

DIW Berlin / SOEP (Ed.)

Research Report

SOEP-Core v35 - HGEN: Household-related status and generated variables

SOEP Survey Papers, No. 831

Provided in Cooperation with:

German Institute for Economic Research (DIW Berlin)

Suggested Citation: DIW Berlin / SOEP (Ed.) (2020) : SOEP-Core v35 - HGEN: Household-related status and generated variables, SOEP Survey Papers, No. 831, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:

<https://hdl.handle.net/10419/219053>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by-sa/4.0/>

SOEP Survey Papers

Series D – Variable Descriptions and Coding

SOEP-Core v35 – HGEN: Household- Related Status and Generated Variables

Running since 1984, the German Socio-Economic Panel (SOEP) is a wide-ranging representative longitudinal study of private households, located at the German Institute for Economic Research, DIW Berlin.

The aim of the SOEP Survey Papers Series is to thoroughly document the survey's data collection and data processing.

The SOEP Survey Papers is comprised of the following series:

Series A – Survey Instruments (Erhebungsinstrumente)

Series B – Survey Reports (Methodenberichte)

Series C – Data Documentation (Datendokumentationen)

Series D – Variable Descriptions and Coding

Series E – SOEPmonitors

Series F – SOEP Newsletters

Series G – General Issues and Teaching Materials

The SOEP Survey Papers are available at <http://www.diw.de/soepsurveyspapers>

Editors:

Dr. Jan Goebel, DIW Berlin

Prof. Dr. Stefan Liebig, DIW Berlin and Freie Universität Berlin

Dr. David Richter, DIW Berlin

Prof. Dr. Carsten Schröder, DIW Berlin and Freie Universität Berlin

Prof. Dr. Jürgen Schupp, DIW Berlin and Freie Universität Berlin

Dr. Sabine Zinn, DIW Berlin

Please cite this paper as follows:

SOEP Group, 2020. SOEP-Core v35 – HGEN: Household-Related Status and Generated Variables. SOEP Survey Papers 831: Series D – Variable Descriptions and Coding. Berlin: DIW Berlin/SOEP



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

© 2020 by SOEP

ISSN: 2193-5580 (online)

DIW Berlin
German Socio-Economic Panel (SOEP)
Mohrenstr. 58
10117 Berlin
Germany

soepapers@diw.de

SOEP-Core v35 – HGEN: Household-Related Status and Generated Variables

SOEP Group

2020

Contents

| | | |
|----------|--|-----------|
| 1 | General Information | 4 |
| 2 | Sample Information | 4 |
| | hid – Current wave household number | 4 |
| | cid – Original household number | 5 |
| | year – Survey Year | 6 |
| 3 | General Housing Information | 6 |
| | hgacquis – Means Of Acquiring Dwelling | 6 |
| | hgcstyrmin – Earliest Possible Construction Year of Dwelling | 7 |
| | hgcstyrmax – Latest Possible Construction Year of Dwelling | 8 |
| | hgcondit – Condition Of House | 9 |
| | hgreval – Rent Of Dwelling Vs Comparable Dwellings | 9 |
| | hgnuts1 – NUTS-Systematic-1 (Federal State) | 10 |
| 4 | Equipment of the Dwelling | 11 |
| | hgroom – Number Of Rooms Larger Than 6 Sq M | 11 |
| | hgeqpbas – Dwelling Has Basement | 12 |
| | hgeqpfhea – Equipment: Floor Heating | 12 |
| | hgeqpgar – Dwelling Has Garden | 12 |
| | hgeqphea – Dwelling Has Central,Floor Heat | 13 |
| | hgeqpinsul – Equipment: Thermal Insulation | 13 |
| | hgeqpiwc – Dwelling Has Indoor Toilet | 13 |
| | hgeqpkit – Dwelling Has Kitchen | 14 |
| | hgeqplif – Dwelling Has An Elevator | 14 |
| | hgeqpmglass – Equipment: at Least Double Glazing | 15 |
| | hgeqpnoar – Equipment: Barrier Free Furnishing | 15 |
| | hgeqpnrj – Dwelling Has Alternative Energy Source | 15 |
| | hgeqppark – Equipment: Garage/ Parking Space | 16 |
| | hgeqpshw – Dwelling Has Indoor Bath,Shower | 16 |
| | hgeqpsol – Dwelling Has Solar System | 16 |
| | hgeqptel – Dwelling Has Telephone | 17 |
| | hgeqpter – Dwelling Has Balcony,Terrace | 17 |
| | hgeqpwat – Dwelling Has Water, Bioler | 18 |
| | hgeqpair – Dwelling Has Air Conditioner | 18 |
| | hgeqpalm – Dwelling Has Alarm System | 18 |
| 5 | Costs of Living, Size, and Rooms (Imputed Variables) | 19 |
| | hgrent – Amount Of Rent Minus Heating Costs | 19 |
| | hgrentinfo – Pays Rent and/or Utilities | 21 |
| | hgnorent – Does Not Pay Rent | 21 |
| | hgelectr – Costs of Electricity (gen) | 22 |
| | hgelectrinfo – Reason for Missing Electricity Costs | 23 |
| | hggas – Costs of Gas (EUR) | 23 |
| | hggasinfo – Reason for Missing Gas Costs | 24 |
| | hgheat – Costs Of Warm Water, Gen | 24 |
| | hgheatinfo – Reason for Missing Heating Costs | 25 |
| | hgmoveyr – Year Moved Into Dwelling | 26 |
| | hgutil – Other Monthly Utility Costs | 27 |

| | |
|--|-----------|
| hgutilinfo – Reason for Missing Additional Utility Costs | 27 |
| hgseval – Adequacy Of Living Space In Housing Unit | 28 |
| hgsize – Size Of Housing Unit In Sq M | 28 |
| hgf2rent – Detailed Imputation Flag, Rent Minus Heating Costs | 29 |
| hgfelectr – Imputation Flag, Costs of Electricity | 30 |
| hgfheat – Imputation Flag, Costs Of Heating And Warm Water | 30 |
| hgfrent – Imputation Flag, Amount Of Rent Minus Heating Costs | 31 |
| hgffroom – Imputation Flag, Number Of Rooms Lager Than 6 Sq M | 31 |
| hgfsiz – Imputation Flag, Size Of Housing | 31 |
| hgffutil – Imputation Flag, Other Monthly Utility Costs | 32 |
| 6 Subsidization of Housing Costs | 32 |
| hgossubs – Amount Of Subsidies Last Year | 32 |
| hgreduc – Dwelling At A Reduced Prize | 33 |
| hgrrsubs – Government Subsidized Housing | 33 |
| hgrrsubsid – Government Subsidizes Housing Payments | 34 |
| 7 Typology of Household | 34 |
| hgowner – Tenant Or Owner Of Dwelling | 34 |
| hgtyp1hh – Household Typology, One Digit | 35 |
| hgtyp2hh – Household Typology, Two Digit | 35 |
| 8 Household Income | 36 |
| hgfhinc – Monthly Household Net Income (EUR) | 36 |
| hgi1hinc – 1. Imputed Monthly Net Household Income (EUR) [1/5] | 37 |
| hgi2hinc – 2. Imputed Monthly Net Household Income (EUR) [2/5] | 39 |
| hgi3hinc – 3. Imputed Monthly Net Household Income (EUR) [3/5] | 40 |
| hgi4hinc – 4. Imputed Monthly Net Household Income (EUR) [4/5] | 40 |
| hgi5hinc – 5. Imputed Monthly Net Household Income (EUR) [5/5] | 41 |
| hgffhinc – Imputation Flag, Monthly Net Household Income | 42 |
| 9 Time and Method of Interview | 42 |
| iyear – Household Related Survey Year | 42 |
| hghmonth – Month Of Interview | 42 |
| hghmode – Interview Method | 43 |
| inputdataset – Source dataset | 43 |

1 General Information

The HGEN file provides a set of time-consistent variables generated from the SOEP household questionnaire. It only includes households who participated in the respective year.

Up to Wave G, “old” households already known to the SOEP were surveyed at the old address with a “green” questionnaire; old households that had moved and new households received a “blue” questionnaire. While a number of questions in the blue questionnaire surveyed information for the first time (e.g. living space in square meters), this information was only asked for again in the green questionnaire in case of changes. Otherwise, the information collected in the previous wave was still valid. The variables described in the following are in part status variables in this sense: Information collected once is carried forward to subsequent waves if no address change has taken place since the previous wave. This is the case for: HGCNSTYRMIN, HGCNSTYRMAX, HGCONDIT (for Wave B to G), HG-SIZE, HGROOM, HGEQPKIT, HGEQPSHW, HGEQPIWC, HGEQPHEA, HGEQPTEP, HGEQPBAS, HGEQPGAR, HGEQPWAT, HGEQPALM, HGEQPSOL, HGEQPAIR, HG-MOVEYR, HGACQUIS, HGSUBSID, HGRSUBS, HGREduc and HGSEVAL (for Wave C to G). Furthermore, identical information is recorded in the green and blue questionnaire in separate variables (e.g. housing tenure as owner or tenant). The corresponding status variables are therefore just a compilation of these more dispersed pieces of information. Since only one common questionnaire has been used since Wave H rather than the former “blue” and “green” versions, the necessity for the aforementioned status variables disappears. However, this user-friendly redundancy is maintained for reasons of consistency.

In addition, we impute missing values of the variables HGROOM, HGSIZE, HGRENT, HGHEAT, HGUTIL and HGHINC to ensure the completeness of the dataset and facilitate its usage. Details of the imputation process can be found in the description of the respective variable.

Since Wave BG, new migration samples (M3 and M4) have been added. These samples provide less household information which explains the high incidence of missing values (=5). Therefore, in the samples M3 and M4 hgen variables are generally not imputed.

