Understanding Society: Linked Nest Auto-enrolment Pensions Dataset, 2014-2022: Secure Access

USER GUIDE

Version 1, July 2023

Table of Contents

1. Introduction

Data linkage is one of the key innovative features of Understanding Society, the UK Household Longitudinal Study (UKHLS) housed at the Institute for Social and Economic Research (ISER), and allows researchers to develop and implement new research agendas. It includes linkage to administrative datasets in key policy areas as well as other external data sources. Understanding Society currently pursues data linkage in a wide range of topic areas, including education, health, economic circumstances, and transport, as well as area characteristics. Further information is provided on Understanding Society website.

The Nest linked dataset provides data from the Nest pensions accounts of 1672 Understanding Society respondents who have consented to have these data linked and made available for research. Nest (the National Employment Savings Trust) is a large workplace pension provider set up by the UK government to accept enrolments from workers who are eligible for enrolment into a workplace pension, under the UK's national programme of pensions auto-enrolment.

The linked data are available via the UK Data Service under their Secure Access category. Researchers are required to attend a Safe Researcher training course to access the Study. Further information is provided in the Data Access section below.

2. Background to Nest and auto enrolment

2.1 Employers' automatic enrolment duties

The UK government introduced workplace pensions auto enrolment on a phased basis, starting in October 2012. Larger employers were brought into the programme first, with employers of all sizes becoming covered by the rules by 2017. Under the legislation, every employer must automatically enrol all their eligible workers into a suitable pension scheme. To be eligible, a worker must:

- Be aged between 22 and State Pension age, and
- Earn a salary equivalent to £10,000 or more a year from this employer[1].

[1] The earnings threshold for enrolment is subject to annual review, and the other criteria may be modified in future by the Department for Work and Pensions. The criteria listed here are true as at the time of writing, March 2023.

Workers who do not meet either, or both, of these eligibility criteria can also ask to be enrolled by their employer if they wish. Employers may also choose to enrol these lower earners.

Under the legislation, employers must make contributions into the pension accounts of all auto-enrolled workers, every pay period. For defined contribution (DC) schemes like Nest, the minimum level for these contributions is set out in legislation as a percentage of the worker's eligible ('pensionable') earnings. The level of minimum mandatory contributions was increased in phases:

- During the initial employer rollout, the minimum total contribution was 2% of qualifying earnings, with at least 1% contributed by the employer
- From April 2018, this rose to 5%, with at least 2% contributed by the employer
- From April 2019, this rose again, to 8%, with at least 3% contributed by the employer.

Workers usually receive tax relief from the government on their contributions, so that their gross contribution of 5% equates to a net contribution of 4%. Most employers apply these percentages not to a worker's total earnings, but rather to a band of earnings set out in legislation[2].

Workers can opt out of auto enrolment within one month of being enrolled, meaning that no contributions are paid into the scheme by this employer, and any contributions already made are refunded. The effect of this opt-out is to leave the worker in the same position they would be in if they'd never been enrolled. They can stop contributions at any time, but if they do this after this initial one-month window, all the contributions already made remain in the pension scheme. Workers can't usually access the money in their pension pots until they reach age 55.

To learn more about the auto-enrolment policy, and in particular the rules surrounding enrolment and opting out, see the PDF document Essentials of the UK retirement system or visit thepensionsregulator.gov.uk.

2.2 Employers using Nest

Many employers have turned to Nest to fulfil their automatic enrolment duties. As of 31 March 2022 around one million had registered with the scheme. This figure is based on all employers who have registered with Nest and are believed to be active. However, it also includes employers who have no current eligible workers enrolled and a small number who may have ceased trading or become insolvent which Nest is not yet aware of. Employers

[2] For the tax year 2023-24, the annual lower limit of this earnings band is £6,240/yr, and the upper band £50,000/yr. Expressed in terms of monthly earnings, the band falls between £540/m and £4,189/m. So monthly earnings of £300, when expressed as pensionable earnings, would equate to £0, monthly earnings of £2000 would equate to £1460, and monthly earnings of £5000 would equate to £3,649.

appearing in the Nest data module, though, have by definition enrolled at least one individual because at least one Understanding Society respondent has an administrative record containing data on this enrolment.

3. Consenting to linkage to the Nest data: question and outcomes

The consent question concerning linkage to the Nest data was asked during Understanding Society's Wave 11 (NESTLINK).

