

Max-Planck-Institut für demografische Forschung
Max Planck Institute for Demographic Research
Konrad-Zuse-Strasse 1 · D-18057 Rostock · GERMANY
Tel +49 (0) 3 81 20 81 - 0; Fax +49 (0) 3 81 20 81 - 202;
<http://www.demogr.mpg.de>

MPIDR WORKING PAPER WP 2010-023
APRIL 2010 (REVISED NOVEMBER 2010)

**Fertility and union histories from
German GGS data:
Some critical reflections**

Michaela Kreyenfeld (kreyenfeld@demogr.mpg.de)
Anne Hornung
Karolin Kubisch
Ina Jaschinski

© Copyright is held by the authors.

Working papers of the Max Planck Institute for Demographic Research receive only limited review. Views or opinions expressed in working papers are attributable to the authors and do not necessarily reflect those of the Institute.

Fertility and Union Histories from German GGS Data: Some Critical Reflections

Michaela Kreyenfeld¹, Anne Hornung, Karolin Kubisch and Ina Jaschinski

Abstract

This paper tries to validate the fertility and marriage histories of the German Generations and Gender Survey (GGS). One major result from this validation is that the fertility of the older GGS cohorts is too low, while it is too high for the younger cohorts. For marriage histories, we find a similar bias. In sum, the GGS gives wrong cohort fertility and marriage trends for Germany. We speculate on various sources for this bias in the data.

Keywords: Fertility, data validation, German Generation and Gender Survey

¹ Correspondence address: Max Planck Institute for Demographic Research; Konrad-Zuse-Straße 1; D-18057 Rostock; Email: kreyenfeld@demogr.mpg.de

Das Deutsche Generations and Gender Survey:

Einige kritische Reflektionen zur Validität der Fertilitäts- und Partnerschaftsverläufe

Zusammenfassung

Ziel dieses Beitrags ist es die Geburts- und Partnerschaftsverläufe des deutschen Generations and Gender Survey (GGS) zu validieren. Ein wesentliches Ergebnis dieser Validierung ist, dass die Fertilität der älteren Kohorten über- und die der jüngeren Kohorten untererfasst wird. Auch für das Heiratsverhalten finden wir einen ähnliches Muster der Unter- bzw. Übererfassung. Insgesamt liefert das deutsche GGS damit ein falsches Bild des Kohortenwandels in Deutschland. Verschiedene Ursachen für die Schiefelage werden im Beitrag diskutiert.

Stichwörter: Fertilität, Datenvalidierung, deutsches Generations and Gender Survey

1 Introduction

In recent years, fertility and family researchers have greatly profited from a growing availability of large-scale data sets. This is particularly true for Germany. With the Generations and Gender Survey (GGS), the Panel Study of Intimate Relationships (PAIRFAM), the German Socio-Economic Panel (SOEP) and the National Educational Panel Study (NEPS), Germany will soon have four large panel data sets that can be used for the study of family and fertility dynamics. With growing data availability, it will also become easier to compare findings across the different surveys and to validate the quality of the survey data, i.e. by studying how stable empirical findings are across surveys.

This paper is mainly concerned with the validation of fertility histories. Birth histories are commonly regarded as hard facts that can be collected retrospectively without much decrement in data quality. Even though there might be under-reporting of children of unmarried men (Rendall et al. 1999), the overwhelming majority of respondents will be happy to report whether they are mother or father to a child. Furthermore, the collection of fertility histories should not be subject to substantial recall errors which make the collection of employment careers a cumbersome venture. However, previous research has shown that fertility information derived from survey data might nevertheless be subject to various types of distortions (Murphy 2009). Unit non-response is a particular issue for fertility researchers given that, in standard surveys, childless respondents are more difficult to reach than families with children (Festy and Prioux 2002: 23). Item non-response is a source of additional bias. For older cohorts, selection on survival might also create a problem. If fertility histories are not equally well collected across cohorts, surveys will not be able to mirror cohort change in behavior in a satisfying manner. Since long-term behavioral changes are central themes for social science researchers, it seems, however, fundamental to answer the question how well standard surveys depict cohort changes in fertility behavior. Murphy (2009: 115) concludes rightly in this respect that “there have been few recent evaluations of the quality of fertility histories, despite the crucial importance of the accuracy of such data for demographic research and policy decision-making”.

This contribution tries to fill parts of this gap by validating fertility histories of the first wave of the German Generations and Gender Survey (GGS). The paper is structured as follows. In part 2, we describe how we generated fertility histories from the GGS data. Next follows the validation of the fertility histories of the German GGS. Here we compare patterns that we find for the GGS with results from other surveys and from vital statistics (part 3). The following chapter speculates about possible sources for bias (part 4). Part 5 concludes our investigation.

2 Data, variables and method

2.1 General description of German GGS

The Generations and Gender Survey (GGS) is part of the Generations and Gender Programme (GGP) that has been conceptualized by the Population Activities Unit (PAU) within the United Nations Economic Commission for Europe (UNECE) in Geneva, Switzerland. The data collection is directed at the demographic and social behavior in developed countries and mainly in Europe (Vikat et al. 2008). The aim of the project is to initiate analyses on the developments and the determinants of the diverse demographic and social phenomena related to family formation processes, demographic change, intergenerational relationships and relationships between partners. One of the main focuses within the first wave is to capture the complete fertility and partnership histories of the respondents in a detailed way.

Our investigation is based on the German GGS. For our purpose, we used fertility and partnership histories that were cleaned within the “Nonmarital Childbearing Project” (see Perelli-Harris et al. 2009). The first wave of the German GGS was carried out between February and May 2005. Respondents aged 18-79 were included in the survey. The survey agency TNS Infratest Sozialforschung conducted the survey and a random-route design was applied to draw the sample. Interviews were realized using CAPI (Ruckdeschel et al. 2006).

Sampling weights which particularly account for the fact that certain subpopulations are underrepresented in the GGS were provided by the survey agency (see Ruckdeschel et al. 2006: 22). Weights were estimated based on standard socio-economic information (such as age, education, region and sex) generated from the

German microcensus and from population statistics. In our investigation, we present weighted and unweighted estimates.

2.2 Sample for our investigation

For our investigation, we limit the analysis to women born between 1930 and 1969 who live in the western states of Germany (excluding Berlin). We do not consider the eastern states of Germany because the demographic situation in eastern Germany is still very different from the one in the western areas. Respondents are defined as western Germans depending on the region they live in at the time of interview. This classification does obviously not account for the migration histories of the respondents. Despite these limitations we still find it appropriate enough to use this sample selection for our investigation of fertility behavior in western Germany.

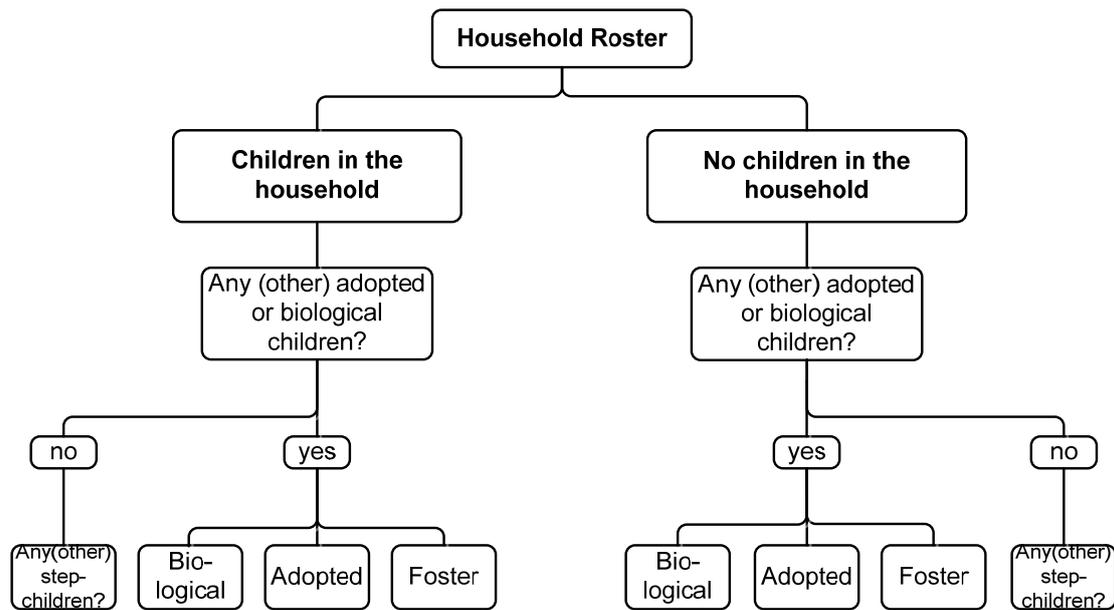
The German GGS contains 10,017 respondents. After selecting females of the cohorts 1930-69 who are resident in the western states of Germany, the sample size is reduced to 2,851 respondents.

2.3 Generating fertility histories from GGS data

Within the GGS questionnaire, the fertility histories of the respondents are recorded in two different sections. In a first section, the respondent is asked to give information on all persons living in her or his household (including children). Here, the year and month of birth of the household members are reported. A distinction between biological children with the current partner, biological children with a former partner, stepchildren, adopted children and foster children is made. From this, one can derive the number of biological and non-biological children who still live in the same household unit as the respondent. In a second section of the questionnaire, the respondent is asked to name any biological children, adopted or foster children who are not living in the same household (see Figure 1). Likewise there is a question asking for stepchildren outside the respondent's household. Again, information on the year and month of birth of these children is collected.

