

July 2008/25

Core funding/operations

Assurance and funding

This report is for information and guidance

This document is a guide to HEFCE's web facility which generates several types of analysis from the Learning and Skills Council's Individualised Learner Record (ILR). Use of the web facility will help further education colleges to return accurate higher education data to the Learning and Skills Council, prepare for the 2009 National Student Survey and to identify errors and forecasting discrepancies in their student data return to HEFCE.

2007-08 statistics derived from ILR data

Guide to HEFCE web facility

Contents

Executive summary.....	4
Introduction	7
Annex A – Summary of changes since publication of HEFCE 2007/13	13
Annex B – Using the HEFCE web facility.....	18
Annex C – Comparison of HEIFES07 and the HEIFES07 re-creation	25
Annex D – Derived statistics that may inform the 2009-10 widening participation allocations	28
Annex E – 2007-08 co-funded employer engagement student numbers (CFEE07) re-creation	34
Annex F – Information for leads and members of HEFCE-recognised funding consortia	36
Annex G – 2009 Unistats summaries.....	41
Annex H – 2009 National Student Survey (NSS) target list	44
Annex I – Lifelong Learning Network (LLN) student summaries	46
Annex J – Submitting overrides to derived fields	47
Annex K – HEIFES08 non-completion toolkit.....	52
List of abbreviations	54

Appendices (see separate downloads)

Appendix 1	HEIFES07 re-creation algorithms
Appendix 2	Troubleshooting the differences between HEIFES07 and the HEIFES07 re-creation
Appendix 3	Problems of fit with the HEIFES07 re-creation algorithms
Appendix 4	Derived statistics that may inform the 2009-10 widening participation allocation algorithms
Appendix 5	Co-funded employer engagement 2007-08 (CFEE07) re-creation algorithms
Appendix 6	2009 Unistats data algorithms
Appendix 7	Generating 2009 Unistats outputs
Appendix 8	2009 National Student Survey (NSS) target list algorithms
Appendix 9	Lifelong Learning Network (LLN) summary algorithms
Appendix 10	HEIFES08 non-completion toolkit algorithms

2007-08 statistics derived from ILR data

Guide to HEFCE web facility

To	Heads of further education colleges directly funded by HEFCE Heads of lead institutions of HEFCE-recognised funding consortia Heads of member colleges of HEFCE-recognised funding consortia
Of interest to those responsible for	Student data, Audit, Finance
Reference	2008/25
Publication date	July 2008
Enquiries to	For enquiries on the use of ILR data to inform the 2009-10 widening participation allocations, contact: Christine Daniel tel 0117 931 7373 e-mail ilr_heifes_stats@hefce.ac.uk For all other enquiries, contact: Marie-Helene Nienaltowski tel 0117 931 7493 Lydia Booker tel 0117 931 7338 e-mail ilr_heifes_stats@hefce.ac.uk

Executive summary

Purpose

1. This document is a guide to HEFCE's web facility, which further education colleges (FECs) should use to check data for higher education (HE) students on the Learning and Skills Council (LSC) Individualised Learner Record (ILR). It is expected that colleges will use this facility before submitting their 2007-08 ILR F04 and F05 returns to the LSC. It generates the following derived statistics:
 - a. Higher Education in Further Education: Students Survey 2007-08 (HEIFES07) re-creation.
 - b. Derived statistics that may be used to inform the 2009-10 widening participation (WP) allocations.
 - c. Re-creation of the aggregate return to monitor 2007-08 co-funded employer engagement student numbers (CFEE07).
 - d. Summaries of data for the Unistats web-site.
 - e. A target list of students to be included in the 2009 National Student Survey (NSS).
 - f. Lifelong Learning Network (LLN) student summaries.
 - g. A non-completion toolkit for the Higher Education in Further Education: Students Survey 2008-09 (HEIFES08).

Key points

2. The use of the web facility is strongly encouraged because it will help colleges to:
 - return accurate data to the LSC
 - reduce the likelihood of selection for the derived statistics exercise for 2007-08 (selected colleges are typically subject to considerable additional work and potential funding adjustments)
 - identify discrepancies between forecasts in HEIFES07 and the outturn position for 2007-08 as generated from their ILR
 - identify errors in HEIFES07
 - prepare for the 2009 NSS
 - ensure that Unistats summaries are fit for publication in 2009
 - ensure that LLN student summaries are accurate
 - make informed estimations of non-completion rates for HEIFES08.
3. This document (including the technical appendices which can be downloaded separately) provides:
 - guidance on using the web facility
 - details of the algorithms used for generating the derived statistics
 - guidance on troubleshooting the differences between HEIFES07 and the HEIFES07 re-creation
 - details of problems of fit with HEIFES07 re-creation algorithms

- guidance for submitting overrides to our algorithms, and information on when this is appropriate
- guidance for using the HEIFES08 non-completion toolkit.

Action required

4. HEFCE-funded colleges and members of HEFCE-recognised consortia are required to use the web facility to generate 2009 NSS target lists. Further details about the arrangements for the 2009 NSS will be published later this summer.
5. More generally, we strongly encourage colleges to use the web facility as part of their data quality processes. Past reviews have confirmed that it is an essential element of most HEFCE-funded colleges' ILR data quality processes, and that colleges that do not use the facility are more likely to be selected for the monitoring exercise described in paragraphs 7 and 8.
6. Colleges that are members of a HEFCE-recognised funding consortium should also use this web facility to ensure that ILR data errors do not cause their consortium to be selected for our monitoring exercises.

Relationship with 'Statistics derived from ILR data'

7. Each year we publish 'Statistics derived from ILR data for the monitoring and allocation of funding in FECs' (see HEFCE 2008/08) in which we monitor FECs' HEIFES and CFEE returns using ILR data which we receive from the LSC. If we find, either through comparisons with ILR data or any data audit, that erroneous data have resulted in colleges receiving incorrect funding allocations (including for WP) then we will adjust their funding accordingly (subject to the appeals process and the availability of our funds).
8. Any funding adjustments following the comparison of 2007-08 ILR data with HEIFES07 or CFEE07 are likely to affect the funding previously announced for 2007-08 and all subsequent years, and WP funding for 2008-09. Colleges selected to make a response to the exercise must typically undertake a substantial amount of work to reconcile their data, which may take several months to complete. The web facility is provided to complement the monitoring exercise, but it does not replace it.

New outputs

9. We will collect the CFEE07 return in August 2008 to inform funding for co-funded employer engagement. In the meantime colleges will be able to use the web facility to produce a CFEE07 re-creation based on 2007-08 ILR data. The CFEE07 re-creation is a new output that has not been produced by past web facilities. It provides colleges with an early opportunity to verify that 2007-08 ILR data are accurate for their co-funded

employer engagement students. Upon receipt of 2007-08 ILR data from the LSC later this year we will use the CFEE07 re-creation to monitor the CFEE07 return.

Introduction

10. This document provides guidance on using HEFCE's web facility and its outputs. The primary purpose of the web facility is to help further education colleges (FECs) to return accurate Individualised Learner Record (ILR) data to the Learning and Skills Council (LSC). It provides colleges with an opportunity to identify, and therefore rectify, any errors in data that affect the outputs generated by the web facility, before these data are submitted to the LSC.

11. Both the LSC and HEFCE strongly encourage colleges to use this facility before submitting their ILR return. Both organisations regard it as an essential element of HEFCE-funded FECs' data quality processes for higher education (HE) students.

12. Last year almost 150 FECs used the web facility, which is a significant majority of HEFCE-funded FECs and member colleges of HEFCE-recognised funding consortia. We believe that since the introduction of the web facility there has been a significant improvement in ILR data quality for HE students. In previous years, colleges that did not use the facility before submitting their F04 return to the LSC were more likely to be selected for the derived statistics exercise later in the year. This typically generated a significant amount of additional work for them.

13. Use of the web facility can help identify errors and discrepancies between the forecasts made in the HEIFES07 and the outturn position for 2007-08 as recorded in the 2007-08 ILR. Where discrepancies occur, we expect colleges to take full account of the outputs from this facility when preparing future HEIFES returns. We encourage colleges to analyse the web facility outputs as part of their planning and audit processes.

14. Changes since the 2007 version of the HEFCE web facility (HEFCE 2007/13) are described in Annex A.

15. The web facility generates seven outputs. These are:

- a HEIFES07 re-creation
- derived statistics that may inform the 2009-10 allocations for widening participation (WP)
- a CFEE07 re-creation
- summaries of data for the Unistats web-site to be published in 2009
- a target list of students to be included in the 2009 National Student Survey (NSS)
- summaries for students taught as part of a Lifelong Learning Network (LLN)
- a non-completion toolkit for the Higher Education in Further Education: Students Survey 2008-09 (HEIFES08).

Using the web facility

16. The web facility allows colleges to upload their ILR data to a server at HEFCE. Instructions on how to use the web facility are provided in Annex B.

17. Typically colleges submit their data to the web facility and retrieve the derived statistics several times before all identified errors are remedied. Therefore, we will not restrict colleges' use of the facility. However, users should be aware that response times may be slower when there is high demand.

18. Colleges should allow adequate time to make full use of the web facility without jeopardising ILR quality arrangements or timetables.

19. Generally, we will not view ILR data submitted to, or outputs generated by, the web facility unless permission is given by the college. However, HEFCE staff will access outputs relating to Unistats summaries and NSS target lists. We will monitor use of the web facility in order to offer assistance and to promote the facility to colleges who have not used it.

HEIFES07 re-creation

20. The web facility generates a HEIFES07 re-creation from 2007-08 ILR data submitted by the college. This includes the calculation of formulaic adjustments to teaching grant and WP allocations. This output is coupled with a copy of the original HEIFES07 outputs for comparison and reconciliation.

21. The algorithms used to generate the re-creation are intended to be the same as those to be used for the 'Statistics derived from ILR data for the monitoring and allocation of funding in FECs' exercise (see HEFCE 2008/08). We may, however, make changes to improve the algorithms. The HEIFES07 re-creation is generated using the methods described in Annex C and the algorithms in Appendix 1.

22. We strongly advise colleges to use this opportunity to identify the cause of all discrepancies between their HEIFES07 return and the HEIFES07 re-creation, so that where errors are caused by ILR data, these can be corrected before submitting the F04 and F05 returns to the LSC. The removal of such errors from 2007-08 ILR data will reduce the likelihood of colleges being selected for the 2007-08 'Statistics derived from ILR data' exercise (see paragraphs 23 and 24 below).

23. Each year we publish 'Statistics derived from ILR data for the monitoring and allocation of funding in FECs' (see HEFCE 2008/08) in which we monitor FECs' HEIFES and CFEE returns using ILR data which we receive from the LSC. If we find, either through comparisons with ILR data or any data audit, that erroneous data have resulted in colleges receiving incorrect funding allocations (including for WP) then we will adjust

their funding accordingly (subject to the appeals process and the availability of our funds).

24. Any funding adjustments following the comparison of 2007-08 ILR data with HEIFES07 or CFEE07 are likely to affect the funding previously announced for 2007-08 and all subsequent years, and WP funding for 2008-09. Colleges selected to make a response to the exercise must typically undertake a substantial amount of work to reconcile their data, which may take several months to complete. The web facility is provided to complement the monitoring exercise, but it does not replace it.

Derived statistics that may inform 2009-10 WP allocations

25. 2007-08 ILR data may be used to inform the following WP allocations for 2009-10:

- widening access for students from disadvantaged backgrounds: full-time and part-time
- improving retention: full-time
- widening access and improving provision for disabled students.

26. The derived statistics used to inform these allocations may be generated using the methods described in Annex D and the algorithms in Appendix 4.

