Blurred Memory, Deliberate Misreporting, or "True Tales"?
How Different Survey Methods Affect Respondents' Reports of Partnership Status at First Birth

Michaela Kreyenfeld (kreyenfeld@demogr.mpg.de)
Sonja Bastin (bastin@demogr.mpg.de)

Blurred Memory, Deliberate Misreporting, or "True Tales"?

How Different Survey Methods Affect Respondents' Reports of Partnership Status at First Birth

Michaela Kreyenfeld (Max Planck Institute for Demographic Research)

Sonja Bastin (Max Planck Institute for Demographic Research)

Abstract

This paper examines the reliability of biographical information gathered retrospectively. It draws on data from the German Family Panel (pairfam), which collected information on the partnership status at first birth using two different methods. The first method is based on data on partnership and fertility histories collected retrospectively. The second method uses data gathered through the use of a "landmark question" on the respondents' partnership status when their first children were born. We find that in almost 20 percent of the cases, the information collected using the first method did not correspond with the information collected using the second method. Partnership dissolution and "turbulence" in the partnership biography are strong predictors for discrepancies in the information gathered through the two different survey methods. We conclude by drawing attention to the limitations of the retrospective collection of partnership histories at a time when divorce and separation rates are increasing.

Keywords: Family change, recall bias, retrospective surveys, partnership status at childbirth

1 Introduction[1]

In life course analysis, researchers regularly use retrospective surveys to study social and demographic behavior. In this type of survey, central life events—like the birth of a child, the termination of a partnership, or a residential move—are reconstructed based on the memories of the respondents. During the interview, the respondents are asked to report the start and end dates of certain states or activities that they have experienced over their life course, typically using life history calendars (Freedman, Thornton, Camburn, Alwin, & Young-DeMarco, 1988). Because many statistical methods, like event history modeling or sequence analysis, require monthly data, respondents are regularly asked to recall the year and the month of life course events. This level of precision enables researchers to reconstruct detailed life histories, and to establish links between the timing of events in various domains of the life course.

Recall bias is one of the key problems that retrospective surveys have to grapple with (Beckett, Da Vanzo, Sastry, Panis, & Peterson, 2001; Bradburn, Rips, & Shevell, 1987; Dex, 1995; Groves et al., 2009). Recall bias may, however, vary across different domains of the life course. More salient events, like the birth of a child, are easier to remember than less significant events, like the start of a new job. More unpleasant episodes, like spells of unemployment, may not be reported at all, because they are forgotten or concealed from the interviewer (Jürges, 2007). If life histories are recorded with different degrees of precision across different domains of the life course, these discrepancies will affect investigations of the timing and sequencing of life course events. While this type of bias is of general concern for any investigation based on retrospective data, it is particularly relevant for family research. Whether people cohabit before they marry, whether they have their first child before they enter the labor market, and whether they leave the parental home before they have their first partner are among the classic research questions that help us to understand family behavior in

contemporary societies (Billari, 2001; Elzinga & Liefbroer, 2007). In order to answer these questions, we need to have reliable information on the timing of events across the life course, especially because family events often take place within a narrow time frame. So far, however, there have been very few attempts to identify the biases that may lead to imprecision in the recording of events in different domains of the family life course.

In this paper, we seek to fill parts of this research gap by focusing on two methods used in the German Family Panel (pairfam) to collect information on the *partnership status at first birth*. In the first method, the information comes from fertility and partnership biographies, which were collected separately. In the second method, information was gathered using a "landmark question," in which the respondents were asked to report whether they were married, cohabiting, or partnered when their first child was born. Providing temporal landmarks during an interview situation is commonly believed to generate more reliable information (Dhum, 1998; Gaskell, Wright, & O'Muircheartaigh, 2000). Unfortunately, we are unable to evaluate whether this method indeed provides more trustworthy results, as we lack the "gold standard" for describing real behavior (Beckett et al., 2001, p. 598). Nevertheless, we are able to compare the results from both methods which enables us to get a firm understanding to what extend they are sensitive to the two different modes of survey.

2 Theoretical Considerations

2.1 Retrospective Surveys and Family Research

Vital statistics data from across Europe show that, in recent decades, marriage rates have been declining, and the shares of unmarried mothers have been increasing.[2] These trends have fuelled ongoing debates about the significance of marriage as an institution in general, and as

a setting f or raising children in particular (Cherlin, 2004; Goldstein & Kenney, 2001). Among the questions that have been raised in response to these developments are: Does the growing share of births to unwed parents indicate a retreat from marriage? Do these trends show that lone parenthood is on the rise, or that cohabiting unions are becoming increasingly common? Is marriage postponed to later phases of the life course, or is it forgone entirely? And, how do marriage and cohabitation patterns vary across the different strata of a society?

Retrospective family surveys, like the Family and Fertility Survey or the Generations and Gender Survey, have led the way in providing answers to these types of research questions (Heuveline & Timberlake, 2004; Perelli-Harris et al., 2012; Perelli-Harris et al., 2010; Sobotka & Toulemon, 2008). These surveys contain partnership histories which enable researchers to distinguish periods of cohabitation from spells of lone parenthood. The German Family Panel (pairfam), which is used in this investigation, also includes episodes of "living apart together" (LAT), and thus allows for even more subtle distinctions to be made between partnership forms across the life course (Bastin, Kreyenfeld, & Schnor, 2012; Huinink, Kreyenfeld, & Trappe, 2012).[3] In the German Family Panel, as well as in the abovementioned surveys, partnership and fertility histories are gathered in a modularized way. This means that the fertility and partnership histories are collected separately in different (but adjacent) sections of the questionnaire. Landmark surveying, in which, for example, dates of childbirths are provided during the collection of the partnership histories, is not applied.

2.2   Recall Bias in Retrospective Surveys

Retrospective surveys crucially rely on the ability of respondents to recall when various life course events took place. When gathering fertility histories, it generally suffices for people to remember the dates of birth of their children.[4] In order to reconstruct partnership histories, however, respondents must provide the dates of their marriages, the dates when they moved in

with each of their partners, and the dates when each of these households was dissolved. For the LAT episodes, respondents have to recall the start and end dates of each partnership. Obviously, a respondent's ability to retrieve this information from his or her memory will vary greatly depending on the type of partnership, the time that has elapsed since the relationship began, and whether the partnership continues up to the present or has been dissolved.

Recall bias, or the inability to provide accurate and complete information during an interview, can occur for a variety of reasons. *Salience* is generally agreed to be an important determinant of whether the information provided is reliable. Significant life course events, like the birth of a child, can be surveyed with a fair degree of precision, as most people are able to recall accurately how many biological children they have, and the dates when these children were born.[5] Similarly, the dates of marriage or of the death of a partner are not subject to a high degree of recall bias. Apart from the personal and emotional salience of events, the *regularity* with which respondents are asked to recollect events has been found to influence recall bias. Dates of marriage and of the birth of children tend to be swiftly remembered because people are asked to provide this information routinely during administrative processes, and because anniversaries and birthdays are routinely celebrated. This "process of memory rehearsal is thought to strengthen the memory trace and thus increase the ease of recalling an event" (Beckett et al., 2001, p. 595). *Elapsed time* is another factor that is associated with recall bias (MacDermid, 1989). Time might have erased the entire union from the memory of the respondent, or at least dimmed the precise start and end dates of the union. People may also fail to report a past partnership because more salient events have crowded out the memory of the relationship. Short unions are often disregarded if respondents have been in several partnerships, cohabitations, or marriages over their life course (Mitchell, 2012). Moreover, there is strong evidence that separation leads cohabiting respondents to redefine their

partnership history, with disrupted unions often being unmentioned (Teitler, Reichman, & Koball, 2006).