For the subsequent investigation, we only consider the biological children of the female respondents. Women with biological and non-biological children stay in the sample, but only the biological children are used to generate the fertility careers. If, for example, a woman reports to have a biological and an adopted child, we only take the biological child into consideration. Altogether, we observe 2,353 first births in our sample of western German women of the cohorts 1930-69.

Figure 1: Flow chart on how children are surveyed in the Generations and Gender Survey



Source: own representation

2.4 Methods

As a basis for validation, we generate three types of fertility measures: (1) the number of children at interview, (2) the share of childless respondents at interview, (3) the share of childless women by age. For all indicators, we have generated 95 percent confidence intervals around the estimates.

The share of childless women by age is calculated by means of Kaplan-Meier survival curves. In our analysis, the process time for the transition to first birth is defined as the duration between the 15th birthday of the respondent until first birth, counted in months. If the month of birth was unknown, we imputed it by a uniformly distributed random number. Cases are censored at the date of interview or at age 45.

Unfortunately, no exact interview date is given in the German GGS which is why we censored the cases end of 2004. Births that happened in 2005 are therefore not accounted for in this part of the investigation. For measuring the share of childless women, cases with missing information on the year of birth are omitted, too (this applies to 30 cases).

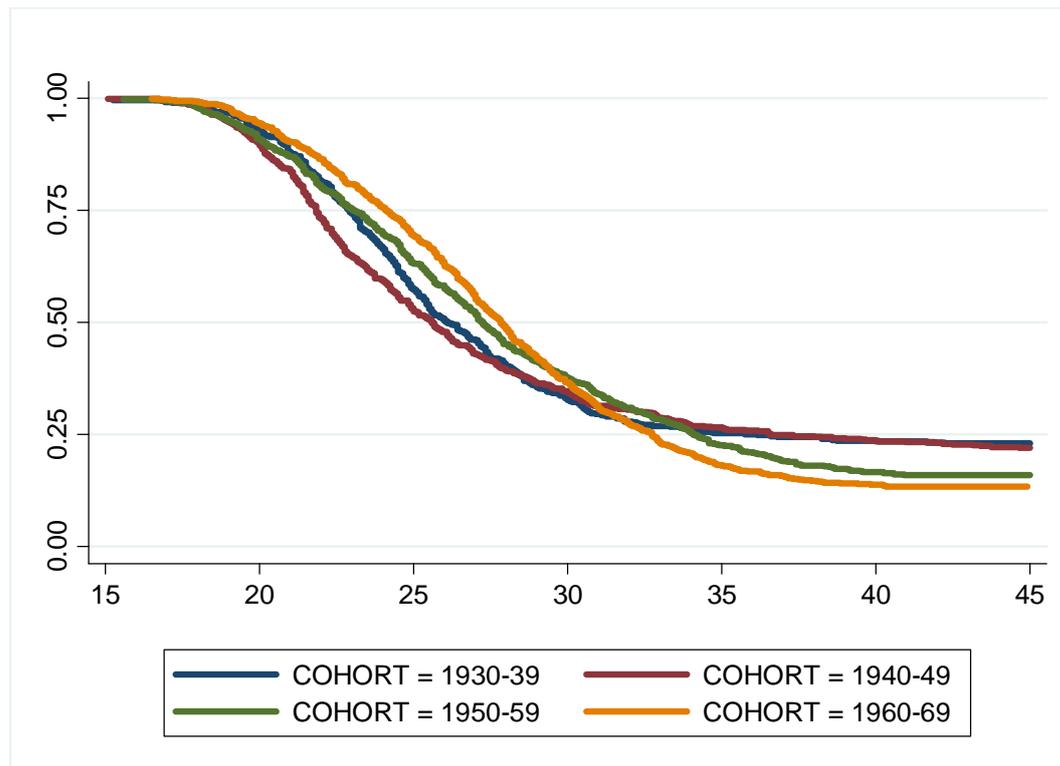
3 Fertility Behavior in western Germany

3.1 Survival curves to first birth

3.1.1 Results from GGS

Previous studies on fertility dynamics in western Germany have revealed a steady postponement of first births and an increase in the share of lifelong childless women since the 1940s cohorts (Duschek and Wirth 2005; Kreyenfeld and Konietzka 2007: 28; Statistisches Bundesamt 2007, 2009a). To our knowledge, until today there is no study which reports a reversal of this trend. Figure 2 presents the survival functions to first birth in western Germany based on GGS data. The graph shows a continuous postponement of first births since the 1940s cohorts which fits prior research. However, focusing on age 40, it is surprising to see that the 1950s and 1960s cohorts display relatively low shares of childlessness. Based on this graph, one would conclude that childlessness has been declining in western Germany since the 1940s cohorts. But this is in sharp contrast to previous findings on childlessness.

Figure 2: Survival curves to first birth, western German women, unweighted estimates



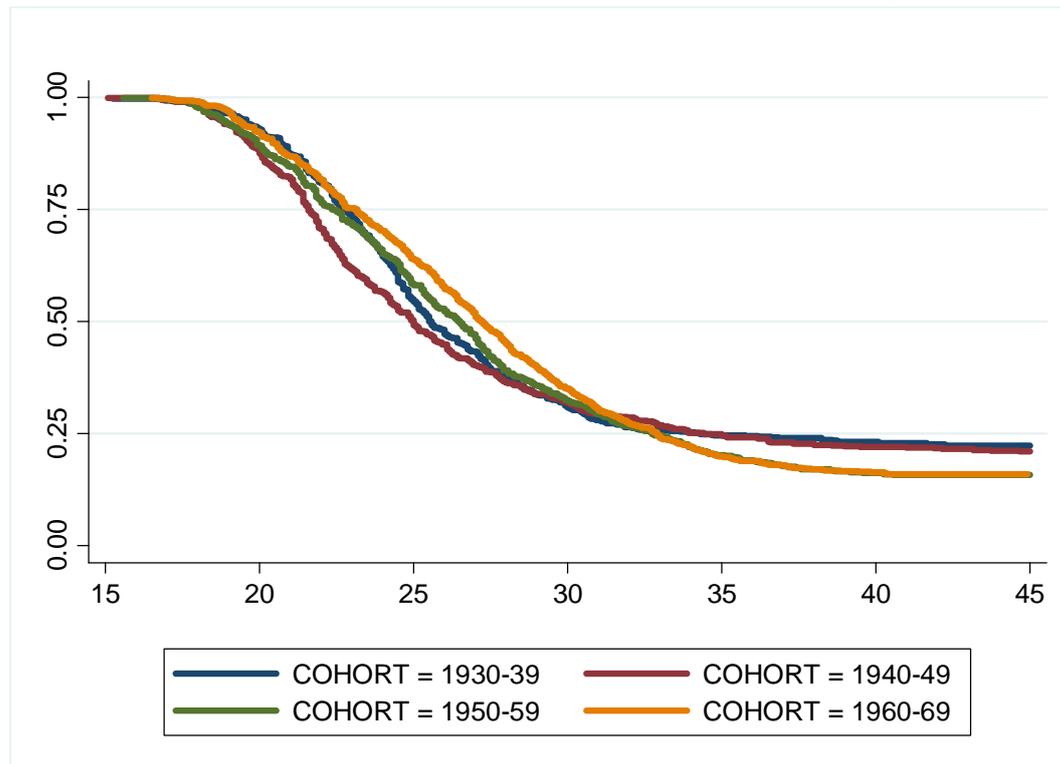
Source: German GGS (own calculations)

3.1.2 How does weighting affect the results?

The first obvious question that arises from this diverging result is whether weighting of the data improves the estimates and gives a more realistic pattern of fertility behavior in western Germany. Figure 3 therefore displays the weighted estimates. As can be seen from this figure, weighting does not change the results much. Both the weighted and the unweighted GGS estimates reveal that women of the western German 1960s cohorts have the lowest share of childlessness compared to their predecessors.

It might be surprising that weighting does not affect the results substantially. But maybe this can be explained by the way that the weights have been constructed. While the weighting factor considers some basic population characteristics (such as age, region, sex, education), it does not include fertility information. A bias in the data that is caused by an oversampling of women with children can probably not be fully corrected by this weighting scheme.

Figure 3: Survival curves to first birth, western German women, weighted estimates



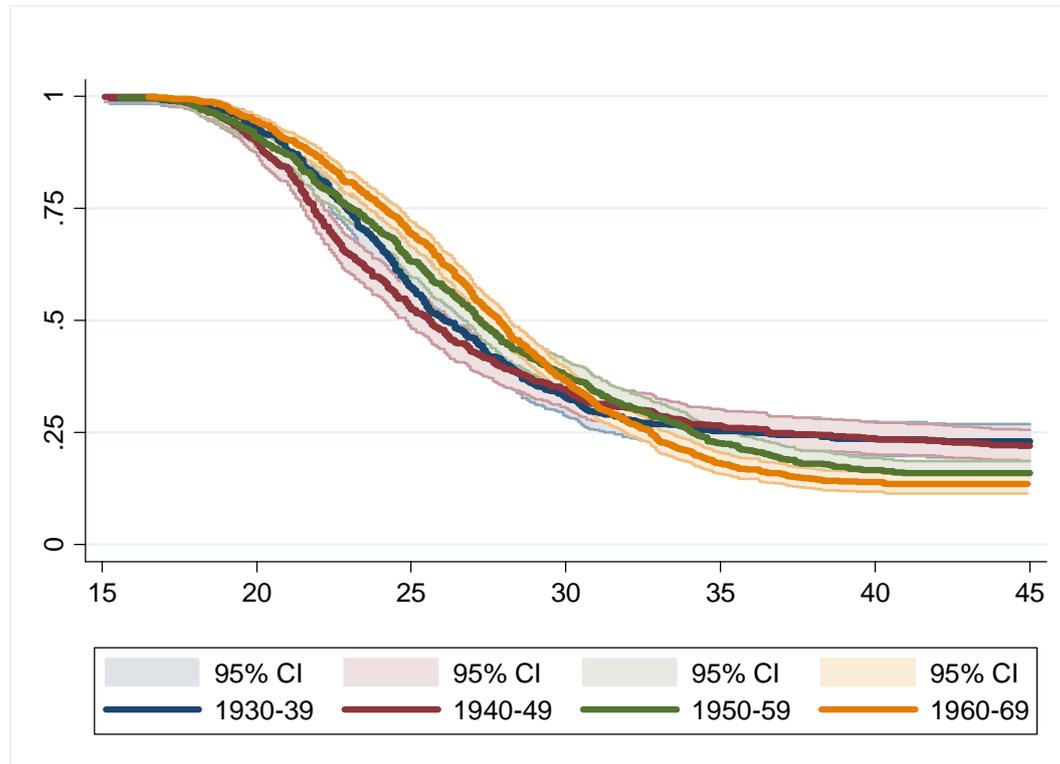
Notes: The data has been weighted by the variable “ppgew”.