CFEE07 re-creation

27. The web facility generates a CFEE07 re-creation from 2007-08 ILR data submitted by the college. Where appropriate this includes an indication of the funding associated with the recruited full-time equivalences (FTEs) and the difference in funding between the CFEE07 re-creation and the original co-funded employer engagement allocation.

28. The CFEE07 return will not be collected until August 2008; however, the web facility provides an early opportunity to verify that the 2007-08 ILR data are correct for these students.

29. The algorithms used to generate the re-creation are intended to be the same as those to be used for the 'Statistics derived from ILR data for the monitoring and allocation of funding in FECs' exercise. We may, however, make changes to improve the algorithms. The CFEE07 re-creation is generated using the methods described in Annex E and the algorithms in Appendix 5.

2009-10 Unistats summaries

30. We intend to publish Unistats summaries from 2007-08 ILR F05 data in 2009. The web facility provides an opportunity for colleges to check, before submission, that their ILR F05 data are fit for publication on the Unistats web-site. The methods used to generate these statistics are described in Annex G and the algorithms are given in

Appendix 7. Any comments on the suitability for publication of colleges' Unistats summaries, or the method used to generate them, should be sent to Marie-Helene Nienaltowski (e-mail ilr_heifes_stats@hefce.ac.uk) by **Friday, 19 December 2008**.

31. From February 2009 HEFCE staff may view Unistats summaries generated by the web facility to investigate any general issues of data quality.

2009 NSS target lists

32. As for the 2008 survey we will use 2007-08 F04 ILR data to produce a target list of students to be included in the 2009 NSS. Colleges are encouraged to use the list to start preparing the contact details for these students so that these can be passed to the agency running the NSS. Contact details for students to be included in the 2009 NSS will need to be returned to the agency in late November 2008. We will extract a final target list for the 2009 NSS from data submitted to the web-facility on 14 November 2008, colleges should ensure that the most recent data submitted to the web-facility on this date generates a complete NSS target list.

33. Further details about the arrangements for the 2009 NSS will be published later this summer. Further details about the target lists are in Annex H and the algorithms used to generate the lists are given in Appendix 8. HEFCE staff will access NSS target lists generated by the web facility."

Destinations of leavers from HE

32. We intend to survey recent graduates from FECs in order to establish what they are doing roughly six months after completion of their courses. In order to facilitate this we will require colleges to provide the most recent contact details for recent graduates. We will publish details of this process in the autumn. Colleges should inform their students that this will happen and allow them the option to opt out of this survey. Colleges should not rely on the general consent to survey work given in field L27 of the ILR but instead contact students regarding this survey in particular.

LLN student summaries

33. For colleges participating in an LLN during 2007-08 we intend to identify LLN students using 2007-08 ILR data. These data will be used as part of our evaluation of the LLN initiative. The web facility generates summaries for LLN students to enable colleges to verify that we are identifying their LLN students correctly. The method for generating these statistics is described in Annex I and the algorithms are given at Appendix 9.

HEIFES08 non-completion toolkit

34. The web facility provides colleges with a spreadsheet (toolkit) to assist them in estimating non-completion rates for HEIFES08. A detailed explanation of how to use the toolkit is given in Annex K.

Data confidentiality

35. The web facility is a tool for colleges. Therefore, HEFCE staff will not access data uploaded to, or derived by, the web facility except in the limited circumstances described below. Only those fields detailed in the appendices will be retained on our system. We will remove individualised files and any outputs from the web facility shortly after the launch of next year's web facility. Uploaded data files are deleted immediately after the outputs have been generated. Details about our commitments to handling personal data can be found on the HEFCE web-site, www.hefce.ac.uk, under About us/Standards of service/How we do it – our commitments.

36. HEFCE staff may examine outputs from the web facility in any of the following circumstances:

- where a college explicitly gives permission
- to study general data quality and algorithmic issues with the Unistats summaries
- to assess the suitability of NSS target lists.

HEFCE-recognised funding consortia

37. Member colleges of HEFCE-recognised funding consortia are strongly encouraged to use this web facility before submitting their 2007-08 ILR data to the LSC. The removal of any errors in their ILR data reduces the likelihood that their consortium will be selected for the 2007-08 monitoring exercises.

38. The web facility outputs that will be made available to leads and members of HEFCE-recognised funding consortia are described in Annex F. Annex F also describes how some data submitted by member colleges will be incorporated in outputs generated for the lead.

Guidance

Instructions on how to use the web facility and its outputs

39. Full instructions on how to get started with the web facility and how to access the outputs can be found in Annex B.

Troubleshooting

40. Appendix 2 has been produced to help colleges identify the causes of discrepancies between 2007-08 ILR data and the HEIFES07 re-creation.

FAQs

41. Frequently asked questions (FAQs) regarding the web facility are on the HEFCE web-site, www.hefce.ac.uk, under Learning & teaching/Data collection/HEFCE web facility for 2007-08 statistics derived from ILR data. We encourage colleges to refer to the FAQs for guidance in the first instance. We will only use our e-mail list of HEIFES contacts to notify colleges of significant changes or updates, not for routine additions to the list of FAQs.

HEFCE contacts

42. Support is available for the web facility via telephone and e-mail; see page 2 for contact details. This will be available between 0900 and 1700, Monday to Friday.

Annexes and technical appendices

43. The annexes provide guidance on using the web facility and its outputs. The technical appendices describe the algorithms used to generate the derived statistics. The appendices are available as separate downloads on the HEFCE web-site, with this document, under Publications.

Comments

44. All colleges are invited to comment on any of the methods described in this publication. Comments should be e-mailed to ilr_heifes_feedback@hefce.ac.uk marked for the attention of Marie-Helene Nienaltowski.

Annex A – Summary of changes since publication of HEFCE 2007/13

Purpose

1. This annex describes the changes that have been made between the release of last year's web facility for Individualised Learner Record (ILR) data, HEFCE 2007/13, and this year's facility.

New output

2. The web facility now generates a CFEE07 re-creation using 2007-08 ILR data submitted by the college. A description of the re-creation is given in Annex E and the CFEE07 re-creation algorithms are explained in Appendix 5.

Indicative 2009-10 WP outputs

3. The indicative 2009-10 widening participation (WP) outputs (including an indicative 2009-10 WP allocation calculation) are based on rates and full-time equivalences (FTEs) used for the 2008-09 WP allocation made in March 2008. Where possible, the calculations also incorporate any subsequent changes arising, for example, from amendments to 2008-09 additional student numbers (ASNs) and transfers. During 2008 and 2009 we may update the rates and FTEs used by the web facility for the indicative WP outputs as more current information becomes available.

4. The indicative WP allocations now treat students studying for an equivalent or lower qualification (ELQ) as non-fundable. The provisional assumed FTEs for 2008-09 and the statistics derived from 2007-08 ILR data which inform the indicative 2009-10 WP allocations have been adjusted to reflect this change.

5. We have made some changes to the measures used to inform the widening access allocations. These changes are described in paragraphs 6 to 9 below. They are a necessary update to the underlying data and do not represent a change in the allocation method.

6. The young participation rates for small areas that underpin the widening access allocation for young full-time students have been updated. Previously we used participation rate quintiles based on the participation in higher education (HE) of people who reached 18 between 1997 and 1999. These rates were calculated for the 1991 Census small-area statistics wards geography. We have updated the participation quintiles to cover young people who reached 18 between 2000 and 2004, and changed the underlying geography used to calculate the rates to 2001 Census area statistics wards. This classification of areas covers the UK, and the quintiles are formed with reference to the UK young population.

7. We have also made some adjustments to the area groupings based on educational qualification levels used for the widening access allocation for mature full-time, young part-time and mature part-time students. These groupings now cover 2001 Census area statistics wards throughout the UK, rather than just England and Wales, though the ranges for the quintiles are determined with reference to the English population only.

8. The postcode to 2001 Census area statistics wards look-up file that is used in the allocation has been updated to reflect the August 2007 release of the National Statistics Postcode Directory (NSPD), which is maintained by the Office for National Statistics (ONS). This look-up file may be updated as more current information becomes available. Changes to the way we calculate young participation rates have led to a revision of the postcodes that are excluded from the allocation because we cannot obtain robust participation rates. Additionally, for the purposes of the widening access allocations, we have formally excluded postcodes that the NSPD indicates as non-geographic (mostly large user and PO boxes).

9. The assignment of postcodes to the young participation and educational attainment quintile groupings are provided in a file on the HEFCE web-site under Widening participation/POLAR and participation rates/POLAR2. This file includes postcodes which are excluded from the quintile mapping along with the reason for exclusion.

10. Guidance on how to generate the outputs from the individualised files for WP can be found in Appendix 4.

Changes to algorithms and field names

11. The changes to algorithms listed here are those that have changed since HEFCE 2007/13.

HEIFES07 re-creation

12. The algorithm for identifying Home and EC non-fundable students has changed to include new members to the EC (Bulgaria and Romania), overseas territories (a list of overseas territories affected can be found on the HEFCE web-site, www.hefce.ac.uk, under Questions/HEIFES FAQs) and the new code values for country of domicile, ST_DOMIC (L24), which affect Cyprus, France, Italy and Spain.

13. The algorithm for assigning price groups has been improved. The approximation based on Superclass II codes is replaced for a direct mapping between Learndirect codes and price group. Price group is assigned by mapping the three Learning Directory Classification System (LDCS) fields, LDCS_CO1 – LDCS_CO3, to price groups as indicated in the table in below. For Initial Teacher Training (ITT) and In-Service Education and Training (INSET) students the distribution based upon LDCS codes is not used and all activity is assigned to the ITT and INSET price groups respectively.

Field name	Learndirect code (LDCS_CO1, LDCS_CO2, LDCS_CO3)	Value*
PRGB	M [‡] , NL [‡] , NM [‡] , PB, PC.1, PC.5, PE.6, PE.7, PF.1, PF.2, PG.1, PG.2, QA.3, QH.6, R (except [†] RA.3, RA.5, RA.6, RB, RF.4, RG), S (except [†] SE, SF, SJ.5, SJ.61, SM, SN.4, SP, SQ), TL, TM, VE, VF.4, VG, WA, WC.1, WC.2, WC.3, WC.4, WE, X (except [†] XA, XD, XE, XF, XN, XS, XQ.45), Y (except [†] YA, YB, YD.3)	sum of HQ_PERSX/100
PRGC ^{ψφ}	C (except [†] CY.3, CY.4, CY.6, CY.7, CY.8, CY.9, CZ), DC, FN.3, FN.4, FN.5, FN.6, FN.7, FN.9, G [†] , J (except [†] JA.11, JA.22, JA.23, JA.32, JA.33, JA.34, JA.5, JA.7, JA.8, JD, JE), L (except [†] LF, LG), M [‡] , N [‡] (except [†] NG, NK, NN), P (except [†] PA, PB, PC.1, PC.5, PE.6, PE.7, PF.1, PF.2, PG.1, PG.2), Q (except [†] QA.3, QB, QH, QJ), RA.3, RA.5, RA.6, RB, RF.4, RG, SE.1, SN.4, SQ, T (except [†] TC.44, TC.5, TC.6, TF, TL, TM), VF.1, VF.2, VF.3, VF.5, VF.6, W (except [†] WA, WC.1, WC.2, WC.3, WC.4, WE), XA, XD, XE, XF, XN, XS, YA, YB, YD.3, Z (except [†] ZX.3, ZX.4, ZX.5)	sum of HQ_PERSX/100
PRGD	A, B, D (except [†] DC), E, F (except [†] FN.3, FN.4, FN.5, FN.6, FN.7, FN.9), G ^ψ , H, JA.11, JA.22, JA.23, JA.32, JA.33, JA.34, JA.5, JA.7, JA.8, JD, JE, KB, KC, LF, LG, M [‡] , NG, NK, NL [‡] , NM [‡] , NN, PA, QB, QH.1, QH.2, QH.3, QH.4, QH.5, QH.7, QH.8, QH.9, QJ, SE.2, SE.3, SE.4, SE.5, SE.7, SE.8, SE.9, SF, SJ.5, SJ.61, SM, SP, TC.44, TC.5, TC.6, TF, U, V (except [†] VE, VF.1, VF.2, VF.3, VF.4, VF.5, VF.6, VG), XQ.45, ZX.3, ZX.4, ZX.5	sum of HQ_PERSX/100
PRGMEDIA	CY.3, CY.4, CY.6, CY.7, CY.8, CY.9, CZ, K (except [†] KB, KC)	sum of HQ_PERSX/100
PRGITT	College specific	1
PRGINSET	QA_FEHE1 (A11A) = 025 or QA_FEHE2 (A11B) = 025	1

*Where HQ_PERSX is HQ_PERS1 (H33), HQ_PERS2 (H34), HQ_PERS3 (H35).

