

ISR 2000-01 derived statistics for funding allocations and monitoring in FECs

To Heads of further education colleges directly funded by the HEFCE
Of interest to those responsible for Student data, Audit, Finance
Reference 02/10
Publication date February 2002
Enquiries to For general enquiries contact:
Scott Chu tel 0117 931 7338
Anthony Ryan tel 0117 931 7297
e-mail isr_heifes_stats@hefce.ac.uk

Executive summary

Purpose

1. This document describes how we will use the Learning and Skills Council's July 2000-01 individualised student record (ISR) to monitor aggregate returns made to the HEFCE and inform funding for 2002-03.
2. This document also outlines the procedures we will adopt where our monitoring process identifies significant differences between aggregate and individualised returns.
3. The data included in the re-creation of the Higher Education in Further Education: Students survey 2000-01 (HEIFES2000) from the July 2000-01 ISR (ISR20) will form the basis of calculating premiums based on the following:
 - full-time mature undergraduates
 - the neighbourhood types of students
 - students receiving a Disabled Students' Allowance.

These premiums will be used in the funding allocations for 2002-03.

Key points

4. Where we identify significant differences in the comparison of ISR20 data with HEIFES2000, we will write to the institutions concerned by 14 March 2002 and ask for an explanation. We may audit institutions that are unable to explain adequately the reasons for any differences between the two data returns.
5. If we find, either through the reconciliation with ISR data or any data audit, that institutions have received higher allocations than they should have done, then the funding for those institutions will be reduced accordingly. Any funding adjustments following the reconciliation of ISR and HEIFES data for 2000-01 are likely to affect the funding previously announced for 2000-01, and may be consolidated into revised funding levels for all subsequent years.

Action required

6. Where we require a response, an action plan should be sent to Scott Chu detailing how the institution will reconcile the two data sources by 28 March 2002.

Introduction

7. This document describes how we will use the Learning and Skills Council's (LSC's) July 2000-01 ISR to:

- monitor aggregate returns made to the HEFCE
- inform funding for 2002-03.

8. This document also outlines the procedures we will adopt where our monitoring process identifies significant differences between aggregate and individualised returns. We believe this exercise helps us, and institutions, to better understand the data supplied to the LSC and the HEFCE, and ensures that higher quality, more consistent data are delivered to both the LSC and the HEFCE.

9. The data included in the re-creation of the Higher Education in Further Education: Students survey 2000-01 (HEIFES2000) from the July 2000-01 ISR (ISR20) will form the basis of calculating premiums based on the following:

- full-time mature undergraduates
- the neighbourhood types of students
- students receiving a Disabled Students' Allowance.

These premiums will be used in the funding allocations for 2002-03.

10. Where we identify significant differences in the comparison of ISR20 data with HEIFES2000, we will write to the institutions concerned and ask for an explanation. We may audit institutions that are unable to explain adequately the reasons for any differences between the two data returns.

11. If we find, either through the reconciliation with ISR data or any data audit, that institutions have received higher allocations than they should have done, then the funding for those institutions will be reduced accordingly. Any funding adjustments following the reconciliation of ISR and HEIFES data for 2000-01 are likely to affect the funding previously announced for 2000-01, and may be consolidated into revised funding levels for all subsequent years.

Selection of institutions asked to respond

12. We will ask for a response from institutions if the total difference in holdback calculated using HEIFES2000 and ISR20 exceeds £400,000.

13. In calculating the comparison holdback we have ignored any appeals for mitigation of holdback. Therefore, the holdback shown may differ from the agreed final holdback notified for 2000-01. Before making adjustments to an institution's funding as a result of this exercise, we will take full account of any previously agreed mitigation. We have adopted this approach to allow us to apply consistent monitoring procedures to all institutions irrespective of individual circumstances that have affected the agreed holdback.

14. Holdback was chosen as the basis of selection because:

- a. Holdback is calculated at the highest level of aggregation used in operating our funding process.
- b. It shows major changes in funding while remaining neutral to variations in other institutions' data.

15. With the introduction of the higher education (HE) data set for the July 2001-02 ISR (ISR23) we expect to see a decrease in the discrepancies between ISR and HEIFES data. Therefore, in future years, we will introduce more rigorous criteria.

16. Changes will only be made to our copy of the ISR20 data if large discrepancies that can easily be corrected become apparent. If such changes are necessary we shall expect institutions to adopt procedures designed to ensure that, in future years, the data submitted to the LSC are correct.

17. The LSC will not amend its version of ISR20 unless institutions re-submit their ISR20 directly to the LSC. We expect any changes to be included in the institution's December 2000-01 return (ISR22). Where institutions have already submitted their ISR22, the LSC provides the following guidance:

'Where a revised final return generates the same number of funding units as the original audited final return there is no automatic requirement for the external auditors to confirm they are content, although

the institution may wish to inform their auditors of the changes that have been made.

‘Where the revised return generates a different number of funding units to the original audited final return then the Learning and Skills Council would expect the external auditors to confirm the revision has been made in accordance with audit guidelines.’

18. Our copy of ISR20 data was extracted on 6 December 2001. It is a condition of HEFCE grant that accurate data are supplied on time to the LSC. If an institution has failed to meet this requirement we will write to it asking both that a valid ISR20 is submitted to the LSC, and that an action plan is submitted to us detailing how the institution will ensure it meets this requirement in future. The institution may be charged for any costs we incur in ensuring accurate data are submitted to our timescales.

HEFCE-recognised funding consortia

19. For the lead institution of an HEFCE-recognised funding consortium the re-creation will incorporate ISR20 data for each consortium member.

Responses required

20. We will write to all institutions indicating whether we require a response.

21. An action plan will be required for all institutions asked to make a response. Guidance for submitting an action plan will be included in the letter to heads of institution informing them that a response is required.

22. Responses should address one, or more, of the following problems and quantify the extent to which it contributes to the overall discrepancy:

- a. Errors in HEIFES2000.
- b. Errors in ISR20.
- c. Errors in the Qualification Database.
- d. Problems of fit with the HEFCE algorithm.

23. Annex A aims to help institutions identify the cause of discrepancies between their ISR data and HEIFES2000.

24. If a response indicates that the institution has received higher allocations than they should have done,

then the funding for the institution will be reduced accordingly. Any funding adjustments following the reconciliation of ISR20 and HEIFES2000 are likely to affect the funding previously announced for 2000-01, and may be consolidated into revised funding levels for all subsequent years.

Errors in HEIFES2000

25. Where errors are found in the HEIFES return a revised return will be required. Institutions will be informed of the outcome of these changes by their HEFCE higher education adviser.

Errors in ISR20

26. The implications of processing and accepting amendments to ISR data are significant, so we want to keep amendments to a minimum. In particular we would not normally expect to make amendments where:

- institutions have already identified similar weaknesses in last year’s data
- the error does not significantly affect the outcome of this exercise, or other HEFCE analysis.

27. Where an institution identifies similar weaknesses in two years’ data we may audit the institution’s HEIFES and ISR data to allow us, and the institutions, to better understand why the error has recurred and how similar problems can be avoided in future.

28. If amendments to the ISR20 data are necessary, we require that the corrections are submitted to us in a standard format. Details of the standard format are given in Annex C. This is essential in order to establish an audit trail of data changes, and to ensure that corrections can be processed quickly and accurately. If amendments to ISR20 data are received we will use this information to re-create HEIFES2000 tables. Where we are content that the amendments result in a reasonable comparison with the HEIFES2000 return, we will ask the institution to confirm the accuracy of the amendments. A confirmation form will be sent to the institution via e-mail. Provided that the institution is content with the changes, the form should be printed and signed by the member of staff with responsibility for signing off ISR data. Where we are not content that the amendments result in a reasonable comparison, we

will ask for a further response. Details of this process are given in Figure 1.

Errors in Qualification Database

29. ISR20 data has been linked to version 14.3 of the LSC's Qualification Database to obtain information about the qualification aim of study. Where it is identified that information contained on the Qualification Database is incorrect, institutions should notify the LSC of the error, copied to Scott Chu at the HEFCE, and request that the relevant entry is corrected for the next practical release of the database.

30. Where it is identified that the student is incorrectly linked to a qualification in the Qualification Database then:

- a. If the correct qualification exists a link should be made to it.
- b. If a correct qualification does not exist a new qualification code should be requested from the LSC.

31. Both cases will require an amendment to the qualification aim reference in ISR20 data. Where a new qualification code is requested we will require evidence that the request has been made and details of the new qualification.

32. Where changes to the Qualification Database are requested, institutions should notify Scott Chu when the request is accepted by the LSC.

Problems of fit with the HEFCE algorithm

33. Where the response identifies a difference of fit between our algorithm and data supplied, evidence of where this occurs will be required.

34. Annex B gives full details of the methods used in generating the comparisons, including details of where assumptions have been made due to poor fit between the data collected on HEIFES2000 and that supplied to the LSC. In addition, the SAS code used to generate the comparisons can be found on our web-site, www.hefce.ac.uk under 'Learning and teaching', 'Data collection'.

