Student income and expenditure survey 2021 to 2022

Technical Report

November 2023

Authors: NatCen Social Research & Institute for Employment Studies

Contents

Please note Appendices 1 – 4 are provided as a separate document.

List of tables

3

Background to the study

This is a report on the research methods used in the Student Income and Expenditure Survey 2021 to 2022 (SIES 2021 to 2022) carried out on behalf of the Department for Education (DfE), the main source of data in this report.

SIES is a large-scale comprehensive survey that collects detailed information on the income, expenditure, and debt levels of higher education (HE) students.

The 2021 to 2022 survey is the latest in a series of surveys carried out at approximately three-year intervals. However, the current survey was conducted after a longer period due to the COVID-19 pandemic when it was decided to not proceed with the 2019/20 SIES. In previous waves of SIES (up to and including 2007 to 2008), the survey consisted of a 60-minute face-to-face interview, while from and including 2011 to 2012 a 30-minute web/ telephone survey was used. Each survey has included an expenditure diary which students, who are willing to do so, complete.

There have been changes to the student finance system since the 2014 to 2015 survey (see the Executive Summary). The current survey questionnaire and results reflect that.

Collaboration

The National Centre for Social Research (NatCen) and the Institute for Employment Studies (IES) conducted the 2021 to 2022 SIES in close collaboration. NatCen had overall responsibility for the delivery of the survey, lead responsibility for the sample design, questionnaire design, fieldwork with students and data preparation and IES had lead responsibility for collecting sample data from providers. The data analysis and report writing was shared between the two organisations.

Overview of the methodology

This SIES 2021 to 2022 technical report provides detailed descriptions of all aspects of the survey and data collection in England, including the development phases. Information relating to the overall methodology, such as sampling, fieldwork approach and weighting relate to students living in both England and Wales. However, detail on the profile of responding participants relates only to England domiciled students - a separate technical report will be published accompanying the Welsh report, which will detail the profile of Wales domiciled students. To give an overview of the research process, the key activities within the main-stage of the survey are outlined within the overall project timetable shown in Table 1.1.

Table 1: Project timeline

Timeline	Task
2019	Start of the 2019/20 survey contract
2019	Pilot for the aborted 2019/20 survey
Disruption by the COVID-19 pandemic	
October - November 2021	Start of the 2021 to 2022 contract Design of sampling plan Development of interview questionnaire and expenditure Diary (for pilot) Initial contact with providers
November 2021 – March 2022	Sample for pilot drawn Pilot of interview Recruitment of providers for main-stage Sample for main-stage drawn Interview questionnaire and expenditure diary finalised for main-stage
March – June 2022	Main-stage fieldwork: web and telephone interviews with students and diary completion

Source: NatCen/IES 2021 to 2022

Originally the current survey was planned for the 2019 to 20 academic year but was cancelled after the pilot due to the COVID-19 pandemic.

For the 2021 to 2022 survey, the main stage of fieldwork was carried out between March and June 2022, broadly corresponding to the Spring and Summer terms in the academic year. As the student spending diary was intended to capture term-time spending only, the diary was not available to students over the Easter holiday period of 8th April to 22nd April 2022.

Contacting students and fieldwork approach

Contacting providers

NatCen selected a number of providers in England and Wales based on Higher Education Statistics Agency (HESA)[1] figures about the student populations at each.

Letters were sent from DfE and Welsh Government (WG) to the Vice Chancellors and Principals at selected providers containing information about the research and an invitation to take part. IES made contact with individual providers, explained their role in sampling and secured their agreement to take part.

Contacting students

NatCen identified the numbers of full-time and part-time students to be sampled from each provider taking part (numbers differed by type and country of provider). The sample of students in their second year and above was drawn directly from HESA records. Once random selection of second year plus students was accomplished, providers were given a list of the HESA Unique Student Identifier codes (HUSID) of the sample of students in their second year of study or above and asked to provide the research team with contact details for each of these[2].

For students in their first year, HESA records were not available at the time when the sample was drawn. For these students, IES instructed providers about the numbers of students to sample and helped providers to do this using random selection. Providers then produced a list of sampled students containing names and addresses.

Fieldwork and data collection

Students selected for the study were sent an advance letter containing information about the study and invited to take part. The letter included the information needed to access the on-line questionnaire. No further input from providers to encourage participation was required. A few days later the selected students were contacted by email to remind them about the study and provide them with a web link to the questionnaire and the unique password that enabled them to access the on-line survey.

Students received two emails and one text message to remind them to complete the main questionnaire. If they did not fill out the questionnaire on-line (or only partially completed it), they were contacted by a telephone interviewer and given the opportunity

---

[1] Skills Funding Agency data were used to provide Further Education College (FEC)/Further Education Institution (FEI) information.
[2] The processing of SIES student data was compliant with Data Protection Legislation. Details about the legal basis for processing 2021 to 2022 student data can be found at these links: Collection notices (historical) | HESA How DfE shares personal data - GOV.UK (www.gov.uk)

to complete the survey over the phone. A further e-mail was sent to students who agreed to complete the questionnaire on the web following contact with a telephone interviewer to ensure that they had the information needed to access the survey on-line.

On completion of the main survey questionnaire, students were asked to fill out a seven-day on-line spending diary. They received an email and text message on the first day of completing the diary and on days 3, 4 and 7 of the seven-day diary to remind them to do so.

If the respondent completed the questionnaire with a telephone interviewer, they were sent an e-mail the same day, which included a link to the web diary, prior to any reminders being sent. Students completing the diary and the questionnaire were offered a £20 Love2Shop voucher for their time.

Sampling

Target numbers of students

The sampling was developed for English and Welsh students together so the following information focusses on both when discussing the sampling frame development and approach. The initial aim was to achieve the following numbers of interviews:

• 2,500 responses from English-domiciled full-time students (including those studying at Welsh HEPs);

• 1,050 responses from English-domiciled part-time students (including those studying at Welsh HEPs and the OU);

• 1,000 responses from Welsh-domiciled full-time students (including those studying at English HEPs);

• 500 responses from Welsh-domiciled part-time students (including those studying at English HEPs and the OU).

The following table shows more detail on the target numbers by type of student and provider.

Table 2: Target number by type of provider

	English HEPs	English FECs	Welsh HEPs	Welsh FEIs	OU	Total
English-domiciled full-time	1,730	470	300	0	0	2,500
English-domiciled part-time	642	250	25	0	133	1,050
Welsh-domiciled full-time	350	0	575	75	0	1,000
Welsh-domiciled part-time	58	0	350	25	67	500
Total	2,780	720	1,250	100	200	5,050

Source: NatCen/IES 2021 to 2022

Sampling providers

Higher Education Providers in England and Wales and Further Education

establishments (Further Education Colleges (FECs) in England and Further Education Institutions (FEIs) in Wales) as well as Open University (OU) students were included in the issued sample of students.