The exact phrasing of the consent question is as follows:
"We would like to add any records held by the National Employment Savings Trust (NEST), the government-backed pension scheme, to the responses you have given to this study. If you have an account with NEST, these records will include information they hold on your pension savings through employers who've enrolled you into NEST, and any savings you've made yourself into the scheme. Linking the information will help us to get a fuller picture about who saves with NEST and how their retirement savings are building up. All information will be used for research purposes only. It will be used by academic or policy researchers under restricted access arrangements which make sure that the information is used responsibly and is held securely. If you decide that you no longer wish to take part in the study you can withdraw your consent at any time. If you withdraw consent, we will retain your NEST information up until that point after which time no further links to the personal data will be made.

"Please read this leaflet explaining how we would like to attach your NEST records to the answers you have given in this study and ask me any questions.

"Under Data Protection Legislation we need you to give your consent for us to pass your name, post code, sex and date of birth to the National Employment Savings Trust (NEST) Corporation or a contractor conducting the linkage, so that they can identify whether you have a NEST account, and if so, to link this information to your survey responses. Are you happy to give us your consent?"

12,832 Wave 11 survey respondents consented to linkage to their Nest records.

4. Linkage of consenting survey respondents to their Nest data

Where study participants provided consent (and had not revoked this consent since providing it), ISER generated an anonymised ID and extracted Forename, Surname, Date of Birth, Postcodes and Address details from our respondent database. This information was then sent to Nest who matched the records to their files using the identifying information. The identifying information was then deleted before the file was sent across (with the anonymised ID only) to ISER.

1672 of the 12,701 (continuing) consenters (13.2%) were linked to their Nest records.

There are a variety of reasons why a respondent was not linked to a Nest account. Whilst there will be genuine failures to match a respondent to their account, chief among the reasons is likely to be that whilst a respondent agreed to the linkage, they may never have had a Nest account. Nest account holders are much different from the general UK population profile - and, indeed, the profile of consenting Understanding Society participants - being both younger and - by the scheme's design - working in roles and employment sectors where saving into workplace pensions has been traditionally less common.

5. Weights for analysis of linked data

As noted in the previous sections, the consent and linkage processes potentially lead to non-random selection of Understanding Society participants for whom linked data exists. Hence, weights are also supplied with the dataset that map linked respondents to the population of Nest pension holders, enabling inference about that population. If you undertake an unweighted analysis of the data, you should be clear on the assumptions that justify such an analysis. A discussion of these assumptions in the context of survey non-response (an analogous process to non-linkage) in the Understanding Society main survey can be found in the Main Survey User Guide (Institute for Social and Economic Research 2019).

Two sets of weights are supplied. These weight all relevant (see definition below) Understanding Society participants linked to Nest pension information aged under 70 at December 2020 (the time of consent to linkage): the scheme only started in 2014, so very few older individuals hold such pensions, and only 5 were linked, negatively affecting weight estimation when included in computations. The first set maps all linked Understanding Society participants aged under 70 to the Nest population. These weights (k_indpeukui_lxw), which are supplied for 1667 linked individuals in the dataset, are termed cross-sectional weights. The second set of weights map only linked individuals aged under 70 who responded to all UKHLS main survey waves from 2014 to 2020 (specifically, waves 6, 7, 8, 9, 10 and 11) to the Nest population. These latter weights (k_indpeukui_llw), which are supplied for 912 linked individuals, are termed longitudinal weights. They enable full panel analyses of subjects for whom there are Understanding Society survey responses over the period covered by the Nest dataset.

Weight computation incorporated both the consent and linkage elements of the linkage process, and was as follows. First, (cross-sectional or longitudinal) participant consent to linkage to Nest information, conditional on main survey Wave 11 response, was modelled using a probit model. Model selection was undertaken via an a priori Lasso procedure. Predictors were taken from subject main survey Wave 11 responses, and included demographics, household composition, previous survey outcomes and other paradata, and economic, health and attitudinal variables. Second, participant linkage to Nest information, conditional on consenting to linkage, was similarly modelled, using the same predictors.

Third, inverse propensity (IP) non-linkage weights for linked subjects were computed as the product of their supplied main survey Wave 11 non-response weights (which account for selection into and non-response to survey), the inverse of their consent propensities as estimated above, and the inverse of their linkage propensities as also estimated above (see Benzeval et al. 2023 for the use of these methods in the context of survey non-response). The main survey Wave 11 cross-sectional weight (k_indinui_xw) was used as the 'input' weight in the calculation of the cross-sectional weights, and the main survey Wave 10 longitudinal weight (k_indinui_lw) as the 'input' weight in the calculation of the longitudinal weights.

