Destination of Leavers from Higher Education Longitudinal Survey 2002/3 Cohort: assessment of robustness and fitness for purpose

Sarah Tipping and Rebecca Taylor

Destination of Leavers from Higher Education Longitudinal Survey 2002/3 Cohort: assessment of robustness and fitness for purpose

Sarah Tipping and Rebecca Taylor

Prepared for Higher Education Statistics Agency

June 2007

P2724

Contents

List of tables 4
List of figures 5
Acknowledgements 5
Introduction 6

1 SUMMARY OF DLHE LONGITUDINAL SURVEY 7
1.1 Sample design 7
1.2 Data collection methods 8
1.3 Topic coverage 8

2 NON-RESPONSE 10
2.1 Response rates 10
2.1.1 Contact, refusal and co-operation rates for Longitudinal Survey 12
2.1.2 Graduates from Scottish institutions 13
2.2 Evidence of non-response bias 14
2.2.1 Profiles 15
2.2.2 Response rates for sub-groups to the Longitudinal Survey 15
2.2.3 Multivariate analysis of non-response to the Longitudinal Survey 16

3 DATA QUALITY 19
3.1 Sampling variance 19
3.2 Data collection mode 22
3.3 Item non-response 22
3.3.1 Item non-response for specific variables 23
Income 23
Postcode of employer's location 24
Standard industry and occupation codes 24
Subject area 25
Differences by data collection mode 26
3.3.2 Item non-response in longitudinal data 26
3.4 Evidence of routing errors and missed sections 27
Section B 28
Section C 28
Sections B & C 28
Section D 29
Section E 29
Section F 29

4 SUB-GROUPS 30
4.1 Methods used 30
4.2 Results of sub-group analysis 31
4.2.1 Sample sizes available for cross-border analysis 33
4.2.2 Routed sections 36
4.3 Statistical power 38

5 ACTIVITY HISTORIES 39
5.1 Questionnaire versions sent out 39
5.2 Questionnaire versions returned 39
5.3 Non-response and non-response bias 40
5.4 Sample size – long activity history 40
5.5 Data collection mode – long activity history 41
5.5.1 Coding inconsistencies 42
5.6 Item non-response – short activity summary 42
5.7 Item non-response – long activity history 43

5.8 Date consistency – long activity history 43
5.9 Comparison of short and long versions 43
5.10 Conclusions 45

6 RECOMMENDATIONS 47
6.1 Sample bias 48
6.2 Response process information 49
6.3 Methods to increase response 49
6.4 Review of sub-group requirements 50
6.5 Section C 50

APPENDIX A SAMPLING FRACTIONS 52

APPENDIX B RESPONSE RATE DEFINITIONS 55

APPENDIX C NON-RESPONSE PROFILES 61

APPENDIX D SUB-GROUP RESPONSE RATES 65

APPENDIX E NON-RESPONSE MODEL 67

APPENDIX F LONG ACTIVITY HISTORY ITEM NON-RESPONSE 69

List of tables

Table 1-1 Questionnaire versions and sample sizes 9
Table 2-1 Response rates 12
Table 2-2 Contact, refusal and co-operation rates at Longitudinal Survey 13
Table 2-3 Weighted contact, refusal and co-operation rates at Longitudinal Survey for graduates from Scottish institutions by domicile 14
Table 2-4 Weighted contact, refusal and co-operation rates at Longitudinal Survey for graduates from Scottish institutions by activity at Early Survey 14
Table 2-5 Characteristics associated with response 17
Table 3-1 Standard errors and confidence intervals for a selection of key estimates 21
Table 3-2 Item non-response for income by questionnaire mode 23
Table 3-3 Item non-response for employer postcodes by questionnaire mode 24
Table 3-4 Item non-response for SIC by questionnaire mode 25
Table 3-5 Item non-response for SOC by questionnaire mode 25
Table 3-6 Item non-response for subject area by questionnaire mode 26
Table 3-7 The longitudinal questionnaire: sections, topics and eligibility. 28
Table 4-1 Sample size and design effects for respondent sub-groups in the Longitudinal Survey 32
Table 4-2 Country of HEI sample sizes by domicile 34
Table 4-3 Work location sample size by domicile 35
Table 4-4 Work location sample size per country of HEI 35
Table 4-5 Weighted, unweighted and effective sample sizes for routed sections of the questionnaire, by sub-group 36
Table 4-6 Effective sample sizes required for different levels of precision 38
Table 5-1 Sample size of returned questionnaires by version 39
Table 5-2 Section E long activity history – proportion completed 41
Table 5-3 Summary of the number of activity spells by questionnaire mode 42
Table 5-4 Number of jobs, short and long version 44
Table 5-5 Ever unemployed, short and long version 44

Apx. Table A.1 Sampling fractions for the Longitudinal Survey 52
Apx. Figure B.1 Derivation of final outcome for Longitudinal Survey 55
Apx. Table B.2 Final outcome for Early Survey (unweighted) 56
Apx. Table B.3 Final outcome for Longitudinal Survey for selected sample (unweighted) 56

Apx. Table B.4 Available, selected and responding sample sizes 57
Apx. Table B.5 Non-response categories for Longitudinal Survey..... 58
Apx. Table B.6 Non-response categories for Longitudinal Survey for Scottish institution sub-group by domicile 59
Apx. Table B.7 Non-response categories for Longitudinal Survey for Scottish institution sub-group by activity at Early Survey..... 59
Apx. Table B.8 Contact, refusal and co-operation rates at Longitudinal Survey for graduates attending Scottish institutions by domicile 60
Apx. Table B.9 Contact, refusal and co-operation rates at Longitudinal Survey for graduates attending Scottish institutions by activity at Early Survey..... 60
Apx. Table C.1 Profile of population and samples at different stages..... 61
Apx. Table C.2 Profile of population and non-respondents..... 63
Apx. Table D.1 Response rates to Longitudinal Survey for key sub-groups..... 65
Apx. Table E.1 Results from the logistic regression model..... 67
Apx. Table F.1 Item non-response for employment activities, by questionnaire mode 69
Apx. Table F.2 Item non-response for study/training activities, by questionnaire mode..... 70
Apx. Table F.3 Overall item non-response for key items by spell order 71

List of figures

Figure 1-1 Issued and responding sample sizes at each stage 7
Figure 2-1 Eligibility and non-response at each stage 11

Acknowledgements

We are grateful for the advice received from Catherine Benfield and Sarah Grimes at HESA; and for information provided by IFF Research who were responsible for the DLHE Longitudinal survey data collection.

Introduction

The aim of this work is to assess the 2002/3 Destinations of Leavers from Higher Education (DLHE) Longitudinal survey for fitness for purpose. DLHE data are to be used to provide evidence about the graduate labour market and careers to inform policy and research.

The assessment covers a number of different aspects:

- the response process and profile of responding graduates to determine whether they are representative of the population (chapter 2),
- identification of any groups that are less likely to respond (chapter 2),
- the quality and robustness of the data collected from respondents (chapter 3),
- whether there are sufficient numbers to conduct the desired analyses of sub-groups of interest (chapter 4),
- the quality of the activity history data (chapter 5).

The recommendations based on the findings are presented in the final chapter (6). These cover how the current data can be used, identify problematic elements, and consider changes for the future.

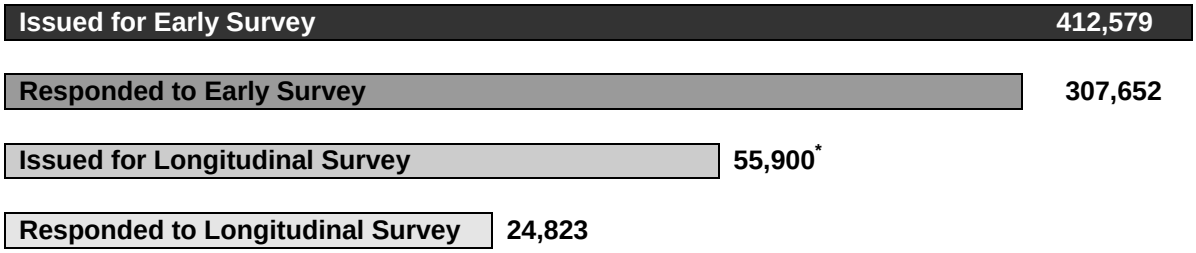
1 SUMMARY OF DLHE LONGITUDINAL SURVEY

The DLHE is carried out by the Higher Education Statistics Agency (HESA), the central source for the collection and dissemination of statistics about publicly funded higher education in the UK. The DLHE survey is the successor to the First Destination Survey series. There are two stages to the survey. The first stage is a census of individuals who have completed Higher Education courses in the UK. This stage is carried out six months after the courses end, (or longer period for a minority of eligible leavers) and is referred to as the Early Survey. The second stage of the survey is conducted on a sample of around 60,000 leavers who responded at the first stage and takes the form of a follow-up around three and a half years after leaving. This stage is referred to as the Longitudinal Survey[1].

1.1 Sample design

There were 412,579 students who graduated from full-time and part-time programmes from UK Higher Education Institutions (HEI) between 1st August 2002 and 31st July 2003 who were also domiciled (prior to joining their HEI) in the UK, Channel Islands, Isle of Man or other EU countries, and were not known to have died. This was the eligibility criteria for the Early Survey. The Early Survey was a census and, as such, was issued to all 412,579 graduates. It achieved responses from 307,652 graduates (a response rate of 75%).

Figure 1-1 Issued and responding sample sizes at each stage



* 62,039 graduates were sampled by HESA, IFF was provided with useable contact details for 55,900 of these.

The sample for the Longitudinal Survey was selected from the 307,652 graduates who responded to the Early Survey. All HEIs were asked to participate. 62,039 graduates were selected by HESA. Selected graduates could become ineligible if they had died. The HEIs provided contact details - at least one of email or mail address, or telephone number - for 55,900 of the sampled graduates; 11,817 email addresses, 55,521 postal address and 46,838 phone numbers. A response was achieved from 24,823 graduates.

There were certain sub-groups of graduates that were of specific interest. These groups were over-sampled to provide adequate numbers for analysis[2].

[1] More information on the sample design, plus other details about the survey, can be found in the Destinations of Leavers from Higher Education (DLHE) Longitudinal Study: Technical Report, put together by IFF Research.

[2] The unequal selection probabilities mean selection weights need to be applied to the data before any analyses can be carried out. Survey estimates based on unweighted data will be biased towards the over-sampled sub-groups.

The sub-groups were based on country of origin (domicile), country of study, ethnicity of graduate and type of degree course. The actual sampling fractions are given in APPENDIX A. In summary:

- All Foundation degree and RC-funded students were selected,
- All Black, Mixed and 'Other ethnic group' students were selected and 35% of Asians,
- 60% of NI domiciled students and 60% of students who studied at a NI HEI (non-NI domiciled),
- 40% of Welsh domiciled students and 25% of students who studies at Welsh HEIs (non-Welsh domiciled),
- 25% of Scottish domiciled students and 25% of students at Scottish HEIs (non-Scottish domiciled),
- 35% of TDA funded students,
- 15% of healthcare profession students, and
- 5% of all remaining students not already covered.

1.2 Data collection methods

The data have been collected using a mixture of postal, telephone and online questionnaires. The majority of respondents were contacted by post. The approach used depended on the contact details provided by the institution. The different modes were used sequentially:

1. All graduates with an email address were invited by email to fill in an online questionnaire. A week after the initial email a reminder was sent out.
2. Seven weeks after the initial email, graduates not responding plus all other graduates for whom they had a postal address were sent a postal questionnaire. A reminder and second questionnaire were sent out a month after the initial mailing.
3. Contact attempts were then made by telephone for graduates who had not responded online or by post, and for whom a telephone number had been provided. Up to seven calls were made to each respondent.

1.3 Topic coverage

There were two versions of the questionnaire. Both versions covered a set of key data items. These were; the attitudes of the graduate towards their original course, details of any other qualifications since obtained and what activities the graduate were doing on the 27th November 2006, with a detailed description for graduates where the main activity was either employment or education.

In addition, data were collected on periods of employment and unemployment the graduate had had between graduation and the 27th November 2006. For the 'snapshot' version of the questionnaire this was restricted to a basic count of the number of jobs and periods of unemployment. For the 'long' version of the questionnaire this information was collected as a detailed activity history documenting detailed job characteristics and the employing organisation for example.

All respondents who completed an online version of the questionnaire or who were interviewed by telephone were given both the snapshot and the long version. The majority of postal respondents were given the snapshot version of the questionnaire;

a sub-set were selected at random to receive the long version. A summary of the number of responses by questionnaire version is given in Table 1-1, below.

Table 1-1 Questionnaire versions and sample sizes

Version	Number of responses	Percentage of total
Long – online	2,462	10%
Long – telephone	11,418	46%
Long – postal	925	4%
Snapshot – postal	10,018	40%
Total	24,823	100%

Base: All respondents.

2 NON-RESPONSE

This chapter provides a review of the response process (using response rates) and the characteristics of the respondents to determine whether they are representative of the population or biased.

Response rates need to be calculated whenever non-response occurs in surveys. By 'non-response' we mean unit non-response which is where a unit (such as an individual) has been selected for a survey but has not taken part. The achievement of high response rates is a key criterion in the assessment of survey performance. The response rates are presented in section 2.1. 74.6% of invited graduates responded to the Early Survey. 40.0% of invited graduates responded to the Longitudinal Survey. After adjusting for the over-sampling and taking non-response at both surveys into account, the response rate is 31.0%.

The combined loss from non-response over both surveys is not insignificant. A low response rate can reduce data quality but it may not. The response rates have been broken down into components that focus on different elements of non-response (specifically non-contacts and refusals) to improve understanding of the data collection process and the quality of the data, both overall and for graduates from Scottish institutions (sections 2.1.1 and 2.1.2).

Furthermore if non-response was completely at random it would not bias survey estimates. However, this is rarely, if ever, the case. Non-respondents tend to be systematically different in some way to respondents. These systematic differences mean certain groups (or characteristics) in the sample will be over-represented and the survey estimates will be biased towards these. Section 2.2 identifies groups that are less likely to respond. In fact we find that the profile of the Early Survey respondents is very close to that of the population. As a result much of the exploration concentrates on the potential for bias from the non-response occurring at the Longitudinal Survey rather than the Early Survey. The exploration also demonstrates that the sample selected for the Longitudinal Survey - when weighted by the selection weight - is representative of the population.

The characteristics of those least likely to respond to the Longitudinal Survey are:

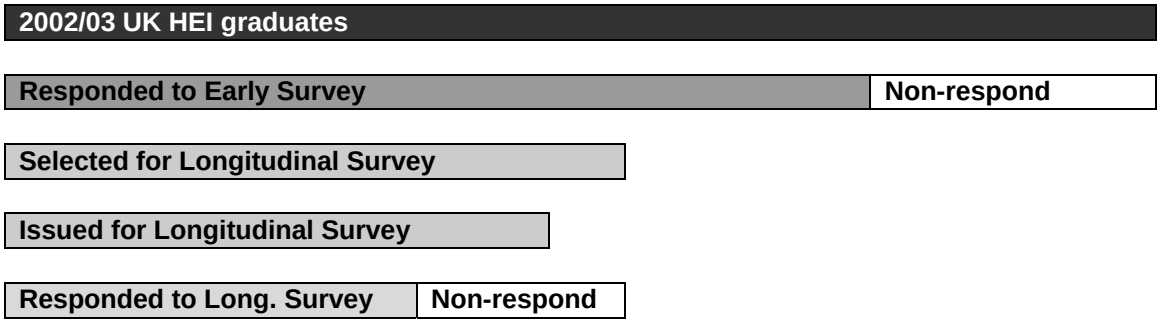
- no telephone number available for Longitudinal Survey,
- at Early Survey responded to later invitations (by phone or online),
- completing a qualification at graduation other than Doctorate or First Degree,
- no email address available for Longitudinal Survey,
- attended an HEI in England,
- obtained a lower class degree at graduation,
- younger (<30 years old),
- at Early Survey engaged in an activity of "assumed unemployed",
- were not funded by either TDA or Healthcare profession,
- male.

2.1 Response rates

The response rates are calculated from the outcomes recorded for each survey. An outcome (respond / non-respond) is needed for each eligible graduate to enable us

to understand how many graduates provided a response compared to the population of graduates represented by the data. A summary of the eligibility for each stage of the surveys is given in Figure 2-1. All graduates from UK HEIs in 2002/3 were eligible for the Early Survey.[3] The outcomes recorded for the Early Survey indicate whether a graduate responded or not, and for non-respondents whether it was a refusal or not. Surveys often record more detailed information about the outcome of the invitation to participate, however in this case there is no other information available. Partly this is due to data being collected separately by each HEI which processes the information in different ways. Also the specification of data that is compulsory for HEIs to provide may not include detailed outcomes.

Figure 2-1 Eligibility and non-response at each stage



Eligibility is different for the Longitudinal Survey. All respondents to the Early Survey 'qualified' for the Longitudinal Survey. A sample was selected from the pool of 'qualifiers'. This selected sample are the group of graduates who were eligible for the Longitudinal Survey. Some of the selected graduates could not be issued because (a) their contact details were not complete or (b) they chose not to take part during an opt-out procedure conducted in advance.[4] This group of eligible graduates that were not issued are included in the response rate calculations as non-respondents.

The outcomes from the Longitudinal Survey give more detailed information than the outcomes from the Early Survey. They indicate whether a graduate responded or not, and for many non-respondents we know whether the non-response was due to non-contact, refusal or some other reason.[5] Whether graduates were actually contacted is not always known (this is a feature of some of the data collection modes and cannot be helped).