2 Sample Information

hid – Current wave household number

| | |
|--------------------------|--------|
| 19 | 6 |
| 27 | 25 |
| 35 | 10 |
| 43 | 3 |
| 51 | 9 |
| 60 | 29 |
| 78 | 3 |
| 86 | 2 |
| 94 | 35 |
| 108 | 4 |
| 116 | 15 |
| 124 | 23 |
| 132 | 14 |
| 140 | 7 |
| 159 | 35 |
| ... (51313 rows omitted) | 375303 |

| | |
|---------|---|
| 3923487 | 2 |
| 3923541 | 1 |
| 3923550 | 1 |
| 3923568 | 2 |
| 3923576 | 2 |
| 3923584 | 2 |
| 3923614 | 2 |
| 3923649 | 1 |
| 3923746 | 2 |
| 3923762 | 1 |
| 3923789 | 1 |
| 3923797 | 1 |
| 3923800 | 2 |
| 3923835 | 2 |
| 3923878 | 1 |

cid – Original household number

| | |
|--------------------------|--------|
| 19 | 9 |
| 27 | 50 |
| 35 | 10 |
| 43 | 3 |
| 51 | 9 |
| 60 | 89 |
| 78 | 3 |
| 86 | 2 |
| 94 | 35 |
| 108 | 4 |
| 116 | 15 |
| 124 | 23 |
| 132 | 16 |
| 140 | 7 |
| 159 | 35 |
| ... (42232 rows omitted) | 375213 |
| 3923487 | 2 |
| 3923541 | 1 |
| 3923550 | 1 |
| 3923568 | 2 |
| 3923576 | 2 |
| 3923584 | 2 |
| 3923614 | 2 |
| 3923649 | 1 |
| 3923746 | 2 |
| 3923762 | 1 |
| 3923789 | 1 |
| 3923797 | 1 |
| 3923800 | 2 |
| 3923835 | 2 |
| 3923878 | 1 |

syear – Survey Year

| | |
|----------------------|-------|
| 1984 | 5921 |
| 1985 | 5322 |
| 1986 | 5090 |
| 1987 | 5026 |
| 1988 | 4814 |
| 1989 | 4690 |
| 1990 | 6819 |
| 1991 | 6699 |
| 1992 | 6665 |
| 1993 | 6637 |
| 1994 | 6795 |
| 1995 | 6968 |
| 1996 | 6894 |
| 1997 | 6810 |
| 1998 | 7668 |
| ... (5 rows omitted) | 57337 |
| 2004 | 11796 |
| 2005 | 11440 |
| 2006 | 12499 |
| 2007 | 11689 |
| 2008 | 11058 |
| 2009 | 11889 |
| 2010 | 15414 |
| 2011 | 16819 |
| 2012 | 16508 |
| 2013 | 18093 |
| 2014 | 16037 |
| 2015 | 15991 |
| 2016 | 17775 |
| 2017 | 19701 |
| 2018 | 18682 |

3 General Housing Information**hgacquis** – Means Of Acquiring Dwelling

| | | |
|----|---|--------|
| 1 | Bought From Owner | 34696 |
| 2 | Inheritance, Gift | 23125 |
| 3 | Bought, Built New | 39681 |
| 4 | Got Back From Public Property | 6 |
| -1 | No answer | 54235 |
| -2 | Does not apply | 200304 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 23499 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 0 |

Waves: all

If a household did not provide any information on HGACQUIS, the information from previous waves is carried forward under the condition that a change of address or ownership status (HGOWNER) has not taken place.

In Wave A to Wave G, HGACQUIS was asked only if a household was new to the SOEP or if it had changed its address. From Wave H to Wave R, HGACQUIS was also asked if a change in ownership status had taken place in the last year. Since Wave S, HGACQUIS is only asked if a change in ownership status has taken place. In consequence, households new to the SOEP without a change in ownership in the last year, do not state any information on HGACQUIS anymore. Furthermore, since Wave S, the category „Bought New or Built“ is no longer included in the household questionnaire but is carried forward from last years. Given these changes in the categories and the sample that provides information for HGACQUIS, we recommend not to compare the shares of the different categories of HGACQUIS over time.

„Returned to private ownership“ was surveyed only in Wave I in East Germany and is carried forward.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgcnstyrmin – Earliest Possible Construction Year of Dwelling

| | |
|---|--------|
| 0 | 50739 |
| 1918 | 7547 |
| 1919 | 47158 |
| 1949 | 107180 |
| 1972 | 50857 |
| 1973 | 904 |
| 1974 | 1123 |
| 1975 | 996 |
| 1976 | 733 |
| 1977 | 430 |
| 1978 | 636 |
| 1979 | 554 |
| 1980 | 926 |
| 1981 | 27699 |
| 1982 | 583 |
| ... (28 rows omitted) | 51581 |
| 2011 | 654 |
| 2012 | 360 |
| 2013 | 151 |
| 2014 | 287 |
| 2015 | 203 |
| 2016 | 176 |
| 2017 | 75 |
| 2018 | 9 |
| -1 No answer | 11881 |
| -2 Does not apply | 0 |
| -3 Not valid | 0 |
| -4 Forbidden Multiple Response | 0 |
| -5 Not Included In Questionnaire Version | 12104 |
| -6 Questionnaire Version With Modified Filter | 0 |

-8 Question not part of the survey program this year 0

Waves: 1984–2013, 2015–2016

HGCNSTYRMIN provides the lower limit of the time period in which the household's building was constructed. HGCNSTYRMAX provides the upper limit. E.g. if a household's HGCNSTYRMIN is 1984 and HGCNSTYRMAX is 1990, the building was built between 1984 and 1990.

Households new to the SOEP and households who have moved since the last interview are asked for the time period of construction of the building they live in. With this information HGCNSTYRMIN and HGCNSTYRMAX are constructed. For old households the variables are carried forward from the last years. Since Wave X, households are also asked for the exact construction year. In these cases, HGCNSTYRMIN and HGCNSTYRMAX contain the same value. If a building was built before 1919, HGCNSTYRMIN contains a zero and HGCNSTYRMAX 1919.

In Wave BE, the question about the construction year was not included in the SOEP. All information for HGCNSTYRMIN and HGCNSTYRMAX derives from the previous year.

HGCNSTRMIN and HGCNSTRMAX replace CNSTYR\$\$ that has previously been distributed in \$HGEN files. The latter was a categorical variable grouping about one decade each into one category. Hence, with the last category having a flexible upper limitation and being split up regularly, sorting was ambiguous in some cases. For example, a household responding in Wave L that his house has been built "1981 or later" was not unambiguously assignable to either "1981-1990" or "1990 or later". Using HGCNSTYRMIN and HGCNSTYRMAX, changing questions can be far better described.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgcnstyrmax – Latest Possible Construction Year of Dwelling

| | |
|-----------------------|--------|
| 1917 | 6546 |
| 1918 | 44193 |
| 1948 | 54705 |
| 1971 | 107180 |
| 1972 | 1610 |
| 1973 | 904 |
| 1974 | 1123 |
| 1975 | 996 |
| 1976 | 733 |
| 1977 | 430 |
| 1978 | 636 |
| 1979 | 554 |
| 1980 | 38174 |
| 1981 | 650 |
| 1982 | 583 |
| ... (28 rows omitted) | 90535 |
| 2011 | 641 |
| 2012 | 403 |
| 2013 | 164 |
| 2014 | 287 |
| 2015 | 217 |
| 2016 | 192 |
| 2017 | 83 |

| | | |
|------|---|-------|
| 2018 | | 22 |
| -1 | No answer | 11881 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 12104 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 0 |

Waves: 1984–2013, 2015–2016

HGCNSTYRMAX provides the upper limit of the time period in which the household's building was constructed. The generation of HGCNSTYRMAX is analogous to HGCNSTYRMIN.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgcondit – Condition Of House

| | | |
|----|---|--------|
| 1 | In A Good Condition | 212908 |
| 2 | Some Renovations | 92371 |
| 3 | Full Renovations | 12333 |
| 4 | Dilapidated | 478 |
| -1 | No answer | 1298 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 0 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 56158 |

Waves: 1984–2015

Respondent's subjective assessment of the condition of the building. In the West German sub-samples from Wave B to Wave G, the information on HGCONDIT was collected only for new households and for households with a residential move since the previous year (households with "blue" questionnaires). For households who had not moved ("green" questionnaire), the information collected in previous years was carried forward. The wording in the questionnaire was changed in the first wave of the East German sub-sample (Wave G) to better capture the rundown condition of some residential buildings in East Germany. Since Wave H the wording is identical for the entire SOEP-sample in East and West Germany.

Since Wave BG, HGCONDIT is not asked anymore and information is not carried forward.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgcreval – Rent Of Dwelling Vs Comparable Dwellings

| | | |
|----|--------------------|--------|
| 1 | Very Inexpensive | 12957 |
| 2 | Inexpensive | 39152 |
| 3 | Reasonable | 67492 |
| 4 | Slightly Expensive | 27088 |
| 5 | Too Expensive | 5335 |
| -1 | No answer | 1643 |
| -2 | Does not apply | 125873 |

| | | |
|----|---|-------|
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 72149 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 23857 |

Waves: 1984–2002, 2005–2014

Subjective assessment by respondent (household head). This variable was not surveyed in Wave T and Wave U. The corresponding information from the previous year is not carried forward longitudinally due to the possibility of changes in rent and income, residential moves, and change in the person responding.