Cognitive psychology tells us that the ability of people to recall past events may be improved if links between different domains of the life course and significant events are established during the interview (Beckett et al., 2001, p. 595; Matthes et al., 2005; Reimer, 2004, pp. 18-20). In this context, researchers have posited that autobiographical memory is structured as a "hierarchical network" consisting of three main levels: knowledge of life time periods, knowledge of general events, and event-specific knowledge (Belli, 1998; Conway & Pleydell-Pearce, 2000). On the first level, life time periods are stored as broad life time phases with "identifiable," but still "fuzzy" start and end dates (Conway & Pleydell-Pearce, 2000, p. 262). These are conceptualized by stretches in a person's life in which he or she has, for example, been in school, living in town x, employed at company y, or living with partner z. Although life time periods are essential components that structure the autobiographical memory, they are not necessarily memorized in connection with precise dates. General events are stored on the second level. Unlike life time periods, they are connected to concrete dates. This is particularly the case for landmark events, like the birth of a child or a marriage. More specific events (like the first kiss) are stored on the third level. In order to construct autobiographical memory, information from all of these levels must be retrieved and connected. The more "temporal and thematic cues" that are given during the retrieval process, the better and more accurate the biographical information provided is likely to be (Belli, 1998, p. 385). We can therefore assume that errors in retrospective surveying may be minimized if "cues" are offered that assist respondents in establishing links between different domains of the life course, and in constructing their biographical memories (Beckett et al., 2001, p. 600; Matthes et al., 2005, pp. 8, 11). Belli (1998, p. 394) has argued in this context that because survey questionnaires "typically move from topic to topic, respondents are encouraged to segment their paths into units that are largely unrelated to one another." If, however, landmark events

are provided in the interview situation, or links between different domains of the life course are established, the ability of respondents to retrieve information during the interview may be expected to increase (Glasner, Vaart, & Belli, 2012; Teitler et al., 2006).

In this paper, we try to shed more light on the ability of survey respondents to recall biographical information. We do so by contrasting two methods used to collect information on partnership status at first birth. In the first method, a respondent's partnership and fertility information were collected separately. In the second method, each respondent was directly asked a "landmark question" about his or her partnership status at first childbirth.

3   Data Source

3.1   The German Family Panel (pairfam)

The data used in this investigation came from the German Family Panel (pairfam).[6] The German Family Panel is an annual panel survey. The first wave was conducted in 2008/2009 and included about 12,000 respondents from the birth cohorts 1971-73, 1981-83, and 1991-93. In 2009/2010, an eastern German subsample was added to the data, which included an additional 1,500 respondents from eastern Germany of the cohorts 1971-73 and 1981-83 (Huinink et al., 2012; Kreyenfeld, Huinink, Trappe, & Walke, 2012). An advantage of the German Family Panel is that it collects detailed partnership and fertility histories for each respondent. These histories contain the start and end dates of all of a respondent's partnerships, including any living apart together relationships (LAT). In the third wave of the German Family Panel, an additional "landmark" question asked a subsample of respondents a direct question about their partnership status at first childbirth.[7]

In our investigation we used data from the survey round 2010/11 (wave 3), in which this additional "landmark" question was posed. Since only the eastern German subsample was

asked this question, we had to restrict our investigation to this group of respondents. The original sample included 1,173 respondents. For our investigation, we selected men and women who reported having given birth to or fathered a child prior to the date of interview, which narrowed the sample to 713 respondents. Of these 713 respondents, 17 provided incomplete information on their partnership status at childbirth. They have been retained in the descriptive statistics, but omitted from the multivariate analysis.

3.2 Variables

3.2.1 Partnership Status at Childbirth

Based on our sample, we generated a variable that gives the partnership status at first childbirth. This variable distinguishes the following states:

(1)   Married: The respondent was in a marital union.

(2)   Cohabiting: The respondent was unmarried, but was living with a partner.

(3)   LAT (living apart together): The respondent was unmarried and partnered, but was not sharing a household with the partner.

(4)   Single: The respondent had no partner.

In order to generate this variable, we used the two types of survey methods, which we refer to in the following as Method I and Method II.

Method 1: Independent Collection of Partnership and Fertility Biographies

In Method I, we drew upon the separately gathered partnership and fertility histories to generate the partnership status at first childbirth. The fertility and partnership histories were surveyed in the first wave of the study, and were updated every year thereafter based on a

computerized event history calendar (EHC). In a first step, the partnership history was recorded by the interviewer. Respondents were asked to provide the names of all of their partners, and the start and end dates of each partnership (beginning with the most recent one). The interviewer instructed the respondents to list all partnerships after age 14 which lasted at least six months, or which were significant for other reasons (i.e., because the birth of a child resulted from this partnership). Respondents could report gaps in the partnership as well as overlaps with other partnerships. Further information (like episodes of cohabitation and dates of marriage, separation, and divorce) were nested within each partnership. After the partnership histories had been completed, the fertility histories were surveyed. For each child, the gender, the year and month of birth, the past cohabitation with the parents, and the relationship to the anchor respondent (biological, step-, or adopted child) were collected. Furthermore, the names of the second biological parent of these children were reported, which made it possible to link the children to the partners named in the partnership history. An additional question asked the respondent whether she or he had been in a "serious" relationship with the second biological parent of the child. These fertility and partnership histories provided during the first wave were "pre-loaded" in the subsequent wave. This means that the past partnership history was shown on the computer screen, and the respondent was then asked to update the history to include any changes that had occurred since the last interview.

Based on this information, we generated complete fertility and partnership histories, which we merged into a single file (Schnor & Bastin, 2012). To generate the fertility histories, only small modifications of the original data were needed. One modification concerned the selection of biological children and the imputation of missing information on the month of the first birth, which was imprecise for 10 cases because only the year of birth of the child or the season of birth was reported. In the partnership histories, more significant imputation was required. For LAT-partnerships that "surrounded" the period of the first birth, about 20

percent of the start or end dates needed to be imputed. For cohabitation, this amounted to about 17 percent and for marriage less than 5 percent. If information was missing, we relied on a random number generator to impute missing or inaccurate time information. A decision also had to be made regarding "tied events." For example, we had to decide whether a birth was marital or non-marital if the marriage and the childbirth occurred in the same month. In this case we assumed that the marriage or the start of a marriage or a new union always preceded the childbirth.

Method II: Temporal Landmarks and Partnership Status at Childbirth

Our second source of information consisted of answers to a landmark question (Method II) that asked respondents directly to report their family, cohabitation, and partnership status at first birth. More specifically, the following questions were included in the questionnaire:[8]

- When your first child was born, did you have a partner?

  Response categories: yes, no, refusal

- When your first child was born, did you cohabit with a partner?

  Response categories: yes, no, refusal

- What was your family status when your first child was born?