35. All institutions are invited to comment on the methods described in Annex B, and to suggest how they can be improved. Comments should be sent to Scott Chu.

Data used in allocating funding premiums

36. ISR20 data will form the basis of calculating the mature, disability and widening participation funding premiums that will be used in the funding allocations for 2002-03.

37. Details of the methodology and fields likely to be used in generating these data are given in paragraphs 59-77 of Annex B.

Information supplied

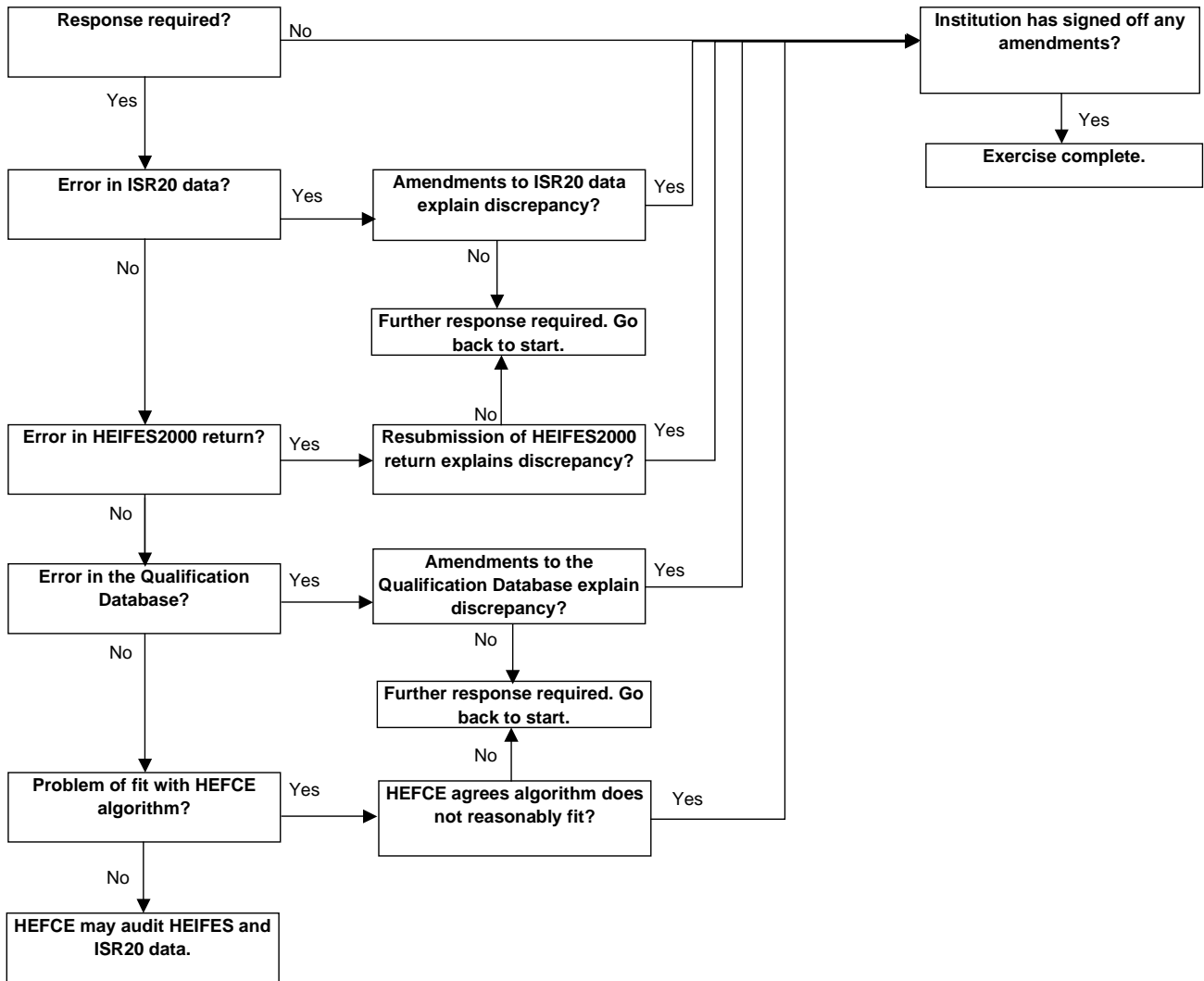
38. We will write to heads of institution, copied to HEIFES01 data contacts by 14 March 2002 indicating whether a response is required, and enclosing the following information:

- a. A copy of the HEIFES2000 return, including all amendments made during the funding process.
- b. HEIFES2000 as re-created using ISR20 data, and the coding methods described in Annex B.
- c. Holdback reports generated using HEIFES2000 data and ISR20 data.
- d. The numbers and proportions used to inform the mature, disability and widening participation premiums.

Annex A

Troubleshooting

Figure 1 Responses process diagram



- e. A summary of the ISR20 records excluded from the HEIFES re-creation.
 - f. Summary information comparing HEIFES2000 with ISR data.
39. An individualised file to support the comparison is available from our web-site. Details on how to access this file are given in Annex D.

Frequently asked questions

40. There is a web page featuring answers to frequently asked questions (FAQ). The web page will also contain any amendments or clarifications that need to be made. It can be found on the HEFCE web-site under 'Learning and teaching', 'Data collection'. The web page is updated regularly, and institutions are expected to look here for guidance in the first instance. We will only use our e-mail list of HEIFES01 contacts to notify institutions of any significant changes or updates.

Deadline for responses

41. Action plans should arrive no later than 28 March 2002 and should be sent to:

Scott Chu
Analytical Services Group
HEFCE
Northavon House
Coldharbour Lane
BRISTOL
BS16 1QD

42. We will agree details of the action plan, including the timetable for full response, by 8 April 2002. The agreed deadline for full response will not be later than 31 May 2002.

Annex A

Troubleshooting

1. This annex aims to help institutions identify the cause of any discrepancies between their ISR20 data and HEIFES2000 return. It is expected that institutions will have worked through this annex and consulted the web-based FAQ page (see paragraph 40) before seeking detailed assistance on resolving discrepancies.

Using this annex

2. Figure 2 describes a systematic method for identifying the highest level at which discrepancies between the returns occur. The subsequent paragraphs give possible causes for each discrepancy. These causes can be grouped into two categories:

- errors in completing specific fields on the ISR return
- weaknesses in the algorithm used in the re-creation.

The flowchart indicates which of the subsequent paragraphs are relevant when identifying weaknesses at each point.

3. Throughout this annex, fields taken from the ISR20 return or derived as part of the re-creation are shown in capitals, using the names given in Tables 1 and 2 of Annex B.

4. The match between HEIFES and ISR data is unlikely to be exact, due to estimates made when returning HEIFES and assumptions made in the re-creation algorithms (see paragraphs 8-20 of Annex B for further details). Therefore, when using the flowchart we expect institutions to exercise their own judgement to decide when small differences between the two data sources are not the result of an error. However, institutions need to be aware that small errors may accumulate and become significant. When the cause of a significant error cannot be determined, it may be necessary to backtrack to find the root of the problem.

