller, though they still remain, the increase in education endowments is expected to raise women's employment chances after a break due to childbirth.

The alleged increased differentiation might consequently generate particular problems for workers belonging to the older cohorts, who invested less in human capital and who therefore face more difficulties in finding a new suitable job. People belonging to younger cohorts encounter fewer problems to access employment due to their higher investments in education. Most of the barriers to employment they face occur at the start of the career, when no prior career record is present.

The scarring thesis suggests that the longer unemployment spells are, the lower the chances to re-enter and to move into steady employment are. This might be attributed to stigma effects of unemployment (in the eye of future employers) but also to the loss of human capital and its depreciation (Gangl, 2004, 2006; Barbieri, 2009). On the other hand, the longer one stays in the labor market, the less likely exit occurs due to a higher level of accumulated human capital. This reflects the life-course perspective, which stresses the importance of earlier experiences and relevant life events. These life course events are presumed to affect subsequent outcomes and to have important implication for workers' labor market prospects.

Addressing these issues properly, as any analysis on social change would ask for, calls for a dynamic approach focusing on flows and using longitudinal data. Such data allow us to examine to what extent changes, if any, are related to compositional changes of the workforce or are due to changes over time in the choices people make belonging to different cohorts associated with changes in conditions and/or opportunities.

3. Outline of the main findings

Answering the methodological questions

For answering the methodological research questions, as mentioned before, we first tried to determine the impact of memory bias in retrospective survey designs on the observed transitions and employment sequences over the life course (Chapter 2). We referred to the theoretical debate on the advantages and disadvantage of both panel and life-history survey designs based on an extensive review of the state of the art as far as the literature on longitudinal data reliability is concerned (Chapter 3).

Retrospective data have some practical advantages: they costs less and have a longer time horizon. However, they suffer from errors mainly due to memory bias. The length of the recall period, the complexity of the reporting task, the salience, pleasantness and social desirability of the events being recalled are among the factors which have been shown to influence the degree of recall error. Non-standard episodes (like unemployment), have markedly higher risk to be less well remembered (Dex, 1991; Elias, 1997; Jacobs, 2002; Paull, 2002; Reimer, 2004). Short episodes are more likely to be omitted or merged, leading to simpler, more conventional and stable careers (Reimer and Künster, 2004).

Prospective designs suffer much less from recall bias than retrospective surveys, although they also often rely on some sort of retrospective information to fill in gaps between interviews. However, the length of the recall period is much shorter and usually the longer the recall period, the more the concerns about reliability, even though there are important differences by subject matter and salience (Scott and Alwin, 1998). Yet, biases in recall might occur in the short-term as well (Jürges, 2005; Paull, 2002).

Furthermore, panel data face considerable problems with population representativeness, due to panel mortality and attrition. Panel conditioning may affect the quality of data reported but panel participation may also improve the quality of data, training respondents to keep records and documents that will help them report past events and changes with more accuracy. The multi measurement occasions in panel studies introduce problems related to changing measurement and seam effects.

Our contribution, however, was not limited to a descriptive exploration of the possible sources of biases. On the contrary, we added to the literature, since we identified to

what extent memory bias affects the results and we tried to correct for various sources of measurement error and memory bias in the retrospective surveys we analyzed. We also identified the real differences and similarities in the estimates of the substantive determinants across the two designs after modeling and correcting for memory bias by disentangling memory bias effects from more general design differences (Chapter 3).
We were able to assess measurement error in retrospective surveys for different recall distances and to correct for bias and survey differences when combining different survey designs. We specifically investigated the effect of the length of the recall period, the underreporting of short spells and seam effects.

Our findings confirmed that in retrospective reports over a longer period careers appear simpler, less heterogeneous, characterized by fewer episodes and transitions than in reality, and that short spells are underreported. Unemployment seemed more often mistakenly reported. Complex careers and spells out of employment were found to increase misreporting at longer recall distance, but to significantly affect recall accuracy also of events which had to be recalled at shorter distance, closer to their occurrence. This seems to imply that careers of older cohorts occurring further in the past tend to be reported as more stable than they are in reality and that, due to that misreporting, the alleged shifts in the career volatility might be overestimated when these inferences are derived from life course data. Respondents' socio-demographic characteristics never appeared to affect recall accuracy retrospectively, which is a very reassuring result (Chapter 2). Increasing length of the recall period appeared not to enlarge the gap between the findings from retrospective and panel data, whereas seam effects were found to lead to an overestimation of the transition probabilities at particular time points in the panel data.
Although part of the difference in the level of transition rates between the two survey designs could be explained by short spells recall bias, our findings suggested that dissimilarities cannot be accounted for by memory bias associated with the length of the recall period and short unemployment spells only, nor by compositional differences. We were able to account for most of the factors renown in the literature to affect the reliability of longitudinal survey data; Dealing with the length of the recall period, seam effects, and the occurrence of short non-employment spells allowed us to control for

some differences in the design of panel and life-course surveys, although they did not account for all the differences (Chapter 3).
Overall, the findings warned us against too easily formulated claims about the caveats of life-course studies. Also, comparing them to panel studies, the differences across the two designs seemed more associated with rather general design differences than with the alleged larger recall or memory bias in retrospective designs.
Investigating differences across survey designs in the estimated substantive effects (Chapter 4) brought us to conclude that the main results obtained from the life-history study were very similar to the ones obtained from the panel, supporting the reliability of inferences made from life history data. We found little evidence to support the concerns about the reliability of retrospective data and their alleged lower quality because of memory bias. Correcting for design differences seemed to lead to more reliable estimates and to reduce spurious differences in outcomes between the two surveys with respect to labor market mobility. Only the strength of the substantive effects under consideration turned out sometimes different.