[†] Including all sub-levels of the hierarchy.

[‡] Students on sports science courses with LDCS codes MA to MJ, NL and NM should be allocated to price groups B, C or D according to the outcome of the HEFCE review carried out in 2004-05 or as subsequently agreed.

^ψ All Certificate of Education activity (HEFQAIM = CERTED) is assigned to price group C.

^φ All students on a sandwich year-out are assigned to price group C.

Widening participation

14. As a result of the change described in paragraph 7 of this annex, the algorithms for indicating inclusion in the mature full-time widening access population (EDPOPM) and inclusion in the part-time widening access population (EDPOPPT) have been modified to include students domiciled in Scotland and Northern Ireland (ST_DOMIC (L24) = 099, 299). Details of this change can be found in paragraphs 44 and 45 of Appendix 4.

15. We are now able to map directly from postcode (ST_POSTC (L17)) to the young higher education participation quintile and the adult educational attainment quintile (WAQUIN), so we have removed the derived fields displaying the 1991 Census ward (WARD6_C) and 2001 Census ward (CASWARD) of the student's home postcode from the WP individualised file.

16. Two new groupings of highest qualification on entry for use in the full-time improving retention allocation have been created which identify separately students entering with ordinary national certificate (ONC) or ordinary national diploma (OND) (including Business and Technology Education Council (BTEC) and Scottish Qualifications Authority (SQA) equivalents) or entering with no formal qualifications. (ENTQUAL = BTEC and ENTQUAL = NONE) Users should note that this change will not affect the funding allocation; it is solely for technical purposes. Details of this change can be found in paragraph 11 of Appendix 4.

17. Students in the full-time improving retention population (EQPOP ≠ 0) whose highest qualification on entry is a National Vocational Qualification (NVQ), Scottish Vocational Qualification (SVQ) at level 5 or a graduate equivalent qualification not elsewhere specified (H11 = 15 or 16) are now placed into the 'degree and higher' grouping of highest qualifications on entry (ENTQUAL = DEG) instead of the 'other HE' grouping of highest qualification on entry (ENTQUAL = OHE) as in previous years. As a result of this change such students are now placed into a low risk group (EQGRP = Y_L or M_L) instead of a medium risk group (EGGRP = Y_M or M_M). Details of this change can be found in paragraph 11 of Appendix 4.

18. As a result of the change described in paragraph 4 of this annex, the derived fields that indicate inclusion in the widening access and full-time improving retention populations (YNGPART, EDPOPM, EDPOPPT and EQPOP) have been modified so that students who are studying for an ELQ are not included. Additionally for a student in any of these populations where the level of entry qualifications is not known, an appropriate proportion is removed from the field indicating inclusion in the population. Appendix 4 provides further detail on the changes to these fields and also provides information on how we identify which students are studying for an ELQ, and which students have an unknown-level entry qualification (ELQQENT = UNK). It also explains how we have calculated the proportion that should be treated as non-fundable (ELQ_PROP) for students with unknown-level entry qualifications. Note that the field

which indicates inclusion in the disability allocation population (DISPOP) has not changed as home and EC students who are classed as non-fundable (HEFTYPE = HOMENF) may be included in this population.

19. The derived fields which indicate inclusion in each of the widening participation populations (YNGPART, EDPOPM, EDPOPPT, EQPOP and DISPOP) have been modified to reflect the addition of two new UK-related country of domicile codes (ST_DOMIC (L24) = 907 and 984) to the ILR. Details of this change can be found in paragraphs 43, 44, 45, 47 and 36 respectively of Appendix 4.

20. Three new values of highest qualification on entry have been introduced to the ILR record for 2007-08: HQ_QUAL_ (H11) = 31 (Foundation Degree), 57 (NVQ/SVQ Level 2) and 72 (Diploma in Foundation Studies (Art and Design)). Consequently we have updated our algorithms for the flag indicating whether or not a student has previously studied for their qualification aim, or a higher qualification aim (HIGHQUAL) and the grouping of the student's highest qualification on entry (ENTQUAL). Details of this change can be found in paragraphs 41 and 49 respectively of Appendix 4.

Annex B – Using the HEFCE web facility

Purpose

1. This annex contains detailed instructions on how to use the web facility.

Data preparation

2. The web facility can process data files that are structured in the format described in the Learning and Skills Council (LSC) publication 'Specification of the Individualised Learner Record for 2007/08', which can be downloaded from the LSC's web-site, www.lsc.gov.uk under Providers/Data/Data collection/ILR.
3. The web facility has been designed to provide outputs even when it has been supplied with invalid data, so that colleges can use the facility as early as possible in their data collection process. Every effort has been made to ensure the web facility is robust. However, it may not function reliably when the input file contains invalid data. To ensure reliable outputs are generated, colleges should first correct any errors identified by running the Learner Information Suite software on their data (www.lsc.gov.uk under Providers/Data/Provider software/LIS).
4. The web facility can receive and process zip files, provided that the contents when unzipped are in the format specified by the LSC for the 2007-08 ILR.

Access to the facility

5. The web facility can be accessed via the HEFCE extranet at <https://extranet.hedata.ac.uk>. On viewing the screen below you will be prompted to either log in or register for access to the HEFCE extranet as a new user.

HE data - login - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Links

HE data - Login

Please enter your email address and chosen password, then login:

- if this is your first visit, please [register](#).
- if you have forgotten your password, [request a new one](#).
- you may want to read the [online help](#).

e-mail address :

password :

[Login via ATHENS SSO](#) If you have an AthensSSO account, you can use it to log in.

For support, please mail extranet@hefce.ac.uk. Please note that your browser must accept session cookies to use this site (most browsers do so by default).

Local intranet

Registering a new account

6. New users of the HEFCE extranet will first need to register an e-mail address and extranet password. This can be done by clicking the 'Register' link on the log-in screen. In order to register, you will require an 'organisation key' and a 'group key'. Details of these were sent to heads of institutions and our HEIFES contacts in July 2008, in the letter from Marie-Helene Nienaltowski titled 'HEFCE web facility for 2007-08 statistics derived from ILR data'. We can re-issue the information if necessary, but only to the head of institution or HEIFES contact.

7. Once registered you should be able to log in by entering the e-mail address you used during registration, and the password that you created.

Existing users of the extranet

8. If you have used the HEFCE extranet for other HEFCE returns (for example HEIFES07) you will be required to log in and join the data group for this HEFCE web facility. Follow the log-in procedure by entering your e-mail address and password. You will be directed to a page for 'HEFCE extranet – All resources'; under 'Applications' click 'Join a group'. Enter the group key supplied in Annex A of the letter from Marie-Helene Nienaltowski mentioned in paragraph 6 above, and select 'Join group'.

9. If you have registered in the past but your account has expired, you will be required to refresh your account using the organisation key referred to in paragraph 6 of this annex.

Athens Single Sign On account

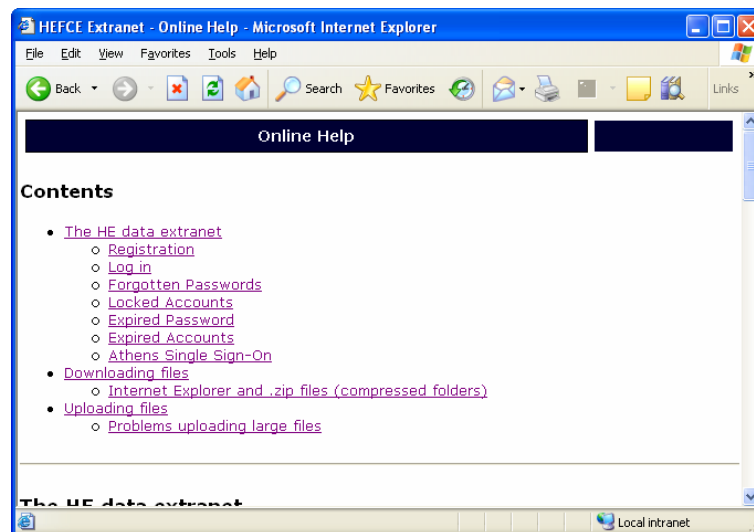
10. You can also log in to the HEFCE extranet using Athens Single Sign On (SSO) account, if this is available at your institution.

- a. Follow the 'Log in via Athens SSO' link on the log-in page.
- b. Log in to Athens as normal (if you have not already done so).

- c. When Athens has authenticated you, your browser will be directed to the 'HEFCE extranet - All resources' page.
- d. You will be required to join the group. Under 'Applications' click 'Join a group'. Enter the group key sent in the letter from Marie-Helene Nienaltowski referred to in paragraph 6 above, and select 'Join group'.

Help

11. An information page can be accessed by clicking the 'Online Help' link, located above the log-in box, or on the right of the 'HEFCE extranet – All resources' page, under 'Navigation'.



Forgotten password

12. If you forget your password, select 'Request a new one' on the log-in page. Enter your e-mail address in the space provided and click 'Request password'. A new password will be generated and sent to you. You will be required to change this password when you next access the HEFCE extranet.

