Interviewer effects and the reliability of retrospective event histories within the German Generations and Gender Survey

Extended abstract

Almost all European countries are characterized by an ageing society, low birth rates, considerable childlessness and changing family and partnership structures. Against this background the Population Activities Unit (PAU) within the United Nations Economic Commission for Europe (UNECE) has conceptualized the Generations and Gender Survey (GGS) as one pillar of the Generations and Gender Programme (GGP). Its main objective is to initiate analyses with particular attention given to relationships between children and parents (generations) and relationships between partners (gender) in order to contribute to the understanding of the determinants of demographic choices at the individual level, thereby achieving a better understanding of the causal mechanisms that underlie demographic change.

There is wide agreement among social scientists that a better understanding of demographic behaviour – especially of family formation – is based on the life-course approach (for an overview on life-course research, see e.g. (Elder Jr. 2009; Mayer 2009). Under this approach, one looks at family and fertility behaviour as processes that evolve interdependently with each other and with other processes in an individual's life course. Also cultural frames as well as institutional and structural settings play an important role in life-course research by relating micro-, macro- and meso-levels of analysis (Vikat, et al. 2007: 394). There are two main ways to undertake life-course research. One possibility is a prospective approach, which implicates a panel design of the survey and the inclusion of questions about expectations and intentions in the questionnaire. The other possibility is asking retrospective questions to cover a period of an individual's lifetime (Blossfeld, Huinink 2001: 10f.). The first wave questionnaire of the GGS collects retrospective information on partnerships, fertility, the parental home, and home-leaving.

The importance of collecting retrospective data that cover a long period of an individual's life is counterbalanced by questions about the quality and reliability of such data. Despite the crucial importance of the accuracy of such data for demographic research and policy decision-making, there have been few recent evaluations of the quality of fertility or partnership histories in developed countries. "Landmark" events like the birth of children or marriage are commonly regarded as reliable memories and therefore are assumed to be applicable retrospectively without much decrement in data quality (Swicegood et al. 1984; Wu et al. 2001; Hayford, Morgan 2008). However, more recent studies suggest that even fertility information derived from survey data might be subject to various types of distortions (Murphy 2009).

An inevitable difficulty in the validation of retrospective data is the availability of an objective source of information against which to compare survey responses. One possibility to validate retrospective information is the comparison with data which has been produced over life time like diary entries. If surveys containing identical retrospective questions are repeated, it is also possible to compare experiences of different cohorts (von Babka Gostomski, Hartmann 1997: 131). Another possibility is to compare reports from retrospective and panel observations on the same individuals (Peters 1988; Teitler et al. 2006). If events in life-course are generating a certain state like being childless or being without a partner, then it is also possible to validate this information on an aggregate level with official statistics.

The possibilities for extensive validation of the fertility and union formation histories in the German GGS were quite limited, due to the lack of comparable surveys or vital statistics. This

has changed in 2008 when for the first time the German microcensus contains questions on childlessness and the total number of children of women (Statistisches Bundesamt 2009; Pötzsch 2010). This now allows comparisons on this important part of the German GGS. Recent research on this basis shows various types of distortions: Fertility of the older GGS-cohorts is too low, while it is too high for the younger cohorts. For the transition to partnership there is a similar bias with a too low share of married women in the older cohorts (Kreyenfeld et al. 2010; Kreyenfeld et al. 2011).

Based on these initial research results this paper addresses three issues: First, the identification of the main distortions in the fertility history and in marital status and the identification of the most affected groups by making further differential analyses. Second, the paper provides possible reasons for this bias particularly with regard to the quality and reliability of retrospective data. As a third issue this paper discusses possible solutions for further handling of the German GGS.

For the validation of the fertility histories and marital status in the German GGS the paper focuses on three possible sources of distortions: the sampling procedure, the instrument and the interview situation. In Germany two GGS-samples have been realized by using different sample designs: the basic survey of German speaking population by using a random route design, a second supplementary survey of Turkish nationals living in Germany by using a multi-stage sampling procedure based on the foreigner and population register. It is therefore possible to analyse the effects of different sampling methods on the bias in fertility histories and marital status.