Providers in Wales

The small number of providers in Wales meant that no sampling of institutions was required.

Providers in England

Sampling of providers in England is described below. As with the 2014 to 2015 survey, a larger number of providers were sampled than was required in order to allow for refusals.

English HEPs

Of the 129 HEPs in England, 57 were selected for the survey with the expectation that 40 would agree to take part. The 57 were selected systematically with probability proportional to a weighted size (the calculation of which is described further below).

The stratifiers used were Region, tariff grouping and (weighted) size i.e. the 129 HEPs were sorted first by region, then, within region by tariff grouping, and within fee structure, by (weighted) size. A sample of 57 HEPs was then drawn systematically from this sorted list (that is, with a random start and then a fixed interval down the whole list) with probability proportional to their weighted size. The weighted size per HEP was calculated as a weighted sum of three figures (based on enrolment data for academic year 2019/20):

0.59 * (English dom F/T students) + 4.76 * (Welsh dom F/T students) + 6.79 * (P/T students)

The weights (0.59, 4.76, 6.79) were calculated as the ratio of the percentage of students in the group to be issued for the survey to the percentage of students in the group across the whole set of 129 HEPs. The weights per group were also used in the selection of students within providers, described below, with the result that all students were selected with equal probability as far as was possible. Table 3.2 gives the figures.

Table 3: Issued English HEP sample by domicile and full/part-time status

	English-domiciled full-time	Welsh-domiciled full-time	Part-time	Total
Issued for survey (approx.)	5,868	1,077	3,843	10,788
%	54.4%	10%	35.6%	100%
Numbers in all HEPs	103,7660	23,500	58,750	1,119,910
%	92.7%	2.1%	5.2%	100%
Weight	54.4/92.7=0.59	10/2.1=4.76	35.6/5.2=6.79	-

Source: NatCen/IES 2021 to 2022

The largest 13 HEPs (according to their weighted size) were all selected for the survey, the 13 being those with a weighted size larger than the sampling interval. A further 44 HEPs were selected from the remaining providers with probability proportional to their weighted size, thus providing a total sample of 57 HEPs.

English FECs

There were 173 English FECs with eligible students. Of these, the 42 containing less than 100 eligible students were excluded, leaving 131 FECs from which to sample. This is in line with previous SIES surveys and means that around five per cent of students were excluded. A total of 35 FECs were selected for the survey, with the expectation that 23 would agree to take part.

The selection followed a very similar procedure to that for English HEIs. The 35 FECs were selected from a stratified list with probability proportional to weighted size.

The stratifiers used were Region and weighted size i.e. the 131 FECs were sorted first by region, then, within region by weighted size. The sample of 35 was drawn from this sorted list systematically (that is, with a random start and then a fixed interval down the list).

The weighted size per FEC was calculated as follows (based on enrolment data for academic year 2019 to 2020):

0.89 * (Full-time students) + 1.36 * (Part-time students)

As with the HEP sample, the weights were then used in the selection of students within providers, described below, with the result that all students were selected with equal probability as far as was possible. Table 3.3 shows how the weights were derived.

Table 4: Issued English FEC sample by full/part-time status

	Full-time	**Part-time**
Issued for survey (approx.)	2,271	1,079
%	67.8%	32.2%
Numbers in 126 FECs	51,785	16,040
%	76.4%	23.6%
Weight	67.8/76.4=0.89	32.2/23.6=1.36

The largest three FECs (according to their weighted size) were all selected for the survey, the three being those with a weighted size larger than the sampling interval. A further 32 FECs were selected from the remaining providers with probability proportional to the weighted size, thus providing a total sample of 35 FECs.

Sampling students within providers

In all selected providers, first year students were sampled by the provider in accordance with instructions provided by NatCen/IES. Students in their 2nd year or above were sampled from HESA and (Individualised Learner Records (ILR) databases. Students who began their courses before September 2016 were dropped from the sampling frame, due to a change in funding structure. We now describe the sampling process for each provider type.

English HEPs: first year students

A sample of 110 first year students was requested from each English HEP. The figure of 110 was based on a calculation involving anticipated student response rates (from SIES 2014 to 2015) and the expectation that around 40 providers would take part out of the 57 selected. It also included 20 cases as contingency/reserve if response rates were lower than anticipated.

The ratio of English-domiciled full-time, Welsh-domiciled full-time and all part-time students was calculated by dividing the total number into three in proportion to the weighted totals used in the sampling of institutions. For example, the number of first year full-time English-domiciled students requested was calculated as follows:

{ 110 * 0.59 * (English dom F/T students) }

/{ 0.59 * (English dom F/T students) + 4.76 * (Welsh dom F/T students) + 6.79 * (P/T students) }

In theory this method should produce an equal probability sample. In practice, however, the actual number of students in each provider varied from the historical figures used and therefore the actual probabilities of selection varied to some degree.

English HEPs: second/third year students

For 2nd year and above students, 220 students were sampled from each institution. The ratio of English-domiciled full-time, Welsh-domiciled full-time and all part-time students was calculated in the same manner as for first year students. Whilst this should have produced an equal probability sample, as with first year students, the selection probabilities varied when the actual number of students deviated from the historical numbers.

The sample was selected from HESA records. Within each provider these were stratified by Student type (English-domiciled full-time, Welsh-domiciled full-time, Part-time), Year, Level of study, Sex, Course Aim and finally by Date of birth. A systematic sample using a random start and fixed interval was taken within each student type in each institution.

English FECs: first year students

A sample of 60 first year students was requested from each English FEC. This figure was based on a calculation involving anticipated student response rates and the expectation that around 23 providers would take part out of the 35 selected.

A similar calculation to that used in English HEPs was used to divide the totals in proportion to the weighted totals for Full-time and Part-time students. For example, the number of first year Full-time English-domiciled students to request was calculated as follows:

{ 60 * 0.89 * (Full-time students) }

/

{ 0.89 * (Full-time students) + 1.36 * (Part-time students) }

English FECs: second/third year students

For 2nd year and above students, 120 were sampled from each FEC. The ratio of English-domiciled full-time, Welsh-domiciled full-time and all part-time students was calculated in the same manner as for first year students.

The sample was selected from ILR records. Within each provider these were stratified by Student type (Full-time/Part-time), Year of study, Domicile status, Sex, and finally by Date of birth. A systematic sample using a random start and fixed interval was taken within each student type in each institution.