We also investigated whether recall bias might result from the anchoring of employment transitions close to major life events such as marriage or the birth of a child (Chapter 4). We contended that major life events such as marriage or childbirth might act as implicit anchors, using the term 'implicit' to mean that such events are not explicitly mentioned in the questionnaire as landmark events to help people recall their work history, but respondents might use them as such in their mind. Our hypothesis was that people may be more likely to anchor events around childbirth or marriage in the life history study, when they have to remember events over a longer time horizon. The results showed that relevant events in one's life (like marriage or childbirth) do not seem to serve as (implicit) anchors; they didn't appear to lead respondents to temporally connect events, which would affect their recall accuracy and therefore could explain differences between survey designs. Correcting for possible anchoring of labor market transitions around marriage or childbirth was not found to have downsizing effects on design differences. This is a quite reassuring result, in that it provides little evidence that labor market mobility patterns over time are strongly affected by the use of retrospective data of allegedly lower quality.

Answering the substantive questions

Building on such findings, we combined data coming from different designs. While correcting for design differences, we studied the labor mobility patterns of five birth cohorts of West German men and women covering the period from the late 1930s to 2005 (Chapter 4). This led to robust substantive results on the observed changes in mobility patterns by birth cohort, education level and prior career investments. We confirmed the alleged rising volatility on the labor market in West Germany. Younger birth cohorts indeed displayed more mobility than older cohorts. This seemed to be the case for all education levels, in contrast with the notion of skill-biased technological change, according to which workers with higher levels of education and skill should be favored over those with lower levels, due to the increased demand for high skilled workers.

Tertiary education turned out to serve less as a guarantee for a successful career for the younger cohorts than it used to be for the older generations.

The main features of the German labor market as an insider-oriented, institutionalized and highly stratified occupational system coupled to relatively low levels of mobility were instead found to be retained. Prior career investments through work experience or education appeared to pay off especially for the insiders on the internal labor market and less so for the outsiders, such as the young or part-time working women. Due to their caring responsibilities, women were found to face higher exit and reduced re-entry chances than men, supporting the idea that they face a more 'patched career'. Men, instead, showed an opposite behavior, with an increase in their participation on the labor market due to the financial needs implied with having a family.

Our results also corroborated the thesis about 'scarring' effects in the German labor market, with longer time out of the labor market decreasing the chances to re-enter in a stable job later in the career.

Analyzing the determinants of labor market transitions, an important role of familial responsibilities for the labor supply of women emerged. Furthermore, life-history surveys are especially suited to capture cohort effects across time, allowing changes over the entire life course of different cohorts to be investigated. In a panel study, instead, respondents belonging to different cohorts are observed at different life stages; this may blur the cohort effects, due to the difficulty in separating cohorts and life-stages.

For these reasons, using retrospective data and looking at three clearly distinct labor markets and gender employment regimes –i.e. Germany, the Netherlands and Great Britain- we analyzed women's labor market participation over the life course for several cohorts, with a specific attention on how it is affected by their fertility history (Chapter 5). Here, we addressed our substantive questions on the importance of the stage in life course to explain labor market transitions and on the role played by events in respondents' life. The focus is on childbirth events and differences across regimes are investigated. We studied the long-term effects of childbirth on the labor supply in a cohort and cross-country perspective, exploiting the features of life-history data and reflecting the life-course perspective.

By means of an empirical model controlling for the selectivity of women into motherhood and the number of children, we explained mothers' labor supply taking account of the birth sequence and the timing of births over the life-course, revealing that in Great Britain more mothers withdraw from the labor market around childbirth, but also return faster into paid employment in the 20 years following childbirth compared to the other countries. Successive cohorts displayed lower withdrawal rates before and at childbirth and higher participation rates 20 years following childbirth, especially in the Netherlands and Great Britain. Younger birth cohorts showed a larger diversity in participation patterns over the life course.

In all three countries, younger cohorts showed increasing participation rates also when having more children, suggesting that younger cohorts exhibit stronger preferences for combining work and care. Birth sequence and timing of births appeared endogenous to the different career preferences of mothers being primarily oriented at work, the family or a combination of both.

Overall, it emerged that existing income and employment support systems need to become more life-course tuned or targeted to take account of the different career preferences of women belonging to the younger cohorts.

The three countries we analyzed represent three distinct employment and gender regimes. They historically shared the 'male breadwinner' model of welfare regime that evolved over time into a 'dual earner' regime or, in the case of the Netherlands, 'one-and-a-half earner' regime, though with rather different gender labor supply patterns (Lewis, 1992; Lewis et al., 2008). The three regimes significantly differ in the institutional support they provide: The British regime is characterized by a lower level of family

support as indicated by the availability of childcare services (De Henau et al., 2006), which leaves the work and family reconciliation predominantly a private responsibility (Lewis et al., 2008). Public support for mothers' employment is much stronger in Germany and the Netherlands, where parental leave schemes are more generous.

The role of institutions emerged to be very important, in that they can be more or less favorable in accommodating individuals' preferences, contributing to explaining different patterns. In this context, regimes providing a wider portfolio of income and employment support options might also help to create more and better part-time jobs because of shifting social norms in favor of part-time employment.

More in general, some policy implications might be drawn from our study. Life course sociology could be helpful not only for researchers but also for policy makers, especially when the attention is directed to structural long term changes. Life-course policies are still in their infancy and there is no clear view as to how policy might address these life course changes. Policy might however learn from the results of this type of research investigating long-term and structural changes. The substantive findings corroborated the alleged cohort changes over time related to shifts in mothers' labor market behavior around childbirth (Chapter 5) but contrasted the alleged labor market effects of the shift in the demand for higher education and therefore provided substantive evidence that seems highly relevant for policy (Chapter 4).