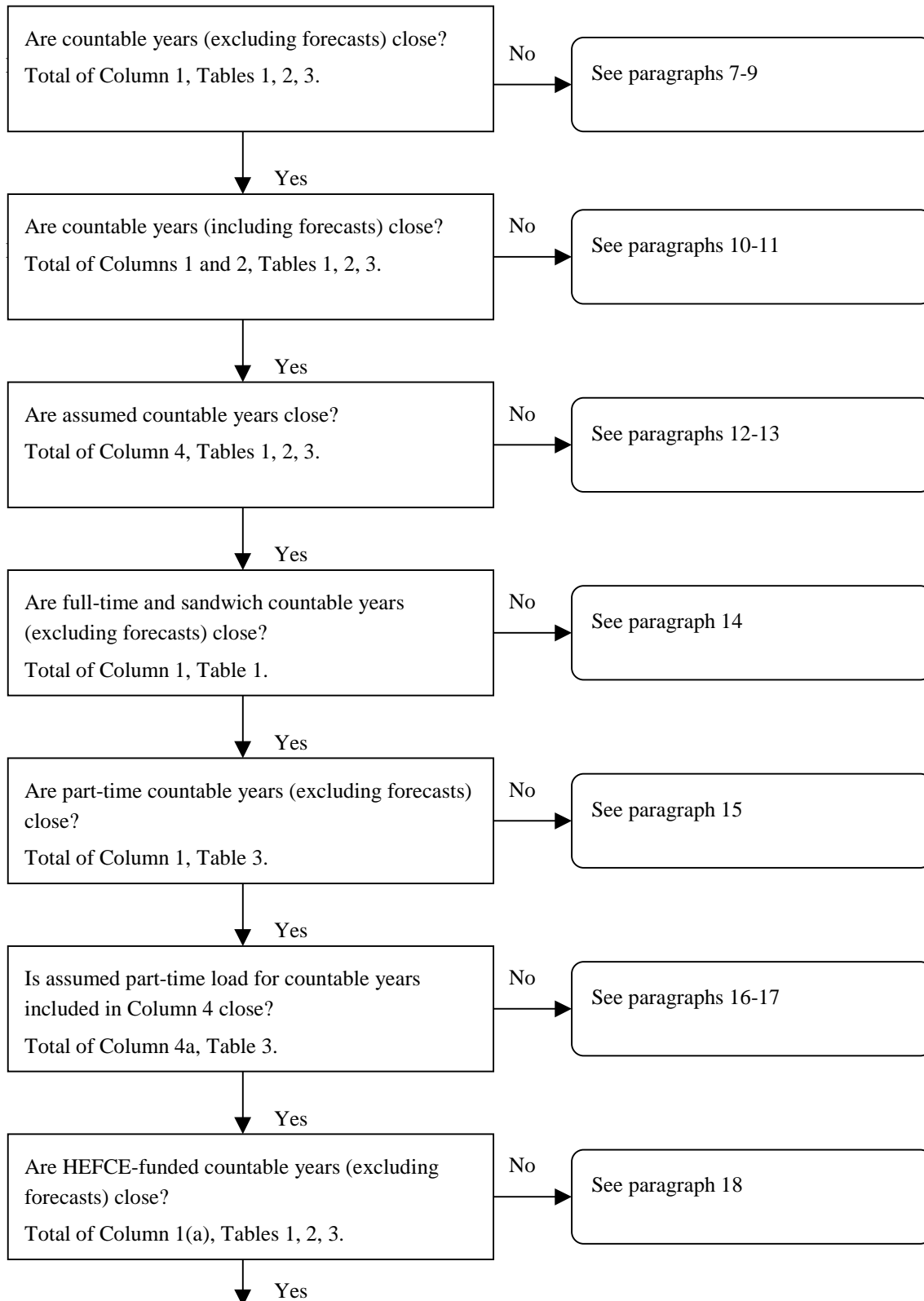
5. The description in this annex is aimed at identifying weaknesses in the ISR data rather than HEIFES. Therefore it is possible to follow the flowchart given in Figure 2 and yet still have discrepancies due to errors in HEIFES2000. If institutions find that a significant error in their HEIFES2000 return is the cause of discrepancies, a revised return will be required. Institutions will be informed of the outcome of these changes by their HEFCE higher education adviser.

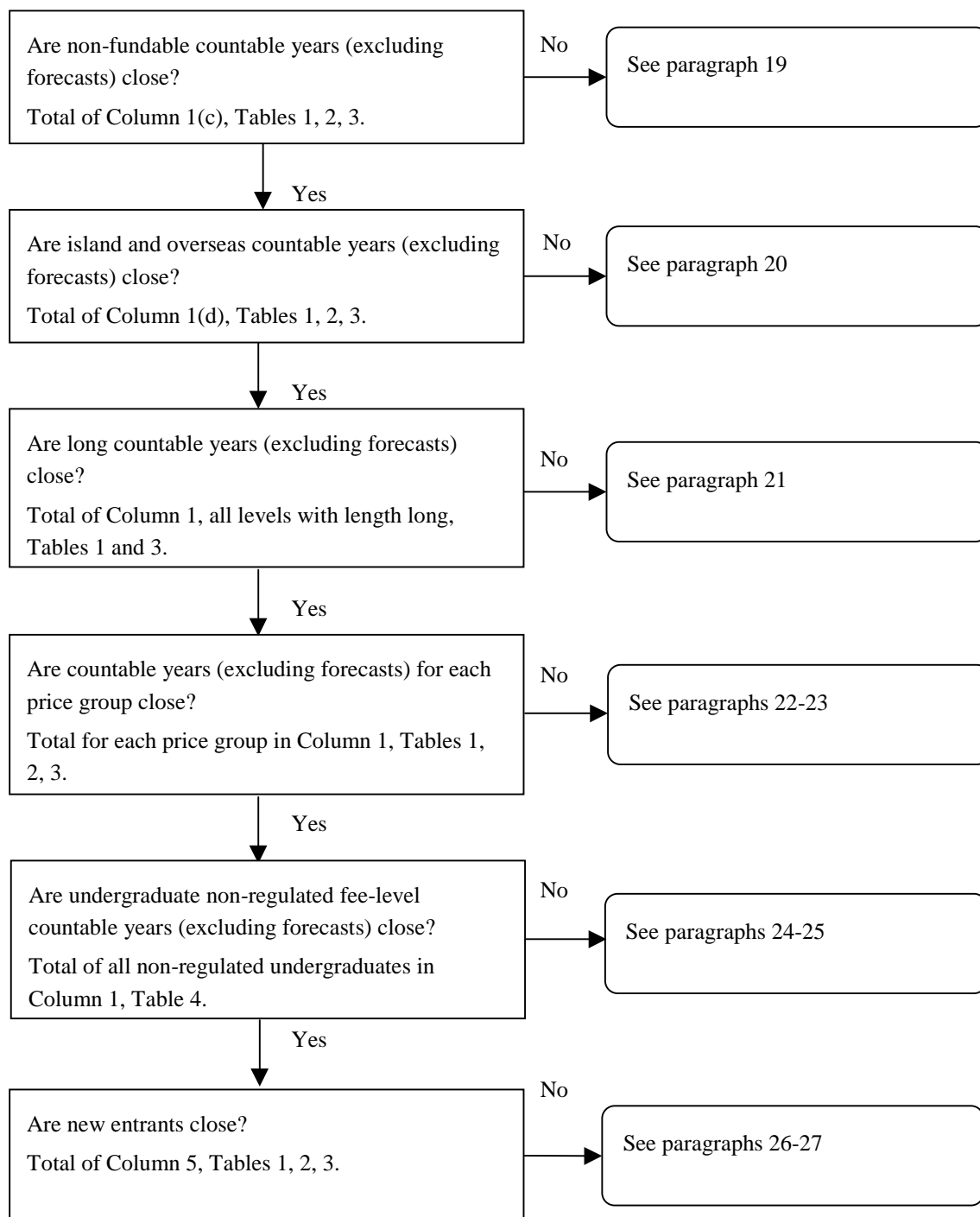
Using the individualised file

6. When working through this annex it is necessary to use the individualised file, indXXXXXX.csv, where XXXXX is the LSC code for the institution. Full details on how to access this file are given in Annex D. This will show the allocation of students to cells within the ISR-HEIFES2000 re-creation tables and, where relevant, details of why they were excluded. Records with specific values of a particular field in the individualised file can be found in the following way:

- a. Open the individualised file indXXXXXX.csv, containing student data in Excel.
- b. Select the row containing the field headings.
- c. Select <filter> from the data menu and then <autofilter>.
- d. Click on the arrow in the column containing the data which you want to filter.
- e. Either select a specific value or select <custom> to apply a comparison operator other than equality.

Figure 2 **Diagnostic flowchart** (paragraph numbers refer to Annex A)





Countable years (excluding forecasts)

7. To identify countable years (excluding forecasts) from the individualised file select HEFEXCL = 0 and HEFREG = 1. The algorithms for deriving HEFEXCL and HEFREG are given in paragraphs 50-51 and 44 respectively of Annex B.

8. Exceptionally a student generates two countable years of programme of study on the HEIFES return (see paragraph 11, Annex F of 'HEIFES2000', HEFCE 00/42, for further details). Students generating multiple years of programme of study are identified by STUBID = 1, 2. The algorithms for deriving STUBID are given in paragraphs 28-31 of Annex B.

9. Franchised-in students should not be returned in the individualised records. Paragraph 55 of the ISR 2000-01 Institution Support Manual refers to the return of franchised-in students.

Countable years (including forecasts)

10. To identify countable years (including forecasts) from the individualised file select HEFEXCL = 0. Forecast countable years are identifiable by HEFEXCL = 0 and HEFREG = 2. The algorithms for deriving HEFEXCL and HEFREG are given in paragraphs 50-51 and 44 respectively of Annex B.

11. There may be a small variance as a result of forecasting countable years on the HEIFES2000.

Assumed countable years

12. To identify assumed countable years from the individualised file select HEFCOL4 = 1. Forecast non-completions are identifiable by HEFEXCL = 0 and HEFCOMP = 3. The algorithms for deriving HEFCOL4, HEFEXCL and HEFCOMP are given in paragraphs 58, 50-51 and 45 respectively of Annex B.

13. There may be a small variance as a result of forecasting non-completions on HEIFES2000.

Full-time and sandwich countable years (excluding forecasts)

14. To identify full-time and sandwich countable years (excluding forecasts) from the individualised file select HEFEXCL = 0 and HEFREG = 1 and HEFMODE = FTS. The algorithms for deriving HEFEXCL, HEFREG and HEFMODE are given in paragraphs 50-51, 44 and 32 respectively of Annex B.

Part-time countable years (excluding forecasts)

15. To identify part-time countable years (excluding forecasts) from the individualised file select HEFEXCL = 0 and HEFREG = 1 and HEFMODE = PT. The algorithms for deriving HEFEXCL, HEFREG and HEFMODE are given in paragraphs 50-51, 44 and 32 respectively of Annex B.