  Response categories: single, married, widowed, refusal

Compared to Method I, we may expect to find that Method II generated less recall bias. Respondents were asked about the partnership status in direct reference to the landmark event of the birth of the first child, which should have helped them recall their partnership status at the time they started their family (Dhum, 1998). While the degree of recall bias may indeed have been lower than it was in Method I, this does not mean that the information collected was valid and reliable. Bias may have occurred because of the way the questions were

phrased, as respondents may have been tempted to draw upon response sets (Teitler et al., 2006). As divorce or separation are generally seen as undesirable and socially less acceptable events, respondents may have a tendency to hide such information during an interview situation. Conversely, people who were unmarried when they had a child might report that they were married if they believe that unwed parenthood is stigmatized. The respondents in our investigation were relatively young, and we may assume that unmarried parenthood was associated with less stigma for them than it would have been for older respondents. Nevertheless, we cannot rule out the possibility that the response sets invalidate some of our conclusions in our comparison of the two methods.[9]

3.2.2 Research Strategy and Control Variables

Our research strategy is to compare the responses from Method I and Method II. In a first step, we generate descriptive tables that provide an overview of the discrepancies between the two methods. In a second step, we employ binary logistic regression models designed to help us understand how socioeconomic background influences the response patterns. The final step contains a sequence analysis that tries to shed light on the question of whether the discrepancies between the information generated by the different methods could be explained by the turbulence in the partnership transitions that occurred around the first birth.

In the multivariate analysis, we account for several factors which may have affected recall bias (see Table A1 in the appendix for the descriptive statistics). The most obvious of these factors is elapsed time. We account for this factor by controlling for the *age of the first child* as a continuous variable. The other control variables are *gender, education*, and *citizenship*, which have been shown in previous investigations to influence recall bias (Auriat, 1993; Coughlin, 1990; Mitchell, 2012; Thompson, Herrmann, Bruce, Read, & Payne, 1998). We assume that turbulence in the partnership biography is the main factor which affects the

reliability of the information gathered from retrospective surveys. We control for this factor by accounting for the *number of partnership disruptions* after the first birth. We also control for the total *number of biological children* and the number of *non-biological children*. The latter group mostly includes respondents with stepchildren, and thus also signifies turbulence in the family biography. A binary variable controls for whether the *date information was imputed* in the birth or partnership history, based on the assumption that some of the differences between the two methods are attributable to the inability of respondents to correctly specify the start and end dates of their partnerships.

A variable that is particularly relevant for our investigation is the *partnership status at first birth*. This variable is assumed to tell us which partnership types would be associated with the greatest discrepancies in the information generated by the two methods. However, the variables for the number of separations and the partnership status at childbirth are closely related. Only respondents who have a partner are able to separate. To avoid multicollinearity between the two variables, we estimate a second model that does not include the number of separations, but only the partnership status at the first birth (according to the landmark question).

4   Results

4.1   *Descriptive Analysis*

Tables 1a and 1b provide a descriptive account of the differences between the two methods of collecting information on the partnership status at first childbirth. Table 1a gives the absolute counts. The diagonal, gray-shaded cells show the cases for which the two methods generated the same results. Out of our sample of 713 observations, there is a match between the two methods in 579 cases. Thus, in 19 percent of all of the observations, the results from Method I

and from Method II conflict. Table 1a also tells us about the distribution of missing values. For the landmark question, only one respondent refused to provide an answer. For the biographical questions, there is a much higher share of missing cases, of two percent.

Table 2b provides the column percent. Assuming the answer to the landmark question is more likely to have been reliabel, the table suggests that discrepancies were least prevalent among respondents in marital unions. In 89 percent of these cases, the two methods provide the same results. For respondents in non-marital unions, the two methods generate the same results in 82 percent of the cases. Discrepancies are most frequently found among respondents who reported single births and births in living apart together relationships (LAT). For the respondents in LATs, there are discrepancies in 54 percent of the cases. For the singles, conflicting information is found in 56 percent of all of the cases. In most of these cases, the discrepancies are due to a mix-up of non-marital cohabitations, singlehood, and living apart together relationships.

Table 1a: Partnership status according to landmark question and according to biographies, absolute values

		Partnership status according to landmark question					
		Single	LAT	Cohab.	Married	Missing	Total
Partnership status according to biography	Single	24	6	16	3	-	49
	LAT	13	20	35	3	-	71
	Cohabiting	4	8	296	21	1	330
	Married	-	2	6	239	-	247
	Missing	2	1	10	3	-	16
	Total	43	37	363	269	1	713

Table 1b: Partnership status according to landmark question and according to biographies, column percent

		Partnership status according to landmark question					
		Single	LAT	Cohab.	Married	Missing	Total
Partnership status according to biography	Single	56%	16%	4%	1%	-	7%
	LAT	30%	54%	10%	1%	-	10%
	Cohabiting	9%	22%	82%	8%	100%	46%
	Married	-	5%	2%	89%	-	35%
	Missing	5%	3%	3%	1%	-	2%
	Total	100%	100%	100%	100%	100%	100%

4.2 Multivariate Results

Table 2 reports the results from the binary logistic regression. The dependent variable equals one if the information of the two methods are in conflict, and zero otherwise. We first estimate a model that contains the major control variables, including the number of separations. Model 2 does not include the number of separations, but it takes into account the partnership status at first birth (according to the landmark question).

The results of Model 1 show that gender has a strong and significant influence on the probability that information from Method I and Method II conflict. Men appear to be more likely to have provided inaccurate information. The odds of a mismatch occurring are more than 40 percent higher among male respondents. Surprisingly, we do not find that the amount of time that has elapsed is significant in this multivariate model. It should be noted, however,

that some of the control variables, such as the number of children as well as the number of separations, are closely related to elapsed time. After these aspects are accounted for, we see no significant results for elapsed time in the multivariate investigation.[10] We neither find that the level of education influences people's ability to provide matching information on their partnership status at childbirth in the two methods. This is at odds with prior findings, which showed that more highly educated individuals were better able to provide reliable information in retrospective surveys (e.g. Coughlin, 1990, p. 88).

Our results further show that citizenship strongly contributes to discrepancies in the information generated by the different survey methods. Foreigners are substantially more likely than Germans to have given inconsistent responses. This may be attributed to a lack of German proficiency among the foreigners, and to an inability to understand the meaning of the complex biographical or landmark survey questions. It may also be explained by cultural response sets. The migrants in our sample mostly came from countries where unwed parenthood is uncommon and is less socially accepted. The landmark question directly asked respondents whether they had a birth outside of marriage, which may have led some respondents to draw upon response sets and provide a more socially acceptable answer.

We also find that respondents with three or more (biological) children are more likely to have given inaccurate answers than parents with only one child. In this case, we might assume that as the number of children increases, people's ability to correctly remember the birth dates of each child declines. Given, however, that childbirth is a rather salient event, this explanation seems a little far-fetched. An interaction of the number of children and the respondent's gender (see the appendix) shows that this problem is mainly found among male respondents, which suggests that men have more difficulties than women in correctly recalling the birth date of their first-born child and/or their partnership situation at the time of the first birth. The

results may also be affected by the fact that some men with larger families live in higher order unions, and no longer have close contact with their first-born children.

The degree of turbulence in the partnership history, measured by the number of separations since the first birth, seems to be strongly related to recall bias. For respondents who had experienced more than one disruption since the first birth, we find that the odds of a mismatch are 65 percent higher relative to the reference group of respondents who did not experience a disruption of their unions. For respondents with two or more disruptions, the odds are 170 percent higher. The number of non-biological children, which is also indicative of a turbulent family biography, is closely related to discrepancies in the information provided by the two methods. Surprisingly, the variable that indicates the imputation of date information in the biographies is insignificant.