Source: German GGS (own calculations)

3.1.3 Is the difference in childlessness between cohorts significant?

The next question concerns the differences in the survival functions and their statistical significance. Maybe the disturbing first birth pattern is induced by the fact that the sample size is just too low to give statically meaningful results. To verify this assumption, Figure 4 contains the 95 percent confidence bounds around the survival functions. This figure shows that the confidence bounds of the 1940s and 1950s cohorts do not overlap at age 40. This suggests that the decline in childlessness at higher ages since the 1940s cohorts is statistically significant.

Figure 4: Survival curves to first birth, western German women, with 95 percent confidence bounds



Source: German GGS (own calculations)

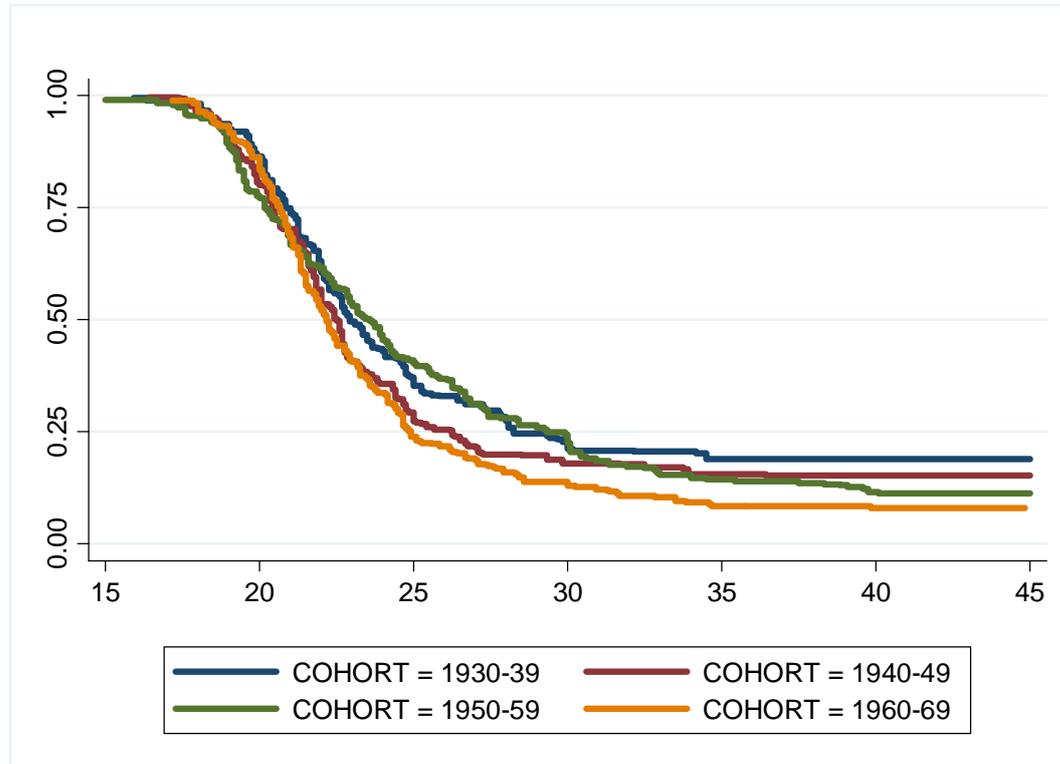
3.1.4 Are the results affected by East-West-German migrants?

The next question that we want to address is whether the odd cohort pattern in the GGS data might be induced by young East-West migrants who have higher first birth rates than western Germans who never migrated. Unfortunately, the first wave of the GGS does not contain migration histories nor does it contain information on the region of birth. This makes it difficult to test this hypothesis. However, there are several arguments that speak against this idea.

Firstly, it is unlikely that East-to-West migrants are able to strongly influence western German fertility patterns due to their relatively small numbers. Another aspect that refutes this hypothesis is that other surveys, which have reported an increase in shares of childlessness over the cohorts, such as the German microcensus, also use information on the place of residence at time of survey to distinguish eastern and western Germans. A final reason objecting to this hypothesis is that, also for eastern Germany, we find a decline in childlessness for the most recent cohorts (Figure 5). For eastern Germany, there is no reason to believe that West-to-East migrants may

bias the level of childlessness downwards in the eastern German states. Neither is there any evidence that the eastern German women of the 1960s cohort have had extraordinary low levels of childlessness.

Figure 5: Survival curves to first birth, eastern German women, weighted estimates²



Source: German GGS (own calculations)

3.2 Childlessness in GGS and microcensus 2008

So far the investigation has shown that survival curves of the transition to first birth derived from GGS data do not match previous findings from other surveys well. One could argue here that prior research could be wrong, too, because it is only based on surveys with relatively small sample sizes. Ideally, one would compare GGS results to results from vital statistics which are not subject to sampling errors. After all, such kind of comparison is not possible for the share of childlessness because order-specific fertility data for Germany is not available from the birth registers up to now.

However, the German microcensus contains a question on childlessness for the year 2008. The microcensus is the largest survey for Germany, covering one percent of the

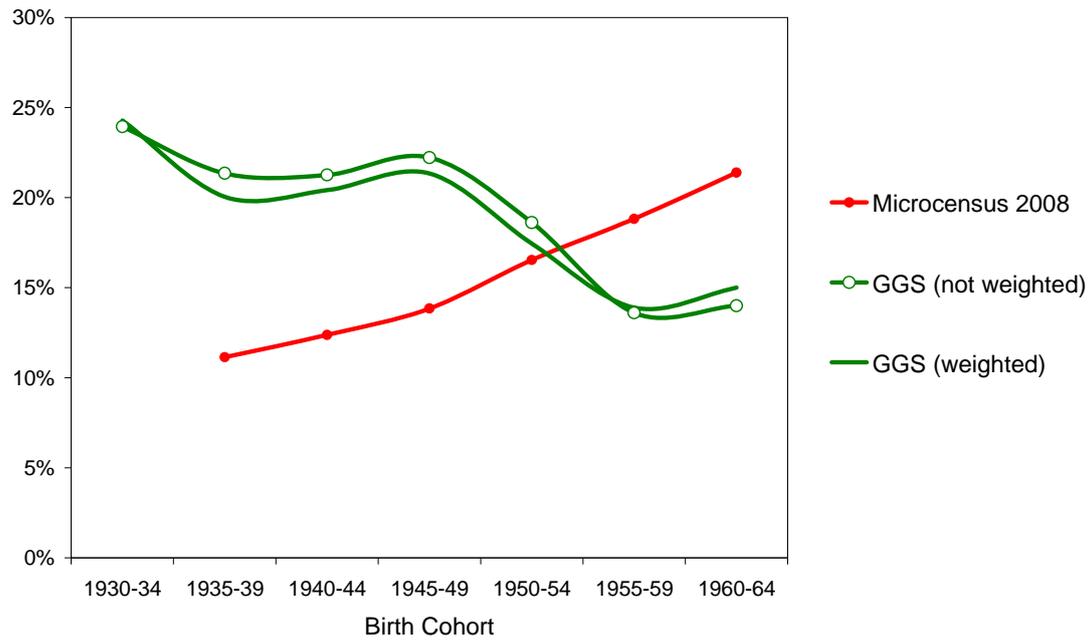
² Eastern German women are defined here as respondents who live in the eastern states of Germany (excluding Berlin) at the time of interview in 2005.

population. A great merit of this data set is the low degree of unit non-response due to the fact that participation in this survey is compulsory. Unfortunately, item non-response for the question on the number of children is quite high which induced the German Statistical Office to apply an imputation scheme to correct this bias (Statistisches Bundesamt 2009a, 2009b). Despite these caveats, the microcensus 2008 is probably the best data source currently available to generate indicators of childlessness for Germany.

Figure 6 graphs the share of childless western German women from the microcensus 2008 and the GGS 2005. Since the cohorts were at least 40 years old at interview, the figure gives an impression of the ultimate level of childlessness. While childlessness has steadily increased over the cohorts from what one can derive from the microcensus estimates, it has declined considerably according to the GGS estimates. Table 1 also gives the 95 percent confidence bounds around the GGS estimates.³ Only for the cohorts 1950-54, the microcensus estimates are captured by the confidence bounds of the estimates from the GGS. For the other cohorts, one must conclude that the two samples stem from different populations.

