

Technical note

THE DANISH GENERATIONS **AND GENDER SURVEY 2020:** **DATA COLLECTION, ACCESS** **AND QUALITY OF THE DATA**

Emil Simonsen, ROCKWOOL Foundation

Peter Fallesen, ROCKWOOL Foundation & Stockholm University

Lisbeth Loft, Loft Group & University of Copenhagen

Laust Hvas Mortensen, Statistics Denmark & University of Copenhagen

The Danish Generations and Gender Survey 2020:

Data collection, access and quality of the data

Emil Simonsen, ROCKWOOL Foundation

Peter Fallesen, ROCKWOOL Foundation & Stockholm University

Lisbeth Loft, Loft Group & University of Copenhagen

Laust Hvas Mortensen, Statistics Denmark & University of Copenhagen

Table of Contents

1. Introduction	3
2. Data collection strategy	3
Sampling frame.....	3
Response rates	4
3. Data access.....	4
Types of access	4
Linkage to administrative registry data from Statistics Denmark	5
4. Representativeness.....	5
Representativeness: Core demographic characteristics	5
Representativeness: Education level and family demographic characteristics	7
5. Non-response.....	8
Non-response. core demographic characteristics.....	8
Non-response. Education level and family demographic characteristics	9
6. Consistency between Danish GGS responses and registry data.....	11
Consistency in relationship status.....	11
Consistency in Cohabiting history	12
Consistency in the number of children living the household.....	13
7. Conclusion	14
Appendix 1. Observed characteristics among invited individuals measured against the total Danish population within age range. Measured in percentages.....	15
Appendix 2. GGS representativeness on observed characteristics measured against the total Danish population within age range. Measured in percentages.....	16
Appendix 3. GGS non-response on observed characteristics measured against the total Danish population within age range. Measured in percentages.....	18
Appendix 3. Consistency between Danish GGS responses and Danish registry data. Measured in percentages.....	20

1. Introduction

The Danish Gender and Generations Survey (GGS) 2020 was funded and coordinated by the ROCKWOOL Foundation¹. The survey is the first Danish GGS installment in the international series of panel surveys, administered by the Gender and Generations Programme (GGP)². The GGP was initiated in 2000 and managed by the Population Unit of the United Nations Economic Commission for Europe (UNECE). Since 2009 the GGP has been managed by the Netherlands Interdisciplinary Demographic Institute (NIDI).

Internationally, the first set of GGS was collected in 2004, and at least one wave has been collected in 24 countries (20 European, 4 non-European). The GGS is the successor to the Fertility and Family Survey (FFS) which was likewise administered by UNECE and collected for 24 countries. FFS contains survey questions qualitatively similar to those of the GGS, making it possible to obtain longer time-series by using both the FFS and the GGS where available.

The Danish GGS questionnaire contains information concerning family, gender relations and life trajectories of individuals, and of particular novelty is the added information regarding fertility intentions.

2. Data collection strategy

The Danish GGS was collected as a web-based survey by Statistics Denmark between March 15th and June 10th 2021. Participants were invited through e-Boks, which, at the time of data collection, was the secured web portal for communication between Danish residents and Danish authorities. e-Boks came with its own smart phone app and SMS notification, and all invited individuals were contacted a total of three times or until participation in the survey.

Incentives in the form of lottery prizes were offered to all invited individuals if they participated. Participants had the opportunity to win one of two prizes of 5000Dkr. (roughly €667) and one of five prizes of 2,000Dkr. (roughly €267). Incentives of this kind has been shown to have a significantly positive effect on response rates and accuracy³.

Sampling frame

Due to the high quality of Danish population registers, is possible for Statistics Denmark to draw exact population samples based on prespecified demographic measures. The target population for the Danish GGS was legal residents of Denmark age 18 to 49. At the time of the of the data collection the total pool of legal residents of Denmark within the age range of 18 to 49 was 2,359,801 individuals.

A total of 42,116 individuals were invited to participate in the Danish GGS, and comparison between the all Danish legal residents age 18 to 49 and the invited sample with regard to observed demographic characteristics is available in Appendix 1. The sample of invited individuals was the result of two separate subsamples, of which the second subsample was drawn to increase the overall response rate and to correct for a low response rate among males.

- Subsample 1. The first subsample consisted of 18,060 individuals randomly drawn among all Danish legal residents age 18 to 49.

¹ <https://www.rockwoolfonden.dk/en/>

² See <https://www.ggp-i.org> and <https://www.ggp-i.org/about>.

³ Bonke, Jens, and Peter Fallesen. 2010. " The Impact of Incentives and Interview Methods on Response Quantity and Quality in Diary- And Booklet-Based Surveys." *Survey Research Methods*, 4 (2): 91-101.

- Subsample 2. The second subsample consisted of 24,056 individuals randomly drawn among all Danish legal residents age 18 to 49 (excluding those already invited in subsample 1). This second subsample was stratified with 2/3 males and 1/3 females, due to a low response rates among males in the first subsample.

Response rates

A total of 10,268 responses were collected resulting in an overall response rate of 24 pct. Among the 10,268 collected responses, 7,275 (71 pct.) were complete responses and 2,993 (29 pct.) were partly complete responses. Counting only the 7,275 complete responses, the response rate equals 17 pct. Taken together, the data collection strategy resulted in an *overall response rate* of 24 pct. and a *completed response rate* of 17 pct.

The response rate of the Danish GGS was significantly lower than expected, which is likely due to a combination of the long length of the questionnaire, generally declining survey response rates, increased competition for responses among surveyors, and factors related to the Covid-19 pandemic.

3. Data access

The Danish GGS questionnaire consists of a total of eight sections including demographics (DEM), lifehistories (LHI), household (HHDI), generation (GEN), wellbeing (WEL), work (WRK), income (INC) and attitudes (ATT). The full questionnaire with all eight sections is available for download together with the dataset.