Welsh HEPs/FEIs: first year students

As we had taken a census of providers, a constant sampling fraction was applied across providers to each of the three types of student (English-domiciled full-time, Welsh-domiciled full-time and part-time). Therefore, the numbers of first year students requested from Welsh HEIs varied by provider (and by type).

The same principle applied to 2nd year and above students in sampling from HESA records.

OU students

The OU provided us with a sample of 800 English domiciled and 400 Welsh domiciled students taken across all years of study.

Questionnaire and diary of spending

All students – whether at a higher education provider (HEP), further education establishments (FEC/FEI) or the Open University – were surveyed using the same methods. First, they were asked to complete a main survey, which was conducted either on-line (as a self-completion questionnaire; see Appendix 1, in separate document) or over the phone with an interviewer. All students were then asked to complete a seven-day diary of spending after the interview. The spending diary was only available on-line; (see Appendix 2, in separate document).

The combination of the main questionnaire and the seven-day diary of spending meant that all areas of income and spending could be monitored. For example, the questionnaire was able to pick up on larger and more memorable spending such as rent, travel, childcare and holidays whilst day-to-day spending on items such as food and entertainment was recorded in the diary of spending.

Initial development

The 2014 to 2015 questionnaire was taken as the starting point for the 2021 to 2022 survey. Changes were made to the questionnaire content to reflect changes to tuition fees and student loans since 2015, and improvements were made to the content following feedback on the 2014 to 2015 survey. Otherwise, the content of the questionnaire was largely similar to that employed in 2014 to 2015.

Pilot

SIES was piloted to check the wording and routing of new and amended questions, pilot fieldwork was conducted between February 2022 and March 2022.

The sample was provided to us directly by four universities: Durham University, University of Cambridge, University of Sussex, and the University of Hull. There were a total of 522 cases in the pilot sample for the questionnaire and six people piloted the spending diary.

Participants were offered a £20 Love2Shop voucher for completing the pilot study.

E-mail and text reminders were sent to encourage responses and more questionnaires were completed following the reminders, but the level of participation was low following each contact. In total, 90 students fully completed the questionnaire with the median time taken to complete being 40 minutes. The analysis of the pilot data did not suggest any major issues with the routing nor question content.

Fieldwork

Briefing and interviewer numbers

A team of telephone interviewers were briefed to work on SIES 2021 to 2022. The briefings covered the background to the survey, the sample of respondents, use of the study documents (for example, the advance letters and reminder letters and emails), approaching the sample, an overview of the questionnaire content and the use of the seven-day diary of spending. The interviewers were informed of the content of the diary to enable them to explain this part of the survey to respondents although the diary would be completed by the student on-line and not administered by the interviewers.

Contact procedures

Students were informed about the study via an advance letter. The letter introduced the study, emphasised its importance and provided respondents with a unique access code to log on to the study website and complete the survey.

An email was then sent out to students providing them with information about the study, the unique access code and a web link to the survey. This email was followed with a text message (where we had students' mobile numbers) to alert potential respondents to the fact they had been sent an email about the study – and encouraging them to complete it.

Telephone contacts began soon after sending the email. Interviewers telephoned respondents who had not yet completed the main survey on-line and attempted to arrange an appointment either to complete the survey over the phone, or a reminder call to do the survey on-line if they preferred.

At the end of the main survey students were asked whether they would be willing to complete the seven-day on-line spending diary. Students were directed automatically to the diary where they agreed to complete it. Reminder e-mails and text messages were sent at intervals during the seven-day diary period to encourage recording of spending for all seven days.

Incentives

Respondents received a £20 Love2Shop voucher for completing all seven days of the diary alongside the questionnaire. They did not receive an incentive for just completing the main questionnaire, or for only partially completing the diary.

Fieldwork period and monitoring

The mainstage of fieldwork was carried out between 20th of April and 22nd June 2022, broadly corresponding to the Summer term in the academic year.

Response

Survey response

Table 4.1 shows the final completion rates for the main questionnaire for the full sample, including both English and Welsh domiciled students. Overall, 22 per cent of the issued sample of students completed the questionnaire. A further seven per cent of students were found to be ineligible when the interviewer contacted them or when they entered their details at the start of the questionnaire (for example, they had dropped out of their course). Table 4.2 shows the completion rates of the questionnaire for English domiciled students.

Table 5: Full sample main questionnaire completion rates, for English and Welsh domiciled students

	N	%
Issued	28,071	100
Ineligible	1,896	6.8
Productive	6,148	21.9
Partial productive	148	0.5
Refusals	995	3.5
No contact/ other non-response	18,884	67.3

Source: NatCen/IES 2021 to 2022

Table 6: English HEP and FEC/FEI questionnaire completion rates (excluding OU)

	Completed	Incomplete
English HEP	3,003	9,870
English FEC/FEI	785	3,103
Full-time	2,898	8,646
Part-time	876	4,195
Male	1,283	5,288
Female	2,182	6,583

Base: Students attending English HEP or FEC/FEI sampled for SIES 2021 to 2022
Source: NatCen/IES 2021 to 2022

Diary response

Table 4.3 shows the level of diary returns for English domiciled students. The seven-day spending diary was completed by 42.6 per cent of English-domiciled students who completed the questionnaire.

Table 7: English domiciled diary completion rates (excluding OU)

	Completed	Incomplete
All	1,553	2,158
English HEP	1,084	1,483
FEC	317	464
Full-time	1,226	1,641
Part-time	327	517
Male	477	935
Female	1,076	1,217

Base: Students sampled for SIES 2021 to 2022   Source: NatCen/IES 2021 to 2022

Fieldwork and quality control procedures

As with all surveys carried out at NatCen, a programme of live monitoring of the work of our telephone interviewers was carried out on SIES. For each project, five per cent of productive interviews are monitored and the aim is to monitor all interviewers working on the project at least once.

Data checking, coding and editing

Data checking

Checks in the questionnaire program helped to limit the number of data discrepancies. Within the program, each numeric answer was given a set range of possible answers. This allowed only potentially valid answers. For example, if the maximum amount of Adult Dependants' Grant received by a student is £3,190, this would be the upper limit of the range within a question asking about this. Given that spending may legitimately be very high or very low in a given week, it was not feasible to set validation checks on individual entries in the spending diary. Interim data were inspected by researchers from NatCen.

Coding and editing of data

A data processing team carried out the coding and editing of questionnaires. Data coding was necessary to enable the analysis of information collected by verbatim answers.

Factsheets were used to assist coding and editing of the data. These provided a summary of a productive interview and alerted editors to possible errors or inconsistencies that needed to be dealt with. For guidance on how coding was implemented please see Appendix 3 (in separate document).