Especially the retrospective part seems to have been too difficult for some of the respondents or interviewers. The paper therefore addresses the question if information based on fertility history and marital status is entirely or only partly distorted. Differential analyses show that in cases in which the existence of children or partners has been mentioned, the results are more or less in line with the results of the micro-census 2008, see e.g. the results of the parity distribution of live births by birth years (Naderi et al. 2009: 16, Sauer et al. 2012). Our basic hypothesis is that one of the main reasons for the problems of the German GGS is in the interaction of interviewer and respondent (and vice versa). On the one hand this includes the reaction to difficult questions. Respondents as well as interviewers might have decided to skip difficult parts of the questionnaire by refusing an answer or by answering “no”. These refusals and invalid answers again might have been wrongly coded as “no non-residential children” or “no former partner, at any time”. There are two hints which are in line with the presumption of wrong coding. In the German GGS there are nearly no refusals and invalid answers in the household grid and the module on non-residential children, while in the other GGS-countries the share of this item non-response is notably higher. Furthermore, in some cases in which no children or no partner are coded in the fertility and partner history, suddenly children or partners have been mentioned in some other modules of the questionnaire (e.g. health or support in the household). Furthermore in the fertility history a sort of „memory gap“ can be found, as respondents answer that they have children outside the household, but the questions concerning these children haven’t been answered (Sauer et al. 2012). On the other hand in any survey in which data emerge from the interaction between interviewer and respondent there is a risk of falsification by interviewers which could result in the contamination of the data. Several factors may affect the prevalence of interviewer falsification: design factors relating mainly to questionnaire characteristics like length, complexity and difficult questions, organizational factors like inadequate remuneration and training of the interviewers as well as external factors like bad weather or bad neighbourhood (Crespi 1945; Schräpler, Wagner 2003; Groves et al. 2009).

In the German GGS the interviewers work on a free-lance basis. This means they were paid per completed interview (regardless of the length of the interview) plus compensation for

travel expenses. In order to save time this could lead in some cases to the incentive to miscode the answers to questions to avoid follow-up questions. Several tests have been performed to investigate this problem further (Sauer et al. 2012). In order to check this item non-response problem on the part of the interviewers this paper shows the results of an extensive interviewer control taking into account several kind of “learning effects” to shorten parts of the interview.

Taken as a whole our findings contribute to the knowledge and understanding of the problems and challenges of retrospective data on event histories in the German GGS with particular attention given to the interaction between interviewer and respondent. These results are used to provide recommendations for further handling with the GGS.

References

Babka Gostomski, Christian von; Hartmann Josef (1997): Zur Problematik von Retrospektivbefragungen. In: Kopp, Johannes (Hg.): Methodische Probleme der Familienforschung. Zu den praktischen Schwierigkeiten bei der Durchführung einer empirischen Untersuchung. Frankfurt, New York: Campus, S. 113–146.

Blossfeld, Hans-Peter; Huinink, Johannes (2001): Lebensverlaufsforschung als sozialwissenschaftliche Forschungsperspektive. Themen, Konzepte, Methoden und Probleme. In: BIOS, Jg. 14, H. 2, S. 5–31.

Crespi, L. P. (1945): The Cheater Problem in Polling. In: Public Opinion Quarterly, S. 431–445.

Elder Jr., Glen H. (2009): Perspectives on the Life Course. In: Heinz, Walter R.; Huinink, Johannes; Weymann, Ansgar (Hg.): The Life Course Reader. Individuals and Societies Across Time. Frankfurt, New York: Campus, S. 91–110.

Groves, Robert M.; Fowler, Floyd J.; Couper, Mick P.; Lepkowski, James M.; Singer, Eleanor; Tourangeau, Roger (2009): Survey Methodology. Second Edition. Hoboken: Wiley.