³ We did not report the confidence bounds for the microcensus data. They would be very narrow due to the high sample size in the microcensus.

Figure 6: Share of childless western German women in microcensus and GGS



Source: German GGS (own calculations), Statistisches Bundesamt (2009a)

Table 1: Share of childless western German women in microcensus and GGS, with 95 percent confidence bounds

Cohorts	Generations and Gender Survey					Micro-census 2008
	Mean (weighted)	Mean (unweighted)	Standard error	95 percent confidence bounds		
1930-34	0.24	0.24	0.03	0.18	0.30	--
1935-39	0.20	0.21	0.02	0.17	0.26	0.11
1940-44	0.20	0.21	0.02	0.17	0.26	0.12
1945-49	0.21	0.22	0.03	0.17	0.27	0.14
1950-54	0.17	0.19	0.02	0.14	0.23	0.17
1955-59	0.14	0.14	0.02	0.10	0.17	0.19
1960-64	0.15	0.14	0.02	0.11	0.17	0.21

Source: German GGS (own calculations), Statistisches Bundesamt (2009a)

3.3 Total number of children in GGS and vital statistics

In this section, the total numbers of children from the vital statistics and from the GGS data are compared to further investigate the deficiencies of the GGS estimates. As mentioned before, it is not possible to compare the results from the survival curves of the transition to first birth directly to the data from the vital statistics due to lacking information on the order of the births. Therefore we compare the levels of cohort fertility in order to diagnose the distortions in the GGS data.

There are some shortcomings regarding the comparability of the two data sources which need to be clarified further: The cohort data from the vital statistics are generated from period data that are “stacked together” in a way that they build a real cohort. This has some implications for the definition of the individual’s regional affiliation. While in the GGS data, a western German is somebody living in western Germany at time of interview, the data from the vital statistics give the fertility pattern of a “synthetic person” who has lived his or her whole life in western Germany.⁴

Moreover, it is a bit fuzzy to compare cohorts across these two types of data. For the GGS, it is necessary to group several cohorts into one category to assure a sufficiently high sample size. For the vital statistics, we only have the average number of children for single cohorts at our disposal.⁵ In order to make data sources comparable, we used the cohort from the vital statistics which falls in the middle of the cohort grouping of the GGS.

Figure 7 displays the results from the comparison. Moreover, the data from the microcensus 2008 are included in this figure. The estimates on the total number of children from the microcensus 2008 are a bit higher than from the vital statistics, but the trend is the same: The average number of children per woman has declined since the 1930s cohort in western Germany. The GGS results, however, show that the total

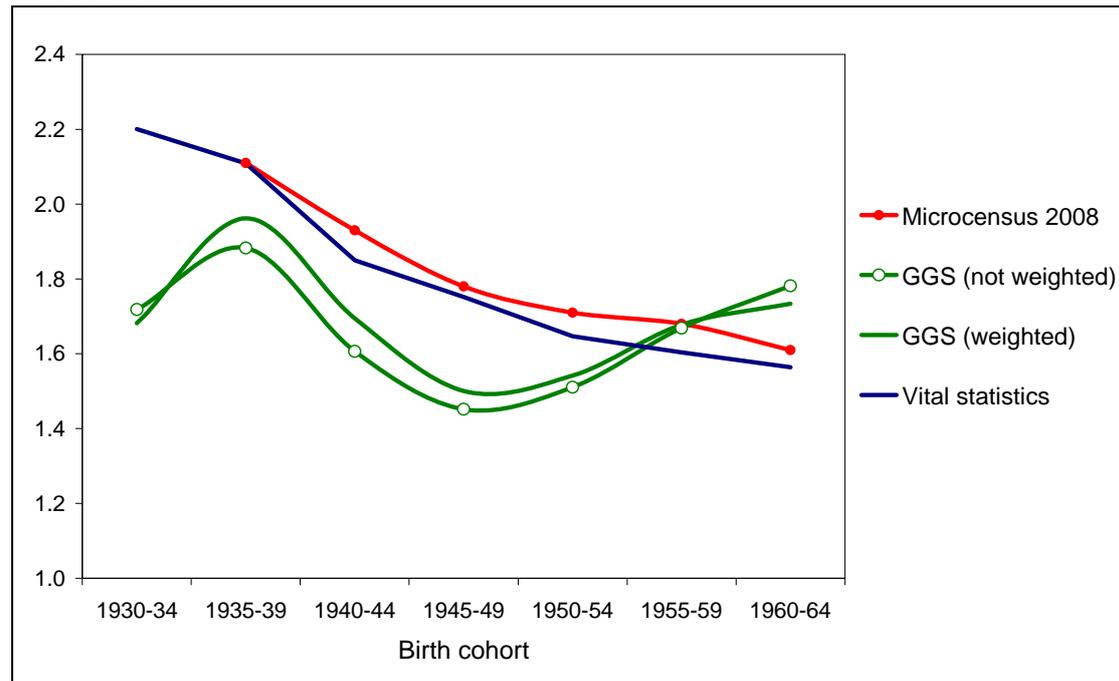
⁴ For the time prior to 1999, a western German is a person who lives in the western states of Germany including West Berlin. For the time after 1999, a western German is a person who lives in western Germany excluding West Berlin. The cut point in the vital statistics is 2000, because this is the year of the regional reform which made it difficult to furthermore distinguish East and West Berlin.

⁵ There are other differences which should be mentioned, too: The vital statistics include fertility information for all individuals for the time individuals live in Germany. The GGS contains retrospective fertility information for respondents who live in Germany at the time of interview. Selection on survival and migration affects the comparability of both sources.

number of children per woman has increased since the cohorts born between 1945 and 1949.

Table 3 reports the 95 percent confidence bounds around the GGS estimates. It is only for the cohorts 1955-59 that we can conclude that the GGS is similar to the data from the vital statistics.

Figure 7: Average number of children, western German women, microcensus, vital statistics and GGS



Notes: see Table 2

Table 2: Average number of children of western German women in microcensus, vital statistics and GGS, with confidence bounds

Cohorts	Generations and Gender Survey				Vital statistics	Micro-census 2008
	Mean (weighted)	Mean (unweighted)	Standard error	95 percent confidence bounds		
1930-34	1.68	1.72	0.10	1.53 1.91	2.20	--
1935-39	1.96	1.88	0.09	1.70 2.06	2.11	2.11
1940-44	1.69	1.61	0.08	1.46 1.75	1.85	1.93
1945-49	1.50	1.45	0.06	1.33 1.58	1.75	1.78
1950-54	1.54	1.51	0.06	1.39 1.63	1.65	1.71
1955-59	1.68	1.67	0.05	1.57 1.77	1.60	1.68
1960-64	1.83	1.82	0.05	1.70 1.95	1.56	1.61

Note: The cohort fertility rates from the vital statistics were compiled from various sources provided by the Statistisches Bundesamt. The values from the vital statistics refer to the class middle. E.g. the value 2.20 in the first row is the total number of children of western German women from the cohort 1932.

Source: German GGS 2005 (own calculations), Statistisches Bundesamt (2009a)

4 What are the reasons for the bias?

Our investigation so far suggests that there is an over-reporting of births for the younger cohorts and an under-reporting of births for the older ones. How can such a pattern emerge? Our best guess is here that two counteracting mechanisms are at play. On the one hand, the GGS includes too few childless respondents due to the fact that this group is difficult to reach in standard surveys, in particular the ones which are based on a random route design like it is the case for the German GGS (Koch 1998; Festy and Prioux 2002: 23). On the other hand, we suspect that children who have left parental home have not been adequately reported by all of the respondents. This kind of bias should particularly affect the older cohorts whose children have more often left parental home. These two mechanisms taken together could produce a pattern that gives too few births for the older cohorts and too many ones for the younger cohorts. If this interpretation is correct, this would suggest that item non-response is higher for the older than for the younger cohorts.

4.1 Item non-response

The question how item non-response is distributed over cohorts seems to be a simple one. However, the answer to it is more difficult. Due to the complex nature in which fertility careers have been surveyed in the GGS it is very cumbersome to detect item non-response for those respondents who did not provide reliable information on their fertility careers.

The number of children in the GGS must be generated by relying on two types of sources, namely the household grid and the module that surveys information on children who do not live in the household of the respondent anymore (see Figure 1). Information on the number of children can be missing if a person does not fill in the household grid or if he or she refuses to answer the module where information on the non-residential children is surveyed. In the German GGS, there is no item non-response for the household grid⁶ and only 23 persons (out of 10,017) refused to

⁶ In the original variable for the household grid, there are some few missing cases (f010100a, f010100b etc.). However, the Bundesinstitut für Bevölkerungsforschung (BiB) has provided a cleaned variable for the household grid that does not contain any missing values (hh_rasta, hh_rastb etc.).

answer the module on non-residential children. In our sample of western German women of the cohorts 1930-1969, there are only 4 persons who did not give valid answers on non-residential children.

If there is no item non-response, one might suspect whether respondents have given wrong information. One possible reason for this is that the retrospective part was too complex for some respondents (or for some interviewers) so that they decided to skip this section of the questionnaire. These refusals might have been wrongly coded as “no residential children”. If this assumption is correct, the strong bias should be restricted to the fertility indicators and should not appear for other socio-economic characteristics.

4.2 Is there a bias for other socio-economic characteristics?

Table 3a and Table 3b compare data from the microcensus 2005 with data from the GGS 2005. If our assumption is right that mainly the retrospective module creates a problem in the GGS, then there should not be equal distortion for other socio-demographic characteristics. Table 3b shows that, in the data, highly-educated people are over- and foreigners are under-represented in the German GGS. This is very much in line with what is reported in the documentation of the German GGS (Ruckdeschel et al. 2006).