Response rates have been calculated separately for each survey as well as for the surveys combined (called the longitudinal rate). Table 2-1 presents the unweighted and weighted response rates. The exact definitions used for this study are given in APPENDIX B. The unweighted response rates are calculated from the data without any adjustments; all eligible graduates are included and treated equally. While this measures the process of response for the selected sample it is not an accurate reflection of the population the sample aims to represent because some sub-groups were over-represented deliberately. The weighted response rates account for this, giving eligible graduates varying influence depending on their chance of selection,

[3] Any graduates known to have died before the Early Survey were not eligible and have been removed.
[4] A small number of HEIs conducted an opt-out exercise. HEIs did not supply the data collection organisation with the contact details of graduates who chose to opt out of the survey. Graduates who opted out can be identified from one HEI but not others.
[5] Additional refusals were recorded but were not included in the supplied outcomes. Seventeen graduates are known to have died – and therefore were not eligible. They remain in the data (classified as non-contacts) due to the late receipt of these information. Other deaths remain unknown. See APPENDIX B for further information.

(see section 1.1 for a description of the selection). Therefore the unweighted rates can be used to understand the success of the data collection process, whereas the weighted rates correct for the unequal selection and show response for a sample that is representative of 2002/3 HE leavers.

Table 2-1 Response rates

	Response rate (unweighted)	Response rate (weighted)
Early Survey	74.6%	n/a
Longitudinal Survey	40.0%	41.6%
Longitudinal – 1st & 2nd survey	n/a	31.0%

Bases: All eligible graduates.

The unweighted rates in Table 2-1 show that a higher percentage of graduates responded to the Early Survey (74.6%) than the Longitudinal Survey (40.0%). There are a number of reasons for this, including contact details becoming out of date and reduced saliency of their HE experience, both as a result of the length of time since graduation. Once the composition of the sample selected for the second sample has been adjusted to represent the sample the response rate increases slightly (41.6%).

The weighted longitudinal rate is of interest to anyone using information collected from both surveys whose aim is to say something about 2002/3 HE leavers. It considers the cumulative non-response to each stage by applying the response rate for the Early and Longitudinal Surveys to the sample selected for the Longitudinal Survey. The rate shows that 31.0% responded to both surveys.[6] Since the response rates show that the loss from non-response at each wave is substantial, we explore the non-response further in the following sections.

2.1.1 Contact, refusal and co-operation rates for Longitudinal Survey

Here we consider the different elements of non-response at the Longitudinal Survey[7]. Table 2-2 presents weighted contact,[8] co-operation and refusal rates at the Longitudinal Survey. Given the uncertainty about the outcome data (including some excluded information), these rates should be treated with caution. The contact rate gives an indication of the quality of the contact details. The co-operation rate shows the level of response for the sample who could be contacted. This is the sub-sample who we can be sure received the invitation to participate in the survey and made an active decision to respond or not. The co-operation rate is calculated because the response rate includes the sub-sample of non-respondents who could not be contacted, but may not have actively decided not to participate. The refusal rate is calculated because refusal, as a reason for non-response, is of particular interest. It is thought that the refusers have more potential to cause bias because often there is a closer link between the refusal and the survey topic than for other non-response.

The weighted rates in Table 2-2 show that just over half of graduates eligible to participate at the Longitudinal Survey were definitely contacted (51.3%). This is relatively low. However it is likely that some of those for whom contact is unknown and have been classified as non-contacts did receive their invitation. We cannot

[6] This should not be interpreted as data from both surveys being available for 31% of all 2002/3 HE leavers (24,283 or 6% of all leavers provided data at both surveys).
[7] Equivalent information is not available for the Early Survey.
[8] Non-contact rate = 100% - contact rate.

ascertain whether people received their postal questionnaire or email invitation to the online questionnaire when no response - either positive or negative - was received.

Table 2-2 Contact, refusal and co-operation rates at Longitudinal Survey

Definition	Rate (unweighted)	Rate (weighted)
Longitudinal Survey contact	49.5%	51.3%
Longitudinal Survey refusal	7.8%	8.0%
Longitudinal Survey co-operation	80.8%	81.0%

Bases: Contact = all eligible graduates. Refusal = all eligible graduates. Co-operation = all contacted graduates.

A refusal to participate was received from 8% of those eligible. Refusals were recorded for non-response to telephone contacts only. The inability to identify refusals for other questionnaire modes is a feature of the data collection process. Given this, and the uncertainty about the outcome data, we suggest that the level of refusals indicated by the rate calculated is thought of as a minimum.

Where contact could be confirmed, it can be seen that 81.0% co-operated with the survey request.

2.1.2 Graduates from Scottish institutions

Two specific sub-groups of interest were identified by HESA's statutory customers:

- graduates who attended Scottish institutions broken down by domicile (Scottish domiciled versus the rest),
- graduates who attended Scottish institutions who stated they were unemployed at the Early Survey.

There are 38,266 graduates identified as having attended a Scottish institution. 7,287 were selected for the Longitudinal Survey and are analysed here. 25% of selected graduates who attended Scottish institutions were domiciled outside of Scotland (1,767). 6% declared themselves to have been unemployed in the Early Survey (452).

The tables below show the weighted contact, refusal and co-operation rates. Given the uncertainty about the outcome data these rates should be treated with caution. See APPENDIX B for the unweighted rates and the definition of unemployment. Each sub-group of interest has been compared to the remaining graduates, i.e. Scottish domiciled compared to other UK/EU domiciled, and those unemployed at the Early Survey are compared to those doing other activities. Table 2-3 shows that Scottish domiciled graduates from Scottish institutions had a significantly higher contact rate than other UK/EU domiciled graduates although the rates are not sizeably different (51.8% compared to 47.3%).[9] The refusal and co-operation rates were not significantly different.

None of the rates shown in Table 2-4 are significantly different when tested, however the sub-group who were unemployed is relatively small.

[9] Significance testing conducted in Stata to account for the complex sample design (stratification and weighting).

Table 2-3 Weighted contact, refusal and co-operation rates at Longitudinal Survey for graduates from Scottish institutions by domicile

Definition	Scottish domicile	Other UK/EU domicile
Longitudinal Survey contact	51.8%	47.3%
Longitudinal Survey refusal	7.7%	7.6%
Longitudinal Survey co-operation	82.4%	81.5%

Bases: Contact = all eligible graduates from Scottish institutions. Refusal = all eligible graduates from Scottish institutions. Co-operation = all contacted graduates from Scottish institutions.

Table 2-4 Weighted contact, refusal and co-operation rates at Longitudinal Survey for graduates from Scottish institutions by activity at Early Survey

Definition	Unemployed at 1st survey	Other activity at 1st survey
Longitudinal Survey contact	53.0%	50.7%
Longitudinal Survey refusal	9.2%	7.5%
Longitudinal Survey co-operation	80.3%	82.4%

Bases: Contact = all eligible graduates from Scottish institutions. Refusal = all eligible graduates from Scottish institutions. Co-operation = all contacted graduates from Scottish institutions.

2.2 Evidence of non-response bias

The loss from non-response at the Longitudinal Survey and the combined loss over both surveys is not insignificant. A low response rate can reduce data quality but it may not invalidate the data completely. As a result an investigation of the bias caused by the non-response is presented in this section. Non-respondents tend to be systematically different in some way to respondents. These systematic differences mean certain groups (or characteristics) in the sample will be over-represented and the survey estimates will be biased towards them.

The very nature of non-response bias makes it difficult to assess. Essentially we need information about the non-respondents, but being non-respondents means that we do not have the information needed. Another important aspect of non-response bias is that it can only be measured relative to the estimate of interest. This implies, for example, that a survey's estimate of the average income earned may be biased but the estimate of the proportion of women currently studying may not be biased. Although it is not possible to calculate a general measure of bias we can explore where bias may exist using the information we have about non-respondents from the Student Record and Early Survey. If any of this information is highly correlated to the survey estimates we could calculate the relative bias. With limited variables (as is the case here) we are restricted to using it to provide evidence about possible bias.[10]

The exploration identifying groups that are less likely to respond follows a number of steps: comparing the profile of the respondent samples with the population (2.2.1), exploring differences in response rates for sub-groups (2.2.2), and modelling the likelihood of non-response (2.2.3). We had also anticipated exploring other data sources to examine bias of survey estimates. This has not been conducted because it is not likely to provide strong evidence of bias because of definition differences between data sources and because of time constraints.

[10] Any variables that are correlated are likely to be used in the weighting strategy, thereby adjusting for the bias.

2.2.1 Profiles

The profiles are shown in APPENDIX C.

The profile of the Early Survey respondents is very close to that of the population. The high response rate to the Early Survey may be a contributing factor. No further analysis has been conducted of the non-response to the Early Survey.

Similarly, the sample selected for the Longitudinal Survey is also very close to the Early Survey sample profile (when weighted by the selection weight). This is reassuring as it demonstrates that once the known bias from over-sampling certain sub-populations is accounted for, the Longitudinal Survey selected sample is representative of the population. This also means that any differences between the profile of the Longitudinal Survey respondents and the profile of the Early Survey respondents are due to non-response at the Longitudinal Survey, rather than bias from the sample selection process.

The weighted profile of the Longitudinal Survey respondents is the most different to the population profile. This has been caused by differential non-response to the Longitudinal Survey. In fact the profiles are broadly similar, with some fluctuation especially where there are smaller sample sizes. This is encouraging as it suggests that there is not any *strong* differential bias in the characteristics considered and any bias relating to these characteristics is likely to be marginal. Further explorations of the non-response to the Longitudinal Survey are presented in the following sections. A discussion of the characteristics in the profiles that differ is withheld since the findings in the next section mirror those found here.

2.2.2 Response rates for sub-groups to the Longitudinal Survey

Here we present an exploration of the differences in response rates for sub-groups to the Longitudinal Survey. If response rates vary by respondent sub-group and there is also variation in the survey estimates by sub-group, then there is the potential for the estimates to be biased. Differences in response rates for sub-groups are provided as an indicator of potential bias. It is up to the analyst to establish the link with their survey estimates of interest. The response rates are provided as a resource for analysts to use because all possible survey estimates cannot be considered here.

Response rates for key sub-groups have been generated using the method outlined in Section 2.1. Full results can be found in APPENDIX D.

There is some variation in response by sub-group compared to the average of 41.6%. This is not unexpected since response to surveys often varies in predictable ways by sex, age and ethnicity for example. Some of the expected patterns are repeated here. Male graduates were less likely to respond than female graduates, with a response rate of 39% compared to 44%. The oldest graduates, aged 30+ years, have a response rate of 47% and were more likely to respond than graduates of other ages; the worse response being from the group aged 25-29 years (35%). Response was poorer amongst graduates from a non-white ethnic background, particularly black graduates (32%).

There were some differences in response rates by country of domicile and study. Graduates who were domiciled in or studied in Northern Ireland were most likely to respond, (46% and 48% respectively). The response rates for other countries were very similar, around the average. We can speculate that this could be because these

groups are less mobile or because the Northern Ireland HEIs simply hold better contact details; making them easier to trace. This could be confirmed by checking the relative contact rates of these sub-groups.

The response rate varied by their graduating course subject. The two subjects exhibiting the largest differences are: (i) graduates who studied “Architecture, Building and Planning” were around 10% less likely to respond than the average, and (ii) graduates who studied “Eastern, Asiatic, African, American and Australasian language/literature/other” were around 10% more likely to respond than the average. It is likely that the differences in the gender balance of each subject is related to the differences in response (which can be tested by a multivariate analysis).

Any differences in the response rate by the level of study are not apparent until the level is dis-aggregated. Even dis-aggregated the rates are similar, with the exception of PhD graduates, who comprised a small group with a high response rate (47%). Graduates who had studied part-time also had a higher response rate (46%).

Graduates who combined work with further study at the Early Survey were relatively more likely to respond (46%) and those assumed to be unemployed were relatively less likely (37%).

The higher rates for sub-groups were studied part-time or combined work with study may have been less mobile, meaning their contact information would be better.

This exploration has shown how response to the Longitudinal Survey varied by the characteristics of their original course as well as by demographic characteristics. There is potential for this variation to cause bias in survey estimates calculated from the respondents, if there is a link between the characteristics and the estimates. Up to this point, the relationship of each factor with response behaviour has been looked at separately. This is known as univariate analysis. To increase the scope of our analyses we also carried out multivariate analyses. Multivariate analysis is where the relationship of a number of variables is analysed at the same time, allowing us to take a number of characteristics into account simultaneously.

2.2.3 Multivariate analysis of non-response to the Longitudinal Survey

The aim of the multivariate analysis was to look into the relationships between response behaviour and the graduates’ characteristics in more detail. It is possible that significant associations can appear between variables because of their relationship to other variables (for example, mode of study may be related to level of study and both variables may be related to response). A multivariate analysis looks at all the variables in relation to each other as well as in relation to the outcome variable (in this case, response). In instances where two variables are both strongly related to the outcome variable, but also strongly related to each other, the model will exclude the variable with the weaker relationship. A number of variables from the Student Record and Early Survey were used to predict whether or not the graduate responded to the Longitudinal Survey.

An overview of the non-response model is given below, with the characteristics ordered by strength of the effect (strongest first). The full model can be found in APPENDIX E. The regression modelled the propensity of graduates who had been eligible for the Longitudinal Survey.

Table 2-5 Characteristics associated with response

More likely to respond	Less likely to respond
- Telephone number available	- No telephone number available
- Responded to 1st mailing of Early Survey	- At Early Survey responded to later mailings, by phone or online
- Completed a doctorate	- Completed qualification that wasn't doctorate or first degree
- Email address available	- No email address available
- HEI in NI, Scotland or Wales	- HEI in England
- Higher class degrees	- Lower class degrees
- Older (30+ years)	- Younger (25-29 years)
- In part time work at Early Survey	- Assumed unemployed at Early Survey
- Funded by TDA or healthcare	- Other funding
- Female	- Male

The multivariate analysis excluded some of the characteristics that were shown to have a relationship with response earlier: country of domicile, ethnicity, mode of study and course subject. Their exclusion will have been because of their relationship with other included characteristics or because their relationship with response was weaker. For the included characteristics, if their relationship with response has been affected by the inclusion of other characteristics the change is highlighted in the discussion below.

The propensity of a graduate to respond was positively related to the provision by the HE institute of either a telephone number or email address, regardless of the mode the graduate used to respond by. It should be noted that around 10% of the sample included in the model had no contact details for any mode, which may have influenced this finding. So that while this appears to be an important bias in the sample we are not able to determine whether the email and telephone contacts were significantly more successful at securing graduates response with the information we have available to us.

The method and speed of the graduate's response to the Early Survey was also significantly related to the graduate's response behaviour at the Longitudinal Survey. Respondents who had responded to the first mailing for the Early Survey were much more likely to respond to the Longitudinal Survey. Respondents who responded late, or responded by phone or online were less likely to respond at the Longitudinal Survey. This may be a proxy measure of motivation to respond.

A number of demographic characteristics were associated with response, although the magnitude of the variation was not as great as for some of the other types of characteristics. Graduates who attended institutions in Northern Ireland, Scotland and Wales were more likely to complete the Longitudinal Survey than those attending in England. This is different to earlier analyses that showed graduates in Northern Ireland only to be particularly likely to respond. Graduates in the oldest age category and females were more likely to respond than other age groups and males.

The characteristics of the course studied at the HEI were also related to response. Graduates who had completed doctorates were most likely to respond, although their numbers in the sample were small. Graduates completing their first degree were also

very likely to respond (not noticeably different in previous analyses), graduates completing all other levels of course were less likely to respond. The classification of the first course was significantly related with response. Respondents whose course did not fit the degree classification – “not applicable” - were most likely to respond. This was not seen previously. Graduates with lower class degrees were generally less likely to respond than graduates with higher-class degrees, although the lowest response was for graduates with thirds, rather than unclassified degrees. There were also differences by funding; graduates being funded by the TDA and Healthcare profession were more likely to respond than graduates being funded from other sources. The increased likelihood of response from Healthcare funded graduates was seen previously, but not from TDA funded graduates.

Finally, the model shows that the respondent's activity at the Early Survey was significantly related to their likelihood of response at the Longitudinal Survey. Graduates in part-time paid work were most likely to respond, while those assumed to be unemployed were least likely.

Although the profile comparison showed that the Longitudinal Survey respondents were broadly similar to the population the multivariate analysis has identified characteristics that differ significantly. In situations like this it is usual to try to account for the differences by calculating weights from the non-response model to apply to the responding sample.

3 DATA QUALITY

In order to inform HESA and its Statutory Customers about the fitness for purpose of the DLHE Longitudinal Survey we have assessed the survey for data quality and robustness. The analyses sought evidence of sampling variance, effects of the data collection mode, routing errors, coding errors, item non-response (both to the Longitudinal Survey and across both surveys), and quality of information for coded variables.

This section concentrates on assessing the data available for respondents to the Longitudinal Survey only. The Longitudinal Survey received responses from 24,823 graduates. This is a sufficiently large enough sample for the key estimates planned on the full sample. The sample design reduced the efficiency of the sample to 41% (equivalent to an effective sample size of 10,120). However the loss is not so great that it undermines the sample, the large sample size means precision of estimates based on the whole sample are not unduly affected (section 3.1).

Any affect from the mode of data collection (online, postal or telephone) on responses is likely to be small (section 3.2). A robust analysis of the mode effects on responses could not be conducted in the time given. Item non-response rates and the prevalence of missed sections differed by mode.

The level of item non-response was low; less than 5% for the majority of questions in the survey (section 3.3). Since the level of item non-response for the Early Survey was also low (less than 1% for most questions), item non-response should not have a significant impact on most analyses that combine data from both surveys (section 3.3.2). Levels tended to be higher for more complex questions and in Section C. Questions identified with higher rates of item non-response were: income, employer location, Standard Occupation Code and subject of study/training which has reduced the associated sample sizes (section 3.3.1). Overall, the proportion of graduates missing sections was small and should not compromise data quality (section 3.4). The proportion of eligible respondents missing out the section of the questionnaire covering study/training activities (Section C) was relatively high compared to other sections.