Model 2 shows that the partnership status at the first birth is strongly related to a mismatch. Compared to married respondents, individuals who were in LAT-relationships or who were single are much more likely to have discrepancies in the information provided (in response to the landmark question). The very high odds ratio of more than five for single respondents and more than eight for respondents in LAT-partnerships are most likely indicative of the inability of retrospective surveys to correctly capture the start and end dates of less established relationships like LAT-partnerships.

Table 2: Results from logistic regression. Determinants of mismatch (1) versus match (0) between Method I and Method II. Odds ratios

	Model 1		Model 2	
Gender				
Female	1.48	*	1.64	**
Male	1		1	
Age of first child (continuous)	1.02		1.03	
Level of education				
Low	1		1	
Medium	1.18		1.30	
High	0.95		1.14	
Citizenship				
German	1		1	
Other	2.77	***	3.45	***
Number of children				
One child	1		1	
Two children	0.82		0.87	
Three children	1.77	*	1.94	**
Number of separations				
None	1		--	
One	1.68	*	--	
Two and more	2.79	***	--	
Non-biological children				
No	1		1	
Yes	2.02	*	1.73	
Imputation of dates				
No imputation	1		1	
Imputation	0.92		1.15	
Partnership status at childbirth				
Single	--		5.82	***
LAT	--		8.26	***
Cohabiting	--		1.93	***
Married	--		1	
Constant	0.09	***	0.04	***
N	695		695	

Note: * p<0.1; ** p<0.5; *** p<0.01

Source: German Family Panel (pairfam) wave 3, own estimates

4.3 Sequence Analysis

Our analysis so far has focused only on the time of the first birth. The birth of the first child is a significant transition in a person's life course. This transition is regularly preceded by other major transitions in the partnership domain of the life course. Normative pressures to get married prior to having a child may have weakened in recent decades. Nevertheless, there are many economic as well as social reasons to institutionalize a partnership in anticipation of family formation. Thus, we expected to find that the union transitions in the partnership domain of the life course accelerate around the time of the first birth. For the investigation of family behavior with retrospective surveys, this may have important implications. If there are small inaccuracies in how people remember the start and end dates of their partnerships, this may produce large differences in the partnership status that we measure at first birth. In order to rule out the possibility that our prior findings were produced solely by the fact that we were focusing on a very particular point in time (namely, the first birth) we now turn to sequence analysis, which is a more explorative method for mapping birth and fertility histories across time. Unlike the logistic regression model, which focused on one point in time only, sequence analysis maps the entire life course, or (as in our case) a relevant section of it.

Sequence analysis has its origins in biology, where it was first used to analyze protein and DNA sequences (Abbott & Tsay, 2000). It has since been applied in social science research, such as in the analysis of employment careers (Brzinsky-Fay, 2007; Fasang, 2012) and of partnership and fertility trajectories (Berghammer, 2012; Elzinga & Liefbroer, 2007; Fasang, in press). The method of sequence analysis is a collection of techniques used to describe sequential data which in the social sciences are usually arranged in person-month units (Abbott, 1995). In this paper, we use sequence index plots which arrange the monthly partnership biographies for each respondent in horizontal bars.[11] The analysis includes individuals for whom the information from Method I and from Method II conflicted. In

addition, we restrict the observation period to the time period around the first birth; i.e., the 12 months before and after the first birth. The sequence index plots are shown by partnership status at the first childbirth according to the response to the landmark question. The figures give us a visual impression of the degree of turbulence of the partnership biography around the first birth. This allows us to evaluate whether the information generated by the two methods conflict because the response to the landmark question captures the respondent's partnership status at a single point in time only, even though transitions to other states may have occurred in the months surrounding the first birth.

We begin by looking at the respondents who reported having been single when they were asked directly about their partnership status at the first birth (Panel 1 in Figure 1), but who reported having another status in their biography. The figure shows that the overwhelming majority of these women were in an LAT-partnership, according to their biography. More than three-quarters of these women experienced partnership formation or dissolution in the year before or after the first birth. This result supports the notion that partnership transitions accelerate around the time of the first birth. It also suggests that respondents may have had difficulties in clearly differentiating between LAT and being single.

Panel 2 in Figure 1 provides information for respondents who were in an LAT-arrangement according to their answer to the landmark question, but who had another status according to their biography. In this case as well we can see the difficulties the respondents had in distinguishing singlehood from LAT-relationships. This finding also shows the difficulties the respondents experienced in distinguishing between LAT and cohabiting. About half of the respondents who reported in their answer to the landmark question that they were in an LAT-partnership at the time of the first birth, but had another status according to their biography, were cohabiting according to the biography. In many cases, the respondents reported having had an LAT-period just before the first birth.

Panel 3 of Figure 3 also supports the notion that periods of LAT and cohabiting were difficult to distinguish for the respondents. More than half of all of the respondents who were in a non-marital union according to their answer to the landmark question, but had another status according to their biography, were in an LAT-relationship according to their biography. In about half of these cases, we observe episodes of cohabitation after the first birth. It is, however, noteworthy that around one-quarter of the respondents were classified single over the entire period. These respondents may have erased an unsuccessful union from their partnership biography.
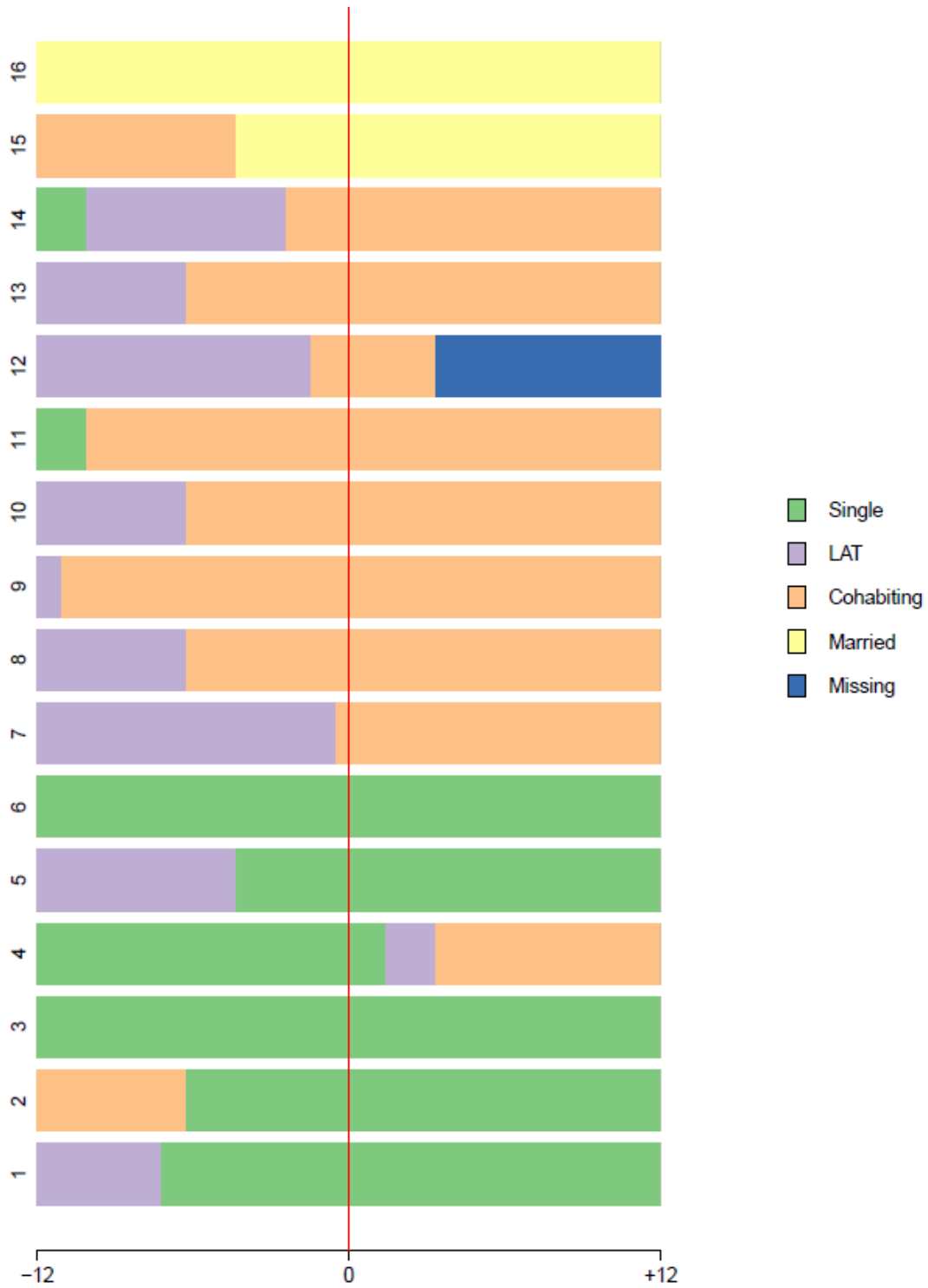
The last figure (Panel 4) gives the sequence index plots for respondents who reported having been married in their answer to the landmark question, but reported having been in other states according to the biographies. The overwhelming majority of discrepancies found in the information generated by the two methods were due to respondents who were living in non-marital cohabitations at the time of the first birth according to their biography. About one-third of them married within a year after the first child was born (according to the biographical information). However, for the marital unions, it is much more difficult to discern a consistent pattern.