Since Wave BF, HGREVAL is not asked and information is neither carried backward nor forward.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgnuts1 - NUTS-Systematic-1 (Federal State)

| | | |
|----|---|-------|
| 1 | Baden-Wuerttemberg | 45454 |
| 2 | Bavaria | 54245 |
| 3 | Berlin | 15781 |
| 4 | Brandenburg | 14421 |
| 5 | Bremen | 3155 |
| 6 | Hamburg | 6426 |
| 7 | Hesse | 26787 |
| 8 | Mecklenburg-West Pomerania | 8548 |
| 9 | Lower Saxony | 33098 |
| 10 | North Rhine-Westphalia | 76700 |
| 11 | Rhineland-Palatinate | 18351 |
| 12 | Saarland | 3153 |
| 13 | Saxony | 23642 |
| 14 | Saxony-Anhalt | 13975 |
| 15 | Schleswig-Holstein | 11856 |
| 16 | Thuringia | 14031 |
| -1 | No answer | 2 |
| -2 | Does not apply | 5921 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 0 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 0 |

NUTS (“Nomenclature of Statistical Territorial Units”) URL is a hierarchical system for dividing up the economic territory of the European Union. It was introduced by Eurostat more than 30 years ago in order to provide a single uniform breakdown of territorial units for the production of regional statistics. NUTS 1 especially contains the major socio-economic regions for analyzing regional Community problems. It subdivides the European Union by now into 97 regions, whereas in Germany there are equivalent to the German Federal States 16 regions. Before Wave Q, Rhineland-Palatinate and Saarland were defined as one region.

For more information, contact: Jan Goebel (Tel. +49-30-89789-377), <jgoebel@diw.de>

4 Equipment of the Dwelling

hgroom – Number Of Rooms Larger Than 6 Sq M

| | |
|--|--------|
| 0 | 14 |
| 1 | 13724 |
| 2 | 60212 |
| 3 | 114670 |
| 4 | 78003 |
| 5 | 49426 |
| 6 | 29950 |
| 7 | 14348 |
| 8 | 7101 |
| 9 | 2924 |
| 10 | 1363 |
| 11 | 455 |
| 12 | 232 |
| 13 | 87 |
| 14 | 65 |
| ... (8 rows omitted) | 98 |
| 25 | 1 |
| 35 | 3 |
| 42 | 1 |
| 45 | 2 |
| 51 | 11 |
| 60 | 19 |
| 80 | 20 |
| 82 | 12 |
| -1 No answer | 2805 |
| -2 Does not apply | 0 |
| -3 Not valid | 0 |
| -4 Forbidden Multiple Response | 0 |
| -5 Not Included In Questionnaire Version | 0 |
| -6 Questionnaire Version With Modified Filter | 0 |
| -8 Question not part of the survey program this year | 0 |

Waves: all

Up to Wave R (except for Wave O), HGROOM was collected only in the first interview with new households. In case a household had moved or (from Wave H onwards) whether a household's housing unit size changed due to renovations or additions (up to Wave G, these households filled out a "green" questionnaire). In Wave O, the information had been asked again in order to correct for mistakes that may have occurred in carrying forward or in the process of imputation. From Wave S onwards, the question on the number of rooms has been asked to all households annually.

For old households still residing at their old address and neither having moved nor renovated their dwelling, the information on the number of rooms is carried forward as a status variable in order to provide valid current information. Analogously, the information is carried back to fill gaps if a household missed to state the number of rooms in the first year(s) they moved to a new housing.

In case, the information on the number of rooms is still missing, it is imputed simultaneously with other variables using Stata's chained imputation procedure. Imputed values are

not carried forward. For information on the imputation procedure, see the description of HGRENT.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgeqpbas – Dwelling Has Basement

| | | |
|----|---|--------|
| 1 | Yes | 345960 |
| 2 | No | 24261 |
| -1 | No answer | 5243 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 82 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 0 |

Waves: all

Except for Wave H, Wave O, and from Wave V onwards, HGEQPBAS has been collected only from new households and households who have moved since the previous interview. For this reason, in case no address change has taken place, the information for HGEQPBAS is carried forward from the previous years.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgeqpfhea – Equipment: Floor Heating

| | | |
|----|---|--------|
| 1 | Yes | 30633 |
| 2 | No | 132573 |
| -1 | No answer | 3703 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 0 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 208637 |

Waves: Since 2014

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgeqpgar – Dwelling Has Garden

| | | |
|----|---|--------|
| 1 | Yes | 214789 |
| 2 | No | 152636 |
| -1 | No answer | 6953 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 1168 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 0 |

Waves: all

Except for Wave H, Wave O, and from Wave V onwards, HGEQPGAR has been collected only from new households and households who have moved since the previous interview. For this reason, in case no address change has taken place, the information for HGEQPGAR is carried forward from the previous years.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgeqphea – Dwelling Has Central,Floor Heat

| | | |
|----|---|--------|
| 1 | Yes | 274723 |
| 2 | No | 25603 |
| -1 | No answer | 2203 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 868 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 72149 |

Waves: 1984-2014

Only in Wave H, Wave O, and from Wave V onwards, HGEQPHEA is asked from every household. In previous years, the variable was only collected from new households and households who had moved since the previous interview. For this reason, in case no address change has taken place, the information for HGEQPHEA is carried forward from the previous years. Additionally, from Wave G onwards the information is updated if the household has stated that it modernized its heating since January of the previous year. Since Wave BF, HGEQPHEA is not asked and the information is neither carried backward nor forward.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgeqpinsul – Equipment: Thermal Insulation

| | | |
|----|---|--------|
| 1 | Yes | 37190 |
| 2 | No | 24033 |
| -1 | No answer | 801 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 10125 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 303397 |

Waves: Since 2015

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgeqpiwc – Dwelling Has Indoor Toilet

| | | |
|----|-----------|--------|
| 1 | Yes | 202210 |
| 2 | No | 4371 |
| -1 | No answer | 1832 |

| | | |
|----|---|--------|
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 224 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 166909 |

Waves: 1984-2008

Only in Wave H, Wave O, and from Wave V to Wave Y, HGEQPIWC is asked from every household. In previous years, the variable was only collected from new households and households who had moved since the previous interview. For this reason, in case no address change has taken place, the information for HGEQPIWC is carried forward from the previous years. Additionally, from Wave B onwards, the information is updated if the household has stated that it modernized its bathroom since January of the previous year. In a second step, the information of the variable HGEQPSHW is used to replace missing and inconsistent values of EQPIWC\$. For every household with EQPSHW\$=1, EQPIWC\$ is set to 1. Since Wave Z, HGEQPWC is not asked. Since Wave BA, the information is neither carried backward nor forward.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgeqpkit – Dwelling Has Kitchen

| | | |
|----|---|--------|
| 1 | Yes | 282602 |
| 2 | No | 3187 |
| -1 | No answer | 1558 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 13 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 88186 |

Waves: 1984-2013

Only in Wave H, Wave O, and from Wave V onwards, HGEQPKIT is asked from every household. In previous years, the variable was only collected from new households and households who had moved since the previous interview. For this reason, in case no address change has taken place, the information for HGEQPKIT is carried forward from previous years. Additionally, from Wave B on, the information is updated if the household has stated that it modernized its kitchen since January of the previous year. Since Wave BE, HGEQPKIT is neither asked nor carried forward.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgeqplif – Dwelling Has An Elevator

| | | |
|----|-----------------------------|--------|
| 1 | Yes | 14728 |
| 2 | No | 148667 |
| -1 | No answer | 3514 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |

| | | |
|----|---|--------|
| -5 | Not Included In Questionnaire Version | 0 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 208637 |

Waves: Since 2009

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgeqpmglass – Equipment: at Least Double Glazing

| | | |
|----|---|--------|
| 1 | Yes | 57347 |
| 2 | No | 4537 |
| -1 | No answer | 140 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 10125 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 303397 |

Waves: Since 2015

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgeqpnoar – Equipment: Barrier Free Furnishing

| | | |
|----|---|--------|
| 1 | Yes | 14905 |
| 2 | No | 88041 |
| -1 | No answer | 3333 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 0 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 269267 |

Waves: Since 2013

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgeqpnrj – Dwelling Has Alternative Energy Source

| | | |
|----|---|--------|
| 1 | Yes | 5370 |
| 2 | No | 150829 |
| -1 | No answer | 585 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 10125 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 208637 |

Waves: Since 2009

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgeqppark – Equipment: Garage/ Parking Space

| | | |
|----|---|--------|
| 1 | Yes | 44685 |
| 2 | No | 24283 |
| -1 | No answer | 3181 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 0 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 303397 |

Waves: Since 2015

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgeqpswh – Dwelling Has Indoor Bath,Shower

| | | |
|----|---|--------|
| 1 | Yes | 280620 |
| 2 | No | 4860 |
| -1 | No answer | 1662 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 218 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 88186 |

Waves: 1984-2013

Only in Wave H, Wave O, and from Wave V onwards, HGEQPSHW is asked from every household. In previous years, the variable was only collected from new households and households who had moved since the previous interview. For this reason, in case no address change has taken place, the information for HGEQPSHW is carried forward from the previous years. Additionally, from Wave B onwards, the information is updated if the household has stated that it modernized its bathroom since January of the previous year. Since Wave BE, HGEQPSHW is not asked anymore. Since Wave BF, the information is not carried forward.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgeqpsol – Dwelling Has Solar System

| | | |
|----|--|--------|
| 1 | Yes | 15510 |
| 2 | No | 163262 |
| -1 | No answer | 759 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 10125 |
| -6 | Questionnaire Version With Modified Filter | 0 |

-8 Question not part of the survey program this year 185890

Waves: Since 2007

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgeqptel – Dwelling Has Telephone

| | | |
|----|---|--------|
| 1 | Yes | 265719 |
| 2 | No | 6768 |
| -1 | No answer | 4603 |
| -2 | Does not apply | 4640 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 7083 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 86733 |

Waves: Since 1990 (east) or 1991 (west) to 2008, 2010–2015

The question on whether a household has a telephone or not has varied over the years. In some years, households were asked very generally if they had a “telephone”. In other years they were asked more specifically for a “landline telephone” and a “cellphone”. HGEQPTEL is set to 1 if a household stated that it had a telephone, landline telephone, or cellphone. Additionally, the information from HBRUTTO (HTEL) is used to complete missing information. In Wave N, Wave P, and Wave Z, the household questionnaire did not ask about the possession of a telephone and HBRUTTO is the only source of data. For Wave K, no information is available.