Fourth, cross-sectional IP weights could not be computed for 15 relevant individuals because they lacked main survey 'input' weights (see Institute for Social and Economic Research 2019 for reasons). To weight these subjects, weight sharing methods that share with unweighted subjects the available IP weights of subjects with similar characteristics were used (see Institute for Social and Economic Research 2022 for more details of these methods in the context of non-response weights). The shared weights were then scaled to have a mean of one.

The (shared) IP weights map linked subjects to the UK population, of which Nest pension holders are a non-random subset (see previous section). Hence, fifth, these weights (in effect, selection weights in this scenario) were adjusted to meet Nest sex, age and region entire 'population' benchmark totals in December 2020. Post-stratification methods could not be used (see Valliant & Dever 2018) for this because they adjust weights to target population conditional totals (sex * age * region conditional totals in the Nest case). Whilst such information exists for the Nest population, small dataset sizes meant that some linked dataset categories included no subjects. Consequently, Entropy Balance Weighting (EBW) methods (Hainmeuller 2012; Hainmeuller & Xu 2013; see Rothbaum et al 2020 for an empirical application involving non-response to the American Community Survey) were utilised. These employ an iterative re-weighting procedure to enable weights to instead be adjusted to population marginal totals. In the Nest case, these were population sex (2 categories), age (5 categories) and region (12 categories) totals. Information on the (high) performance of the EBW weights in meeting NEST population benchmark totals is reported in Table 1.

As of this release, these weights are made available as beta-versions. Refinement of the weighting models continues, and updates with future data releases are likely. The weights scale linked subjects to the Nest population in December 2020. That is, using the weights will provide estimates that are representative of all Nest pension holders at that point in time.

Table 1. Understanding Society – Nest linked dataset cross-sectional and longitudinal EBW non-linkage weight performance in recovering Nest population benchmarks.

Column (i) reports the (entire) Nest population benchmarks to which the EBW weighting process adjusted the computed IP weights and against which the EBW weights were evaluated. Columns (ii) (for the cross-sectional weights) and (v) (for the longitudinal weights) report the linked dataset EBW weighted estimates of these quantities, columns (iii) and (vi) the P-values of T-tests comparing these weighted estimates to the Nest population benchmarks, and columns (iv) and (vii) the absolute relative biases of the weighted estimates compared to the benchmarks (i.e. abs ((EBW estimate – Nest benchmark) / Nest benchmark)). No differences between the EBW estimates and the Nest benchmarks were statistically significant, and no absolute relative bias was above 7%.

		Cross-sectional weights			Longitudinal weights		
	(I)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
Quantity	Nest popn.	EBW wt. est.	P-value	Abs. rel. bias	EBW wt. est.	P-value	Abs. rel. bias
Males	0.530	0.531	0.945	0.002	0.530	0.983	0.000
Females	0.469	0.469	0.984	0.000	0.470	0.937	0.002
Age: 16-29	0.264	0.263	0.914	0.004	0.261	0.852	0.011
Age: 30-39	0.301	0.302	0.946	0.003	0.300	0.951	0.003
Age: 40-49	0.202	0.202	0.997	0.000	0.202	0.975	0.000
Age: 50-59	0.169	0.169	0.994	0.000	0.170	0.931	0.006
Age: 60-69	0.063	0.064	0.819	0.016	0.067	0.630	0.063
North East	0.033	0.033	1.000	0.000	0.033	0.989	0.000
North West	0.101	0.103	0.791	0.020	0.103	0.839	0.020
Yorkshire	0.075	0.076	0.872	0.013	0.076	0.892	0.013
East Midlands	0.078	0.079	0.846	0.013	0.079	0.892	0.013
West Midlands	0.088	0.090	0.778	0.023	0.090	0.824	0.023
E. England	0.091	0.093	0.782	0.022	0.093	0.838	0.022
London	0.173	0.176	0.757	0.017	0.175	0.854	0.012
South East	0.129	0.131	0.814	0.016	0.131	0.872	0.016
South West	0.083	0.085	0.765	0.024	0.086	0.784	0.036
Wales	0.041	0.043	0.733	0.049	0.043	0.799	0.049
Scotland	0.064	0.066	0.745	0.031	0.066	0.822	0.031
N. Ireland	0.025	0.025	0.945	0.000	0.025	0.941	0.000

6. Overview of the Nest data

The Nest data module is a panel with 3 dimensions:

- Individuals
- Employers
- Time in months

along with time invariant data related to each individual and employer. The next section of this guide contains a detailed account of the variables contained in this panel. This section covers some high-level considerations about the data.