On the whole, the sequence index plots support our assumption that family transitions tend to accelerate around the time of the first birth. This can lead respondents to make inaccurate statements about their partnership status at the first birth, often because they do not have a clear recollection of the start and end dates of more loose relationships, like LAT-partnerships. The transition to an LAT-partnership and to cohabitation is not a qualitative change, but is in most cases a transition period during which the partnership is being established, which can make it difficult for respondents to draw a clear line between being single, LAT, and cohabiting.
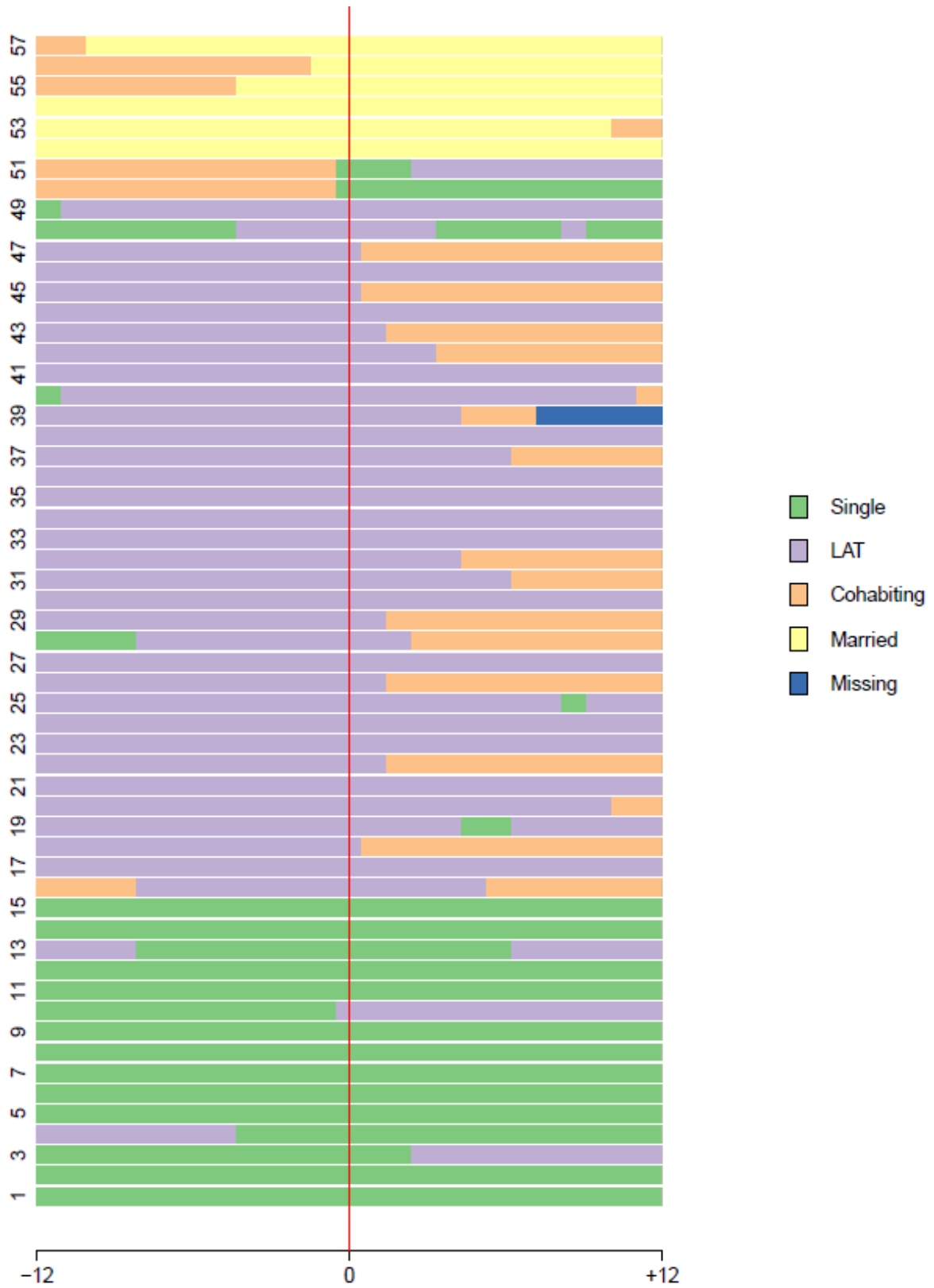
Figure 1: Sequence index plots by partnership status at birth according to the landmark question response and the age of the first child (in months), only women with conflicting answers. Red line: Date of first birth