Code frames used in editing were developed by the researchers based in part on those used for the 2014 to 2015 study. Where no previous list existed, researchers inspected question responses from the first completed interviews. Any complex editing decision was referred to the researchers for adjudication. These cases were documented, and instructions relayed to the data processing team.

Diary of spending

A data processing team also carried out the coding and editing of the seven-day diaries of spending.

One major purpose of the diary editing was to allocate a code to any spending that students had entered as 'other' spending. Similar rules that were developed for the 2014 to 2015 study were adopted here. For each day filled in on the diary there were four 'other' items which may have required editing. The aim of these 'other' codes was to allow students to write in spending that did not fit into the existing categories in the diary. The aim of editing the diary was to code these 'other' categories into: a) existing diary categories; b) a new 'other' category not asked about in the questionnaire; or c) a new other category already asked about in the questionnaire. (Option c was particularly important, as we did not want to double-count items included in the diary that had already been covered in the interview.)

Summary measures of income, expenditure, debt and savings

Within the summary report the majority of monetary figures refer to the total amounts of money spent, received or owed over the whole academic year. However in the questionnaire and diary, these monetary amounts may have been recorded referring to a week, a month, a term or over the whole calendar year in order that students could give as accurate figures as possible. It was therefore necessary to create summary derived variables which totalled the amount of money spent or received over the full academic year, assuming that answers given in the questionnaire or diary represented average weeks. Details of how specific derived variables were derived can be found in Appendix 4 (see separate document). The derived variables relating to the day-to-day spending recorded in the diary follow the same principles. Again it was assumed that spending within the recorded week was an average week. The weekly amount spent on different types of item was therefore multiplied by 39 to give the spending for the academic year.

Dataset

Dataset

Interview and diary data were merged together to form one complete dataset. This dataset also included all the derived variables for the interview and diary instruments. A list of key variables, including break variables, can be found in Appendix 5.

Extreme values

Once the summary measures of income, spending, borrowing, and savings were created and tested, they were reviewed by the research team. This allowed them to correct any unfeasible answers (e.g. amounts being recorded as annual amounts when they were obviously intended to be term time amounts or vice versa) and also trim any outliers that would skew the analyses if left untreated. Trimming involved identifying outliers and then trimming these outliers to the highest amount within the accepted range.

Details about the variables that have been trimmed can be found in Appendix 6.

Missing values and imputation

Missing values occur when a respondent provides no answer, or when they opt 'refuse to answer' or 'don't know'. A different approach to dealing with missing values has been used for the income section than for the expenditure and savings and debt sections.

The SIES questionnaire includes a number of question 'sets' which build to provide a figure for each element of student income. For example, in most cases students were asked whether or not they received a particular source of income (such as student maintenance loan), how frequently they received this income, and the regular amount received. The answers to these questions were then used to calculate the total amount received for that particular source of income (this is a derived variable).

Missing values could occur in any one of the questions that make up the set and would lead to a missing value for the derived variable and any other derived variables higher up the scale.

Left untreated missing values would have meant that a large amount of useful data would have been lost, as in many cases respondents might have only answered 'don't know' or 'refuse to answer' to one question lower down in the hierarchy. It was decided therefore to give missing values/data an imputed value in order to retain all the cases for analysis, and to make full use of the data that students did provide. Imputed values were either a zero value or a median recipient value (based on the median value of a similar group of recipients). It should be noted that for any one derived variable the number of imputed values was relatively small. However, over 40 derived variables, each made up of several

items, the cumulative impact of missing values was significant enough to warrant such an approach as noted above.

Zero values were used when there was insufficient additional data to be able to assume a non-zero value (either from the respondents' other answers to the questions in that 'set' or from the answers to that specific question/variable from other similar respondents). Nonzero values were used when there was sufficient additional data to be able to estimate a likely response value.

For example, in the section on student loan for maintenance, full-time students were asked: a) whether or not they received a loan; b) how much they received; c) whether their answer was per term or per year; d) whether they got additional weeks; e) how many additional weeks; and f) how much per week. A missing value could occur at any one of these six questions making it impossible to calculate the total received for student loan for maintenance. Therefore, to make use of the data the respondent provided in response to the other questions in the set it is necessary to impute a value for the missing item. If they answered 'don't know' or 'refuse to answer' for 'a)' or 'd)' above a value of zero was imputed for the respective part of the total for their maintenance loan (as it was assumed that if they received a maintenance loan they would know about it and be able to answer the question). On the other hand if they answered 'don't know' or 'refuse to answer' for 'b)', 'e)' or 'f)' in the example above a value for the missing item was imputed based on the median value for the same group of students (e.g. full-time English students in receipt of a loan who gave a termly/annual answer). In the case of a missing value for 'c)' in the example above it was normally possible to work out whether the respondent had given a termly or annual amount based on their answer to 'b)' and so this would be corrected accordingly.[3]

This approach follows that of the previous wave of SIES. It ensures that a consistent base is used throughout the analysis of income, and has the added benefit that the mean values of each element of student income sum to the mean value of the total student income and that it is possible to estimate the proportion of income among students coming from each source. For the SIES analysis dealing with expenditure, missing values were treated as missing for the analysis (i.e. excluded from each relevant statistical calculation), and different bases were used depending on the most appropriate sample to use (e.g. diary responses or survey responses). This approach was driven by

---

[3] In the case of income from paid work a slightly different approach was used for treatment of missing values. In the relatively few cases where students answered 'don't know' or 'refuse to answer' how much they earned it was assumed that the income from the job was not significant (otherwise they would be able to say roughly how much they were paid) and so they were assigned an income of zero for this part of their work income. In most of these cases respondents answered how much they were paid during term-time and not how much they earned during vacations or vice-versa, in which case any pay reported was assigned to the appropriate job. In a few cases students told us how much they earned from a given job but answered 'don't know' or 'refuse to answer' when the job started. In these latter cases a median start date for similar students (e.g. part-time Welsh students with job 1) was imputed in order to make use of the answers given. Again this only applied to a relatively small number of respondents.

the relatively small size of the cumulative missing data (less than ten per cent across the entire section) and the different data sources available (diary and main-stage questionnaire).

Similarly, with the analysis relating to savings and debt, missing values were also treated as missing for the analysis, and therefore different bases were used for each derived variable. Where there were overlaps between income variables and debt variables, the cleaned and imputed variables from the income section were used to ensure consistency. The overall approach was driven by the relatively small size of cumulative missing data in income section (less than ten per cent across the income section), the difficultly in making a 'best guess' for missing data, and the case that there were relatively few contributory variables for each derived variable (i.e. few questions in each 'set').