4. Discussion and suggestions for further research

In the following, we proceed with a more critical evaluation of our work, stimulating a discussion on possible limitations with respect to the way we tackled our research questions. We outline some possible improvements on our approach, in addition to specific ones which were acknowledged in each single chapter. We believe that they might stimulate future research into the topic.

Although we were able to answer the main research questions we formulated at the start of the study, any research, due to its very nature, always leads to a (virtuous) circle, in which new interests are stimulated. Sometimes due to data restriction, sometimes to space constraints, we couldn't deal with all the interesting questions which arose in the course of our research. This generates some room for further investigation, which we

would be eager to carry out ourselves, but which of course may also constitute a source of inspiration for other scholars interested in such issues.

Limitations to our research might arise from the nature of the data used and from the choices of methods as well. Our analyses are based on specific datasets and on a number of methodological decisions, which are arbitrary, to a certain extent, and represent just one of many possible alternatives. Our conclusions might therefore be driven by the arbitrary choices we (necessarily) made in the course of the analyses.

Studying reliability

Starting from our methodological question, ideally, there would be an accessible external validating source of information. However, in most cases no such external source is available, since social reality is always represented with error in any kind of data. Register data are generally considered highly valid and consequently the comparison with them may offer a way to assess (survey) data reliability. However, some kinds of error occur in each kind of data. Register data are usually gathered for different purposes, and tracking individual history longitudinally is not their original goal. Problems might arise according to the specific focus of the comparison and the changes which are to be tracked. Besides that, such data are not always easy to obtain and law restrictions regarding privacy might make it hard to link respondents.

Consequently, an alternative approach is needed. We compared information coming from different sources, both of which might not be free from error. One comparison related to data obtained from two different interviews which were part of the same retrospective survey (the Swedish Level of Living Survey) and were carried out at different points in time (Chapter 2) but with overlapping periods; another comparison related to panel versus retrospective survey data (Chapter 3 and 4).

The availability of two indicators for each observation (person-month) in the Swedish data provided us with a rich source of information because it allowed us to overcome compositional differences as well as any difference which might be related to the question phrasing or the questionnaire structure (Chapter 2). Applying a latent Markov model to such data, we were able to correct for measurement error, which could arise from several circumstances during the interview and the data handling process as well. This might affect each report differently, and, with a specific view to memory bias, it relaxed the assumption that errors are independent across time points. In this way, we

could make the process closer to reality, so to arrive at a better definition of the true (latent) state.

Using panel and retrospective data (Chapter 3 and 4), we were able to control for seam effects in the first and for length of the recall period in the second, as well as for the underreporting of short unemployment spells. When combining the two sources, we were also faced with compositional differences in the samples coming from different sources. However, although residual differences might remain, we were able to (partly) control for them by including most of the (observed) characteristics which may supposedly account for such differences. The inclusion of observed sample characteristics also allowed us to isolate survey effects: Residual differences were captured introducing a survey dummy. Therefore, we dealt with the factors renown in the literature to affect memory bias and, although we couldn't rule out all design differences, we certainly did the best we could do, modeling recall bias as good as possible.

To obtain monthly information, also panel surveys rely on some sort of retrospective questions referring to the previous year. Therefore, we also conducted analyses on the reliability of such short-term retrospective information (which basically serve to fill in gaps between interviews), comparing calendar information on the previous year with panel information on the current status. However, such analyses go beyond our specific interest in the reliability of life-history data (and are therefore not treated in the thesis), and just provided the authors with an additional check.

Event history analysis as the reference paradigm

Moving the focus to the substantive questions, we grounded our discussion in the event history framework. In the last three decades, the sociology of the life course has emerged as a substantial field in sociology. Event history analysis (EHA) has been the method of choice from the emergence of life course sociology as a substantial field in sociology and today it is basically established as the exclusive method to analyze life course transitions (Aisenbrey and Fasang, 2007). However, while event history analysis focuses on transitions, life course research also focuses on trajectories, which might suggest a different approach.

Sequence analysis might offer a valuable insight into the identification of life course patterns across cohorts or societies, answering substantial theoretical questions that

cannot adequately be conceptualized in the form of discrete transitions or single durations. This method was introduced to the social sciences by Andrew Abbott over two decades ago (Abbott and Forrest, 1986), and it has been emphasized by several authors to lend itself to life course related problems (Abbott, 1995; Abbott and Tsay, 2000; Billari and Piccarreta, 2005; Han and Moen, 1999). To date, however, it has not been broadly accepted as a complementary approach in life course sociology. Many concerns still lurk behind the application of sequence analysis, because of which it is not a real alternative yet to event history analysis. It might become a serious alternative, but much research still needs to be done before that happens (Aisenbrey and Fasang, 2007). This approach, although it can be used mainly in an explorative way, may offer new insight into the explanation of career patterns. We applied it at least partially in Chapter 2. [92]

However, although we recognized the potential of sequence analysis, which might offer a different and very interesting perspective to the issues considered in this thesis, we believe that it doesn't yet offer a sufficiently developed framework to address the (substantive) research questions we are interested in.[93]

Therefore, we opted for event history analysis as the proper method to analyze our data. Within this framework, we analyzed labor market dynamics focusing on single job spells. We concentrated on specific transitions (Chapter 3 and 4) on the labor market, i.e. exit, re-entry and job-to-job moves and looked at participation in the labor market at specific stages in the life course (Chapter 5).

Both the occurrence and the timing of transitions are relevant, for methodological as well as substantive reasons. Respondents might entirely forget episodes but they may also just misdate their occurrence. Moreover, in a life-course perspective, the interest could be on the link between the different events and their timing.

[92] In the first chapter, we do introduce sequence analysis, as an algorithmic approach to investigate memory bias in retrospectively reported employment trajectories, stressing its assumptions and shortcomings, opposing it to a model based approach. However, we confine our use of such technique to a tool to compare reported biographies while, to investigate our substantive question, we only move within the EHA framework, for the reasons exposed above.