Since Wave BG, HGEQPTEL is no longer asked in the household questionnaire (except for the samples M3 and M4 as HGEQPTEL was asked in the migration household questionnaire). For this reason, it is no longer provided since Wave BG for any other samples than M3 and M4.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgeqpter – Dwelling Has Balcony,Terrace

| | | |
|----|---|--------|
| 1 | Yes | 274894 |
| 2 | No | 92613 |
| -1 | No answer | 6764 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 1275 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 0 |

Waves: all

Except for Wave H, Wave O, and from Wave V onwards, HGEQPTER was only collected from new households and households who have moved since the previous interview. For this reason, in case no address change has taken place, the information for HGEQPTER is carried forward from the previous years.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgeqpwat – Dwelling Has Water, Bioler

| | | |
|----|---|--------|
| 1 | Yes | 246907 |
| 2 | No | 3103 |
| -1 | No answer | 1847 |
| -2 | Does not apply | 4640 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 0 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 119049 |

Waves: Since 1990 (east) or 1991 (west) to 2013

Except for Wave H, Wave O, and from Wave V onwards, HGEQPWAT has been collected only from new households and households who have moved since the previous interview. For this reason, in case no address change has taken place, the information for HGEQPWAT is carried forward from the previous years. Since Wave BE, HGEQPWAT is not asked and since Wave BF, the information is neither carried backward nor forward.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgeqpair – Dwelling Has Air Conditioner

| | | |
|----|---|--------|
| 1 | Yes | 3234 |
| 2 | No | 175464 |
| -1 | No answer | 833 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 10125 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 185890 |

Waves: Since 2007

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgeqpalm – Dwelling Has Alarm System

| | | |
|----|---|--------|
| 1 | Yes | 9540 |
| 2 | No | 201133 |
| -1 | No answer | 14718 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 0 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 150155 |

Waves: Since 2004

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

5 Costs of Living, Size, and Rooms (Imputed Variables)

hgent – Amount Of Rent Minus Heating Costs

| | |
|--|--------|
| 0 | 35 |
| 1 | 6 |
| 2 | 5 |
| 3 | 8 |
| 4 | 3 |
| 5 | 18 |
| 6 | 7 |
| 7 | 16 |
| 8 | 25 |
| 9 | 26 |
| 10 | 60 |
| 11 | 33 |
| 12 | 73 |
| 13 | 72 |
| 14 | 60 |
| ... (1727 rows omitted) | 197130 |
| 4950 | 2 |
| 5000 | 1 |
| 5033 | 1 |
| 5052 | 1 |
| 5700 | 1 |
| 6000 | 1 |
| 6700 | 2 |
| 7318 | 1 |
| -1 No answer | 130 |
| -2 Does not apply | 167704 |
| -3 Not valid | 0 |
| -4 Forbidden Multiple Response | 0 |
| -5 Not Included In Questionnaire Version | 10125 |
| -6 Questionnaire Version With Modified Filter | 0 |
| -8 Question not part of the survey program this year | 0 |

Waves: all

HGRENT is a measure of gross rent, i.e. it equal to the base rent plus utility costs (HGUTIL) but does not include heating (HGHEAT), electricity costs (HGELECTR) and additional gas costs (HGGAS). HGRENT is converted into Euro values for all waves including those prior to Wave S.

The questions for HGRENT and utility costs have changed considerably since the first SOEP questionnaire in Wave A. From Waves A to Wave G, the amount of rent stated by the households in SOEP is in principle the desired concept of gross rent, i.e. basic rent excluding heating and electricity costs but including utility costs. However, in these waves the information on utility costs was not collected.

From Wave H onwards, households simply state the amount of rent they pay. Following this question, it is asked whether heating and utility costs are included in that amount of rent and what the exact costs for heating and utilities eventually are (in the latter case only if they are included).

In Wave BE and Wave BF, the questionnaire design was different to all other waves. In a first step, households were asked for their overall monthly housing costs (sum of base rent, heating costs, additional utility costs, electricity and gas). In a second step, they stated the respective expenses. In order to obtain RENT\$\$, heating, gas, and electricity costs have to be deducted from these overall monthly housing costs. In Wave BE, some households state monthly housing costs that presumably do not include all possible expenses. They might do so because they do not know all their expenses or because they misunderstood the question. We attribute the drop of rent from Wave BD to Wave BE and Wave BF to the different questionnaire design and do no longer adjust the values as we did in data release v31. Therefore, the average rent of Wave BE (data release v31) is about 10 Euros lower in data release v32 and v33 than it was in v31. Both, in Wave BE and Wave BF, implausible values of rent (e.g. if utility costs are above total living costs) have been set to missing to be imputed in the following step.

The questionnaire design in Wave BG was set back to the format of Wave BD.

Missing values of HGRENT are treated as follows: In a first step, past values of the two last years are carried forward and adjusted for inflation given that the household still lives at the same address and the dwelling's size has remained the same. The type of rent (heating/utility costs included or not) is then also carried forward. In a second step, if HGRENT is still missing, values of the two subsequent years are carried backwards in the same manner. In a third step, the remaining missing values of HGRENT are imputed by Stata's chained imputation procedure. Imputed values are not carried forward.

General Information on the Imputation Procedure Since data release v31, missing values of HGROOM, HGSIZE, HGRENT, HGHEAT, HGUTIL, and HGELECTR are imputed with Stata's chained imputation procedure (`mi impute chained`). We impute separately for each wave, for East and West and for owners and tenants. The regression specifications vary between these groups but in general they include:

- For owner:
 - HGROOM, HGSIZE, HGCONDIT, HGCNSTYRMAX
 - Number of persons in household
 - Household net income as stated in questionnaire
 - Type of dwelling (family house, apartment, etc.)
- For tenants:
 - HGROOM, HGSIZE, HGRENT, HGHEAT, HGUTIL, HGELECTR, HGCONDIT, HGCNSTYRMAX, HGEQPHEA, HGEQPFHEA, HGEQPTER, HGEQPGAR, HGEQPSOL, HGEQPAIR, HGCNSTYRMAX, HGREVAL, HGREduc
 - HGSIZE squared
 - Number of persons in household
 - Household net income as stated in questionnaire
 - Type of dwelling (family house, apartment, etc., \$\$wum1)
 - Type of tenant
 - Dummy: HGHEAT included in HGRENT (Wave G to Wave BD)
 - Dummy: HGUTIL included in HGRENT (Wave H to Wave BD)
 - Dummy: HGUTIL partly included in HGRENT (Wave H to Wave BD)
 - Dummy: HGHEAT includes electricity costs (since Wave BE)
 - Dummy: household pays only utility costs (since Wave BE)
 - Household has children younger than 16

In addition, the regression of HGRENT contains HGRSUBS, a dummy for children under 16, information on the residential area (\$\$wum3) and BIK regions. The regression of HGHEAT includes a dummy for children under 16. The regression of HGUTIL includes HGEQPLIF and omits HGRENT, HGCONDIT, the type of dwelling and the dummy for whether HGHEAT is included in HGRENT. The regression of HGSIZE omits HGSIZE squared.

HGRENT is not normalized across households before imputation and may include heating and utility costs or not. Solely partially included utility costs are subtracted. We use dummies for HGHEAT/HGUTIL included in HGRENT to account for the different types of rent. Note that these dummies are also imputed if they are missing. If HGUTIL contains a partial amount, it is set to missing and then also imputed.

Residents of homes (HGOWNER==5) are excluded from the imputation. Values larger than four times the 99th percentile in one of the target variables are also excluded.

For all six target variables, the predictive mean matching imputation method is applied. We use 175 iterations as burn-in period. Furthermore, we only distribute one imputed value for each observation.

Gross warm rent including electricity (Bruttowarmmiete inkl. Strom) could be computed from this data as follows: HGRENT + HGHEAT + HGELECTR.

HGRENT is set missing (= -5) for all M3 and M4 respondents although the migration household questionnaire surveyed rent information. However, the surveyed information for M3 and M4 is not comparable with the information for the other samples as information on HGHEAT, HGUTIL, HGELECTR, etc. is not surveyed.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgrentinfo – Pays Rent and/or Utilities

| | | |
|----|---|--------|
| 1 | Pays Rent | 41751 |
| 2 | Pays Utility Costs, but No Rent | 1361 |
| 3 | Does Not Pay either Rent or Utilities | 798 |
| -1 | No answer | 37 |
| -2 | Does not apply | 34114 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 10125 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 287360 |

Waves: Since 2014

HGRENTINFO is a filter question for housing costs. It was introduced in Wave BE and replaces NORENT\$\$ in \$HGEN files.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgnoorent – Does Not Pay Rent

| | | |
|----|---------------------------------------|--------|
| 1 | Pays No Rent | 10395 |
| -1 | No answer | 129 |
| -2 | Does not apply | 354896 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 10126 |

| | | |
|----|---|---|
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 0 |

Waves: 1984–2013

HGNORENT marks tenants who do not pay rent. Until Wave BD, it was a filter question for all housing costs such that non-rent paying tenants were not asked for HGHEAT, HGUTIL and HGELECTR. In Wave BE, the filter question was rephrased and tenants were asked for their utility costs even if they stated not to pay a rent. To maintain consistency over time, we set HGNORENT to 1 for all tenants who only pay utility costs. RENTINFO\$\$ contains the information of the new filter question.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgelectr – Costs of Electricity (gen)

| | | |
|-----|---|--------|
| 0 | | 5 |
| 1 | | 7 |
| 2 | | 2 |
| 3 | | 4 |
| 4 | | 9 |
| 5 | | 15 |
| 6 | | 9 |
| 7 | | 6 |
| 8 | | 12 |
| 9 | | 9 |
| 10 | | 127 |
| 11 | | 23 |
| 12 | | 31 |
| 13 | | 31 |
| 14 | | 50 |
| ... | (277 rows omitted) | 76416 |
| 450 | | 2 |
| 500 | | 2 |
| 530 | | 1 |
| 540 | | 1 |
| 580 | | 1 |
| 690 | | 1 |
| 840 | | 1 |
| 950 | | 1 |
| -1 | No answer | 2602 |
| -2 | Does not apply | 65527 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 10125 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 220526 |

Waves: Since 2010

Costs for electricity (excluding arising expenses for heating if applicable) have been collected starting from Wave BA if tenants pay a rent. From Wave BE onwards, they have been collected from all tenants. In Wave BF, total electricity costs were explicitly asked for so that costs for

heating and warm water were subtracted to generate HGELECTR that year. Furthermore, due to changes in the questionnaire design in Wave BF, HGELECTR is neither carried backward nor forward. See HGHEAT for further explanations on housing costs in Wave BF.