Imputation of fee loan amount

Due to an issue in the design of the questionnaire, it was possible for respondents to the online survey to submit a response of "*Don't know*" when asked how much tuition fee loan they had used or expected to use. This resulted in missing values for fee loan amounts for 16 per cent of English full-time students and for 14 per cent of English part-time students.

To enable total income values to be derived for these students, it was necessary to impute amounts for their fee loan. Some students had given an amount for their course fees, and in these instances the fee loan amount was set at the amount of their course fees (for the vast majority of respondents who gave amounts for their course fees and fee loan, the two were the same). However, for students who did not give an amount for their course fees, a value for course fees and fee loan had to be assigned, on the basis of their course characteristics.

Table 8: Imputation of fee loan amount

	Full-time		Part-time	
	Number	Percentage	Number	Percentage
Not imputed - gave fee loan amount	2,414	84.2	965	85.5
Imputed from course fee	250	8.7	43	3.8
Assigned value	203	7.1	120	10.6
Total	2,867	-	1,128	-

The amounts assigned were based on the median values by type of provider and whether or not the student was on an accelerated degree for full-time students, and by type of provider and intensity of study for part-time students, as follows:

- English full-time students
    - Students at English HEPs on normal or accelerated degrees - £9,250
    - Students at FECs on normal degrees - £7,650
    - Students at FECs on accelerated degrees - £7,450
    - Students at Welsh HEPs - £9,000

- English part-time students
    - Students at HEPs on a high intensity course - £5,000
    - Students at HEPs on a medium or low intensity course - £4,500
    - Students at FECs - £3,714.

Treatment of student loan

During the analysis of the 2014 to 2015 data it was uncovered that the estimates for student borrowing and debt were lower than could have been expected given the levels of loan that were available. Investigation of the component variables revealed that the value of student loans from previous years was underestimated. This was the result of a substantial proportion of second and third year students reporting that they did not have outstanding student loan debt from previous years when in fact they probably had. Receiving student loan payments one year but not the other is uncommon given the way student loans are paid.

It was therefore decided to only use the existing data and restrict the analysis of second and third year full-time students to those who had reported taking out a student loan before. This is likely to lead to slight overestimation in the mean value for this group, as students who genuinely did not have an outstanding loan from previous years were also excluded. All part-time students and full-time students in their first year are included as their results were unlikely to be substantially affected by this error (it is possible that they too had outstanding student loan debt from previous full-time courses that they did not report, but proportion of students affected by this is likely to be small). While less pronounced in the 2021 to 2022 data, this issue is still present. As such the same approach has been taken for 2021 to 2022.

Construction of strata variable

For 2021 to 2022 data, the strata variable was created to reflect the stratification used in sampling providers. As such it is appropriate for use in analysis, to account for the complex survey design (along with the PSU and weight variables). This approach is consistent with the strata variable used in 2014 to 2015. The change does not affect the survey estimates, only the standard errors, and due to the relatively small effect of the

stratification for most outcomes, it is unlikely to have had any significant impact on the comparisons between the two surveys.

Weighting

Summary of approach

As in previous years, the main SIES weights were calculated in three stages. First the probability of selection was calculated for each respondent based on the selection probability for the provider and the individual student within the institution. Second, response to the survey was modelled using logistic regression and a non-response weight was calculated as the inverse of the estimated probability of responding; this was then multiplied by the selection weight to create a final non-response weight. Extreme weights were checked at each stage and trimmed where necessary to reduce their impact. Finally, the weights were adjusted so that the characteristics of the weighted sample matched that of the student population (as recorded by HESA) in terms of age and sex. For SIES 2021 to 2022, an additional set of weights were also created for respondents who completed the diary.

As in previous years the overall aim was to generate a weighted sample that matched the population of students as closely as possible, whilst at the same time generating weights that were not so variable that the standard errors of survey estimates were unnecessarily inflated. The diary weights were created to facilitate analysis of the diary data.

Selection weighting

The first stage of weighting accounted for differential probabilities of selection as determined by the type of provider and student. Details of how providers and students were selected can be found in Sampling (section 3).

All selection probabilities were calculated as the product of the probability of selection of the provider and the probability of selection of the student within each institution. In each case the selection weights were the inverse of this overall probability of selection.

As noted in the section 3 ,Sampling, the data used in sampling providers (enrolment data for academic year 2019 to 2020) were not consistent with the HESA/ILR/LLWR data (individual records for academic year 2020to 2021) from which 2nd year and above students were sampled, therefore the selection probabilities varied where student totals did not match.

HEPs

For English HEPs, selection probabilities were taken directly from the sampling file used in selecting these providers. Selection probabilities for 2nd year and above students were also taken directly from sample files, in this case from the files (based on HESA) that were used for sampling students. For first year students the same approach was used as in previous years. In addition to the details of selected students, co-operating providers were asked to provide the number of students that were selected per group and the size

of the population groups from which the samples were selected. Selection probabilities were then calculated, for each category of student (English-domiciled full-time), as the ratio of these two figures:

Number of students selected / Total number of students in group

If the denominator (the total number of students in the group) was missing, this was estimated using 2020 to 2021 HESA data. For HEPs that did provide population figures, the correlation with the previous years' HESA figures was calculated and was found to be reasonably high. This meant that the computed selection probabilities were not highly variable, which is consistent with the previous SIES.

FECs

For English FECs, selection probabilities were calculated in the same manner as described above (for English HEPs) within student category (Full-time; Part-time), year and institution. The denominators for first year students were estimated from the ILR if not provided by the institution. For FECs the correlation between the ILR counts and the population figures provided by the providers was very slightly lower than was the case for HEPs, but still reasonable.

Open University

Finally, for Open University (OU) students, selection probabilities were calculated directly from the information on total English domiciled students provided by the OU.

Trimming extreme selection weights

The distribution of selection weights was examined within each provider and student type where sampling took place. Unlike previous years of SIES, there were no big outliers so trimming of extreme selection weights was not required.

Non-response weighting

Final outcome codes for all issued cases were used to categorise each student as a respondent, non-respondent or ineligible. After removing the ineligible respondents, response to the survey was modelled using logistic regression. The following variables were used as predictors:

• Student type (nine categories, see below)

• Age group

• Sex

• Year

• Living in London (based on address)

The model was weighted by the selection weights and all variables were found to be pre-dictive of response. The interaction between age and sex was tested and found to be significant so age group and sex were combined into one variable (age group by sex) in the final model.

A non-response weight was calculated as the inverse of the predicted probability of response. This weight was trimmed at the 1st and 99th percentiles then multiplied by the selection weight to create a pre-calibration weight. Two outliers were trimmed from the pre-calibration weight.