For each individual, the data show[3]:

- **Involvement in and interaction with the auto enrolment system**, including: the start and end months of all enrolments into the Nest scheme; contributions paid by the individual, the employer and the government (in tax relief); savings balances; withdrawals; and other behaviours within the Nest scheme.
- **Start and end date[4] of any employments** with employers using Nest as their auto enrolment pension scheme.
- **Monthly pensionable earnings** received in each of these employments[5] – as normalised from weekly, fortnightly or monthly payroll data received by Nest from the employer.

Given this scope, research applications for the data are likely to involve the study of one or more of the following areas:

- participation in, and levels of savings and withdrawals within, the UK's pensions auto enrolment system
- workers' behaviour within the workplace pensions system
- labour market dynamics, including earnings patterns broken down into individual employments.

---

[3] More detailed descriptions of the data, including definitions of variables, appear later in this document.

[4] End dates of employment periods are not explicitly captured in the Nest data. Snapshots of the data are taken at each month-end. A change in the employment status from the previous month indicates if the employment has ended.

[5] In certain cases an employers does not need to provide the earnings data on a worker, but in the great majority of cases, these data are included.

6.1 Considerations and limitations

The Nest population comprises over 11 million people who are employed in the UK. This is a significant proportion of the UK working population. However, researchers should note that the Nest dataset has a number of structural idiosyncrasies and limitations of scope. For information on the reasons behind these structural factors, see 2. Background to Nest.

These factors can be managed and controlled for, provided they are properly understood by the researcher, so a careful reading of this manual, and the documentation it links to, is recommended to avoid misinterpretation of the data.

Three key issues to bear in mind when considering use of the dataset are:

- **Sample of individuals** Although Nest is the largest automatic enrolment scheme, its population is not fully representative of the UK workforce, or of pension scheme participants. Due to the structure of the pensions market and the design of the auto enrolment policy (see 2. Background to Nest) the Nest membership over-represents certain subgroups who were previously under-represented in the workplace pension market:
    - Younger people
    - Employees at small firms
    - Lower- and middle-income employees at large firms
    - Those in short-tenure roles at large firms

For details on the composition of this population, see Nest's annual statistical report.

- **Sample of employments** The Nest data relate to employments that this group of individuals have had with employers who are using Nest as their auto enrolment pension ('Nest employers'). Data are supplied by Nest employers as part of the administration of the Nest scheme. The dataset does not contain any data from employers who are **not** using Nest ('Non-Nest employers'); or on any workers who are employed by Nest employers, but who are enrolled into another pension scheme[6] or have opted out of saving with Nest. As a result, there are gaps in the employment and contribution histories of most individuals. This includes:

[6] Large employers often have more than one pension scheme, with different schemes offered to different segments of their workforce. In these cases, Nest might be the scheme for lower-income or more transient segments of the workforce, while more senior staff, or head office professionals, are enrolled into another scheme, often at higher levels of employer contribution.

- Employments with Nest employers where the individual never met the eligibility criteria for enrolment (see section 2.)
- Employments with Nest employers where the individual opted out of saving in Nest
- Employments with non-Nest employers, whether these occurred **between** employments with Nest employers, or **concurrently** *with* employments with Nest employers; and
- Self-employed work[7].

- **Use of (banded) pensionable earnings** The earnings data in the Nest dataset is highly granular, as it is sourced directly from employers' payroll systems on a pay-period-by-pay-period basis. As such it gives a detailed picture of ongoing levels of, and fluctuations in, pensionable earnings that each individual has received from each employer. However, researchers should note that these data do not usually represent the **total** earnings received from Nest employers.
  - As noted in section 2, in the auto-enrolment system, employers generally calculate contribution amounts using a band of their workers' gross earnings[8]. As a result, if a researcher wishes to know an individuals' **total** gross earnings from a given employment, they will need to infer this by adding on the earnings that have been excluded because of the band.
  - There are limitations to this approach: for instance, in a pay period where an individual earns below the lower end of the earnings band, their pensionable earnings will appear as zero, meaning it is not possible to distinguish between a month of zero gross earnings, and a month of non-zero gross earnings that fell below the band.
  - This also means that estimates of monthly gross earnings that are calculated in this way will have a floor of £540 a month (using 2023–24 figures). Annually, this implies a floor of £6,240.
  - Similarly, any level of earnings an individual receives over the monthly upper limit of the band would be translated into £3,649 of pensionable earnings[9]. This is therefore a ceiling on estimated earnings, and any earnings that appear as this upper limit will need to be treated as greater than or equal to this amount.