Missing values of HGELECTR are treated as follows: In a first step, past values of the two last years are carried forward and adjusted for inflation given that the household still lives at the same address, the dwelling's size has remained the same and the rent has not changed considerably. In a second step, if HGELECTR is still missing, values of the two subsequent years are carried backwards in the same manner. In a third step, the remaining missing values of HGELECTR are imputed by Stata's chained imputation procedure. Imputed values are not carried forward. We do not impute HGELECTR for years before Wave BA. For more information on the imputation procedure, see the description of HGRENT.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgelectrinfo – Reason for Missing Electricity Costs

| | | |
|----|---|--------|
| 1 | Electricity Costs Unknown | 9 |
| 2 | Electricity Costs Included in Heating Costs | 282 |
| 3 | No Electricity Costs | 617 |
| -1 | No answer | 45320 |
| -2 | Does not apply | 15665 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 10302 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 303351 |

HGELECTRINFO was introduced in Wave BE and indicates why HGELECTR is missing. If a household stated that it did not know its electricity costs, we impute HGELECTR and set HGELECTRINFO to -2.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hggas – Costs of Gas (EUR)

| | |
|------------------------|-----|
| 0 | 40 |
| 1 | 1 |
| 3 | 1 |
| 4 | 1 |
| 5 | 9 |
| 6 | 6 |
| 7 | 10 |
| 8 | 13 |
| 9 | 12 |
| 10 | 15 |
| 11 | 6 |
| 12 | 7 |
| 13 | 7 |
| 14 | 3 |
| 15 | 16 |
| ... (115 rows omitted) | 790 |

| | | |
|-----|---|--------|
| 202 | | 1 |
| 210 | | 1 |
| 213 | | 1 |
| 220 | | 1 |
| 234 | | 1 |
| 250 | | 5 |
| 300 | | 3 |
| 550 | | 1 |
| -1 | No answer | 7234 |
| -2 | Does not apply | 7852 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 287360 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 72149 |

HGGAS states the monthly costs for gas not used for heating in Euros. HGGAS was only collected in Wave BA. Missing values of HGGAS are not imputed due to the low number of households which state gas costs.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hggasinfo – Reason for Missing Gas Costs

| | | |
|----|---|--------|
| 1 | Gas Costs Unknown | 1059 |
| 2 | Electricity Costs Included in Heating Costs | 1547 |
| 3 | No Gas Costs | 5256 |
| -1 | No answer | 0 |
| -2 | Does not apply | 8175 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 287360 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 72149 |

HGGASINFO has only been asked in Wave BE and indicates why a household did not state additional gas costs that were not included in the heating costs.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgheat – Costs Of Warm Water, Gen

| | |
|---|-----|
| 0 | 457 |
| 1 | 12 |
| 2 | 22 |
| 3 | 16 |
| 4 | 27 |
| 5 | 143 |
| 6 | 80 |
| 7 | 195 |
| 8 | 96 |
| 9 | 104 |

| | |
|--|--------|
| 10 | 762 |
| 11 | 146 |
| 12 | 414 |
| 13 | 192 |
| 14 | 182 |
| ... (387 rows omitted) | 187363 |
| 595 | 1 |
| 600 | 5 |
| 700 | 5 |
| 723 | 1 |
| 756 | 1 |
| 800 | 3 |
| 900 | 2 |
| 930 | 1 |
| -1 No answer | 8623 |
| -2 Does not apply | 155325 |
| -3 Not valid | 0 |
| -4 Forbidden Multiple Response | 0 |
| -5 Not Included In Questionnaire Version | 21368 |
| -6 Questionnaire Version With Modified Filter | 0 |
| -8 Question not part of the survey program this year | 0 |

Waves: Since 1986

HGHEAT contains the heating costs of tenants. It is collected since Wave C. Until Wave BD, only tenants who stated to pay a rent were asked. Since Wave BE, every tenant was asked. Heating costs are reported in Euro for all years.

In Wave BF, HGHEAT has not been asked for directly. Instead, respondents stated their costs for different energy sources (district heating, gas, liquefied gas, electricity, heating oil, coal, wood, biomass, solar, and geothermal heat) and whether they use them for heating, warm water, cooking, and/or electronic devices respectively (if applicable). To generate HGHEAT, costs for cooking and electricity costs other than heating have been excluded by regressing the usage-dummies and controlling for household-specific (number of people living in the household, type of tenant) and dwelling-specific variables (number of rooms, size of the dwelling, construction year, living environment). The shares for heating and warm water were then added to generate HGHEAT. Due to changes in the questionnaire design, firstly, HGHEAT is lower in Wave BF than it has been in previous and following years and secondly, HGHEAT from Wave BF is neither carried backward nor forward. In file HGEN, only HGHEAT is distributed. All energy cost variables can be found in the dataset BFH.

Missing values of HGHEAT are treated as follows: In a first step, past values of the two last years are carried forward and adjusted for inflation given that the household still lives at the same address, the dwelling's size has remained the same and the rent has not changed considerably. In a second step, if HGHEAT is still missing, values of the two subsequent years are carried backwards in the same manner. In a third step, the remaining missing values of HGHEAT are imputed by Stata's chained imputation procedure. Imputed values are not carried forward. In contrast to previous versions of the SOEP, we do not impute HGHEAT for Wave A and Wave B anymore, since data release v32. For more information on the imputation see the description of HGRENT.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

[hgheatinfo](#) – Reason for Missing Heating Costs

| | | |
|----|---|--------|
| 1 | Heating Costs Unknown | 24 |
| 3 | No Heating Costs | 1871 |
| -1 | No answer | 44377 |
| -2 | Does not apply | 15730 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 10193 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 303351 |

Waves: Since 2014

HGHEATINFO was introduced in Wave BE and indicates why HGHEAT is missing. If a household stated that it did not know its heating costs, we imputed HGHEAT and set HGHEATINFO to -2.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgmoveyr - Year Moved Into Dwelling

| | | |
|------|---|--------|
| 1900 | | 1 |
| 1901 | | 34 |
| 1902 | | 11 |
| 1903 | | 7 |
| 1904 | | 18 |
| 1905 | | 4 |
| 1906 | | 24 |
| 1907 | | 93 |
| 1908 | | 8 |
| 1909 | | 45 |
| 1910 | | 19 |
| 1911 | | 18 |
| 1912 | | 69 |
| 1913 | | 22 |
| 1914 | | 33 |
| ... | (97 rows omitted) | 337245 |
| 2012 | | 6594 |
| 2013 | | 5633 |
| 2014 | | 5215 |
| 2015 | | 6439 |
| 2016 | | 6293 |
| 2017 | | 3430 |
| 2018 | | 1130 |
| 2019 | | 3 |
| -1 | No answer | 2834 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 324 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 0 |

Waves: all

HGMOVEYR contains the year of moving into the household of the person who answers the household questionnaire. For old households at their old address data is carried forward for up to two years. For new households in SOEP and for old households that have moved the variable is based on newly collected data. In case the information is missing and an old household has moved that year or the previous year, HGMOVEYR is given the value of the year of the respective wave.