Types of access

Depending on the type of dataset the user is interested in, there are two ways to access the Danish GGS data:

- A. For users only interested in the survey data: a single dataset can be downloaded from the GGP website:
 - *Generations and Gender Survey 2020 Denmark Wave 1*
 - Contains all survey responses, both those completed and those partially completed (N = 10,268).
 - The full list of variables and corresponding coding of answers are available in the questionnaire. Thus, the questionnaire doubles as codebook.
 - Individual level population weights are provided in the variable “pervgt”. It is strongly recommended to apply the population weights when using the Danish GGS. See also section on Representativeness.
- B. For users interested in linking the data to Statistics Denmark’s registry data: data must be obtained from the Danish National Archives. See also section on Linkage to registry data from Statistics Denmark. A total of three datasets are available at the Danish National Archives:
 - *rffggp2020_svar*
 - Contains all survey responses, both those completed and those partially completed (N = 10,268).
 - The full list of variables and corresponding coding of answers are available in the questionnaire. Thus, the questionnaire doubles as codebook.
 - Individual level population weights are provided in the variable “pervgt”. It is strongly recommended to apply this population weight when using the Danish GGS 2020. See also section on Representativeness.
 - *rffggp2020_stik*
 - Contains all invited participants (N = 42,116).
 - *rffggp2020_pop*

- Contains the full population of 18–49-year-old legal residents of Denmark at the time the sample was drawn (N = 2,359,746)

Linkage to registry data from Statistics Denmark

The Danish GGS stored at the Danish National Archives includes a key-file necessary to link the Danish GGS data to Danish registry data maintained by Statistics Denmark. The key-file makes it possible for Statistics Denmark to link unique identifiers in the Danish GGS dataset with unique identifiers within the Danish registry data. It is a standardized process and only Statistics Denmark are able and allowed to undertake this linking process.

Only researchers affiliated with Danish universities or Danish private scientific organizations can obtain access to registry data through requests made to Statistics Denmark and (depending on the project) the Danish National Archives.

All rules and regulations guarding access and work with survey and registry data apply, and is the responsibility of the researcher(s) acquiring access to the data. Any fees associated with obtaining, linking and storing data are at the expense of the researcher(s) or research project requesting access.

4. Representativeness

The sampling frame of the Danish GGS was created in order to match the target population of legal residents of Denmark age 18 to 49 (total population). The resulting Danish GGS dataset of 10,268 respondents does exhibit some observed issues with regard to representativeness when compared to the total population. Comparisons were also made between only the complete Danish GGS responses (7,275 respondents), only the partly complete Danish GGS responses (2,993) and the total population.

Representativeness of the Danish GGS was assessed descriptively with regard to observable core demographic characteristics (age, gender, origin and geographical region), education level and family demographic characteristics (marital status, transition to parenthood and parity). Differences of 5 percentage points or greater between the GGS responses and the total population were considered a potential threat to the representativeness. All comparisons are available as tables in Appendix 2.

The observed issues of representativeness in the Danish GGS pertain to gender, origin and education level. Specifically, among the Danish GGS respondents we observed smaller shares of men, of persons with immigrant or descendant origin, and of persons with primary education as their highest completed educational level (and a larger share of persons with tertiary education).

To correct for observed differences between the Danish GGS respondents and the total population, a variable with individual-level population weights has been included with the Danish GGS dataset. It is recommended that users of the Danish GGS apply this variable (pervgt) when analyzing the data. Depending on the mode of analysis, an alternative method is to add control-variables for the observable characteristics known to be related to reduced representativeness when building a statistical model.

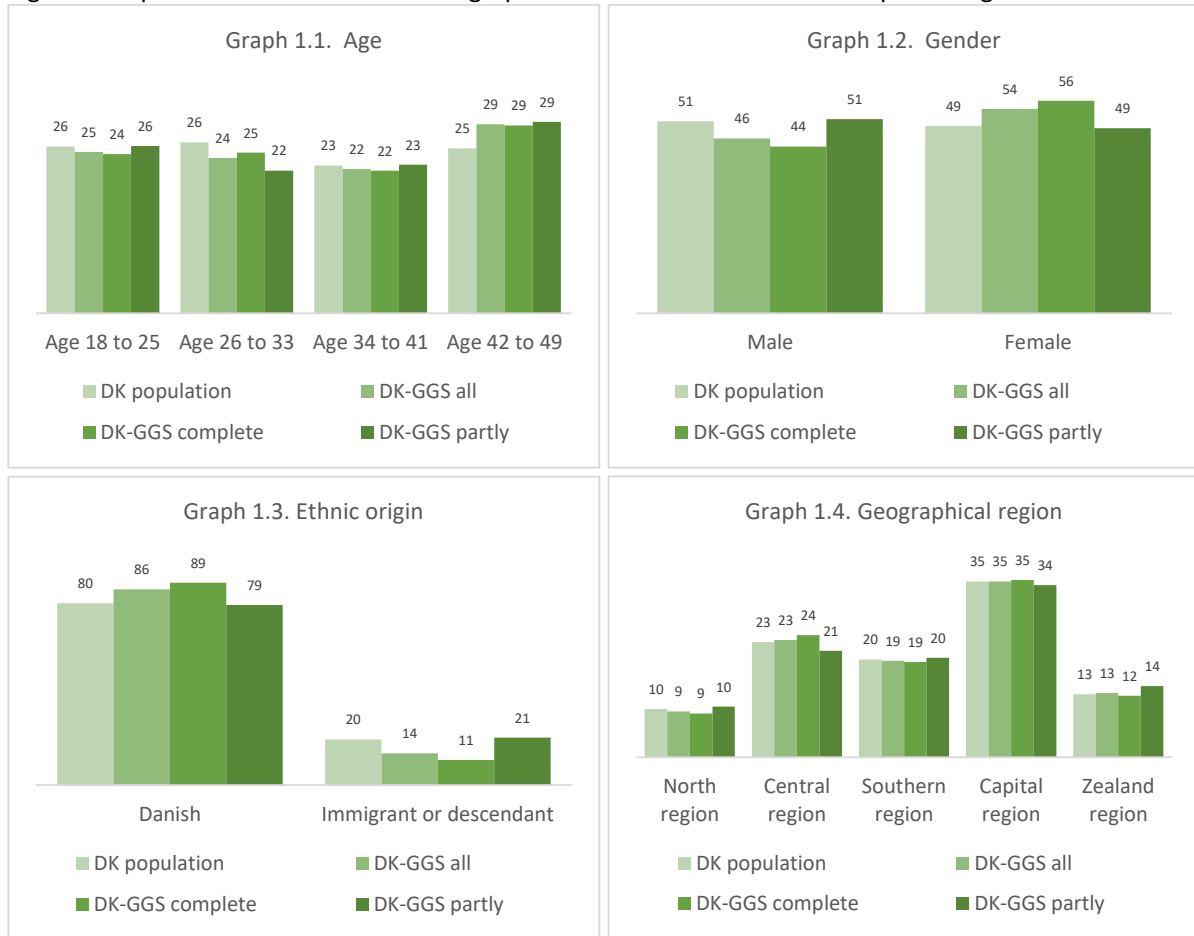
Representativeness: Core demographic characteristics

To investigate core demographic characteristics, differences with regard to age, gender, origin and geographical region were investigated.