---

[7] Some members of Nest are self-employed, and contribute some self-employed earnings into their Nest retirement accounts. However, earnings from self-employment does not appear in the Nest administrative data, as this only contains data supplied by employers using Nest.

[8] See note 3 and the PDF document Essentials of the UK retirement system for further details on banded earnings.

[9] In 2023-24 terms – see note 2.

7. Nest data panel structure and variables

The Nest data, as linked to the Understanding Society data, are derived from management information (MI) data supplied by Nest's scheme administrator, Tata Consultancy Services (TCS). The MI data are compiled into prespecified reports based on topic areas such as set up of employer accounts, enrolments of workers, payment of contributions and transactions within member accounts.

Many of the variables from the MI that are appended to the Understanding Society dataset are unprocessed input data, such as the age of the worker enrolled or the value of the member's contribution, but some are derived or manipulated by analysts at Nest before inclusion.

An important feature of the MI data is the ability to match and merge records across reports from different topic areas. Each employer registering with Nest, and every individual enrolled into Nest, is given a unique reference number, and these reference numbers are used to match and merge records. So, for instance, an individual is linked to all employers that have enrolled them in Nest at some point in time.

In the panel, we make use of reference numbers to attach data to either members, or to the unique combination of members and employers. In the interest of anonymisation, the Nest member reference has been replaced by the unique Understanding Society ID and the Nest employer reference number has been replaced by a simple randomised 9-digit numeric value.

7.1 Panel introduction and structure

The design is an unbalanced panel provided as rows of information, where a row represents a set of observations for a unique Member-Employer-Month (MEM). In other words, a row describes a particular member at a particular month, in relation to their employment with a particular employer.

The panel is **unbalanced** in that we include members that enter the panel (or exit) at any month. Balanced panels can later be created from the panel by selecting for members that meet given criteria, such as being a member for the entire calendar year 2019.

Table 2 describes the concept of variable levels. Note, there is sometimes repetition of values within the so-called "long panel" structure. An MM level variable could be repeated in the panel, if for example, that member has more than one employer. This is illustrated at the end of this section.

Table 2: Variable levels

Member-Month (MM) variables include MEMSCHSTAT, MEMAGE, FUNDCODE, POTVAL

Member-Employer (ME) variables include MEMEMPSTAT, MEMENROLDT

Member (M) variables include GENDER, MEMTITLE

Employer (E) variables include ANONEMPREF, INDUSTRY

Level	Abbreviation	Meaning
Member-Employer-Month	MEM	Values for that member vary for each month and vary by employment.
Member-Month	MM	Values for a member vary by month
Member-Employer	ME	Values for a member are fixed for each of their employments
Member	M	Values for a member are fixed
Employer	E	Values for an employer are fixed
Month	T	Variables indexing time (monthly increments)

7.2 Variable descriptions

7.2.1 Keys and references

Name	Description	Level	Data Type
ANONEMPREF	Anonymised employer reference number.	E	Nine character string
key_mm	Uniquely identifies a set of rows for a month and member. This set could be empty if the member has not joined Nest at that point in time. If the member has had more than one employer at Nest the set will contain more than two rows (the minimum will be two: at least one actual employer and one 'placeholder' employer, see section 7.3).	MM	String concatenating anon_id , and month_index (with "_" separators)

key_mem	Uniquely identifies a single row for a triplet of month, employer and member.	MEM	String concatenating anon_id , ANONEMPREF and month_index (with "_" separators)
month_index	Index time (month). Note, the panel starts with month 1 being January 2014 although no enrolment data is observed until July 2014 – month 7 due to the lack of reliable historic snapshot data until this time. The panel currently ends with month 108 being December 2022.	T	Numeric

7.2.2 Member variables

Name	Description	Level	Data Type
RETIREAGE	Age the member intends to retire at. [Note, this is defaulted to the member's state pension age (SPA) unless the member has made an active choice to change that, or if the member reaches their SPA and has not retired, in which case RETIREAGE is defaulted to 75. If they reach 75 and have still not taken their money then their RETIREAGE is defaulted to 104.]	MM	Numeric
MEMCOMM	States how the member wishes Nest to communicate with them.	MM	Factor w/ 2 levels "Electronic","Paper"
MEMSCHSTAT	A member's status with the Nest scheme.[10] All those members with a status other than "active" will not have a Nest investment pot. Active does not necessarily mean employed, still contributing or below retirement age.	MM	Factor w/ 6levels "","Active", "Inactive", "Retired", "Death in service", "RDM

[10] Blanks in MEMSCHSTAT are self-employed people who are active – their status appears in MEMEMPSTAT.