The carrying forward of data entails the possibility that the year of moving into the new dwelling may lie before the year of birth of the oldest household member.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgutil - Other Monthly Utility Costs

| | |
|--|--------|
| 0 | 5 |
| 1 | 263 |
| 2 | 498 |
| 3 | 184 |
| 4 | 127 |
| 5 | 551 |
| 6 | 158 |
| 7 | 314 |
| 8 | 266 |
| 9 | 136 |
| 10 | 2135 |
| 11 | 154 |
| 12 | 674 |
| 13 | 173 |
| 14 | 177 |
| ... (335 rows omitted) | 152944 |
| 480 | 1 |
| 481 | 1 |
| 500 | 6 |
| 600 | 2 |
| 700 | 1 |
| 710 | 1 |
| 900 | 2 |
| 930 | 1 |
| -1 No answer | 8047 |
| -2 Does not apply | 146786 |
| -3 Not valid | 0 |
| -4 Forbidden Multiple Response | 0 |
| -5 Not Included In Questionnaire Version | 24257 |
| -6 Questionnaire Version With Modified Filter | 0 |
| -8 Question not part of the survey program this year | 37682 |

hgutilinfo - Reason for Missing Additional Utility Costs

| | |
|-------------------------------------|-------|
| 1 Utility Costs Unknown | 35 |
| 3 No Other Additional Utility Costs | 7933 |
| -1 No answer | 41842 |

| | | |
|----|---|--------|
| -2 | Does not apply | 28251 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 10125 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 287360 |

hgseval – Adequacy Of Living Space In Housing Unit

| | | |
|----|---|--------|
| 1 | Much Too Small | 10095 |
| 2 | A Bit Too Small | 56398 |
| 3 | Just Right | 263204 |
| 4 | A Bit Too Large | 36594 |
| 5 | Much Too Large | 4632 |
| -1 | No answer | 4623 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 0 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 0 |

Waves: all

Subjective assessment by respondent (household head). From Wave C to Wave G, the information on HGSEVAL was only collected for new households or households that had moved (households with “blue” questionnaires) and immobile households whose HGSIZE had changed. In these Waves, HGSEVAL is carried forward from previous years.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgsize – Size Of Housing Unit In Sq M

| | |
|------------------------|--------|
| 6 | 16 |
| 7 | 1 |
| 8 | 13 |
| 9 | 26 |
| 10 | 121 |
| 11 | 42 |
| 12 | 309 |
| 13 | 38 |
| 14 | 128 |
| 15 | 312 |
| 16 | 313 |
| 17 | 81 |
| 18 | 248 |
| 19 | 77 |
| 20 | 1026 |
| ... (305 rows omitted) | 369492 |
| 500 | 11 |
| 540 | 4 |
| 550 | 3 |

| | | |
|-----|---|------|
| 564 | | 2 |
| 650 | | 1 |
| 754 | | 1 |
| 838 | | 2 |
| 938 | | 1 |
| -1 | No answer | 3278 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 0 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 0 |

Waves: all

Up to Wave R (except for Wave O), HGSIZE was collected only in the first interview with new households. In case a household had moved or whether a household's housing unit size had changed due to renovations or additions (up to Wave G, these households filled out a "green" questionnaire). From Wave S onwards, the question on the size of the housing unit has been posed to all households annually. For households still residing at their old address and neither having moved nor renovated their dwelling, the information on the size of the housing unit is carried forward as a status variable in order to provide valid current information. Analogously, the information is carried back to fill gaps if households missed to state the size of the housing in the first year after moving.

In the case the information on the size of the housing unit is still missing, it is imputed simultaneously with other variables using Stata's chained imputation procedure. Imputed values are not carried forward. For information on the imputation procedure, see the description of HGRENT.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgf2rent – Detailed Imputation Flag, Rent Minus Heating Costs

| | | |
|----|--|--------|
| 0 | Not imputed | 301315 |
| 1 | Reported Rent Corrected for Imputed Heat | 28078 |
| 2 | Reported Rent Corrected for Imputed Util | 35315 |
| 3 | Reported Rent Corrected for Imputed Electr | 482 |
| 4 | Reported Rent Corrected for Heat and Util | 3515 |
| 5 | Reported Rent Corrected for Heat and Electr | 1493 |
| 6 | Reported Rent Corrected for Util and Electr | 4 |
| 7 | Reported Rent Corrected for Heat, Util and Electr | 3 |
| 8 | Reported Rent, at least one Sub-Aggregate Imputed | 3612 |
| 9 | Rent Imp., Sub-Aggregates Not Imputed | 567 |
| 10 | Rent Imp. and Corrected For Imp. Heat | 438 |
| 11 | Rent Imp. and Corrected For Imp. Util | 312 |
| 12 | Rent Imp. and Corrected For Imp. Electr | 4 |
| 13 | Rent Imp. and Corrected For Imp. Heat and Util | 112 |
| 14 | Rent Imp. and Corrected For Imp. Heat and Electr | 233 |
| 15 | Rent Imp. and Corrected For Imp. Util and Electr | 7 |
| 16 | Rent Imp. and Corrected For Imp. Heat, Util and Electr | 2 |
| 17 | Rent Imp. & at least one Imp. Sub-Aggregates | 54 |
| -1 | No answer | 0 |

| | | |
|----|---|---|
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 0 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 0 |

Waves: all

HGF2RENT indicates whether or not HGRENT has been imputed (=1) by Stata's chained imputation procedure.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgfelectr – Imputation Flag, Costs of Electricity

| | | |
|----|---|--------|
| 0 | Not imputed | 149163 |
| 1 | Imputed by PMM | 3306 |
| 2 | Electricity Costs from Previous or Subsequent Years | 2551 |
| -1 | No answer | 0 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 0 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 220526 |

Waves: Since 2010

HGFELECTR indicates whether HGELECTR value has been imputed (=1) by Stata's chained imputation procedure or whether it has been carried for- or backwards (=2) from past/subsequent waves.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgfheat – Imputation Flag, Costs Of Heating And Warm Water

| | | |
|----|---|--------|
| 0 | Not imputed | 322646 |
| 1 | Imputed by PMM | 22666 |
| 2 | Electricity Costs from Previous or Subsequent Years | 18991 |
| -1 | No answer | 0 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 11243 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 0 |

Waves: Since 1986

HGFHEAT indicates whether HGHEAT has been imputed (=1) by Stata's chained imputation procedure or whether it has been carried for- or backwards (=2) from past/subsequent waves.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgfrent – Imputation Flag, Amount Of Rent Minus Heating Costs

| | | |
|----|---|--------|
| 0 | Not imputed | 301496 |
| 1 | Rent and/or Sub-Aggregate Imputed by PMM | 57952 |
| 2 | Rent and/or Sub-Aggregate from Previous or Subsequent Years | 13695 |
| 3 | Rent and/or Sub-Aggregate From Prev./Subs. Years and/or PMM | 2403 |
| -1 | No answer | 0 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 0 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 0 |

Waves: all

HGFRENT indicates whether HGRENT or expenses added or subtracted to HGRENT has been imputed (=1) by Stata's chained imputation procedure, whether HGRENT itself or expenses added or subtracted to HGRENT was carried for- or backwards (=2) from past/subsequent waves or whether HGRENT is the result of a mixture of imputed and past/subsequent waves' values. The fourth category marks households who very likely did not include all utility costs in their overall housing costs in Wave BE (the question from which HGRENT is generated). We identified these households by their longitudinal information on housing costs and did not subtract HGHEAT and/or HGELECTR to obtain their HGRENT. HGF2RENT contains more detailed information about which imputed expenses were added or subtracted.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgfrroom – Imputation Flag, Number Of Rooms Lager Than 6 Sq M

| | | |
|----|---|--------|
| 0 | Not imputed | 373185 |
| 1 | Imputed | 2361 |
| -1 | No answer | 0 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 0 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 0 |

Waves: all

HGFROOM indicates whether or not HGROOM has been imputed (=1) by Stata's chained imputation procedure.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgfsz – Imputation Flag, Size Of Housing

| | | |
|----|----------------|--------|
| 0 | Not imputed | 374347 |
| 1 | Imputed | 1199 |
| -1 | No answer | 0 |
| -2 | Does not apply | 0 |

| | | |
|----|---|---|
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 0 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 0 |

Waves: all

HGFSIZE indicates whether or not HGFSIZE has been imputed (=1) by Stata's chained imputation procedure.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgfutil – Imputation Flag, Other Monthly Utility Costs

| | | |
|----|---|--------|
| 0 | Not imputed | 229638 |
| 1 | Imputed by PMM | 80237 |
| 2 | Electricity Costs from Previous or Subsequent Years | 13857 |
| -1 | No answer | 0 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 14132 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 37682 |

Waves: Since 1991

HGFUTIL indicates whether HGUTIL value has been imputed (=1) by Stata's chained imputation procedure or whether it has been carried for- or backwards (=2) from past/subsequent waves.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

6 Subsidization of Housing Costs

hgsubs – Amount Of Subsidies Last Year

| | | |
|----|---|--------|
| 1 | Yes | 13048 |
| 2 | No | 80190 |
| -1 | No answer | 1225 |
| -2 | Does not apply | 108728 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 0 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 172355 |

Waves: 2000–2014

Statement by respondent. HGOSUBS contains information on cash housing subsidies received by owners from the government during the year prior to the interview. Wave W was the last wave house owners could apply for the 8-years-long financial support such that since Wave BF, HGOSUBS is not asked any longer. Information is not carried forward.

Please note: The old variable \$FOERD (available until SOEP data release v25 in \$HGEN)

is discarded. Homeowner subsidies in Germany have been subject to major revisions and fluctuations over time. The corresponding question in SOEP was in some waves only posed to new households and those that have moved, in some waves it was not surveyed at all. For these reasons, the question for government housing subsidies was changed in Wave Q to cover direct subsidies received the previous year. HGSUBSID and HGOSUBS replace the old variable \$FOERD.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgreduc – Dwelling At A Reduced Prize

| | | |
|----|---|--------|
| 1 | Yes | 16273 |
| 2 | No | 169972 |
| -1 | No answer | 19821 |
| -2 | Does not apply | 153445 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 4792 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 11243 |

Waves: 1986–2002 (west), 1990 & 1992–2002 (east), since 2008 (both)

HGREduc states whether the flat is offered by the owner at a reduced rent. Information is carried forward from previous waves for old households residing at their old address. For newly surveyed households and for old households that have moved, newly collected data is used. From Wave T to Wave X, this information was not collected. It is carried forward from Wave S for households who have not moved and whose stated amount of rent vary only slightly. The new information from Wave Y is then carried backward for households with the same characteristics if HGREduc is still missing after carrying forward from Wave S.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgrsubs – Government Subsidized Housing

| | | |
|----|---|--------|
| 1 | Yes, With Due Diligence | 19240 |
| 2 | Yes, With Run Out Due Diligence | 4896 |
| 3 | No | 182451 |
| -1 | No answer | 3805 |
| -2 | Does not apply | 155607 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 9547 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 0 |

Waves: Since 1984 (west), 1990 & since 1994 (east)

HGRSUBS states whether the rent is subsidized by government. Up to Wave K (1994), the question was asked only to new households and households who had moved since last year. For the remaining households, the information is carried forward from previous waves. East German households were asked for the first time in Wave J (1993).