Panel 1: Single (according to the response to the landmark question)
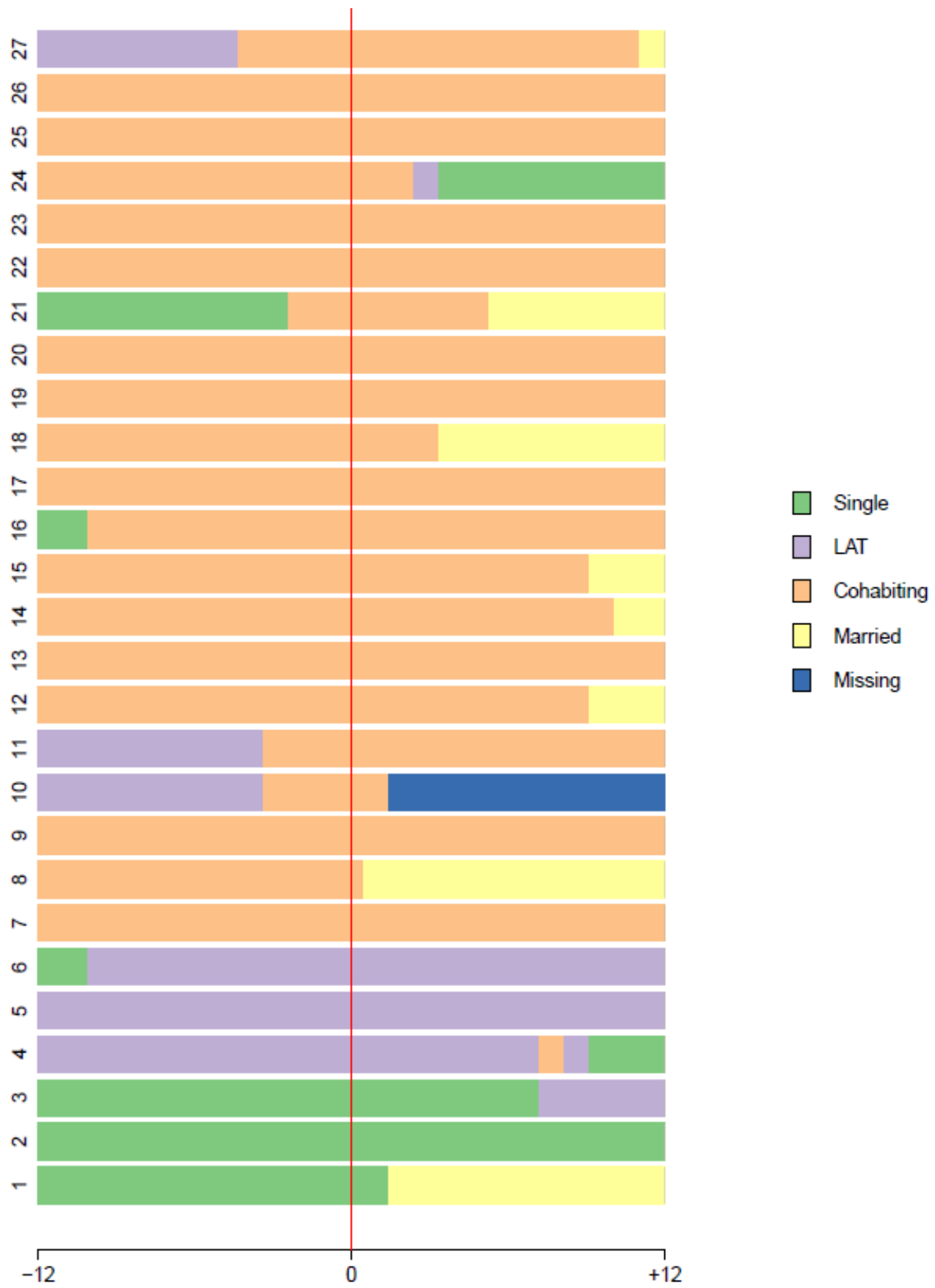
Panel 2: LAT (according to the response to the landmark question)

Panel 3: Cohabiting (according to the response to the landmark question)

Panel 4: Married (according to the response to the landmark question)

5 Conclusion

In recent decades, most European countries have witnessed large increases in non-marital childbearing and the prevalence of "non-standard" family forms. Family surveys have been especially important in this context, enabling us to expand our understanding of the changes in the partnership and family domain of the life course. In these surveys, partnership and fertility histories are often gathered based on retrospectively collected information on the dates of childbirth and the start and end dates of partnerships, cohabitations, and marriages. Although it is generally acknowledged that different types of events are collected with different levels of precision, it is unclear whether the lack of precision also biases our investigations of family change. This paper has explored this issue by contrasting two methods used to survey partnership status at first birth. In the first method, fertility and partnership histories were collected independently; while in the second method, a landmark question asked respondents directly about the partnership status they had when their first child was born.

The descriptive analysis has shown that in almost 20 percent of the cases, the results from the two methods conflicted. We were unable to evaluate which of the two methods produced more accurate data, as we lack register-based evidence to compare our findings to. While recall bias might have affected the quality of the information we gathered from the biographies, response sets might have biased the results from the landmark question. Nevertheless, our results provide us with a clear pattern of the factors that dilute people's biographical memory, and are thus obvious sources of recall bias in retrospective family surveys.

In line with prior research, our results showed that the family biographies of male respondents are less trustworthy than those of female respondents (Sorensen, 1997). We also found that large discrepancies in the information generated by the two survey methods were especially

frequent among respondents with non-German citizenship. This finding is of particular relevance for studies that compare partnership dynamics by ethnic origin (Phillips & Sweeney, 2005). One of the important findings of our investigation is that turbulence in the partnership biography can affect a person's ability to provide consistent biographical information. Respondents who had experienced multiple separations since the first birth often provided inconsistent information on their partnership status at first birth. In addition to the number of separations, stepfamily membership (measured by whether a person has non-biological children) was shown to have influenced the consistency of biographical information.

In this paper, we also used sequence analysis in order to gain an understanding of how partnership transitions around the time of the first birth might have led to discrepancies. Our sequence index plots showed that family transitions accelerate around the time of the first birth. In many cases, the respondent's family status according to his or her response to the landmark question matched the family status we measured two or three months before or after the childbirth in the respondent's biography. This was particularly true for singlehood and LAT-partnerships, but also for LAT and cohabiting. From our investigation, we are unable to tell whether discrepancies between the answers to the landmark question and the biographies were due to the deliberate withholding of information, recall error, or difficulties in distinguishing between the start and the end dates of loose relationships like LAT- and cohabiting unions. However, the results strongly suggest that the inability of respondents to draw a clear line between singlehood, LAT-partnership, and cohabitation is part of the problem.

The findings from our paper have implications for the analysis of family change. Most importantly, simple summary measures that try to map the partnership status at a particular moment in time, like at the time of the first childbirth, are highly misleading. The months around the first birth constitute a period during which major transitions in the partnership

domain of the life course accelerate. For some couples, the pregnancy marks the moment in the relationship that initiates the transition from a cohabiting union into a marriage. For other couples, having a child means that they need to start cohabiting instead of having an LAT-relationship. Still other couples might separate in response to an unplanned pregnancy. The ability of respondents to recollect precisely the start and end dates of looser relationships, like LAT-partnerships and cohabitations, is limited. Thus, their ability to report their exact living arrangement at a particular point in time may also be limited. Adopting a longitudinal perspective that captures life course periods instead of particular points in time could improve our understanding of family behavior. However, even taking a longitudinal view of the question will not alleviate the difficulties that arise because people are unable to specify the correct start and end dates of their past unions. Statistical methods, like event history or sequence analysis, generally rely on the notion that we have clearly defined states that are unambiguously situated in time. The growing importance of LAT-partnerships and cohabiting unions, with their fuzzy start and end dates, create an obvious challenge for researchers using these kinds of techniques.

6   Acknowledgment

7 References

Abbott, A. (1995). Sequence analysis: New methods for old ideas. *Annual Review of Sociology, 21* 93-113.

Abbott, A., & Tsay, A. (2000). Sequence analysis and optimal matching methods in Sociology. *Sociological Methods and Research, 29*(1), 3-33.

Auriat, N. (1993). "My wife knows best" - A comparison of event dating accuracy between the wife, the husband, the couple and the Belgium population register. *Public Opinion Quarterly 57*, 165-190.