Assumed part-time load for countable years included in Column 4

16. To identify part-time countable years included in Column 4 from the individualised file select HEFCOL4 = 1 and HEFMODE = PT. To obtain the load for these countable years sum the values of HEFESFTE. The algorithms for deriving HEFCOL4, HEFMODE and HEFESFTE are given in paragraphs 58, 32 and 48-49 respectively of Annex B.

17. The calculation of HEFESFTE may result in differences between the two data sources for students on non-standard academic years. This is due to an assumption that years of programme of study are in a steady state – that is, the same load is applicable to all years of programme of study. In particular, we make assumptions about the average load for cohorts of students in their final academic year (see the description of AVRGLOAD in the footnote of the table in paragraph 48 of Annex B for further details). This may cause a discrepancy for some records where ST_PYTYP = 5 and QA_EN_DA ≤ 31 July 2001 and ST_COMPY ≠ 2. (See Table 1 in Annex B for details of these fields.)

HEFCE-funded countable years (excluding forecasts)

18. To identify HEFCE-funded countable years (excluding forecasts) from the individualised file select HEFEXCL = 0 and HEFREG = 1 and HEFTYPE =

HOMEF. The algorithms for deriving HEFEXCL, HEFREG and HEFTYPE are given in paragraphs 50-51, 44 and 34 respectively of Annex B.

Non-fundable countable years (excluding forecasts)

19. To identify non-fundable countable years (excluding forecasts) from the individualised file select HEFEXCL = 0 and HEFREG = 1 and HEFTYPE = HOMENF. The algorithms for deriving HEFEXCL, HEFREG and HEFTYPE are given in paragraphs 50-51, 44 and 34 respectively of Annex B.

Island and overseas countable years (excluding forecasts)

20. To identify island and overseas countable years (excluding forecasts) from the individualised file select HEFEXCL = 0 and HEFREG = 1 and HEFTYPE = ISOV. The algorithms for deriving HEFEXCL, HEFREG and HEFTYPE are given in paragraphs 50-51, 44 and 34 respectively of Annex B.

Long countable years (excluding forecasts)

21. To identify long countable years (excluding forecasts) from the individualised file select HEFEXCL = 0 and HEFREG = 1 and LENGTH = L. The algorithms for deriving HEFEXCL, HEFREG and LENGTH are given in paragraphs 50-51, 44 and 41 respectively of Annex B.

Countable years (excluding forecasts) by price group

22. To identify countable years (excluding forecasts) by price group from the individualised file select HEFEXCL = 0 and HEFREG = 1 and PRICEGRP = X, where X is the price group of interest. The algorithms for deriving HEFEXCL, HEFREG and PRICEGRP are given in paragraphs 50-51, 44 and 42-43 respectively of Annex B.

23. Where qualifications on the LSC's Qualification Database can be assigned to multiple Superclass II codes, the re-creation assumes the code that maps to the lowest price group based on the following hierarchy:

- B
- C
- Psychology

- Media studies
- D.

Therefore if the institution has assumed a Superclass II code that maps to a higher price group than an alternative, this will result in a discrepancy. Records that may be affected can be identified where SUPERCL2 is not empty or SUPERCL3 is not empty.

Undergraduate non-regulated fee-level countable years (excluding forecasts)

24. To identify undergraduate non-regulated fee-level countable years (excluding forecasts) from the individualised file select HEFEXCL = 0 and HEFREG = 1 and HEFLEVEL = UG and HEFFEELV = NONREG. The algorithms for deriving HEFEXCL, HEFREG, HEFLEVEL and HEFFEELV are given in paragraphs 50-51, 44, 33 and 39 respectively of Annex B.

25. Because fee level was not collected on ISR20, the allocation of students to HEFFEELV may result in differences between the two data sources. In particular, students taking multiple qualification aims where the amount of tuition fees received is returned at the student level (i.e. ST_FAMNT > 0), are likely to be assigned to an incorrect fee level.

New entrants

26. To identify new entrants from the individualised file select HEFEXCL = 0 and YEARONE = 1. The algorithms for deriving HEFEXCL and YEARONE are given in paragraphs 50-51 and 40 respectively of Annex B.

27. Students taking foundation years at HE level (year 0) should be counted as new entrants only during their foundation year. The re-creation does not identify records returned with QA_PROGY = 0 as new entrants, however a student is always counted as a new entrant when QA_PROGY = 1. (See Table 1 in Annex B for details of this field.)

Annex B

Technical supplement

Purpose

1. This annex describes the method used to generate the data distributed to institutions as part of this exercise. It also gives details of known discrepancies between the data sources.
2. This annex is aimed at expert readers with in-depth knowledge of the data. Readers are advised to have a copy of the 2000-01 ISR institution support manual, HEIFES2000 and the guidance notes to version 14.3 of the LSC's Qualification Database to hand when using this document. In addition, users may wish to consult the SAS code used in the comparison which can be found on our web-site under 'Learning and Teaching', 'Data collection'.
3. The details listed below are similar to those in HEIFES2000 Annex N, but some alterations have been made to improve the matching and clarity.

Extraction and manipulation of ISR20 data

4. All ISR20 data returned to and passed as valid by the LSC before 6 December 2001 have been processed using the methods described in this document. These

data have been denormalised so that there is one record, including student data, for each qualification aim.

5. We will write separately to all institutions in receipt of HEFCE funding in 2001-02 for which we have not received ISR20 data, asking them to submit a valid ISR20 return to the LSC as soon as possible. These institutions may be charged for any costs we incur in ensuring the receipt of timely and accurate data for them.

Fields used in the re-creation

6. Only certain fields, detailed below, were used to generate the comparison between the HEIFES and ISR20 data. Fields from the student data set part of ISR20 take the prefix ST_; those from the qualification aim data set have the prefix QA_; and those from the Qualification Database use the same names as described in the guidance notes for version 14.3.
7. Throughout this annex, fields taken from ISR20 and version 14.3 of the Qualification Database are shown in capitals using the names given in the table below.

Table 1 Fields used in the re-creation

Field code	Description	Name	Data set	Column in individualised file*
AWARD_BOD1	First qualification awarding body	AWARD_BO	Qualification Database	N
AWARD_BOD2	Second qualification awarding body	AWARD_B2	Qualification Database	O
ENG_LEVEL	Level of study (FE or HE) England specific	ENG_LEVE	Qualification Database	L
QUAL_TYPE	Type of qualification	QUAL_TYP	Qualification Database	M
QUAL_TITLE	Qualification title	QUAL_TIT	Qualification Database	H
SUPERCLASS, SUPERCLAS2, SUPERCLAS3	The Superclass II subject classification	SUPERCL1, SUPERCL2, SUPERCL3	Qualification Database	Y, Z, AA
Q01	Qualification aim data set reference	ST_INST / QA_SEQNO	Qualification aim	A C
Q02	Qualification reference code	QA_AIM_R	Qualification aim	I
Q07A	Annual fees indicator	QA_F_IND	Qualification aim	AV
Q07B	Amount of tuition fees received or expected for the student	QA_FAMNT	Qualification aim	AX
Q08	Reason for partial or full non-payment of tuition fees	QA_NON_P	Qualification aim	AT

Q10	LSC or HEFCE funding	QA_FEHE_	Qualification aim	V	
Q11	Major source of funding other than tuition fees and LSC/HEFCE funding	QA_OTHER	Qualification aim	AB	
Q12	Minor source of funding other than tuition fees and LSC/HEFCE funding	QA_MINOR	Qualification aim	AO	
Q16	Start date	QA_ST_DA	Qualification aim	AE	
Q17	Expected end date	QA_EXP_E	Qualification aim	BV	
Q18	Actual end date	QA_EN_DA	Qualification aim	AH	
Q24	Institution-specified data 1	QA_COLL1	Qualification aim	F	
Q25	Institution-specified data 2	QA_COLL2	Qualification aim	G	
QHE01	Major source of tuition fees - HE specific	QA_FEEHE	Qualification aim	AP	
QHE02	Year of programme	QA_PROGY	Qualification aim	AM	
S01	Student data set reference	ST_INST / ST_REF	Student	A	B
S04	Date of birth	ST_DOB	Student	BI	
S06	Home postcode	ST_POSTC	Student	BP	
S07	Country of domicile	ST_DOMIC	Student	W	
S11	Additional support assessment	ST_SUPPA	Student	BU	
S14A	Annual fees indicator	ST_F_IND	Student	AW	
S14B	Amount of tuition fees received or expected for the student	ST_FAMNT	Student	AY	
S15	Reason for partial or full non-payment of tuition fees	ST_NON_P	Student	AU	
S17	Institution-specified data 1	ST_COLL1	Student	D	
S18	Institution-specified data 2	ST_COLL2	Student	E	
SHE01	Highest qualification on entry	ST_QUAL_	Student	BS	
SHE09	Type of programme year	ST_PYTYP	Student	AD	
SHE10	Mode applicable to HEIFES	ST_MHESE	Student	Q	
SHE11	Level applicable to HEIFES	ST_LHESE	Student	S	
SHE12	Completion of year of programme of study	ST_COMPY	Student	AG	
SHE13	Student FTE	ST_FTEHE	Student	AJ	

* The individualised data file, indXXXXXX.csv, downloadable from the web (see Annex D for further details).

