May 2012/10
Issues paper

This report is for information

This report introduces a new approach to measuring rates of qualification from research degree study. It provides qualification rates of students on postgraduate research degree programmes at higher education institutions in England.

Rates of qualification from postgraduate research degrees

Projected study outcomes of full-time students starting postgraduate research degrees in 2008-09 and 2009-10

HIGHER EDUCATION FUNDING COUNCIL FOR ENGLAND

Contents

Rates of qualification from postgraduate research degrees

Projected outcomes of full-time students starting postgraduate research degrees in 2008-09 and 2009-10

To	Heads of publicly funded higher education institutions in the UK
Of interest to those responsible for	Research management, Supervising and managing research degree programmes, Quality assurance, Student data
Reference	2012/**10**
Publication date	May 2012
Enquiries to	Alison Brunt, tel 0117 931 7166, e-mail a.brunt@hefce.ac.uk

Executive summary

Purpose

1. This report introduces a new approach to measuring rates of qualification from research degree[1] study. It provides qualification[2] rates of students on postgraduate research degree programmes at higher education institutions (HEIs) in England, projected over periods of 7 and 25 academic years. HEFCE published research degree qualification rates in 2007 and 2010, but the new method means the results included in this publication are not directly comparable with them.

Key points

2. We expect that this information will be of interest to HEIs and others interested in the quality of research degree programmes and qualification rates of research degree students.

3. The rates of qualification from research degree study presented here are calculated on the basis of a method[3] derived from that currently employed by table series T5 of the UK higher education performance indicators (PIs). An overview of our new method is described below and a full technical description is given in the annexes of this publication. It is intended to improve the relevance and production of information regarding qualification from research degree study.

[1] In this document the term 'research degree' is used to refer to all 'doctorate degrees obtained primarily through advanced supervised research written up as a thesis/dissertation'.

[2] A student is defined as qualifying when they have been awarded a research degree and the 'qualification obtained' has been returned through the individualised HESA student record. This will typically be up to a year after the student submitted their thesis for assessment.

[3] A previous method generated two publications titled 'Research degree qualification rates', HEFCE 2007/29 and HEFCE 2010/21, which were published in 2007 and 2010 respectively.

4. HEIs in England reviewed this new method in winter 2011 to see how it calculated their own qualification rates. They were invited to provide feedback on the method used and it has subsequently been modified to take account of concerns they raised.

5. Projected rates of qualification from research degree study are provided in this document for full-time students registered at English HEIs who started research degree programmes in 2008-09 and 2009-10. It is restricted to students who were resident in the EU, including the UK, Guernsey, Jersey and the Isle of Man.

6. This report shows qualification rates calculated from data submitted by HEIs to the Higher Education Statistic Agency (HESA) in its annual data collections. For each HEI in England the proportions of their full-time research degree starters that are likely to be in each of three end states have been projected. That is, the proportion of the cohort expected to have qualified, transferred to another institution, or become absent from HE. Recent patterns of students at an institution are used in the projection of these outcomes. Each HEI's rate is presented alongside a sector-adjusted average (benchmark) and, where appropriate, an indicator which highlights significant variation from the sector-adjusted average.

Action required

7. No action is required in response to this document.

Introduction

Background: existing measures of retention and study outcomes

8. The benefits of higher education (HE) are substantial, wide ranging and impact upon not only students themselves but employers, the economy and society in general. The retention of students in HE, in addition to their study outcomes, has therefore been a focus and a policy concern for several years. A number of different measures have previously been used to consider retention and study outcomes of students studying at different levels of HE.

Undergraduates

9. For full-time first degree students retention has been measured and published annually since 1999 in the UK higher education performance indicators (PIs)[4]. To date, cohorts of students registered at UK higher education institutions (HEIs) in each year from 1996-97 to 2009-10 have been considered. Additionally, the PIs have included retention information relating to full-time 'other undergraduate'[5] entrants since 2004 (the 2004 publication considered the 2001-02 starting cohort) and relating to part-time first degree entrants since 2010 (considering the 2006-07 starting cohort).

10. The PIs make use of two methods in respect of retention among undergraduates[6]. The first method generates table series T3 reports on the proportion of first degree and other undergraduate entrants in a particular year who are found to be present in HE one year later (for full-time students) or two years later (for part-time first degree students). This measure aims to show how effective an institution is at retaining the students it recruits.

11. The second method generates table series T5 using a census cohort approach. This PI calculates the expected study outcomes of a cohort of full-time first degree starters[7] in a particular year by making a projection of the proportion of those starters likely to be in each of three 'end states' after a period of 15 years (which are: gained a qualification, transferred to another institution, or been absent from HE for at least two consecutive years).

---

[4] For further information see www.hesa.ac.uk under Performance Indicators.

[5] 'Other undergraduate' students include those studying for, among others: foundation degrees; higher national diplomas or certificates (HND/HNCs); professional qualifications; and certificates or diplomas of higher education. Those studying for institutional credit are not included.

[6] Definitions and technical explanations of the methods are available alongside the PIs at www.hesa.ac.uk under Performance Indicators.

[7] While table series T3 of the Performance Indicators considers entrants to programmes of study, table series T5 considers first degree starters. A student is defined as an entrant in a given year if the commencement date of their study falls within that academic year. With regard to table series T5, starters are full-time first degree students at a particular institution who were not studying at that institution full-time for a first degree in either of the two years prior to the academic year of interest, unless they had already qualified with a first degree. For further information, see www.hesa.ac.uk under Performance Indicators and Non-continuation rates (including projected outcomes) – definitions (tables T3, T4, T5).

Taught postgraduate programmes

12. Since 2005 retention rates of students who start a full- or part-time taught postgraduate programme of study have been published. Initially they were on the Teaching Quality Information (TQI) web-site, before they moved to the Unistats[8] web-site in 2007. Rates of continuation from the year of entry to the following academic year are provided at the level of a course at an institution.

13. For a cohort of students starting a particular programme of taught postgraduate study at an institution, the Unistats web-site has reported the proportion of that cohort seen to: take a break from study; continue at the institution; gain the intended award or higher; gain another award; or leave with no award.

Research degree programmes

14. In keeping with a commitment to support excellence in the national research base, HEFCE has required the research degree programmes we support through our grant to meet minimum standards set out in the Quality Assurance Agency for Higher Education (QAA) code of practice[9]. In September 2004 HEFCE announced plans to monitor how long it takes postgraduate research degree students to obtain their qualification[10]. This monitoring was intended as one measure by which the quality of research degree programmes could be assessed.

15. HEFCE published research degree qualification rates (RDQRs) for the first time in 2007 in 'Research degree qualification rates' (HEFCE 2007/29). This publication provided qualification rates for full-time students in English HEIs who had commenced their research degree in 1999-2000. In 2010 we published 'Research degree qualification rates' (HEFCE 2010/21), which presented rates for full-time students who started research degree programmes in 2000-01, 2001-02 or 2002-03; and part-time and mode-switch students who started their research degree in 1999-2000.

16. The analyses informing the RDQR publications sought to identify institutions whose returns to the Higher Education Statistics Agency (HESA) indicated a low proportion of research students qualifying within a given time, compared to the average in the sector, having taken into account factors which we know have a significant bearing on qualification rates[11].

---

[8] http://unistats.direct.gov.uk/

[9] The QAA's 'Code of practice for the assurance of academic quality and standards in higher education' is available at www.qaa.ac.uk/Publications/InformationAndGuidance/Pages/Code-of-practice-section-1.aspx.

[10] HEFCE Circular letter 18/2004, 'Postgraduate research degree programmes: minimum standards and funding' is available at www.hefce.ac.uk/pubs/circlets/2004/cl18_04/.

[11] A number of factors are known to impact on a student's chances of completing their doctoral degree within a given time frame. These factors were described in 'PhD research degrees: Entry and completion' (HEFCE 2005/02) and considered again in HEFCE 2007/28, and included: mode of study; domicile; subject area of study; source of funding; age; ethnicity and qualifications on entry.

17. Full-time qualification rates were calculated over seven academic years[12]. A cohort of full-time research degree students were tracked from the year they entered an institution across seven consecutive academic years, with the RDQR being the number of that cohort found to have qualified within the time frame, divided by the total population. RDQRs for part-time and mode-switch students (those who switched from full-time to part-time, or vice versa) were calculated over 10 academic years.

18. There are limitations to the calculation of RDQRs, as well as to the current coverage of the PIs in respect of research degree retention. Given the need to consider at least seven academic years' data, the most recent cohorts that the 2010 publication of RDQRs was able to consider were those of full-time research degree students at English HEIs, who commenced their programme of study in 2000-01, 2001-02 or 2002-03, and these cohorts may be considered somewhat out of date. The type and attitudes of students choosing to embark upon a research degree may have changed in recent years, and behaviours observed among students commencing in 2000-01 may be less relevant to research degree starters active today. Changes may also have been made to institutional practices in more recent years, and it is important to note that RDQRs would be very slow to respond to these: any improvements made to these practices would be unlikely to become fully apparent in RDQRs until seven academic years after the change had been made.

19. Further to the limitations outlined in paragraph 18, the generation of RDQRs was particularly resource-intensive for both HEFCE and HEIs. Resulting from the need to use historical data, institutions were tasked with verifying the accuracy of the underlying HESA data used, and making amendments where errors in that data influenced the accuracy of the rates calculated for them. Enhancement of HESA's validation processes, and other improvements made by institutions in terms of data quality, mean that more recent HESA data is likely to be more robust.