Figure 1 provides an overview of the differences between the total population (DK population, N = 2,359,746), all respondents in the Danish GGS regardless of whether the respondents completed or only partially

completed the survey (DK-GGS all, n = 10,268), the respondents who completed the survey (DK-GGS complete, n = 7,275) and respondents who partly completed the survey (DK-GGS partly, n = 2993).

Figure 1. Representativeness. Core demographic characteristics. Measured in percentages.



Note: All characteristics are measured by December 31st 2020. DK population is equal to the total Danish population of 18-49-year-olds at the time of the survey (N = 2,359,746). DK-GGS all is equal to all respondents in the Danish GGS (n = 10,268). DK-GGS complete is equal to respondents who completed the survey (n = 7,275). DK-GGS partly is equal to respondents who started but did not complete the survey (n = 2,993). Immigrant ethnic origin is defined as born outside Denmark and none of the parents are Danish citizens and born in Denmark. Descendant ethnic origin is defined as born in Denmark and none of the parents are Danish citizens and born in Denmark. Danish ethnic origin is defined as not being an immigrant or a descendant.

Examining graph 1.1 in figure 1, no sizeable differences with regard to age was identified. There appear to be, however, a slight trend towards a larger share of Danish GGS respondents in the age category of 42 to 49 in comparison to the total population (29 pct. vs. 25 pct.). Age was thoroughly inspected at single age-points (not shown), and the slight trend towards a larger share of Danish GGS respondents in the age category of 42 to 49 is primarily driven by minor differences (<2 pct.) at age 47 to 49.

Reviewing graph 1.2 in figure 1, some differences were identified with regard to gender. Compared to the total population, there are fewer males among the Danish GGS respondents. Whereas the total population holds 51 pct. males, only 46 pct. of all Danish GGS respondents are male, and among complete Danish GGS responses this figure is 44 pct.

Similar to gender, differences were also detected with regard to ethnic origin (graph 1.3 in figure 1). A larger share of the Danish GGS respondents are of Danish origin relative to the total population (86 pct. vs. 80 pct.). This difference increases in magnitude when the comparison only involves complete Danish GGS responses

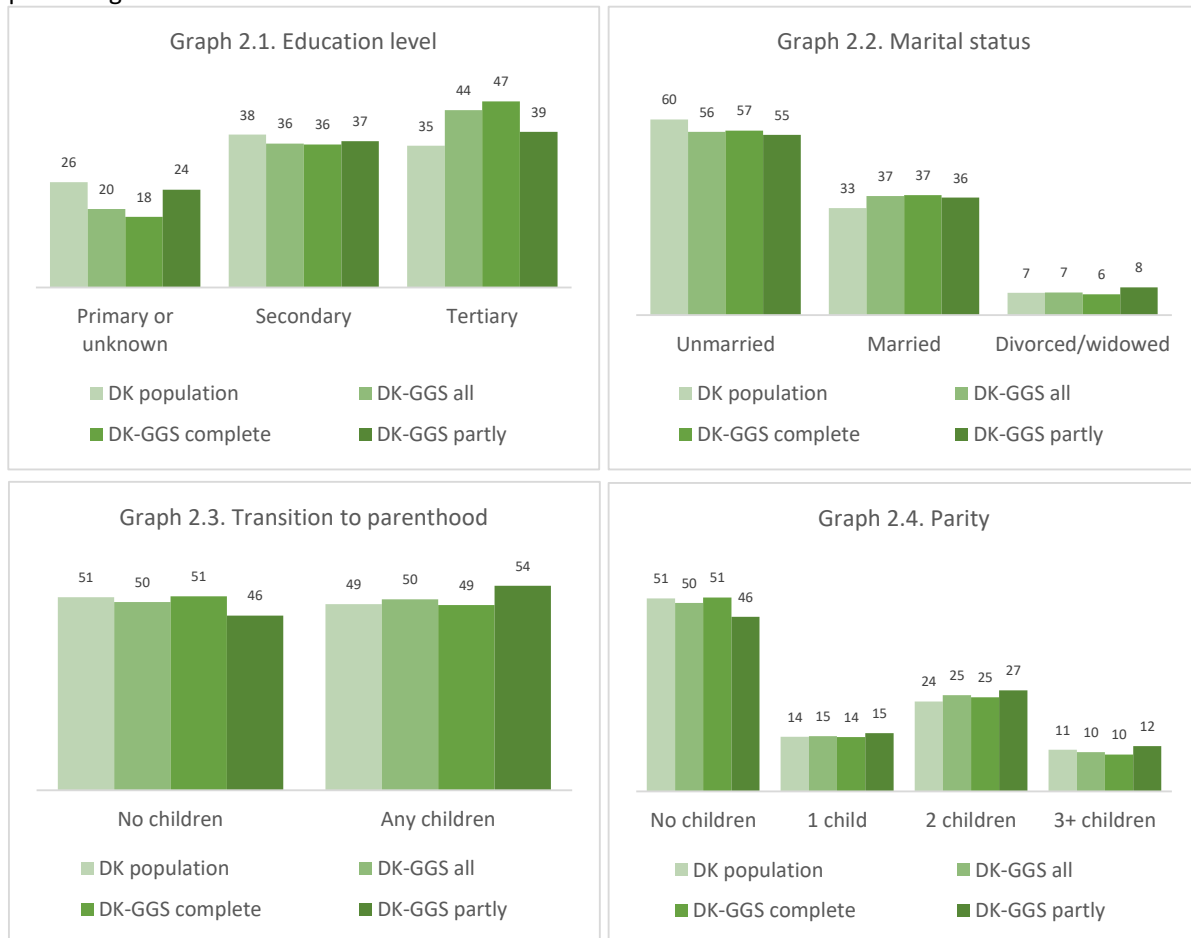
(80 pct. vs. 89 pct.). Separate categories for immigrants and descendants were examined (not shown) and no indication was found of either category contributing more to the observed differences. Across all compared samples descendants account for less than 5 pct.

Denmark consists of five administrative geographical regions. Examining graph 1.4 in figure 1, no noticeable differences were observed with regard to geographical region of the respondents in the Danish GGS. This holds true both when comparing all GGS respondents and when comparing only the respondents with complete GGS responses to the total population.