Calibration weighting

The pre-calibration weights were then calibrated to the following population estimates taken from HESA and ILR databases.

Table 9: Student type by age group

	Under 25	25+
English HEPs: English-domiciled full-time	912,315	167,125
Welsh HEPs: English domiciled full-time	30,200	4,560
English HEPs: All part-time	18,325	33,880
FECs: All full-time	26,372	19,358
FECs: All part-time	5,783	8,609
Open University	24,235	65,195

Table 10: Student type by sex

	Male	Female
English HEPs: English-domiciled full-time	458,455	620,985
Welsh HEPs: English domiciled full-time	17,280	17,480
English HEPs: All part-time	23,410	28,795
FECs: All full-time	18,141	27,589
FECs: All part-time	8,178	6,214
Open University	31,135	58,295

The calibrated weights were scaled to have a mean of one and checked for outliers. No further trimming was required.

Diary weighting

An additional set of weights was created for the respondents that completed a diary (n = 2,553). A model predicting the probability of diary completion was fitted for the responding cases, using stepwise logistic regression. Models were run
weighted by the main SIES weights. The following variables were found to be predictive of diary completion and included in the final model:

- Student type (nine categories, see below)
- Age group
- Sex
- Current year of course studied
- Total length of course or programme studied
- Living in London (based on address)
- Ethnicity
- Qualification aim
- Whether student funding and support available affected decisions about study
- Any outstanding loans from the Students Loans Company
- Any earnings from paid work since start of the academic year
- Paid work over last summer vacation
- Receipt of benefits
- Type of school attended between ages of 11 and 16
- Highest qualification before studying at university
- Whether student had ever been in paid work prior to course

A non-response weight was calculated as the inverse of the predicted probability of diary completion. This weight was trimmed at the 1st and 99th percentiles then multiplied by the main SIES weight. Three outliers were trimmed from this diary weight and it was rescaled to have a mean of 1.

Conducting analysis

In conducting analysis of SIES data for the summary report, all percentages, medians and means quoted are based on weighted data. The unweighted number of cases on which figures are based is also included.

Thirty was taken as the minimum subgroup size for which percentages and means could be quoted. Figures based on subgroup sizes of less than 50 but equal or more than 30 are shown in brackets.

All cross-tabulation, description of medians and significance testing for the summary report was carried out using SPSS. Commands from the Complex Samples procedures were used, to ensure the correct treatment of complex sampling weights in these procedures.

Making comparisons

The methodology for the 2014 to 2015 and 2021 to 2022 surveys was, as far as possible, the same so comparisons between the two are appropriate. However, there are a number of caveats that should be borne in mind when making such comparisons.

Up-rating 2014 to 2015 results for inflation

To allow for comparisons between 2014 to 2015 and 2021 to 2022 figures, all 2014 to 2015 monetary values have been adjusted for inflation. All monetary values (with the exception of values relating solely to income from paid work or tuition fees) have been up-rated by 1.297, reflecting the changes in the Retail Price Index (RPI) between April 2015 and April 2022 (the start of the relative fieldwork periods).

- Values relating solely to income from paid work were up-rated by the Average Weekly Earnings (AWE) Total Pay index of 1.255.

- Values relating solely to fees (Tuition Fee Loan income and tuition fees paid) were up-rated by an index of 1.028, since fees increased by only 2.8% between the two surveys, and were not subject to inflation.

Where derived expenditure variables included mid-level variables requiring different adjustment indices, adjustments were made to each mid-level variable. For example, participation costs (XPartic) is comprised of the mid-level variables direct course costs (XBooks), facilitation costs (XParTrav), and tuition fee costs (XFeeCon). 2014 to 2015 participation costs were therefore up-rated by adjusting direct course costs and facilitation costs by 1.297 (reflecting changes in RPI), while tuition fee costs were adjusted by 1.028. The adjusted mid-level variables were then summed to construct adjusted total participation costs. The same approach was taken to construct adjusted total spending.

Weighted up-rating indices were applied to 2014 to 2015 income data to account for the 2.8% increase in tuition fees between the surveys, and the Average Weekly Earnings Total Pay index of 1.255 for paid work. The following indices were therefore applied:

For analysis including Tuition Fee Loan income, for full-time students:

- Full-time main income sources up-rated by an index of 1.1356

- Full-time Tuition Fee Loan income only up-rated by an index of 1.028

- Full-time work up-rated by an index of 1.255

- Full-time all other income sources up-rated by an index of 1.297

- Full-time total income up-rated by an index of 1.1894

For analysis including Tuition Fee Loan income, for part-time students:

- Part-time main sources up-rated by an index of 1.0291

- Part-time Tuition Fee Loan income only up-rated by an index of 1.028

- Part-time work up-rated by an index of 1.255

- Part-time all other sources up-rated by an index of 1.297

- Part-time total income up-rated by an index of 1.2567

Composition of the full-time sample

The key personal characteristics of the weighted responding sample are summarised below and compared with those of the 2014 to 2015 sample (Table 1.2). The profile of the 2021 to 2022 full-time sample was similar in many respects to that of the 2014 to 2015 sample. Where there were differences, these tended to be in terms of HE study characteristics rather than student characteristics.

For **HE study characteristics**, the sample included fewer first year students (which reflects the broader sampling approach in 2021 to 2022, whereas in 2014 to 2015 students were sampled across three years only, see below), and fewer students studying Human and Social Sciences, Business and Law, or Creative Arts, Languages, and Humanities, compared to 2014 to 2015. There was no change in the proportion of the sample studying for a PGCE/ITT qualification, though the percentage of those studying for a Bachelor's degree (rather than an 'other undergraduate' qualification) increased from 82% of the sample in 2014 to 2015 to 91% in 2021 to 2022. This in part reflects real changes in the student population, with fewer students studying for other undergraduate degrees, but also reflects a change in our approach, with those studying for an integrated Master's now categorised as studying for Bachelor's degrees rather than 'other undergraduate' degrees, as they were in 2014 to 2015.

In terms of **student characteristics**, the largest changes were in the ethnic composition of the sample and in the number of students identifying as having a disability. In the 2021 to 2022 sample, there were fewer white students (by 13 percentage points), and more students from mixed or other ethnic backgrounds (by 9 percentage points). The proportion of full-time students with a disability increased by 16 percentage points, rising from 21% of the sample in 2014 to 2015 to 38% in 2021 to 2022. This increase is primarily driven by a large increase in the number of students reporting mental health conditions, and this reflects changes in the wider student population (described above). Other slight differences were that in 2021 to 2022 there were fewer respondents studying in further education colleges and more studying in English HEIs than found in the 2014 to 2015 survey (again reflecting changes in the wider student population, see above).