			(retirement decision made)”
MEMUPDATE	States if the member has agreed to be sent update information from Nest.	MM	Factor w/ 2 levels “Yes”, “No”
MEMACCDT	Date when the member’s account was created. This will match the date of the member’s first enrolment into Nest.	M	Date
MEMAGE	Current age of the member.	MM	Numeric
GENDER	Gender of the member.	M	Factor w/ 2 levels “M” “F”
MEMTITLE	Title of the member.	M	Factor w/ 8 levels “Lord” “Mr” “Dr” “Lady” “Miss” “Mrs” “Ms” “Mx”
MEMREG	States if the member has registered their online Nest account.	MM	Factor w/ 2 levels "Yes","No"

7.2.3 Employer variables

Name	**Description**	**Level**	**Data Type**
INDUSTRY	Industry of the employer.[11] These are related to SIC codes.	E	Factor w/ 29 levels "Administration and support services" etc
EMPDUTYDT	Date at which the employer was called to duties under automatic enrolment legislation.	E	Date
EMPSIZE	Self-declared size of the employer at the time the employer registered with Nest.	E	Factor w/ 7 levels “1-4” ,“5-49” ,“50-249”,“250-499” , “500-999”, “1000-4999” , “5000+”

[11] There are blanks in INDUSTRY as that field is no longer mandatory to complete.

7.2.4 Enrolment variables

Many individuals have been enrolled more than once at different, or the same, employers. However, in the panel we do not have granularity down to an enrolment event. Rather, the lowest unit of observation is at a member-employer-month (MEM) level. Consequently, variables shown below are updated into the MEM structure using information from **the most recent enrolment** at the same employer.

Name	Description	Level	Data Type
ENROLTYPE	Type of enrolment if SELFDUMMY = TRUE	MEM	Factor
ENROLTYPE_USED	Type of enrolment.	MEM	Factor The majority are automatic enrolments but other types include self-employed, opt-ins, voluntary, wwqe and others.
MEMEMPSTAT	Latest known status the member has with the associated enrolment at that employer.	MEM	Factor w/ 24 levels “Active”,“Leaver”, etc.
MEMENROLDT	Date at which the most recent enrolment at that employer was made.	MEM	Date
ENROLAGE	Member’s age at the time of relevant enrolment.	MEM	Numeric
OPTOUTSTDT	Start date of the opt-out window for the most recent enrolment at that employer.	MEM	Date
OPTOUTENDT	End date of the opt-out window for the most recent enrolment at that employer.	MEM	Date
TRELIG	States if the member is eligible to receive tax relief.	MEM	Factor w/ 2 levels “Yes”,”No

WWQE	States if the enrolment is of a member without qualifying earnings.	MEM	Factor w/ 2 levels "Yes", "No"

7.2.5 Contribution variables

Contributions that are not paid calendar-monthly have been adjusted to reflect a more consistent and workable monthly basis. For example, if the payment period runs from 28th March to 3rd April then four 7ths of the contribution value attributed to the month of March and three 7ths are attributed to the month of April. For this reason, the names of adjusted contribution variables end in ".ADJ"; those without this naming were already monthly.

Here is an example to illustrate the process. A January 2020 adjusted gross contribution was made from five weekly transactions:

29/12/2019 04/01/2020 123.29 (4/7 x 123.29=70.45)

05/01/2020 11/01/2020 199.96

12/01/2020 18/01/2020 233.78

19/01/2020 25/01/2020 145.84

26/01/2020 01/02/2020 159.37 (6/7 x 159.37 = 136.60)

This is entered in the panel as a single monthly value for GRS.ADJ of 786.63 with a NUMCONT of 4.43.

Name	**Description**	**Level**	**Data Type**
GRS_ADJ	Payroll gross contribution amount – this sum of the contribution, minus the charges plus the tax relief.	MEM	Numeric (£)
NUMCONT	Number of contributions made.	MEM	Numeric
PPP_ADJ	Pensionable pay that the contribution amount was paid upon. (not always present for legitimate reasons).	MEM	Numeric (£)
EMPC_ADJ	Employer contribution.	MEM	Numeric (£)