Representativeness: Education level and family demographic characteristics

In addition to the four core demographic characteristics of age, gender, origin and region, differences with regard to education level, marital status, transition to parenthood and parity were examined. Figure 2 presents these differences. Again, the comparisons were done across the total population (DK population, N = 2,359,746), all respondents in the Danish GGS regardless of whether the respondents completed or only partially completed the survey (DK-GGS all, n = 10,268), the respondents who completed the survey (DK-GGS complete, n = 7,275) and respondents who partly completed the survey (DK-GGS partly, n = 2993).

Figure 2. Representativeness. Education level and family demographic characteristics. Measured in percentages.



Note: All characteristics are measured by December 31st 2020. DK population is equal to the total Danish population of 18-49-year-olds at the time of the survey (N = 2,359,746). DK-GGS all is equal to all respondents in the Danish GGS (n = 10,268). DK-GGS complete is equal to respondents who completed the survey (n = 7,275). DK-GGS partly is equal to respondents who started but did not complete the survey (n = 2,993). All characteristics, including transition to parenthood and parity, is measured for both men and women. Married marital status includes both married and registered partnerships. Divorced/widowed marital status includes both divorced and dissolved registered partnerships.

With regard to education level (graph 2.1 in figure 2) sizeable differences were identified. In comparison to the total population, a smaller share has primary education (26 pct. vs. 20 pct.) and a larger share has tertiary education (35 pct. vs. 44 pct.) among Danish GGS respondents. These difference increases in magnitude when restricting the Danish GGS sample to respondents who completed the survey. (26 pct. vs 18 pct. and 35 pct. vs. 47 pct. respectively).

Examining the family demographic characteristics of marital status, transition to parenthood and parity (also shown in figure 2), no notable differences were detected.

In sum, the representativeness analysis identified observed differences with regard to gender, ethnic origin and education level. Relative to the total population, none of the identified differences in the Danish GGS exceeds 12 percentage points across the three difference samples (all respondents, respondents that completed the survey, and respondents that partly completed the survey).

5. Non-response

The purpose of the non-response analysis is to assess if noticeable differences exists between those who responded to the Danish GGS and those who did not. Such comparisons can be a useful supplement to the representativeness analysis. Of the 42,116 individuals invited to participate in the Danish GGS, 10,268 individuals participated with a complete or partly complete response. The remaining 31,848 invited individuals who did not participate is considered the non-response sample.

As with representativeness, non-response in the Danish GGS was assessed descriptively and with regard to core demographic characteristics (age, gender, origin and geographical region) as well as education level and family demographic characteristics (marital status, transition to parenthood and parity). Only differences of 5 percentage points or greater are considered a potential issue. Corresponding tables are available in Appendix 3.

Generally, the differences identified in the non-response analysis mirrors those found in the representativeness analysis with regard to gender, ethnic origin and education level.

Non-response. Core demographic characteristics

Figure 3 presents differences in non-response for age, gender, origin and geographical region across those who were invited but did not provide any response (N = 31,848), all respondents in the Danish GGS regardless of whether the respondents completed or only partially completed the survey (DK-GGS all, N = 10,268), the respondents who completed the survey (DK-GGS complete, n = 7,275) and only the partly complete Danish GGS responses (2,993) and the total population.

Reviewing the graph 3.1 in figure 3, there is a smaller share in the 42 to 49 age category among the non-respondents when compared to Danish GGS respondents (24 pct. vs. 29 pct.), which mimics the slight age trend discussed with regard to representativeness.

The fact that the sampling frame included a second sample stratified by gender (2/3 males and 1/3 females) and that this effort did not translate into a gender-balanced Danish GGS sample, is evident when examining the graph 3.2 in figure 3. Among the non-respondents 61 pct. is male and a corresponding 39 pct. is female. Among those who responded to the survey, 46 pct. was male and 54 pct. was female. Only considering respondents who completed the Danish GGS, these figures were 44 pct. and 56 pct. respectively.

Graph 3.3 in figure 3 provides information on the distribution of ethnic origin. Among non-respondents the share with Danish origin is 78 pct. which is close to the observed population share of 80 pct. (not shown in

figure 3, but available in graph 1.3 in figure 1). As seen in the representative analysis, this share was noticeable larger among the Danish GGS respondents with 86 pct. among all Danish GGS respondents and 89 pct. among those who completed the survey.

No notable differences between Danish GGS non-respondents and those who responded to the Danish GGS were present with regard to geographical region (graph 3.4, figure 3).

Figure 3. Non-response. Core demographic characteristics. Measured in percentages.



Note: All characteristics are measured by December 31st 2020. DK-GGS-nonresponse is equal to GGS nonresponses (N = 31,848). DK-GGS all is equal to all respondents in the Danish GGS (N = 10,268). DK-GGS complete is equal to respondents who completed the survey (n = 7,275). DK-GGS partly is equal to respondents who started but did not complete the survey (n = 2,993). Immigrant ethnic origin is defined as born outside Denmark and none of the parents are Danish citizens and born in Denmark. Descendant ethnic origin is defined as born in Denmark and none of the parents are Danish citizens and born in Denmark. Danish ethnic origin is defined as the not being an immigrant or a descendant.