Objectives of this report

20. Given the constraints of the RDQR method for measuring qualification from research degree study, and the need to provide information of this nature, it was desirable to look for an alternative approach. At the same time, a programme of work for the Performance Indicators Steering Group (PISG) has been seeking to establish extensions of existing indicators to additional cohorts, including those of postgraduates. This publication is intended to bridge the two agendas and inform discussion regarding the future approach to the measurement of qualification from research degree study.

21. This report introduces an alternative approach to measuring rates of qualification from research degree study. Rather than the RDQR method, here we make use of a method similar to that currently employed by table series T5 of the PIs to project the study outcomes of full-time first degree study. In the sections that follow we present an overview of the new method used to measure rates of qualification from research degree study, along with the outcomes for each

[12] This time frame was derived from HEFCE 2005/02, which showed that after seven academic years the number of qualifiers and of students ceasing to be active had started to level off. The update of that report (HEFCE 2007/28) showed that qualification rates continued to rise beyond a seven-year period, but not significantly.

institution. A technical explanation of the method is at Annex A and data definitions are at Annex B.

22. For a cohort of full-time research degree starters[13] registered at English HEIs, we have projected the proportions that are likely to be in each of three 'end states' after a period of 25 years, based on the recent patterns of students at the institution. That is, the proportion of the cohort expected to have qualified, transferred to another institution, or become absent from HE in the 25 years following commencement of their research degree. This report also provides projected outcomes after a seven-year period, to enable a higher degree of comparison with the outcomes produced by the previous RDQR method[14].

23. The method in this report makes no consideration of research degree students who were domiciled outside of the European Union (EU), or those who studied part-time in the year they commenced their research degree.

24. The diversity of, and flexibility found in, part-time provision more generally make the data difficult to capture accurately and interpret. Coupled with small cohorts of part-time students and more flexible programme structures in study at postgraduate level, this means that recent patterns of study for part-time research degree students are particularly disparate and do not form a robust basis from which to project study outcomes.

25. For research degree students domiciled outside of the EU, numbers were found to be too small for study outcomes to be projected and reported in the manner proposed by this publication as a distinct group. Previous HEFCE analysis has shown that the aggregation of such students with their UK- and EU-domiciled counterparts is not appropriate to the proposed method on account of the differing nature of progression patterns through postgraduate research degree study.

Institutions' feedback

26. HEIs in England reviewed the outcomes generated by the new method in winter 2011. Each HEI was provided with its own results, along with overall results for the sector and an explanation of the method. Institutions were invited to provide feedback on the method and the results generated. A number of responses were received.

27. A number of the concerns we received related to errors in the institution's underlying data submitted to HESA and the impact of these errors on the qualification rates calculated. Where these concerns were upheld, the institution could choose to have this publication include a suppression of their results along with an explanation of the reasons for this suppression.

28. Other concerns related to perceived flaws or technical oversights of the method used to calculate the qualification rates. Where these concerns were justified, the method has been modified to account for them. In particular, institutions were concerned that the method did not properly take account of the progression of students who undertook one or more periods of part-time study, or one or more periods of writing up, in projecting the outcomes. We have modified

[13] For the purposes of this analysis a research degree starter is specifically defined. A full definition of a research degree starter is given at paragraphs 49 to 50.

[14] Further details of this approach are provided at paragraph 55.

the method to ensure that the outcomes of students who undertake extended periods of part-time study or writing-up are included in the projections. Institutions were re-contacted and advised of the changes and their re-calculated qualification rates early in 2012.

29. As described at paragraphs 23 to 25 there are a number of reasons why the method currently makes no consideration of research degree students who were domiciled outside of the European Union, or those who studied part-time in the year they commenced their research degree. Completion of students engaged in other forms of postgraduate research study (that is, not a research degree programme) is beyond the scope of the current method. Additionally, and as paragraph 62 later describes, if an institution's population of research degree starters contains fewer than 20 individuals, percentages calculated by the method for that institution have been suppressed. A number of institutions noted these limitations.

30. For this first publication, the intention has been to establish the new method as an objective, fair and reliable measure of postgraduate research degree completion, and to embed understanding of its underlying principles. It is intended that these qualification rates will be calculated regularly going forward, and in implementing such regular reporting a number of the issues outlined at paragraph 29 above would be addressed. In due course, and on the basis of an enhanced understanding of the study patterns of such students, it is likely that further development would lead to a complementary version of this method that would consider a population of non-EU domiciled students, and/or small institutional populations of research degree starters. Additional measures may also be developed that are appropriate for consideration of part-time research degree starters and starters on other forms of postgraduate research programmes of study.

Data source and the cohort considered

Data source

31. The method described in this report draws on data from the HESA individualised student records from 2003-04 through to 2010-11, the most recent data available.

32. Individual students are tracked within and through each of the individualised HESA student data sets from 2006-07 through to 2010-11 using a number of personal characteristics, linking instances of HE study for each student within our cohorts. The process of linking students in this way is the same as that used by HESA in the generation of the PIs.

33. We have also linked back via the Husid, Institution and Numhus (HIN) variables[15] to the HESA student records for the five years preceding the academic year for which the study outcomes are projected. Minimal information is retained from the HESA student records for each of these five years: namely, the HIN information, and the student's commencement date of study, mode of study, qualification aim and any qualification obtained.

[15] The HIN is a combination of three fields from the HESA individualised student records which uniquely identify a student on a course (or instance of study). These are the HESA unique student identifier (Husid), the UK provider reference number (Institution) and the student instance identifier (Numhus). The HIN forms a year-on-year linking mechanism which can be used to track the student's progression on the instance from one year to the next. It follows that the HIN enables a student to be tracked from commencement of the instance through to completion.

34. This process enables us to determine if the student was active on an instance of full-time postgraduate research study at any time during the five-year period, and thus ensures that a full-time research degree student continues to contribute to successive executions of the method[16] until they can be considered to have reached one of the end states.

35. Additionally, the five-year HIN linking allows us to consider an instance of full-time postgraduate research degree study from either the academic year in which they commenced the instance of study, or the academic year five years prior to the one for which study outcomes are projected. This in turn enables us to exclude from the method students who have previously reached one of the end states. For example, a student identified as having been inactive for two or more consecutive academic years are considered throughout the method to have 'become absent from HE': that is, they have reached one of the end states. Similarly a student who was awarded the qualification prior to the academic year for which the study outcomes are projected has reached a 'qualified' end state. In both cases if the student returned to full-time study, inclusion of their onward study patterns would distort the method. This process seeks to ensure that a full-time research degree student's contribution to successive executions of the method ceases beyond their arrival at one of the end states.

36. For data definitions and further explanation of how students are tracked see Annexes B and C respectively.

The cohort considered

37. At present all populations considered by this method are restricted to students who were resident of the EU, including the UK, Guernsey, Jersey and the Isle of Man. Incoming and visiting exchange students are excluded, as are those studying for the whole of their programme of study outside the UK. Unless it is further clarified, the term 'student' hereafter is taken to refer to individuals with these characteristics.

38. In an approach consistent with that used for the PIs, students are not included if they left the programme of study within 50 days of commencement. There is no adjustment for individual reasons for non-qualification.

39. This report considers a population of students who commenced a full-time research degree registered at an English HEI. All populations discussed within the main body of the report focus on the 2009-10 cohort in respect of both the technical explanation of the method, and its outcomes.

Overview of the method

40. The method used and described in this report is derived from that currently employed by table series T5 of the PIs to project the study outcomes of full-time first degree study. It seeks to project the study outcomes of full-time research degree students.

[16] That is, when the method is later used to project the outcomes of students starting a full-time research degree in 2010-11, 2011-12, and so on.

41. The description that follows makes use of terminology, which is defined as follows for the purposes of this report:

a. 'Research degree' – refers to all 'doctorate degrees obtained primarily through advanced supervised research written up as a thesis/dissertation'.

b. 'MPhil' – refers to all 'masters degrees obtained primarily through advanced supervised research written up as a thesis/dissertation' and all 'other postgraduate qualifications obtained primarily through advanced supervised research at masters degree level written up as a thesis/dissertation'.

c. 'Postgraduate research qualification' – refers to research degrees along with MPhil qualifications, with each as defined at a and b above.

d. 'Qualified' – a student is defined as qualifying when they have been awarded a postgraduate research qualification and the 'qualification obtained' has been returned through the individualised HESA student record.

42. Making use of a **transition matrix** that describes the progression patterns of the **transition population**, the method projects the study outcomes of a **starting population**. This publication makes available the projected outcomes of two cohorts of full-time research degree students: those who started their research degree in 2008-09 and those who started in 2009-10. **Sector-adjusted averages** are provided to support interpretation of the projected outcomes. Each of these components is described further at paragraphs 43 to 60.

Transition matrix and the transition population

43. We seek to project the proportions of a cohort of full-time research degree starters that are likely to be in each of three 'end states' 25 years following commencement of their research degree[17]. That is, the proportion expected to have qualified, transferred to another institution, or become absent from HE. Recent patterns of student behaviour at an institution are used in these projections.