[93] See chapter 1 for an overview of the main critiques to sequence analysis and Optimal Matching in particular.

Suggestions for further investigation

The complexity of the matter certainly invites additional research on the topic. Here we hint to some possible directions. When dealing with exit, we did not specify the destination to which individuals moved. However, in this way we pooled a number of very different situations, since people might leave employment for many different reasons: While older people mostly exit to retire, others might want to go back to education, or take care of their family, to mention some examples. Furthermore, people may go out of the labor market permanently or just temporarily, for a shorter or longer period of time.

Also, looking at re-entry, many different situations could be distinguished, since people may re-enter employment not only after a longer or shorter period out (which we take into consideration), but also after having been out for different reasons, and they may reenter in a different position compared to their previous one. The mechanisms behind occupational mobility might be different for upward and downward moves, which would be an interesting aspect to consider in the case of job to job mobility as well, but unfortunately the data we are using do not offer comparable information to investigate this further.

Looking at (mothers') labor market participation around childbirth, we limited the analysis to the choice between employment and non-employment. A further interesting and challenging extension would involve investigating different occupational choices across and within cohorts across countries. Descriptive analyses (not shown in the thesis) on our life-history data on Germany, Great Britain and The Netherlands, for example, showed an upward shift in occupational level of mothers across cohorts (that is explained by the rising educational level). The occupational level in the years after childbirth appeared to decrease for younger cohorts (who are the least selective in terms of employment chances), while hardly any change seemed to occur for older cohorts (who are the most selective in terms of employment chances). However, cohort differences could be due to changes in the self-selection process (in employment, childbearing and occupational choices), which would require complicated panel regression estimation techniques, but the data at hand were not rich enough to model the selection processes into paid employment and into occupations alongside with selection into motherhood simultaneously. Nevertheless, this would be worth to be

further studied, preferably with data from more countries, so to be able to argue about the reasons for the observed differences within cohorts across countries.

Another very important issue about labor market participation concerns the distinction between part-time and full-time employment. The decision about labor market participation regards not only whether people participate or not in the labor market, but also the amount of hours worked. Furthermore, also from the demand side perspective, the availability of different types of jobs might play a role, especially for women who wish to combine work and care. The availability of part-time employment, which is both supply and demand driven, has indeed been shown to play an important role in explaining participation, especially for women. Countries do differ significantly as to the supply as well as to the demand for part-time workers. We acknowledge the importance of such issues but were unable to capture it properly in the empirical analysis. [94]

Our analyses didn't dig into the extent to which the debate on stability versus volatility might be due to the rise of part-time and temporary jobs.[95] We just looked into transition patterns for all the workers, without making a distinction between temporary and permanent workers. It might however be the case that stability or volatility patterns involve groups of workers differently, depending on their type of contract.
Moreover, all types of transitions might be the result of a voluntary choice or a constraint, and voluntary or involuntary moves might account for different mechanisms. Increased labor market flexibility might point to a greater freedom of choice for individuals to determine their careers but also to more constraints put on them since it gives more room to employers to hire and dismiss employees at will. This debate is still open.

[94] This was mainly due to data restrictions, since information on the amount of hours worked was not always available or anyway not comparable across datasets, but also to avoid additional complications of our models, which were already quite complex. Furthermore, our decision not to focus specifically on working-time arrangements was also driven by the belief that (the growth of) part-time employment is more important when looking at the behavior of more recent cohorts, especially when looking at the German employment regime.
[95] Statements on the employment contract were not collected in the GLHS (besides in the last wave), which makes such analysis not possible, at least in a comparative perspective.

In all our analyses, we mainly focused on the individual determinants of labor market dynamics and not on the contextual ones. It might be worth paying more heed to these context effects.

We mainly looked into the within country dynamics, focusing a large part of the substantive conclusions on the German context. Germany has a highly rigid and protective employment system compared to other OECD countries, which might result in more stable careers in Germany. Although Germany is a major Western economy and a typical case with a view to the impact of economic globalization, and the increasing cultural, societal, and political differences, the situation might vary across other countries, where the cultural, societal and political context might have a different effect on the underlying dynamics. The generalizability of the results to other countries remains therefore disputable. It would be interesting to replicate the analyses of patterns and determinants of employment careers also in different contexts and with different data, so as to verify whether the same mechanisms are at play in different contexts and whether results are not specific to the survey data we used.

However, we concluded the thesis adopting a cross country perspective (Chapter 5) for the analysis of women's labor supply around childbirth. There, we showed differences across welfare and gender regimes, referring to the gender regime literature as the frame of reference to deal with the impact of institutions. We believe that the agency-structure approach is the appropriate reference in which to frame the discussion but, unfortunately, we were not able to formally test the impact of structure/institutions, due to having only three countries.

As we mentioned, lives are embedded in a cultural and historical context and are therefore subject to change. At the same time, lives are not simply the result of external forces but are also shaped by the actions of the individuals themselves and are linked to other people as well. We recognize the importance of preferences, alongside with constraints: both preferences and structural constraints are important driving forces behind labor participation.

In particular, when analyzing mothers' behavior on the labor market around childbirth, we used age at first birth as an (imperfect) proxy for work orientation, since we lacked information on career orientations or preferences. However, career orientations could change over the life course. Furthermore, one may contend that women who enjoyed

more education will, on average, become first time mothers at a later age. This might be related to career orientations but it might also merely signal a postponement effect, because it takes more time to complete higher education. Yet, we found no evidence that the timing of childbirth is conditioned by education level (see Chapter 5).
Education choices might also be endogeneous because career oriented women might choose for extending their enrolment in education to improve their chances for a career, whereas family oriented women for setting up a family choose for shorter education. Younger cohorts might also behave differently from older cohorts because of having more opportunities to work even when they have lower education. We controlled for education level in the model but, if education is endogenous, we need good instruments for educational choices (which are not readily available in the data).
Tentative analyses on our data seemed to indicate that the relationship between age at first birth, education and career orientation is more complex than assumed and would need more scrutiny. Unfortunately, we had no comparative direct information on career orientations. Further evidence on career preferences of women and support policies for a larger number of countries would be needed to be able to disentangle the effects of changes in fertility behavior, career orientations over time and policies and gender regimes.