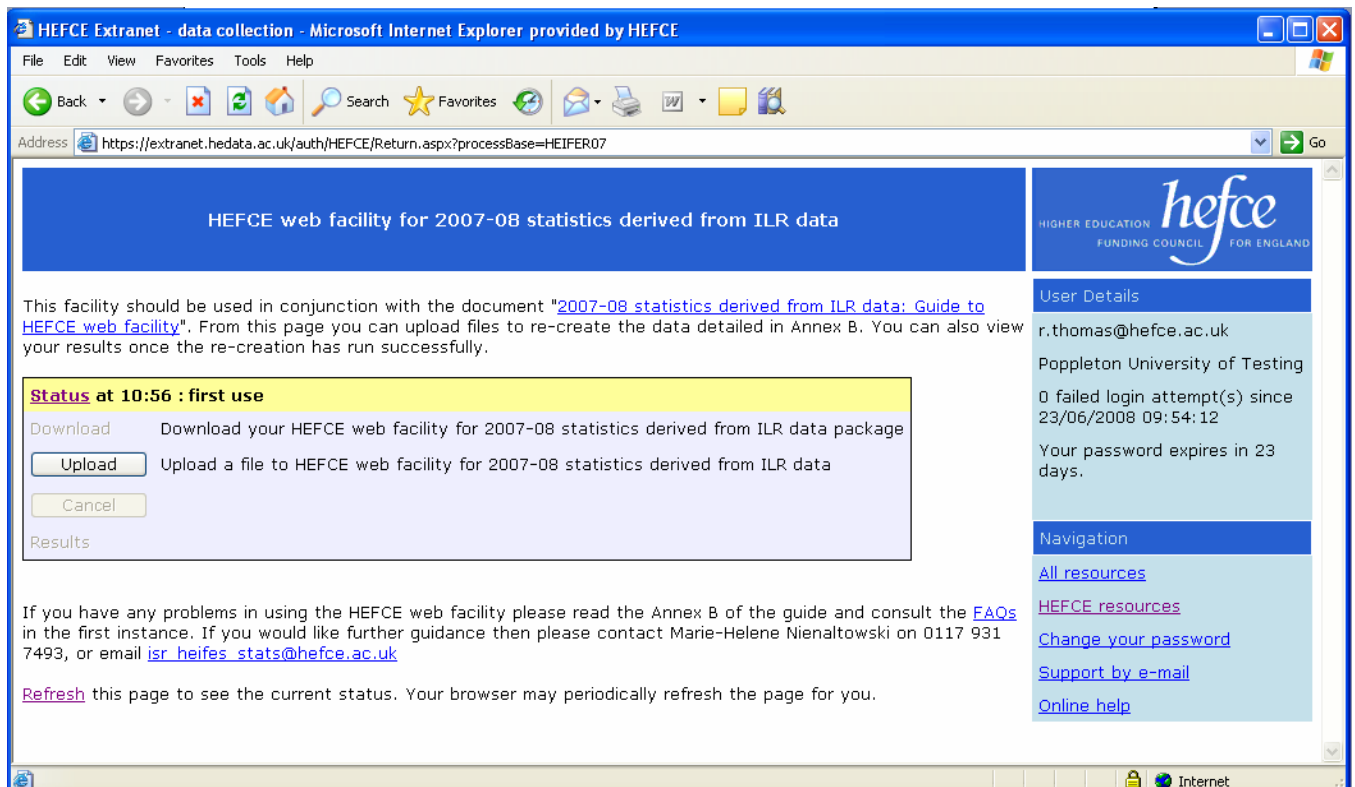
13. You have three attempts to log in before your account is locked. If you use all three attempts and still do not have access to the web facility, contact Marie-Helene Nienaltowski or Lydia Booker (e-mail ilr_heifes_stats@hefce.ac.uk).

All resources

14. After verifying the user name and password, your browser will be directed to the 'HEFCE extranet – All resources' page. Click 'HEFCE resources'.

15. From the 'Applications' list click on the link to the 'HEFCE web facility for 2007-08 statistics derived from ILR data'.

16. This option will allow you to upload the ILR file. Click the upload button, use the 'Browse' button to find the file and then select 'Upload file'. This places the file into the processing queue.



17. You can leave the extranet while the derived statistics are being generated. To do this you can exit the web browser in the usual way, or enter a new address in the address bar. You may be prompted to log in again when you return to the page, but no information will be lost.

18. Each time you upload a file for re-creation, you should refresh the page using either the refresh icon in the toolbar or the link on the page. Refreshing ensures that the page showing 'History of HEFCE web facility for 2007-08 statistics derived from ILR data submissions' reflects the current stage of processing, because it may not update automatically.

19. Our system allows only one submission to be processed at a time; therefore if more than one college uploads data, the submissions will be placed in a queue. A submission by an average-sized college typically takes 10 minutes to process (once it reaches the front of the queue). The total time taken to complete the re-creation, including waiting time, will depend on the number of colleges in the queue.

20. If the uploaded file is waiting in the queue it is possible to cancel the submission by selecting 'Cancel'.

21. If you do not use the extranet for 60 minutes, you will automatically be logged out of the system. To regain entry, select any option and re-enter your e-mail address and extranet password.

Outputs

22. Once the file has been processed, you can access the outputs by selecting 'Results'.

23. If the outputs were not generated, and the reason is known, it will be given under 'History of HEFCE web facility for 2007-08 statistics derived from ILR data submissions'. If the outputs were not generated because an unspecified error occurred, this will typically be a result of invalid ILR data (see paragraph 3 of this annex).

Accessing results

24. The results are made available as a zip archive containing the following files (where YYYYYY is the Unique Provider Identification Number (UPIN) for the college):

- HEIFER07YYYYYY.ind – HEIFES07 re-creation individualised student data file. For leads of HEFCE-recognised funding consortia this file does not include data for the member colleges
- HEIFER07YYYYYY.xls – HEIFES07 re-creation tables (HEIFERC07YYYYYY.xls for members of consortia). For leads of HEFCE-recognised funding consortia this includes the combined 2007-08 ILR data uploaded by both the lead and member colleges (see paragraph 25 below and Annex F)
- IHWP07YYYYYY.ind – 2009-10 WP individualised student data file. For leads of HEFCE-recognised funding consortia this file does not include data for the member colleges (see paragraph 25 below and Annex F)
- IHWP07YYYYYY.xls – 2009-10 WP allocation tables (IHWPC07YYYYYY.xls for members of consortia). For leads of HEFCE-recognised funding consortia this includes the combined 2007-08 ILR data uploaded for both the lead and member colleges
- CFEE07YYYYYY.ind - CFEE07 re-creation individualised student data file
- CFEE07YYYYYY.xls – CFEE07 re-creation tables
- TQI07YYYYYY.ind – Unistats individualised student data file
- TQI07YYYYYY.xls – Unistats tables

- NSSP09YYYYYY.ind – an individualised student data file containing the derived fields that generate the 2009 NSS target list
- NSSP09YYYYYY.xls – 2009 NSS target list
- LLN07YYYYYY.ind – an individualised student data file containing the derived fields that will be used to generate the LLN summary
- LLN07YYYYYY.xls – LLN student summary
- IHNCTK07YYYYYY.xls – HEIFES08 non-completion toolkit.

25. Lead colleges of HEFCE-recognised funding consortia will also be able to access the following files:

- LEAD07YYYYYY.xls – an Excel workbook containing the HEIFES07 re-creation tables for lead colleges of HEFCE-recognised funding consortia (includes 2007-08 ILR data uploaded by the lead further education college (FEC) only)
- HEIFERC07YYYYYY.xls – Excel workbooks containing the HEIFES07 re-creation tables for each of the member colleges (includes the 2007-08 ILR data uploaded by the member college only). These tables will only become available after member colleges have submitted data to the web facility
- IHWPLEAD07YYYYYY.xls – Excel workbooks containing the indicative 2009-10 WP tables for the lead college (includes 2007-08 ILR data uploaded by the lead FEC only)
- IHWPC07YYYYYY.xls – Excel workbooks containing the indicative 2009-10 WP tables for each of the member colleges (includes the 2007-08 ILR data uploaded by the member college only). These tables will only become available after member colleges have submitted data to the web facility.

Using the individualised files

26. The individualised files will show the allocation of students to cells within the equivalent re-creation tables and, where relevant, details of why they were excluded. The following guidance will assist colleges in the troubleshooting process. All records with a specific value in a specific field can be filtered and located in the following way:

- a. To open an individualised file, from the list above in Microsoft Excel, load Excel and then click <File>, <Open>. You will need to specify 'All files' in the 'Files of type' box before the individualised file will appear in the list. Once you have selected the file, the 'Text Import Wizard' will appear. Ensure that 'Delimited' is

selected near the top of the window, then click 'Next'. On the next page, uncheck 'Tab' and check 'Comma'. Click 'Finish' to open the file.

- 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 that you want to filter.
- e. Either select a specific value or select <Custom> to apply a comparison operator other than equality.
- f. To select records using multiple fields, repeat steps d and e for each field.

Using the output

27. Owing to a limitation of Excel, institutions who submit more than 65,530 records will find that they are unable to view the entire individualised file in Excel. At this point it may be useful to check whether the tables reflect the number of students you would expect to see. If they do not, please check that the file is the one that you intended to upload. We are not responsible for the accuracy of the files uploaded to the web facility.

28. The output files will be overwritten each time new data are uploaded and therefore you will need to make a copy of an output if you are likely to need it for future reference.

Technical issues

29. Although there are no limits on the size of files that can be uploaded to the facility, it is common for individual colleges to have limits imposed by their network administrators. Attempts to upload a file beyond the size limit imposed within the college are unlikely to generate a meaningful error message. They will, however, cause the browser to 'stall' or 'hang'.

30. If you are encountering difficulties in uploading due to file size or any other issue relating to connectivity, we will normally be able to advise. If you are not submitting your data as a zip file, we suggest that you try this first, to reduce the size of the transmission. We recommend that you involve your network administrator in discussions regarding connectivity issues, and that you tell them that the Internet Protocol (IP) address for the web facility is 195.194.167.114.

Annex C – Comparison of HEIFES07 and the HEIFES07 re-creation

Purpose

1. This annex describes the methods we use to generate a provisional HEIFES07 re-creation via the web facility. It also describes how we will use the comparison of HEIFES07 and the HEIFES07 re-creation in the 2007-08 ‘Statistics derived from ILR data for the monitoring and allocation of funding in FECs’ exercise. The outputs are provisional because they are generated before FECs submit final F04 data to the LSC.

Web facility outputs

2. The web facility generates a HEIFES07 re-creation. Annex B describes how to access the provisional HEIFES07 re-creation Excel workbook (HEIFER07YYYYYY.xls, where YYYYYY denotes the UPIN for the college). The provisional HEIFES07 re-creation workbook contains the worksheets listed in Table A.

Table A Excel workbook HEIFER07YYYYYY

Page number	Worksheet (see tabs on spreadsheet)	Description
1	Coversheet	Title page
2	Summary	HEIFES07 re-creation comparison summary table
3	PRGCMP	HEIFES07 re-creation price group comparison summary table
4	Excl	HEIFES07 re-creation exclusion table
5	FTS	Provisional HEIFES07 re-creation Table 1: Full-time years of programme of study
6	SWOUT	Provisional HEIFES07 re-creation Table 2: Sandwich year-out years of programme of study
7	PT	Provisional HEIFES07 re-creation Table 3: Part-time years of programme of study
8	FEE	Provisional HEIFES07 re-creation Table 4: Home and EC fees
9	CONS	Provisional HEIFES07 re-creation Table 6: HEFCE-recognised funding consortia 2007-08
10	HBK	Provisional HEIFES07 re-creation grant adjustment report
11	STD	Provisional HEIFES07 re-creation standard resource table
12	F07	Provisional HEIFES07 re-creation assumed fee income table
13	WP	Indicative 2008-09 WP allocation based on provisional HEIFES07 re-creation assumed FTEs

14	hFTS	HEIFES07 Table 1: Full-time years of programme of study
15	hSWOUT	HEIFES07 Table 2: Sandwich year-out years of programme of study
16	hPT	HEIFES07 Table 3: Part-time years of programme of study
17	hFEE	HEIFES07 Table 4: Home and EC fees
18	hCONS	HEIFES07 Table 6: HEFCE-recognised funding consortia 2007-08
19	hHBK	HEIFES07 grant adjustment report
20	hSTD	HEIFES07 standard resource table
21	hF07	HEIFES07 assumed fee income table
22	hWP	2008-09 WP allocation, based on HEIFES07
23	Credibility	Identifies areas for recognised HE qualification aims on 2007-08 ILR where data is potentially inaccurate
24	FTSDIFF	Identifies discrepancies in student numbers between the FTS and hFTS sheets
25	SWOUTDIFF	Identifies discrepancies in student numbers between the SWOUT and hSWOUT sheets
26	PTDIFF	Identifies discrepancies in student numbers and FTE between the PT and hPT sheets
27	FEEDIFF	Identifies discrepancies in student numbers between the FEE and hFEE sheets

Note: The workbook contains a version number that will increment by one following each successful submission of data to the web facility. There are also contact details for any queries relating to data in the workbook.