44. To do this we firstly define two populations:

a. **Postgraduate students who were studying full-time on a research degree programme in 2009-10.** Their activity in this year is categorised according to the year of study recorded for the student, and we consider their status in 2010-11. This provides information about the progression, or transitions, of students from a period of full-time study on a research degree programme.

b. **A population of postgraduate students in 2008-09. This population includes those who have been identified as studying full-time on a postgraduate research**

[17] HEFCE 2007/28 showed that 10 years after the commencement of a full-time research degree, only 1 per cent of students with Research Council funding remained active on a research degree programme having not yet completed a research degree. The equivalent figure for full-time students who did not receive Research Council funding was 4 per cent. The 25 year period has been chosen as an overestimate of the amount of time that the majority of full-time research degree students should require to have reached one of the end states. This approach is intended to be generous in its attempt to capture successful study outcomes, and is consistent with that used by table series T5 of the Performance Indicators.

qualification in one of the five academic years prior to 2009-10[18], but excludes those who have been identified as having reached one of the end states prior to 2009-10[19]. These data help establish progression patterns that span all possible pathways into, through and out of full-time research degree study. For example, we require information about transitions made by research degree students following an extended period of writing up, or a period of part-time study. The status of these students in 2009-10 is identified, and we consider transitions from this status to a status we identify for them in 2010-11.

45. Together the populations described at paragraph 44 form the **transition population**. The activity of students in the transition population is considered in each of the two consecutive academic years in terms of the available classifications shown at Table A1 at Annex A. They enable us to see, for example, that one student may be classified as studying full-time in year one of a research degree programme in 2009-10 and as studying full-time in year two of a research degree programme at the same institution in 2010-11. Another student may be classified in the same way in 2009-10 and as studying part-time on a research degree programme in 2010-11. Another may be classified as studying part-time in year two of a research degree programme in 2009-10 and as having transferred to another institution in 2010-11.

46. These classifications allow us to arrange the progression patterns of the transition population in matrix form. For each institution, we observe the numbers and proportions of students who moved from one classification in a particular year to another classification in the following year. Arranging these observations in matrix form shows the patterns of movement between any and all combinations of classifications. Making an assumption that the progression patterns identified are typical for the institution, this matrix (referred to as a **transition matrix**) can be used to anticipate a student's state 25 years after the commencement of their studies.

47. We justify the assumption that progression patterns identified are typical for the institution by defining a transition population that is of sufficient size to ensure stability and robustness in respect of postgraduate populations. To this end we require supplementary information regarding transitions from 2008-09 to 2009-10. Equivalent populations to those described at paragraph 44 are derived to enable an additional focus on transitions between 2008-09 and 2009-10. These are incorporated within the transition population and are treated accordingly with respect to the derivation of the transition matrix.

48. Consider the examples given in paragraph 45.

a. The student classified as studying full-time in year one of a research degree programme in 2009-10 and as studying full-time in year two of a research degree programme at the same institution in 2010-11 is captured with STATE0 = 'Full-time research degree, year of study 1' and STATE1 = 'Full-time research degree, year of study 2'. Suppose an institution has seven students like this.

---

[18] As noted at paragraph 34, HIN information has been used to identify students who were active on a full-time postgraduate research qualification in one of the five academic years prior to 2009-10.

[19] Paragraph 35 describes how HIN information has been used to identify students considered to have reached an end state prior to 2009-10.

b. The student classified as studying full-time in year one of a research degree programme in 2009-10 and as studying part-time on a research degree programme in 2010-11 is captured with STATE0 = 'Full-time research degree, year of study 1' and STATE1 = 'Part-time research degree'. Suppose the institution has two students like this.

c. The student classified as studying part-time in year two of a research degree programme in 2009-10 and as having transferred to another institution in 2010-11, is captured with STATE0 = 'Part-time research degree' and STATE1 = 'Transfer'. Suppose the institution has three students like this.

If these three possibilities were the only ones that existed for students at an institution and we were to construct a transition matrix to contain the total of 12 students, we would obtain:

STATE 1 / STATE 0	Full-time research degree, year of study 2	Part-time research degree	Transfer
Full-time research degree, year of study 1	7	2	0
Part-time research degree	0	0	3

Starting population

49. For a cohort of full-time research degree starters in 2009-10 our analysis of transition into 2010-11 provides information on their actual movement between classifications over this period. The population (referred to as a **starting population**) can then be iterated through a further 24 years using the transition matrix. For each institution this process of iteration will tell us the number and proportion of the starters who are projected to have reached each of the end states 25 years after commencement of their studies. That is, the number and proportion of the starting population expected to have qualified, transferred to another institution, or become absent from HE.

50. We note that the starting population is specifically defined for the purpose of this analysis. It includes entrants to a full-time research degree programme of study in 2009-10: in other words, those whose commencement date for their programme of study fell within that academic year. It further includes those who were registered as an MPhil[20] student in 2009-10, and who showed evidence of studying for a doctoral qualification in 2010-11 while having had no MPhil qualification awarded.

[20] As described at paragraph 3 of Annex A, in this document the term 'MPhil' is used to refer to all 'masters degrees obtained primarily through advanced supervised research written up as a theses/dissertation' and all 'other postgraduate qualifications obtained primarily through advanced supervised research at masters degree level written up as a thesis/dissertation'.

Sector-adjusted averages

51. The interpretation of the projected outcomes described above is supported through the provision of sector-adjusted averages. A student's chance of completing their research degree is affected by a number of factors, and the sector-adjusted averages take into account the institution's research degree student profile in respect to some of these factors. They are intended to help avoid comparisons between institutions whose research degree student profiles are so different from one another that they should not be compared directly[21].

52. The sector-adjusted averages are not targets. If no factors were taken account of in the sector-adjusted average calculations each institution would have the same sector-adjusted average: the sector average. A number of factors are accounted for so rather they show the outcomes that might be projected if they reflected the sector averages after taking into account the impact of variations in the subject area of study as well as students' domicile, highest qualification held on entry and source of funding[22]. They give information about the figure that might be expected for an institution if no factors other than those allowed for were important. Where differences do exist, this may be due to the institution's performance, or due to some other factor which is not included in the sector-adjusted average.

53. Two symbols are used to show whether the difference between the projected outcome and the sector-adjusted average is significant: that is, where there is sufficient variation to be noteworthy. A plus sign, '+', indicates that the institution's projected outcome is significantly better than its sector-adjusted average and a minus sign, '-', indicates that the projected outcome is significantly worse than its sector-adjusted average. Such markings should be taken as an invitation to the institution to investigate possible causes for the differences that have been identified. If neither symbol is used, the projected outcome is similar to the sector-adjusted average, allowing for subject areas of study, and students' domiciles, highest qualifications held on entry and sources of funding.

Results

54. This section of the report provides results arising from the implementation of the method described at paragraphs 40 to 53 (and in detail at Annex A). The tables that accompany this section relate separately to the 2008-09 and 2009-10 cohorts of full-time research degree starters, and the study outcomes that have been projected for these two cohorts of starters. A number of institutions have identified errors in their underlying data submitted to HESA which led to the results generated for them being misleading. Results for these institutions have been suppressed throughout, and their figures have been excluded from all sector-level results provided in the accompanying tables, as well as from calculations to determine sector-adjusted averages.

---

[21] To compare an institution's projected outcomes to the sector, the benchmark should be used in preference to the overall sector average.

[22] In order for a factor to be accounted for in the calculation of the sector-adjusted averages it must have a number of characteristics. In particular they should: be associated with what is being measured; vary significantly from one institution to another; and not be in the institution's control, and not be part of their performance.

55. Below is a list of tables that show the results generated by the method described in this publication. Each table listed has been provided for both the 2008-09 and 2009-10 cohort of starters:

a. Table 1 shows the complete transition matrix for the sector. It considers patterns of progression for research degree students across all English HEIs with full-time research degree students. This transition matrix is in a simplified form that does not take account of the institutions at which students were registered.

b. Table 2 shows the distribution of the starting population across year of study on the programme recorded for students in that population.

c. Table 3 shows projected outcomes for full-time research degree starters at English institutions, along with the sector-adjusted averages and significance markers calculated to assist in the interpretation of the projected outcomes. This is the main output of the method.

d. Table 4 shows the projected study outcomes of full-time research degree starters across English HEIs as a whole. The method used to produce this table is the same as that used to produce Table 3, with the exception that all institutions in the sector are treated as one large institution, rather than being considered individually.

e. Tables 5 and 6 show the equivalents to Tables 3 and 4, where projections have been made over a seven-year period (as opposed to the 25-year period considered for the main output of the method). Again they are provided along with the sector-adjusted averages and significance markers calculated to assist in the interpretation of the projected outcomes.

As described at paragraph 7 of Annex A, we acknowledge that the time scale we have used for the main output of the method is much longer than that considered by the earlier RDQR method (seven years for full-time research degree students). The 25-year period is intended to be generous in its attempt to capture successful study outcomes. Tables 5 and 6 seek to provide some comparability with the RDQR method, though, by design, the new method is very different in its consideration of a different population of research degree students. The new method may also be less statistically robust when projecting over a shorter time scale, particularly for institutions with small populations of research degree students.

56. The tables described at paragraph 55 are provided as a separate Excel file relating to each of the 2008-09 and 2009-10 starting cohorts for ease of access and interpretation. Their layout and content are described in paragraphs 57 to 60.

57. The first two columns of Table 3 show the total number of full-time research degree students at each institution in the starting population in the 2009-10 reference year, and what percentage of these are mature. The next columns detail the projected outcomes. The proportions of the starters expected to obtain a research degree award, obtain an MPhil award, transfer to another institution, or to neither transfer nor obtain any qualification are all provided. Sector-adjusted averages are included for each of these outcomes, and should be interpreted as described at paragraphs 51 to 53.