3.1 Sampling variance

Sampling variance measures the amount of sampling error. This is the error present in all samples that results from the estimate being based on a sample of people from the population, rather than a full census. The error is said to be random, since the results obtained for any single sample may, by chance, vary from the true values for the population but the variation would be expected to average to zero if the survey was repeated enough times. The amount of sampling variance depends on the size of the sample, the sample design and the weighting method.

The sample for the Longitudinal Survey was designed to meet the needs of a number of data users. Their needs were different and competing. To meet these needs as fully as possible, key sub-groups had to be over-sampled. The resulting data for analysis has to be corrected for the over-sampling using selection weights. Selection weights are generated as the inverse of the selection probabilities. The effect of this design element is considered here.

Sampling error means the estimates produced from a sample will rarely be identical to the population value. However, we are able to generate standard errors to measure the accuracy of any survey estimate. When, as in the case of the Longitudinal Survey, the sample design is not a simple random sample, the standard error needs to be multiplied by a design factor (deft). The deft is the ratio of the standard error with a complex sample design to the standard error from a simple random sample (SRS) of the same size[11]. It measures the extra error in the sample that results from there being a complex sample design. The standard error and deft can be estimated using sample data and allow the calculation of confidence intervals. The confidence intervals give an indication of the range in which the true population value is likely to fall.

The design effects (deffs) can be used to estimate the effective sample size. This is the size a simple random sample would need to be to give estimates of the same precision as the complex design in question. The effective sample size can be estimated by dividing the unweighted sample size by the deff. The Longitudinal Survey has an actual sample size of 24,823 and an effective sample size of 10,120. The closer the effective sample size is to the actual sample size, the more efficient the sample design. An efficient sample has smaller standard errors and smaller confidence intervals. This gives the estimates greater precision, since we can be more confident the actual population value lies close to the survey estimate. The ratio of the effective sample size to the actual unweighted sample size, expressed as a percentage, is a useful measure of sample efficiency. The sample for the Longitudinal Survey has a sample efficiency of 41%, this is relatively low.

Table 3-1 shows a number of key estimates from the Longitudinal Survey, along with their standard errors, confidence intervals and design factors. The average deft for the Longitudinal Survey is 1.57, this is fairly constant across different estimates. This means, on average, the standard errors for survey estimates are inflated by 57% because of the sample design.

The main cause of this loss of efficiency is the variable selection weights (that correct for over-sampling certain areas and sub-groups)[12]. There were large differences in the proportion of the population selected in different sub-groups, resulting in large differences in the size of the weights. Sample design often involves a trade-off between statistical efficiency and practical necessity. The different needs of the data users meant that key sub-groups had to be over-sampled at such high rates.

However, the loss of efficiency is not so great that it undermines the sample. Table 3-1 shows that the standard errors and confidence intervals are still relatively small, although there is a loss of efficiency, the information available is still sufficient for requirements. This is due to the large sample size. For example the data from the Longitudinal Survey show that the main activity for 87.6% of the sample is employment. The 95% confidence interval for this estimate is 86.9 to 88.2. These confidence intervals are narrow; the true value of the population is likely to be within only 0.65 of a percentage point of the survey estimate. A second example is taken for a question from the routed section of the questionnaire that has a lower sample size. The data show 21% of the respondents that were studying or training on the reference date were doing a higher degree, mainly by taught research. The 95%

[11] SRS is statistically the most efficient method of selecting samples, giving the most precise estimates. However, the method is not used in most real life situations as it is usually impractical. Complex samples are compared to SRS to evaluate their statistical efficiency and precision

[12] The selection weights range from 1 to 20.

confidence interval for this estimate is 18.9% to 23.3%, hence the true value of the population is likely to be within only 2 percentage points of the survey estimate.

Table 3-1 Standard errors and confidence intervals for a selection of key estimates

	Base unwtd	Estimate (%)	Std Err	Lower 95% CI	Upper 95% CI	Deft	Deff
Activities on 27 November 2006							
	24,823						
Employed either full time or part time		89.7	0.31	89.0	90.3	1.61	2.59
Unemployed looking for work		2.3	0.15	2.0	2.6	1.58	2.50
Engaged in study or training		16.0	0.36	15.3	16.7	1.56	2.43
Main activity on 27 November 2006							
	24,823						
Employed either full time or part time		87.6	0.34	86.9	88.2	1.61	2.58
Unemployed looking for work		2.1	0.14	1.8	2.4	1.58	2.48
Engaged in study or training		6.5	0.25	6.0	7.0	1.61	2.60
Retired		0.9	0.11	0.7	1.1	1.82	3.33
Housewife/homemaker/carer		0.6	0.08	0.4	0.8	1.60	2.57
Maternity leave		0.7	0.07	0.5	0.8	1.34	1.79
Travelling/break/gap year/time out		0.8	0.10	0.7	1.0	1.67	2.79
Long term sick		0.2	0.04	0.1	0.3	1.58	2.50
Doing something else		0.6	0.08	0.5	0.8	1.66	2.74
Which best describes the employment you were in on 27 November 2006?							
	22,021						
Employed full time		85.4	0.37	84.7	86.1	1.56	2.42
Employed part time		9.5	0.29	8.9	10.1	1.49	2.22
Self employed or freelance		4.6	0.24	4.2	5.1	1.67	2.80
Voluntary work/other unpaid work		0.5	0.08	0.3	0.6	1.66	2.76
Annual salary in Pounds Sterling (banded)							
	19,079						
Under £10,000		4.1	0.23	3.7	4.6	1.57	2.48
£10,000 or more to £12,500		5.2	0.26	4.7	5.7	1.62	2.64
£12,500 or more to £15,000		6.4	0.28	5.9	7.0	1.60	2.56
£15,000 or more to £17,500		9.2	0.33	8.5	9.8	1.58	2.51
£17,500 or more to £20,000		10.7	0.36	10.0	11.4	1.59	2.54
£20,000 or more to £22,500		14.3	0.39	13.6	15.1	1.56	2.42
£22,500 or more to £25,000		10.1	0.33	9.5	10.8	1.49	2.23
£25,000 or more to £30,000		17.2	0.41	16.4	18.0	1.51	2.27
£30,000 or more to £50,000		18.6	0.44	17.7	19.5	1.55	2.41
£50,000 or more		4.1	0.24	3.6	4.6	1.67	2.78
Whether studying full time or part time on 27 November 2006							
	3,545						
Full time		48.9	1.32	46.3	51.5	1.57	2.46
Part time		51.1	1.32	48.5	53.7	1.57	2.46
Which of the following best describes the type of qualification you were aiming for?							
	3,787						
Higher degree mainly by research		21.0	1.08	18.9	23.2	1.64	2.68
Higher degree mainly by taught course		18.7	0.98	16.8	20.6	1.55	2.40
Postgraduate diploma or certificate		14.9	0.92	13.1	16.7	1.58	2.51
First degree		12.2	0.76	10.7	13.7	1.43	2.06

Other diploma or certificate		7.5	0.64	6.2	8.7	1.50	2.24
Professional qualification		15.6	0.95	13.7	17.4	1.61	2.59
Other qualification		6.8	0.62	5.6	8.0	1.52	2.30
Not aiming for a qualification		3.3	0.46	2.4	4.2	1.59	2.54
Did you obtain any other qualifications between the time you completed your course in 2002/03 and 27 November 2006?							
Yes	24,526	38.8	0.49	37.9	39.8	1.57	2.47

3.2 Data collection mode

The use of different data collection instruments, or modes, can result in mode effects, where the same question asked in different modes could produce different results. This is because respondents may be interpreting and answering questions differently when they are presented to them in different modes[13].

Studies have shown mode effects to be smaller for behavioural questions than attitudinal questions, with the effects being particularly large for sensitive questions[14]. This would imply the majority of the questions asked in this survey would not be greatly affected. The exceptions are two questions in Section F, which could be classed as attitudinal, and the income questions in Section B and the activity history (Section E).

If mode effects are present you would expect to see differences in the responses by mode. However for this survey there are other mechanisms that could be causing differences. The modes were allocated on the basis of the availability of contact details and subsequently many graduates could choose their preferred mode. These are selective processes that imply that the underlying characteristics of respondents using each mode could differ. It's these characteristics that could explain any differences in responses. If groups with similar characteristics could be created from each mode respondent group this would eliminate the selective processes to leave only mode effects. Given more time and resources this might be possible using propensity score matching techniques. A comparison of the profiles of the respondents and survey estimates by questionnaire mode can provide an indication of differences in characteristics (omitted due to time constraints). Item non-response by questionnaire mode is considered in the next section.

3.3 Item non-response

Item non-response is the term given to absence of information on individual data items for respondents who have provided information elsewhere. Item non-response is generally higher for postal self-completion methods than electronic or interviewer assisted methods, where the questionnaire can be designed to force an answer or where an interviewer can encourage an interviewee to respond.

For the purposes of this report we have designated item non-response as missed questions/missing information and refusals. For many questions an explicit 'don't know' answer category has been provided. In these instances we have not included

[13] Dillman D. A. (2000) Mail and Internet Surveys: The Tailored Design Method, 2nd Edition, John Wiley and Sons, Inc.
[14] Reference as previous footnote.

this category as item non-response, since an existing answer category was used by the respondent. This means the focus of the analysis is on refusals and skipped questions, where no answer category at all has been ticked.

We only have clean data, so we are unable to comment fully on some aspects of non-response, for example it was not possible to tell if respondents has responded more than once to single coded questions. In addition we are unable to comment on the amount of work that was required to clean and code variables. What we can comment on is the amount of usable data we have in the sample and whether this fits the requirements.

In general the level of item non-response was low; less than 5% for the majority of questions in the survey. This is a positive sign, since low item non-response is an indicator of good data quality. Item non-response tables for each variable are given in the Longitudinal Survey technical report[15]. The tables show there is quite a lot of fluctuation in the amount of non-response, with levels being higher for more complex questions. Levels of item non-response are also generally highest for Section C of the questionnaire, which covers study and training carried out by the graduate on the reference date.

3.3.1 Item non-response for specific variables

This section looks in more detail at a sub-set of questions that were expected to be problematic: income, employer location, industry, occupation and subject of study/training. These questions tend to be more sensitive in nature or to place a higher cognitive burden on the respondent. Each of these variables is looked at in turn below. Information provided on occupation, industry and subject of study/training must be coded before it can be used. This allows easier and better comparison of the data. If adequate information is not provided then the data cannot be coded or used easily.

Income

Respondents in employment were asked to provide their annual gross income in Section B of the questionnaire. Income is generally a sensitive topic and item non-response for income questions is always high.

Table 3-2 Item non-response for income by questionnaire mode

	Telephone	Online	Postal – long	Postal - snapshot	Total
	%	%	%	%	%
Refused	*15.1*	*15.3*	*0.0*	*0.0*	*8.3*
Question not answered	*1.9*	*4.0*	*15.0*	*9.1*	*5.6*
Total item non-response	18.0	19.3	15.1	9.1	14.3
Total (unweighted)	10,013	2,181	842	9,233	22,269

Base: All respondents in employment on the reference date.

[15] Item non-response for each variable in the Longitudinal Survey is given in Appendix J of the Destinations of Leavers from Higher Education (DLHE) Longitudinal Study: Technical Report.

14.3% of the eligible respondents did not answer the income question. On top of this less than one per cent said they did not know their income. This leaves nearly 86% of the sample with a usable income variable. Table 3-2 shows the item non-response to income by questionnaire mode.

Item non response for this question is relatively high, and ranges across questionnaire modes from 9.1% to 19.3%. It is highest for the online questionnaire and lowest for the snapshot postal questionnaire. The lower level of item non-response for the snapshot postal questionnaire is difficult to explain. It might have been anticipated that respondents experienced less fatigue with the snapshot questionnaire. However this can be ruled out because the income question is positioned early in the questionnaire, at the same point in the two versions.

Postcode of employer's location

Respondents in employment were asked to provide details on their employer's location in Section B. Respondents were asked to provide the town they were based in, along with the postcode district [16]. 9.4% of respondents working in the UK did not provide sufficient information to allow a postcode district to be looked up (Table 3-3). For the cases where postcode districts are available it is not clear whether the respondent provided them directly or whether they were derived by the data collection organisation using the name of a town given by the respondent, For most cases there is a postal district available for respondents employed in the UK, although in less than one per cent of cases this was only the postcode area.

Table 3-3 Item non-response for employer postcodes by questionnaire mode

	Telephone	Online	Postal – long	Postal - snapshot	Total
	%	%	%	%	%
Item non-response	16.8	0.8	4.0	3.8	9.4
Total (unweighted)	10,013	2,181	842	9,233	22,269

Base: All respondents in employment on the reference date.

The Longitudinal Pilot demonstrated that compared to other modes the postal respondents were more likely to provide employer postcodes and suggested that this was probably because they had the time to look them up. This would also be the case for online respondents. This finding was replicated in the Longitudinal Survey and is demonstrated in Table 3-3.

Standard industry and occupation codes

All respondents in employment on the reference date were asked for details about their occupation and industry. The responses from these questions were used to allocate respondent's Standard Industry Classification (SIC) and Standard Occupational Classification (SOC) codes. These questions require a degree of effort from the respondents in order to provide information at the required level, hence we expect item non-response for these questions to be relatively high. Item non-response can be due to responses being inadequate, as well as being skipped completely.

[16] More information on postal geography can be found on the ONS web-site at http://www.statistics.gov.uk/geography/postal_geog.asp

The levels of item non-response by mode are shown in Table 3-4 and Table 3-5. 4.8% of respondents either missed the industry questions or did not provide enough information to allow their industry to be classified using the SIC codes. 10.3% of respondents either did not provide enough information to allow their occupation to be classified using SOC codes or skipped the question altogether. 1.1% of the respondents did not provide enough information to code either variable. For the majority of these cases the respondent did answer the question but did not provide enough information to allow their answer to be coded. The proportion of respondents skipping the question altogether was low, around 1% for both variables.

It can be seen that for both variables item non-response is highest for the snapshot postal questionnaire and lowest for the online survey. As expected, the postal respondents were more likely to skip questions than the telephone or online respondents.

Table 3-4 Item non-response for SIC by questionnaire mode

	Telephone	Online	Postal – long	Postal - snapshot	Total
	%	%	%	%	%
Not enough info for coding	*3.7*	*1.9*	*3.6*	*4.4*	*3.8*
Question skipped	*0.2*	*0.3*	*1.9*	*2.0*	*1.0*
SIC total item non-response	3.9	2.2	5.5	6.4	4.8
Total (unweighted)	10,013	2,181	842	9,233	22,269

Base: All respondents in employment on the reference date.

Table 3-5 Item non-response for SOC by questionnaire mode

	Telephone	Online	Postal – long	Postal - snapshot	Total
	%	%	%	%	%
Not enough info for coding	*9.9*	*7.0*	*8.1*	*9.7*	*9.5*
Question skipped	*0.2*	*0.4*	*1.2*	*1.6*	*0.8*
SOC total item non-response	10.1	7.3	9.3	11.3	10.3
Total (unweighted)	10,013	2,181	842	9,233	22,269

Base: All respondents in employment on the reference date.

Subject area

All respondents studying or in training were asked to provide information about their course in Section C of the questionnaire. The information was then coded using the JACS subject classification code. Respondents that gave inadequate information could not be coded. A quarter of the eligible respondents either missed this question entirely or gave information that could not be coded. This is relatively high. However it varied by mode from almost 6% of online respondents to a third of respondents who filled in the snapshot postal questionnaire (see Table 3-6). The dramatically higher level of item non-response overall is caused by the higher item non-response

for postal respondents. This level of non-response, combined with the smaller number of respondents who were studying or in training means that there will only be a small sample for analysis.

Table 3-6 Item non-response for subject area by questionnaire mode

	Telephone	Online	Postal – long	Postal - snapshot	Total
	%	%	%	%	%
Not enough info for coding	*8.8*	*5.1*	*12.7*	*13.9*	*11.9*
Question skipped	*0.5*	*0.6*	*15.1*	*19.9*	*13.3*
Total subject item non-response	9.3	5.8	27.8	33.9	25.3
Total (unweighted)	996	311	245	2,475	4,027

Base: All respondents in study or training on the reference date.

The difference between modes is most apparent when questions were skipped rather than not giving enough information for coding (which was only slightly higher for postal respondents). The techniques used in the online and telephone questionnaires seem to be more successful at persuading respondents to answer questions. Both of these modes have a very low level of skipped questions. Despite this, 9% of telephone respondents did not give enough information to allow their subject to be successfully coded. This suggests the interviewers could have been more successful at probing respondents until they had adequate information. While the level of item non-response for this mode is not ruinous, more training for interviewers could reduce the level of un-coded telephone responses if required.

Differences by data collection mode

For the specific questions considered in this section the online questionnaire respondents had the lowest rates of item non-response on the whole. An exception being the income question where online respondents had the highest level of item non-response. There is no evidence available to unpick this, even so we can speculate that online respondents may have been more concerned about the security of sensitive information being transmitted electronically.

Postal respondents generally had the highest rates of item non-response. The subject area question had a particularly high rate of non-response. The higher rates in general were to be expected because postal questionnaires rely on the respondent's understanding and motivation to complete it fully and correctly. The employer location question was an exception to this (this may have been because respondents had the opportunity to look up the information requested). The version of the postal questionnaire did not appear to affect the levels of item non-response in different ways, except for the income question.