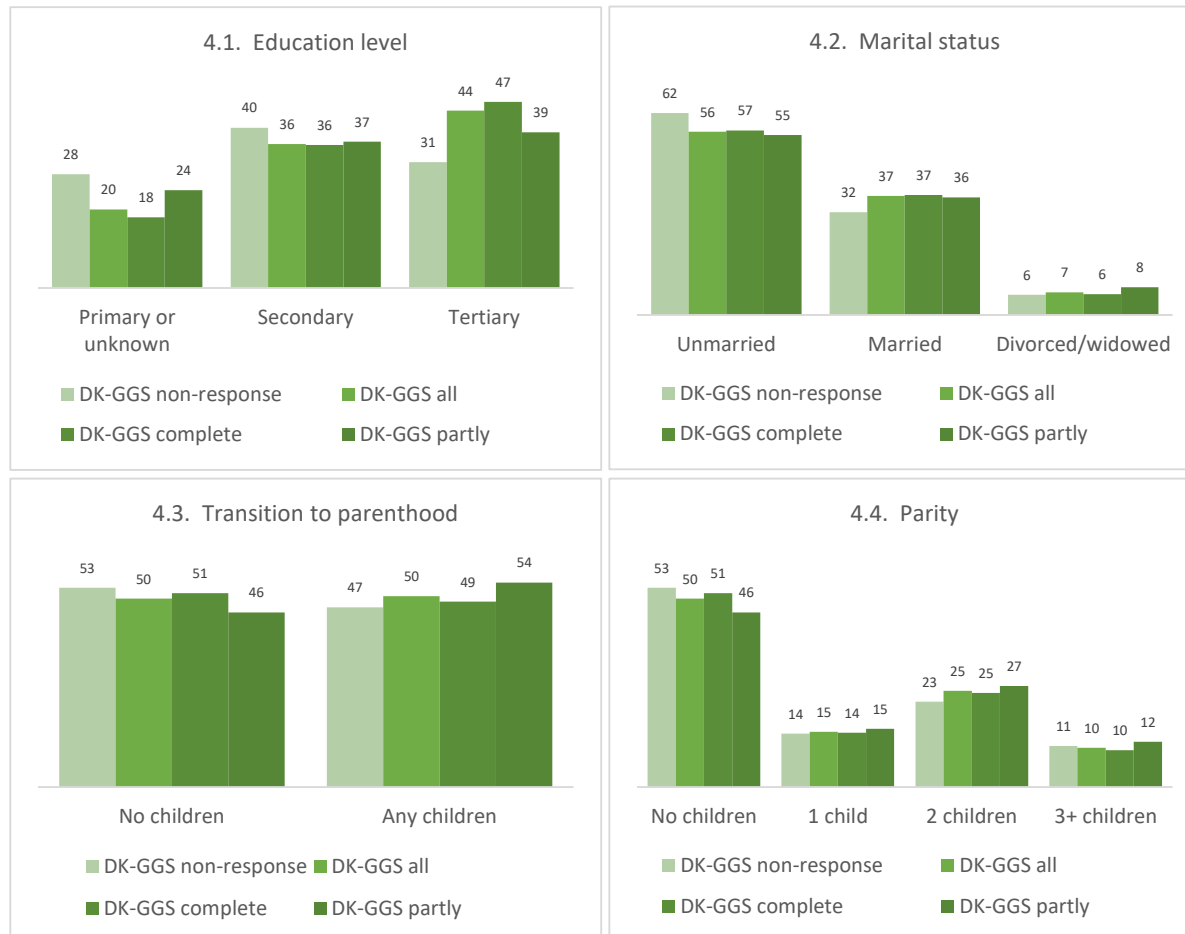
Non-response. Education level and family demographic characteristics

The four observed characteristics of education level, marital status, transition to parenthood and parity were also examined with regard to non-response, and is presented in figure 4.

Graph 4.1 in figure 4 shows that, in comparison to Danish GGS respondents, a larger share of non-respondents has primary education as their highest completed level of education (28 pct. vs. 20 pct.). Correspondingly, a smaller share of non-respondents has tertiary education as their highest completed level of education (31 pct.

vs. 44 pct.). As in the representativeness analysis, these differences in education level between non-respondents and Danish GGS respondents increase in magnitude when comparing only to respondents who completed the survey (28 pct. vs. 18 pct. and 31 pct. vs. 47 pct. respectively).

Figure 4. Non-response. Education level and family demographic characteristics. Measured in percentage



Note: All characteristics are measured by December 31st 2020. DK population is equal to the total Danish population of 18-49-year-olds at the time of the survey (N = 2,359,746). DK-GGS all is equal to all respondents in the Danish GGS (N=10,268). DK-GGS complete is equal to respondents who completed the survey (N = 7,275). DK-GGS partly is equal to respondents who started but did not complete the survey (n = 2,993). All characteristics including transition to parenthood and parity is measured for both men and women. Married marital status includes both married and registered partnerships. Divorced/widowed marital status includes both divorced and dissolved registered partnerships.

Examining marital status (graph 4.2 in figure 4) it is apparent that a larger share among non-respondents are unmarried (62 pct. vs. 56 pct.) which, to some degree, corresponded with a slightly smaller share of married non-respondents relative Danish GGS respondents. No observed sizeable differences with regard to the share of divorced or widowed individuals where detected across non-respondents and respondents.

With regard to transition to parenthood or parity (graph 4.3 and 4.4 in figure 4), sizeable differences were only observed between non-respondents and respondents who only partly completed the Danish GGS.

In sum non-respondents deviates from those that did respond to the Danish GGS on the same dimensions identified in the representativeness analysis, i.e., in terms of gender, education, and ethnic origin. Additionally, a larger share of non-respondents was unmarried.

6. Consistency between Danish GGS responses and registry data

For some questions in the Danish GGS, the information provided by the respondent can be compared to information available in the Danish registry data, which, depending on whether one is willing to assume that either the survey responses or the information contained in the registry data is correct, allows for a measure of the accuracy of the other data, or at least overlap between data sources. Given the GGS focus of family life, consistency in reports on relationship status, cohabitation history and number of children in the household was examined.

Under the assumption that people know whether they are living with a partner at time of answering the survey, the analysis of relationship status allowed us to assess the quality of cohabitation information in the registry data (with the caveat that cohabitation information is measured as of December 31st 2020, and the Danish GGS was collected during the spring/summer of 2021).

The overlap between cohabitation histories provided in the Danish GGS with those available from Danish registry data going back 10 years (or, in the case of young respondents or immigrant, going back to age 16 or year of migration).

Last, the overlap between reports in the Danish GGS and the Danish registry data with regard to number of children less than 16 years of age living in a household was also examined.

Consistency in relationship status

Figure 5 shows the results with regard to consistency for relationship status (cohabiting and single status). A corresponding table is available in Appendix 4.

Graph 5.1 in figure 5 shows the fraction reporting to be living in a cohabitating relationship initiated before 2021 while also being registered as such in the registry data⁴ by December 31st 2020, by age of the Danish GGS respondents. The accuracy of the registry data for cohabitating status appears to be very high, assuming respondents are reporting their relationship status correctly. Consistency across the two data sources ranges from 80 pct. among Danish GGS respondents age 18 to 20, to 96 pct. among Danish GGS respondents age 42 to 44. Add to that, that consistency as measured here should be biased downward, as dissolution of cohabitations could have occurred between December 31st 2020 and the day of the interview (spring/summer 2021).