58. For the percentage projected to neither transfer nor obtain any research qualification, and for the percentage projected to qualify with a research degree, the standard deviation and significance marker described at paragraph 53 both accompany the projected outcome and the sector-adjusted average. The proportion of starters for whom the projected outcome is 'not known' is provided at the last column of Table 3. The calculation of a sector-adjusted average is not appropriate for this outcome.

59. Columns containing the projected outcome and the sector-adjusted average are shown in bold. The numbers of starters shown in the first column are rounded to the nearest five. Percentages are calculated from unrounded data, and themselves are not subject to rounding. If a percentage is calculated on the basis of a population containing fewer than 20 individuals it is suppressed and represented as '...'.

60. Table 4 shows the projected study outcomes of full-time research degree students across the HE sector as a whole. Again we consider those who started a course in 2008-09. The method used to produce this table is the same as that used to produce Table 3. However, all institutions in the sector are treated as one large institution, rather than being considered individually. It follows that the concept of a student transferring from one institution to another is not relevant to this analysis. Those who transfer between HEIs are treated as continuing, and the percentage of students projected to transfer will be zero. Students may be registered at multiple institutions, so the method is modified further to remove duplicates of students across institutions. As a result, the transition and starting populations used to produce Table 4 will be smaller than those used to produce Table 3.

Annex A: Technical explanation of the method

1. This annex provides a technical description of the method outlined in the main body of this report to project qualification rates from postgraduate research degree study. It is aimed at readers with in-depth knowledge of the data.

2. The description makes reference to a number of student and course characteristics, defined using variables collected in the HESA individualised student records[23]. Algorithms and data definitions relating to these characteristics, and the HESA variables associated with them, are provided at Annex B.

3. The description that follows makes use of terminology which is defined as follows for the purposes of this report:

- 'Research degree' – refers to all 'doctorate degrees obtained primarily through advanced supervised research written up as a thesis/dissertation'.
- 'MPhil' – refers to all 'masters degrees obtained primarily through advanced supervised research written up as a thesis/dissertation' and all 'other postgraduate qualifications obtained primarily through advanced supervised research at masters degree level written up as a thesis/dissertation'.
- 'Postgraduate research qualification' – refers to research degrees along with MPhil qualifications, with each as defined at a and b above.
- 'Qualified' – a student is defined as qualifying when they have been awarded a postgraduate research qualification and the 'qualification obtained' has been returned through the individualised HESA student record.

Overview

4. The method described is based on that currently employed by table series T5 of the UK higher education (HE) performance indicators (PIs)[24] to project the study outcomes of full-time first degree study. We make use of 2009-10 as the academic year of interest to describe the manner in which the study outcomes of full-time EU-domiciled research degree starters are projected. For clarity, the academic year of interest is referred to as the 'reference year' for the remainder of this document.

5. In an approach consistent with that used for the PIs, students are not included if they left the programme of study within 50 days of commencement. There is no adjustment for individual reasons for non-qualification.

6. Recent progression patterns at the institution are used to project what proportion of the starting cohort from the reference year are likely to be in each of the 'end states' after a period of

[23] Full descriptions of the variables collected in the HESA individualised student records are available at www.hesa.ac.uk, under Student Stream, Student Collection (for the academic year of interest), and Field list and detail.

[24] For further information see www.hesa.ac.uk under Performance Indicators.

25 years (that is, having gained a qualification, transferred to another institution, or been absent from HE for two consecutive years).

7. The 25-year period has been chosen as an over-estimate of the amount of time that the vast majority of full-time research degree students should require to have reached one of the end states[25]. Having a 25-year period increases the statistical robustness of the method, and is consistent with the approach used in the generation of table series T5 of the PIs[26]. However, we acknowledge that the time scale we have used is much longer than that considered by the earlier research degree qualification rates (RDQR) method (seven years for full-time research degree students). A projection of the proportion of the starting cohort that are likely to be in each of the 'end states' after a period of seven years is given alongside the results of the new method to provide some comparability with the RDQR method. However, by design the new method is very different in its consideration of a different population of research degree students and as such projected outcomes are not directly comparable to the RDQRs.

8. To ensure that the data used in the projections are robust, we define the progression pattern of a research degree student by reference to four years of data. These are:

a. The 'reference year'– the year in which the cohort of interest commence their research degree programmes, and to which the projections relate.

b. The two academic years prior to the reference year.

c. The academic year following the reference year.

As stated at paragraph 4, the reference year for the purposes of the technical explanation given here is 2009-10.

9. For each institution, the method described by this publication project the proportions of the reference year starting cohort that are likely to be in each of the end states after 25 years. Alongside these projected proportions we present a sector-adjusted average (sometimes called a benchmark) and, where appropriate, an indicator which highlights significant variation of the actual projection from the sector-adjusted average.

Implementation of the method

10. Implementation of the method involves:

a. The identification of the starting population, as defined at paragraphs 11 to 14.

b. The identification of the transition population, as defined at paragraphs 15 to 43, and derivation of the 'states' of students within that population in each year of the transition period.

c. The construction of a transition matrix for each institution to describe recent patterns of progression, in accordance with the description given at paragraphs 44 to 51.

---

[25] HEFCE 2007/28 showed that 10 years after the commencement of a full-time research degree, only 1 per cent of students with Research Council funding remained active on a research degree programme having not yet completed a research degree. The equivalent figure for full-time students who did not receive Research Council funding was 4 per cent.

[26] Table series T5 of the PIs projects the outcomes of full-time first degree study after 15 academic years.

d. Projection of study outcomes through the iterative process described at paragraphs 52 to 58, whereby a cohort of starters are projected through the system over a period of 25 years to reach a 'final state' that can be categorised into one of five potential study outcomes.

e. Calculation of sector-adjusted averages and significance markers to accompany the projected outcomes and aid interpretation of institutional results, as outlined at paragraphs 59 to 64.

The starting population

11. A population of students who commence their research degree in the reference year is established so that their study outcomes can be projected. Such a population is referred to as the starting population and is defined with reference to the reference year and the academic year following the reference year.

Definition of the starting population

12. The starting population consists of entrants to a full-time research degree in the reference year. That is, students who were studying full-time for a research degree in the reference year whose commencement date for their programme of study fell within that academic year.

13. Students who convert a full-time MPhil qualification (defined at paragraph 3 of this annex) to a research degree are also included within the starting population. We seek to include these students at the earliest time that doctoral study can be identified, and do this by locating them on a full-time MPhil in the reference year and looking for evidence of them studying on a research degree in the following year while having had no MPhil qualification awarded. Because one academic year following the reference year is required to identify this conversion of their qualification aim, they are retrospectively added to the starting population. We do not take account of the commencement date of the programme of study for these students, and consider their qualification aim in the reference year to be that of a research degree for the purposes of this analysis.

14. A student who was registered as a full-time MPhil student in the reference year, and shows evidence of studying for a doctoral qualification in the following year, is not included within the starting population if that student is recorded as having qualified with an MPhil award from the reference year. In such cases, the award of an MPhil qualification suggests that the student's progression through postgraduate study should be treated as a standalone MPhil, followed by a research degree. Patterns of progression will differ to those of students who convert from an MPhil to a research degree qualification aim, and exclusion from the starting population defined with regard to the reference year is therefore appropriate. Such students would fall within the starting population of research degree students in the subsequent year.

The transition population

15. To establish the progression patterns of research degree students we define a 'transition population'. This population consists of students who were registered with a postgraduate research qualification aim (defined at paragraph 3 of this annex), for whom we consider their 'states' in each of the relevant years described at paragraph 8. For each institution, we observe the numbers of students who moved from one state in a particular year to another state in the following year. This enables the construction of a transition matrix, showing the recent patterns of

progression between any and all combinations of states. An assumption that the pattern of progression identified is typical for the institution allows us to make use of the transition matrix to anticipate a student's state 25 years after the commencement of their studies.

16. In mathematical terms, the projection of a student's progression through a research degree over a period of 25 years is equivalent to multiplying a scalar matrix of starters by a matrix of students in the transition population a total of 24 times.

Definition of the transition population

17. To determine the transition population we initially define three broad conditions in which a student with a postgraduate research qualification aim can be in any one year. These are informed by the student's mode of study, level of study and qualification aim, and institution:

a. Condition **1**: The student is active at the institution as a full-time student with a research degree qualification aim.

b. Condition **2**: The student is active at the institution with a postgraduate research qualification aim but is not in condition **1**. They could be writing up, studying full- or part-time with an MPhil qualification aim, or writing up or studying part-time with a research degree qualification aim.

c. Condition **3**: The student is not active at the institution as a student with a postgraduate research qualification aim. They could be active at the institution as a student with a postgraduate taught or undergraduate qualification aim, they could be at another institution, or they could be absent from HE entirely (including being absent as a result of having qualified with a postgraduate research qualification).

18. For students in condition **2** or **3** in the academic year prior to the reference year we make use of HIN[27] information in relation to the five-year period prior to the reference year[28]. This process enables us to determine if the student was active on an instance of full-time postgraduate research study at any time during the five-year period, and thus ensures that a full-time research degree student continues to contribute to successive executions of the method[29] through their inclusion in the transition population until they can be considered to have reached one of the end states.

19. Additionally, the five-year HIN linking allows us to consider an instance of full-time postgraduate research degree study from either the academic year in which they commenced

[27] The HIN is a combination of three fields from the HESA individualised student records (the HESA unique student identifier, the UK provider reference number and the student instance identifier) which uniquely identify a student on a course (or instance of study). The HIN forms a year-on-year linking mechanism which can be used to track the student's progression on the instance from one year to the next. It follows that the HIN enables a student to be tracked from commencement of the instance through to completion.