Differences between HEIFES and ISR20 data

8. Following the revision of the 1998-99 ISR record, the HEIFES data can be more easily re-created.

However, there remain some data returned in HEIFES that cannot be re-created exactly using ISR data because of differences in definition. In such cases, reasonable approximations have been made. Listed below are areas where there may be some uncertainty about the correspondence of ISR records to HEIFES cells. Where possible, we have indicated the likely effects of the uncertainties.

Student load

9. The calculation of HEFESFTE described in paragraphs 46-49 ensure that the full-time equivalence (FTE), over the whole programme of study, is consistent with HEIFES, while attempting to minimise variance within years of programme of study.

10. Exceptionally, for non-standard academic years, assumptions are made for the final year of programme of study. The average FTE of students in the first academic year returned on the ISR, for all similar programmes of study at the same institution, is added to ST_FTEHE. A constant FTE for the course has been assumed. Therefore HEFESFTE will be deflated where the intensity of the course is increasing over time, and inflated when it is decreasing.

Two countable years of programme of study – first countable year

11. Where two countable years of programme of study are generated, data returned to ISR20 should reflect the year of programme of study at the end of the academic year. Therefore ISR20 data relate to the second countable year when two years are generated. For the first countable year we assumed ISR20 data with the following exceptions:

- ST_PYTYP = 4
- ST_COMPY = 1
- QA_PROGY = QA_PROGY - 1.

Fee level

12. Fee level is not currently collected on the ISR. Therefore if partial or full non-payment of fees has

occurred we assume the following fee has been charged:

- NHS – students whose major source of tuition fees are either the Department of Health, or, if studying for nursing and health related studies, their employer
- £1,050 – full-time undergraduates and PGCE students
- £520 – sandwich year-out undergraduates
- £0 – ERASMUS students
- Non-regulated – any students not included above.

13. Students generating two countable years are assumed as above for the first countable year and as non-regulated for the second countable year.

14. If full payment of tuition fees has been made we use the QA_F_IND, QA_FAMNT, ST_F_IND or ST_FAMNT pairs to assign fee level. We take fee data at qualification aim level in preference to fee information at the student level. However, where fee data are returned at the student level we assume that the fee information applies to a group of qualification aims that consist of one programme of study. Where fee information is returned for the whole programme of study we divide by an approximation of the expected length of programme of study derived by rounding to the nearest whole year QA_EXP_E - QA_ST_DA. Due to the imprecise nature of the approximation we assume a range of greater than or equal to the fee minus £50. Therefore HEFFEELV will be assigned to 'Non-regulated fee level' where we overestimate course length. From 2001-02 onwards fee level information is collected on the HE data set of the ISR.

Price group

15. Price group is assigned using the following hierarchy, where we assume the lowest price group the SUPERCL1 - SUPERCL3 fields map to:

- B
- C
- Psychology
- Media studies
- D.

16. In-Service Education of Teachers (INSET) is always assigned where QA_FEHE_ = 9 and QA_OTHER = 025.

17. From 2001-02 onwards proportion of study in each of the three Superclass II subject fields is collected on the HE data set of the ISR. We will use this information to apportion students between price groups.

Qualification aims information at the student level

18. The ISR format involves collecting HE qualification aims information at the student level. Where a student is studying for two or more qualifications simultaneously we assume information returned at the student level applies to all the student's qualification aims.

19. From 2001-02 this information is collected at the qualification aim level. Fields that will be affected in this year's exercise are ST_MHESE, ST_LHESE, ST_PYTYP, ST_COMPY and ST_FTEHE.

Areas of uncertainty in completing HEIFES2000

Forecasts of countable years of programme of study and non-completions

20. HEIFES2000 required institutions to provide forecasts of countable years and non-completions between 2 November 2000 and 31 July 2001. Forecasts are by their nature inexact. Unless there

are exceptional circumstances, it is expected that these figures will, when considered as a whole, be an accurate reflection of actual non-completions and forecasts of countable years.

Description of derived fields

21. This section contains details of the derived fields contained on the individualised data file. These fields are used to build the key dimensions of the HEIFES return.

Table 2 Derived fields

Field name	Description	Paragraph	Column in individualised data file*
ST_INST	LSC institution code	22	A
ST_REF	Unique student identifier	23	B
QA_SEQNO	Sequential numbering of each student's qualification aims	24	C
HEFOAIM	Recognised HE qualification aim	26	K
STUBID	Unique countable year of programme of study identifier	28	AZ
HEFMODE	Mode of study	32	P
HEFLEVEL	Level of study	33	R
HEFTYPE	Fundability status	34	U
TU_CASE	Indicator showing how HEFFEELV was calculated	35	AS
HEFTUIT	Estimate of fee level	38	AQ – AR†
HEFFEELV	Fee level used in HEIFES	39	AN
YEARONE	New entrant flag	40	AL
LENGTH	Long or standard length years of programme of study indicator	41	T
PRICEGRP	Price group	42	X
HEFREG	HEIFES column 1 or 2 indicator	44	AC
HEFCOMP	HEIFES column 3 or 4 indicator	45	AF
FTE_CASE	Indicator showing how HEFESFTE was calculated	46	AK
HEFESFTE	FTE consistent with HEIFES definitions	48	AI
HEFEXCL	Reason for exclusion from the HEIFES population	50	J
EXCL1 – EXCL32	Flags indicating reason(s) for a student's exclusion	52	BB – BG
HEFCOL4	Flag indicating whether the student was included in HEIFES Column 4	58	BA
ENTRYAGE	Age of student on entry to programme of study	60	BH
MATPOP	Flag indicating whether student was in the mature premium population	62	BJ
MATPREM	Flag indicating whether student was used to inform mature premium	63	BK
DISPOP	Flag indicating whether student was in the disability premium population	65	BL
DISPREM	Flag indicating whether student was used to inform disability premium	66	BM
SPC	Neighbourhood type indicator	71	BQ
WPWEIGHT	Widening participation weighting assignment	72	BR

NORMENT	Flag indicating normal entrants	73	BT
WPPPOP	Flag indicating whether student was in the widening participation funding stream population	75	BN
WPFUND	Flag indicating whether student was used to inform widening participation funding stream	76	BO

* The individualised data file, indXXXXXX.csv, downloadable from the web (see Annex D for further details).

† This field has been split across two columns in the individualised file. Column AP contains character values of HEFTUIT, column AQ contains numeric values of HEFTUIT.

ST_INST

22. This field is derived by the LSC and contains the LSC institution code. It is the first seven characters of the 'Student data set reference' (field S01).

ST_REF

23. This field is derived by the LSC and uniquely identifies students on ISR20. It is the first 19 characters of the 'Student data set reference' (field S01).

QA_SEQNO

24. This field is derived by the LSC. It contains a sequential number of the qualification aim records for each student. It is characters 24 and 25 of the 'Qualification aim data set reference' (field Q01).

Record uniqueness in ISR20

25. Record uniqueness in ISR20 is provided by taking the ST_REF, QA_SEQNO pair.

HEFQAIM

26. This field allocates qualification aims to broad qualification types.

HEFQAIM	Description	Definition
FIRST	First degree	QUAL_TYP = 0394, 1406, 1407, 1408, 1409, 9000, 9002, 9107, E007, and ENG_LEVE = H
MASTER	Masters	QUAL_TYP = 0393, 1410, 2001, 9100, 9101, 9109, and ENG_LEVE = H
HIGHER	Higher degree	QUAL_TYP = E008, 1411, 1412 and ENG_LEVE = H
DIPHE	DipHE	QUAL_TYP = 9112 and ENG_LEVE = H
PGCE	PGCE	QUAL_TYP = 9103 and ENG_LEVE = H
CERTED	CertEd	QUAL_TYP = 9111 and ENG_LEVE = H
FOUDEG	Foundation degree	QUAL_TYP = 9110 and ENG_LEVE = H
HNC	HNC	QUAL_TYP = 0031 and ENG_LEVE = H and (AWARD_BO or AWARD_B2 = EDEXCEL, SQA)
HND	HND	QUAL_TYP = 0032 and ENG_LEVE = H and (AWARD_BO or AWARD_B2 = EDEXCEL, SQA)
OTHER	Other qualifications	Otherwise

Second countable years of programme of study

27. Non-standard academic years where all activity for the final year of programme of study falls entirely within an academic year, generate two countable years of programme of study.

STUBID

28. This field uniquely identifies years of programme of study when two years are generated. Where a student generates two countable years of programme of study within a single academic year we create two records for that student. These records are distinguished using STUBID.

Value	Description
1	First countable year of programme of study
2	Second countable year of programme of study
0	One countable year of programme of study

29. Undergraduates (HEFQAIM = FIRST, DIPHE, CERTED, FOUDEG, HNC, HND) generated two countable years if:

- ST_PYTYP = 1
- QA_ST_DA < 1 August 2000
- QA_EN_DA < 1 August 2001
- month of QA_ST_DA = January to July inclusive and
- QA_EN_DA > anniversary of QA_ST_DA + 14 days.