Our central question is, however, if there is a systematic bias that varies by the cohorts. All we can say here is that there is some over-representation of university graduates in the older cohorts. This bias disappears, however, after weighting the data. In essence, we do not detect a very systematic bias in these socio-economic variables. This again raises the question whether it is mainly the fertility histories that create a problem in the GGS.

Table 3a: Distribution of sample size by cohort groups, GGS 2005 and microcensus 2005

	GGG 2005 (unweighted)	GGG 2005 (weighted)	Microcensus 2005 (unweighted)
Cohorts 1930-39	0.19	0.19	0.21
Cohorts 1940-49	0.22	0.23	0.23
Cohorts 1950-59	0.26	0.26	0.26
Cohorts 1960-69	0.32	0.32	0.30
Total	1.00	1.00	1.00
N	6,981	--	260,262

Source: German GGS 2005, German microcensus 2005 (own calculations)

Table 3b: Socio-demographic characteristics by cohort groups, GGS 2005 and microcensus 2005

	GGG 2005 (unweighted)	GGG 2005 (weighted)	Microcensus 2005 (unweighted)
<i>Eastern Germans</i>			
Cohorts 1930-39	0.25	0.22	0.24
Cohorts 1940-49	0.24	0.22	0.23
Cohorts 1950-59	0.24	0.22	0.23
Cohorts 1960-69	0.20	0.21	0.21
<i>Females</i>			
Cohort 1930-39	0.51	0.54	0.54
Cohort 1940-49	0.48	0.50	0.50
Cohort 1950-59	0.56	0.50	0.51
Cohort 1960-69	0.57	0.49	0.50
<i>University graduates</i>			
Cohort 1930-39	0.17	0.10	0.08
Cohort 1940-49	0.20	0.15	0.13
Cohort 1950-59	0.24	0.19	0.16
Cohort 1960-69	0.21	0.17	0.17
<i>Foreign nationals</i>			
Cohort 1930-39	0.02	0.02	0.02
Cohort 1940-49	0.04	0.04	0.06
Cohort 1950-59	0.04	0.04	0.07
Cohort 1960-69	0.06	0.07	0.09
N	6,981	6,981	

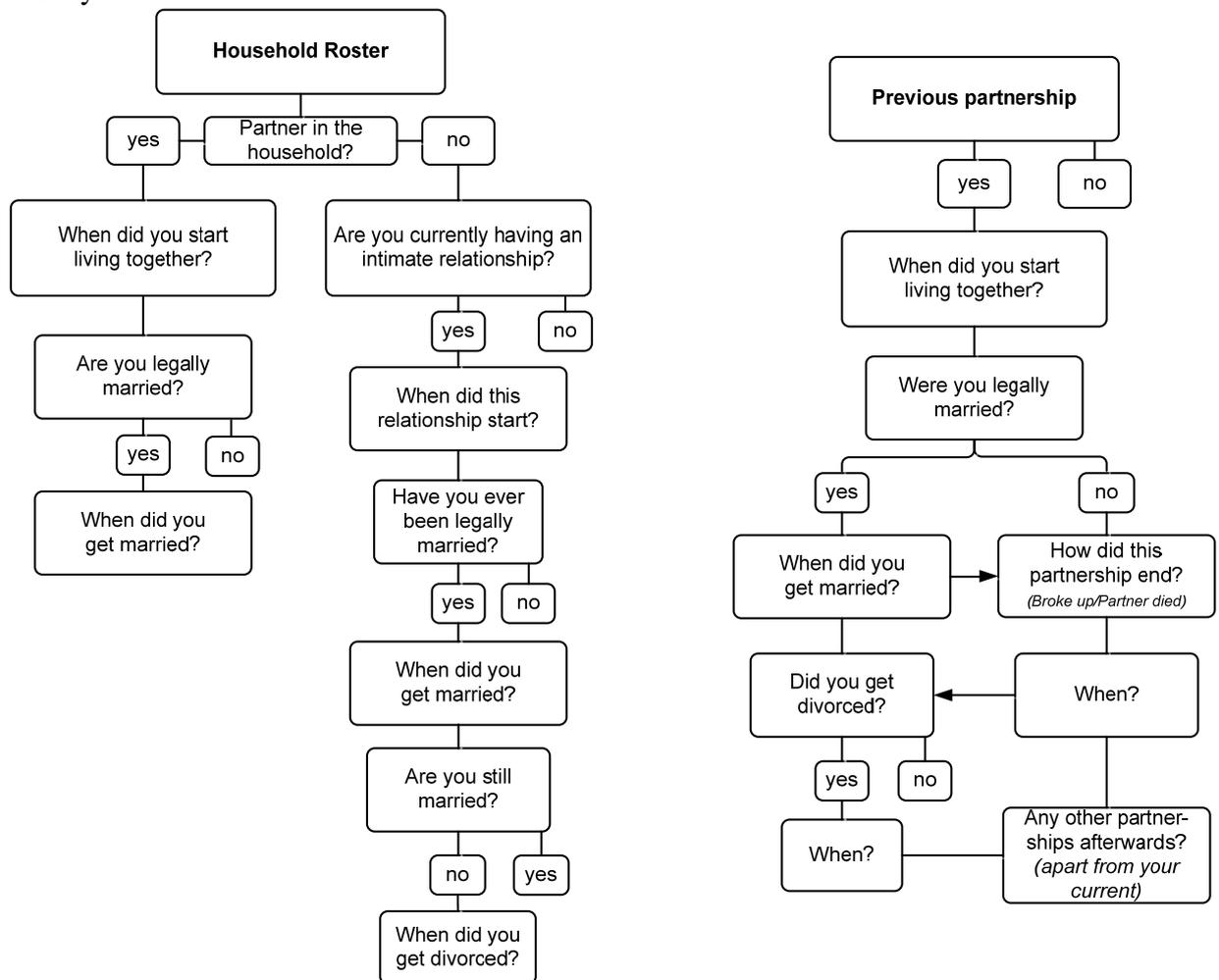
Note: For this table, we have included West Berlin into West Germany. Respectively, East Germany does not include West Berlin. For this part of the analysis, we used the national file of the GGS, provided by the BiB

Source: German GGS 2005 (national file), German microcensus 2005 (own calculations)

4.3 Is there a bias for other vital events?

The main “oddity” of the fertility careers of the GGS is that the bias in the fertility careers is very unevenly distributed across cohorts which is apparently not the case for other socio-economic variables. Fertility histories are surveyed differently depending on how old the respondents are at interview. Older respondents, whose children have left parental home, must provide their fertility information through the retrospective histories. This might have created a problem for some respondents (or interviewers). If this assumption is correct, then the bias should not only matter for fertility careers, but also for other events which have been surveyed in a similarly complex fashion. Partnership histories have been surveyed according to a similar logic as fertility histories (see Figure 8). If our hypothesis is correct that the retrospective part that collects information on non-residential children “did not work”, then it should have failed also for the partnership histories.

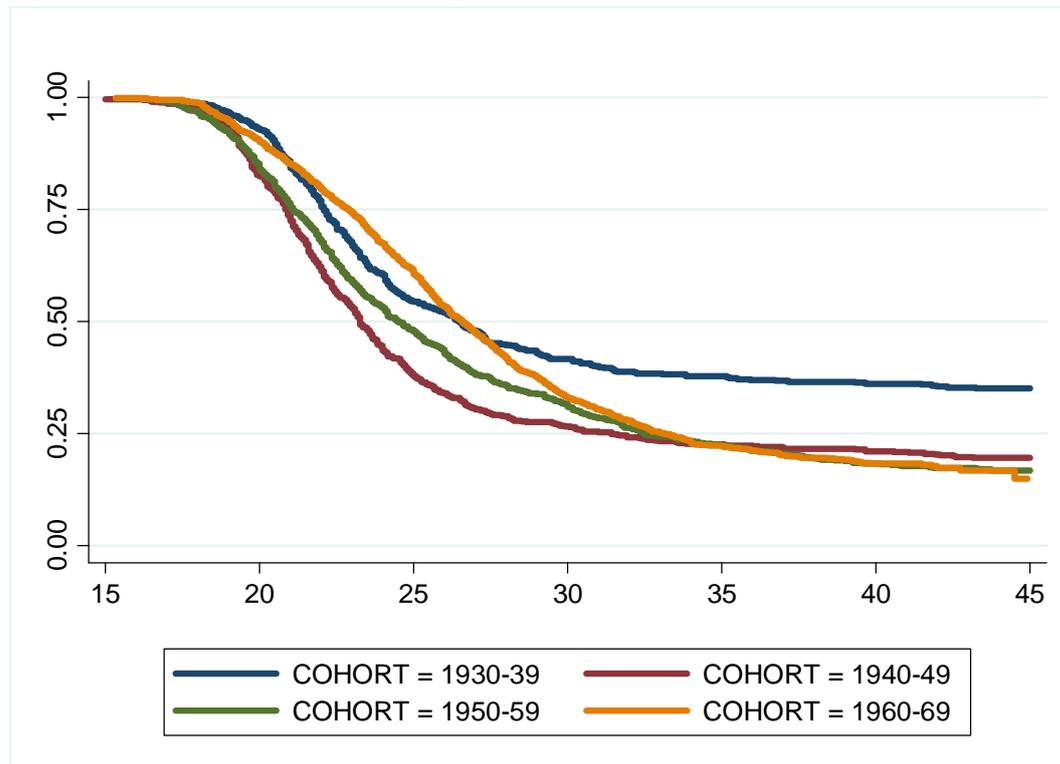
Figure 8: Flow chart on how partnerships are surveyed in the Generations and Gender Survey



Source: own representation

Figure 9 therefore displays estimates of the proportion of unmarried women by age and cohort based on the GGS. In line with other studies, we find a postponement of first marriage since the 1940s cohorts. However, what comes as a surprise is that the share of women unmarried at age 35 did not change over the cohorts. What very much sticks out are the 1930s cohorts showing a very slow transition to first marriage and low shares of women married at age 45. This is not in line with previous findings on marriage patterns for Germany at all which show rather high shares of ever married women for the cohorts born in the 1930s (e.g. Blossfeld and Jaenichen 1992: 308; Huinink 1995: 227).