3. The 'DIFF' sheets (see items 24-27 in Table A) will indicate where differences in cell totals between the provisional HEIFES07 re-creation and HEIFES07 exceed a given threshold. The size of this threshold can be altered by entering the required value where indicated on the worksheets. These sheets are provided to assist institutions in reconciling differences between HEIFES07 and the provisional HEIFES07 re-creation.

4. The web facility will also generate a provisional HEIFES07 re-creation individualised file (HEIFER07YYYYYY.ind) to supplement the provisional HEIFES07 re-creation. The file contains details of how each student was classified in the re-creation. A full description of the data in the individualised file is given in Appendix 1.

Methods

5. The algorithms used to generate the provisional HEIFES07 re-creation are described in Appendix 1.

6. Appendix 2 contains guidance on troubleshooting the causes of differences between HEIFES07 and the provisional HEIFES07 re-creation.
7. In some circumstances, we have to make approximations when transforming the ILR data into the HEIFES07 re-creation. These 'problems of fit' are described in Appendix 3.
8. By early 2009 we will generate a HEIFES07 re-creation from the 2007-08 ILR F04 data submitted by colleges to the LSC. We will compare this re-creation to HEIFES07. We intend to use the algorithms described in Appendix 1 to generate the HEIFES07 re-creation; however, we may change them if we identify any improvements that can be made.
9. We expect to select colleges that we require to make a formal response to the comparison on the basis of thresholds. These are likely to be set in terms of funding differences arising from HEIFES07 and the HEIFES07 re-creation. This includes the 2008-09 WP allocation reports. The recalculations of 2008-09 WP funding in the 2007-08 'Statistics derived from ILR data' exercise may differ from those in the web facility, because of subsequent changes to, for example, 2008-09 ASN allocations, transfers and 2006-07 ILR data. Notwithstanding the thresholds, we may also ask for further information about the comparison from any college.
10. Each college required to respond formally will be asked to reconcile the two data sources. After reconciling the two sources and confirming that the HEIFES07 re-creation reasonably reflects the outturn position for 2007-08, the HEIFES07 re-creation will supersede HEIFES07. Any consequent grant adjustments (including 2008-09 WP allocations) will then be based on the superseding re-creation (subject to the appeals process and the availability of our funds).

Annex D – Derived statistics that may inform the 2009-10 widening participation allocations

Purpose

1. This annex describes how we may use 2007-08 ILR F04 data to inform the WP allocations for 2009-10. Further details of the algorithms we use for these data are given in Appendix 4.
2. The web facility offers colleges the opportunity to check and, where errors would affect these allocations under current proposals, correct their 2007-08 ILR F04 data before submission to the LSC. We strongly encourage colleges to use the facility in this way, as correction of data after submission is exceptional and time-consuming.

Web facility outputs and indicative funding calculations

3. The web facility will generate an indicative 2009-10 funding calculation for each of the WP allocations. These calculations use provisional 2008-09 allocation rates (announced in March 2008) applied to provisional 2008-09 assumed FTEs in most cases, and may be updated as more current information becomes available. They may not incorporate all 2009-10 additional student numbers, transfers or mergers. Annex B describes how to access these indicative funding calculations and derived statistics produced by the web facility in an Excel workbook (IHWP07YYYYYY.xls, where YYYYYY is the UPIN for your college).
4. The indicative allocations for 2009-10 are provided solely to highlight potential 2007-08 ILR F04 data errors, so they can be corrected by colleges before submission to the LSC. They should not be considered as any kind of funding commitment by HEFCE, and are without prejudice to what the HEFCE Board may agree to be the final allocations for any college. The final allocations for 2009-10 may be higher or lower than the illustrations given in the output from the web facility as a result of changes to data by the college or to the data provided by any other college, or changes to the total sum available for allocation, or as a result of any changes to the funding methods.
5. WP funding allocations are informed by the data provided by colleges. If we find that data errors have resulted in colleges receiving incorrect funding allocations, then we will adjust their funding accordingly. In particular, where reconciliations with 2008-09 ILR data or HEIFES08 data audit highlight that the assumed FTEs used to allocate 2009-10 WP funding were incorrect, then we will adjust WP funding accordingly, subject to the availability of HEFCE funds.
6. The WP derived statistics can be re-built from the individualised file we provide. (IHWP07YYYYYY.ind – see Annex B for how to access this file) This file contains details of how each student was categorised in the indicative WP outputs and, where relevant, details of why they did not contribute. A full description of the data in the individualised file is given

in Appendix 4, along with instructions on how to rebuild the figures in the WP derived statistics and indicative allocations workbook (IHWP07YYYYYY.xls).

Widening access for students from disadvantaged backgrounds

7. This is a formula-based allocation of funding for teaching which recognises the extra costs associated with recruiting and supporting undergraduate students from disadvantaged backgrounds, who are currently under-represented in higher education. The proposed method of allocating funds is as follows.

8. Using postcode information from 2007-08 ILR F04 data, each student is mapped to a 2001 Census area statistic ward. These wards are themselves assigned to quintiles based on young participation rates (used for young full-time students) and quintiles based on the proportion of 16-74 year olds with an HE qualification (for mature full-time, young part-time and mature part-time undergraduates). Each student is weighted according to the relevant quintile assignment of their ward:

Table B Widening access weights

Quintile	Weight
1 Lowest HE participation (young full-time) or lowest average educational achievement (part-time and mature full-time)	2
2	1
3, 4, 5	0

9. Young students are those aged under 21 on entry to the programme of study; mature students are those aged 21 and over on entry.

10. The young HE participation quintiles are based on an extension to our 2005 report on measuring young participation, 'Young participation in higher education' (HEFCE 2005/03) that will be published in 2008. These updated rates are based on young people who reached 18 between 2000 and 2004 and entered a higher education course in the UK while aged 18 or 19. Young participation rates are calculated for each 2001 Census area statistics ward in the UK and are used to rank the wards into five participation quintiles, each containing 20 per cent of the UK young population.

11. The adult HE qualification quintiles are based on 2001 Census area statistics. We use the national equivalents of the 2001 Census Key Statistics table 13 (KS013, 'Qualifications and students') for 2001 Census Output Areas (subsequently aggregated to 2001 Census area statistics wards). These tables can be obtained from the ONS, the General Register Office for Scotland (GROS) and the Northern Ireland Statistics and Research Agency (NISRA). We calculate the proportion of 16-74 year-olds with an HE qualification for UK 2001 Census small-area statistics wards. These wards are then ranked by this proportion to give

the adult HE qualification quintiles, with each quintile covering 20 per cent of the English 16-74-year-old population.

12. We allocate postcodes to 2001 Census area statistics wards using the August 2007 release of the ONS's NSPD. A file containing the allocation of postcodes to young participation and adult HE attainment quintiles is available on the HEFCE web-site, www.hefce.ac.uk, under Widening participation/POLAR and participation rates/POLAR2. This file includes postcodes which are excluded from the quintile mapping along with the reason for exclusion (including non-geographic postcodes) and may be updated as more current information becomes available.

13. Part-time and mature students who already hold an HE qualification at the same level as, or higher than, their current qualification aim, or have unknown entry qualifications¹, are given a weight of zero, irrespective of their postcode.

14. We calculate a 'widening access average weight' (separately for full-time and part-time students) as follows:

$$\frac{\text{Total weight for all students in the population}}{\text{Total students in the population}}$$

15. The population is defined as full-time or part-time (as appropriate) HEFCE-funded UK-domiciled new entrants that generate a Column 4 countable year in the HEIFES07 re-creation.

16. Some students are excluded from the population defined above:

- those with a postcode that has been identified in our young participation analysis as being associated with an unfeasible number of young entrants in relation to our population estimates. Typically this would be a postcode relating to a boarding school
- those whose postcode is marked as a non-geographic postcode in the NSPD
- those with a postcode that, although valid, is not mapped to the required Census 2001 geography in the NSPD
- those studying for an ELQ.

Additionally students with entry qualifications of unknown level are partially excluded from the population defined above. The amount we remove for each student is the proportion of their activity which we treat as ELQ (for further details see the algorithm for ELQ_PROP in paragraphs 32 of Appendix 4).

¹ This may differ slightly from students who study for an ELQ, where students with entry qualifications of unknown level are partially removed from the population – see paragraph 16 of this annex

17. These excluded students (with the exception of those studying for an ELQ and a corresponding proportion relating to students with entry qualifications of unknown level who are assumed to be studying for an ELQ) are counted in the FTEs in the next step (see paragraph 18), and therefore receive an average weight for the purpose of allocating funds.

18. Each average weight that was derived in paragraph 14 is London-weighted (8 per cent for inner London and 5 per cent for outer London) and applied to the assumed undergraduate (including foundation degree) FTEs for 2008-09 (which may not incorporate all 2009-10 additional student numbers, transfers or mergers).

Improving retention

Full-time students

19. As well as allocating funding to widen access, we allocate funding to improve retention. For full-time undergraduate students, this is likely to be based on their entry qualifications and age, and is likely to be calculated as follows.

20. Using age and entry qualification information from the 2007-08 ILR F04 data, full-time UK-domiciled undergraduate new entrants are assigned to one of six risk categories (see paragraph 21 for further information on how students are assigned to risk categories). These are then weighted as shown in Table C below. Students are only part of the population if they generate a HEFCE-fundable Column 4 countable year in the HEIFES07 re-creation and are not studying for an ELQ. Additionally students with entry qualifications of unknown level are partially excluded from this population. The amount we remove for each student is the proportion of their activity which we treat as ELQ (for further details see the algorithm for ELQ_PROP in paragraph 32 of Appendix 4).

Table C Risk category weights

	Young	Mature
Low risk	0	0
Medium risk	1	1.5
High risk	1.5	2.5

21. For this allocation, mature students are those aged 21 or over on entry. The assignment of students to one of the six risk categories based on entry qualifications and age is shown in Table D below.

Table D Assignment of students to risk categories

	Young	Mature
Low risk	<ul style="list-style-type: none"> • A-levels/Highers with more than 18 A-level points or unknown* A-level points • Baccalaureate • Degree or higher • Unknown qualifications[†] 	<ul style="list-style-type: none"> • A-levels/Highers with more than 24 A-level points or unknown* A-level points • Degree or higher • Unknown qualifications[†]
Medium risk	<ul style="list-style-type: none"> • A-levels/Highers with between 9 and 18 A-level points • Foundation course • Other HE qualification (below degree level) 	<ul style="list-style-type: none"> • A-levels/Highers with fewer than 25 A-level points • Other HE qualification (below degree level) • Foundation course • Access course
High risk	<ul style="list-style-type: none"> • A-levels/Highers with fewer than 9 A-level points • BTEC • Access course • Other qualifications • No qualifications 	<ul style="list-style-type: none"> • BTEC • Baccalaureate • Other qualifications • No qualifications

* New entrants whose highest qualification on entry are A-levels but who did not enter via UCAS (the universities and colleges admissions body) and do not have A-level points recorded are allocated to medium risk.

[†] New entrants with unknown entry qualifications are given a zero weight, and are identified in a separate category in the individualised file and indicative allocations spreadsheet to aid with data checking. Colleges should ensure that highest qualification on entry is recorded to ensure students are weighted appropriately in the allocation method for this stream of funding.

22. We calculate a 'full-time improving retention average weight' as:

$$\frac{\text{Total weight for all students in the population}}{\text{Total students in the population}}$$

23. The average weight derived from paragraph 22 is given a London weighting (8 per cent for inner London, 5 per cent for outer London) and applied to the assumed full-time undergraduate (including foundation degree) FTEs for 2008-09 (which may not incorporate all 2009-10 additional student numbers, transfers or mergers).

Part-time students

24. The part-time allocation is likely to be distributed pro rata to London-weighted part-time undergraduate (including foundation degree) assumed FTEs for 2008-09 (which may not incorporate all 2009-10 additional student numbers, transfers or mergers).