30. Postgraduates (HEFQAIM = HIGHER, MASTER, PGCE) generated two countable years if:

- ST_PYTYP = 1
- QA_ST_DA < 1 August 2000
- QA_EN_DA < 1 August 2001 and
- QA_EN_DA > anniversary of QA_ST_DA + 14 days.

31. When STUBID = 1, we made the following assumptions:

- ST_PYTYP = 4
- ST_COMPY = 1
- QA_PROGY = QA_PROGY - 1.

HEFMODE

32. This field allocates students to mode of study.

Value	Description	Definition
FTS	Full-time and sandwich	ST_MHESE = 01
SWOUT	Sandwich year-out	ST_MHESE = 02
PT	Part-time	ST_MHESE = 03

HEFLEVEL

33. This field allocates students to level of study.

Value	Description	Definition
UG	Undergraduate	ST_LHESE = 10, 11
PG	Postgraduate	ST_LHESE = 20, 21, 30, 31

HEFTYPE

34. This field allocates students to the four categories of fundability and residential status.

Value	Description	Definition
HOMEF	Home and EC HEFCE funded	QA_FEHE_ = 2
HOMEIF	Home and EC independently funded	QA_FEHE_ = 5
HOMENF	Home and EC non-fundable	ST_DOMIC = EC* and QA_FEHE_ ≠ 2, 5
ISOV	Island and overseas	Otherwise

* EC domiciled is identified where ST_DOMIC = 099, 299, 399, 599, 610, 614, 641, 651, 653, 656, 659, 661, 671, 676, 678, 693, 710, 718, 728, 751, 755, 827.

TU_CASE

35. This field contains the case description for calculating HEFTUIT, as given in the table below.

36. The method used to generate the level of tuition fee charged to the student is dependent on the following factors:

- Whether tuition fees have been waived.
- Whether tuition fees have been returned for the current academic year or for the whole programme of study.
- Number of countable years of programme of study generated in HEIFES2000.

37. The table below shows how we identify different values of TU_CASE.

Value	Description	Definition
1	Tuition fees have been waived	QA_NON_P or ST_NON_P* = 04, 08, 10, 13, 14, 15
2	One year generated in HEIFES2000 and annual fees returned	STUBID = 0 and (QA_F_IND = A or ST_F_IND* = A)
3	One year generated in HEIFES2000 and whole programme of study fees returned	STUBID = 0 and (QA_F_IND = W or ST_F_IND* = W)
	Two years generated in HEIFES2000	
4a	1st year	STUBID = 1
4b	2nd year	STUBID = 2

* Data returned in the student fields will only be used where data have not been returned in the qualification aim fields.

HEFTUIT

38. This field contains an approximation of the tuition fee charged to the student as in HEIFES2000. The table below shows the method of calculating HEFTUIT for different methods of returning tuition fee data.

Value of HEFTUIT	Definition
NHS	TU_CASE = 1, 4a and QA_FEEHE = 31 or (QA_FEEHE = 81 and SUPERCL1 = PB, PF, PH, PJ)
1050	TU_CASE = 1, 4a and HEFMODE = FTS and (HEFLEVEL = UG or HEFQAIM = PGCE)
520	TU_CASE = 1, 4a and HEFMODE = SWOUT and HEFLEVEL = UG
0	TU_CASE = 1, 4a and (QA_OTHER = 017 or QA_MINOR = 017)
NONREG	(TU_CASE = 1, 4a and not included above) or TU_CASE = 4b
TUITION ^o	TU_CASE = 2
TUITION ^o /CRSELGTH*	TU_CASE = 3

^o TUITION = QA_FAMNT unless QA_FAMNT is empty when TUITION = ST_FAMNT.

* CRSELGTH is an approximation of the expected length of the programme of study derived by rounding up to the nearest whole year $QA_EXP_E - (QA_ST_DA + 14 \text{ days})$.

HEFFEELV

39. This field contains the level of tuition fee charged to the student as in HEIFES2000.

Value	Description	Definition
NHS	NHS bursaried courses	HEFTUIT = NHS
1050	Undergraduate full fee	HEFTUIT = 1050 or (HEFMODE = FTS and (HEFLEVEL = UG or HEFQAIM = PGCE) and HEFTUIT ≥ 1000)
520	Undergraduate half fee	HEFTUIT = 520 or (HEFMODE = SWOUT and HEFLEVEL = UG and HEFTUIT ≥ 470)
0	Regulated zero	HEFTUIT = 0 or QA_OTHER = 017 or QA_MINOR = 017
NONREG	Non-regulated fee	Otherwise

YEARONE

40. This indicates whether a student is a new entrant as defined in HEIFES2000.

Value	Description	Definition
1	New entrant	(ST_PYTYP = 1 and QA_PROGY = 1) or (ST_PYTYP = 4, 5 and QA_PROGY = 2)
0	Not new entrant	Otherwise

LENGTH

41. This field indicates whether the student is on a standard or long year of programme of study.

Value	Description	Definition
L	Long	ST_LHESE = 11, 21, 31
S	Standard	Otherwise

PRICEGRP

42. Price group is assigned by mapping the three Superclass II fields, SUPERCL1 - SUPERCL3, to price groups as indicated in the table below. We assume the lowest price group using the ordering given in the table, that is, PRICEGRP = B will only be assigned if all non-empty Superclass II fields map to price group B.

PRICEGRP	Superclass II code
B	PB, PE, RA, RC - RF, RH, SA, SB, SK, TL, TM, XH - XL, XP - XR, XT, YC - YE
C	CA - CH, CY, DC, FK (excluding FK.343), J, L, M, NA - NH, PA, PC, PD, PF - PJ, PL - PQ, RB, RG, SC - SJ, SL - SP, TA - TK, W, XA - XF, XM, XN, XS, YA, YB
PSYCH	PK
MEDIA	KE - KG
D	A, B, CX, CZ, DA, DB, DD, DE, E, FB, FC, FJ, FK.343, FL, G, H, KA - KD, KH, NK - NN, Q, V, Z
INSET	See paragraph 43

INSET and ITT

43. ITT(QTS) and INSET(QTS) students are identified by QA_FEHE_=9 and QA_OTHER=025. We assume all such students are INSET.

HEFREG

44. This field indicates whether the student will appear in Column 1 or 2 of the appropriate HEIFES2000 tables. It should be noted that if the student is excluded this field is not used to populate the tables.

Value	Description	Definition
1	Included in Column 1	(ST_PYTYP = 1 and anniversary of QA_ST_DA < 2 November 2000) or ST_PYTYP = 4, 5
2	Included in Column 2	Otherwise

HEFCOMP

45. This field indicates whether the student will appear in Column 3 or 4 of the appropriate HEIFES2000 tables. It should be noted that if the student is excluded this field is not used to populate the tables.

Value	Description	Definition
3	Included in Column 3	(ST_PYTYP = 1 and ST_COMPY = 2) or (ST_PYTYP = 4, 5 and ST_COMPY = 2 and QA_EN_DA < anniversary of QA_ST_DA)
4	Included in Column 4	Otherwise

FTE_CASE

46. This field contains the case description as given in the table below.

47. For non-standard academic years or when two years of programme of study are generated, the method used to calculate HEFESFTE is dependent on the number of countable years of programme of study generated in HEIFES2000 and whether the year of programme of study is the last or not.

Value	Description	Definition
1	One year generated in HEIFES2000 and the programme of study is not in the final academic year	ST_PYTYP = 4 and STUBID = 0
2	One year generated in HEIFES2000 and the programme of study is in the final academic year	ST_PYTYP = 5 and STUBID = 0
	Two years generated in HEIFES2000	
3a	First year	STUBID = 1
3b	Second year	STUBID = 2
0	Standard academic year	Otherwise

HEFESFTE

48. This field contains the FTE we assume for the year of programme of study in Column 4a of HEIFES2000. When the year of programme is contained within a standard academic year and one year of programme of study is generated, HEFESFTE is taken to be ST_FTEHE. The table below shows the method of calculating HEFESFTE for different groups of non-standard academic years of programme of study.

FTE_CASE	Definition
1	ST_FTEHE
2	ST_FTEHE+AVRGLOAD*
3a	(ST_FTEHE+AVRGLOAD*) - (ST_FTEHE×PROP [†])
3b	ST_FTEHE×PROP [†]

* AVRGLOAD is the arithmetic mean of ST_FTEHE for all non-standard academic years of programme of study in their first academic year, with the same ST_MHESE and HEFQAIM at the same institution.

[†] PROP = (QA_EN_DA - anniversary of QA_ST_DA) / (QA_EN_DA - 1 August 2000).

49. HEFESFTE is capped at 100. HEFESFTE is set to 50 for all sandwich year-out years of programme of study (HEFMODE = SWOUT). HEFESFTE is set to 100 for all full-time and sandwich years of programme of study (HEFMODE = FTS).