3.3.2 Item non-response in longitudinal data

In addition to levels of item non-response in the Longitudinal Survey, we looked at aggregate levels of item non-response to key variables related to the graduates' career progression in both surveys. These variables were; SIC, SOC, income and location of employment. The analysis is focussed on a sub-set of respondents who

were in employment at both the Early and Longitudinal Survey. The amount of item non-response to both surveys effects the longitudinal analyses that can be carried out on the data[17]. It is likely that the respondents who had high levels of item non-response in the Early Survey will also have higher levels in the Longitudinal Survey. The level of item non-response for the Early Survey was low, less than 1% for most questions. Therefore the amount of item non-response seen for most questions was so small that it would not have a significant impact on most analyses that combine data from both surveys.

There were 22,269 respondents who were employed on the reference date in the Longitudinal Survey. Of these, 18,005 had also been in employment in the Early Survey. The analysis here concentrates on these graduates.

We compared the levels of item non-response for SIC and SOC codes for the 18,005 graduates and found only a tiny proportion with missing codes at both waves. The number of missing codes for SIC and SOC for the Early Survey in general was very low; 44 cases for SIC and 22 for SOC. The number missing codes at both waves was even smaller; only 6 cases that were missing a SIC code at both surveys and 2 were missing SOC codes at both surveys. This left 16,178 graduates with SOC codes at both waves and 17,140 with SIC codes at both waves, both are large enough sample sizes on which to analyse career progression in some detail.

We also looked at information on employment at country level – required for analysis of cross-country movement of graduates in work. Information on employment location at country level was only missing for 19 cases out of the 18,005 graduates in employment at both surveys. However, the levels of item non-response at the Longitudinal Survey were higher, nearly 13% for this particular sub-set of graduates. This leaves 15,729 graduates with employer locality information at both surveys.

The final variable included in the analysis was income. The number of graduates with missing information on income at the Early Survey was higher than for the previous variables, with 11% of this group of graduates missing income at the Early Survey. There is also a relatively high proportion of item non-response at the Longitudinal Survey; 14% for this particular sub-set. This leaves 7,898 graduates with a valid income at both surveys.

The majority of item non-response is at the Longitudinal Survey. For the data on SIC, SOC and employment location the amount of item non-response at the Early Survey was sufficiently small that it would not have a significant impact on any analysis that combined data from both surveys. The amount of missing income is higher but the sample size of graduates with income information is still relatively high at nearly 8,000. We are unable to say whether the sub-group who recorded incomes at both surveys are a biased group however.

3.4 Evidence of routing errors and missed sections

A problem that is more serious than a respondent missing an item on a questionnaire is when whole sections have been omitted. This can be deliberate or accidental. Routing errors occur when respondents either skip sections of a questionnaire that they are eligible for or answer sections they should exclude. Not all respondents

[17] It also has implications for the suggested non-response weights, as it effects the amount of information available on non-respondents.

were required to fill in all sections of the questionnaire. We found that routing errors were higher for the postal questionnaires, this is expected since the routing is not automated and there is no interviewer to give guidance.

The sections of the questionnaire and respondent eligibility are set out in Table 3-7, below. There were two sections, B and C, where only a sub-set of respondents were required to give answers. In addition, Section D contained a filter question; respondents who answered 'no' to that question were not obliged to complete the rest of the section.

Table 3-7 The longitudinal questionnaire: sections, topics and eligibility.

Section	Topic	Eligibility
Section A	Respondents asked to indicate what activities they were doing on the reference date	All respondents
Section B	Respondent's employment activity on the reference date	All respondents in work on the reference date, even if this is not their main activity
Section C	Study, research or training on the reference date	All respondents engaged on study or training on the reference date, even if it is not their main activity
Section D	Other qualifications	All respondents Q25 Remainder of the section = All respondents with other qualifications.
Section E	Snapshot questionnaire – a set of 4 questions about previous activities	All respondents
	Long questionnaire – activity history	All respondents
Section F	Satisfaction with original course	All respondents
Section G	Respondent contact details	All respondents

Section B

Respondents who were doing any work, either full-time or part-time, on the reference date were asked to give details about their employment in Section B. Of the 22,269 eligible respondents, less than 1% of these respondents missed the entire section and did not complete any questions in Section B (94). Although the proportion missing the section was under 1% overall, the majority of missed sections were found in the postal responses.

Section C

Respondents who were engaged in any study, research or training on the reference date were eligible for Section C. There were 4,017 respondents eligible for Section C, around 3% of these did not complete any of the questions in Section C (125). Again, this was more prevalent in the postal responses (4% for the snapshot version and 6% for the longer version) compared to the other data collection modes (less than 1%).

Sections B & C

Findings from the Longitudinal Survey Pilot (conducted by IPSOS MORI, subsequently referred to as the Longitudinal Pilot) suggest that some respondents

did not read the instructions properly and only filled in sections relating to their main activity. Graduates were asked to complete Sections B and C if they were doing any work or study, regardless of whether it was their main activity or not. If the same type of omissions were also being made in the Longitudinal Survey we would expect to see respondents who have missed the employment section (Section B) to have had a main activity that was not employment, and missed Section C if their main activity was not study/training.

Section C was not completed by 125 of the 4,017 respondents who were eligible. Of these, 86% had said employment was their main activity and had already completed Section B. This implies respondents may not have realised they needed to complete both sections and had simply filled in the first relevant one.

However, the pattern is different for Section B. 94 out of 22,269 eligible respondents did not complete Section B. Of these, 57% had said employment was their main activity and none were eligible for Section C. This suggests the routing error for these respondents is for different reasons.

There was only one person who did not fill out either Section when they were eligible for both.

The proportion of eligible respondents missing Section C was higher than the proportion missing Section B. It is possible that the employment section was somehow more salient to respondents, or was in some way easier to fill in. A difference this large was not found in the Longitudinal Pilot, which suggests further investigation may be required.

Section D

The proportion of eligible respondents missing other sections was small (less than 1%). In Section D there were 9,299 respondents who said they had gained other qualifications since graduating, all went on to fill in some of the remaining section.

Section E

Section E asked respondents to record the activities they had been engaged in since graduation. The activity history is looked at in more detail in Chapter 5.

Section F

All respondents were eligible for Section F, this section was short and was skipped by less than 1% of the sample. Again, this was more prevalent in the postal responses (3% for the snapshot version and 5% for the longer version) compared to the other data collection modes (less than 1%).

4 SUB-GROUPS

This section examines sub-groups that are of specific interest to HESA and its Statutory Customers. Its purpose is to assess the level of dis-aggregation, and by implication analysis, the cell counts in certain sub-groups will allow. The precision of survey estimates based on all respondents to the Longitudinal Survey has already been considered in Section 3.1. However, the sample sizes, and hence the precision, are reduced for analyses that involve sub-groups which need to be examined explicitly. The sub-groups are:

- UK administration in which the graduate lived before HE
- What body funded the original course (i.e. RCs, TDA, DH)
- Original course subject
- Class of original degree
- Mode of study for original course
- Level of study for original course
- PhD graduates
- Sex
- Age at time of Longitudinal Survey
- Ethnicity
- Border flows.

Our examination shows that many of the key sub-groups identified above are of sufficient size to conduct analysis upon them. The over-sampling in the sample design has increased the sample size of some of the sub-groups to make them large enough to analyse. However it has also reduced the effective sample size of other groups such that analysis of some sub-groups will be restricted. The efficiency of the whole sample was 41%, sub-group efficiency ranges from 33% to 100%.

Many of the sub-groups in the Longitudinal Survey have an effective sample size of at least 1,000, meaning differences around 5% or more are needed between two estimates before the results can be interpreted with confidence. The sub-groups to be cautious about are: Ph.D. graduates, graduates from particular course subjects (D, J and T), graduates who were doing voluntary/unpaid work at the Early Survey, and border flows (particularly between Scotland and Wales). There are additional sub-groups to be cautious about when analysing the responses to questionnaire Section C.

4.1 Methods used

The same methods used in Section 3.1 will be applied to the sub-group analysis; we will generate sample errors and confidence intervals for estimates at the sub-group level and use the results to evaluate the effects of the sample design on the sub-group sample sizes. The effective sample size will also be generated for each sub-group and used to comment on the levels of precision the sub-groups analysis will allow. Smaller sample sizes mean a reduction in precision; as sample sizes decrease, standard errors and confidence intervals tend to increase. In addition, the sample sizes will be reduced further for certain topics because of the routing in the questionnaire (section 4.2.2).

Power calculations will also be used to calculate the statistical power of a test for a given sample size (section 4.3). This shows how large the difference between two

estimates would need to be before we are able to state that that difference is genuine and not there by chance. This indicates how large a sub-group needs to be before statistical tests can be carried out on the data

4.2 Results of sub-group analysis

To recap, the effective sample size is the size a simple random sample would need to be to give estimates of the same precision as the complex design in question. A sample is said to be efficient if the effective sample size is close to the actual sample size. The measure of efficiency is the ratio of the effective sample size to the actual sample size (given as a percentage). The estimates calculated on the entire Longitudinal Survey sample had a good level of precision despite the low level of efficiency (41%).

Table 4-1 shows the actual sample size, effective sample size and efficiency for key sub-groups of respondents to the Longitudinal Survey. The table shows that the efficiency of the sub-group samples varies; from 33% for PhD graduates to 100% for groups of graduates from Black, Mixed and Other ethnic backgrounds.

The level of sample efficiency is higher for sub-groups based on variables used to stratify the Longitudinal Survey sample. These are; ethnic group of graduate, country of HEI, domicile, foundation degree graduates and main source of funding. For these particular sub-groups the selection probabilities, and therefore the size of the selection weights, do not vary greatly within the sub-group. This makes the sub-groups more efficient. The exception to this is the white ethnic category, this group cut across other stratifiers and, as a result, has more variable weights. It has a sample efficiency of 47%. However, this group also has a relatively large effective sample size of 8,409; meaning estimates from this group will still have a reasonable level of precision.

The remainder of the sub-groups have a level of precision similar to that of the overall Longitudinal Survey sample. This is because each of the groups contains variable weights. Care must be taken when looking at sub-groups not related to the stratification variables listed above. If the effective sample size is small, and precision is low, there will be less certainty about the estimates. The standard errors will be larger and the confidence intervals wider, hence we can be less certain that the survey estimate accurately reflects the true population value.

If sample sizes are small then only the most basic analyses can be carried out. This would mean analysing single variables separately (by generating simple frequencies and averages), rather than groups of variables (such as cross tabulations or regressions). Cross tabulating two variables with a small sample would result in tables with small cell sizes. Groups that are too small for all but the most basic estimates include the PhD graduates (effective sample size of 208), and some of the course subjects and activity categories, (D, J and T and the Voluntary/Unpaid Work activity group, all have effective sample sizes of less than 100 graduates). Caution should be exercised even for basic analyses (such as frequencies of question responses). Small sample sizes are also unsuitable for statistical tests, since we would not be able to confidently interpret the results. This is discussed further in Section 4.3.

Where there is uncertainty it may be possible to make the sample sizes larger by merging categories. For example, the 51+ years category for age group could be

merged with the 41-50 years category. This is not possible for all variables because there may not be a sensible category to merge smaller categories with.

Table 4-1 Sample size and design effects for respondent sub-groups in the Longitudinal Survey

Sub-group	Unweighted sample size	Deft	Deff	Efficiency	Effective sample size
Sex					
Male	8867	1.55	2.41	41%	3678
Female	15956	1.58	2.49	40%	6408
Ethnicity					
Asian	2748	1.00	1.01	99%	2728
Black	2528	1.00	1.00	100%	2528
Mixed	457	1.00	1.00	100%	457
Other	1257	1.00	1.00	100%	1257
White	17833	1.46	2.12	47%	8409
Domicile					
England	16611	1.52	2.30	43%	7224
Northern Ireland	2968	1.00	1.00	100%	2964
Scotland	2708	1.01	1.03	97%	2640
Wales	2536	1.01	1.01	99%	2510
Country					
England	16887	1.53	2.35	42%	7174
Northern Ireland	2558	1.00	1.00	100%	2556
Scotland	3018	1.02	1.05	95%	2876
Wales	2360	1.03	1.07	93%	2203
Age at Longitudinal Study (grouped)					
25 or under	8814	1.53	2.35	43%	3755
26-30	7724	1.60	2.55	39%	3031
31-40	3695	1.60	2.55	39%	1451
41-50	3273	1.53	2.34	43%	1398
51+	1308	1.49	2.23	45%	586
PhD graduates	637	1.75	3.06	33%	208
Course subject					
A Medicine and Dentistry	589	1.14	1.31	76%	450
B Subjects allied to Medicine	3453	1.16	1.34	75%	2576
C Biological Sciences	1695	1.58	2.50	40%	679
D Veterinary Sciences, Agriculture and related subjects	190	1.42	2.01	50%	95
F Physical Sciences	903	1.57	2.47	40%	365
G Mathematical and Computer Sciences	1957	1.64	2.70	37%	725
H Engineering	1094	1.61	2.60	38%	420
J Technologies	120	1.61	2.59	38%	46
K Architecture, Building and Planning	418	1.63	2.67	38%	157
L Social studies	1920	1.56	2.44	41%	788
M Law	762	1.64	2.70	37%	282
N Business and Administrative studies	2955	1.62	2.63	38%	1123
P Mass Communications and Documentation	409	1.64	2.68	37%	153
Q Linguistics, Classics and related subjects	854	1.59	2.51	40%	340

R European Languages, Literature and related subjects	268	1.59	2.54	39%	105
T Eastern, Asiatic, African, American and Australasian Languages	72	1.47	2.15	47%	34
V Historical and Philosophical studies	871	1.54	2.37	42%	367
W Creative Arts and Design	1586	1.59	2.52	40%	630
X Education	1978	1.58	2.51	40%	789
Z joint honours - various subjects	2144	1.59	2.52	40%	850
Mode of study					
Full time	17713	1.58	2.49	40%	7110
Sandwich	1502	1.62	2.62	38%	574
Part time	4676	1.58	2.51	40%	1866
Writing up	932	1.53	2.35	43%	397
Class of degree					
First class honours	1447	1.48	2.19	46%	659
Upper second honours	6912	1.68	2.84	35%	2437
Lower second honours	4609	1.60	2.56	39%	1804
Third class honours	825	1.60	2.55	39%	324
Unclassified	1269	1.60	2.55	39%	497
Activity at Early Survey					
Full-time paid work only (including self-employed)	14286	1.57	2.45	41%	5823
Part-time paid work only	2082	1.57	2.46	41%	847
Voluntary/unpaid work only	165	1.61	2.60	38%	63
Work and further study	2757	1.58	2.50	40%	1102
Further study only	3208	1.61	2.61	38%	1230
Assumed to be unemployed	1291	1.57	2.45	41%	527
Not available for employment	831	1.70	2.88	35%	289
Other	203	1.64	2.70	37%	75
Main source of funding					
TDA	3110	1.02	1.03	97%	3013
Health: DH/NHS/Social care	3379	1.15	1.32	76%	2562
RC	653	1.00	1.00	100%	653
Foundation degree	300	1.00	1.00	100%	300

4.2.1 Sample sizes available for cross-border analysis

In addition to the sub-groups set out above HESA and it's Statutory Customers also required information on the sample sizes available for the analysis of graduate flows across country borders. The majority of graduates were domiciled and studied in the same country (87%); of all graduates in work at the Early Survey the majority were domiciled, studied and worked in the same country (83%). The sample sizes for cross border analyses are small because they are based on the remaining minority (13% or 17%). For analyses of graduates in work, many groups will be too small for rigorous analysis and care needs to be taken when interpreting the results of even basic analyses.

Tables 4-2, 4-3 and 4-4 show the sample sizes available for the movement of graduates between countries for (i) domicile and HEI location, (ii) domicile and work location in the Early Survey and (iii) HEI location and location of work in the Early Survey, respectively. These tables show that some groups are too small for any

analyses. Effective sample sizes and sample efficiency have been generated for the larger sub-groups.

Since country of HEI and domicile were both used to stratify the sample the efficiency of the smaller countries is generally good. However, the small sample sizes for some flows, for instance from England domicile to Northern Ireland HEI, are too small for any substantial analysis and would only support basic estimates, such as simple frequencies and cross tabulations. Some are even too small for basic estimates, for instance from Scotland domicile to Northern Ireland HEI.

Table 4-2 Country of HEI sample sizes by domicile

Country of HEI	Unweighted sample size	Deft	Deff	Efficiency	Effective sample size
England domicile					
England HEI	15250	1.51	2.29	44%	6662
Northern Ireland HEI	176	1.00	1.00	100%	176
Scotland HEI	459	1.03	1.06	94%	432
Wales HEI	726	1.02	1.05	96%	694
Northern Ireland domicile					
England HEI	381	1.00	1.00	100%	379
Northern Ireland HEI	2371	1.00	1.00	100%	2369
Scotland HEI	206	1.00	1.00	100%	206
Wales HEI	10	-	-	-	-
Scotland domicile					
England HEI	353	1.03	1.06	94%	333
Northern Ireland HEI	6	-	-	-	-
Scotland HEI	2342	1.01	1.02	98%	2297
Wales HEI	7	-	-	-	-
Wales domicile					
England	903	1.01	1.01	99%	890
Northern Ireland	5	-	-	-	-
Scotland	11	-	-	-	-
Wales	1617	1.00	1.01	99%	1605

Base: all respondents

Table 4-3and Table 4-4 show location of work in the Early Survey by country of domicile and HEI, respectively. The sample sizes for these flows are generally smaller than those in Table 4-2 because not all graduates responding to the Longitudinal Survey were in work at the Early Survey. The efficiency for the flows involving England are also lower because location of work was not a stratification variable. The implication of this is that the effective sample sizes are less than 100 for many of the flow groups. We urge caution for any type of analysis based on these small groups.