MEMC_ADJ	Member contribution.	MEM	Numeric (£)
TRA_ADJ	Tax relief component.	MEM	Numeric (£)
CONCH_ADJ	Charges levied on the contribution.	MEM	Numeric (£)
TRCH_ADJ	Charges levied on the tax relief.	MEM	Numeric (£)
CONTL_ADJ	Contribution less the charges.	MEM	Numeric (£)
GRSCONTAMT	Direct gross contribution amount – this sum of the contribution, minus the charges plus the tax relief.	MEM	Numeric (£)
NUMCONT_SELF	Number of direct contributions	MEM	Numeric
CONTTYPE_adhoc CONTTYPE_Permitted Party CONTYPE_Regular	Iindicator of the number of each type of direct contribution in that month.	MEM	Numeric
CASHVALUE	Cash value of the withdrawal	MEM	Numeric (£)
CONTRECAMT	Value of the direct contribution that's been received.	MEM	Numeric (£)
CONTCHGAMT	Value of the charges levied on the contribution.	MEM	Numeric (£)
TRAMT	Value of tax relief claimed against the contribution.	MEM	Numeric (£)

TRCHGAMT	Value of charges levied on the tax relief component of the contribution.[12]	MEM	Numeric (£)

7.2.6 Transfer and retirement variables

Name	Description	Level	Data Type
TOTDISVAL_U	Value of the UFPLS (Uncrystallised funds pension lump sums[13]).	MM	Numeric (£)
MEMAMTPAID_U	Value of money paid out to the member.	MM	Numeric (£)
TAXPAID_U	Monetary amount of tax paid.	MM	Numeric (£)
NUMUFPLS	Number of UFPLS withdrawals.	MM	Numeric
TOTDISVAL_O	Value of the annuity.	MM	Numeric (£)
TAXPAID_O	Tax paid.	MM	Numeric (£)
NUMOMO	Number of annuities taken.[14]	MM	Numeric

[12] Here an NA would indicate that it is not applicable and a zero would indicate there is a value but the rounded value is £0. A £0 entry here would not be surprising as this is the charge element (1.8%) applied to the tax relief of a contribution.
[13] An uncrystallised funds pension lump sum (UFPLS) is a one-off withdrawal of money from a pension account that leaves the account open for future contributions and investment growth. It is one of a range of ways that money can be taken out of a UK defined contribution pension.
[14] An annuity is a financial product that pays the saver a whole-life income across their retirement. This is one of the options available to UK pension savers after the age of 55. They can convert their Nest account value into an annuity to secure this income.

TOTDISVAL_TRIV	Value of trivial payment out.[15]	MM	Numeric (£)
TAXPAID_TRIV	Tax paid.	MM	Numeric (£)
NUMTRIV	Number of trivial payments.	MM	Numeric
RETIREVALBEFTAX_TRIV	Value of retirement pot before trivial commutation.	MM	Numeric (£)
MEMAMTPAID_TRIV	Value of money paid out to the member.	MM	Numeric (£)
TFRREQAMT	Monetary amount of the transfer requested.	MM	Numeric (£)
TFRAMT_IN	Monetary amount of the transfer in.	MM	Numeric (£)
NUMTRSIN	Number of transfers in.	MM	Numeric
TFRAMT_OUT	Value of transfer out.	MM	Numeric (£)
NUMTRSOUT	Number of transfers out.	MM	Numeric
RETIREVAL	Total OMO PAIDAMT.	MM	Numeric (£)

[15] A ‘trivial’ payment is another way to withdraw money from a defined contribution pension account. In contrast to an ‘UFPLS’ lump sum (see note 15), a trivial payment is always 100% of the account value, and is restricted to pension pots of no more than £30,000 value.

7.2.7 Pension pot variables

Name	Description	Level	Data Type
FUNDCODE	Notation for the fund the pot is invested in.	MM	Factor w/ 16 levels "LS_HGTH" , "TGT_FND" etc
FUNDSERCODE	For target date funds this target date year of the fund.	MM	Factor
FUNDPRICE	Price per unit of the fund.	MM	Numeric (£)
UNITMOVED	Number of units held.	MM	Numeric
MEMCURFUND	Member current fund.	MM	Factor
POTVAL	Value of the retirement pot (also found by FUNDPRICE x UNITMOVED).	MM	Numeric (£)
PREV_POTVAL	Pot value in the previous month.	MM	Numeric (£)

7.2.8 Convenience variables

We provide several variables that are convenient aggregations or indicators of core Nest data.