Figure 9: Transition to first marriage, western German women, GGS



Source: German GGS 2005 (own calculations)

4.4 Is there any systematic under-reporting of non-residential children?

The analysis so far has suggested that the bias that we have found for the fertility career is similar to the one related to partnership careers. This supports our idea that maybe the module where information on non-resident persons is collected did not work. To finally assess this assumption, we compared GGS results with other German surveys which contain information on co-residential and non-residential children. For comparison, we used data from the DJI-Family Survey 2000. The difference in the survey year might distort the results. However, for the moment we do not have a more recent survey for this kind of analysis at our hands.⁷

Table 4 gives the share of women whose children no longer live in the household (out of all women). If our assumption is correct, the share of non-residential children

⁷ The ALLBUS 2006 would provide more recent evidence, but case numbers are too small to provide reliable estimates.

should be lower in the GGS than in other surveys. In principle, this assumption is supported by the results from the table.

Table 4: Share of respondents who have non-residential children

Age at interview	GGS 2005	DJI-Family-Survey 2000
30-34	1.7%	2.5%
35-39	3.2%	4.8%
40-44	6.8%	10.6%
45-49	24.2%	36.3%
50-54	49.4%	61.4%
55-59	74.9%	83.5%

Source: German GGS 2005, DJI-Family Survey 2000 (own calculations)

5 Conclusion

The aim of this paper was to validate fertility histories from the German Generations and Gender Survey. Survival functions for the transition to first birth were estimated by different cohort groups of western German women. We compared GGS results to estimates from other surveys as well as to fertility indicators from vital statistics. Our main finding is that fertility of the older GGS cohorts is too low, while it is too high for the younger cohorts.

We investigated possible reasons for the odd fertility patterns that we find in the GGS. In particular, we followed the idea that the bias might be caused by two counteracting mechanisms. On the one hand, the GGS includes too few childless respondents due to the fact that this group is difficult to reach in standard surveys. On the other hand, we suspect that children who have left parental home have not been adequately reported by all of the respondents due to the complex nature of how non-residential children are surveyed in the GGS. These two mechanisms taken together have produced a pattern that gives too low numbers of children for the older cohorts and too large ones for the younger cohorts.

There are two aspects that support this idea. Firstly, we find a similar bias for the partnership careers of the German GGS. Our investigation gives unreasonably low

shares of married women for the older cohorts (especially, the ones born 1930-39). Since the partnership histories are surveyed in a similarly complex fashion as the fertility history, the source of bias might be similar. Moreover, comparison with other data gives some mild support for the idea that vital events are under-reported in the German GGS, because the retrospective module did not capture the non-residential children and partners in a satisfying manner.

Nevertheless, this is only a best guess of what the source of bias in the GGS could be. It was very cumbersome to assess the quality of the fertility information from the GGS because item non-response is very difficult to track. We had wished, at various occasions for a simple question asking respondents how many children they had born/fathered over their life course. Unlike the predecessor of the GGS, the Family and Fertility Survey (FFS), the GGS did not contain such a simple question any more. The same can be said for partnership or marriage. Here, one dearly misses a simple question that surveyed the marital status at interview. Such a simple question would have been highly valuable to check the quality of the partnership information given in the retrospective partnership histories.

However, a lot of aspects remain unresolved. Particularly, it remains unclear why such a pattern evolved only for Germany, while such a bias does not seem to have emerged for other GGS countries that used similar questionnaires. Furthermore, it remains unclear how to cure this bias. Cohort changes in partnership formation and fertility are central themes for family and fertility researchers. Therefore it is very regrettable that for several important demographic aspects the behavioral changes over the cohorts cannot be examined. In addition, it is central for policy makers to have a clear-cut answer to the question whether there is drastic postponement of first births and first union formations or not. It just seems unfortunate that the German GGS is unable to give an answer to these very basic questions.

6 Acknowledgment

We thank Robert Herter-Eschweiler from the Statistisches Bundesamt for providing data from the German microcensus 2008. We thank the Bundesinstitut für Bevölkerungsforschung (BiB) for providing us with the GGS data. For valuable comments we thank Krystof Zeman. This paper was presented at the 2nd Workshop of the Nonmarital Childbearing Network held in Rostock in January 2010. We also thank the participants of this venue for their comments on this paper.

7 References

- Blossfeld, H.-P./ Jaenichen, U. (1992): Educational expansion and changes in women's entry into marriage and motherhood in the Federal Republic of Germany. *Journal of Marriage and the Family* 54 (2): 302-315.
- Duschek, K.-J./ Wirth, H. (2005): Kinderlosigkeit von Frauen im Spiegel des Mikrozensus. Eine Kohortenanalyse der Mikrozensus 1987 bis 2003. *Wirtschaft und Statistik* (8): 800-820.
- Festy, Patrick/ Prioux, France (2002): An evaluation of the Fertility and Family Surveys Project. New York and Geneva: United Nations.
- Huinink, J. (1995): *Warum noch Familie?* Frankfurt: Campus.
- Koch, A. (1998): Wenn „mehr“ nicht gleichbedeutend mit „besser“ ist: Ausschöpfungsquoten und Stichprobenverzerrungen in allgemeinen Bevölkerungsumfragen *ZUMA-Nachrichten* 42: 66-93.
- Kreyenfeld, M./ Konietzka, D. (2007): Die Analyse von Kinderlosigkeit in Deutschland: Dimensionen – Daten – Probleme. In: Konietzka, D./ Kreyenfeld, M. (eds.): *Ein Leben ohne Kinder. Kinderlosigkeit in Deutschland*. Wiesbaden: VS Verlag für Sozialwissenschaften: 11-41.
- Murphy, M. (2009): Where have all the children gone? Women's reports of more childlessness at older ages than when they were younger in a large-scale continuous household survey in Britain. *Population Studies* 63 (2): 115-133.
- Perelli-Harris, B./ Kreyenfeld, M./ Kubisch, K. (2009): Manual for the cleaning of the German GGS. *MPIDR-Working paper 2010-08*.
- Rendall, M. S./ Clarke, L./ Peters, E. H./ Ranjit, N./ Verropoulou, G. (1999): Incomplete Reporting of Men's Fertility in the United States and Britain: A Research Note. *Demography* 36 (1):135-144.
- Ruckdeschel, K./ Ette, A./ Hullen, G./ Leven, I. (2006): *Generations and Gender Survey. Dokumentation der ersten Hauptbefragung in Deutschland*. Materialien zur Bevölkerungswissenschaft, Heft 121a. Wiesbaden: Bundesinstitut für Bevölkerungsforschung.
- Statistisches Bundesamt (2007): *Geburten in Deutschland 2006*. Wiesbaden: Statistisches Bundesamt.
- Statistisches Bundesamt (2009a): *Mikrozensus 2008. Neue Daten zur Kinderlosigkeit in Deutschland*. Begleitmaterial zur Pressekonferenz am 29. Juli 2009 in Berlin. Wiesbaden: Statistisches Bundesamt. (Link: http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Content/Publikationen/Fachveroeffentlichungen/Bevoelkerung/GeburtenSterbefaelle_property=file.pdf)
- Statistisches Bundesamt (2009b): *Imputation von Werten bei fehlenden Angaben zur Mutterschaft und zur Zahl der geborenen Kinder im Mikrozensus 2008*. Internal Document from Statistisches Bundesamt (provided by Julia Weinmann).
- Vikat, A. et al. (2008): Generations and Gender Survey: Concept and Design. In: UNECE/ UNPF (eds.): *Generations & Gender Programme. Survey Instruments*. Geneva: United Nations. 1-32.