Bastin, S., Kreyenfeld, M., & Schnor, C. (2012). Diversität der Familienformen in Ost- und Westdeutschland [Diversity of family forms in Eastern and Western Germany]. In D. Krüger, H. Herma & A. Schierbaum (Eds.), *Familie(n) heute – Entwicklungen, Kontroversen, Prognosen* (pp. 126-145). Wiesbaden: Juventa.

Beckett, M., Da Vanzo, J., Sastry, N., Panis, C., & Peterson, C. (2001). The quality of retrospective data: An examination of long-term recall in a developing country. *The Journal of Human Resources 36*(3), 593-625

Belli, R. F. (1998). The structure of autobiographical memory and the event history calendar: Potential improvements in the quality of retrospective reports in surveys. *Memory 6* 383-406.

Berghammer, C. (2012). Family life trajectories and religiosity in Austria. *European Sociological Review, 28*(1), 127–144.

Billari, F. C. (2001). The analysis of early life courses: Complex descriptions of the transition to adulthood. *Journal of Population Research 18*(2), 119-142.

Bradburn, N. M., Rips, L. J., & Shevell, S. K. (1987). Answering autobiographical questions: The impact of memory and inference on surveys. *Science*, 208-216.

Brzinsky-Fay, C. (2007). Lost in transition? Labour market entry sequences of school leavers in Europe. *European Sociological Review, 23*(4), 409-422.

Cherlin, A. J. (2004). The deinstitutionalization of American marriage. *Journal of Marriage and the Family, 66*(4), 848-861.

Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review 107* 261-288.

Coughlin, S. S. (1990). Recall bias in epidemiologic studies. *Journal of Clinical Epidemiology, 43*(1), 87-91.

Dex, S. (1995). The reliability of recall data: A literature review. *Bulletin of Sociological Methodology, 49*(1), 58-89.

Dhum, M. S. (1998). The role of temporal landmarks in autobiographical memory processes. *Psychological Bulletin, 124*(3), 423-442.

Elzinga, C. H., & Liefbroer, A. C. (2007). De-standardization of family-life trajectories of young adults: A cross-national comparison using sequence analysis. *European Journal of Population, 23*(3-4), 225-250.

Fasang, A. E. (2012). Retirement patterns and income inequality. *Social Forces, 90*(3), 685-711.

Fasang, A. E. (in press). New perspectives on family formation: What can we learn from sequence analysis? In P. Blanchard, F. Bühlmann & J.-A. Gauthier (Eds.), *Advances in Sequence Analysis: Methods, Theories and Applications*: Springer: Series Life Course Research and Social Policies.

Freedman, D., Thornton, A., Camburn, D., Alwin, D., & Young-DeMarco, L. (1988). The life history calendar: A technique for collecting retrospective data. *Sociological Methodology, 18*, 37-68.

Gabadinho, A., Ritschard, G., Studer, M., & Müller, N. S. (2011). *Mining Sequence Data in R with the TraMineR Package: A User's Guide*. University of Geneva: http://mephisto.unige.ch/traminer/.

Gaskell, G. D., Wright, D. B., & O'Muircheartaigh, C. A. (2000). Telescoping of landmark events. Implications for survey research. *Public Opinion Quarterly 64*(1), 77-89.

Glasner, T., Vaart, W. v. d., & Belli, R. F. (2012). Calendar interviewing and the use of landmark events − implications for cross-cultural surveys. *Bulletin of Sociological Methodology, 115*(1), 45-52.

Goldstein, J. R., & Kenney, C. T. (2001). Marriage delayed or marriage forgone? New cohort forecasts of first marriage for U.S. women. *American Sociological Review, 66*(4), 506-519.

Groves, R. M., Jr., F. J. F., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey Methodology* (2nd edition ed.). Hoboken, New Jersey: Wiley.

Heuveline, P., & Timberlake, J. M. (2004). The role of cohabitation in family formation: The United States in comparative perspective. *Journal of Marriage and Family, 66*(5), 1214-1230.

Huinink, J., Kreyenfeld, M., & Trappe, H. (2012). *Familie und Partnerschaft in Ost- und Westdeutschland. Ähnlich und doch immer noch anders. [Family and Partnership in East and West Germany. Alike, but different]*. Leverkusen: Babara Budrich.

Jürges, H. (2007). Unemployment, life satisfaction and retrospective error. *Journal of the Royal Statistical Society. Series A (Statistics in Society), 170*(1), 43-61.

Kreyenfeld, M., Hornung, A., & Kubisch, K. (2012). The German Generations and Gender Survey: Some critical reflections on the validity of fertility histories. *Comparative Population Studies, 38*(1), 3-28.

Kreyenfeld, M., Huinink, J., Trappe, H., & Walke, R. (2012). DemoDiff: A Dataset for the Study of Family Change in Eastern (and Western) Germany. *Schmollers Jahrbuch, 132*(4), 653-660.

MacDermid, R. H. (1989). The recall of past partisanship: Feeble memories or frail concepts? *Canadian Journal of Political Science, 22*(2), 363-375.

Matthes, B., Reimer, M., & Künster, R. (2005). *TrueTales - Ein neues Instrument zur Erhebung von Längsschnittdaten [TrueTales - A New Instrument for Collecting Longitudinal Data]*. Berlin Max-Planck-Institut für Bildungsforschung.

Mitchell, C. (2012). Are divorce studies trustworthy? The effects of survey nonresponse and response errors. *Journal of Marriage and Family, 72*(4), 893–905

Ní Bhrolcháin, M., Beaujouan, É., & Murphy, M. (2011). Sources of error in reported childlessness in a continuous British household survey. *Population Studies, 65*(3), 305-318.

Perelli-Harris, B., Kreyenfeld, M., Sigle-Rushton, W., Keizer, R., Lappegård, T., Jasilioniene, A., . . . Di Giulio, P. (2012). Changes in union status during the transition to parenthood in eleven European countries, 1970s to early 2000s. *Population Studies 66*(2), 167-182.

Perelli-Harris, B., Sigle-Rushton, W., Kreyenfeld, M., Lappppegård, T., Keizer, R., & Berghammmmer, C. (2010). The educational gradient of childbearing within cohabitation in Europe. *Population and Development Review, 36*(4), 775–801.

Phillips, J. A., & Sweeney, M. M. (2005). Premarital cohabitation and marital disruption among White, Black, and Mexican American women. *Journal of Marriage and Family, 67*(2), 296-314.

Reimer, M. (2004). *Autobiografisches Gedächtnis und retrospektive Datenerhebung: die Rekonstruktion und Validität von Lebensverläufen. [Autobiographical memory and retrospective data collection: A Reconstruction and validation of life courses]*. Berlin: Freie Universität.

Schnor, C., & Bastin, S. (2012). Partners and Babies. Partnership and Fertility Histories of the German Family Panel (pairfam) as an Event History Data Set. *Internal Report. Max Planck Institute for Demographic Research. Rostock.*

Sobotka, T., & Toulemon, L. (2008). Changing family and partnership behaviour. Common trends and persistent diversity across Europe. *Demographic Research, 19*(6), 85-138

Sorensen, E. (1997). A national profile of nonresident fathers and their ability to pay child support. *Journal of Marriage and Family, 59*(4), 785-797.

Teitler, J. O., Reichman, N. E., & Koball, H. (2006). Contemporaneous versus retrospective reports of cohabitation in the fragile families survey. *Journal of Marriage and Family, 68*(2), 469-477.

