**Codebook of the Panel Study of Social Integration in the Netherlands 1987-2009**

**Aat Liefbroer, Matthijs Kalmijn, Anne-Rigt Poortman**

**October 2015**

**Background**

PSIN is a longitudinal study that investigates the transition from youth to adulthood in the Netherlands. Originally, the study focused on the transition from the educational system to the labour market, how socialization into work occurred, and how the transition from the parental home to adult living arrangements occurred. The study started in 1985, funded by a large grant from the Ministry of Education and Science. The first three waves of data-collection (1987, 1989, 1991) were organized and conducted by a research team established at the Vrije Universiteit in Amsterdam. Since then various other institutions have kept the panel alive. Currently, there are 7 rounds of data with intervals of about 4 years between consecutive waves.

**Sample**

The PSIN was based on a stratified random sample of the population of young adults born in 1969, 1965, and 1961. First, a sample of municipalities was drawn. 19 clusters of municipalities were defined based on the the region of the Netherlands and a typology of urbanization. A cluster consisting of the three largest Dutch cities (Amsterdam, Rotterdam, The Hague) was added. Second, within each of the 20 clusters, a municipality was randomly chosen. In addition, a second municipality was chosen that could replace the first one, if the local authority refused to cooperate. Third, persons within municipalities were drawn randomly based on age and sex as provided by the registers. The number of interviews per municipality was based on the proportion of the population living in municipalities with a specific level of urbanization and region. In some clusters too few young adults would be living in the selected municipality to ensure the required number of interviews. In that case, the second municipality that was sampled from that cluster and was included in the study as well. Three municipalities refused to participate, 2 used too strict privacy criteria. In the end, samples were drawn in 25 municipalities.

**Funding and organization**

Various persons have been involved in the organization of the panel and various universities and institutes have contributed to its funding. The table below gives an overview.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **year** | **P.I.’s** | **Funding** |
| **wave 1** | 1987 | Gierveld, Dijkstra, Feij, Liefbroer | VU, Ministerie OCW |
| **wave 2** | 1989 | Gierveld, Dijkstra, Feij, Liefbroer | VU, Ministerie OCW |
| **wave 3** | 1991 | Gierveld, Dijkstra, Feij, Liefbroer | VU, Ministerie OCW |
| **wave 4** | 1995 | Kalmijn, Liefbroer, Weesie, Bernasco | UU, NIDI, NWO |
| **wave 5** | 1999/2000 | Jansen, Liefbroer & Kalmijn | NIDI, UU |
| **wave 6** | 2005/2006 | Jansen, Liefbroer & Kalmijn, Soons | NIDI , UU, Tilburg |
| **wave 7** | 2009/2010 | Liefbroer, Kalmijn, Poortman | NIDI, UU, Tilburg |

**Interviews and fieldwork**

Due to budget restrictions and other practical considerations, the mode has changed across waves. The table below gives an overview of the various methods that were used in each wave. In two occasions, an alternative mode was used for nonrespondents in the initial wave.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **year** | **main mode** | **additional mode**  **(for same respondents)** | **extra mode**  **(for nonreponse)** |
| **wave 1** | 1987 | face-to-face | self-administered during interview |  |
| **wave 2** | 1989 | self administered by mail |  |  |
| **wave 3** | 1991 | face-to-face | self-administered during interview |  |
| **wave 4** | 1995 | face-to-face | self-administered during interview | self-administered by mail |
| **wave 5** | 1999 / 2000 | phone | self-administered by mail | self-administered by mail |
| **wave 6** | 2005 / 2006 | self-administered by mail or internet |  |  |
| **wave 7** | 2009 / 2010 | self administered by mail |  |  |

In 1995, the PSIN was done in conjunction with the Households in the Netherlands survey (HIN95), which was a couple-oriented survey where both partners were interviewed. The PSIN respondent could be either the husband, the wife, or a single respondent. The data have been restructured so that the respondent in the file is the original PSIN respondent. The partner data are not included but can be given upon request.

**Respons**

The numbers of cases interviewed each year as well as respons rates are given below. In the first waves, respondents were not approached again if they did not participate in the previous wave. After wave 4, respondents who did not participate in a prior wave were approached again. About half of the original respondents are still participating in the panel 22 years after the first round. No refreshment samples were set up.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **respons (from previous wave)** | **respons (from earlier wave)** | **total respons** | **no respons** |
| wave 1 | 1775 |  | 1775 |  |
| wave 2 | 1419 |  | 1419 | 356 |
| wave 3 | 1257 |  | 1257 | 518 |
| wave 4 | 969 |  | 969 | 806 |
| wave 5 | 715 | 127 | 842 | 933 |
| wave 6 | 607 | 163 | 770 | 1005 |
| wave 7 | 631 | 199 | 830 | 945 |
|  | **wave-to-wave retention** | **wave-to-wave attrition** | **cumulative retention** | **cumulative retention after wave 2** |
| wave 1 | nap | nap | 100.0 |  |
| wave 2 | 79.9 | 20.1 | 79.9 | 100.0 |
| wave 3 | 88.5 | 11.4 | 70.1 | 88.6 |
| wave 4 | 77.1 | 22.9 | 54.6 | 68.3 |
| wave 5 | 73.8 | 26.2 | 47.4 | 59.3 |
| wave 6 | 72.1 | 27.9 | 43.3 | 54.3 |
| wave 7 | 81.9 | 18.1 | 46.8 | 58.5 |