The second response category, which indicates expired subsidization, was added in Wave L (1995). In Wave BG (2016), the category was removed. In Germany, rent subsidization is

typically set up as a system of subsidized loans to real estate investors and an obligation to offer below market rents. The obligation expires after one to several decades. In principle, there is the possibility that households say to live in a subsidized flat, although subsidization has expired.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgsubsid – Government Subsidizes Housing Payments

| | | |
|----|---|--------|
| 1 | Yes | 5274 |
| 2 | No | 8767 |
| -1 | No answer | 11769 |
| -2 | Does not apply | 66348 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 21422 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 261966 |

Waves: 1984–1999 (west), 1994–1999 (east)

Statement by respondent. HGSUBSID contains information on government subsidies at the time the housing was built or bought. From Wave B to Wave N, this was only asked to new households or in case an old household had moved. Information is then carried forward. In Wave O and Wave P, the question was again posed to the whole population.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

7 Typology of Household

hgowner – Tenant Or Owner Of Dwelling

| | | |
|----|---|--------|
| 1 | Owner | 157315 |
| 2 | Main Tenant | 202700 |
| 3 | Sub-Tenant | 8716 |
| 4 | Tenant | 103 |
| 5 | [de] Heimbewohner oder Gemeinschaftsunterkunft | 1891 |
| -1 | No answer | 29 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 4792 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 0 |

Waves: all

Up to Wave H, the information for HGOWNER was collected in separate questionnaires for “old” and first-time respondents, respectively (“blue” and “green” questionnaires). In all waves and for each household with the original value of -1 (“no answer”), codes are set to =1 and =4 if at least one owner- or tenant-specific item is given. Code 4 is also used if a change in ownership (from owner to tenant) has taken place but no original information for HGOWNER was given. Code 5 (‘resident of a home or institutional living facility’) has only

been assigned by interviewers during fieldwork since Wave P.
 For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgtyp1hh – Household Typology, One Digit

| | | |
|----|---|--------|
| 1 | 1-Pers.-HH | 89979 |
| 2 | Couple Without Children | 105355 |
| 3 | Single Parent | 30718 |
| 4 | Couple With Children LE 16 | 84539 |
| 5 | Couple With Children GT 16 | 34135 |
| 6 | Couple With Children LE And GT 16 | 20197 |
| 7 | Multiple Generation-HH | 4547 |
| 8 | Other Combination | 6076 |
| -1 | No answer | 0 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 0 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 0 |

Waves: all

HGTYP1HH is an aggregation of HGTYP2HH. For further information on HGTYP1HH, please refer to HGTYP2HH.

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

hgtyp2hh – Household Typology, Two Digit

| | | |
|-----|--|--------|
| 11 | 1-Person HH Male LE 35 Y. | 12886 |
| 12 | 1-Person HH Male 35- LT 60 Y. | 15577 |
| 13 | 1-Person HH Male GE 60 Y. | 10358 |
| 14 | 1-Person HH - Female LE 35 Y. | 9377 |
| 15 | 1-Person HH Female 35- LT 60 Y. | 11800 |
| 16 | 1-Person HH Female GE 60 Y. | 29981 |
| 21 | Couple Without Children | 105355 |
| 31 | Single Parent,1 Child, LE 16 | 9778 |
| 32 | Single Parent,2 Or More Children, LE 16 | 5959 |
| 33 | Single Parent,1 Children, GT 16 | 9112 |
| 34 | Single Parent,2 Or More Children, GT 16 | 2207 |
| 35 | Single Parent,2 Children, LE and GT 16 | 2080 |
| 36 | Single Parent,3 Or More Children, LE and GT 16 | 1582 |
| 41 | Couple, 1 Child, LE 16 | 31317 |
| 42 | Couple, 2 Children, LE 16 | 36108 |
| ... | (3 rows omitted) | 49445 |
| 53 | Couple, 3 Or More Children, GT 16 | 1804 |
| 61 | Couple, 2 Children, LE and GT 16 | 9706 |
| 62 | Couple, 3 Or More Children, LE and GT 16 | 10491 |
| 71 | 3-Generation-HH | 4095 |
| 72 | 4-Generation-HH | 31 |

| | | |
|----|---|------|
| 73 | GrandParents-GrandChildren-HH | 421 |
| 81 | Other Combination Without K. LE 16 | 6076 |
| 82 | Other Combination With K. LE 16 | 0 |
| -1 | No answer | 0 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 0 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 0 |

Waves: all

Generated variable created by combining the relationships of all persons living in the household to the head of household (Variable \$STELL in the file \$PBRUTTO) at the time of the survey. With data release v26 the data production process switched to a standardized procedure for all Waves to ensure longitudinal consistency, resulting in minor changes compared with older distributions. HGTYP1HH is an aggregation of HGTYP2HH and corresponds to the first digit of the two-digit code of HGTYP2HH. Single households are differentiated in HGTYP2HH according to both gender and age. The following abbreviations are used in the labels: LE=less or equal, LT=less than, GE=greater or equal, GT=greater than, u.m. = and more, 1-P-HH/1-Pers.-HH = one-person household, K. = children, Alleinerz. = single parent, J. = years, HH = Haushalt, sonst. = sonstige.

Help for old friends: Starting with data distribution v26 (Wave A to Wave Z), the category "(88) Other combination" has been further differentiated into (81) households without vs. (82) those with children (up to the age of 16).

For more information, contact: Carsten Schröder, <cschroeder@diw.de>

8 Household Income

hghinc – Monthly Household Net Income (EUR)

| | |
|-------------------------|--------|
| 0 | 67 |
| 1 | 1 |
| 8 | 1 |
| 10 | 1 |
| 20 | 1 |
| 26 | 5 |
| 28 | 1 |
| 31 | 2 |
| 40 | 1 |
| 49 | 1 |
| 50 | 2 |
| 51 | 5 |
| 52 | 1 |
| 62 | 1 |
| 66 | 2 |
| ... (5498 rows omitted) | 356114 |
| 60000 | 6 |
| 65000 | 2 |
| 70000 | 1 |

| | |
|--|-------|
| 85000 | 1 |
| 90000 | 1 |
| 99999 | 2 |
| 200000 | 1 |
| 999999 | 1 |
| -1 No answer | 17718 |
| -2 Does not apply | 301 |
| -3 Not valid | 1306 |
| -4 Forbidden Multiple Response | 0 |
| -5 Not Included In Questionnaire Version | 0 |
| -6 Questionnaire Version With Modified Filter | 0 |
| -8 Question not part of the survey program this year | 0 |

This variable contains the current monthly net household income asked for in the household questionnaire, always provided in Euros which was introduced in January 2002 (1 Euro = 1.95583 DM). Income is reported by the respondent (head of household).

For more information, contact: Jan Goebel (Tel. +49-30-89789-377), <jgoebel@diw.de>

hgi1hinc - 1. Imputed Monthly Net Household Income (EUR) [1/5]

| | |
|--|--------|
| 0 | 39 |
| 1 | 1 |
| 10 | 1 |
| 20 | 1 |
| 28 | 1 |
| 40 | 1 |
| 50 | 2 |
| 51 | 1 |
| 70 | 1 |
| 72 | 1 |
| 77 | 1 |
| 78 | 1 |
| 80 | 1 |
| 84 | 3 |
| 87 | 1 |
| ... (5376 rows omitted) | 310988 |
| 55000 | 2 |
| 60000 | 7 |
| 65000 | 2 |
| 70000 | 1 |
| 85000 | 1 |
| 90000 | 1 |
| 99999 | 2 |
| 999999 | 1 |
| -1 No answer | 4 |
| -2 Does not apply | 64481 |
| -3 Not valid | 0 |
| -4 Forbidden Multiple Response | 0 |
| -5 Not Included In Questionnaire Version | 0 |
| -6 Questionnaire Version With Modified Filter | 0 |
| -8 Question not part of the survey program this year | 0 |

Multiple imputation procedures provide a way to deal with missing values on the variable Current Monthly Net Household Income by using information about components and determinants of the household income and replacing item-nonresponse with multiply imputed data. The first five imputations are available within the HGEN dataset: the variables HGI1HINC-HGI5HINC.

The imputations were calculated using multiple imputations by chained equations. Up to Wave BB the program ICE of Stata which was written by Patrick Royston (see Royston 2004, 2005a, 2005b) and which is based on the program MICE in S-Plus and R was used. Since Wave BC the Stata command `mi impute` is used. The missing observations are assumed to be missing at random. We set the number of imputations $m=10$ and get 10 multiple imputed values for IHINC\$\$\$. For a discussion on the choice of m , see Rubin (Wave D) and Royston (Wave U).

The dataset MIHINC contains the complete imputation results and is separately available. To be compatible with methods for analyzing multiply imputed data, MIHINC is constructed in the so called stacked or MIM Dataset Format. It contains the following variables: HHNRAKT, SVYYEAR, MJ, MI, IHINC and IMPFLAG. For all household from Wave L to Wave BD, there are ten imputed values for the current household income. MJ identifies the individual dataset to which each observation belongs while MI identifies the observations within each individual dataset. To distinguish between the original data containing missing values and the imputed values, the dummy variable IMPFLAG is added. In the \$HGEN files five of these imputed incomes are stored in the conventional wide format.

The number of iterations carried out in each prediction model was specified to be 500. For East and West Germany, imputations were done separately. Furthermore, the option for predicted mean matching was chosen which means that for each missing observation on income, the particular non-missing observation is found whose prediction on observed data is closest. This closest observation is used to impute the missing value.

Most important variables for modelling the current household net income consist in the household net income of the previous year, in basic information about the household and changes in its composition as well as all relevant income components received.