[28] It is important to note that, to this end, minimal information has been retained from the HESA student records for each of these five academic years: namely, the HIN information, and the student's commencement date of study, mode of study, qualification aim and any qualification obtained.

[29] That is, when the method is later used to project the outcomes of students starting a full-time research degree in 2010-11, 2011-12, and so on.

the instance of study, or the academic year five years prior to the one for which study outcomes are projected. This in turn enables us to exclude from the method students who have previously reached one of the end states and should not therefore be included within the transition population. For example, a student identified as having been inactive for two or more consecutive academic years are considered throughout the method to have 'become absent from HE': that is they have reached one of the end states. Similarly a student who was awarded the qualification prior to the academic year for which the study outcomes are projected has reached a 'qualified' end state. In both cases if the student returned to full-time study inclusion of their onward study patterns would distort the method. This process seeks to ensure that a full-time research degree student's contribution to successive executions of the method ceases beyond their arrival at one of the end states.

20. If a student is recorded more than once at an institution, with each instance relating to a postgraduate research qualification aim, duplicate records are removed. A record that indicates progression or qualification is prioritised over others, with the most relevant record retained in the definition of the transition population for onward analysis.

21. A transition population then consists of students who were active on a full-time research degree in the reference year, plus students who were active on a full-time research degree course in one of the five years prior to the reference year, excluding those who had previously obtained a research degree award or reached an end state. In other words, it is defined as:

i. All students who were in condition **1** at the institution in the reference year; plus

ii. Students who were in condition **2** or **3** in the reference year and whose HIN information indicates that they were active on a full-time postgraduate research qualification in one of the five years prior to the reference year. It is restricted to students who had not obtained a postgraduate research qualification or previously reached an end state.

22. The approach described in paragraphs 18, 19 and 21 ensures that the transition population for whom we consider progression patterns is all and only those who have been active on a full-time postgraduate research qualification in one of the five years prior to the reference year, having not reached one of the end states prior to the reference year. Students who studied the entirety of a research degree on a part-time basis are likely to have recent patterns of study that are particularly disparate to those of students who studied full-time. This leads to a requirement to draw on the HIN information rather than to include all students who were in condition **2** or **3** in the reference year. The approach enables consideration of students whose progression through a full-time postgraduate research degree involved two or more consecutive years of writing up or part-time study on a research degree in the year prior to the reference year and the reference year.

23. For analysis of first degree outcomes the definition of a transition population is similar to the one given at paragraph 21, but for full-time first degree students and restricted to consider only the year prior to the reference year in respect of criteria outlined in (ii) of the definition. This definition generates a population that is large enough to establish recent patterns of progression among first degree students at an institution that are complete, stable and robust. Relatively small populations of research degree students in a given year mean that a transition population informed by only the reference year and one year prior to the reference year is of insufficient size for analysis of research degree outcomes. Additionally, a transition population defined on this

basis would be less able to account for postgraduate research programmes having longer periods of writing up or inactivity prior to completion.

24. To ensure stability and robustness of a transition matrix and associated assumptions regarding the progression patterns it identifies, we require an extended definition of the transition population: one that is inclusive of a further year's research degree students. The transition population defined for this analysis is the one given at paragraph 21 plus a population defined using the same process as described at paragraph 21 but treating the year prior to the reference year as if it were the reference year. In other words, it is defined as:

iii. All those students who were in condition **1** in the year prior to the reference year and were in condition **1** in the reference year, or had obtained a research degree; plus

iv. Students who were in condition **2** or **3** in the year prior to the reference year and whose HIN information indicates that they were active on a full-time postgraduate research qualification in one of the four years prior to the year prior to the reference year. It is restricted to students who had not obtained a postgraduate research qualification or previously reached an end state.

25. The resulting transition matrix is then informed by progression patterns between the reference year and the year that follows, and between the year prior to the reference year and the reference year.

Definition of the 'states' used to construct the transition matrix

26. To calculate the transition matrix the transition population must be considered further. For each student in the transition population we define their 'states' in different years. These need to be specified more fully than the broad conditions described at paragraph 17: they are informed by the student's mode of study, level of study and qualification aim, institution, year of study on the programme and if applicable, qualifications obtained.

27. Paragraphs 23 to 25 describe the need to consider two transition periods: that is, to consider transitions both between the reference year and the year that follows, and between the year prior to the reference year and the reference year. This process is simplified by splitting the transition population into two groups of students.

<u>Transitions between the reference year and the following year</u>

28. Knowledge of transitions between the reference year and the year that follows is required for those who satisfied the criteria (i) and (ii) as described at paragraph 21. For those who were in condition **1** in the reference year we define their state in the reference year ('state 0') and consider the state that they moved into in the following year ('state 1'). Similarly for those who were in condition **2** or **3** in the reference year and met the criteria outlined in (ii) at paragraph 21.

29. For some of these students we require additional information regarding the year of study on the programme recorded before we can assign them to a 'state 0' value. For others, their 'state 0' value is determined through closer consideration of their condition in the reference year.

30. The year of study on the programme that was recorded in the reference year must be identified for students who were in condition **1** in that year. Further, if a student was recorded as writing up and in condition **2** in the reference year, having met the criteria outlined in (ii) at

paragraph 21, we need to know the year of study on the programme that was recorded for the student in the reference year.

31. The values of 'state 0' that are possible for the students described at paragraph 21 are shown in Table A1. They classify the students' activity in the reference year.

Table A1 'State 0' values available for classification of a student's activity in the reference year

	'State 0' value	**Further detail (as required)**
a.	Full-time research degree, year of study 1	
b.	Full-time research degree, year of study 2	
c.	Full-time research degree, year of study 3	
d.	Full-time research degree, year of study 4	
e.	Full-time research degree, year of study 5	
f.	Full-time research degree, year of study 6	
g.	Full-time research degree, year of study 7 or above	
h.	Writing-up research degree, year of study 0, 1 or 2	(having previously been recorded on a full-time research degree)
i.	Writing-up research degree, year of study 3 or above	(having previously been recorded on a full-time research degree)
j.	Full- or part-time MPhil	(having previously been recorded on a full-time research degree)
k.	Part-time research degree	(having previously been recorded on a full-time research degree)
l.	Transfer	Transferred to another HEI
m.	Inactive	Not currently in postgraduate research HE, but was recorded on a full-time research degree in previous year

32. A 'state 1' value is assigned to the each of the students meeting the criteria described at paragraph 21, with this assignment made on the basis of the student's activity in the year following the reference year. In the year that follows the reference year we need to know:

- if they had qualified with a research degree or an MPhil qualification in the reference year
- which year of study on the programme was recorded if they were still at the institution and whether they were still studying full-time for a research degree, or recorded on a writing-up mode of study
- if they had transferred to an MPhil or part-time programme
- if they had transferred to another institution.

33. If a student was not recorded as a full-time, part-time or writing-up research degree student in the reference year ('state 0' is one of j., k., or m.) and they gained a qualification in the reference year then that student is considered to have qualified in terms of their 'state 1' value. Otherwise, if such a student did not gain a qualification in the reference year, and there was no record of them as a student with a postgraduate research qualification aim in the following year (condition **3**) then that student is assumed to be 'absent' in terms of their 'state 1' value. A student who was recorded as a full-time, part-time or writing-up research degree student in the reference year, but for whom no record can be found the following year, is initially deemed to become 'inactive' in terms of their 'state 1' value in the year following the reference year. Such students had the possibility of returning to the institution in the following year. It is only if students cannot be traced for a second year that they are categorised as 'absent'.

34. Note that a student who is reported to have qualified with a research degree or an MPhil qualification in the year that follows the reference year does not have a 'state 1' value indicating that they have qualified. The method described here seeks to project outcomes in relation to starters in academic year 2009-10 as the reference year and is informed by students' progression patterns between this year and the following 2010-11 academic year. As such, it will identify students who were awarded a research degree qualification in the 2009-10 academic year and assign them with a 'state 1' value of 'Qualify with research degree'. Students who were awarded a research degree award in the 2010-11 academic year will be identified by the method with a 'state 1' value of 'Qualify with research degree' when it is calculated at a future date in relation to starters in academic year 2010-11.

35. The approach described at paragraph 34 seeks to minimise the potential to overstate the propensity to qualify with a research degree. That is, rates of qualification would be inflated should the method count all awards of research degree qualifications in either the reference year or the year following. The latter would be counted in each of two successive applications of the method. This application of the method requires the information about their activity at the institution during the year following the reference year to ensure that the transition matrix is fully informed about progression into a final year of study, up to and including the award of a research degree.

36. The values of 'state 1' that are possible for the students described at paragraph 21 are shown in Table A2.

Table A2 'State 1' values available for classification of a student's activity in the year following the reference year

	'State 1' value	**Further detail (as required)**
A.	Full-time research degree, year of study 1	
B.	Full-time research degree, year of study 2	
C.	Full-time research degree, year of study 3	
D.	Full-time research degree, year of study 4	
E.	Full-time research degree, year of study 5	
F.	Full-time research degree, year of study 6	

G.	Full-time research degree, year of study 7 or above	
H.	Writing-up research degree, year of study 0, 1 or 2	(having previously been recorded on a full-time research degree)
I.	Writing-up research degree, year of study 3 or above	(having previously been recorded on a full-time research degree)
J.	Full- or part-time MPhil	(having previously been recorded on a full-time research degree)
K.	Part-time research degree	(having previously been recorded on a full-time research degree)
L.	Transfer	Transferred to another HEI
M.	Inactive	Not currently in postgraduate research HE, but was recorded on a full-time research degree in previous year
N.	Qualified with research degree award	
O.	Qualified with MPhil award	
P.	Absent	Absent from postgraduate research HE

37. Several of the states shown in Table A2 are what are known as ‘sink states’. Such states are characterised by the fact that once a student enters that state they do not progress any further through the system as far as the institution is concerned. By default, four of the 16 states shown are sink states, namely ‘qualified with a research degree award’, ‘qualified with an MPhil’, ‘transfer’ and ‘absent’. Students cannot return to the institution, as a continuous part of their progress towards a research degree study intention, from any of these states.