Table 4-3 Work location sample size by domicile

Work location	Unweighted sample size	Deft	Deff	Efficiency	Effective sample size
England domicile					
England	12115	1.51	2.27	44%	5328
Northern Ireland	31	-	-	-	-
Scotland	110	1.34	1.79	56%	61
Wales	167	1.36	1.85	54%	91
Northern Ireland domicile					
England	201	1.00	1.00	100%	200
Northern Ireland	1920	1.00	1.00	100%	1919
Scotland	67	1.00	1.00	100%	67
Wales	7	-	-	-	-
Scotland domicile					
England	199	1.03	1.07	94%	187
Northern Ireland	4	-	-	-	-
Scotland	1840	1.01	1.02	98%	1804
Wales	3	-	-	-	-
Wales domicile					
England	447	1.01	1.02	98%	439
Northern Ireland	2	-	-	-	-
Scotland	6	-	-	-	-
Wales	1505	1.00	1.01	99%	1494

Base: All respondents in employment at Early Survey

Table 4-4 Work location sample size per country of HEI

Work location	Unweighted sample size	Deft	Deff	Efficiency	Effective sample size
England HEI					
England	12106	1.52	2.31	43%	5252
Northern Ireland	165	1.29	1.66	60%	99
Scotland	186	1.26	1.60	63%	117
Wales	332	1.51	2.27	44%	146
Northern Ireland HEI					
England	96	1.00	1.00	100%	96
Northern Ireland	1720	1.00	1.00	100%	1719
Scotland	12	-	-	-	-
Wales	6	-	-	-	-
Scotland HEI					
England	287	1.04	1.09	92%	264
Northern Ireland	70	1.00	1.01	99%	69
Scotland	1818	1.01	1.03	97%	1766
Wales	10	-	-	-	-
Wales HEI					
England	473	1.03	1.07	94%	444
Northern Ireland	2	-	-	-	-
Scotland	7	-	-	-	-
Wales	1334	1.02	1.03	97%	1289

Base: All respondents in employment at Early Survey

4.2.2 Routed sections

We now turn our attention to the sections that were routed, that is, where only certain respondents were required to provide responses. The routing reduces the sub-group sample sizes further.

Table 4-5 shows the sample sizes within each sub-group for questionnaire Sections B and C. Respondents who were in paid employment on the reference date were eligible for Section B. Respondents who were in study or training on the reference date were eligible for Section C. There are not many additional sub-groups to be cautious about for analyses of Section B (course subject R and those in an 'other' activity at the Early Survey). However there are many more sub-groups with effective sample sizes of less than 100 for analyses of Section C (relating to course subject, mode of study, class of degree, activity at Early Survey and those doing Foundation degrees). This is not unexpected given the smaller respondent sample size for Section C.

Table 4-5 Weighted, unweighted and effective sample sizes for routed sections of the questionnaire, by sub-group

Sub-group	Employed (Section B)		In study/training (Section C)	
	Unweighted sample size	Effective sample size	Unweighted sample size	Effective sample size
Sex				
Male	7886	3271	1329	551
Female	14383	5776	2698	1084
Ethnicity				
Asian	2400	2382	438	435
Black	2193	2193	490	490
Mixed	385	385	84	84
Other	1055	1055	232	232
White	16236	7656	2783	1312
Domicile				
England	14860	6463	2676	1164
Northern Ireland	2685	2682	508	507
Scotland	2447	2385	416	405
Wales	2277	2254	427	423
Country				
England	15101	6416	2682	1139
Northern Ireland	2310	2308	446	446
Scotland	2713	2585	501	477
Wales	2145	2002	398	372
Age at Longitudinal Study (grouped)				
25 or under	7720	3289	1560	665
26-30	6990	2743	1166	458
31-40	3376	1326	521	205
41-50	3071	1312	549	234
51+	1103	495	228	102
PhD graduates	592	193	65	21

Course subject				
A Medicine and Dentistry	557	425	126	96
B Subjects allied to Medicine	3174	2368	717	535
C Biological Sciences	1390	557	432	173
D Veterinary Sciences, Agriculture and related subjects	169	84	31	15
F Physical Sciences	753	305	202	82
G Mathematical and Computer Sciences	1758	651	241	89
H Engineering	986	379	146	56
J Technologies	107	41	17	7
K Architecture, Building and Planning	380	142	67	25
L Social studies	1746	717	364	149
M Law	680	251	125	46
N Business and Administrative studies	2735	1039	391	149
P Mass Communications and Documentation	374	140	42	16
Q Linguistics, Classics and related subjects	758	302	114	45
R European Languages, Literature and related subjects	239	94	36	14
T Eastern, Asiatic, African, American and Australasian Languages	59	27	14	7
V Historical and Philosophical studies	734	309	152	64
W Creative Arts and Design	1401	557	169	67
X Education	1859	742	224	89
Z joint honours - various subjects	1912	758	359	142
Mode of study				
Full time	15742	6319	2905	1166
Sandwich	1360	520	252	96
Part time	4311	1720	746	298
Writing up	856	365	124	53
Class of degree				
First class honours	1178	537	393	179
Upper second honours	6103	2152	1255	443
Lower second honours	4106	1607	693	271
Third class honours	717	281	112	44
Unclassified	1179	462	191	75
Activity at Early Survey				
Full-time paid work only (including self-employed)	13446	5481	1850	754
Part-time paid work only	1864	759	296	120
Voluntary/unpaid work only	133	51	27	10
Work and further study	2562	1024	633	253
Further study only	2503	960	846	325
Assumed to be unemployed	1025	418	186	76
Not available for employment	603	210	150	52
Other	133	49	39	14
Main source of funding				
TDA	2982	2889	268	260
Health: DH/NHS/Social care	3102	2352	742	563
RC	577	577	122	122
Foundation degree	268	268	40	40

Bases: all respondents in employment on reference date, all respondents studying/training on reference date

4.3 Statistical power

Analysts may wish to test the difference between two estimates to allow them to comment on how significant, or likely, the difference is. Power calculations estimate the statistical power of a test for a given sample size. This shows how large the difference between two estimates would need to be before we are able to state that the difference is statistically significant, i.e. it is genuine and not there by chance.

Table 4-6 shows the different sample sizes required to test for different levels of precision. The power of a statistical test varies by sample size, the difference in size of the two estimates being tested and how close the mid-point of the two estimates is to 50%. Larger sample sizes are required to test smaller differences between two estimates. The sample sizes given in the table are for each sub-group and are the effective sample size rather than the unweighted sample size. For example, we would need two sub-groups of 916 graduates to test whether the difference between an estimate of 27% and an estimate of 33% was statistically significant (i.e. a difference in estimates of 6% or more was and where the mid–point of the two estimates is 30%). If the two sub-groups are uneven in size then the size of the smallest sub-group should be used as a guide.

Table 4-6 Effective sample sizes required for different levels of precision

	Mid-point of the two estimates		
	10% or 90%	30% or 70%	50%
Level of precision			
+ or - 1%	14,129	32,966	39,246
+ or - 2%	3,532	8,242	9,812
+ or - 3%	1,570	3,663	4,361
+ or - 4%	466	2,061	2,453
+ or - 5%	299	1,319	1,570
+ or - 6%	393	916	1,090
+ or - 7%	289	673	801
+ or - 8%	221	515	613
+ or - 9%	175	407	485
+ or - 10%	142	330	393

Tests used a 95% confidence level and 80% power.

Many of the sub-groups in the Longitudinal Survey have an effective sample size of at least 1,000, meaning differences around 5% or more are needed between two estimates before the results can be interpreted with confidence.

5 ACTIVITY HISTORIES

In this chapter we evaluate the information recorded about the activities respondents had been engaged in since graduation. The activity history is Section E of the questionnaire, covering activities between graduation and the activity being carried out on the reference date respondents recorded in Section A. Two versions of questions were tried in Section E. First of all we describe how the activity histories were collected to be able to comment on how many respondents provided data (5.1 and 5.2) and whether they are representative of the population (5.3). The quality of the data was explored to understand whether respondents missed out the activity history section (5.4), whether the data collection modes affected the data (5.5), the extent of item non-response (5.6 and 5.7), the consistency of the chronology of activities (5.8), and whether the two versions produce different findings (5.9). Our conclusions about the value of the two versions of activity histories are presented in section 5.10.

All analysis has been conducted on unweighted data.

5.1 Questionnaire versions sent out

Activity histories were collected in two different ways. In the 'snapshot' version Section E contained a set of four questions asking the respondent to count all periods of work and periods of unemployment lasting more than a month. The 'long' version asked for a detailed career history; collecting details of up to eight activities, each lasting more than a month.

The versions varied between and within data collection mode. The postal questionnaires contained either the snapshot or long version of Section E. One of the two versions was allocated to graduates at random. 10% of the sample sent a postal questionnaire were sent the long version of the questionnaire.[18] The online and telephone questionnaires contained both versions, the short summary questions followed by the longer career history.

5.2 Questionnaire versions returned

Table 5-1 gives the sample sizes of the different versions of the questionnaire returned by respondents.

Table 5-1 Sample size of returned questionnaires by version

Version	Number of responses	Percentage of total
Long & snapshot – online	2,462	10%
Long & snapshot – telephone	11,418	46%
Long – postal	925	4%
Snapshot – postal	10,018	40%
Total	24,823	100%

Base: All respondents.

[18] 54,652 postal questionnaires sent out; 5,465 long and 49,187 snapshot.

The long version of the activity history was returned by 60% of respondents (14,805). Most of these were online and telephone questionnaires that had contained both versions. The snapshot version was returned by almost all respondents (96%). Just over half of these were online or telephone respondents. The respondent sample is likely to contain a large enough sample size to conduct analyses on both versions of the activity history data collected.

5.3 Non-response and non-response bias

Response rates for the different versions of the questionnaire and modes have not been calculated. This is because they are difficult to calculate because the sample were often contacted by and responded to different modes. For example, responses were received from 21,895 of the 49,187 graduates sent the snapshot postal questionnaire, however only half responded using the postal questionnaire. The other half responded via the online or telephone where the long version of the activity history had been completed. The interpretation of any response rates calculated would be difficult. It had been anticipated that the long version of the questionnaire might have generated a lower response rate. Response for postal questionnaires tends to be better when the questionnaire is short.

The Longitudinal Survey technical report contains an analysis of response to the long and snapshot versions of the postal questionnaire by a number of key variables. In general the results indicate that most graduates were more likely to respond to the snapshot version than the long version. PhD students were the only sub-group more likely to respond to the longer version. The technical report showed some differences in response by age and sex. There were no consistent patterns for ethnicity or degree course.

The profile of postal respondents completing the snapshot and long versions was compared as a further examination of possible non-response bias. The profiles are broadly similar, although there are some differences. There is no strong differential bias suggested by the characteristics available.

5.4 Sample size – long activity history

The activity history was placed near the end of the questionnaire. It was a cognitively demanding section, asking detailed information on all activities carried out between graduation and the reference date. Respondents who said they were engaged in a single activity during this time period were not obliged to fill in the activity history. Table 5-2 shows the proportion of respondents completing the activity section by questionnaire mode. Overall, 70% of respondents completed Section E. Graduates responding via the online were less likely to state that they had been doing a single activity from graduation to the reference date and subsequently a higher percentage completed the activity history (81% compared to 68% and 67%).

The effective sample size will allow some basic analyses to be conducted (assuming an efficiency of 41%). Some sub-group analyses will be restricted, as will statistical testing unless the differences are relatively large (see Table 4-6).

Table 5-2 Section E long activity history – proportion completed

	Telephone	Online	Postal – long	Total
All respondents	11,418	2,462	925	14,805
Respondents doing one activity since graduation	3,620	459	304	4,383
Respondents completing Section E	7,798	2,003	621	10,422
Proportion of respondents completing Section E	68%	81%	67%	70%

Base: All respondents returning long version questionnaire

Those not completing Section E are coded in the data as doing 'one continuous activity since graduation'. They were investigated further to establish whether there was any evidence that they had been engaged in more than one activity and therefore should have completed the activity history. A comparison of the main activities recorded at the Early and Longitudinal Surveys showed that this activity had changed for 32% of respondents. This suggests that some respondents may have been eligible to complete the activity history although the discrepancy may be due to error-prone memories or mistakes made recording their activity.

A second comparison was made using the responses from the online and telephone questionnaires. In these questionnaires the summary variables captured in the snapshot version can be compared to the long activity history variables. 19% of the online and 28% of the telephone respondents who are coded as having carried out a single activity since graduation in the long version also reported having two or more jobs in the snapshot version. This is a strong indication that the coding of respondents doing one continuous activity may be flawed, over-estimating the number of graduates doing only one activity.

There is a knock-on effect for the calculation of non-response to this section of the questionnaire. While the data state that 30% of respondents did not complete Section E because they were doing one activity throughout the period, a more realistic estimate may be 22%.[19] This implies that around 8% may have missed out the section.

5.5 Data collection mode – long activity history

In section 3.2 we noted that the activity history may be subject to mode effects. This is where the same question asked in different modes produces different results. Time restrictions prevented a full analysis, however a comparison of the number of activities recorded in each mode is given as an indicator of mode effects.

Table 5-3 summarises the differences in the number of activity spells recorded by different questionnaire modes. The summary statistics show similar numbers of activities being recorded in the online and telephone questionnaires. Postal respondents recorded slightly more activities on average, with a mean of 2.7

[19] Estimate calculated as [(0.19 x 459) + (0.28 x 3620)] / 13880.

activities compared to 2.0. Therefore the long activity history appears to be noticeably affected by the data collection mode used. It is not clear whether the difference is because the respondents using each mode are genuinely different or whether the mode itself has induced bias.

Table 5-3 Summary of the number of activity spells by questionnaire mode

	Telephone	Online	Postal – long	Total
Total number of activities recorded	15,670	3,941	1,672	21,283
Average number of activities per respondent	2.0	2.0	2.7	2.0
Respondents recording one activity recorded[20]	50%	51%	37%	49%
Respondents recording more than one activity recorded	50%	49%	63%	51%
Bases	7,798	2,003	621	10,422

Base: All respondents returning long version questionnaire

5.5.1 Coding inconsistencies

Some of the coded data are inconsistent across data collection modes. The data may need to be subjected to further editing. Three issues identified are:

- type of HEI was missing for postal questionnaire respondents,
- country of employment/study was missing for postal questionnaire respondents,
- graduates whose main activity is not work have missing SIC and SOC codes; these have different missing values for postal questionnaire respondents and online/telephone respondents which has the potential to mislead analysts.

5.6 Item non-response – snapshot activity summary

A comparison of item non-response was carried out by the data collection organisation and can be found in the Longitudinal Survey Technical Report. Their findings are summarised here. The interpretation is our own.

The summary variables collected in the snapshot version were conducted by online, postal and telephone modes. Item non-response was very low for the telephone and online modes (less than 1%, except for the fourth question where item non-response was 2% for the telephone questionnaire). Item non-response varied by question for postal respondents: question 1 = 3%, question 2 = 2%, question 3 = 6% and question 4 = 23%. The last two questions relate to spells of unemployment which are known to be particularly prone to under-reporting. However the level of item non-response for the last question is surprisingly high (the last question asked respondents to give the total number of months spent in periods of unemployment lasting more than one month).

[20] Not to be confused with the respondents engaged in one continuous activity since graduation. These respondents engaged in two activities in the period between graduation and the reference date, one recorded in the activity history and the other as their activity on the reference date recorded in Section A

5.7 Item non-response – long activity history

We have conducted our own analysis of item non-response in the long version of Section E (none found in the Longitudinal Survey Technical Report). As with our earlier analyses (see section 3.3) we concentrate on refusals and skipped questions, excluding 'don't know' answers.

We would expect higher levels of item non-response for the activity history because it has a high cognitive burden; a lot of detail was requested from respondents for events that happened some years previous. In addition the activity history is at the end of the questionnaire and respondents may have become fatigued. Full details of the analysis can be found in APPENDIX F, a summary is provided here.

For each activity the respondent was required to enter the date they started and ended (month and year). Complete dates - month and year for both dates recorded – are available for 17,745 activities (83% of all activities).

Some key questions were expected to have higher levels of item non-response than others (similar to those identified in other sections of the questionnaire, see section 3.3.1). The levels of item non-response for questions associated with employment, (income, employer location, industry and occupation codes) are relatively high, and are higher than the equivalent questions in other questionnaire sections. With the exception of employer location, the highest levels of item non-response came from the postal questionnaire. Item non-response for the subject of study was lower in the activity histories than in Section C. There is a very high level of missing cases in the postal questionnaire although the sample size is relatively small which reduced its effect on the total level of item non-response. The item non-response for the spells of activity recorded later is generally higher (with the exception of income), suggesting that respondents may have grown tired of filling in their details.

The level of item non-response for other questions in the activity history is substantially lower, less than 3%.

5.8 Date consistency – long activity history

When respondents recorded more than one activity the start date of the later activity should equal the end date of the immediately preceding activity. This was the case for 91% of applicable cases. It appears that respondents were fairly good at making sure activities were consecutive within the history. However, respondents were less good at ensuring dates in the activity history matched up with dates elsewhere in the questionnaire. For the last activity recorded the end date should equal the start date of the activity recorded in either the employment section of the questionnaire (Section B) or the study/training section (Section C). These two dates matched for only 40% of 'last activities'.

5.9 Comparison of snapshot and long versions

This section looks at the correspondence between the summary variables collected in the snapshot version of the questionnaire with responses given in the long version. Generally, the correspondence between the snapshot and long versions was poor. Analysts will need to be very careful using the data. Given the discrepancies found, analysts will need to think carefully before using the responses from each version simultaneously.