**Respons patterns**

The table below gives an overview of which combinations of waves the respondents participated in. The pattern shows that many respondents were lost after the first and second wave. About a quarter of the respondents participated in all waves. The bottom part of the table shows cases in which respondents were ‘recovered’ after not participating in a specific wave. We see that none of the dropouts after the first two waves were ever recovered. We also see that recovery in the last two waves was quite successful, despite the fact that these were mail questionnaires.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **wave1** | **wave2** | **wave3** | **wave4** | **wave5** | **wave6** | **wave7** | **number** |
|  |  |  |  |  |  |  |  |  |
| 1000000 | x |  |  |  |  |  |  | 356 |
| 1200000 | x | x |  |  |  |  |  | 162 |
| 1230000 | x | x | x |  |  |  |  | 93 |
| 1234000 | x | x | x | x |  |  |  | 76 |
| 1234500 | x | x | x | x | x |  |  | 85 |
| 1234560 | x | x | x | x | x | x |  | 74 |
| 1234567 | x | x | x | x | x | x | x | 453 |
|  |  |  |  |  |  |  |  |  |
| 1234507 | x | x | x | x | x |  | x | 103 |
| 1234067 | x | x | x | x |  | x | x | 94 |
| 1230567 | x | x | x |  | x | x | x | 67 |
| 1234007 | x | x | x | x |  |  | x | 57 |
| 1230007 | x | x | x |  |  |  | x | 26 |
| 1230067 | x | x | x |  |  | x | x | 31 |
| 1234060 | x | x | x | x |  | x |  | 27 |
| 1230500 | x | x | x |  | x |  |  | 25 |
| 1230507 | x | x | x |  | x |  | x | 22 |
| 1230560 | x | x | x |  | x | x |  | 13 |
| 1230060 | x | x | x |  |  | x |  | 11 |

**Cumulative file**

All data of all waves are included in a single file. The data are stored in wide format, each person is represented by one line in the data and the variables for each wave are stored next to each other. The first 2 symbols of each variable indicate the wave (e.g., w1xxx to w7xxx).

The questionnaires were different in each year but also contain overlap. When similar concepts were measured, the same instruments were used in each wave as far as possible. Several waves contain unique measures that are not present in other waves. In the file, no attempt was made to make variables consistent across waves. The researcher is responsible for making measures comparable.

The variables and values are labeled so that they can be used easily by the researcher. We strongly advice to not only rely on the labels but to also consult the questionnaires that were used in each wave. The questionnaires are appendices in this codebook. Variable names in the file follow the questionnaire format as much as possible.

From the total sample, cases were removed with inconsistent birth dates. Not counted as inconsistent are the following cases: wrong year but with consistent day and month, inconsistent month but consistent day and year, inconsistent day but consistent month and year, day and month reversed. Cases with inconsistent sex values across waves were also removed. These can be actual sex changes, coding errors, or errors due to the participation of a non-panel member in a given wave.

Places where people live and work are removed from the data but can be requested. An urbanization code for the place of birth is included. Birth dates are removed as well. Only the year and month of birth is listed in the beginning of the file. Wave-specific sex variables were also removed and replace by a single sex variable in the beginning of the file.

**Quotation**

The data are free to use but this codebook as well as the data should be cited in any publication using the data as follows:

Liefbroer, A.C., M. Kalmijn & A.R. Poortman (2015). Codebook of the Panel Study of Social Integration in the Netherlands 1987-2009. NIDI: Unpublished Manuscript.

Liefbroer, A.C., M. Kalmijn & A.R. Poortman (2015). The Panel Study of Social Integration in the Netherlands 1987-2009. DANS/KNAW: Public use data set.

**Earlier versions**

An earlier version of the PSIN was based on the first five waves. These data are located in the DANS Data Archive as p1437. The codebook “PANEL STUDY OF SOCIAL INTEGRATION IN THE NETHERLANDS 1987-1995 (PSIN8795)” by Liefbroer & Kalmijn (1999) contains more details and a nonresponse analysis.

**Appendices**

Questionnaires