HEFEXCL

50. This field indicates whether the student should be included in the HEIFES re-creation and, if not, the reason for the exclusion. The following table gives details of students excluded from HEIFES2000 and the binary exclusion code used (HEFEXCL).

Value	Description	Definition
1	Not active in academic year	(QA_ST_DA > 31 July 2001 or QA_EN_DA < 1 August 2000)
2	Non-recognised HE, FE, NVQ or QTS students	HEFOAIM = OTHER
4	Students explicitly excluded from the HEIFES2000 student population	ST_LHESE = 99, ST_MHESE = 99 or ST_COMPY = 9
8	Students with an FTE of less than 3%	HEFESFTE < 3
16	Students on a non-standard academic year in the first academic year	ST_PYTYP = 3
32	Students on standard academic years who withdrew before 2 November 2000 or students on non-standard academic years who withdrew before the anniversary of their commencement date	QA_EN_DA < 2 November 2000 and ((ST_PYTYP = 1 and ST_COMPY = 2) or (QA_EN_DA < anniversary of QA_ST_DA and ST_PYTYP = 4, 5 and ST_COMPY = 2))

51. The value in HEFEXCL will be the sum of all applicable exclusion codes for a student. For example, if HEFEXCL = 14, then subtracting figures from the above table starting at the bottom, we see that the student has an FTE of less than 3 per cent (HEFEXCL = 8), is explicitly excluded from the HEIFES2000 student population (HEFEXCL = 4) and is a non-recognised HE, FE, NVQ or QTS student (HEFEXCL = 2).

EXCL1

52. Flag indicating students excluded due to non-activity in the academic year.

Value	Description	Definition
1	Not active in academic year	(QA_ST_DA > 31 July 2001 or QA_EN_DA < 1 August 2000)
0	Active in academic year	Otherwise

EXCL2

53. Flag indicating students excluded because they are studying for a non-recognised HE, FE, NVQ or QTS programme of study.

Value	Description	Definition
1	Non-recognised HE, FE, NVQ or QTS qualification aim	HEFOAIM = OTHER
0	Recognised HE qualification aim	Otherwise

EXCL4

54. Flag indicating students excluded because they were returned as 'Not in HEIFES population' on ISR20.

Value	Description	Definition
1	Student explicitly excluded from the HEIFES2000 student population	ST_LHESE = 99, ST_MHESE = 99 or ST_COMPY = 9
0	Student not explicitly excluded from the HEIFES2000 student population	Otherwise

EXCL8

55. Flag indicating students excluded due to an FTE of less than 3 per cent.

Value	Description	Definition
1	FTE of less than 3%	HEFESFTE < 3
0	FTE of at least 3%	Otherwise

EXCL16

56. Flag indicating students excluded because they are on non-standard academic years of programme of study in the first academic year.

Value	Description	Definition
1	Students on non-standard years of programme of study in the first academic year	ST_PYTYP = 3
0	Students not on non-standard years of programme of study in the first academic year	Otherwise

EXCL32

57. Flag indicating students excluded for withdrawing before 2 November 2000.

Value	Description	Definition
1	Students on standard academic years who withdrew before 2 November 2000 or students on non-standard academic years who withdrew before the anniversary of their commencement date	QA_EN_DA < 2 November 2000 and ((ST_PYTYP = 1 and ST_COMPY = 2) or (QA_EN_DA < anniversary of QA_ST_DA and ST_PYTYP = 4, 5 and ST_COMPY = 2))
0	Not above	Otherwise

HEFCOL4

58. This field indicates whether the student is assigned to Column 4 of HEIFES2000.

Value	Description	Definition
1	Included in Column 4 of HEIFES	HEFCOMP = 4 and HEFEXCL = 0
0	Not included in Column 4 of HEIFES	Otherwise

Premium funding proportions

59. Proportions are derived from ISR20 data that drive the 2002-03 allocation of the mature, disability and widening participation premiums. The remainder of this annex details the methodologies likely to be used in generating the proportions that will be used to inform these premiums for 2002-03. Σ is used to denote ‘the sum of’ from this point onwards.

ENTRYAGE

60. This field contains the student’s age at the commencement of the programme of study.

ENTRYAGE = (QA_ST_DA – ST_DOB)/365.25 rounded down to the nearest whole number.

Mature premium

61. A single proportion, $\Sigma\text{MATPREM}/\Sigma\text{MATPOP}$, is calculated for each institution. This proportion is applied to the number of home and EC funded, full-time and sandwich, and sandwich year-out undergraduates returned in Column 4 of HEIFES01.

MATPOP

62. This flag indicates whether the student was included in the denominator of the mature premium proportion calculation.

Value	Description	Definition
1	Included in the denominator of the mature premium proportion calculation	HEFCOL4 = 1 and HEFTYPE = HOMEF, HOMEIF and HEFMODE = FTS, SWOUT and HEFLEVEL = UG
0	Not included in the denominator of the mature premium proportion calculation	Otherwise

MATPREM

63. This flag indicates whether the student was included in the numerator of the mature premium proportion calculation.

Value	Description	Definition
1	Included in the numerator of the mature premium proportion calculation	HEFCOL4 = 1 and HEFTYPE = HOMEF, HOMEIF and HEFMODE = FTS, SWOUT and HEFLEVEL = UG and ENTRYAGE \geq 25
0	Not included in the numerator of the mature premium proportion calculation	Otherwise

Disability premium

64. A single proportion, $\Sigma\text{DISPREM}/\Sigma\text{DISPOP}$, is calculated for full-time and sandwich year-out undergraduates. The relevant proportion is then applied to the number of HEFCE funded Column 4 students returned in HEIFES01, for each of the four combinations of mode (full-time and part-time) and level (undergraduate and postgraduate).

DISPOP

65. This flag indicates whether the student was included in the denominator of the disability premium proportion.

Value	Description	Definition
1	Included in the denominator of one of the disability premium proportions	HEFCOL4 = 1 and HEFTYPE = HOMEF, HOMEIF, HOMENF and YEARONE = 1 and HEFLEVEL = UG and HEFMODE = FTS, SWOUT
0	Not included in the denominator of one of the disability premium proportions	Otherwise

DISPREM

66. This flag indicates whether the student was included in the numerator of the disability premium proportion.

Value	Description	Definition
1	Included in the numerator of one of the disability premium proportions	HEFCOL4 = 1 and HEFTYPE = HOMEF, HOMEIF, HOMENF and YEARONE = 1 and HEFLEVEL = UG and HEFMODE = FTS, SWOUT and ST_SUPPA = 7
0	Not included in the numerator of one of the disability premium proportions	Otherwise

67. 2000-01 is the first year that Disabled Students' Allowance information was collected on the ISR. Therefore we have restricted the calculations to new entrants only.

Neighbourhood type

68. For the purpose of the widening participation funding stream, each student is assigned a weighting based on the estimated participation of young students of the neighbourhood type of the student's home postcode. The weighting assignment was based solely on participation rather than other factors such as estimated income. There is a web page showing the mapping of neighbourhood types to weighting assignments, which can be found on the HEFCE web-site under 'Good practice', 'Classification of students based on postcode'.

69. The funding assignments are essentially pragmatic and have restrictions on the sources of data used and the type of students that are classified. The performance indicators' (PI) grouping of small areas provides a more sensitive and accurate look at the participation of young students. This identifies low participation areas as those with consistently less than two-thirds of the mean national young participation rate. See Annex A2 of 'Performance indicators in higher education in the UK' (HEFCE 01/69) for further details. There is a web page showing the mapping of neighbourhood types to PI assignments, which can be found on the HEFCE web-site under 'Good practice', 'Classification of students based on postcode'.

70. The assignment of postcodes to neighbourhood types is derived through a commercial product called Super Profiles developed by a company called Claritas. Super Profiles is based on 1991 census data, which together with other information such as electoral roll data and market research data, are used to classify households according to the characteristics of the neighbourhood in which they are located. We do not advise institutions to purchase the Super Profiles product purely to try to match our funding methodology, especially given the cost involved. However geodemographic classification systems such as Super Profiles may be useful to help institutions with widening participation policies. Other geodemographic systems available include Mosaic from Experian, and Acorn from CACI.

SPC

71. This field holds the Super Profiles neighbourhood type assignment based on the student's home postcode (ST_POSTC).

WPWEIGHT

72. This field holds the widening participation weighting assignment based on SPC.

Value	Description
0	A widening participation weighting of 0
1	A widening participation weighting of 1
2	A widening participation weighting of 2
3	A widening participation weighting of 3

NORMENT

73. This flag indicates whether the student is a normal entrant.

Value	Description	Definition
1	Normal entrant	ST_QUAL_ = 26 to 98
0	Not normal entrant	Otherwise

Widening participation funding stream

74. Three proportions, $\Sigma WPFUND / \Sigma WPPOP$, are calculated for each combination of full-time and the three non-zero widening participation weightings. A single proportion, $\Sigma WPFUND / \Sigma WPPOP$, is calculated for part-time and all below average participation SPCs. The three full-time proportions are multiplied by the proportion of young entrants ($ENTRYAGE < 25$), and the part-time proportion is multiplied by the proportion of normal entrants ($NORMENT = 1$). The relevant proportions are applied to the number of HEFCE funded Column 4a FTEs in HEIFES01 for each mode. The sum of the three proportions for full-time is assumed.