Thompson, C. P., Herrmann, D. J., Bruce, D., Read, J. D., & Payne, D. G. (1998). *Autobiographical Memory: Theoretical and Applied Perspective* Mahwah, N.J.: Erlbaum.
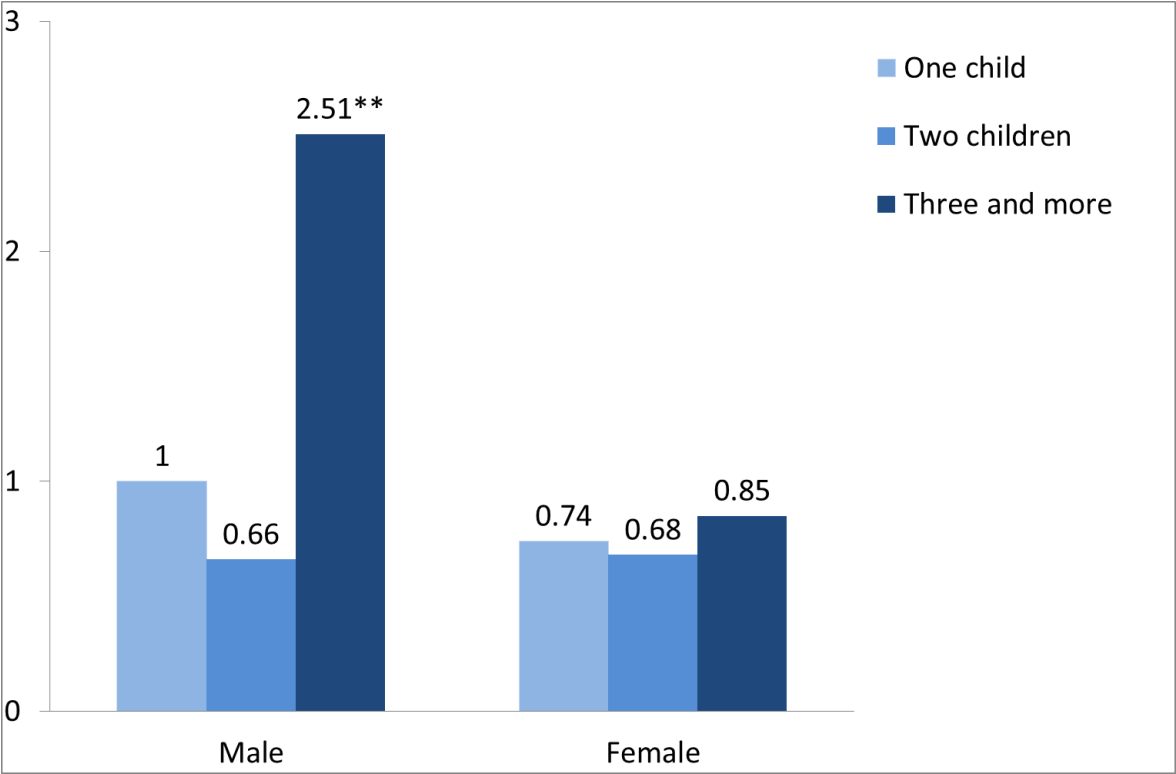
Appendix

Table A1: Composition of sample, column percent

	Number of Cases	Column Percent	Mismatch between Method I & Method II (in %)
Age of first child			
Age 0-2	106	15%	9%
Age 3-5	112	16%	19%
Age 6-9	159	23%	16%
Age 10-14	190	27%	15%
Age 15-32	128	18%	25%
Gender			
Female	270	39%	14%
Male	425	61%	21%
Level of education			
Low	62	9%	19%
Medium	505	73%	17%
High	128	18%	14%
Citizenship			
German	36	5%	31%
Other	659	95%	16%
Number of children			
One child	294	42%	14%
Two children	285	42%	13%
Three children	116	17%	33%
Non-biological children			
No	659	95%	16%
Yes	36	5%	36%
Number of separations			
None	518	75%	13%
One	115	17%	23%
Two and more	62	9%	36%
Imputation of dates			
No imputation	452	65%	16%
Imputation	243	35%	18%
Partnership status at childbirth			
Single	41	6%	42%
LAT	36	5%	44%
Cohabiting	353	51%	16%
Married	265	38%	10%
N	695	100%	17%

Source: German Family Panel (pairfam) wave 3, own estimates

Figure A1: Results from logistic regression. Determinants of a mismatch (1) versus a match (0) between Method I and Method II. Results from the interaction of gender and the number of children. Odds ratios

Note: Further covariates in model are the same as in Table 2 (Model 2)

* p<0.1; ** p<0.5; *** p<0.01

Source: German Family Panel (pairfam) wave 3, own estimates

Endnotes

---

[1]   The term "true tales" in our title quotes from the study by Matthes, Reimer, and Künster (2005).

[2]   See, for example, the databases "Proportion of live births outside marriage" provided by Eurostat http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/themes.

[3]   The GGS includes the start dates of LAT-partnerships that were later transformed into cohabiting unions, as well as the start dates of LAT-partnerships that were ongoing at the time of the interview. However, no full partnership histories were collected.

[4]   Child-related information (such as the childcare usage of younger children and the regional proximity of older children) is frequently collected in conjunction with the gathering of fertility histories. It has been shown that this procedure may generate faulty information in the fertility histories. To shorten the interview, respondents deliberately reported having had fewer children and fewer partnerships (Kreyenfeld, Hornung, & Kubisch, 2012; Ní Bhrolcháin, Beaujouan, & Murphy, 2011).

[5]   Exceptions are non-residential fathers. For the US, it has been shown that surveys do not adequately capture their fertility histories, most likely because they do not report children to whom they have no social contact (Sorensen, 1997). There may also be problems in collecting reliable fertility histories of respondents with deceased children or stillbirths.

[6]   This paper uses data from the German Family Panel pairfam, coordinated by Josef Brüderl, Johannes Huinink, Bernhard Nauck, and Sabine Walper. Pairfam is funded as long-term project by the German Research Foundation (DFG). Analyses are based on data from the eastern German subsample Release 2.0 (doi:10.4232/demodiff.5684.3.0.0).

7   This question was included in the survey because investigations had revealed that there were unusually large shares of women in eastern Germany who were unpartnered when the first child was born. It was thought that this finding could not be taken at face value, but was rather driven by the complexity of the surveys on partnership histories. It was therefore assumed that a simple landmark question that asks respondents about their partnership status at childbirth would provide more reliable results.

8   The wording of the German original is as follows: "Als Sie ihr erstes Kind bekommen haben, waren Sie da mit einem Partner/einer Partnerin zusammen?" (Response categories: „Ja", „Nein", „Keine Angabe"); „Als Sie ihr erstes Kind bekommen haben, haben Sie da mit einem Partner/einer Partnerin zusammen gewohnt?" (Response categories: „Ja", „Nein", „Keine Angabe"); „Als Sie ihr erstes Kind bekommen haben, welchen Familienstand hatten Sie da?" (Response categories: „Ledig", „Verheiratet", „Geschieden", „Verwitwet", „Keine Angabe").

9   The question may also have been ambiguous for respondents who had a partner at the first birth who was not the father or the mother of the child. These respondents must have separated or divorced from the father of the child and re-partnered between the conception and the birth of the first child. The size of this population would have been rather small.

10   The age of the first child was significant if no other variables were accounted for in the model. Please see Table A1 in the appendix for descriptive results.

11   The subsequent analyses were conducted with the R-extension TraMineR (Gabadinho, Ritschard, Studer, & Müller, 2011).