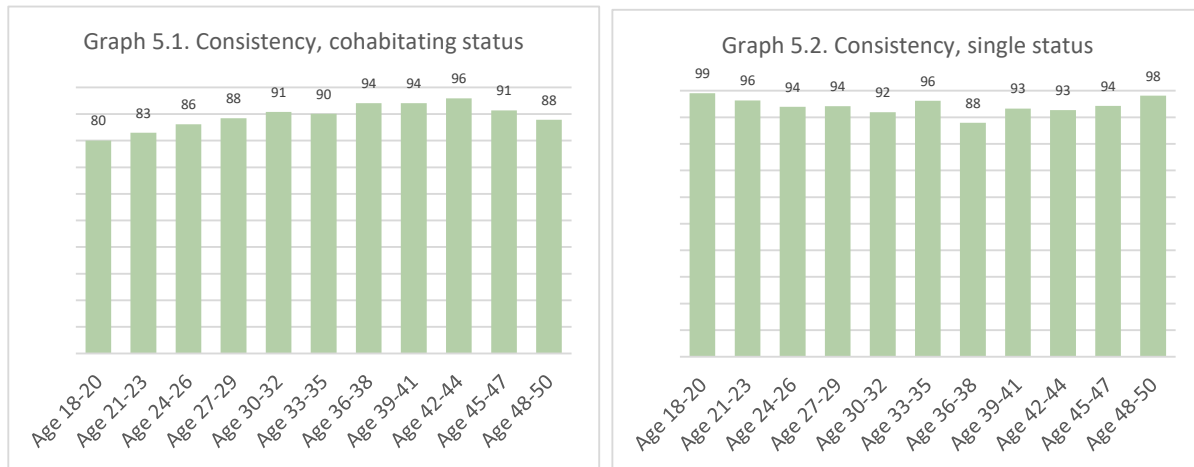
Graph 5.2. in figure 5 shows the fraction of Danish GGS respondents reporting to be single before 2021, while also being registered as such in the registry data⁵ by December 31st 2020, by age of the respondent. In this graph, those aged less than 25 registered as living at home with their parents are assumed to be single⁶. As with cohabiting status, the consistency, and thus accuracy, of single status is very high, ranging from 88 pct. among Danish GGS respondents age 36 to 38 to 99 pct. among Danish GGS respondents age 18 to 20.

⁴ Statistics Denmark's definition of a cohabiting couple in the registry data is two individuals, who share a dwelling, and either a) are unmarried but has a joint child; or b) are of opposite sex, within 15 years of age of each other, and not directly related.

⁵ Statistics Denmark's definition of a cohabiting couple in the registry data is two individuals, who share a dwelling, and either a) are unmarried but has a joint child; or b) are of opposite sex, within 15 years of age of each other, and not directly related.

⁶ Singles in the registry are defined as those that are not cohabiting and not married, with the exception for those less than 25 years old mentioned in the text.

Figure 5. Consistency. Relationship status. Measured in percentages.



Note: The analysis includes respondents who completed the survey and could be matched to the population registry the relevant year. For consistency in cohabitation status the analysis is further restricted to respondents who in the survey reported to be in a cohabiting union. For consistency in single status the analysis is further restricted to respondents who in the survey reported to be single. Individuals less than 25 years of age are considered to be single in the registry data if they live alone or at home with at least one parent. Age is measured on the day of the interview.

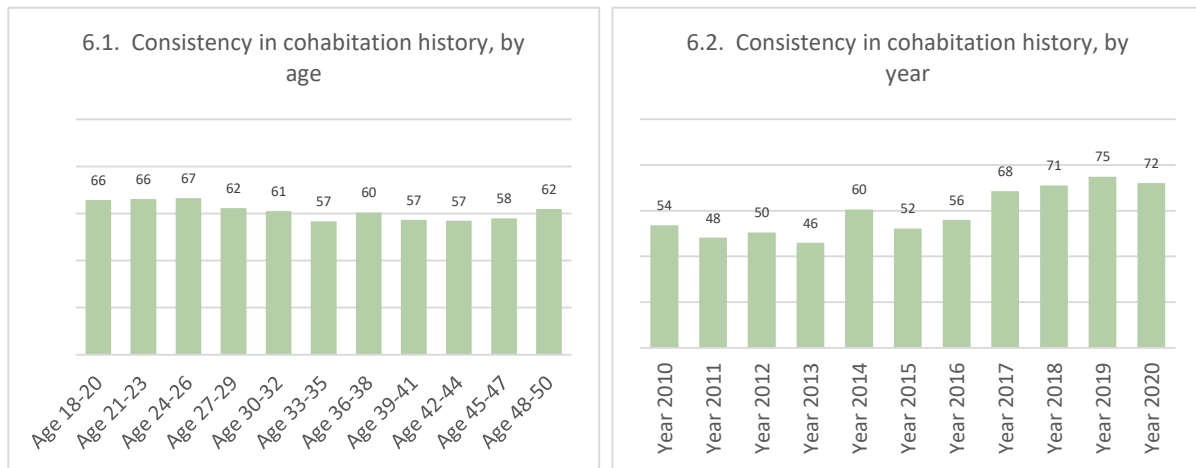
Consistency in cohabitation history

Figure 6 shows the results with regard to cohabitation history by age and year. A corresponding table is available in Appendix 4.

Graph 6.1 in figure 6 shows the fraction of those reporting to have initiated a cohabitation (regardless of whether it is still ongoing or not) in a given year (in the period 2010-2020) while also being registered as such in the same given year, by age of the respondents. Consistency with regard to cohabitation history by age appear moderate ranging from 57 pct. among Danish GGS respondents age 33-35, 39-41 and 42-44 to 67 pct. among Danish GGS respondents age 24-26. Reviewed in concert with the high accuracy with regard to cohabiting status (graph 5.1 in figure 5), this moderate level of consistency with regard to cohabitation history by age likely suggests, that the accuracy of the Danish GGS respondent reports in this case are fairly low.

Graph 6.2 in figure 6 provides the fraction of those reporting to have initiated a cohabitation (regardless of whether it is still ongoing or not) in a given year (in the period 2010-2020), while also being registered as such in the same given year, by year of the initiated cohabitation. Examining the consistency in cohabitation history by year, reveals that the consistency generally is lower for years further back in time, with the reverse being true for more recent years. The range for consistency in cohabitation history by year is from 46 pct. in 2013 to 75 pct. in 2019.

Figure 6. Consistency. Cohabitation history across age and years. Measured in percentages.