####The complete list of the variables used for modelling

- Description of household:
 - size, number of children, sample
 - head of household: not German, age, sex
 - changes in household composition between years: births, deaths, persons entering or leaving the household or being temporarily absent
- Financial Situation:
 - Monthly household income previous year
 - Income from employment
 - Pensions
 - Sum of personal incomes (e.g. Support from the “Arbeitsamt”, Maternity benefit, Alimony, etc.)
 - Household related incomes (e.g. Child allowance, Housing assistance, Social assistance, Unemployment benefit, Assets, etc.)
 - Fraction of persons greater than 16 in household who refused answering a component of income (0-1)
- Number of persons not attended survey (PUNR, partial unit nonresponse)
- Cross-sectional weights

####Analyzing multiply imputed data

For analyzing multiple imputed data, you do not necessarily need special methods, however such tools exist and simplify the use of multiply imputed data. Below is given a short overview of some useful tools for various statistical packages. These tools estimate the parameters of a regression model by combining the estimates across the several replicates of imputation. Point estimates from multiple imputations are then the arithmetic mean of the several point estimates obtained from analysis on each imputed data. Standard errors are obtained by combining the average of the squared standard errors of the several (m) estimates with the within- and between-imputation variance.

- Stata provides various built-in functionality called mi.
- Within SAS, the MIANALYZE procedure combines the results of the analyses of imputations and generates valid statistical inferences: <http://support.sas.com/rnd/app/stat/procedures/mianalyze.html>
- IVEware is a set of routines that can be launched from SAS or run independently using data from many sources. You can use the IVEware module regress to perform multiple imputation analysis.

Royston, Patrick (2004): *Multiple imputation of missing values*. In: *Stata Journal* 4(3): 227-241.

Royston, Patrick (2005a): *Multiple imputation of missing values: update*. In: *Stata Journal* 5(2): 188-201. Royston, Patrick (2005b): *Multiple imputation of missing values: Update of ice*. In: *Stata Journal* 5(4): 527-536. Rubin, D.B. (1987): *Multiple imputation for non-response in surveys*. New York.

For more information, contact: Jan Goebel (Tel. +49-30-89789-377), <jgoebel@diw.de>

hgi2hinc - 2. Imputed Monthly Net Household Income (EUR) [2/5]

| | |
|-------------------------|--------|
| 0 | 40 |
| 1 | 1 |
| 10 | 1 |
| 20 | 1 |
| 28 | 1 |
| 40 | 1 |
| 50 | 2 |
| 51 | 1 |
| 70 | 1 |
| 72 | 1 |
| 77 | 1 |
| 78 | 2 |
| 80 | 1 |
| 84 | 3 |
| 87 | 1 |
| ... (5376 rows omitted) | 310987 |
| 55000 | 2 |
| 60000 | 6 |
| 65000 | 2 |
| 70000 | 1 |
| 85000 | 1 |
| 90000 | 1 |
| 99999 | 2 |

| | | |
|--------|---|-------|
| 999999 | | 1 |
| -1 | No answer | 4 |
| -2 | Does not apply | 64481 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 0 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 0 |

hgi3hinc – 3. Imputed Monthly Net Household Income (EUR) [3/5]

| | | |
|--------|---|--------|
| 0 | | 41 |
| 1 | | 1 |
| 10 | | 1 |
| 20 | | 1 |
| 28 | | 1 |
| 40 | | 1 |
| 50 | | 2 |
| 51 | | 1 |
| 70 | | 1 |
| 72 | | 1 |
| 77 | | 1 |
| 78 | | 1 |
| 80 | | 1 |
| 84 | | 2 |
| 87 | | 2 |
| ... | (5376 rows omitted) | 310986 |
| 55000 | | 2 |
| 60000 | | 7 |
| 65000 | | 2 |
| 70000 | | 1 |
| 85000 | | 1 |
| 90000 | | 1 |
| 99999 | | 2 |
| 999999 | | 1 |
| -1 | No answer | 4 |
| -2 | Does not apply | 64481 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 0 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 0 |

hgi4hinc – 4. Imputed Monthly Net Household Income (EUR) [4/5]

| | | |
|----|--|----|
| 0 | | 40 |
| 1 | | 2 |
| 10 | | 1 |
| 20 | | 1 |
| 28 | | 2 |

| | | |
|--------|---|--------|
| 40 | | 1 |
| 50 | | 3 |
| 51 | | 1 |
| 70 | | 1 |
| 72 | | 1 |
| 77 | | 1 |
| 78 | | 1 |
| 80 | | 1 |
| 84 | | 2 |
| 87 | | 1 |
| ... | (5376 rows omitted) | 310985 |
| 55000 | | 2 |
| 60000 | | 7 |
| 65000 | | 2 |
| 70000 | | 1 |
| 85000 | | 1 |
| 90000 | | 1 |
| 99999 | | 2 |
| 999999 | | 1 |
| -1 | No answer | 4 |
| -2 | Does not apply | 64481 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 0 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 0 |

hgi5hinc – 5. Imputed Monthly Net Household Income (EUR) [5/5]

| | | |
|-------|---------------------|--------|
| 0 | | 39 |
| 1 | | 1 |
| 10 | | 1 |
| 20 | | 1 |
| 28 | | 1 |
| 40 | | 1 |
| 50 | | 2 |
| 51 | | 1 |
| 70 | | 1 |
| 72 | | 1 |
| 77 | | 1 |
| 78 | | 1 |
| 80 | | 1 |
| 84 | | 2 |
| 87 | | 3 |
| ... | (5376 rows omitted) | 310987 |
| 55000 | | 2 |
| 60000 | | 6 |
| 65000 | | 3 |
| 70000 | | 1 |
| 85000 | | 1 |

| | | |
|--------|---|-------|
| 90000 | | 1 |
| 99999 | | 2 |
| 999999 | | 1 |
| -1 | No answer | 4 |
| -2 | Does not apply | 64481 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 0 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 0 |

hgfhinc – Imputation Flag, Monthly Net Household Income

| | | |
|----|---|--------|
| 0 | Not imputed | 359006 |
| 1 | Imputed | 16533 |
| -1 | No answer | 3 |
| -2 | Does not apply | 4 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 0 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 0 |

HGFHINC is a dummy variable indicating whether an observation was missing on HGH-INC and was therefore imputed or not.

For more information, contact: Jan Goebel (Tel. +49-30-89789-377), <jgoebel@diw.de>

9 Time and Method of Interview

iyear – Household Related Survey Year

| | | |
|------|---|--------|
| 2017 | | 18771 |
| 2018 | | 19033 |
| 2019 | | 579 |
| -8 | Question not part of the survey program this year | 337163 |

hghmonth – Month Of Interview

| | | |
|----|-----------|-------|
| 1 | January | 23569 |
| 2 | February | 93682 |
| 3 | March | 94602 |
| 4 | April | 55352 |
| 5 | May | 33018 |
| 6 | June | 25211 |
| 7 | July | 17969 |
| 8 | August | 12064 |
| 9 | September | 8693 |
| 10 | October | 5935 |
| 11 | November | 3421 |

| | | |
|----|---|------|
| 12 | December | 1992 |
| -1 | No answer | 37 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 1 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 0 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 0 |

Waves: all

The month of participation in the survey is generated using data from the household questionnaire. Missing information is filled in using data from the HBRUTTOL file. Interviews that took place in the month of December, and prior to the 20th of that month, were recoded to -3.

For more information, contact: Peter Krause (Tel. +49-30-89789-690), <pkrause@diw.de>

hghmode – Interview Method

| | | |
|-----|---|--------|
| 100 | With Interviewer Assistance | 4262 |
| 110 | Oral Interview | 101115 |
| 120 | Written Ques. Interviewer | 57988 |
| 131 | Written Ques. No Interviewer | 6026 |
| 132 | Oral And Written | 7082 |
| 133 | Proxy | 12 |
| 134 | Third Person Present | 1334 |
| 135 | No Third Person Present | 9175 |
| 140 | CAPI - Since 1998 (O) | 142577 |
| 150 | Cawi Since 2014 (BE) | 4286 |
| 210 | Written, By Mail | 41416 |
| 220 | Telephone Interview | 262 |
| -1 | No answer | 11 |
| -2 | Does not apply | 0 |
| -3 | Not valid | 0 |
| -4 | Forbidden Multiple Response | 0 |
| -5 | Not Included In Questionnaire Version | 0 |
| -6 | Questionnaire Version With Modified Filter | 0 |
| -8 | Question not part of the survey program this year | 0 |

Waves: all

The interview method is generated through data from the household questionnaire. Missing information is filled in with data from the corresponding HBRUTTO file.

For more information, contact: Peter Krause (Tel. +49-30-89789-690), <pkrause@diw.de>

inputdataset – Source dataset

| | |
|--------|-------|
| ahgen | 5921 |
| bahgen | 15414 |
| bbhgen | 16819 |
| bchgen | 16508 |
| bdhgen | 18093 |

| | |
|----------------------|-------|
| behgen | 16037 |
| bfhgen | 15991 |
| bghgen | 17775 |
| bhgen | 5322 |
| bhhgen | 19701 |
| bihgen | 18682 |
| chgen | 5090 |
| dhgen | 5026 |
| ehgen | 4814 |
| fhgen | 4690 |
| ... (5 rows omitted) | 33615 |
| lhgen | 6968 |
| mhgen | 6894 |
| nhgen | 6810 |
| ohgen | 7668 |
| phgen | 7388 |
| qhgen | 13249 |
| rhgen | 11947 |
| shgen | 12692 |
| thgen | 12061 |
| uhgen | 11796 |
| vhgen | 11440 |
| whgen | 12499 |
| xhgen | 11689 |
| yhgen | 11058 |
| zhgen | 11889 |