All the mentioned improvements notwithstanding, we believe that we came to interesting new results in the study. If nothing else, our results removed the debate from the realm of speculation and hypothetical data to that of empirical proof.
The analyses carried out in this thesis provide empirical evidence to answer the methodological research questions about longitudinal data reliability and about the effect of memory bias specifically. We were able to investigate it extensively and to correct for it, which constituted the basis on which we built to arrive at a sound empirical investigation of our substantive research questions. Although we cannot always control for all the sources of bias, we certainly increased our knowledge on that.

An interdisciplinary approach
We adopted a sociological perspective, but we also recognized the value of an interdisciplinary approach and of potential contributions from other fields, which may

expand our understanding of career processes. Accordingly, we sometimes referred to economic and psychological theories and borrowed concepts from demography as well. While psychological theories pervaded the argumentation on recall bias, economic theories contributed to explain the determinants of career transitions and the link with demographic issues emerged more clearly when accounting for mothers' participation. We believe that crossing disciplinary boundaries and expanding our conceptualization of careers beyond the traditional models we have much to gain, since the study of the life course is a true interdisciplinary endeavor.

Given that the findings on labor market mobility patterns over time did not turn out to be strongly affected by the use of retrospective data, we make a strong argument for life-history data, encouraging researchers' confidence in using them to study long-term changes in the labor market.

Life-course data do not seem to be as bad as most people think. Problems seem limited and specific to some cases, i.e. unemployment, complex careers and short spells, which implies that some contexts might be better suited than others to be analyzed with retrospective data. Furthermore, panel data are not immune from bias either and each type of survey design might actually generate some problems. However, once aware of possible biases, it appears possible to correct for them, at least to some extent, so to arrive at more robust conclusions about substance.

Overall, different types of data seem to complement each other. The accuracy of recall is very likely to be specific to the type of information retrieved, which should guide researchers in the choice of the appropriate data and also warn them to always acknowledge the possible limitations of the data at their disposal.

Researchers should choose their data sources according to their specific research questions: Different research interests might make different sources of data more suitable than others. Furthermore, the combination of prospective and retrospective data sources appears useful and promising with a view to extending the time horizon of and the cohort sizes involved in longitudinal data. It might therefore be a good way to proceed, especially given that the combination of prospective and retrospective data sources seemed to give us a broad view into long-term trends in labor market mobility and the occurring shifts in mobility patterns across birth cohorts over time (Chapter 4).

Although we cannot infer from these results that one type of data collection method is superior to the other, for which reason a concluding assertion about the comparative added-value of retrospective data would be premature, some advice to survey designers can be drawn out of this study. We found for example that while a focus on the levels of mobility, especially in flexible labor markets where intermittent spells occur frequently and careers are more complex, might make panel data more suitable, the use of retrospective data appears less problematic if the interest is in its determinants. Moreover, retrospective data appear particularly advantageous in some circumstance, for example when the focus is on a cohort perspective.
Overall, we conclude that, to make plausible inferences about labor dynamics, a deep insight into the implications of the survey design for the research findings and into the way the information has been recalled and reported appears needed. Researchers should always be aware of the specific design features of their data sources and acknowledge the possible impact that their data design and quality might have on the inferences they want to make.

We also want to remark that a greater clarity of terms is needed, as well as more uniformity in the survey practice and the way different topics are addressed in data collecting processes, to avoid confusion and to assure the comparability of different analyses. It is often the case that either many different concepts come under the same definition or, on the contrary, the same concept is referred to with a variety of terms.
This is particularly relevant as regards the recent ongoing debate on the (in)stability of labor market careers. The first step in trying to solve such debate would therefore be a more precise definition of what is meant with stability versus volatility and above all a uniform way to operationalize the concept, to make sure that the various research endeavors look into the same processes and mechanisms on the labor market.

Eventually, we want to stress the need to engage in comparative work in multiple ways: adopting a cross-country perspective, using multiple methods of data collection, but also different approaches and perspectives in the analysis, applying our theoretical and empirical tools to a variety of research settings.

References

Abbott, A. (1995) Sequence Analysis: New Methods for Old Ideas. *Annual Review of Sociology*, Vol. 21., pp. 93-113.

Abbott, A. and Forrest, J. (1986) Optimal matching methods for historical sequences, Journal of Interdisciplinary History, XVI(3), 471–494.

Abbott, A. and Tsay, A. (2000) Sequence analysis and optional matching methods in sociology, Sociological Methods and Research, 29(1), 3–33.

Acemoglu, D. (2002). Technical Change, Inequality, and the Labor Market. Journal of Economic Literature 40, 7-72.

Aisenbrey, S. and Fasang, A. (2007) Beyond Optimal Matching: The 'Second Wave' of Sequence Analysis. CIQLE WP 2007-2, Yale University, New Haven.

Billari, F. C. and Piccarreta, R. (2005). Analyzing Demographic Life Courses through Sequence Analysis. *Mathematical Population Studies* 12:81-106.

Brückner, H. and Mayer, K. U. (2005). The De-Standardization of the Life Course: What It Might Mean and Whether It Actually Took Place. Advances in Life Course Research 9:27-54.

Buchholz, S., Hofacker, D., Mills, M., Blossfeld, H. P., Kurz, K., Hofmeister, H., 2009. Life Courses in the Globalization Process: The Development of Social Inequalities in Modern Societies. European Sociological Review 25, 53-71.