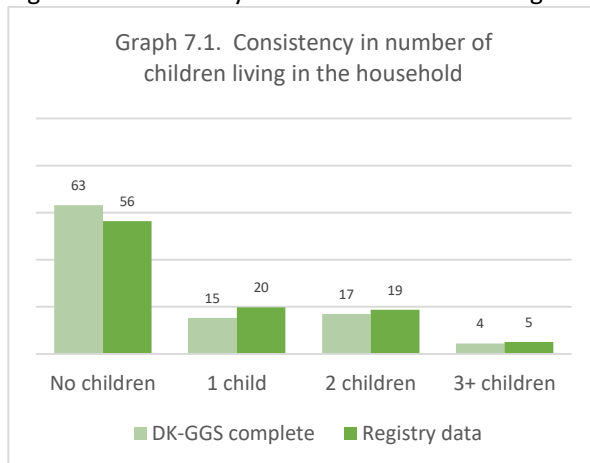
Note: The analysis includes respondents who completed the survey and could be matched to the population registry the relevant year. For consistency in cohabitation history, the share is calculated as the share of the total number of reported cohabitations. Age is measured on the day of the interview.

Consistency in the number of children living the household

In Figure 7, the consistency with regard to the number of children age 15 or less to be living in the household is presented. A corresponding table is available in Appendix 4.

Graph 7.1 in figure 7 shows that the accuracy of the registry data is fairly high, assuming the Danish GGS responses are correct. As can be seen, there appears to be an overrepresentation of those reporting to have no children living with them (63 pct.) when compared to the information from the registry (56 pct.), and an underrepresentation of those that report having one child living with them (15 pct. vs. 20 pct.). Still, the overall consistency with regard to number of children living in the household is considered to be fairly high.

Figure 7. Consistency. Number of children living in the household. Measured in percentages.



Note: DK-GGS complete is equal to respondents who completed the survey (N = 7,275). Registry data is equal to the corresponding records in the Danish administrative registry data. Children are age 15 or younger. In the registry data child age is measured per March 31st 2021.

In sum, the comparison between reports from Danish GGS respondents and records in the Danish registry data exhibited very high consistency with regard to relationship status (cohabiting and single status) indicating a high level of accuracy of the registry data records for these measures. Examining the consistency for cohabitation history across age and year, the findings suggest that the accuracy of the GGS respondent reports for cohabitation status may be fairly low. In addition, reviewing the consistency in cohabitation histories by year, indicated some level of recall error among the Danish GGS respondents for cohabiting relationships dating further back in time. The quality of the Danish registry data did not change across the observed period (2010 – 2020). Finally, the overall consistency with regard to number of children living in the household is considered to be fairly high.

7. Conclusion

The Danish GGS 2020 is the first Danish installment in the international series of panel surveys, GGS, administered by the GGP. The Danish GGS data provides researchers with the novel opportunity to analyze a wide range of demographic challenges that cannot be fully understood by the use of registry data alone. One such example, is the added information about fertility intentions.

The Danish GGS contains 10,268 responses, with 7,275 completed responses, sampled from all legal Danish residents age 18 to 49 corresponding to an overall response rate of 24.4 pct. and a complete response rate of 17.3 pct. .

Examinations of representativeness and non-response in the Danish GGS revealed that some issues pertaining to gender, ethnic origin and education level exist. Although representativeness in the Danish GGS is similar to other surveys, it is recommended to apply the provided individual-level population weights or control for observable characteristics when analyzing the data.

Consistency of the Danish GGS responses with records in the Danish registry data for contemporaneous variables for cohabiting status, single status, and the number of children in the household appears either very high or fairly high.

Appendix 1. Observed characteristics among invited individuals measured against the total Danish population within age range. Measured in percentages.

CORE DEMOGRAPHIC CHARACTERISTICS			
	DK population (A)	DK-GGS invited (B)	Difference (A - B)
Age			
Age 18 to 25	25,66	26,16	-0,50
Age 26 to 33	26,28	26,22	0,06
Age 34 to 41	22,72	22,41	0,31
Age 42 to 49	25,35	25,22	0,13
Gender			
Male	50,67	57,41	-6,74
Female	49,33	42,59	6,74
Ethnic origin			
Danish	79,88	80,07	-0,19
Immigrant or descendant	20,12	19,93	0,19
Geographical region			
North region	9,66	9,67	-0,01
Central region	23,02	23,06	-0,04
Southern region	19,56	19,59	-0,03
Capital region	35,13	34,91	0,22
Zealand region	12,63	12,78	-0,15
EDUCATION LEVEL AND FAMILY DEMOGRAPHIC CHARACTERISTICS			
	DK population (A)	DK-GGS invited (B)	Difference (A - B)
Education level			
Primary or unknown	26,37	26,29	0,08
Secondary	38,21	39,11	-0,90
Tertiary	35,42	34,60	0,82
Marital status			
Unmarried	60,19	60,71	-0,52
Married	32,92	32,86	0,06
Divorced/widowed	6,89	6,43	0,46
Transition to parenthood			
No children	50,90	51,87	-0,97
Any children	49,10	48,13	0,97
Parity			
No children	50,90	51,87	-0,97
1 child	14,38	14,21	0,17
2 children	23,72	23,21	0,51
3+ children	11,00	10,72	0,28

Appendix 2. GGS representativeness on observed characteristics measured against the total Danish population within age range. Measured in percentages.

CORE DEMOGRAPHIC CHARACTERISTICS							
	DK population (A)	DK-GGS all (B)	DK-GGS complete (C)	DK-GGS partly (D)	Difference 1 (A - B)	Difference 2 (A - C)	Difference 3 (A - D)
Age							
Age 18 to 25	25,66	24,83	24,47	25,73	0,83	1,19	-0,07
Age 26 to 33	26,28	23,90	24,70	21,95	2,38	1,58	4,33
Age 34 to 41	22,72	22,20	21,94	22,85	0,52	0,78	-0,13
Age 42 to 49	25,35	29,06	28,89	29,47	-3,71	-3,54	-4,12
Gender							
Male	50,67	46,08	43,96	51,22	4,59	6,71	-0,55
Female	49,33	53,92	56,04	48,78	-4,59	-6,71	0,55
Ethnic origin							
Danish	79,88	86,07	88,93	79,12	-6,19	-9,05	0,76
Immigrant or descendant	20,12	13,93	11,07	20,88	6,19	9,05	-0,76
Geographical region							
North region	9,66	9,16	8,77	10,12	0,50	0,89	-0,46
Central region	23,02	23,49	24,40	21,28	-0,47	-1,38	1,74
Southern region	19,56	19,29	19,04	19,91	0,27	0,52	-0,35
Capital region	35,13	35,18	35,48	34,45	-0,05	-0,35	0,68
Zealand region	12,63	12,87	12,32	14,23	-0,24	0,31	-1,60