Widening access and improving provision for disabled students

25. We also allocate funding for widening access and improving provision for disabled students. This allocation is likely to be calculated using 2007-08 ILR F04 data as follows.

26. Firstly, we calculate for each college the proportion of eligible home and EC students who received the Disabled Students' Allowance (DSA). Students are only part of the population if they generate a Column 4 countable year in the HEIFES07 re-creation.

27. Next, each college is assigned to one of four quartiles, according to the proportion of students in receipt of the DSA as calculated in paragraph 26, although this is smoothed to ensure that no college falls by more than one quartile since the previous year. Separate weightings are attached to each of the four quartiles, as shown in Table E.

Table E Quartile weights

Quartile	Weighting
A (lowest proportion)	1
B	2
C	3
D (highest proportion)	4

28. Finally, each college's share of the funding is allocated pro rata to the assumed FTEs for 2008-09 (which may not incorporate all 2009-10 additional student numbers, transfers or mergers), weighted according to the quartile in which they fall and a London weighting (8 per cent for inner London, 5 per cent for outer London) although the following minimum allocations apply.

Table F Minimum allocations for widening access and improving provision for disabled students

FTE	Minimum allocation
<50	£500
50 to 249	£1,000
250 to 499	£5,000
500 or more	£10,000

Annex E – 2007-08 co-funded employer engagement student numbers (CFEE07) re-creation

Purpose

1. This annex describes the methods we use to generate a provisional CFEE07 re-creation via the web facility. It also describes how we will use the comparison of CFEE07 and the CFEE07 re-creation in the 2007-08 'Statistics derived from ILR data for the monitoring and allocation of funding in FECs' exercise. The outputs are provisional because they are generated before FECs submit final F04 data to the LSC.

Web facility outputs

2. The web facility generates a CFEE07 re-creation. Annex B describes how to access the provisional CFEE07 re-creation Excel workbook (CFEE07YYYYYY.xls, where YYYYYY denotes the UPIN for the college). The provisional CFEE07 re-creation workbook contains the following worksheets.

Excel workbook CFEE07YYYYYY.xls

Page number	Worksheet*	Title
1	Coversheet	Title page
2	CFEE	CFEE07 re-creation table

*This worksheet reference corresponds to the spreadsheet tabs.

Note: The workbook contains a version number that will increment by one following each successful submission of data to the web facility. There are also contact details for any queries relating to data in the workbook.

3. All the information contained in the CFEE07 re-creation tables can be re-built by categorising and aggregating the data contained in the individualised file which we provide. The file (CFEE07YYYYYY.ind) contains details, in the form of ILR and derived fields, of how each student was classified in the re-creation. A full description of the data in the individualised file is given in Appendix 5.

Methods

4. The algorithms used to generate the provisional CFEE07 re-creation are described in Appendix 5.

5. The CFEE07 return will not be collected until August 2008, however the web facility provides an early opportunity to verify that the 2007-08 ILR data are correct for these

students. By early 2009 we will generate a CFEE07 re-creation from the 2007-08 ILR F04 data submitted by institutions to the LSC. We will compare the CFEE07 re-creation to the CFEE07 return. Where appropriate, we will also re-calculate the funding associated with recruited FTEs and any differences from the original co-funded employer engagement funding based on the CFEE07 re-creation by applying the same formulae that were used to calculate the funds for CFEE07. We intend to use the algorithms described in Appendix 5 to generate the CFEE07 re-creation; however, we may change them if we identify any improvements that can be made.

6. The CFEE07 return will include data on the employers involved in the co-funded employer engagement provision, including characteristics of the employer such as sector and size of company. We are not able to re-create these data in the CFEE07 re-creation.

7. We expect to select colleges that we require to make a formal response to the comparison on the basis of thresholds. These are likely to be set in terms of funding differences arising from CFEE07 and the CFEE07 re-creation. Notwithstanding the thresholds, we may also ask for further information about the comparison from any college.

8. Each college required to respond formally will be asked to reconcile the two data sources. After reconciling the two data sources and confirming that the CFEE07 re-creation reasonably reflects the outturn position for 2007-08, the CFEE07 re creation will supersede CFEE07 and any consequent funding adjustments (including 2008-09 WP allocations and other targeted allocations) will be made (subject to the availability of our funds).

Annex F – Information for leads and members of HEFCE-recognised funding consortia

Purpose

1. This annex provides information for FECs that are part of a HEFCE-recognised funding consortium about the outputs that will be made available to them via the web facility.
2. This annex also describes how the data submitted by member colleges within a consortium will be incorporated in any outputs generated for the lead.

Data collection arrangements for HEFCE-recognised funding consortia

3. Every student within a HEFCE-recognised funding consortium is always recognised as being a student of the appropriate consortium member college, rather than of the lead college. There is only one lead institution in each consortium.
4. Within the sector, the leads of HEFCE-recognised funding consortia are both higher education institutions (HEIs) and FECs. Where an FEC is the lead, the institution must submit ILR data to the LSC only for students that are registered with it, and not for students registered with the member colleges. Students registered at a member college will be returned on that college's 2007-08 ILR.
5. All member colleges' data for students funded under the consortium arrangement should have been returned by the lead institution on its 2007-08 Higher Education Students' Early Statistics survey (HESES07) or HEIFES07 return. Therefore, for the lead college of a HEFCE-recognised funding consortium, the HEIFES07 re-creation will incorporate its 2007-08 ILR F04 data with 2007-08 ILR F04 data from each member college in the consortium. The lead and member college data are similarly aggregated for the indicative 2009-10 WP outputs.
6. In order for the web facility to produce a HEIFES07 re-creation for leads of HEFCE-recognised funding consortia that incorporates the students from member colleges, we have designed the facility so that it incorporates the latest data from the member colleges when it generates the re-creation for the lead.
7. The lead's web facility outputs will only incorporate member college data after these data have been submitted to the web facility by the member colleges. Where the lead of the consortium is an HEI, the member colleges' data will be passed to the HEI's re-creation output that is produced from an equivalent web facility for HEIs.

Web facility outputs

HEIFES07 and the HEIFES07 re-creation

8. The web facility generates a provisional HEIFES07 re-creation. This is described in Annex C. In addition, Annex B describes how to access the HEIFES07 re-creation Excel workbook from the HEFCE extranet. The information in the provisional HEIFES07 re-creation will vary depending on whether the college is a lead or member college.

9. Lead institutions of a HEFCE-recognised funding consortium will be able to access the following files via the web facility:

- LEAD07YYYYYY.xls – an Excel workbook containing the HEIFES07 re-creation tables for the lead college (includes the 2007-08 ILR data uploaded by the lead FEC only)
- HEIFER07YYYYYY.xls – an Excel workbook containing the HEIFES07 re-creation tables. This file includes the combined 2007-08 ILR data uploaded by both the lead and member colleges. This workbook combines the data contained in LEAD07YYYYYY.xls with the 2007-08 ILR data from the member colleges where they have submitted data to the web facility
- HEIFERC07YYYYYY.xls – Excel workbooks containing the HEIFES07 re-creation tables for each of the member colleges where member colleges have submitted ILR data to the web facility
- HEIFER07YYYYYY.ind – the individualised file to supplement the provisional HEIFES07 re-creation tables for the lead college (LEAD07YYYYYY.xls). All of the information in the HEIFES07 re-creation tables for the lead college can be re-built by categorising and aggregating the data in this file. It contains details, in the form of 2007-08 ILR fields and derived fields, of how each student was classified in the re-creation. A full description of the data in the individualised file is given in Appendix 1.

10. Member colleges of a HEFCE-recognised funding consortium will be able to access the following files:

- HEIFERC07YYYYYY.xls – Excel workbooks containing the HEIFES07 re-creation tables for the member college
- HEIFER07YYYYYY.ind – the individualised file to supplement the provisional HEIFES07 re-creation tables for the member college (HEIFERC07YYYYYY.xls). All of the information in the provisional HEIFES07 re-creation tables for the member college can be re-built by categorising and aggregating the data in this file. It contains details, in the form of 2007-08 ILR fields and derived fields, of

how each student was classified in the re-creation. A full description of the data in the individualised file is given in Appendix 1.

Derived statistics that may inform the 2009-10 WP allocations

11. The web facility generates indicative 2009-10 WP allocations. These are described in Annex D, and Annex B provides instructions on how to access the IHWP07YYYYYY.xls re-creation Excel workbook from the HEFCE extranet. The information in the IHWP07 re-creation will vary depending on whether the college is a lead or member college.

12. Lead institutions of a HEFCE-recognised funding consortium will be able to access the following files:

- IHWPLEAD07YYYYYY.xls – an Excel workbook containing the indicative WP tables for the lead college only
- IHWP07YYYYYY.xls – an Excel workbook containing the indicative WP tables that combines the data from the lead of the consortium with all member colleges that have submitted ILR data to the web facility
- IHWPC07YYYYYY.xls – Excel workbooks containing the indicative WP tables for each of the member colleges that have submitted ILR data to the web facility. These tables will only become available when member colleges submit data to the web facility
- IHWP07YYYYYY.ind – the individualised file that supplements the indicative 2009-10 WP allocations (IHWPLEAD07YYYYYY.xls). All of the information in the indicative WP tables for the lead college can be re-built by categorising and aggregating the data in this file. It contains details, in the form of 2007-08 ILR and derived fields, of how each student was classified in the re-creation. A full description of the data in the individualised file is given in Appendix 4.

13. Member colleges of a HEFCE-recognised funding consortium will be able to access the following files:

- IHWPC07YYYYYY.xls – an Excel workbook containing the indicative WP tables for the member college
- IHWP07YYYYYY.ind – the individualised file that supplements the indicative 2009-10 WP allocations (IHWPC07YYYYYY.xls). All of the information in the indicative WP tables for the member college can be re-built by categorising and aggregating the data in this file. It contains details, in the form of 2007-08 ILR and derived fields, of how each student was classified in the re-creation. A full description of the data in the individualised file is given in Appendix 4.

Access to outputs

14. Lead colleges will be given automatic access to all outputs except the separate HEIFES07 re-creation (HEIFER07YYYYYY.ind) and indicative WP (IHWP07YYYYYY.ind) individualised files for each of the member colleges. Each member college will receive access to its own separate HEIFES07 re-creation, and the accompanying HEIFES07 re-creation individualised file that will contain only its own 2007-08 ILR data. A similar arrangement will apply for the indicative WP outputs.

15. Where a member college gives us permission to allow the lead college access to its HEIFES07 re-creation or indicative WP re-creation individualised files, we will write to the lead college to describe how it can access these files via the HEFCE extranet.

Other outputs available to the lead and member colleges

16. In addition to the HEIFES07 re-creation and indicative WP outputs listed above, each lead and member college will also have access to the other outputs relating to its own data listed in paragraph 25 of Annex B.

Worksheets

16. The provisional HEIFES07 re-creation workbook for lead colleges, LEAD07YYYYYY.xls contains the worksheets listed in Table G.

Table G Excel workbook LEAD07YYYYYY

Page number	Worksheet (see tabs on spreadsheet)	Title
1	Coversheet	Title page
2	FTS	Provisional HEIFES07 re-creation Table 1: Full-time years of programme of study
3	SWOUT	Provisional HEIFES07 re-creation Table 2: Sandwich year-out years of programme of study
4	PT	Provisional HEIFES07 re-creation Table 3: Part-time years of programme of study
5	FEE	Provisional HEIFES07 re-creation Table 4: Home and EC fees
6	CONS	Provisional HEIFES07 re-creation Table 6: HEFCE-recognised funding consortia 2007-08
7	Excl	Provisional HEIFES07 re-creation exclusion table for the lead college
8	Credibility	Identifies areas for recognised HE qualification aims on 2007-08 ILR where data is potentially inaccurate

17. The provisional HEIFES07 re-creation workbook for member colleges, HEIFERC07YYYYYY.xls, contains the worksheets listed in Table H.