Responses equivalent to the summary questions were derived from the activity history data in the long version, and can be compared directly for the telephone and online respondents. The online and telephone respondents were asked both sets of questions which means that we can be certain that any differences in the estimates from the two sets of questions are due to the questions themselves rather than by any variations in the sample. 9,786 graduates reported both a summary and an activity history (there were a number of respondents who reported jobs in the summary (snapshot version) but did not do so in the activity history (long version)).

Of the 9,786 graduates who gave responses to both versions only 17% reported the same number of jobs in the period between graduation and the reference date in both versions. There was very little difference between the two modes (19% telephone, 17% online). For this question the number reported in the snapshot version was consistently higher than the number derived from the long version (Table 5-4). We speculated that the respondents were incorrectly including their current activity in the snapshot version and created a new summary variable for comparison (original minus one). We cannot be certain that this is what actually occurred, however the proportion of respondents giving the same number increases to 51% for the new variable (49% telephone, 52% online). Even this increased correspondence is less than ideal.

Table 5-4 Number of jobs, snapshot and long version

How many jobs did you have between graduating in 2002/03 and 27th November 2006?	Sample size	Minimum	Maximum	Mean	Median
Long version (derived)	9,801	0	8	1.43	1
Snapshot version	9,786	0	64	2.73	2
Snapshot version, minus one	9,786	0	63	1.75	2

Base: All telephone and online respondents answering the relevant question in the activity history and the activity summary.

The second question in the snapshot version summary asked respondents to report whether they had had any spells of unemployment lasting longer than one month (Table 5-5).

Table 5-5 Ever unemployed, snapshot and long version

		Long version		
		Yes	No	Total
Snapshot version	Yes	**1,742**	1,676	3,418
	No	223	**6,150**	6,373
	Total	1,965	7,826	9,791

Base: All telephone and online respondents answering the relevant question in the activity history and the activity summary.

The same information was obtained from both versions for 80% of respondents. Nearly half of the respondents who said they had had a period of unemployment lasting at least a month in the snapshot version did not report the spell in the long version. The snapshot version estimates the proportion of graduates who had experienced a period of employment for at least one month as 35%, whereas the

estimate from the long version is 20%. Whilst unemployment spells are known to be particularly prone to being under-reported it is not clear whether the long version is underestimating unemployment or the snapshot version is an over-estimate (perhaps including spells shorter than a month).

The final two questions in the snapshot version ask about the number of spells of unemployment and the total amount of time (in months) spent unemployed. There were also discrepancies between versions for these two measures. The correspondence between the two versions for the number of periods spent unemployed was poor. Only a third of respondents gave the same answer. Respondents using the snapshot version gave a higher number of spells. Of the respondents who had been unemployed, 32% had two or more periods of unemployment according to the snapshot version, whereas the long version had only 15%. In spite of this the mean number of months spent unemployed was very similar for both versions, (snapshot = 6.2 and long = 6.4).

5.10 Conclusions

Most respondents completed the snapshot activity history questions. Therefore the sample size should be large enough to conduct analyses on these responses for the whole sample or sub-groups. More respondents than anticipated completed the long version of the activity history because of the high proportion responding via the telephone. As a result the sample size available is sufficient for interrogation of the activity patterns. The low efficiency of the sample has reduced the effective sample size which means that statistical testing is unlikely to be able to detect anything other than large differences between groups and may not permit many of the sub-groups to be analysed separately.

Whether it was worthwhile to use the long version compared to the snapshot version depends on the rigour of the analyses expected for the data. The effective sample sizes are likely to be too small for complex modelling and statistical testing but sufficient to give a flavour of the respondent's activity history. The quality of the information collected in the long version was reasonable although the levels of item non-response for key questions were relatively high which reduces the value of the data. There was evidence that respondents grew tired of filling in their details so that later activities contained more gaps. There was some evidence of inconsistencies in the chronology of the recorded activities, but further investigations are required before any comment can be made about the impact of this.

The two versions produce different estimates (for the small number of statistics tested). It is impossible to say which version is closer to the truth because there is no way that the estimates can be validated. From the exploration of online/telephone respondents and our previous experience of collecting similar data, we conclude that the long version may be under-estimating employment and unemployment spells and the snapshot version may be over-estimating spells.

If all postal respondents had been sent the long version of the questionnaire fewer would have responded. Assuming the rate of conversion by telephone would have been the same we estimate that 9,200 postal responses would have been obtained (rather than 10,943). This implies the total achieved sample size would have been 23,080 rather than 24,823, with around 16,000 activity histories completed (and 7,000 claiming one continuous activity throughout the period) compared to 10,422 activity histories available. The large increase in the activity history sample size

comes at the price of the total sample size being reduced. The reduction will reduce the sub-groups that can be analysed, which may not be acceptable to HESA.

We suggest that the snapshot version continues to be collected from all respondents. We recommend that the long version is also continued if the additional analyses possible from the full activity histories that cannot be conducted on the snapshot version are a high priority.

6 RECOMMENDATIONS

We found that the sample selected for the Longitudinal Survey is representative of the population (when weighted). The 24,823 respondents to the Longitudinal Survey are broadly representative of the population although there are some differences due to non-response. 40% of invited graduates responded to the Longitudinal Survey, compared to nearly 75% at the Early Survey. After adjusting for the over-sampling and taking non-response at both surveys into account, the response rate is 31.0%. Our exploration did not show any strong differential bias in the characteristics of the respondents compared to the population. Any bias relating to these characteristics is likely to be marginal however we recommend the calculation and use of non-response weights (discussed below in 6.1). We also recommend that pilot samples are selected in a different way, to eliminate potential bias to the sample (6.1). Improvements need to be made to the information available to assess the response process (discussed in 6.2). If required, improvements could be made to the data collection to increase the response rate (6.3).

The respondent sample size is large enough for the key estimates planned on the full sample. The sample design reduced the efficiency of the sample to 41% (equivalent to an effective sample size of 10,120). However the loss is not so great that it undermines the sample, the large sample size means precision of estimates based on the whole sample are not unduly affected. Before this survey is repeated on another cohort of graduates we recommend that the sub-group analysis requirements are reviewed (discussed in 6.4).

The level of non-response for individual questions was low. Since it was also low for the Early Survey, this should not have a significant impact on most analyses that combine data from both surveys. Levels tended to be higher for more complex questions: income, employer location, Standard Occupation Code and subject of study/training. Item non-response for these questions has reduced the associated sample sizes (particularly for the activity histories).

Overall, the proportion of graduates missing sections was small and should not compromise data quality. The extent that the activity histories were missed by respondents was harder to be sure about because of the coding of ‘one continuous activity’. The proportion of eligible respondents missing out the section of the questionnaire covering study/training activities (Section C) was relatively high compared to other sections. This section is considered in more detail below (6.5). As a result of the relatively small number of graduates who were studying on the date in question, combined with the higher proportion of eligible graduates who missed out these questions (either completely or some of them) this sub-group is quite small and may be difficult to analyse thoroughly.

Our exploration also demonstrated that many of the key sub-groups identified by HESA and their Statutory Customers are of sufficient size to conduct analysis upon them. The over-sampling in the sample design has increased the sample size of some of the sub-groups to make them large enough to analyse. However it has also reduced the effective sample size of other groups such that analysis of some sub-groups will be restricted. Many of the sub-groups have an effective sample size of at least 1,000, meaning differences around 5% or more are needed between two estimates before the results can be interpreted with confidence. The sub-groups to be cautious about are: Ph.D. graduates, graduates from particular course subjects

(D, J and T), graduates who were doing voluntary/unpaid work at the Early Survey, and border flows (particularly between Scotland and Wales). There are additional sub-groups to be cautious about when analysing the responses to Section C and E. Sub-groups are discussed further in section 6.4.

The experimental allocation of two versions of the activity history questions did not compromise the survey. The two versions produce different estimates, neither of which can be validated. The sample sizes obtained for each version are sufficient for analyses of the whole sample. Many sub-groups will be too small to analyse the detailed activity histories separately. The quality of the data collected in the longer version was reasonable although not without errors. We suggest that the snapshot version continues to be collected from all respondents. We recommend that the long version is also continued if the additional analyses possible from the full activity histories that cannot be conducted on the snapshot version are a high priority.

We recommend that the effects from data collection modes are monitored. Any effect on responses is likely to be small. A robust analysis of the mode effects on responses could not be conducted in the time given. Item non-response rates and the prevalence of whole sections being missed differed by mode. We believe that the increased response obtained from using a variety of data collection modes in sequence outweighs the possible bias introduced. Given the time elapsed between the two surveys and difficulty ascertaining the validity of the contact details, response is improved if all forms of contact details can be used in conjunction with multiple attempts to contact graduates.

If a third survey were to be carried out, inviting participation from respondents to the Longitudinal Survey, the resulting sample size is likely to be much smaller and bias from non-response is likely to increase.

6.1 Sample bias

Differences in the characteristics of the respondents (compared to the population) can be addressed using weighting. The Early Survey and Student Record contain information on the respondents and non-respondents to the Longitudinal Survey. This information can be used to model the likelihood of responding to produce a non-response weight. The weights can reduce the differences in the characteristics included in the model. We strongly recommend that non-response weights are produced and applied to the data for all analyses. A strategy has already been suggested for the remaining reports to be written by NatCen.

During the development phase of the DLHE Longitudinal Study a pilot was conducted to establish the feasibility of the main survey. The pilot sample was selected from the pool of 2002/3 HE leavers. Although the exact selection criteria is not known it is possible that it caused selection bias. This bias occurs if the graduates selected for the pilot are different in some way from the remaining graduates that the main sample is subsequently selected from. To eliminate any bias from future surveys we recommend that any pilots are either conducted on samples that have no chance of being considered for selection into the main study (e.g. using 2001/2 HE leavers), or by ensuring that the pilot cases are a randomly selected sample that represents the population.

The sample for the Longitudinal Survey was selected from a pool of ‘qualifiers’ that had responded to the Early Survey. Therefore the sample does not represent any

graduates who did not respond to the Early Survey. Although our analysis showed that this pool of respondents were fairly similar to the population of 2002/3 HE leavers (section 2.2.1) we cannot be absolutely sure that they do not differ in ways that could bias the sample. The rationale for selecting the sample from the pool of respondents is both sensible and practical although it has made it harder to account for the non-response at the Early Survey. We do not recommend any changes.

6.2 Response process information

A greater understanding of the response (and non-response) process can be obtained through an interrogation of the survey outcomes. This was attempted in chapter 2 but was not wholly successful. While respondents can be easily identified from the outcome information non-respondents could not be classified by their reason for non-response very easily. Better outcome information can be built into the requirements for the both surveys. There are standard guidelines that can be used to specify the requirements. Such improvements would enable outcomes to be constructed for response rate calculations with confidence and allow them to be compared to other surveys.

The survey outcome information generated from the Early and Longitudinal Surveys needs to be improved in a number of ways:

- complete and detailed outcomes for non-respondents to the Early Survey should be recorded,
- a single final outcome should be provided for the Longitudinal Survey (usually created by the data collection organisation),
- each stage of the sequential data collection process using multiple modes needs to transparent in the outcome data,
- outcomes should be coded to an accepted UK standard.

6.3 Methods to increase response

The response rate achieved is commendable. Nevertheless, if improvements were required for the Longitudinal Survey there are a variety of tried and tested methods that may increase the overall response rate. The methods fall into two categories, those that focus on improving contact and those that target co-operation after contact has been made. Improving the contact rate is likely to make the most difference to the Longitudinal Survey response rate.

Contact may be improved if graduates are asked to provide a stable address at the Early Survey. A stable address is the name given to contact details offered by the respondent that they believe will enable you to get in touch with them in the future. This can be a longstanding mobile phone or personal email account for example, or a relative's address. Contact may also be improved by getting in touch with graduates on a more regular basis and asking them to provide updated contact details.

Any improvements may also affect the composition of the respondent sample. Therefore the potential effects on non-response bias should be considered if any different measures are implemented.

6.4 Review of sub-group requirements

The sample was designed to accommodate as many of the competing analysis requirements as possible. This was done through the sample design - over-sampling key sub-groups. The over-sampling provided higher sample sizes for some sub-groups although not without reducing the effective sample sizes of other sub-groups. We have demonstrated that the current sample design provided sufficient sample sizes for many sub-groups but not all of them.

The reduced sample size is not something that can be altered in the current data. Nevertheless there are a number of options that might be considered for the future. Each option needs to be considered carefully with the context of the requirements to ascertain whether the likely benefits outweigh any costs. The first option is to increase the size of the selected sample. This would increase costs. Increasing the size of the selected sample would increase the total sample size and increase some sub-group sizes. The existing design already over-samples some sub-groups to the extent that all cases were selected. For these sub-groups selecting a larger number to participate would not have any effect. Assuming this is the case for some of the sub-groups identified as being ‘too small’, this option will not solve the problem completely.

Other options maintain the current size and either improve the overall response rate or alter the design of the over-sampling. Suggestions for improving the response rate have already been described in section 6.3. HESA may like to consider this option because it has the potential to increase the sample size of every sub-group. Our recommended option is to review each sub-group that has been over-sampled. The design of the over-sampling has provided advantages (higher sample sizes for some sub-groups) although not without some disadvantages (reduced effective sample sizes). If there are sub-groups that are currently too small that were not over-sampled then it may be worth considering over-sampling them if this survey is repeated. Any change must be considered carefully because it has the potential to increase the variability of the selection weights further which will reduce the effective sample size of the overall sample and other sub-groups. If there are small sub-groups that were obtained through over-sampling then there may not be any changes to the design that can improve their sample size.

6.5 Section C

It is not completely clear why a relatively high proportion of eligible respondents missed out a question or the whole of Section C (covering study, research and training activities). However there was a marked difference in the questionnaire mode distribution compared to other sections. A higher proportion of respondents had completed the postal questionnaire, and postal responses exhibited much higher rates of item non-response.

Postal respondents generally had the highest rates of item non-response. The subject area question had a particularly high rate of non-response. The higher rates in general were to be expected because postal questionnaires rely on the respondent’s understanding and motivation to complete it fully and correctly. It is possible that they missed the section because they did not realise they needed to complete it after they had already filled in the previous section (B).

Questionnaires can be reviewed in a number ways, an appropriate technique needs to identify how respondents are approaching and understanding both the questions and response categories offered. The amount of missing data maybe reduced by:

- conducting a review of and revising the way the instructions are given on the postal questionnaire if respondents hadn't realised it needed to be completed,
- conducting a review of and revising the questions if respondents found the questions difficult to answer or the response options available were not adequate,
- improved training for telephone interviewers probing for responses, to reduce the proportion of respondents with inadequate information on their subject of study for example.

APPENDIX A SAMPLING FRACTIONS

Sampling fractions applied by HESA when drawing the sample for the Longitudinal Survey.

Apx. Table A.1 Sampling fractions for the Longitudinal Survey

Country of HEI	Domicile of student	Ethnic group	Student type (Healthcare/TDA/ Foundation/RC)	Sampling fraction
England	England/ other UK/ other EU	White/unknown	General	**0.05**
			RC	**1**
			Foundation	**1**
			TDA	**0.35**
			TDA/RC	**1**
			Healthcare	**0.15**
			Healthcare/RC	**1**
			Healthcare/Foundation	**1**
			Healthcare/TDA	**0.36**
		Black	General	**1**
			RC	**1**
			Foundation	**1**
			TDA	**1**
			Healthcare	**1**
			Healthcare/RC	**1**
			Healthcare/Foundation	**1**
			Healthcare/TDA	**1**
		Asian	General	**0.35**
			RC	**1**
			Foundation	**1**
			TDA	**0.35**
			TDA/RC	**1**
			Healthcare	**0.35**
			Healthcare/RC	**1**
			Healthcare/Foundation	**1**
			Healthcare/TDA	**0.35**
		Mixed	General	**1**
			RC	**1**
			Foundation	**1**
			TDA	**1**
			Healthcare	**1**
			Healthcare/Foundation	**1**
		Other	General	**1**
			RC	**1**
			Foundation	**1**
			TDA	**1**
			Healthcare	**1**
			Healthcare/TDA	**1**
	Wales	White/unknown	General	**0.4**
			RC	**1**
			TDA	**0.4**

			Healthcare	**0.4**
			Healthcare/TDA	**0.4**
		Black	General	**1**
			Healthcare	**1**
		Asian	General	**0.39**
			TDA	**0.4**
			Healthcare	**0.41**
		Mixed	General	**1**
			RC	**1**
			TDA	**1**
			Healthcare	**1**
		Other	General	**1**
			Healthcare	**1**
	Scotland	White/unknown	General	**0.25**
			RC	**1**
			Foundation	**1**
			TDA	**0.34**
			Healthcare	**0.25**
			Healthcare/RC	**1**
			Healthcare/TDA	**0.35**
		Black	General	**1**
			Healthcare	**1**
		Asian	General	**0.35**
			TDA	**0.35**
			Healthcare	**0.31**
		Mixed	General	**1**
		Other	General	**1**
			Healthcare	**1**
	NI	White/unknown	General	**0.6**
			RC	**1**
			Foundation	**1**
			TDA	**0.6**
			Healthcare	**0.6**
		Black	General	**1**
			Healthcare	**1**
		Asian	General	**0.56**
			Healthcare	**0.6**
			Healthcare/Foundation	**1**
		Mixed	General	**1**
		Other	General	**1**
			TDA	**1**
			Healthcare	**1**
Wales	England/ other UK/ other EU	White/unknown	General	**0.25**
			RC	**1**
			Foundation	**1**
		Black	General	**1**
		Asian	General	**0.35**
			RC	**1**
		Mixed	General	**1**
			Foundation	**1**
		Other	General	**1**