Composition of the part-time sample

The 2021 to 2022 part-time sample was different from the 2014 to 2015 sample in a number of respects.

HE study characteristics: As identified for the full-time sample, the 2021 to 2022 part-time sample had a smaller proportion of students in their first year, studying for 'other undergraduate' degrees, and studying in FECs. Compared to the 2014 to 2015 sample, more part-time students were studying between 25 and 29% of a full-time equivalent (FTE) course.

Student characteristics: The 2021 to 2022 part-time sample saw a number of changes not reflected in the full-time sample. The 2021 to 2022 part-time sample had a lower proportion of men and students aged 25 or older and did not see any large changes in the ethnic composition of the sample. While the proportion of students in each socio-economic group remained broadly similar between 2014 to 2015 and 2021 to 2022 for full-time students, there were fewer part-time students in the intermediate and routine and manual occupational groups (by 8 and 5 percentage points respectively), while the proportion in the 'no paid work / unable to code' category increased by 13 percentage points.

These changes need to be taken into account when interpreting some of the changes between 2021 to 2022 results and those from previous surveys.

Those on longer study programmes

When making comparisons for income and expenditure between the 2021 to 2022 and 2014 to 2015 surveys, it should also be noted that the samples for the 2021 to 2022 and 2014 to 2015 SIES were slightly different in how they treated students who were in year 4 or above of their studies. The latest survey included all students eligible for student funding regardless of their year of study, that is it includes students who might be in their 4th or subsequent year of a course. Students commencing their studies during or after

the 2016/17 academic year were eligible to take part so the sample will include students up to and including those in their 6th year of study.

In contrast, those in year 4 or above of their studies were excluded from the 2014 to 2015 study and the 2014 to 2015 sample included students who had started their studies during or after the 2012/13 academic year only. This restriction on the sample in 2014 to 2015 was put in place because significant changes in funding were introduced in the 2012/13 academic year, so funding for those in previous years would have reflected the previous funding regime.

Use of medians

In 2014 to 2015 and previous waves of SIES, all figures reported have typically been means, with medians presented in tables to give an indication of the shape of the skew in each of the underlying income / expenditure distributions. This is in part as means can be aggregated, and also to be consistent with the Family Expenditure Survey (FES).

In 2021 to 2022 we instead report median averages, as these are less likely to be skewed by outliers and so give a more reliable indication of the 'typical' value. Means are still presented in the tables.

All reported medians have been calculated using the 'CTABLES' function in SPSS. Different SPSS functions may produce slightly different median amounts. This is due to different functions employing different default methods for calculating percentiles. CTABLES uses the 'aempirical', whereas the frequencies command uses Waverage (a weighted average).

Tuition fees

In 2021 to 2022 expenditure is presented minus tuition fees, and income is presented excluding tuition fee loans, whereas these has been included previously. This is due to the fact that tuition fee loans are typically paid directly to the provider to cover fees so they do not constitute a meaningful element of students' income or expenditure. The accompanying tables include both with and without tuition fees, and it is indicated in the summary report text where these have been included or excluded.

Despite these differences, the report conducts analysis using the full 2021 to 2022 sample rather than restricting attention to students in years 1 to 3 of their studies only. This ensured sample sizes large enough for sub-analysis and enabled comparison of the full-time and part-time samples to be made on the same basis. However, work was undertaken to explore the likely influence of the differing samples, and this was found to be minimal.[4]

---

[4] The proportion of students who were in year 4 or above in 2021 to 2022 was 8% of full-time students domiciled in England and 14% of part-time students.

Appendix 5: Variable List

Variable Name	Variable Label	Source
XIncome	DV:: Income, including fee loan	DVs_Income
XIncXFL	DV:: Total income, minus fee loan	DVs_Income
XFamily	DV:: Family & friends, including fee loan	DVs_Income
XFamXFL	DV:: Family & friends, minus fee loans	DVs_Income
XParents	DV:: Contributions from parents and other relatives	DVs_Income
XShare	DV:: Share of partner's income, including fee loan	DVs_Income
XShareXFL	DV:: Share of partner's income, minus fee loans	DVs_Income
XPartner	DV:: Gifts of money from partner	DVs_Income
XMStud	DV:: Main sources of student support	DVs_Income
XMSXFL	DV:: Main sources minus fee loans	DVs_Income
XFLoan	DV:: Student loan for fees	DVs_Income
XFeeGrw	DV:: Welsh Government Fee Grant	DVs_Income
XMLoan	DV:: Student loan for maintenance	DVs_Income
XMntG	DV:: Maintenance grant	DVs_Income
XAccess	DV:: Access to Learning Funds/ Financial Contingency Funds	DVs_Income
XCouGr	DV:: Course Grant	DVs_Income
XOStud	DV:: Main sources of student support	DVs_Income
XChSup	DV:: Child related support	DVs_Income
XDep	DV:: Dependent grant	DVs_Income
XTeach	DV:: Teaching related support	DVs_Income
XNHS	DV:: NHS related support	DVs_Income
XDisab	DV:: Disabled student allowances	DVs_Income
XEmpl	DV:: Employer support	DVs_Income
XUniCl	DV:: Support from student s university or college	DVs_Income
XOthStud	DV:: Other (e.g. EU program/ Care Leavers, Travel)	DVs_Income
XWorkXV	DV:: Paid work	DVs_Income
XPJob	DV:: Permanent/continuous job	DVs_Income
XOthJXV	DV:: Other work	DVs_Income
XSummer	DV:: Earnings from summer vacation work	DVs_Income
XBens	DV:: Social security benefits: Total from benefits	DVs_Income
XBen01	DV:: Social security benefits: Child Benefit	DVs_Income
XBen02	DV:: Social security benefits: Child Tax Credit	DVs_Income
XBen03	DV:: Social security benefits: Retirement Pension or Widow's pension	DVs_Income
XBen04	DV:: Social security benefits: Pension Credit	DVs_Income
XBen05	DV:: Social security benefits: Carer's Allowance	DVs_Income
XBen06	DV:: Social security benefits: Employment & Support Allowance	DVs_Income
XBen07	DV:: Social security benefits: Any disability/invalidity/incapacity/sickness benefit	DVs_Income
XBen08	DV:: Social security benefits: Working Tax Credit	DVs_Income
XBen09	DV:: Social security benefits: Job Seeker's Allowance	DVs_Income