EDUCATION LEVEL AND FAMILY DEMOGRAPHIC CHARACTERISTICS							
	DK population (A)	DK-GGS all (B)	DK-GGS complete (C)	DK-GGS partly (D)	Difference 1 (A - B)	Difference 2 (A - C)	Difference 3 (A - D)
Education level							
Primary or unknown	26,32	19,64	17,65	24,49	6,68	8,67	1,83
Secondary	38,21	36,02	35,79	36,59	2,19	2,42	1,62
Tertiary	35,42	44,33	46,56	38,92	-8,91	-11,14	-3,50
Marital status							
Unmarried	60,19	56,37	56,77	55,40	3,82	3,42	4,79
Married	32,92	36,65	36,87	36,12	-3,73	-3,95	-3,20
Divorced/widowed	6,89	6,98	6,36	8,49	-0,09	0,53	-1,6
Transition to parenthood							
No children	50,90	49,66	51,13	46,07	1,24	-0,23	4,83
Any children	49,10	50,34	48,87	53,93	-1,24	0,23	-4,83
Parity							
No children	50,90	49,66	51,13	46,07	1,24	-0,23	4,83
1 child	14,38	14,60	14,30	15,34	-0,22	0,08	-0,96
2 children	23,72	25,35	24,82	26,63	-1,63	-1,1	-2,91
3+ children	11,00	10,39	9,75	11,96	0,61	1,25	-0,96

Appendix 3. GGS non-response on observed characteristics measured against the total Danish population within age range. Measured in percentages.

CORE DEMOGRAPHIC CHARACTERISTICS							
	DK-GGS non-response (A)	DK-GGS all (B)	DK-GGS complete (C)	DK-GGS partly (D)	Difference 4 (A - B)	Difference 5 (A - C)	Difference 6 (A - D)
Age							
Age 18 to 25	26,59	24,83	24,47	25,73	1,76	2,12	0,86
Age 26 to 33	26,96	23,90	24,70	21,95	3,06	2,26	5,01
Age 34 to 41	22,47	22,20	21,94	22,85	0,27	0,53	-0,38
Age 42 to 49	23,98	29,06	28,89	29,47	-5,08	-4,91	-5,49
Gender							
Male	61,06	46,08	43,96	51,22	14,98	17,10	9,84
Female	38,94	53,92	56,04	48,78	-14,98	-17,10	-9,84
Ethnic origin							
Danish	78,14	86,07	88,93	79,12	-7,93	-10,79	-0,98
Immigrant or descendant	21,86	13,93	11,07	20,88	7,93	10,79	0,98
Geographical region							
North region	9,83	9,16	8,77	10,12	0,67	1,06	-0,29
Central region	22,92	23,49	24,40	21,28	-0,57	-1,48	1,64
Southern region	19,68	19,29	19,04	19,91	0,39	0,64	-0,23
Capital region	34,82	35,18	35,48	34,45	-0,36	-0,66	0,37
Zealand region	12,75	12,87	12,32	14,23	-0,12	0,43	-1,48

EDUCATION LEVEL AND FAMILY DEMOGRAPHIC CHARACTERISTICS							
	DK-GGS non-response (A)	DK-GGS all (B)	DK-GGS complete (C)	DK-GGS partly (D)	Difference 4 (A - B)	Difference 5 (A - C)	Difference 6 (A - D)
Education level							
Primary or unknown	28,43	19,64	17,65	24,49	8,79	10,78	3,94
Secondary	40,10	36,02	35,79	36,59	4,08	4,31	3,51
Tertiary	31,47	44,33	46,56	38,92	-12,86	-15,09	-7,45
Marital status							
Unmarried	62,12	56,37	56,77	55,40	5,75	5,35	6,72
Married	31,64	36,65	36,87	36,12	-5,01	-5,23	-4,48
Divorced/widowed	6,25	6,98	6,36	8,49	-0,73	-0,11	-2,24
Transition to parenthood							
No children	52,58	49,66	51,13	46,07	2,92	1,45	6,51
Any children	47,42	50,34	48,87	53,93	-2,92	-1,45	-6,51
Parity							
No children	52,58	49,66	51,13	46,07	2,92	1,45	6,51
1 child	14,08	14,60	14,30	15,34	-0,52	-0,22	-1,26
2 children	22,52	25,35	24,82	26,63	-2,83	-2,30	-4,11
3+ children	10,82	10,39	9,75	11,96	0,43	1,07	-1,14

Appendix 4. Consistency between Danish GGS responses and Danish registry data.
 Measured in percentages.

CONSISTENCY IN COHABITING STATUS	
Age 18-20	80,00
Age 21-23	82,94
Age 24-26	86,12
Age 27-29	88,45
Age 30-32	90,83
Age 33-35	90,06
Age 36-38	94,07
Age 39-41	94,12
Age 42-44	95,83
Age 45-47	91,30
Age 48-50	87,84

CONSISTENCY IN SINGLE STATUS	
Age 18-20	99,05
Age 21-23	96,27
Age 24-26	93,90
Age 27-29	94,21
Age 30-32	91,88
Age 33-35	96,15
Age 36-38	87,88
Age 39-41	93,23
Age 42-44	92,74
Age 45-47	94,29
Age 48-50	98,10

CONSISTENCY IN COHABITATION HISTORY BY AGE	
Age 18-20	65,71
Age 21-23	66,18
Age 24-26	66,58
Age 27-29	62,29
Age 30-32	60,97
Age 33-35	56,65
Age 36-38	60,39
Age 39-41	57,26
Age 42-44	56,86
Age 45-47	57,97
Age 48-50	61,90

CONSISTENCY IN COHABITATION HISTORY BY YEAR

Year 2010	53,55
Year 2011	48,18
Year 2012	50,39
Year 2013	45,96
Year 2014	60,49
Year 2015	52,15
Year 2016	55,98
Year 2017	68,47
Year 2018	70,99
Year 2019	74,90
Year 2020	72,11

CONSISTENCY IN NUMBER OF CHILDREN LIVING IN THE HOUSEHOLD

	DK-GGS complete (A)	Registry data (B)	Difference 7 (A - B)
No children	63,19	56,37	6,82
1 child	15,37	19,84	-4,47
2 children	16,96	18,74	-1,78
3+ children	4,48	5,05	-0,57