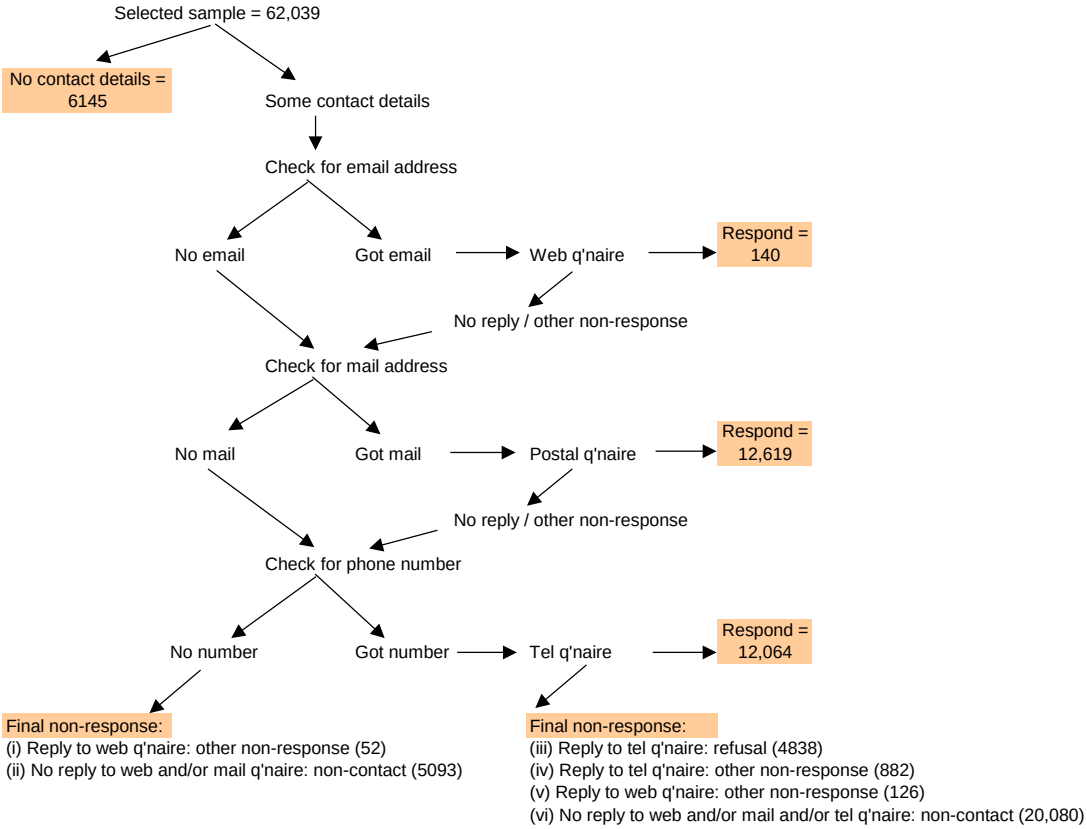
			RC	**1**
	Wales	White/unknown	General	**0.4**
			RC	**1**
			Foundation	**1**
		Black	General	**1**
		Asian	General	**0.4**
			RC	**1**
		Mixed	General	**1**
		Other	General	**1**
	Scotland	White/unknown	General	**0.25**
		Black	General	**1**
		Asian	General	**0.35**
	NI	White/unknown	General	**0.6**
		Asian	General	**0.6**
Scotland	England/ other UK/ other EU	White/unknown	General	**0.25**
			RC	**1**
		Black	General	**1**
		Asian	General	**0.35**
			RC	**1**
		Mixed	General	**1**
			RC	**1**
		Other	General	**1**
	Wales	White/unknown	General	**0.39**
			RC	**1**
		Asian	General	**0.67**
	Scotland	White/unknown	General	**0.25**
			RC	**1**
		Black	General	**1**
		Asian	General	**0.35**
			RC	**1**
		Mixed	General	**1**
		Other	General	**1**
	NI	White/unknown	General	**0.6**
			RC	**1**
		Black	General	**1**
		Asian	General	**0.67**
		Mixed	General	**1**
		Other	General	**1**
NI	England/other UK/ other EU	White/unknown	General	**0.6**
			Foundation	**1**
		Black	General	**1**
		Asian	General	**0.56**
	Wales	White/unknown	General	**0.6**
	Scotland	White/unknown	General	**0.63**
		Asian	General	**0.6**
	NI	White/unknown	General	**0.6**
			Foundation	**0.6**
		Black	General	**1**
		Asian	General	**0.62**
		Other	General	**1**

APPENDIX B RESPONSE RATE DEFINITIONS

The response rates are calculated from the outcomes recorded for each survey. For the Early Survey a single outcome was provided. Our knowledge of the data collection process suggests that this may not be a comprehensive list of all refusals, nor are there any other reasons for non-response recorded (all labelled 'non-contact'). For this reason we refrain from calculating anything other than the response rate.

For the Longitudinal Survey the provisional outcomes were recorded separately for each data collection mode and had to be combined to produce a single 'final' outcome. This is illustrated in the diagram below. Rules were applied to prioritise outcomes for sample members who were approached by more than one mode. Established guidelines from AAPOR have been used to derive the rules.[21] Currently there are no similar published standard guidelines for UK research. The final outcome takes the best information from all available sources using a well-established hierarchy and where there is uncertainty the most recent contact attempt is used. 60% of the final outcomes came from telephone responses, 26% from postal responses, 4% from online/email responses, and the final 10% from all three modes.

Apx. Figure B.1 Derivation of final outcome for Longitudinal Survey



Further information was recorded separately about reasons for non-response (excluded due to the late receipt of this information). The uncertainty that we have

[21] The American Association for Public Opinion Research 2006 "Standard Definitions: final dispositions of case codes and outcome rates for surveys" 4th edition. Available from www.aapor.org.

about the recording of the reasons for non-response leads us to a conclusion that any rates beyond the response rate are only valid as minimum bounds, and even then they should be interpreted with caution since there is some ambiguity about how some of the non-response reasons should be interpreted.

The final outcomes for both surveys are shown in the tables below.

Apx. Table B.2 Final outcome for Early Survey (unweighted)

Outcome code	Frequency	Percentage
Respond	307,652	74.6%
Refusal	4,519	1.1%
Non-contact	100,408	24.3%
Total	412,579	100.0%

Apx. Table B.3 Final outcome for Longitudinal Survey for selected sample (unweighted)

Outcome code*	Frequency	Percentage
Respond	24,823	40.0%
Refusal	3,717	6.0%
Refusal – HEI opt-out	14	0.0%
Refusal – breakdown during interview	395	0.6%
Refusal – broken appointment	726	1.2%
Other non-response	178	0.3%
Other non-response - did not graduate from HE in 2002/3	105	0.2%
Other non-response - claimed had/will complete in other mode	777	1.3%
Non-contact – after at least 4 attempts	7,336	11.8%
Non-contact – no contact details	12,892	20.8%
Non-contact – contact details unknown validity	5,020	8.1%
Non-contact – valid address	48	0.1%
Non-contact – contact details incorrect	6,008	9.7%
Total	62,039	100.0%

* 27 additional refusals are known (including 9 opt-outs). 151 non-contacts have been identified as having moved address with some addresses validated.

The cases labelled in the tables above as ‘non-contacts’ were in fact all ‘unknown eligibility due to non-contact’. However these have been reclassified as eligible because for this sample it is unlikely that many will have become ineligible. The only way that a sample member can become ineligible is through death.[22] The age distribution of the sample selected for the Longitudinal Survey has been estimated (age on 31st July 2003 + 3 years). Around two thirds will have been aged 20-29 years old and a third aged 30+ years (mostly aged 30-49 years). The annual mortality rates in the UK for people of these ages are very low.

[22] 17 non-contacts are known to be deceased graduates.

The response rates have been defined in the following way:

$$\text{Early Survey response rate} = \frac{\text{Respond to Early Survey}}{\text{Eligible for Early Survey}}$$

$$\text{Longitudinal Survey response rate} = \frac{\text{Respond to Longitudinal Survey}}{\text{Eligible for Longitudinal Survey}}$$

$$\text{Longitudinal response rate} = \frac{\text{Respond to 1}^{st}\text{ and 2}^{nd}\text{ surveys}}{\text{Eligible for 1}^{st}\text{ and 2}^{nd}\text{ surveys}}$$

The response rate can be separated into two elements – contact rate and co-operation rate. The relationship between them is:

Response rate = Contact rate x Co-operation rate

And they are defined as:

$$\text{Longitudinal Survey contact rate} = \frac{\text{Contacted for Longitudinal Survey}}{\text{Eligible for Longitudinal Survey}}$$

$$\text{Longitudinal Survey co-operation rate} = \frac{\text{Respond to Longitudinal Survey}}{\text{Contacted for Longitudinal Survey}}$$

The final rate of interest, is the refusal rate defined as:

$$\text{Longitudinal Survey refusal rate} = \frac{\text{Refused at Longitudinal Survey}}{\text{Eligible for Longitudinal Survey}}$$

The rates are calculated using the information provided in the tables below. The stages correspond to Figure 2-1 shown in the main text of the report. The weighted sample sizes have been weighted to correct for the unequal probabilities of selection (only applicable for Longitudinal Survey). The table below indicates that 7,988 were removed from the pool of qualifying graduates for the Longitudinal Pilot. The graduates were from 32 HEIs that had volunteered for the Longitudinal Pilot. The exact selection criteria is not known. Random selection has been assumed, but it should be noted that some criteria have the potential to cause selection bias.

Apx. Table B.4 Available, selected and responding sample sizes

	Unweighted sample size	*Weighted sample size
Eligible for first wave	412,579	
Responded to first wave	307,652	
Qualify for second wave	307,652	
Selected for second wave pilot	7,988	
Available for second wave	299,664	
Selected (eligible) for second wave	62,039	**299,602
Responded to second wave	24,823	124,533

* Weighted using the inverse sampling fraction (1 / f_fracti).

** This does not equal the unweighted available sample exactly because of sampling fraction rounding errors.

Apx. Table B.5 Non-response categories for Longitudinal Survey

	Unweighted sample size	*Weighted sample size
Selected	62,039	299,602
Non-contact	31,304	145,801
Refusal	4,852	23,941
Other non-response	1,060	5,327
Respond	24,823	124,533

* Weighted using the inverse sampling fraction (1 / f_fracti).

Early Survey response rate (unweighted) = $\frac{307,652}{412,579}$ = 74.6%

Longitudinal Survey response rate (unweighted) = $\frac{24,823}{62,039}$ = 40.0%

Longitudinal Survey response rate (weighted) = $\frac{124,533}{299,602}$ = 41.6%

Longitudinal Survey contact rate (unweighted) = $\frac{62,039 – 31,304}{62,039}$ = 49.5%

Longitudinal Survey contact rate (weighted) = $\frac{299,602 - 145,801}{299,602}$ = 51.3%

Longitudinal Survey co-operation rate (unweighted) = $\frac{24,823}{62,039 – 31,304}$ = 80.8%

Longitudinal Survey co-operation rate (weighted) = $\frac{124,533}{299,602 – 145,801}$ = 81.0%

Longitudinal Survey refusal rate (unweighted) = $\frac{4,852}{62,039}$ = 7.8%

$$\text{Longitudinal Survey refusal rate (weighted)} = \frac{23{,}941}{299{,}602} = 8.0\%$$

The Longitudinal response rate is calculated by (i) adjusting the sample selected for the Longitudinal Survey to represent the population using the weights and (ii) grossing up the sample to match the population size to allow the non-respondents from the Early Survey to be added in. This is equivalent to multiplying the two response rates for the Early and Longitudinal Surveys together.

$$\text{Longitudinal response rate (weighted)} = \frac{124{,}533 \times sf}{(299{,}602 \times sf) + (412{,}579 - 307{,}562)} = 31.0\%$$

where sf = scaling factor = 307,562 / 299,602.

The rates for the sub-group of graduates who studied at Scottish institutions were calculated using the same definitions. The sample sizes are given in the first pair tables below, and the rates (unweighted and weighted) given in the second pair of tables.

Apx. Table B.6 Non-response categories for Longitudinal Survey for Scottish institution sub-group by domicile

	Unweighted sample size	*Weighted sample size
Scottish domicile		
Selected for second wave	5,520	21,259
Non-contact	2,679	10,246
Refusal	418	1,626
Other non-response	80	311
Respond	2,343	9,075
Other UK/EU domicile		
Selected for second wave	1,767	5,311
Non-contact	938	2,801
Refusal	132	401
Other non-response	22	63
Respond	675	2,046

* Weighted using the inverse sampling fraction (1 / f_fracti).

Apx. Table B.7 Non-response categories for Longitudinal Survey for Scottish institution sub-group by activity at Early Survey

	Unweighted sample size	*Weighted sample size
Unemployed at Early Survey**		
Selected for Longitudinal Survey	452	1,694

Non-contact	215	796
Refusal	40	155
Other non-response	6	22
Respond	191	721
Other activity at Early Survey* **		
Selected for Longitudinal Survey	6,835	24,876
Non-contact	3,402	12,252
Refusal	510	1,872
Other non-response	96	352
Respond	2,827	10,400

* Weighted using the inverse sampling fraction (1 / f_fracti).
** Unemployed if declared activity to be "starting job within next month" or "unemployed and looking for a job/further study/training".
*** All other activities including "not employed but not looking for a job/further study/training".

Apx. Table B.8 Contact, refusal and co-operation rates at Longitudinal Survey for graduates attending Scottish institutions by domicile

Definition	Scottish domicile	Other UK/EU domicile
Unweighted		
Longitudinal Survey contact	51.5%	46.9%
Longitudinal Survey refusal	7.6%	7.5%
Longitudinal Survey co-operation	82.5%	81.4%
Weighted		
Longitudinal Survey contact	51.8%	47.3%
Longitudinal Survey refusal	7.7%	7.6%
Longitudinal Survey co-operation	82.4%	81.5%

Apx. Table B.9 Contact, refusal and co-operation rates at Longitudinal Survey for graduates attending Scottish institutions by activity at Early Survey

Definition	Unemployed at 1st survey	Other activity at 1st survey
Unweighted		
Longitudinal Survey contact	52.4%	50.2%
Longitudinal Survey refusal	8.8%	7.5%
Longitudinal Survey co-operation	80.6%	82.3%
Weighted		
Longitudinal Survey contact	53.0%	50.7%
Longitudinal Survey refusal	9.2%	7.5%
Longitudinal Survey co-operation	80.3%	82.4%

APPENDIX C NON-RESPONSE PROFILES

The first table shows changes in the profile (compared to the population) as the sample progresses through different stages. The second column from the left gives the weighted profile of respondents to the Longitudinal Survey, the third column shows the weighted graduates selected for the Longitudinal Survey, the fourth column shows the respondents to the Early Survey, and the final column shows the population. The profile can change at each stage. The weighted profile of the Longitudinal Survey respondents is the most different to the population profile. This has been caused by differential non-response to the Longitudinal Survey.

Apx. Table C.1 Profile of population and samples at different stages

	2nd survey responding sample weighted	2nd survey selected sample weighted	1st survey respondent sample unweighted	Population
	%	%	%	%
Sex				
Female	61.3	58.4	58.2	57.1
Male	38.7	41.6	41.8	42.9
Age at July 2003				
Under 20 years	3.1	3.1	3.3	3.1
21 - 24 years	59.2	60.2	60.2	56.7
25 - 29 years	10.0	12.0	11.9	14.0
30 years and over	27.7	24.5	24.5	26.2
Ethnicity				
Asian	6.2	7.1	7.1	7.1
Black	2.0	2.7	2.6	3.0
Mixed	0.4	0.4	0.4	0.5
Other	1.0	1.2	1.2	1.3
White	84.0	81.7	82.0	80.1
Missing	6.3	7.0	6.7	8.0
Domicile				
England	82.7	83.5	83.2	83.7
Northern Ireland	4.0	3.6	3.7	3.3
Scotland	8.3	8.1	8.2	8.2
Wales	5.0	4.8	4.9	4.8
Country				
England	82.2	82.6	82.2	82.3
Northern Ireland	3.4	3.0	3.0	2.7
Scotland	8.9	8.9	9.0	9.3
Wales	5.4	5.5	5.7	5.7
Course subject				
A Medicine and Dentistry	1.8	1.9	1.9	1.8
B Subjects allied to Medicine	11.1	10.4	10.4	10.5
C Biological Sciences	7.3	7.2	7.3	6.9
D Veterinary Sciences, Agriculture and related subjects	1.2	1.1	1.1	1.0
F Physical Sciences	4.3	4.0	3.9	3.8
G Mathematical and Computer Sciences	6.9	7.5	7.4	7.3

H Engineering	4.7	4.8	4.9	5.1
J Technologies	0.5	0.7	0.6	0.6
K Architecture, Building and Planning	1.6	2.1	2.2	2.2
L Social studies	7.7	7.3	7.3	7.5
M Law	3.1	3.8	3.8	3.9
N Business and Administrative studies	12.5	12.8	12.6	13.3
P Mass Communications and Documentation	1.8	2.0	2.0	2.1
Q Linguistics, Classics and related subjects	3.5	3.5	3.6	3.5
R European Languages, Literature and related subjects	1.1	1.2	1.1	1.1
T Eastern, Asiatic, African, American and Australasian Lang	0.5	0.4	0.4	0.4
V Historical and Philosophical studies	4.2	3.8	3.8	3.7
W Creative Arts and Design	7.4	8.2	8.0	8.0
X Education	6.0	5.4	5.6	5.5
Y Missing	3.9	3.3	3.2	3.4
Z joint honours - various subjects	8.9	8.8	8.8	8.4
Level of course (2)				
Doctorate	2.1	1.9	1.7	2.1
Other Higher Degrees	9.5	9.7	10.0	12.3
PGCE	5.5	5.9	5.9	5.7
Other Postgraduate	6.0	6.0	5.9	6.7
First Degree	66.9	65.8	65.8	61.9
HND DipHE	5.4	6.0	6.0	6.1
Other Undergraduate	4.6	4.8	4.6	5.1
Class of degree				
Classification not applicable	33.1	34.1	34.1	38.0
First class honours	7.9	7.1	7.2	6.3
Upper second honours	32.3	30.9	30.7	28.0
Lower second honours	19.0	19.6	19.9	19.3
Third class honours	3.3	3.7	3.7	4.0
Unclassified	4.5	4.6	4.5	4.4
Mode of study				
Full time	70.6	71.8	72.0	70.1
Sandwich	6.2	6.6	6.5	6.0
Part time	20.2	18.3	18.3	19.9
Writing up	3.1	3.2	3.2	4.0
Main source of funding				
TDA	6.7	6.9	6.9	6.4
Health: DH/NHS/Social care	12.2	11.6	11.5	11.5
RC	0.5	0.5	0.5	0.6
Foundation degree	0.2	0.3	0.3	0.2
Activity at Early Survey				
Full-time paid work only (including self-employed)	56.1	56.5	56.8	
Part-time paid work only	8.6	7.6	7.4	
Voluntary/unpaid work only	0.7	0.7	0.7	
Work and further study	11.6	10.5	10.6	
Further study only	13.2	13.8	13.6	
Assumed to be unemployed	4.8	5.4	5.5	

Not available for employment	4.2	4.5	4.5	
Other	0.9	1.0	0.9	
Base (all graduates)				
Weighted	124,533	299,602	na	na
Unweighted	24,823	62,039	307,652	412,579

The next tables compares the profile of non-respondents to the Longitudinal Survey against the population to highlight differences between them.