Table H Excel workbook HEIFERC07YYYYYY

Page number	Worksheet (see tabs on spreadsheet)	Description
1	Coversheet	Title page
2	FTS	Provisional HEIFES07 re-creation Table 1: Full-time years of programme of study for the member college
3	SWOUT	Provisional HEIFES07 re-creation Table 2: Sandwich year-out years of programme of study for the member college
4	PT	Provisional HEIFES07 re-creation Table 3: Part-time years of programme of study for the member college
5	FEE	Provisional HEIFES07 re-creation Table 4: Home and EC fees for the member college
6	Excl	Provisional HEIFES07 re-creation exclusion table for the member college
7	Credibility	Identifies areas for recognised HE qualification aims on 2007-08 ILR where data is potentially inaccurate

Annex G – 2009 Unistats summaries

Purpose

1. This annex describes how we intend to produce Unistats summaries from ILR data for directly funded higher education in FECs and member colleges of HEFCE-recognised funding consortia. We expect colleges to use this opportunity to check that the ILR data used to produce the Teaching Quality Information (TQI) are fit for publication on the Unistats web-site (www.unistats.com) in summer 2009.
2. It is also an opportunity for colleges to comment on the suitability for publication of the Unistats summaries and the algorithms used to generate them. Details of the algorithms used are in Appendix 7. Comments on the suitability for publication of the Unistats summaries and the Unistats algorithms should be sent to Marie-Helene Nienaltowski (e-mail ilr_heifes_stats@hefce.ac.uk) by **Friday, 19 December 2008**.

Background

3. The Unistats web-site was launched to bring together key sources of official information about the quality of higher education in UK universities and colleges, so that it is more accessible to applicants and their advisers. The web-site aims to make it easier to compare institutions and subjects so that students can make more informed choices of where to study. The site contains information such as the number of students completing awards, leaving without awards, and continuing at the institution. It also contains information about the number of leavers who are entering employment or further study, or who are unemployed.
4. The Unistats site currently contains a range of information for HEIs, as well as for HE at FECs that are funded indirectly through an HEI (excluding members of HEFCE-recognised funding consortia). However, from summer 2008, we expect that the Unistats site will also contain a range of information for directly funded FECs and consortium members.

Data verification

5. We are providing this year's web facility to allow institutions to verify the data underlying their Unistats outputs, before submitting their 2007-08 ILR F04 and 2007-08 F05 to the LSC.
6. We intend to publish these statistics on the Unistats web-site in summer 2009. The primary source for the data will be the 2007-08 F05 ILR return. It is our intention, within the boundaries of the available data, to make the information as comparable as possible to the statistics published on the Unistats site for HEIs and FECs that are funded indirectly through an HEI.

7. From February 2009 HEFCE staff may view Unistats outputs generated by the web facility to identify any general issues of data quality and the suitability of the algorithms. Further details about the data confidentiality agreement can be found in paragraph 35 of the introduction.
8. The statistics will be restricted to recognised courses of HE at HEFCE-funded institutions. See Annex D of HEIFES07, HEFCE 2007/25, for the definition of recognised courses of HE.
9. The web facility will give colleges the opportunity to view the proposed Unistats outputs from 2007-08 ILR data that they intend to submit to the LSC, and the chance to identify and correct, where necessary, errors in the underlying data.
10. Comments on the suitability for publication of the Unistats summaries and algorithms should be sent to Marie-Helene Nienaltowski (e mail: ilr_heifes_stats@hefce.ac.uk) by **Friday, 19 December 2008**.

Web facility outputs

11. Annex B describes how to access the derived statistics, in an Excel workbook, that we intend to publish on the Unistats site (TQI07YYYYYY.xls, where YYYYYY is the UPIN for the college).
12. To investigate the underlying data, institutions should use the individualised file (TQI07YYYYYY.ind) that contains details of how each student was classified and, where relevant, details of why they were excluded from the Unistats statistics. A full description of the data in the individualised file is given in Appendix 7.

Descriptions of Unistats statistics derived from ILR data

13. The Unistats statistics that the web facility will generate from the submitted 2007-08 ILR data are as follows.

Entry information

14. Data on the highest qualifications on entry will be provided against the following headings:
- first degree or higher
 - other HE
 - A-levels, Scottish Highers, or other Level 3 qualifications
 - BTEC national diploma
 - access course
 - Baccalaureate
 - other qualifications
 - no formal qualifications/not known.

15. Data are disaggregated according to the level of the student: postgraduate, first degree, and other undergraduate. For undergraduate students the data will be further disaggregated between those who were under 21 on 31 August 2007 and those 21 or over.

Continuation information

16. Continuation data relate primarily to the status in the 2006-07 academic year of students who were active in the 2005-06 academic year. Data are reported against the following categories:

- continuing
- gained intended award
- gained other award
- dormant
- left without award.

17. Students who are recorded as dormant will also have their status in the 2007-08 academic year recorded according to the same categories.

18. Data are disaggregated according to the level of student: postgraduate, first degree, and other undergraduate. These data will be further disaggregated according to whether the student was undertaking their first year of study in 2005-06.

Achievement information

19. These data are only for students whose qualification aim is a first degree; the class of degree is disaggregated as follows:

- first class honours
- upper second class honours
- lower second class and undivided honours
- other honours/pass
- ordinary degree
- unclassified honours.

Annex H – 2009 National Student Survey (NSS) target list

Purpose

1. This annex describes how we will use 2007-08 ILR student data submitted to the web facility to provide FECs with a list of students that should be included in their 2009 NSS target list.
2. This annex also describes how we will use colleges' 2007-08 ILR data to generate an individualised student data file that identifies, by reason, the students that have been excluded from the 2009 NSS target list.
3. The algorithms used to generate the list are explained in Appendix 8.
4. As for the 2008 survey we will use 2007-08 F04 ILR data to produce a target list of students to be included in the 2009 NSS. Colleges are encouraged to use the list to start preparing the contact details for these students so that these can be passed to the agency running the NSS. Contact details for students to be included in the 2009 NSS will need to be returned to the agency in late November 2008. We will extract a final target list for the 2009 NSS from data submitted to the web-facility on 14 November 2008, colleges should ensure that the most recent data submitted to the web-facility on this date generates a complete NSS target list.
5. Further details about the arrangements for the 2009 NSS will be published later this summer. Further details about the target lists are in Annex H and the algorithms used to generate the lists are given in Appendix 8. HEFCE staff will access NSS target lists generated by the web facility."

Background

6. The NSS provides information to prospective students, and their advisers, who are making choices about what and where to study. This information will be published with other information relating to teaching quality on the Unistats web-site (www.unistats.com) in summer 2009. The NSS also aims to contribute to public accountability, and to provide information that will help institutions to enhance teaching quality, by supplementing their internal feedback mechanisms.

Web facility outputs

7. The web facility generates the 2009 NSS target lists from submitted 2007-08 ILR data. Annex B describes how to access the target list (NSSP09YYYYYY.xls – where YYYYYY is the UPIN for the college).

Proposed scope of students to be included

8. For the 2009 survey, we propose that the population of students that will be surveyed comprises all full-time and part-time undergraduate students registered at FECs, who are at or near the point of completion of their programmes of study. Generally, students will be included if the course is expected to end between 1 February 2009 and 31 January 2010 (that is, they are expected to be in their final year of study).

9. In addition to HEFCE-funded students we expect all the following students to be included in the population:

- a. Students that are registered at one institution and taught by another.
- b. International students.
- c. Independently-funded students.
- d. Students funded by Skills for Health (the Sector Skills Council for healthcare) and the Department of Health.
- e. Students funded by the Training and Development Agency for Schools.

10. The following students will be excluded from the population surveyed:

- Students on programmes that do not lead to undergraduate qualifications.
- Students on short courses (of less than one FTE) and on low intensity courses.
- Students who are dormant in 2007-08.

Annex I – Lifelong Learning Network (LLN) student summaries

Purpose

1. This annex describes how we intend to identify students who are part of an LLN using their 2007-08 ILR data. Colleges participating in an LLN in 2007-08 should use this opportunity to verify that we are correctly identifying their LLN students from their data. This annex applies only to colleges that have registered LLN students during 2007-08 at their college as part of an LLN initiative.
2. We have produced an indicative report of 'calculation of notional grant associated with LLN students', based on the number of LLN students that we identify from the ILR data. Within the teaching funding method we monitor LLNs' additional student numbers (ASNs) as part of a college's total ASN targets. These are shown in the HBK worksheet of HEIFER07YYYYYY.xls, as described in Annex C. Therefore the indicative report is provided solely to highlight potential 2007-08 ILR data errors, and should not be considered to be any kind of funding commitment by HEFCE.

Background

3. In John Selby's letter of 2 May 2006, we asked colleges to provide details of how we can identify LLN students on their ILR returns. This letter also explained that, wherever possible, we will use ILR data to evaluate the LLN initiative. If LLN students are identified on the ILR record we can follow them through their HE study and measure their achievement within HE and their evaluation of their HE experience (through the NSS).

Web facility outputs

4. Annex B describes how to access the derived statistics, in an Excel workbook, that are used to generate the LLN student summaries (LLN07YYYYYY.xls, where YYYYYY is the UPIN for the college).
5. To investigate the underlying data, institutions should use the individualised file (LLN07YYYYYY.ind) that contains details of how each student was classified and, where relevant, why they were excluded from the LLN summaries. A full description of the data in the individualised file is given in Appendix 9.

Annex J – Submitting overrides to derived fields

Purpose

1. This annex describes the format and submission process for override files. Erroneous Learning Aims Database (LAD) data or the HEIFES07 re-creation algorithms may occasionally result in correct ILR data generating derived fields that do not reflect a student's actual characteristics. In such cases an override file can be used to correct the derived fields or our version of the LAD data.
2. Colleges should note that we will only apply override files if we agree that they are appropriate. We may need to request further information to determine this. It may not always be possible to determine whether an override is appropriate until we have examined its impact on the derived fields affected. Therefore we may refuse to accept an override after it has been submitted.
3. There are two general cases where overrides may be applied:
 - to correct erroneous or missing LAD data
 - to correct problems of fit with our algorithms.

Erroneous or missing LAD data

4. When a learning aim provided by a college is not included in our copy of the LAD, overrides may be submitted to provide the information needed. Similarly, errors in LAD data for learning aims that are included on our copy of the LAD may require correction through an override file. LAD data can be corrected through an override file for the fields given in Table I.

Table I LAD fields

Field name	Field name in LAD	Column in individualised file
AWARD_BO	Awarding_body_code	AO
ENG_LEVE	England_FE_HE status_code	AR
QUAL_TYP	Learning_aim_type	BQ
LDCS_C01	Learndirect_code_1	BU
LDCS_C02	Learndirect_code_2	BV
LDCS_C03	Learndirect_code_3	BW

Where an override is providing data for a learning aim that is not included on the LAD, all the fields in Table I should be included in the override file.

5. Overrides to correct erroneous or missing LAD data are always temporary. They will only be applied to ILR data submitted to the 2007-08 web facility. Therefore colleges should also liaise with the LSC to request that the necessary corrections are made to the LAD. The LSC asks providers 'to notify the LSC of learning aims missing from the database when they start to enrol learners on the learning aim', (see the guidance notes for the 'learning aim reference' (A09) field in the 'Specification of the individualised learner record for 2007/08' (available from the LSC)).

6. Paragraphs 13 to 17 of this annex describe the format which colleges must use when submitting override files to correct erroneous or missing LAD data. Colleges should note that we will be unable to process files which are not submitted in this format.