38. Other states shown in Table A2 have the potential to become sink states in particular circumstances. These may include an institution having a very small number of students falling into its transition population, or having made major changes, either to the format of its research degree programmes or to the way it records that format. Such circumstances may lead to a student moving into a state and not out of it. For example, an institution did not allow part-time research degree study, but allowed it for the first time in 2010-11. Because information is gained from consideration of movements from 2009-10 to 2010-11 and from 2008-09 to 2009-10, we may have a record of students moving into a part-time research degree in 2010-11, but no record of students moving out of it.

Transitions between the year prior to the reference year and the reference year

39. As described at paragraphs 23 to 25 the transition matrix requires supplementary information regarding progression between the year prior to the reference year and the reference year to ensure sufficient data for stability and robustness of that matrix. To this end, knowledge of transitions between the year prior to the reference year and the reference year is required for those who satisfied the criteria (iii) and (iv) as described at paragraph 24.

40. For those who were in condition **1** in the year prior to the reference year and were in condition **1** in the reference year, or had obtained a research degree award, we define their state in the year prior to the reference year ('state 0'[30]) and consider the state that they moved into in the reference year ('state 1'). Similarly for those who were in condition **2** or **3** in the year prior to the reference year and met the criteria outlined in (iv) at paragraph 24.

41. In line with the transitions described at paragraphs 28 to 31, we require additional information regarding the year of study on the programme recorded in the year prior to the reference year before we can assign some of these students to a 'state 0' value. Namely for those who were in condition **1** in the year prior to the reference year, and for those recorded as writing up and in condition **2** in the year prior to the reference year, having met the criteria outlined in (iv) at paragraph 21. For others, their 'state 0' value is determined through closer consideration of their condition in the year prior to the reference year. The values of 'state 0' that are possible for the students described at paragraph 24 are then consistent with those shown in Table A1.

42. To assign 'state 1' values we need to consider details of their institution, mode of study, level of study and qualification aim, year of study on the programme and qualifications obtained, each in respect of the reference year. Such details are considered in the same manner as outlined at paragraphs 32 to 36, particularly with regards to the assignment of students to the 'inactive' and 'absent' states. The values of 'state 1' that are possible for the students described at paragraph 24 are also consistent with those shown in Table A2.

43. In paragraphs 28 to 42 we have described the definition and consideration of the transition population in terms of two separate groups of students. The transition population itself unites these two groups, removing any duplication that arises from consideration of students meeting criteria (ii) and (iii) as described at paragraphs 21 and 24. The united transition population is then of sufficient size to ensure stability and robustness in onward calculations.

Construction of the transition matrix

44. To summarise the progression patterns of students at an institution requires an aggregation of the students within the transition population and the construction of a transition matrix. The transition population is considered in its entirety, and the transition matrix does not differentiate between the two transition periods informing the patterns of progression: knowledge of transitions between the reference year and the year that follows is summarised in conjunction with knowledge of transitions between the year prior to the reference year and the reference year.

45. Each row of the transition matrix represents a possible state in the first year of a transition ('state 0'), while each column represents a state in the following year ('state 1'). The values in the cells of each row are the proportions of students from 'state 0' who make the transition to 'state 1'.

46. For example, we might know that 100 students in the transition population fell into the 'full-time research degree, year 1' categorisation of 'state 0' (row). We might also know that 10 of

[30] Retention of the 'state 0' and 'state 1' terminology simplifies later construction of the transition matrix where we do not differentiate between the transition periods used to identify patterns of progression.

these students move into the ‘full-time research degree, year 2’ categorisation of ‘state 1’ (column). The value at the intersection of this particular row and column would be the proportion of students who were on year of study 1 of a full-time research degree who then moved to year of study 2: that is, 10 / 100 = 0.1, or 10 per cent.

47. In technical terms, the values of the transition matrix T are found as

$$t_{ij} = \frac{n_{ij}}{n_i}$$

where n_{ij} is the number of students at the institution who were in state i in the first year of a transition and state j in the following year, and n_i is the number of students at the institution who were in state i in the first year of the transition.

48. Rows of the transition matrix which represent ‘sink states’ contain only one non-zero value, on the diagonal, as all students in such a state will, by definition, not move out of that state. However, our definition of other modes and levels of study, and of inactive states, means that we can allow for students who returned to full-time research degree study (for example, from a period of part-time study), or who were awarded a research degree qualification (for example, following a period where the student was recorded as writing up).

49. Construction of the transition matrix assumes that the probability of moving from ‘state 0’ to ‘state 1’ is the same, regardless of how the state was reached. For example, take two students who fell into the ‘full-time research degree, year 4’ categorisation of ‘state 0’, one who spent one of their years of study one to three in a writing-up state, and one who progressed through years one to three continually in full-time research degree study. Each of these students is assumed to have had the same probability of progressing into the ‘qualified with research degree award’ categorisation of ‘state 1’. Although the assumption is not completely accurate, any errors it introduced as a result are likely to be small.

50. We note also that the transition matrices used to project the study outcomes of research degree starters only consider patterns of movement for a particular institution. In this regard, a student who transfers to a different institution is counted as being at the end of their course as far as the original institution is concerned. The method described by this publication takes no account of the number of these transferring students who might go on to obtain a qualification at their new institution.

51. As an example of the transition matrices calculated for each institution, Table 1 – which is described within the main body of this report – shows the transition matrix that considers patterns of progression for research degree students across all HEIs in the sector who identify full-time research degree students. This is a simplified form that does not take account of the institutions at which students were registered. It is calculated in respect of 2008-09 as the reference year. It shows that most students in a given ‘state 0’ were found to move to the next year of the programme, were recorded on a writing-up mode of study, or qualified.

Projected outcomes

52. An underlying principle of this method is the assumption that the transition matrices described at paragraphs 44 to 51 identify patterns of progression that are typical for an institution. Such an assumption then allows us to make use of these transition matrices to anticipate the end states of a cohort of research degree starters 25 years after their studies begin. As stated at paragraph 16, in mathematical terms this projection of a student's progression through a research degree over a period of 25 years is equivalent to multiplying a scalar matrix of starters by a matrix of students in the transition population a total of 24 times.

53. For each institution we describe the starting population defined at paragraphs 11 to 14 by a vector n which gives the number of starters falling into each categorisation of 'state 0'. Given that they have satisfied the definition of a starter, they will be on a research degree programme and most will be in year of study 1. Where an institution has students who fall into the starting population on account of their showing evidence of converting from an MPhil to a research degree qualification aim, they may have starters in later years of study.

54. The starting population defined with regards to 2009-10 is shown at Table 2 – which is described within the main body of this report – split by their year of study on that programme in the reference year. As might be expected, it shows that the majority of students (more than 94 per cent of the cohort) started on the first year of study on a research degree programme.

55. Computing the progression of the starting cohort through a given number of years is achieved by multiplying the vector of starters (n) by the transition matrix (T) a given number of times. After i iterations the starting cohort will be in state $n.T^{i-1}$. For the purposes of this analysis we set i to be 25, where 25 iterations represents a period of 25 years. This is based on the assumption that a 25-year period is an over-estimate of the amount of time that the majority of full-time research degree students require to have reached an end state: all students who are ever going to qualify or transfer from their research degree course will have done so within 25 years from commencement.

56. This method generates a state vector $q = n.T^{i-1}$ which identifies a proportion of the starting population that is expected to arrive in each 'final state' 25 years after commencement in 2009-10. These 'final states' are those in which we might expect to observe the students 25 years after commencement of their course. They are consistent with those shown at Table A2, and can each be attributed to one of five categorisations describing a student's study outcome, or 'end state'. Table A3 shows the categorisations of 'final states' into a projected study outcome.

Table A3 Categorisation of 'final states' as projected study outcomes

	'Final state' value	Description	Projected study outcome
A.	Full-time research degree, year of study 1		Not known
B.	Full-time research degree, year of study 2		Not known

C.	Full-time research degree, year of study 3		Not known
D.	Full-time research degree, year of study 4		Not known
E.	Full-time research degree, year of study 5		Not known
F.	Full-time research degree, year of study 6		Not known
G.	Full-time research degree, year of study 7 or above		Not known
H.	Writing up research degree, year of study 0, 1 or 2	(having previously been recorded on a full-time research degree)	Not known
I.	Writing up research degree, year of study 3 or above	(having previously been recorded on a full-time research degree)	Not known
J.	Full- or part-time MPhil	(having previously been recorded on a full-time research degree)	Not known
K.	Part-time research degree	(having previously been recorded on a full-time research degree)	Not known
L.	Transfer	Transferred to another HEI	Transfer
M.	Inactive	Not currently in postgraduate HE, but was recorded on a full-time research degree in previous year	Not known
N.	Qualified with research degree award		Research degree awarded
O.	Qualified with MPhil award		MPhil awarded
P.	Absent	Absent from postgraduate HE	Neither award nor transfer

57. Table A3 shows that there are values in the state vector (q), or 'final states', which are not 'sink states'. In other words, even 25 years after commencement of their research degree, there are some circumstances when a student has not reached one of the 'final states' relating to either a qualification having been awarded; them having transferred to a different institution; or them having become absent from HE. In such cases the outcome is recorded as 'not known'.