Dex, S. (1991) The Reliability of Recall Data: A Literature Review. ISER Working Paper. Colchester, University of Essex.

DiPrete, T. (2002). Life course risks, mobility regimes, and mobility consequences: A comparison of Sweden, Germany, and the United States. *American Journal of Sociology, 108 No.2*.

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 3:508-532

Elias, P. (1997) Who forgot they were unemployed? ESRC Working Paper

Fouarge, D. and Muffels, R. (2009). Working part-time in The British, German and Dutch Labour Market: Scarring for the Wage Career? , *Schmollers Jahrbuch, 129:* 1-10

Halpin (2003) Tracks Through Time and Continuous Processes: Transitions, Sequences, and Social Structure. Paper presented at the meeting 'Frontiers in Social and Economic Mobility', Cornell University, March 27-29, 2003

Han, S. and Moen, P. (1999). Clocking Out: Temporal Patterning of Retirement. *American Journal of Sociology* 105:191-236.

Jacobs, S. C. (2002) Reliability and recall of unemployment events using retrospective data. Work, Employment and Society, vol. 16, 537–548

Johnson R. A, Gerstein, D. R. and Rasinski, K.A. (1998). Adjusting survey estimates for response bias: An application to trends in alcohol and marijuana use Public Opinion Quarterly 62(3): 354-377.

Jürges, H., 2005. Retrospective error in SOEP calendar data: the case of unemployment. Schmollers Jahrbuch 125, 157-65.

Lindbeck, A., Snower, D. J., 2001. Insiders versus Outsiders. The Journal of Economic Perspectives 15, 165-88.

Lynn, P. (2009). Methodology of Longitudinal Surveys, John Wiley

Magidson, J., Vermunt, J.K., and Tran, B. (2009). Using a mixture latent Markov model to analyze longitudinal U.S. employment data involving measurement error. In: K. Shigemasu, A. Okada, T. Imaizumi, and T. Hoshino (eds.), *New Trends in Psychometrics*. Universal Academy Press, Inc.

Mathiowetz, N. A., Duncan, G. J. (1988). Out of work, out of mind: Response errors in retrospective reports of Unemployment. Journal of Business and Economic Statistics 6, 221-229.

Mills, M. (2009). Globalization and Inequality. European Sociological Review 25, 1-8.

Paull, G (2002). Biases in the reporting of labour market dynamics. London, The Institute for Fiscal Studies.

Powers, E. A., Goudy, W. J. and Keith, P. M. (1978). Congruence Between Panel and Recall Data in Longitudinal Research. Public Opinion Quarterly 42, 380-389.

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, Berlin, Max-Planck-Institut für Bildungsforschung.

Reimer, M., Künster, 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. Berlin, Max-Planck-Institut für Bildungsforschung.

Scott, J., and Alwin, D. (1998). Retrospective Versus Prospective Measurement of Life Histories in Longitudinal Research. In Giele, G. H., Elder, J. Z. (eds.) Methods of Life Course Research: Qualitative and Quantitative Approaches. Thousand Oaks et al.: Sage Publications

Sudman, S., and Bradburn, N. M. (1973). Effects of Time and Memory Factors on Response in Surveys. Journal of the American Statistical Association 68, 805-815.

Vermunt, J. K., Tran, B., and Magidson, J. (2008). Latent class models in longitudinal research. In S. Menard (Ed.), Handbook of Longitudinal Research: Design, Measurement, and Analysis. Elsevier. In press.

Samenvatting

Dit proefschrift bevat een aantal empirische studies die de dynamiek op verschillende Europese arbeidsmarkten zowel inhoudelijk als methodologisch analyseren, daarbij gebruikmakend van twee designs: *panel* en *life history* data. We onderzoeken enerzijds factoren die de arbeidsmarktdynamiek beïnvloeden en anderzijds de mogelijke oorzaken van *bias* in verschillende longitudinale onderzoeksdesigns. We gaan na wat de inhoudelijke effecten van zulke factoren zijn en wat de gevolgen van *bias* zijn, waarbij we proberen te corrigeren voor meetfouten en geheugen *bias*. We richten ons met name op loopbanen, maar bekijken ook de relatie met andere levensdomeinen. Zowel arbeidsmarktparticipatie als arbeidsmarkttransities worden geanalyseerd.

De belangrijkste inhoudelijke vragen gaan over veranderende patronen in arbeidsmarktmobiliteit over de tijd en tussen cohorten. We kijken naar de effecten van investeringen in carrières, maar ook naar de invloed van levensloopgebeurtenissen en veranderingen in arbeidsmarktcondities over de tijd. Tevens verkennen we vanuit een vergelijkend perspectief de invloed van landen en instituties.
Omdat het vrij riskant is om inhoudelijke conclusies te trekken op basis van retrospectieve data die doorgaans minder betrouwbaar worden geacht, vullen we de inhoudelijke vragen aan met relevante methodologische vragen over de kwaliteit van retrospectieve data. We onderzoeken de mogelijke oorzaken van *bias* waarbij we vooral aandacht besteden aan de invloed van geheugen *bias*.
We gaan na in welke mate *life history* en *panel* data tot dezelfde conclusies leiden wat betreft arbeidsmarktmobiliteit over de tijd waarbij we zoveel mogelijk proberen te corrigeren voor geheugen *bias* zodat we de inhoudelijke onderzoeksvragen adequaat kunnen beantwoorden.
We onderzoeken of er verschillen zijn in individuele micro- en macrodeterminanten van loopbanen wanneer we de twee designs vergelijken en behandelen mogelijke implicaties van het design effect.