WPPOP

75. This flag indicates whether the student was included in the denominator of one of the widening participation funding stream proportion calculations.

Value	Description	Definition
1	Included in the denominator of one of the widening participation funding stream proportions	HEFCOL4 = 1 and HEFTYPE = HOMEF and ST_DOMIC = 399 and HEFLEVEL = UG and ((HEFMODE = FTS, SWOUT and ENTRYAGE < 25) or (HEFMODE = PT and NORMENT = 1))
0	Not included in the denominator of one of the widening participation funding stream proportions	Otherwise

WPFUND

76. This flag indicates whether the student was included in the numerator of one of the widening participation funding stream proportion calculations.

Value	Description	Definition
1	Included in the numerator of one of the widening participation funding stream proportions	HEFCOL4 = 1 and HEFTYPE = HOMEF and ST_DOMIC = 399 and HEFLEVEL = UG and ((HEFMODE = FTS, SWOUT and ENTRYAGE < 25) or (HEFMODE = PT and NORMENT = 1)) and WPWEIGHT = 1, 2, 3
0	Not included in the numerator of one of the widening participation funding stream proportions	Otherwise

77. To generate the four widening participation funding stream proportions we subsetted WPPPOP into the two modes (full-time and part-time), and WPFUND into the four combinations: full-time/weighting 1, full-time/weighting 2, full-time/weighting 3, and part-time/weighting 1, 2 and 3.

Annex C

Submitting corrections to ISR20 data

1. This annex details the data structure and format that enable us to efficiently process corrections to ISR20 data. Institutions should supply ISR20 amendment files using the file structure and format detailed here. An example of an amendment file can be found at the end of this annex.

2. These specifications are necessary to ensure we can process corrections to ISR20 data in a timely and accurate manner. We will ask institutions to re-submit amendment files that differ, either in structure or format, from the specifications detailed in this annex.

Format

3. Corrections to ISR20 data should be sent as a comma-separated file either on 3.5-inch floppy disk, CD-ROM or as an e-mail attachment sent to `isr_heifes_stats@hefce.ac.uk`. Amendment files should be given a nine-character file name in the form `amdXXXXXn.csv`, where XXXXX is the LSC institution identifier and n is a sequential number starting at 1.

Structure

4. All amendment files should contain the following information in the first 12 lines of the amendment file.

line 1 – amendment reference number in the form `amdxxxxxn`, where xxxxx is the five-character LSC institution identifier and n is a sequential number starting at 1 (i.e. this will be the file name with the ‘.csv’ file extension removed).

line 2 – creation date of the amendment file in the form `ddmmyyyy` (for example, 01032002 for a file created on 1 March 2002).

line 3 – brief description of the correction (for example corrections to ISR20 fields Q18 and SHE07).

line 4 – this line should be left blank.

line 5 – ISR field number(s) for the field(s) that uniquely identify records on the amendment file, comma-separated (for example, Q01, Q02).

line 6 – ISR field number(s) for the field(s) being corrected, comma-separated (for example, Q18, SHE07).

line 7 – number of records contained in the amendment file, excluding the first 12 lines containing file information.

line 8 – ISR field number for the field used to calculate the file check sum (see paragraph 11 for an explanation of the file check sum).

line 9 – file check sum.

lines 10 and 11 – these lines should contain any notes you wish to include.

line 12 – ordered ISR field numbers for all fields included in the amendment file, comma separated (for example, Q01, Q02, Q18, SHE07).

Correcting data

5. A denormalised record for each qualification aim requiring correction should be returned from line 13 onwards; that is, one line, including any relevant student data, for each ISR record being corrected. Each record should contain complete data for all fields included in the amendment file. For example, in an amendment file containing corrections for the ISR fields SHE09 and SHE10, an entry for SHE10 should be included even if SHE10 is not corrected (albeit the same SHE10 entry that was returned on ISR20 to the LSC). Alternatively, institutions may submit two amendment files, one containing corrections to SHE09 and another containing corrections to SHE10.

6. Only ISR fields should be corrected. For example, to change the HEIFES FTE field, HEFESFTE, corrections to SHE13 should be submitted.

Uniquely identifying records

7. To enable us to merge amendment files to our master ISR data set, we must be able to identify each record on the amendment file in such a way that it uniquely identifies the record in ISR20. The field, or combination of fields, enabling us to achieve this should be listed, comma-separated, in line 5. We recommend institutions use the following combination of seven fields to uniquely identify ISR records:

- S01
- Q01
- Q02

- S17
- S18
- Q24
- Q25.

8. Some institutions have indicated that they are not able to re-generate characters 20-27 of S01. Exceptionally, we will accept the first 19 characters of S01 to enable, as part of a combination of fields, the unique identification of records. Institutions should refer to the first 19 characters of S01 as ST_REF. This is the only derived field we will accept in an amendment file.

Validation

9. We will use the LSC's validation software to ensure that all corrections are valid and do not cause validation failures elsewhere in our master ISR data set. We will ask institutions to re-submit corrections if validation failure occurs.

10. Saving amendment files in Microsoft Excel may result in the loss of leading zeros and the corruption of very large values. We recommend that amendment files are saved in Notepad.

Example of an amendment file

Check sums

11. To ensure amendment files have not been corrupted during transit, we will check that the sum of values in a field is equal to the same calculation made by the institution before transit. The choice of which field has been used to calculate the check sum should be returned in line 8. The sum of the values in this field should be returned in line 9. If an amendment file does not contain any fields suitable for calculating a check sum (i.e. numeric fields), an additional field from ISR20 should be included solely for this purpose (for example, SHE10).

Outcome

12. When we receive a valid amendment file in the structure and format detailed in this annex, we will provide feedback on the impact within five working days. Institutions will be notified by e-mail whether a further response is required, and when revised re-creation tables and individualised file are available from our web-site.

```

amdHEFCE1.csv - Notepad
File Edit Search Help

amdHEFCE1,.....
01032002,.....
Corrections to ISR20 Fields Q18 and SHE07,.....

S01,Q01,Q02,S17,S18,Q24,Q25,.....
Q18,SHE07,.....
10,.....
Q02,.....
2406024,.....
Amendment submitted by Scott Chu.
Working papers stored in File 'ISR-HEIFES01'.
S01,Q01,Q02,S17,S18,Q24,Q25,SHE07
HEFCE00HE00000000110120100,HEFCE00HE00000000130120100,00240602,HE1,,HNDAnimalMan,,30072002,01
HEFCE00HE000000000210000100,HEFCE00HE00000000230000100,00240602,HE2,,HNDAnimalMan,,30072002,01
HEFCE00HE000000000310000100,HEFCE00HE00000000330000100,00240602,HE3,,HNDAnimalMan,,20052002,01
HEFCE00HE000000000410100100,HEFCE00HE00000000430100100,00240603,HE4,,HNCAnimalMan,,20052002,01
HEFCE00HE000000000510070100,HEFCE00HE00000000530070100,00240602,HE5,,HNDAnimalMan,,20052002,01
HEFCE00HE000000000610130200,HEFCE00HE00000000630130100,00240602,HE6,,HNDAnimalMan,FirstAid,30072002,01
HEFCE00HE000000000710030100,HEFCE00HE00000000730030100,00240602,HE7,,HNDAnimalMan,,30072002,01
HEFCE00HE000000000810000100,HEFCE00HE00000000830000100,00240603,HE8,,HNCAnimalMan,,30072002,01
HEFCE00HE000000000910110100,HEFCE00HE00000000930110100,00240603,HE9,,HNCAnimalMan,,30072002,01
HEFCE00HE000000001010100100,HEFCE00HE000000001030100100,00240603,HE10,,HCDAnimalMan,,30072002,01

```

Annex D

Obtaining data from the HEFCE web-site

1. This annex contains details on how to access the individualised file indXXXXXX.csv from our web-site along with an electronic copy of the re-creation tables (heiXXXXXX.xls).
2. These data can be accessed at the web address <http://datacollections.hefce.ac.uk>.
3. On viewing the password screen shown below, institutions should enter the username and password supplied in Annex A of the 'ISR 2000-01 derived statistics for funding allocations and monitoring in FECs' letter sent to heads of institution on 14 March 2002.



4. After verifying the username and password, your browser will be directed to the 'HEFCE Download System' page where institutions will have access to their data files.
5. Files can be accessed via the ISR-HEIFES link. A list of the available files and descriptions of their content will be displayed. Click on the required file name to download a file. Individualised files are usually large so it may take a while to download all the data.
6. Institutions are reminded that the individualised data are covered by the Data Protection Act. Therefore if there are concerns that the password has been revealed to an unauthorised member of staff within the institution, we will be happy to change the password. Please contact Scott Chu on 0117 931 7338 for details.