Problems of fit with our algorithms

7. Where problems of fit with our algorithms result in errors to derived fields we are only allowing certain derived fields to be modified by the override file. These fields, which we refer to as 'primary derived fields', are those which are directly used to generate the HEIFES07 re-creation tables. For example, the value of HEFEXCL determines whether a record is included in the HEIFES population and is a primary derived field. In contrast, the fields EXCL1-EXCL64 (which are used to derive HEFEXCL) are not directly used to generate the re-creation tables; these are secondary derived fields. A list of primary derived fields is given in Table J.

8. Since overrides can only alter primary derived fields, a record's secondary derived fields may be inconsistent with the primary derived fields. Considering the example above, if the value of HEFEXCL is altered by an override file, its value will be inconsistent with the values of EXCL1-EXCL64.

9. We also require that certain primary derived fields are returned as a group, in order to maintain data integrity. If any field in the group is changed, we require that all members of that group are returned, even if the values of certain fields in the group remain unchanged. Details of the fields which must be returned in a group are given in Table J. For example, if an override to correct price groups or FTEs is being submitted, we require that the file contains FTEB-D, FTEMEDIA, FTEITT and FTEINSET, even if the values of some of these fields remain unchanged.

10. Where the cause of an error in a derived field is erroneous ILR data, override files should not be submitted to correct the error – instead the ILR data should be corrected.

Table J Primary derived fields

Field name/group	Description	Column in individualised file
HEFCOMP	HEIFES completion of year of programme of study indicator	X
HEFEXCL	Reason for exclusion from the HEIFES population	M
HEFFEELV	Fee level	AK
HEFLEVEL	Level of study	Z
HEFMODE	Mode of study	U
HEFREG	Column 1 or 2 indicator	V
HEFTYPE	Fundability status	Y
LENGTH	Flag indicating whether long or standard-length years of programme of study	AA
FTEB, FTEC, FTED, FTEMEDIA, FTEITT, FTEINSET	Proportion of FTE in each price group. These fields must be returned as a group	CD-CI

11. To allow colleges to check that an override file has had the desired effect, a new field, 'OVERRIDE', is included in the individualised file. This takes the value 1 if an override has been applied to the record, otherwise its value is 0. This field also allows colleges to determine which records may have inconsistent secondary derived fields, as a result of having an override applied.

12. Paragraphs 13 to 17 of this annex describe the format which colleges must use when submitting override files to correct problems of fit with our algorithms. We will be unable to process files which are not submitted in this format.

Format of override files

13. The override file consists of two parts: a header, which contains details of the override, and the data part, which contains the data which are being corrected. An example of a typical override file can be found at the end of this annex.

14. Overrides to 2007-08 derived fields or LAD data must be sent as a comma-separated file via the HEFCE extranet. To submit via the extranet go to the 'HEFCE resources' page and, after logging in at <https://extranet.hedata.ac.uk>, click on '2007-08 statistics derived from ILR data', then on 'ILROVR07 data collection'. From here click on the 'Upload' button and then, on the subsequent page, click on the

'Browse' button to find the file you wish to submit. Once you have selected this file and the pathname has appeared in the entry field, click on 'Upload File' to complete the upload.

15. Override files must be given a file name in the form ovrYYYYYYn.amd, where:

- YYYYYY is the provider number, ST_UPIN (L01)
- n is a sequential number, starting at 1.

For example, the first override file submitted by institution 999999 would be called ovr9999991.amd. The second file submitted would be called ovr9999992.amd.

The override header

16. The override header should be in the following form:

Line 1: contains the filename (as described in paragraph 15), with the '.amd' extension removed.

Line 2: the date on which the override was submitted, in the form ddmmyyyy. For example 1 August 2008 would be represented as 01082008.

Line 3: a brief description of the purpose of the override.

Line 4: contains the word 'OVERRIDE'.

Line 5: the field(s) used to indicate records which should be corrected by the override, comma separated. These should be named according to Table K, of Appendix 1. For example, QA_AIM_R (A09) could be used to correct a primary derived field for all students on a specific learning aim.

Line 6: the names of the LAD fields or primary derived fields being changed, comma separated. The fields which may be changed are listed in Tables 9 and 10.

Line 7: the number of rows of data (excluding headers) in the override file.

Line 8: the field used to compute the file's check-sum (see paragraph 17, below).

Line 9: the value of the check-sum. Paragraph 17 explains how to choose a suitable field as check-sum and explains how to compute its value.

Lines 10-11: these lines may be used for any notes that the college wishes to include.

Line 12: fields included in the override file. The fields should be specified in the same order as in the data part of the file and must be comma separated.

Line 13: the data must begin on this line. Details regarding its format are given in paragraph 18.

Check-sum

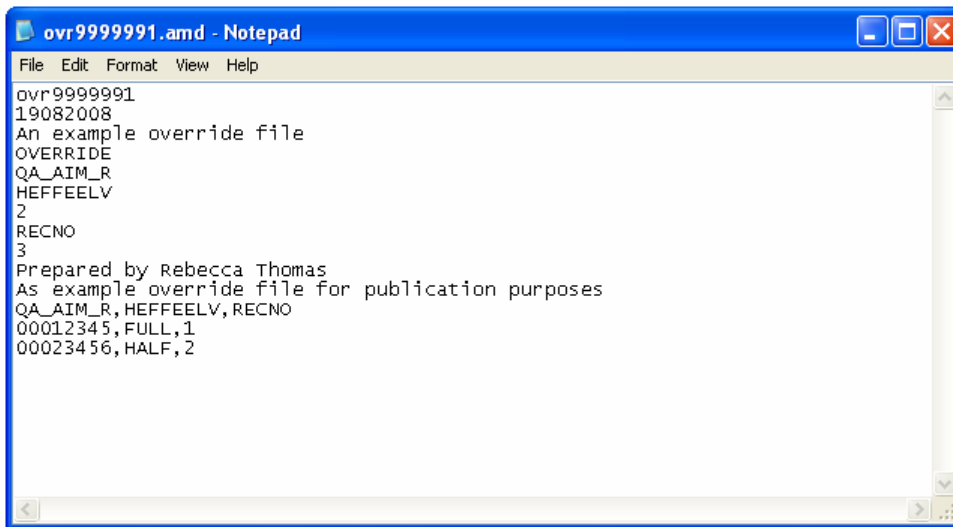
17. To ensure that the override file has been received in its entirety, or has not been otherwise corrupted during transmission, we use a check-sum. The check-sum is calculated by summing the values of the field specified on line 8 over all records in the file. The calculated value should be returned on line 9 of the override file. The field used to compute the check-sum must be numeric, and must not contain any values greater than 20,000. If no suitable field is available, then a sequential field called RECNO may be created. For example, RECNO may contain 1, 2, 3, 4, 5 etc.

The data part of the file

18. The records being changed must begin on line 13 of the override file. The data must consist of comma-separated fields, corresponding to those specified in line 12 of the header. Each record must be separated by a carriage return. A blank line should be placed after the final record.

Figure 1 An example override file

This example will change the fee level of (hypothetical) learner aim 00012345 to FULL and aim 00023456 to HALF. Note the use of RECNO to create a suitable field to allow the check-sum to be computed.

A screenshot of a Notepad window titled 'ovr9999991.amd - Notepad'. The window has a menu bar with 'File', 'Edit', 'Format', 'View', and 'Help'. The text inside the window is as follows:

```
ovr9999991  
19082008  
An example override file  
OVERRIDE  
QA_AIM_R  
HEFFFEELV  
2  
RECNO  
3  
Prepared by Rebecca Thomas  
As example override file for publication purposes  
QA_AIM_R,HEFFFEELV,RECNO  
00012345,FULL,1  
00023456,HALF,2
```

Annex K – HEIFES08 non-completion toolkit

Purpose

1. This annex contains details on how to use the non-completion toolkit. It should be read in conjunction with, and used in accordance with, the forthcoming HEIFES08 publication, which provides full guidance relating to the return of non-completion forecasts.

Description

2. The toolkit is a workbook that allows colleges to generate estimated non-completion rates for all students by course and year for 2007-08, from the ILR data that they submit to the web facility, using fields that they consider influential to this rate. The estimated rates may be used by colleges to inform the forecasts of non-completions returned in HEIFES08.

Reliability of estimated non-completion rates

3. Colleges should note that the toolkit is provided purely as an aid in generating non-completion rates for HEIFES08. We expect colleges to use their own judgement in assessing whether the rates generated are applicable and appropriate.

4. Estimates may be unreliable if the numbers of students involved are small or if the underlying ILR data contain errors or are not completed in line with HEFCE guidance on non-completion status. Colleges should consult the forthcoming HEIFES08 publication for definitive guidance on non-completions as well as information on making HEIFES estimates and keeping records for audit purposes. The HEIFES07 publication (HEFCE 2007/25) also contains guidance on non-completions. While we do not anticipate the guidance changing substantially between HEIFES07 and HEIFES08, institutions should use the most recent available version of the publication.

5. The toolkit will not treat non-exempt ELQ students that are returned as HEFCE-fundable on the 2007-08 ILR return as non-fundable. Therefore colleges should make an adjustment for such students if applying the results of the toolkit to their HEIFES08 return.

Using the toolkit

6. Users should follow the instructions in Annex B to access the file IHNCTK07YYYYYY.xls, and open it in Excel 2000 (or a later version). Owing to the technology used, the toolkit will not function correctly in earlier versions. The creation of tables uses macros, so these should be enabled to use the facility. The worksheet called 'NCdata' contains the raw data used to generate the tables, and the 'Pivot Table' worksheet displays the results. When you open the file you will be prompted to select the fields you want to use to generate the non-completion rates – we advise

summarising by course in the first instance. To create the table click 'OK', and the screen will close, enabling you to view the table. Click on 'Create a new table' to repeat the process – this will delete the previous table from the worksheet.

7. Non-completion rates based on small numbers of students may be unreliable, so any rate based on a population of fewer than 100 is coloured red and noted as unreliable. In these cases we advise using less detailed data to obtain a more reliable estimate.

8. In order to save a table you have created, open the <File> menu and select <Save As>.

List of abbreviations

ASN	Additional student number
BTEC	Business and Technology Education Council
CFEE	Co-funded employer engagement
DSA	Disabled Students' Allowance
EC	European Community
ELQ	Equivalent or lower qualification
FAQ	Frequently asked question
FE	Further education
FEC	Further education college
FTE	Full-time equivalent
FTS	Full-time
GROS	General Register Office for Scotland
HE	Higher education
HEFCE	Higher Education Funding Council for England
HEI	Higher education institution
HEIFES	Higher Education in Further Education: Students Survey
HESA	Higher Education Statistics Agency
HESES	Higher Education Students' Early Statistics survey
INSET	In-service education for teachers
ILR	Individualised Learner Record
IP	Internet Protocol
ITT	Initial teacher training
LAD	Learning Aims Database
LDCS	Learning Directory Classification System
LLN	Lifelong Learning Network
LSC	Learning and Skills Council
NISRA	Northern Ireland Statistics and Research Agency
NSPD	National Statistics Postcode Directory
NSS	National Student Survey
NVQ	National Vocational Qualifications
ONC	Ordinary National Certificate
OND	Ordinary National Diploma
ONS	Office for National Statistics
POLAR	Participation of local areas
QTS	Qualified teacher status
SIVS	Strategically important and vulnerable subjects
SQA	Scottish Qualifications Authority
SSO	(Athens) Single Sign On
SVQ	Scottish Vocational Qualifications
TQI	Teaching Quality Information
UKPRN	UK Provider Reference Number
UPIN	Unique Provider Identification Number
WP	Widening participation