De onderzoeksvragen vatten we als volgt samen:
Inhoudelijke onderzoeksvragen:
IV1: In welke mate veranderen arbeidsmarktloopbanen over de tijd en tussen cohorten?
IV2: Worden arbeidsmarktloopbanen instabieler of niet?
IV3: Zijn er verschillen in de mate waarin arbeidsmarktloopbanen veranderen tussen verschillende groepen in de samenleving?
IV4: Wat zijn de oorzaken van veranderende arbeidsmarktloopbanen? Of meer specifiek, wat is de rol van de toenemende vraag naar menselijk kapitaal, het stijgende aanbod van onderwijskwalificaties en de toenemende arbeidsmarktparticipatie van vrouwen?
IV5: In welke mate spelen levensloopgebeurtenissen (zoals werkloosheid, geboorte van een kind of scheiding) en levensfase een rol?
IV6: Komen veranderingen in mobiliteitspatronen meer voor in landen met een ongereguleerd arbeidsmarktregime dan in landen met een gereguleerd regime?

Methodologische onderzoeksvragen:
MV1: Wat zijn de oorzaken van *bias* in prospectieve en retrospectieve longitudinale data?
MV2: In hoeverre beïnvloeden designkenmerken, en geheugen *bias* in het bijzonder, de effecten van arbeidsmarktdynamiek?
MV3: Hoe kunnen we corrigeren voor geheugen *bias*?
MV4: Leidt het gebruik van prospectieve data tot andere schattingen van niveaus en determinanten van arbeidsmarktdynamiek dan het gebruik van retrospectieve data?
MV5: Wat zijn de voor- en nadelen van het combineren van twee databronnen om tot betere schattingen van de determinanten van arbeidsmarkttransities te komen?

We geven eerst antwoord op de methodologische vragen (Hoofdstuk 2 en 3) om zo een goede basis te scheppen waarop we de antwoorden op onze inhoudelijke vragen (Hoofdstuk 4 en 5) kunnen baseren. Doordat we de methodologische vraagstukken over data betrouwbaarheid en (retrospectieve) geheugen *bias* aanpakken, kunnen we de dynamiek in carrières zuiverder modelleren.
Het schatten van de betrouwbaarheid van retrospectieve data en het bepalen of en in welke mate verschillende dynamische designs in hun bevindingen uiteenlopen zijn de

eerste stappen in onze analyses om hypothesen over veranderingen in carrièrepatronen te toetsen.

We onderzoeken eerst de kwaliteit van retrospectieve data op werkgeschiedenissen waarbij we meetfouten modelleren door te kijken naar retrospectieve data waarin respondenten gevraagd is naar verschillende periodes uit hun leven (Hoofdstuk 2). Daarbij vergelijken we de niveaus (Hoofdstuk 3) en determinanten (Hoofdstuk 4) van arbeidsmarkttransities tussen retrospectieve en prospectieve survey designs.

Hoewel de algemene bevindingen tamelijk hetzelfde blijken te zijn voor de twee survey designs, zijn de geschatte transitiekansen in de levensloopdata significant lager dan in de panel data. Als het om design verschillen in de determinanten van transities gaat, concluderen we dat de belangrijkste resultaten voor beide designs nagenoeg hetzelfde zijn, wat erop wijst dat de bevindingen van de levensloopdata betrouwbaar zijn. We vinden weinig bewijs voor de algeheel gedeelde zorg dat retrospectieve data onbetrouwbaar zouden zijn, noch voor hun zogenaamde mindere kwaliteit vanwege geheugen *bias*. Corrigeren voor designverschillen lijkt tot meer betrouwbare schattingen te leiden en vermindert schijnverschillen in arbeidsmarktuitkomsten tussen de twee surveys. Alleen de sterkte van de inhoudelijke effecten verschillen soms.

Hierop voortbouwend verplaatsen we onze aandacht naar de inhoudelijke vragen, waarbij we corrigeren voor geheugen *bias* op basis van de duur van de teruggevraagde periode, het onderrapporteren van korte periodes van werkloosheid of inactiviteit en 'naad' effecten. Door het combineren van de twee designtypen gebruiken we de voordelen van elk van de designs om de effecten te onderzoeken van inhoudelijke factoren die loopbaanpatronen bepalen (Hoofdstuk 4). De inhoudelijke focus is gericht op individuele werktrajecten, die een belangrijke rol spelen in de levensloop van mensen. We kijken ook naar veranderingen in deze trajecten over de tijd, tussen cohorten en sociale groepen.

Er is een groot aantal studies verricht naar de dynamiek in werk en baan(in)stabiliteit, maar de bevindingen zijn niet altijd consistent; er is nog altijd geen consensus in hoeverre er sprake is van stijgende instabiliteit op de arbeidsmarkt (Biemann et al., 2009). Hoewel we geen duidelijk antwoord op deze vraag kunnen geven, dragen onze analyses zeker bij aan het debat hierover omdat we met nieuwe bevindingen komen over veranderingen tussen cohorten.

Daarnaast verkennen we de oorzaken van zulke veranderingen en de manier waarop ze worden beïnvloed door bepaalde levensloopgebeurtenissen, alsook hun samenhang met en afhankelijkheid van gebeurtenissen in andere levensdomeinen. Tenslotte onderzoeken we – gebruikmakend van retrospectieve levensloop data – cohortverschillen in de loopbanen van vrouwen op bepaalde momenten tijdens hun leven, zoals rond de geboorte van een kind (Hoofdstuk 5).