58. High numbers of students categorised with a 'not known' study outcome may be a result of an institution having made major changes, either to the format of its research degree programmes or to the way it records that format. Such changes may lead to a student moving into a state and not out of it. For example, an institution had no MPhil programmes of study, but introduced such programmes so that the first MPhil study was to be in 2010-11. We may have a record of students moving into an MPhil in 2010-11, but no record of students moving out of it, because this information is gained from consideration of movements from 2009-10 to 2010-11,

and 2008-09 to 2009-10. This would lead to a number of unknown outcomes, and a large proportion of students shown to have a projected outcome of 'not known'. Therefore where large numbers of students are categorised as 'not known' the projected outcomes should be treated with caution.

Sector-adjusted averages and significance markers

59. A student's chance of completing their research degree is affected by a number of factors, and we provide sector-adjusted averages for each institution to support interpretation of the projected outcomes of research degree study. The sector-adjusted averages are intended to help avoid comparisons between institutions whose research degree student profiles are so different from one another that they should not be compared directly. To compare an institution's projected outcomes to others in the sector, the sector-adjusted average should be used in preference to the overall sector average. This will take into account the institution's research degree student profile.

60. The sector-adjusted averages are not targets. If no factors were accounted for in the sector-adjusted average calculations each institution would have the same sector-adjusted average: the sector average. A number of factors are accounted for, so they show the outcomes that might be projected if they reflected the sector averages after taking into account the impact of variations in:

- the subject area of study
- students' domicile
- highest qualification held on entry
- source of funding[31].

They give information about the sort of values that might be expected for an institution if no factors other than those allowed for were important. Where differences do exist, this may be due to the institution's performance, or due to some other factor which is not included in the sector-adjusted average.

61. In terms of the factors accounted for in the calculation of the sector-adjusted averages we consider:

- two categorisations of subject area of study (science-based subject areas and non-science-based)
- two categorisations of students' domicile (UK; and EU, Channel Islands and Isle of Man)
- three categorisations of qualifications on entry (postgraduate qualifications; first degree and graduate equivalent qualifications; and qualifications at neither postgraduate or first degree level)
- two categorisations of source of funding (UK Research Council and other sources).

[31] For a factor to be accounted for in the calculation of sector-adjusted averages it must have a number of characteristics. In particular they should: be associated with what is being measured; vary significantly from one institution to another; and not be in the institution's control, and not be part of their performance.

62. Two symbols are used to show whether the difference between the projected outcome and the sector-adjusted average is significant. A plus sign, '+', indicates that the institution's projected outcome is significantly better than its sector-adjusted average and a minus sign. '-', indicates that the projected outcome is significantly worse than its sector-adjusted average. Where a difference is marked, the difference between the projected outcome and the sector-adjusted average is three or more standard deviations away from zero and greater than three percentage points: this reflects our judgement that variation of this extent is sufficient to be noteworthy. Such markings should be taken as an invitation to the HEI to investigate possible causes for the differences that have been identified. If neither symbol is used, the institution can say that its projected outcome is similar to the sector average – once subject areas of study, students' domiciles, highest qualifications held on entry and sources of funding have been taken into consideration.

63. The matrix used in the calculation of the sector-adjusted averages is produced in a similar way to the institutions' transition matrices, but using all the students in the sector. Weights are given to each student which reflect the relative importance at the institution of their subject area of study, domicile, highest qualification held on entry and source of funding. It is the sums of these weights, rather than the actual number of students, that are used to calculate the values in the matrix. The values in the sector-adjusted average matrix B are defined as

$$b_j = \frac{n_j / n}{N_j / N}$$

where n_j is the number of students in category j at the institution, N_j is the number of students in category j in the sector, and n and N are the total numbers of students in the population in the institution and the sector respectively.

64. The standard deviation of the difference between the actual projected outcome and the sector-adjusted average is calculated for the 'neither transfer nor obtain any qualification' outcome category. If there are n starters at an institution, π_d is the proportion who 'neither transfer nor obtain any qualification', and the sector-adjusted average for the 'neither transfer nor obtain any qualification' outcome category is $(100 \times \pi_d)\%$, then the standard deviation of the proportion in category d has been calculated as

$$\sqrt{\frac{\pi_d \times (1 - \pi_d)}{n}}.$$

This assumes that the n starters have been randomly chosen from a large population in which π_d of the population is in category d, and $(1 - \pi_d)$ is in one of the other categories. A standard deviation is calculated in the same way in relation to the 'qualify with research degree' outcome category.

Annex B: Data definitions

Purpose

1. This annex describes the algorithms and data definitions that were used in the generation of the rates of postgraduate research degree qualification. It is aimed at readers in institutions with in-depth knowledge of the data.

2. Readers are advised to consider this annex in conjunction with Annex A, which provides a technical explanation of the method used to generate the rates of postgraduate research degree qualification. It is also advisable to have access to the HESA student record coding manuals for the years 2006-07 to 2010-11 to hand.

Linking students within and across years

3. Annex C provides an outline of the overall linking process used to track individual students within and across years of the HESA individualised student records. The process of linking students in this way is the same as that used by HESA in the generation of the PIs.

HESA fields used to generate rates of postgraduate research degree qualification

4. The fields detailed in Table B1 were used to generate the rates of postgraduate research degree qualification. Throughout this annex, fields taken from the HESA return or derived as part of the rates calculation are shown in upper case using the names given in Tables B1 and B2.

5. Table B1 provides details of the HESA fields used in the generation of the rates of postgraduate research degree qualification.

Table B1 HESA fields used in the generation of rates of postgraduate research degree qualification

Entity	HESA field name	Description
Course	COURSEAIM / QUALAIM	General qualification aim of course
Course	TTCID	Teacher training course
Entry profile	QUALENT2	Highest qualification on entry
Instance	CAMPID	Campus identifier
Instance	COMDATE	Start date of instance
Instance	ENDDATE / DATELEFT	End date of instance
Instance	EXCHANGE	Exchange programmes
Instance	INSTCAMP	Institution’s own campus identifier
Instance	LOCSDY	Location of study
Instance	MSTUFEE	Major source of tuition fees
Instance	† NUMHUS	Student instance number

Instance	STULOAD	Student instance FTE
Instance	TYPEYR	Type of instance year
Instance	YEARSTU	Year of student on this instance
Institution	† UKPRN	UK provider reference number
N/A	SPCSTU	Special students
Student	† HUSID	Student identifier
HESA derived field	††XDOMHM01	Domicile (participation by nation)
HESA derived field	††XFPE01	Full-person equivalent by subject
HESA derived field	††XJACSA01	JACS subject area
HESA derived field	††XLEV501	Level of study – 5 way split
HESA derived field	††XMODE01	Mode of study
HESA derived field	††XQLEV501	Level of qualification obtained – 5 way split
HESA derived field	††XQOBTN01	Highest qualification obtained

Table B1 notes:

† These fields are only used in the derivation of HIN information, and have been retained to allow easy identification of students. These values are retained in reference to each year of the HESA student records considered, that is, each of 2006-07 through to 2010-11 inclusive.

†† The specifications of the HESA derived fields are available alongside the data supply download from the HESA data submission facility.

Using the individualised data

6. Here we explain the individualised data that we consider in relation to the generation of the rates of qualification from full-time postgraduate research degree study. Our explanations focus on the generation of these rates with respect to the 2009-10 cohort of full-time postgraduate research degree starters. Exactly the same processes, definitions and approaches have been used with respect to the 2008-09 cohort of such starters.

7. The individualised data we consider contains one record for each instance of postgraduate study in a subject area in an academic year. For example, a student who is studying for a postgraduate research degree in biology in 2009-10 will have one record for that instance and will count as one headcount. A student who is studying for a postgraduate masters degree in mathematics and physics in 2009-10 will have two records for that instance; one for each subject area, with each counting as a headcount of a half.

8. We have also linked back via the HIN[32] to the HESA student records for the five years preceding the academic year for which the study outcomes are projected. As explained further at

[32] The HIN is a combination of three fields from the HESA individualised student records (the HESA unique student identifier, the UK provider reference number and the student instance identifier) which uniquely identify a student on a course (or instance of study). The HIN forms a year-on-year linking mechanism which can be used

paragraphs 18, 19 and 22 of Annex A, we seek to determine if the student was active on an instance of full-time postgraduate research study at any time during the five-year period. Additionally, we seek to identify students who have reached one of the end states prior to the academic year for which the study outcomes are projected. To this end, minimal information is retained from the HESA student records for each of these five years: namely, the HIN information, and the student's commencement date of study, mode of study, qualification aim and any qualification obtained.

9. There are two components of the individualised data: one that focuses on identifying progression patterns between the reference year (2009-10) and the year that follows, and another that focuses on identifying progression patterns between the year prior to the reference year (2008-09) and the reference year.

Identifying progression patterns between the reference year and the year that follows

10. The individualised data that enables identification of progression patterns between the reference year and the year that follows contains student records that are considered in the definition both of the starting population[33] in 2009-10, and of the part of the transition population that examines transitions between 2009-10 and 2010-11[34]. As such, data are drawn or derived from the HESA individualised student records for the years 2008-09, 2009-10 and 2010-11.