8 References

***The German Generation and Gender Survey

***Some Critical Reflections on the validity of birth and partnership histories

****1. Part Generation of the Harmonized Histories File from the Original GGS file for Germany

****Shorted version for the paper (background and fertility histories)

```
clear all
set mem 350M
```

```
*Use Original GGS data
use "u\36891_6.dta", clear
```

```
*1.1.Information about the respondent and background information*
```

```
*Identification number
```

```
g ARID=respid
lab var ARID "Original ID number"
label values ARID ARID
```

```
*Sex of respondent
```

```
gen SEX=.
replace SEX=1 if f010300==1
replace SEX=2 if f010300==2
lab var SEX "Sex of the respondent"
lab def SEX 1 "Male" 2 "Female"
label values SEX SEX
```

```
*Year of birth of respondent
```

```
g BORN_Y=f010400j
replace BORN_Y=.a if BORN_Y==.
lab var BORN_Y "Year of birth of respondent"
label values BORN_Y BORN_Y
```

```
*Month of birth of respondent
```

```
g BORN_M=f010400m
replace BORN_M=.a if BORN_M==.
lab var BORN_M "Month of birth of respondent"
lab def BORN_M .a "Unknown" 1 "January" 2 "February" 3 "March" 4 "April" 5 "May" 6 "June" 7 "July" 8 "August" 9
"September" 10 "October" 11 "November" 12 "December" 21 "Winter months at the beginning of the year" 22
"Spring" 23 "Summer" 24 "Autumn" 25 "Winter at the end of the year"
label values BORN_M BORN_M
```

```
*Month of birth of respondent - missings imputed randomly
```

```
g IBORN_M=BORN_M
replace IBORN_M=int(uniform()*12)+1 if BORN_M==.a
lab var IBORN_M "Month of birth of respondent and imputed months"
lab def IBORN_M 1 "January" 2 "February" 3 "March" 4 "April" 5 "May" 6 "June" 7 "July" 8 "August" 9 "September" 10
"October" 11 "November" 12 "December"
label values IBORN_M IBORN_M
```

```
*Birth cohort: Classification I
```

```
g COHORT=-1
replace COHORT=1 if BORN_Y>=1930
replace COHORT=2 if BORN_Y>=1940
replace COHORT=3 if BORN_Y>=1950
replace COHORT=4 if BORN_Y>=1960
replace COHORT=-1 if BORN_Y>=1970
lab var COHORT "Birth cohort"
lab def COHORT 1 "1930-1939" 2 "1940-1949" 3 "1950-1959" 4 "1960-1969" -1 "After 1969"
label values COHORT COHORT
```

```
*Birth cohort: Classification II
```

```
g COHORT2=-1
replace COHORT2=1 if BORN_Y>=1930
replace COHORT2=2 if BORN_Y>=1935
replace COHORT2=3 if BORN_Y>=1940
replace COHORT2=4 if BORN_Y>=1945
replace COHORT2=5 if BORN_Y>=1950
replace COHORT2=6 if BORN_Y>=1955
replace COHORT2=7 if BORN_Y>=1960
replace COHORT2=-1 if BORN_Y>=1965
```

```
lab var COHORT2 "Birth cohort2"
lab def COHORT2 1 "1930-1934" 2 "1935-1939" 3 "1940-1944" 4 "1945-1949" 5 "1950-1954" 6 "1955-1959" 7 "1960-1964"
label values COHORT2 COHORT2
```

```
*Region: West or East Germany
g East=.
replace East=0 if bula>= 1 & bula<= 10
replace East=1 if bula> 11 & bula<= 16
lab var East "West or East Germany"
lab def East 0 "West without Berlin" 1 "East without Berlin"
label values East East
```

```
*Month of survey
g MONTH_S=.c
lab var MONTH_S "Month of survey"
lab def MONTH_S .c "Unavailable in survey" 1 "January" 2 "February" 3 "March" 4 "April" 5 "May" 6 "June" 7 "July" 8 "August" 9 "September" 10 "October" 11 "November" 12 "December" 21 "Winter months at the beginning of the year" 22 "Spring" 23 "Summer" 24 "Autumn" 25 "Winter at the end of the year"
label values MONTH_S MONTH_S
```

```
*Month of survey - imputed radomly
g IMONTH_S=MONTH_S
replace IMONTH_S=int(2+(4)*uniform()) if MONTH_S==.c
lab var IMONTH_S "Month of survey,including imputed dates"
lab def IMONTH_S .c "Unavailable in survey" 1 "January" 2 "February" 3 "March" 4 "April" 5 "May" 6 "June" 7 "July" 8 "August" 9 "September" 10 "October" 11 "November" 12 "December"
label values IMONTH_S IMONTH_S
```

```
*Year of survey
g YEAR_S=2005
lab var YEAR_S "Year of survey"
lab def YEAR_S 2005 "2005"
label values YEAR_S YEAR_S
```

```
*Weights
*Personal weight
g PERSWGT=ppgew
lab var PERSWGT "Personal weight"
lab def PERSWGT .a "Unknown" .b "Does not apply" .c "Unavailable in survey"
label values PERSWGT PERSWGT
```

```
sort ARID
save backgerm,replace
```

```
**SAMPLE NUMBERS 2.2.
```

```
*Total number of respondents GGS -10017 respondents
count
*Total number of women, born between 1930 and 1969, Western Germany excluding Berlin -2851 respondents
count if SEX==2 & COHORT>=1 & COHORT<=4 & East==0
```

```
***1.2.Generating fertility histories (2.3)***
***1.2.1.Biological children in household***
```

```
use "u\36891_6.dta", clear
rename respid ARID
lab var ARID "Original ID number"
label values ARID ARID
keep ARID hh_rast* f01101j* f01101m*
```

```
*correction of mistakes, found by checking procedures
replace hh_rastb=1 if ARID==31941
replace hh_rastd=2 if ARID==31941
replace hh_raste=. if ARID==31941
replace f01101ja=1986 if ARID==366
replace f01101ja=1979 if ARID==42455
replace f01101ja=999 if ARID==44983
replace f01101ja=999 if ARID==33108
replace f01101ja=999 if ARID==44931
*****
```

```
reshape long hh_rast f01101j f01101m, i(ARID)(JUNK) string
```

```
*Only biological children  
keep if hh_rast==2 | hh_rast==3
```

```
sort ARID
```

```
rename hh_rast KID_  
rename f01101m KID_M  
rename f01101j KID_Y
```

```
g type=1
```

```
keep KID_Y KID_M KID_type ARID  
save kidhist1germ,replace
```

```
***1.2.2.Biological children not in household***
```

```
clear  
set mem 250M
```

```
use "u\36891_6.dta", clear  
rename respid ARID  
lab var ARID "Original ID number"  
label values ARID ARID
```

```
keep ARID f021111 f021112* f021200* f021300* f02140j* f02140m* f021500* f02160j* f02160m* f02170j* f02170m*  
f021800* f021900* f02200j* f02200m* f022400* f010400j f010400m
```

```
g obnra=1 if f021112a!=. |f021200a!=. | f021300a!=. | f02140ja!=. | f02140ma!=. | f021500a!=. | f02160ja!=. |  
f02160ma!=. | f02170ja!=. | f02170ma!=. |f021800a!=. | f021900a!=. | f02200ja!=. | f02200ma!=.  
g obnrb=1 if f021112b!=. |f021200b!=. | f021300b!=. | f02140jb!=. | f02140mb!=. | f021500b!=. | f02160jb!=. |  
f02160mb!=. | f02170jb!=. | f02170mb!=. |f021800b!=. | f021900b!=. | f02200jb!=. | f02200mb!=.  
g obnrc=1 if f021112c!=. |f021200c!=. | f021300c!=. | f02140jc!=. | f02140mc!=. | f021500c!=. | f02160jc!=. |  
f02160mc!=. | f02170jc!=. | f02170mc!=. |f021800c!=. | f021900c!=. | f02200jc!=. | f02200mc!=.  
g obnrd=1 if f021112d!=. |f021200d!=. | f021300d!=. | f02140jd!=. | f02140md!=. | f021500d!=. | f02160jd!=. |  
f02160md!=. | f02170jd!=. | f02170md!=. |f021800d!=. | f021900d!=. | f02200jd!=. | f02200md!=.  
g obnre=1 if f021112e!=. |f021200e!=. | f021300e!=. | f02140je!=. | f02140me!=. | f021500e!=. | f02160je!=. |  
f02160me!=. | f02170je!=. | f02170me!=. |f021800e!=. | f021900e!=. | f02200je!=. | f02200me!=.  
g obnrf=1 if f021112f!=. |f021200f!=. | f021300f!=. | f02140jf!=. | f02140mf!=. | f021500f!=. | f02160jf!=. | f02160mf!=. |  
f02170jf!=. | f02170mf!=. |f021800f!=. | f021900f!=. | f02200jf!=. | f02200mf!=.  
g obnrg=1 if f021112g!=. |f021200g!=. | f021300g!=. | f02140jg!=. | f02140mg!=. | f021500g!=. | f02160jg!=. |  
f02160mg!=. | f02170jg!=. | f02170mg!=. |f021800g!=. | f021900g!=. | f02200jg!=. | f02200mg!=.  
g obnrh=1 if f021112h!=. |f021200h!=. | f021300h!=. | f02140jh!=. | f02140mh!=. | f021500h!=. | f02160jh!=. |  
f02160mh!=. | f02170jh!=. | f02170mh!=. |f021800h!=. | f021900h!=. | f02200jh!=. | f02200mh!=.  
g obnri=1 if f021112i!=. |f021200i!=. | f021300i!=. | f02140ji!=. | f02140mi!=. | f021500i!=. | f02160ji!=. | f02160mi!=. |  
f02170ji!=. | f02170mi!=. |f021800i!=. | f021900i!=. | f02200ji!=. | f02200mi!=.  
g obnrj=1 if f021112j!=. |f021200j!=. | f021300j!=. | f02140jj!=. | f02140mj!=. | f021500j!=. | f02160jj!=. | f02160mj!=. |  
f02170jj!=. | f02170mj!=. |f021800j!=. | f021900j!=. | f02200jj!=. | f02200mj!=.
```

```
keep ARID obnr* f02160m* f02160j* f021300*  
reshape long obnr f02160j f02160m f021300, i(ARID)(junk) string
```

```
*Only biological children  
keep if f021300==1  
drop f021300  
sort ARID
```

```
rename obnr KID_  
rename f02160m KID_M  
rename f02160j KID_Y
```

```
g type=2  
save kidhist2germ,replace
```

```
*all biological children  
append using kidhist1germ  
sort ARID KID_ KID_Y KID_M  
save kidhistgerm3,replace
```

```
*manual recoding of implausible cases  
replace KID_Y=.a if ARID==44983 & KID_Y==999  
replace KID_Y=.a if ARID==33108 & KID_Y==999  
replace KID_Y=.a if ARID==44931 & KID_Y==999
```

```

drop type junk
by ARID: g _j=_n
sort ARID KID_Y KID_M

reshape wide KID_ KID_M KID_Y, i(ARID)

keep ARID KID_1 KID_2 KID_3 KID_4 KID_5 KID_6 KID_7 KID_8 KID_9 KID_Y1 KID_M1
order ARID KID_1 KID_2 KID_3 KID_4 KID_5 KID_6 KID_7 KID_8 KID_9 KID_Y1 KID_M1
sort ARID
merge ARID using backgerm

keep ARID KID_1 KID_2 KID_3 KID_4 KID_5 KID_6 KID_7 KID_8 KID_9 KID_Y1 KID_M1 YEAR_S MONTH_S
IMONTH_S SEX BORN_Y BORN_M IBORN_M COHORT* East PERSWGT
save papergerm,replace

*Standardizing and labeling procedures
forvalues i=1/9 {
replace KID_`i'=1 if KID_`i'!=.
replace KID_`i'=0 if KID_`i'==.
}

lab var KID_1 "Indicator of child order 1"
lab def KID_1 .a "Unknown" .b "Does not apply" .c "Not included in survey" 0 "No child of order 1" 1 "Child of order 1"
label values KID_1 KID_1
lab var KID_2 "Indicator of child order 2"
lab def KID_2 .a "Unknown" .b "Does not apply" .c "Not included in survey" 0 "No child of order 2" 1 "Child of order 2"
label values KID_2 KID_2
lab var KID_3 "Indicator of child order 3"
lab def KID_3 .a "Unknown" .b "Does not apply" .c "Not included in survey" 0 "No child of order 3" 1 "Child of order 3"
label values KID_3 KID_3
lab var KID_4 "Indicator of child order 4"
lab def KID_4 .a "Unknown" .b "Does not apply" .c "Not included in survey" 0 "No child of order 4" 1 "Child of order 4"
label values KID_4 KID_4
lab var KID_5 "Indicator of child order 5"
lab def KID_5 .a "Unknown" .b "Does not apply" .c "Not included in survey" 0 "No child of order 5" 1 "Child of order 5"
label values KID_5 KID_5
lab var KID_6 "Indicator of child order 6"
lab def KID_6 .a "Unknown" .b "Does not apply" .c "Not included in survey" 0 "No child of order 6" 1 "Child of order 6"
label values KID_6 KID_6
lab var KID_7 "Indicator of child order 7"
lab def KID_7 .a "Unknown" .b "Does not apply" .c "Not included in survey" 0 "No child of order 7" 1 "Child of order 7"
label values KID_7 KID_7
lab var KID_8 "Indicator of child order 8"
lab def KID_8 .a "Unknown" .b "Does not apply" .c "Not included in survey" 0 "No child of order 8" 1 "Child of order 8"
label values KID_8 KID_8
lab var KID_9 "Indicator of child order 9"
lab def KID_9 .a "Unknown" .b "Does not apply" .c "Not included in survey" 0 "No child of order 9" 1 "Child of order 9"
label values KID_9 KID_9

*Year of birth of first child
replace KID_Y1=.b if KID_1==0
replace KID_M1=.b if KID_1==0
replace KID_Y1=.a if KID_Y1==.
lab var KID_Y1 "Year of birth of child 1"
lab def KID_Y1 .a "Unknown" .b "Does not apply" .c "Not included in survey"
label values KID_Y1 KID_Y1

*month of birth of first child
replace KID_M1=.a if KID_M1==.
lab var KID_M1 "Month of birth of child 1"
lab def KID_M1 .a "Unknown" .b "Does not apply" .c "Not included in survey" 1 "January" 2 "February" 3 "March" 4
"April" 5 "May" 6 "June" 7 "July" 8 "August" 9 "September" 10 "October" 11 "November" 12 "December" 21 "Winter at
the beginning of the year" 22 "Spring" 23 "Summer" 24 "Autumn" 25 "Winter at the end of the year"
label values KID_M1 KID_M1

*standardized months of births of children 1 - imputed
gen IKID_M1=KID_M1
replace IKID_M1=int(1+(2)*uniform()) if KID_M1==21
replace IKID_M1=int(3+(3)*uniform()) if KID_M1==22
replace IKID_M1=int(6+(3)*uniform()) if KID_M1==23
replace IKID_M1=int(9+(3)*uniform()) if KID_M1==24
replace IKID_M1=12 if KID_M1==25
replace IKID_M1=int(uniform()*12)+1 if KID_M1==.a
lab var IKID_M1 "Month of birth of child 1 and imputed months"