XBen10	DV:: Social security benefits: Income Support	DVs_Income
XBen11	DV:: Social security benefits: Housing Benefit	DVs_Income
XBen12	DV:: Social security benefits: Local Housing Allowance	DVs_Income
XBen13	DV:: Social security benefits: Universal credit	DVs_Income
XBen14	DV:: Social security benefits: Council tax benefit	DVs_Income
XBen15	DV:: Social security benefits: Other state benefits	DVs_Income
XOthInc	DV: Other Income	DVs_Income
XSpend	DV:: Expenditure	DVs_Expenditure
XSpend_Ex	DV:: Total expenditure excluding tuition fees	DVs_Expenditure
XLiving	DV:: Living	DVs_Expenditure
XFood	DV:: Food	DVs_Expenditure
XPersT	DV:: Personal	DVs_Expenditure
XPers	DV:: Personal	DVs_Expenditure
XEnterT	DV:: Entertainment	DVs_Expenditure
XEnter	DV:: Entertainment	DVs_Expenditure
XHHGoodT	DV:: Household goods	DVs_Expenditure
XHHGood	DV:: Household goods	DVs_Expenditure
XLivTraT	DV:: Travel	DVs_Expenditure
XLivTrav	DV:: Travel	DVs_Expenditure
XOthLivT	DV:: Other living expenditure	DVs_Expenditure
XOthLiv	DV:: Other living expenditure	DVs_Expenditure
XHouse	DV:: Housing	DVs_Expenditure
XRent	DV:: Rent	DVs_Expenditure
XRetaine	DV:: Retainer	DVs_Expenditure
XOthHous	DV:: Other (e.g. household bills)	DVs_Expenditure
XPartic	DV:: Participation Costs	DVs_Expenditure
XPartic_EX	DV: Participation costs excluding tuition fees	DVs_Expenditure
XBooks	DV:: Books and equipment	DVs_Expenditure
XParTrav	DV:: Travel and study related childcare	DVs_Expenditure
XFeeCon	DV:: Cost of fees	DVs_Expenditure
XChild	DV:: Children	DVs_Expenditure
XBorr	DV:: Borrowing	DVs_Expenditure
XCredit	DV:: Commercial credit	DVs_Expenditure
XOD	DV:: Overdraft	DVs_Expenditure
XArr	DV:: Arrears	DVs_Expenditure
XInfLoan	DV:: Informal Loans	DVs_Expenditure
XStuDebt	DV:: Outstanding Student Loan debt	DVs_Expenditure
XAccDebt	DV:: Outstanding Access to Learning funds (if to be repaid)	DVs_Expenditure
XSave	DV:: Savings	DVs_Expenditure
xdytra12	Diary DV - annual study-related travel expenditure	Diary
xdytrb12	Diary DV - annual leisure-related travel expenditure	Diary
xdytrc12	Diary DV - annual child-related travel	Diary
xdytre12	Diary DV - annual parking for college/work	Diary
xdytrf12	Diary DV - annual parking for leisure	Diary
xdytrx12	Diary DV - annual travel-related - other	Diary
xdylifa1	Diary DV - annual lifestyle -cinema, theatre concerts	Diary
xdylifb1	Diary DV - annual lifestyle - nightclubs, discos etc	Diary

xdylifc1	Diary DV - annual lifestyle - sports etc	Diary
xdylifd1	Diary DV - annual lifestyle - religious activities	Diary
xdylife1	Diary DV - annual lifestyle - national lottery/betting	Diary
xdylifx1	Diary DV - annual lifestyle - other	Diary
xdypera1	Diary DV - annual personal - clothes etc	Diary
xdyperb1	Diary DV - annual personal - CDs music DVDs etc	Diary
xdyperc1	Diary DV - annual personal - cigarettes tobacco	Diary
xdyperd1	Diary DV - annual personal - newspapers books etc	Diary
xdypere1	Diary DV - annual personal - gifts & cards	Diary
xdyperf1	Diary DV - annual personal - prescriptions medicine	Diary
xdyperg1	Diary DV - annual personal - toiletries	Diary
xdyperh1	Diary DV - annual personal - haircuts and grooming	Diary
xdyperx1	Diary DV - annual personal other	Diary
xdyfda12	Diary DV - annual food - meals out	Diary
xdyfdb12	Diary DV - annual food - beer wine spirits	Diary
xdyfdc12	Diary DV - annual food - non-alc drinks	Diary
xdyfdd12	Diary DV - annual food - other	Diary
xdyhoua1	Diary DV - annual household - food drink in	Diary
xdyhoub1	Diary DV - annual household - alcohol	Diary
xdyhouc1	Diary DV - annual household - household goods	Diary
xdyhoud1	Diary DV - annual household - servicing repairs	Diary
xdyhoue1	Diary DV - annual household - laundry dry cleaning	Diary
xdyhouf1	Diary DV - annual household - kitty	Diary
xdyhoux1	Diary DV - annual household - other	Diary
xdyx212	Diary DV - annual - other payments	Diary

Appendix 6: Trimming

Variable	2014 to 2015 trimming	2014 to 2015 value adjusted for inflation 2021 to 2022	New upper limit suggested for trimming 2021 to 2022	Percentage included
XFOOD	9906	11050.18	11050	99.6
XPERST	10000	11155.04	11155	98.2
XENTERT	10000	11155.04	11155	99.3
XHHGOODT	13000	14501.55	14501	99.9
XLIVTRAT	7500	8366.278	8366	97.6
XOTHLIVT	4500	5019.767	5019	1000
XRENT	13500	15059.3	15060	98.9
XRETAINE	5000	5577.519	5580	100
XOTHHOUS	6500	7250.774	7250	96.4
XBOOKS	3000	3346.511	3346	98.5
XPARTRAV	10000	11155.04	11155	99.9
XFEECON	9000	10039.53	11100	100
XCREDIT	35000	39042.63	39042	99.4
XHOUSE	30000	33465.11	33465	99.7
XPARTIC	20000	22310.07	22310	100
XSPEND	40000	44620.15	44620	98.7
xdytra12	4805	5359.995	5359	100
xdytrp12	6513	7265.276	7265	99.9

xdytrx12				
xdylifa1	5967	6656.211	6656	99.9
xdylifb1	2730	3045.325	3045	99.9
xdylifc1	6864	7656.818	7656	99.9
xdylifd1	2730	3045.325	3045	100
xdypera1	17277	19272.56	19272	99.9
xdypere1	6630	7395.79	7395	99.9
xdyperf1	5306	5918.863	5918	99.9
xdyperh1				
xdyfdb12	6045	6743.22	6743	99.9
xdyhoub1	3120	3480.372	3480	100
xdyhouc1	7800	8700.929	8700	99.9
xdyhoud1	12285	13703.96	13703	100
xdyhouf1				
xdyhoux1				
xdyx212	2925	3262.848	3262	100
XSAVE; XSAVEB	80000	89240.3		

Department for Education