Name	Description	Level	Data Type
monthly_contrib	Total monthly gross contribution by the member across all employments and direct contributions.	MM	Numeric (£)
cum_monthly_contrib	Running cumulative total of Monthly.Contrib (starting at the earliest available month in the panel for the member).	MM	Numeric (£)
EMP_USED	ANONEMPREF for the earliest enrolment for this member.	MM	Nine character string
EMPSIZE_used	EMPSIZE for EMP_USED.	MM	Factor w/ 7 levels "1-4" ,"5-49" ,"50-249" ,"250-499" ,"500-999" ,"1000-4999" ,"5000+"

min_EMPDUTYDT	EMPDUTYDT of the earliest enrolment for this member.	MM	Date
min_month	Month that member joined the panel (Nest).	MM	Numeric
rule_cohort	Most relevant staging date.	MM	Factor w/ 91 levels "2015-6","2016-1" etc
tpr_cohort_used	TPR (The Pensions Regulator) validated staging date for small firms (<30 employees)	MM	Factor w/ 13 levels "2015-6","2016-1" etc
TPR_used	TRUE if TPR validated.	MM	Boolean
SELFDUMMY	TRUE for rows that relate to direct contributions, FALSE otherwise. See section 2.2.9.	MEM	
year	Year of the month_index.	T	Numeric
TABLEDATE	Calendar date of the month_index.	T	Date

7.3 'Placeholder' employers, the self-employed and direct contributions

Members may make direct contributions (for example, with a debit card). To allow for this and to preserve the structure of the panel, every member is therefore given a "placeholder" employer, in addition to their actual employer; identified in MEMEMPSTAT as SELF (n=106,934). (You can also use SELFDUMMY = TRUE to achieve the same filter). The reason for this is so that any direct contributions (contributions attributed to SELF) can be made without them being associated with the employer.

Note that the label SELF should not be confused with those members who have enrolled themselves into Nest as a self-employed member. For the self-employed members, direct contributions are the standard means of paying in. To identify members in the panel who are self-employed therefore – and not simply making direct contributions via the placeholder employer construct – look to ENROLTYPE = SELF EMPLOYED (n=457). To labour the point, direct contributions do not imply that the member is self-employed but all self-employed members will pay in via direct contribution.

8. Data access and Statistical Disclosure Control

Due to the sensitive nature of the data, the Secure Access data can only be accessed through the UK Data Service Secure Lab. Full details of the access requirements and the application process can be found on the [UK Data Service website](). It should be noted that access is restricted to researchers registered at a UK institution. In addition, before the data can be accessed, researchers must have attended a Safe Researcher training course. One consequence of this is that the time from applying for the linked data to having access to it can be longer than for data shared under licences. If you have queries regarding access to the data then please get in contact with the UK Data Service in the first instance.

9. Citation

If you use Understanding Society data you must cite every study that you use. The bibliographic reference for this study is as follows:

Nest Corporation, University of Essex, Institute for Social and Economic Research. (2023). Understanding Society: Linked Nest Auto-enrolment Pensions Dataset, 2014-2022: Secure Access [data collection]. UK Data Service. SN: 9127, DOI: http//doi.org/10.5255/UKDA-SN-9127-1.

All works which use or refer to these materials should acknowledge these sources by means of bibliographic citation. To ensure that such source attributions are captured for bibliographic indexes, citations must appear in footnotes or in the reference section of publications.

Citing this User Guide

When citing this User Guide, you can use the citation of this particular version quoted below. Note that where an online version is available on the Understanding Society website it is always the most up to date.

Institute for Social and Economic Research (2023), Understanding Society: Linked Nest Auto-enrolment Pensions Dataset, 2014-2022, User Guide, Version 1, July 2023, Colchester: University of Essex.

10. References

Benzeval, M., Burton, J., Crossley, T.F., Fisher, P., Gardiner, C., Jäckle, A & Moore, J.C. (2023) *Assessing bias prevention and bias adjustment in a sub-annual online panel survey.* Understanding Society Working Paper Series No. 2023 – n (forthcoming)

Hainmeuller, J. (2012) Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies. *Pol. Analysis* 20, 25−46.

Hainmeuller, J. & Xu, Y. (2013) ebalance: a Stata Package for Entropy Balancing. *J. Stat. Soft.* 54, 1-18.

Institute for Social and Economic Research (2022) *Understanding Society COVID-19 User Guide Version 11.0.* University of Essex.

Rothbaum, J., Eggleston, J., Bee, A., Klee, M. & Mendez-Smith, B (2020) *Addressing Nonresponse Bias in the American Community Survey During the Pandemic Using Administrative Data.* ACS21-RER-05 and SEHSD Working Paper 2021-24.

University of Essex & Institute for Social and Economic Research (2019). *Understanding Society: The UK Household Longitudinal Study, Waves 1-9 User Guide.* Institute for Social and Economic Research.

Valliant R. & Dever J. A. (2018) *Survey Weights: A Step-by-Step Guide to Calculation*. College Station, TX: Stata Press.