

Original survey information

Generic information	
Name of survey	German Socio-Economic Panel (GSOEP) / <i>Das Sozio-oekonomische Panel(SOEP) or Leben in Deutschland</i>
Institution responsible	<i>Deutsches Institut für Wirtschaftsforschung (DIW)</i>
Main objective	To collect representative micro-data on persons, households and families in order to measure stability and change in living conditions by following principally a micro-economic approach enriched with sociology and political science variables, mainly determined by the "Social Indicator" movement.
Survey structure	Panel and crossectional.
Frequency	Annual
Survey year / Wave	2005 survey (Wave V, #22)
Coverage	All private (non-group, non-institutional) households in Germany. The institutionalised population (persons living in hospitals, nursing homes, military installations) was not representatively included in the first wave; later, however, persons from the initial households who had taken up residence temporarily or permanently in institutions were followed.
Sampling	
Sampling design	<i>Initial samples:</i> there are 7 different samples, all multi-stage random samples, which are regionally clustered (around federal states, administrative districts and type of community). The respondents (households) are selected by random-walk. <i>Follow-up concept:</i> old households with old and new persons (births and moved in) are followed up as well as new households with old (moved out) and new persons (births and split-offs).
Sampling frame	1982 ADM master tape for sample A, immigrant registration records and local residents' registration lists for sample B, central residents' file of the GDR for sample C
Final sample size	In 2005, the total size of all samples (A-G) was around 22,000 individuals.
Collection	
Collection period	January - October 2005
Instruments and main contents	Data were collected through CAPI interview with the following instruments: - <i>Address log (Adressenprotokoll)</i> : covers, at the household level, the size of the household, some regional information, the survey status, and, at the individual level, the relation to the head, the survey status of the individuals and some demographic information; it is filled in by the interviewer even in case of non-response; - <i>Household questionnaire (Haushaltsfragebogen)</i> : asked to the household head; covers housing, some household incomes and information about child care and elderly/disabled care. - several individual questionnaires: i) the main <i>Individual Questionnaire (Personenfragebogen)</i> , asked to all individuals aged 16 or above; ii) the <i>Youth Supplementary Questionnaire (Jugend Zusatzfragebogen)</i> additionally asked to individuals aged 16 who are interviewed for the first time; iii) the <i>Biographic Supplementary Questionnaire (Lebenslauf Zusatzfragebogen)</i> additionally asked to adults aged 17 or more who are interviewed for the first time; iv) the <i>Child Questionnaire (Ihr Kind im Alter von 2 oder 3 Jahren)</i> asked to mothers for each child aged between 2 and 3; v) the <i>Mother and Child Questionnaire (Mutter und Kind)</i> asked to mothers of newborns.
Definitions	
Household	
Household head	The head of the household is defined as the person who knows best about the general conditions under which the household acts and is supposed to answer the household questionnaire in each given year (to reduce the risk of longitudinal inconsistencies).
Data quality aspects	
Unit response rate	At the household level, the panel stability varies from 96.1 to 97.5% depending on the sample. The response rates of the individual questionnaires vary: 94.6% for the main Adult Questionnaire, 88.8% for the supplementary Youth Questionnaire and 62.1% for the Supplementary adult Questionnaire.
Item non-response / imputation	The Cross-National Equivalent File (CNEF) includes completely simulated taxes and contributions (on the basis of a microsimulation model – the Schwarze routine) and fully imputed missing income information due to item and partial unit non-response.
Weighting	Different longitudinal and cross-sectional weights (both at the individual and household level) to correct for the different sampling probabilities of the subsamples, for non-response (non-willingness to participate in the first wave) and attrition in the subsequent waves.
Labour market information	
Source	Interview.
Reference period	At interview time (calendar year 2004 for calendar of activities).
Definition of employment	Employed are persons whose main self-defined activity is employment (even if irregular or marginal); some information is also collected about employment as secondary activity.
Treatment of military personnel	Conscripts and civil servants are identified (through the self-defined main activity variable) and treated as not employed. Regular Armed Forces (mainly employed) are identifiable through the occupation code.
Income	
Source	Interview.
Reference period	Calendar year 2004.
Unit of collection	Mostly individual, some income sources at the household level.
Unit of time	Mostly monthly income with number of months in 2004, some yearly.
Restrictions	
Gross/net	Variables are recorded gross of taxes and contributions; taxes and contributions are imputed separately (CNEF).