Apx. Table C.2 Profile of population and non-respondents

	Non-resp to 2nd survey	Population
	%	%
Sex		
Female	56.3	57.1
Male	43.7	42.9
Age at July 2003		
Under 20 years	3.2	3.1
21 - 24 years	60.9	56.7
25 - 29 years	13.5	14.0
30 years and over	22.3	26.2
Ethnicity		
Asian	7.7	7.1
Black	3.1	3.0
Mixed	0.5	0.5
Other	1.3	1.3
White	80.0	80.1
Missing	7.4	8.0
Domicile		
England	84.1	83.7
Northern Ireland	3.3	3.3
Scotland	7.9	8.2
Wales	4.7	4.8
Country		
England	82.9	82.3
Northern Ireland	2.7	2.7
Scotland	8.8	9.3
Wales	5.6	5.7
Course subject		
A Medicine and Dentistry	1.9	1.8
B Subjects allied to Medicine	9.9	10.5
C Biological Sciences	7.1	6.9
D Veterinary Sciences, Agriculture and related subjects	1.0	1.0
F Physical Sciences	3.8	3.8
G Mathematical and Computer Sciences	7.9	7.3
H Engineering	4.9	5.1
J Technologies	0.8	0.6
K Architecture, Building and Planning	2.4	2.2
L Social studies	7.1	7.5

M Law	4.2	3.9
N Business and Administrative studies	12.9	13.3
P Mass Communications and Documentation	2.1	2.1
Q Linguistics, Classics and related subjects	3.6	3.5
R European Languages, Literature and related subjects	1.2	1.1
T Eastern, Asiatic, African, American and Australasian Lang	0.3	0.4
V Historical and Philosophical studies	3.6	3.7
W Creative Arts and Design	8.8	8.0
X Education	4.9	5.5
Y Missing	2.9	3.4
Z joint honours - various subjects	8.6	8.4
Level of course (2)		
Doctorate	1.7	2.1
Other Higher Degrees	9.9	12.3
PGCE	6.1	5.7
Other Postgraduate	6.0	6.7
First Degree	65.1	61.9
HND DipHE	6.4	6.1
Other Undergraduate	4.9	5.1
Class of degree		
Classification not applicable	34.9	38.0
First class honours	6.5	6.3
Upper second honours	30.0	28.0
Lower second honours	20.1	19.3
Third class honours	4.0	4.0
Unclassified	4.6	4.4
Mode of study		
Full time	72.7	70.1
Sandwich	6.9	6.0
Part time	17.1	19.9
Writing up	3.3	4.0
Main source of funding		
TDA	7.0	6.4
Health: DH/NHS/Social care	11.1	11.5
RC	0.5	0.6
Foundation degree	0.3	0.2
Activity at Early Survey		
Full-time paid work only (including self-employed)	56.8	
Part-time paid work only	7.4	
Voluntary/unpaid work only	0.7	
Work and further study	10.6	
Further study only	13.6	
Assumed to be unemployed	5.5	
Not available for employment	4.5	
Other	0.9	
Base		
Weighted	175,069	
Unweighted	37,216	

APPENDIX D SUB-GROUP RESPONSE RATES

Apx. Table D.1 Response rates to Longitudinal Survey for key sub-groups

	2nd survey responding sample weighted	2nd survey selected sample weighted	Response rates
Sex			
Female	76376	175015	43.6%
Male	48157	124587	38.7%
Age at July 2003			
Under 20 years	3867	9458	40.9%
21 - 24 years	73765	180395	40.9%
25 - 29 years	12427	36031	34.5%
30 years and over	34474	73718	46.8%
Ethnicity			
Asian	7747	21146	36.6%
Black	2528	7964	31.7%
Mixed	457	1327	34.4%
Other	1257	3503	35.9%
White	104670	244801	42.8%
Missing	7875	20861	37.7%
Country of domicile			
England	102979	250204	41.2%
Northern Ireland	4932	10745	45.9%
Scotland	10392	24214	42.9%
Wales	6231	14439	43.2%
Country of study institution			
England	102407	247515	41.4%
Northern Ireland	4254	8896	47.8%
Scotland	11121	26570	41.9%
Wales	6750	16621	40.6%
Course subject			
A Medicine and Dentistry	2240	5601	40.0%
B Subjects allied to Medicine	13840	31197	44.4%
C Biological Sciences	9069	21508	42.2%
D Veterinary Sciences, Agriculture and related subjects	1528	3360	45.5%
F Physical Sciences	5307	11925	44.5%
G Mathematical and Computer Sciences	8578	22331	38.4%
H Engineering	5836	14359	40.6%
J Technologies	674	2030	33.2%
K Architecture, Building and Planning	1958	6174	31.7%
L Social studies	9604	21994	43.7%
M Law	3881	11284	34.4%
N Business and Administrative studies	15614	38262	40.8%
P Mass Communications and Documentation	2272	6011	37.8%
Q Linguistics, Classics and related subjects	4392	10608	41.4%

R European Languages, Literature and related subjects	1332	3510	38.0%
T Eastern, Asiatic, African, American and Australasian language/literature/other	574	1085	52.9%
V Historical and Philosophical studies	5175	11487	45.0%
W Creative Arts and Design	9179	24663	37.2%
X Education	7522	16031	46.9%
Z joint honours - various subjects	11145	26253	42.5%
Level of course (1)			
Postgraduate PG	28765	70192	41.0%
Undergraduate UG	95768	229410	41.7%
Level of course (2)			
Doctorate	2646	5650	46.8%
Other Higher Degrees	11865	29121	40.7%
PGCE	6832	17545	38.9%
Other Postgraduate	7422	17876	41.5%
First Degree	83294	197253	42.2%
HND DipHE	6762	17916	37.7%
Other Undergraduate	5712	14240	40.1%
Class of degree			
Classification not applicable	41224	102302	40.3%
First class honours	9781	21235	46.1%
Upper second honours	40235	92683	43.4%
Lower second honours	23608	58763	40.2%
Third class honours	4057	10979	36.9%
Unclassified	5629	13640	41.3%
Mode of study			
Full time	87911	215257	40.8%
Sandwich	7664	19782	38.7%
Part time	25098	54971	45.7%
Writing up	3860	9592	40.2%
Main source of funding			
TDA	8404	20604	40.8%
Health: DH/NHS/Social care	15163	34674	43.7%
RC	617	1522	40.5%
Foundation degree	300	787	38.1%
Activity at Early Survey			
Full-time paid work only (including self-employed)	69886	169256	41.3%
Part-time paid work only	10682	22841	46.8%
Voluntary/unpaid work only	824	2127	38.8%
Work and further study	14447	31572	45.8%
Further study only	16403	41266	39.7%
Assumed to be unemployed	5926	16201	36.6%
Not available for employment	5230	13362	39.1%
Other	1134	2978	38.1%
Base (all eligible graduates)	124,533	299,602	41.6%

APPENDIX E NON-RESPONSE MODEL

The response behaviour of the sample members was modelled using logistic regression. The data were weighted by the selection weights before analysis. Some of the variables describe the demographic characteristics of the sample, some the characteristics of their higher education and others their behaviour at the Early Survey and information available for the Longitudinal Survey. These variables were: country of Institution, domicile, ethnicity, nationality, sex, age, religion, disability status, level of qualification obtained - 8 way split, classification of degree, mode of study for original course, subject of original course, funding; TDA, healthcare, foundation and RC, method of data collection for the Early Survey, activity at Early Survey, and contact details available for Longitudinal Survey (postal, email and telephone).

The logistic regression model generates the probability of a graduate participating in the survey given their 'type' (based on the predictor variables described above). Hence respondents who were of a type that were reluctant to take part will have a smaller probability. Not all the variables used made it into the final model. Variables not strongly related to a graduate's propensity to respond to this survey were dropped from the analysis.

Apx. Table E.1 Results from the logistic regression model

	Estimate coefficient	Standard Error of coefficient	Wald test	df	Significance of Wald test	Odds
Country of Institution			99.0	3	0.00***	
England					(baseline)	
Northern Ireland	0.3	0.05	25.6	1	0.00***	1.29
Scotland	0.2	0.03	25.9	1	0.00***	1.18
Wales	0.3	0.04	63.5	1	0.00***	1.37
Level of qualification obtained - 8 way split			85.5	6	0.00***	
Doctorate					(baseline)	
Other Higher Degrees	-0.5	0.07	51.9	1	0.00***	0.60
PGCE	-0.6	0.08	48.2	1	0.00***	0.56
Other Postgraduate	-0.5	0.07	47.0	1	0.00***	0.60
First Degree	-0.1	0.76	0.0	1	0.91	0.92
HND DipHE	-0.7	0.08	78.1	1	0.00***	0.51
Other Undergraduate	-0.6	0.08	67.5	1	0.00***	0.53
Classification of degree			48.8	5	0.00***	
Classification not applicable	0.2	0.75	0.1	1.0	0.75	1.27
First class honours					(baseline)	
Upper second class honours	-0.1	0.04	6.6	1.0	0.01**	0.91
Lower second class honours	-0.2	0.04	31.0	1.0	0.00***	0.81
Third class honours	-0.3	0.06	22.3	1.0	0.00***	0.77
Unclassified	-0.2	0.05	10.7	1.0	0.00***	0.84
Employment categories			63.5	7	0.00***	
E Further study only					(baseline)	
A Full-time paid work only (including self-employed)	0.1	0.03	5.4	1.0	0.02**	1.07
D Work and further study	0.1	0.04	7.2	1.0	0.01**	1.10

O Other	0.0	0.09	0.0	1.0	0.98	1.00
B Part-time paid work only	0.2	0.04	38.9	1.0	0.00***	1.28
F Assumed to be unemployed	-0.1	0.05	2.9	1.0	0.09*	0.93
G Not available for employment	0.0	0.05	0.5	1.0	0.49	0.97
C Voluntary/unpaid work only	0.0	0.11	0.1	1.0	0.73	1.04
TDA funding			4.1	1	0.04**	
No					(baseline)	
Yes	0.1	0.05	4.1	1	0.04**	1.10
Health funding			38.0	1	0.00***	
No					(baseline)	
Yes	0.2	0.03	38.0	1	0.00***	1.20
Method of data collection			1284.8	6	0.00***	
Standard questionnaire first mailing					(baseline)	
Standard questionnaire second mailing	-0.3	0.04	66.8	1	0.00***	0.73
Telephone survey graduate	-0.6	0.02	955.3	1	0.00***	0.53
Telephone survey third party	-0.8	0.03	793.5	1	0.00***	0.47
Electronic reply (email/online)	-0.4	0.08	20.6	1	0.00***	0.68
Own institutions student record	-0.8	0.08	108.8	1	0.00***	0.44
Other	-0.9	0.10	83.6	1	0.00***	0.41
Has email address			567.4	1	0.00***	
No					(baseline)	
Yes	0.5	0.02	567.4	1	0.00***	1.67
Has telephone contact available			3681.1	1	0.00***	
No					(baseline)	
Yes	1.4	0.02	3681.1	1	0.00***	4.21
Sex			50.0	1	0.00***	
Female					(baseline)	
Male	-0.1	0.02	50.0	1	0.00***	0.88
Age on 31st July 2003			201.7	3	0.00***	
Under 20 years					(baseline)	
21 - 24 years	-0.1	0.05	1.6	1	0.20	0.94
25 - 29 years	-0.3	0.06	37.3	1	0.00***	0.70
30 years and over	0.1	0.05	3.2	1	0.08*	1.10
Constant	-0.8	0.14	30.0	1	0.00	0.47

Notes:

1. The response is 1 = graduate responded, 0 = graduate did not respond
2. **Significance** is denoted by asterisks: *** = 1%, ** = 5% and * = 10%.
3. The **odds** are the exponential of the estimate coefficient with standard error **Std.Err.**
4. The **Wald** test measures the impact of the categorical variable on the model. If the test is significant then the categorical variable is considered to be 'significantly associated' with the response variable.
5. R squared = 0.113 (Cox and Snell)
6. The base is all graduates eligible for Longitudinal Survey (n = 62,039)

APPENDIX F LONG ACTIVITY HISTORY ITEM NON-RESPONSE

Item non-response rates for some key questions are given below. The questions selected are those associated with the highest levels of item non-response.

Activity date

For each activity the respondent was required to enter the date they started and ended (month and year). Complete dates - month and year for both dates recorded – are available for 17,745 activities (83% of all activities). The eligible period for the activity history was the period from graduation to the reference date (27th November 2006). This means the activity history should not contain any spells of activity that start before August 2002 or end after November 2006. 98% of the activities with complete dates had given a date within the eligible period (82% of all the activities).

Employment activities

The table below shows the levels of item non-response for questions associated with employment: income, employer location, industry and occupation codes. The levels of item non-response are higher than for the main survey for all questions.

The level of item non-response for income is high. Postal respondents showed a large increase compared to the levels found in the main employment section of the questionnaire (Section B), twice as high. In contrast, telephone respondents had a reduced level (13% for Section E, compared to 18% for Section B).

The level of missing information for employer locations is generally very high; it is missing for a fifth of employment activities. It appears that respondents were having difficulty remembering postcode districts for past employment. Current employer locations were easier to remember (Section B item non-response = 9%). Unlike before the level of item non-response from online respondents is particularly high.

Apx. Table F.1 Item non-response for employment activities, by questionnaire mode

	Telephone	Online	Postal – long	Total
	%	%	%	%
Income	13.1	22.4	36.0	17.2
Employer location (postcode district)	19.7	28.3	17.1	21.0
Industry (SIC)	6.0	11.8	31.5	9.2
Industry not coded	*5.7*	*8.9*	*14.9*	*7.1*
Industry questions skipped	*0.3*	*2.9*	*16.6*	*2.2*
Occupation (SOC)	14.9	11.5	23.0	15.0
Occupation not coded	*14.7*	*10.4*	*12.6*	*13.7*
Occupation questions skipped	*0.3*	*1.1*	*10.4*	*1.3*
Total	11,310	2,696	1,298	15,304

Base: All spells of employment (including self-employed and voluntary)

As before the amount of missing data for SOC is higher than the amount for SIC (15.0% compared to 9.2%).

With the exception of employer location, the highest levels of item non-response came from the postal questionnaire. This was expected.

Study/training activities

The table below shows the item non-response to questions asked about study/training activities: HEI location and subject of study (JACS code).

The item non-response for HEI location was less than for employer location (16.4% compared to 21.0%). The percentage of locations missing was substantially higher for the online questionnaire than other modes.

Item non-response for the subject of study was lower in the activity histories than in Section C. There is a very high level of missing cases in the postal questionnaire although the sample size is relatively small which reduced its effect on the total level of item non-response.

Apx. Table F.2 Item non-response for study/training activities, by questionnaire mode

	Telephone	Online	Postal -long	Total
	%	%	%	%
HEI location (postcode district)	11.8	30.4	18.9	16.4
Subject (JACS code)	10.8	12.1	50.3	14.3
Subject not coded	*9.6*	*10.8*	*18.9*	*10.6*
Subject question skipped	*1.2*	*1.3*	*31.4*	*3.7*
Total	1,455	454	169	2,078

Base: All spells or study/training (including part-time)

Comparison across activities

A comparison was made across the first and last activities recorded for key questions. We hypothesised that item non-response might be higher for activities that happened in the past compared to nearer the reference date because of memory difficulties (also called recall error). However respondents might have grown tired of filling in the details so that later activities might be less complete. These are opposing effects. It’s not clear whether they would cancel each other out or not. The item non-response rates for the first and last activities recorded for key questions are given in the table below. The item non-response for the spells of activity recorded later is generally higher, (with the exception of income), suggesting that the overriding factor may be respondents growing tired of filling in their details.

Apx. Table F.3 Overall item non-response for key items by spell order

	First activity recorded	Last activity recorded if more than one
	%	%
Income	16.7	14.5
Employer location	12.8	28.7
Industry (SIC)	7.6	15.4
Occupation (SOC)	13.9	16.6
Base: All employment activities (including self-employed and voluntary)	7,543	3,880
HEI location	15.4	16.1
Subject (JACS code)	12.8	15.0
Base: All study/training activities (including part-time)	1,170	472