Dit proefschrift vormt een methodologische en inhoudelijke bijdrage aan onderzoek naar individuele arbeidsmarktveranderingen over de tijd.
Enerzijds stelt het empirisch vast in hoeverre verschillende survey methoden *bias* genereren door geheugen of meetfouten. Het helpt onderzoekers om te oordelen of en in hoeverre inhoudelijke conclusies over dynamiek op de arbeidsmarkt mogelijk beïnvloed worden door *bias*.
Onze conclusies zijn robuuster dan eerdere conclusies omdat we controleren voor *bias* en corrigeren voor onderrapportage van korte werkloosheids*spells*, de duur van de teruggevraagde periode en 'naad' effecten. Bovendien controleren we voor designverschillen die overblijven na controle voor deze geheugen *bias*. Ook leveren we een bijdrage aan de reeds bestaande kennis over de aard van *bias* in retrospectief gevraagde arbeidsmarkt*spells* door nieuwe benaderingen te bieden om empirisch te corrigeren voor geheugen *bias*. Onze correctie voor geheugen *bias* vormt ook een uitbreiding van het huidige geheugen *bias* onderzoek – dat tot nu toe vooral van psychologische aard was (e.g., Johnson et al., 1998; Sudman and Bradburn, 1973) – naar levenslooponderzoek.
Anderzijds boeken we inhoudelijke vooruitgang omdat we onze tijdshorizon verbreden door het combineren van *panel* en levensloop data en we een meer robuust antwoord kunnen geven op de vraag in hoeverre arbeidsmarkten veranderen en hoe die veranderingen het gedrag en transitiepatronen van verschillende cohorten beïnvloeden.

De resultaten uit dit proefschrift verschuiven het debat van speculatie en hypothetische data naar dat van empirisch bewijs.
We bieden empirisch bewijs om methodologische onderzoeksvragen te beantwoorden over betrouwbaarheid van longitudinale data en dat van het effect van geheugen *bias* in het bijzonder.

Het blijkt dat onze bevindingen over veranderingen in arbeidsmarktmobiliteit door de tijd slechts beperkt worden beïnvloed door het gebruik van retrospectieve data. Daarom pleiten we sterk voor het gebruik van levensloop data, waarbij we onderzoekers die lange termijn veranderingen op de arbeidsmarkt bestuderen, willen aanmoedigen dergelijke data te gebruiken.

Levensloop data blijken niet zo slecht te zijn als de meeste mensen denken. De problemen lijken zich te beperken tot bepaalde gevallen, zoals werkloosheid en complexe loopbanen, wat impliceert dat sommige contexten beter onderzocht kunnen worden met retrospectieve data dan andere. Verder blijkt het mogelijk om – tot op zekere hoogte – te corrigeren voor *bias*, om tot meer robuuste inhoudelijke conclusies te komen. De combinatie van prospectieve en retrospectieve databronnen blijkt nuttig en veelbelovend als het gaat om een bredere tijdshorizon en de grootte van cohorten in longitudinale data. Deze aanpak kan daarom vaker in de toekomst worden gebruikt, vooral omdat de combinatie van de twee designs ons een brede blik bieden op lange termijn trends in arbeidsmarktmobiliteit en de veranderingen in mobiliteitspatronen tussen cohorten over de tijd.

Samengevat blijken de verschillende typen data elkaar aan te vullen. In hoeverre het geheugen accuraat is, lijkt echter gerelateerd te zijn aan het type informatie dat gevraagd is. Dit biedt enerzijds een leidraad voor onderzoekers in de keuze van de juiste data gegeven hun onderzoeksvragen, en is anderzijds een waarschuwing voor hen om stil te staan bij de mogelijke beperkingen van de data die ze gebruiken.

Hoewel we uit onze bevindingen niet kunnen afleiden dat de ene methode van data verzameling beter is dan de andere – waardoor een conclusie over de comparatieve toegevoegde waarde van retrospectieve data voorbarig zou zijn – kunnen we wel wat aanbevelingen doen voor de makers van surveys. We vonden bijvoorbeeld dat panel data meer geschikt lijken wanneer het onderzoek zich richt op mobiliteitsniveaus, en dan vooral in flexibele arbeidsmarkten waar veel fluctuatie in *spells* voorkomt en loopbanen complexer zijn, terwijl retrospectieve data minder problematisch blijkt te zijn als het onderzoek zich meer op de determinanten van mobiliteit richt. Bovendien blijkt gebruik van retrospectieve data vooral adequaat in specifieke omstandigheden, zoals in cohortstudies.

We concluderen dat om tot plausibele conclusies over dynamiek op de arbeidsmarkt te komen, goed inzicht nodig is in de gevolgen van het gebruikte survey design voor de

onderzoeksbevindingen en in de manier waarop informatie is teruggevraagd en gerapporteerd. Onderzoekers zouden zich altijd bewust moeten zijn van de specifieke kenmerken van het data design dat ze gebruiken en ze moeten zich realiseren dat het design en de kwaliteit van de data hun conclusies kan beïnvloeden.

Anna Manzoni

Labor mobility patterns over the life-course.

A comparison of retrospective and prospective data in different labor markets

This study consists of a collection of empirical studies on labor market dynamics in various European countries using panel and life history data. It contributes to an assessment of the extent to which different survey methods are likely to generate biased results through recall or measurement error. It adds to the existing knowledge on the nature of bias in the recall of labor market spells, proposing new approaches to deal empirically with memory bias.

The main substantive questions refer to the changing labor market mobility patterns over time and across cohorts. Since it is rather risky to derive substantive conclusions from the analysis of allegedly less reliable retrospective data, the substantive questions are supplemented with relevant methodological questions concerning the quality of retrospective data.

The extent by which life history and panel data come to similar or different conclusions with respect to labor market mobility over time is examined, correcting for memory bias in order to be able to properly address the substantive research questions.

The possible implications of the design effect are investigated, focusing on whether individual determinants of careers are different when the two designs are compared, looking both at micro- and macro-level determinants.

Anna Manzoni (1981) studied Sociology at the University of Milano-Bicocca (IT). She conducted the present research between September 2005 and August 2009 at Tilburg University (NL).

As from September 2009, she has been working as a postdoctoral researcher at the Center for Research on Inequalities and the Life-Course at Yale University (USA).