```

```
lab def IKID_M1 .a "Unknown" .b "Does not apply" .c "Not included in survey" 1 "January" 2 "February" 3 "March" 4
"April" 5 "May" 6 "June" 7 "July" 8 "August" 9 "September" 10 "October" 11 "November" 12 "December"
label values IKID_M1 IKID_M1
```

```
*total number of children
gen UNIKID=0
forvalues x=1/9 {
replace UNIKID=UNIKID+1 if KID_`x'>0
}
}
```

```
order ARID MONTH_S IMONTH_S YEAR_S SEX BORN_Y BORN_M IBORN_M COHORT* East UNIKID KID_1
KID_2 KID_3 KID_4 KID_5 KID_6 KID_7 KID_8 KID_9 KID_Y1 KID_M1 IKID_M1
sort ARID
save papergerm,replace
```

***2. Part: Selection

```
*Selection of women, cohorts 1930-1969, WEST without Berlin
keep if East==0
keep if SEX==2
keep if COHORT!=-1
save papergermWEST,replace
```

```
*Observed number of first births - 2353 first births (2.3.)
tab KID_1,m
```

```
*Selection of women, cohorts 1930-1969, EAST without Berlin
clear
use papergerm
keep if East==1
keep if SEX==2
keep if COHORT!=-1
save papergermEast,replace
```

***3. Part Analysis - Fertility Behaviour in Western Germany

```
**Survival curves to first birth- Kaplan Maier 3.1.**
* Western Germany without Berlin
```

```
clear
use papergermWEST
```

```
**Exclude births from 2005 - Interview year****
replace KID_1=0 if KID_Y1==2005
```

```
** TIME VARIABLE UNWEIGHTED****
generate TIME01=.
replace TIME01= (KID_Y1-(BORN_Y+15))*12+IKID_M1-IBORN_M if KID_1==1
replace TIME01= (2004-(BORN_Y+15)) *12+12-IBORN_M if KID_1==0
```

```
drop if TIME01<0
stset TIME01, fail(KID_1) id(ARID) exit(time 360)
```

```
*Figure unweighted - Figure 2 - 3.1.1.
sts graph, by (COHORT)
save papergermWESTUNWGT, replace
```

```
*How does weighting effect the results? 3.1.2.
clear
use papergermWEST
```

```
**Exclude births from 2005 - Interview year****
replace KID_1=0 if KID_Y1==2005
```

```
** TIME VARIABLE WEIGHTED****
generate TIME01=.
replace TIME01= (KID_Y1-(BORN_Y+15))*12+IKID_M1-IBORN_M if KID_1==1
replace TIME01= (2004-(BORN_Y+15)) *12+12-IBORN_M if KID_1==0
```

```
drop if TIME01<0
stset TIME01 [pweight=PERSWGT],id(ARID) failure (KID_1) exit(time 360)
```

```

*Figure weighted
sts graph, by (COHORT)
save papergermWESTWGT, replace

*Is the difference in childlessness between cohorts significant? 3.1.3.
*Figur 4
use papergermWESTUNWGT
sts graph, by (COHORT) ci

*Are the results affected by East-West-German migrants? 3.1.4.
*East Germany with Berlin
clear
use papergermEast

*Weighted estimates
**Exclude births from 2005 - Interview year****
replace KID_1=0 if KID_Y1==2005

** TIME VARIABLE WEIGHTED****
generate TIME01=
replace TIME01= (KID_Y1-(BORN_Y+15))*12+KID_M1-IBORN_M if KID_1==1
replace TIME01= (2004-(BORN_Y+15)) *12+12-IBORN_M if KID_1==0

drop if TIME01<0
stset TIME01 [pweight=PERSWGT],id(ARID) failure (KID_1) exit(time 360)

*Figure 5(weighted)
sts graph, by (COHORT)

*Childlessness in GGS and microcensus 2008 3.2.
*Share of childlessness - women WEST
clear
use papergermWEST

**Exclude births from 2005 - Interview year****
replace KID_1=0 if KID_Y1==2005

*childless women
gen CHLESS=-1
replace CHLESS=1 if KID_1==0

*Figure 6/Table 1

*share of childless western women - unweighted
tab COHORT2 CHLESS,nof row

*share of childless western women - weighted
svyset [pweight=PERSWGT]
svy: tab CHLESS COHORT2

***3.3. Total number of children in GGS and vital statistics
** Figure 7/Table 2

*AVERAGE NUMBER OF CHILDREN unweighted
mean UNIKID, over(COHORT2)

*AVERAGE NUMBER OF CHILDREN weighted
svyset [pweight=PERSWGT]
svy: mean UNIKID,over(COHORT2)

*4. What are the reasons for the bias?
*Table 3a Distribution of sample size by cohort groups

clear
use papergerm
keep if COHORT!=-1
tab COHORT

*Table 3b Socio-demographic characteristics by cohort groups
*This table was generated based on the national GGS

```