Hayford, Sarah R.; Morgan, S. Philip (2008): The Quality of Retrospective Data On Cohabitation. In: Demography, Vol. 45, Nu.1, S. 129–141.

Kreyenfeld, Michaela; Hornung, Anne; Kubisch, Karolin; Jaschinski, Ina (2010): Fertility and union histories from German GGS data: some critical reflections. Max Planck Institute for Demographic Research. Rostock. (MPIDR Working Papers, WP-2010-023).

Kreyenfeld, Michaela; Zeman, Krystof; Burkimsher, Marion; Jaschinski, Ina (2011): Fertility data for German-speaking countries. What is the potential? Where are the pitfalls? MPIDR Working Paper WP 2011-003. Herausgegeben von Max-Planck-Institut für demografische Forschung. Rostock.

Mayer, Karl Ulrich (2009): New Directions in Life Course Research. In: Annual Review of Sociology, H. 35, S. 413–433.

Murphy, Michael (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. In: Population Studies, Vol. 63, No.2, S. 115–133.

Naderi, Robert; Dorbritz, Jürgen; Ruckdeschel, Kerstin (2009): Einleitung - Der Generations and Gender Survey in Deutschland: Zielsetzung, Verortung, Einschränkungen und Potenziale. In: Zeitschrift für Bevölkerungswissenschaft, 34 (1-2): 5–30.

Peters, H. Elizabeth (1988): Retrospective Versus Panel Data in Analyzing Lifecycle Events. In: The Journal of Human Resources, Vol. 23, No. 4, S. 488–513.

Pötzsch, Olga (2010): Kohortenfertilität: Ein Vergleich der Ergebnisse der amtlichen Geburtenstatistik und der Mikrozensuserhebung 2008. In: Comparative Populations Studies - Zeitschrift für Bevölkerungswissenschaft, Jg. 35, H. 1, S. 165–184.

Sauer, Lenore; Ruckdeschel, Kerstin; Naderi, Robert (2012): Reliability of retrospective event histories within the German Generations and Gender Survey. BiB Working Paper 01/2012, Bundesinstitut für Bevölkerungsforschung: Wiesbaden, available online: http://www.bib-demografie.de/cln_099/nn_2295088/SharedDocs/Publikationen/DE/Download/BiB__Working__Paper/GGS.html.

Schräpler, Jörg-Peter; Wagner, Gert G. (2003): Identification, Characteristics and Impact of Faked Interviews in Surveys. An Analysis by means of genuine fakes in raw data of SOEP, IZA Discussion Paper No. 969. Bonn: IZA.

Statistisches Bundesamt (2009): Mikrozensus 2008: Neue Daten zur Kinderlosigkeit in Deutschland. Begleitmaterial zur Pressekonferenz am 29. Juli 2009 in Berlin. Wiesbaden: Statistisches Bundesamt.

Swicegood, C.; Gray, S.; Morgan, Philip; Rindfuss, Ronald R. (1984): Measurement and Replication: Evaluating the Consistency of Eight U.S. Fertility Surveys. In: Demography, Vol. 21, S. 15.

Teitler, J. O.; Reichman, N. E.; Koball, H. (2006): Contemporaneous Versus Retrospective Reports of Cohabitation in the Fragile Families Survey. In: Journal of Marriage and Family, Jg. 68, H. 2, S. 469–477. Online verfügbar unter http://dx.doi.org/10.1111/j.1741-3737.2006.00265.x.

Vikat, Andres; et al. (2007): Generations and Gender Survey (GGS): Towards a better understanding of relationships and processes in the life course. In: Demographic Research, 17 (14), S. 389–440. Online verfügbar unter http://www.demographic-research.org/volumes/vol17/default.htm.

Wu, Lawrence L.; Martin, Steven P.; Long, Daniel A. (2001): Comparing Data Quality of Fertility and First Sexual Intercourse Histories. In: Journal of Human Resources, Jg. 36, S. 520–555.