11. This data contains one record for each instance of postgraduate study (XLEV501 = 1, 2) in a subject area in 2009-10. Where the student identified by this record can be linked to an instance of study in 2010-11, appropriate data from the 2010-11 HESA individualised student records are attached to the 2009-10 record. If the student links to more than one instance of study at the same institution in 2010-11, we select the 2010-11 record with the best outcome in respect of postgraduate research degree.

12. The data also contains one record for each instance of postgraduate study in a subject area in 2008-09. Where the student identified by this record can be linked to an instance of study in 2009-10 and/or 2010-11 we take the same approach as described at paragraph 11.

Identifying progression patterns between the year prior to the reference year and the reference year

13. The individualised data that enables identification of progression patterns between the year prior to the reference year and the reference year contains student records that are considered as part of the transition population that examines transitions between 2008-09 and 2009-10. As such, data are drawn or derived from the HESA individualised student records for the years 2007-08, 2008-09 and 2009-10.

14. This data contains one record for each instance of postgraduate study (XLEV501 = 1, 2) in a subject area in 2008-09. Where the student identified by this record can be linked to an instance of study in 2009-10, we take the approach that is described at paragraph 11. The data

---

to track the student's progression on the instance from one year to the next. It follows that the HIN enables a student to be tracked from commencement of the instance through to completion.

[33] Paragraphs 11 to 14 at Annex A provide a description of the starting population.

[34] Paragraphs 15 to 43 at Annex A provide a description of the transition populations.

also contains one record for each instance of postgraduate study in a subject area in 2007-08. Where the student identified by this record can be linked to an instance of study in 2008-09 and/or 2009-10 we take the approach that is described at paragraph 11.

Description of derived fields

15. Table B2 provides details of the derived fields used in the generation of the rates of postgraduate research degree qualification. These fields are derived from the HESA fields described at Table B1, and details of these derivations are provided at the paragraphs identified in Table B2. We use XX to denote the academic year of the HESA student records from which the field is derived. For example, HUSID09 is a field derived from the 2009-10 HESA student records.

16. Our descriptions focus on the generation of these rates with respect to the 2009-10 cohort of full-time postgraduate research degree starters. Exactly the same definitions and approaches have been used with respect to the 2008-09 cohort of such starters.

Table B2 Derived fields used to generate rates of postgraduate research degree qualification

Derived field name	Description	Academic year for which the HESA variable is retained	Paragraph
†HUSIDXX	HUSID for the academic year	2006-07, 2007-08, 2008-09, 2009-10	18
†NUMHUSXX	NUMHUS for the academic year	2006-07, 2007-08, 2008-09, 2009-10	19
BIRTHDTEXX	BIRTHDTE for the academic year	2006-07, 2007-08, 2008-09	20
CAMPIDXX	CAMPID for the academic year	2006-07, 2007-08, 2008-09, 2009-10	21
COMDATEXX	COMDATE for the academic year	2006-07, 2007-08, 2008-09, 2009-10	22
††DOMXX	XDOMHM01 for the academic year	2006-07, 2007-08, 2008-09	23
EXCHANGEXX	EXCHANGE for the academic year	2007-08, 2008-09	24
MSTUFEEXX	MSTUFEE for the academic year	2006-07, 2007-08, 2008-09	32
††FPEXX	XFPE01 for the academic year	2006-07, 2007-08, 2008-09	25
INSTCAMPXX	INSTCAMP for the academic year	2006-07, 2007-08, 2008-09, 2009-10	26
††JACSXX	XJACSA01 for the academic year	2006-07, 2007-08, 2008-09	27
LEFTXX	ENDDATE for the academic year DATELEFT for the academic year	2007-08, 2008-09, 2009-10 2006-07	28
††LEVXX	XLEV501 for the academic year	2006-07, 2007-08, 2008-09, 2009-10	29
STULOADXX	STULOAD for the academic year	2006-07, 2007-08, 2008-09, 2009-10	38

LOCSDYXX	LOCSDY for the academic year	2006-07, 2007-08, 2008-09	30
††MODEXX	XMODE01 for the academic year	2006-07, 2007-08, 2008-09, 2009-10	31
QAIMXX	COURSEAIM for the academic year QUALAIM for the academic year	2007-08, 2008-09, 2009-10 2006-07	33
QENTXX	QUALENT2 for the academic year	2006-07, 2007-08, 2008-09	34
††QLEVXX	XQLEV501 for the academic year	2006-07, 2007-08, 2008-09, 2009-10	35
††QOBTXX	XQOBTN01 for the academic year	2006-07, 2007-08, 2008-09, 2009-10	36
SPCSTUXX	SPCSTU for the academic year	2006-07	37
TTCIDXX	TTCID for the academic year	2007-08, 2008-09	39
TYPEYRXX	TYPEYR for the academic year	2006-07, 2007-08, 2008-09, 2009-10	40
†UKPRNXX	UKPRN for the academic year	2007-08, 2008-09, 2009-10	41
YEARSTUXX	YEARSTU for the academic year	2006-07, 2007-08, 2008-09, 2009-10	42
TRANSITION	Identification of which of the four component parts of the transition population the record falls into		45
REFYEAR	Reference year for transition analysis, where the transition derivations consider progression across a pair of consecutive academic years		44
POPYEAR	Initial year for transition analysis, in which students are identified as being registered on a full-time research degree programme		43
DURATION	Calculation of the duration of study on the instance where ENDDATE / DATELEFT is completed in the academic year identified by POPYEAR		46

LINK	Flag to indicate whether a student recorded in the academic year identified by POPYEAR can be identified in HESA student records in the following academic year		55
LINK2	Flag to indicate whether a student recorded in the academic year identified by POPYEAR and falls within particular transition populations can be identified in HESA student records in the academic year two years later		56
NEWCAT3	Derived grouping of highest qualification on entry as recorded in the academic year identified by POPYEAR		47
PIDOM	Derived grouping of pre-course domicile as recorded in the academic year identified by POPYEAR		48
PIFUND	Derived grouping of source of tuition fees as recorded in the academic year identified by POPYEAR		49
PIMOD	Derived grouping of mode of study as recorded in the academic year identified by POPYEAR		50
PISUB	Derived grouping of principal subject of course aim as recorded in the academic year identified by POPYEAR		51
COMAGE	Derived grouping of the student's age at 30 September in the year of commencement		52
STARTER	Derived grouping of a student's status as a starter or not as recorded in the academic year identified by REFYEAR		65 to 66
PROGXX	Classification of a student's progression into an academic year from the preceding one		58 to 61

MTOPXX	Flag to indicate whether a student recorded in the academic year identified by REFYEAR shows evidence of an appropriate conversion from a full-time MPhil to a research degree programme		64
STATE0	Classification of a student's activity as recorded in the academic year identified by REFYEAR		62
STATE1	Classification of a student's activity as recorded in the academic year following the one identified by REFYEAR		63
SUBWT	Derivation of a student's full-person equivalence across subject area(s) of study as recorded in the academic year identified by POPYEAR		53
YEAR0	Derived grouping year of student on this instance as recorded in the academic year identified by REFYEAR		54
YEAR1	Derived grouping year of student on this instance as recorded in the academic year following the one identified by REFYEAR		57
YEARHIN	Academic year in which the student was last found to be active on a full-time postgraduate research qualification, not having gained a postgraduate research qualification		67 to 68
MULTIPLE	Flag to indicate whether an individual student was found to have reached the end state of having 'become absent from HE'		69 to 70
EXCL1XX	Flag to indicate whether a student recorded in the academic year identified by POPYEAR has been excluded from the starting and/or transition populations on account of exclusion criteria relating to their location of study		71
EXCL2XX	Flag to indicate whether a student recorded in the academic year identified by POPYEAR has been excluded from the starting and/or transition populations on account of exclusion criteria relating to their pre-course domicile		72

EXCL4XX	Flag to indicate whether a student recorded in the academic year identified by POPYEAR has been excluded from the starting and/or transition populations on account of exclusion criteria relating to incoming or visiting exchange students		73
EXCL8XX	Flag to indicate whether a student recorded in the academic year identified by POPYEAR has been excluded from the starting and/or transition populations on account of exclusion criteria relating to their mode of study		74
EXCL16XX	Flag to indicate whether a student recorded in the academic year identified by POPYEAR has been excluded from the starting and/or transition populations on account of exclusion criteria relating to the general qualification aim of their course		75
EXCL32XX	Flag to indicate whether a student recorded in the academic year identified by POPYEAR has been excluded from the starting and/or transition populations on account of exclusion criteria relating to the duration of their study on the instance		76
EXCL64XX	Flag to indicate whether a student recorded in the academic year identified by POPYEAR has been excluded from the starting and/or transition populations on account of exclusion criteria relating to the student instance FTE recorded		77
EXCL128XX	Flag to indicate whether a student recorded in the academic year identified by POPYEAR has been excluded from the transition population defined in relation to the academic year identified by REFYEAR, on account of progression patterns identified		78 to 80

EXCLUDEXX	Flag to indicate whether a student recorded in the academic year identified by POPYEAR has been excluded from the starting and/or transition populations on account of meeting any of the exclusion criteria defined		81

Table B2 notes:

† These fields are only used in the derivation of HIN information, and have been retained to allow easy identification of students. These values are retained in reference to each year of the HESA student records considered, that is, each of 2006-07 through to 2010-11 inclusive.

†† The specifications of the HESA derived fields on which these fields are based are available alongside the data supply download from the HESA data submission facility.

17. The derived fields shown in Table B2 are each discussed in greater detail at the paragraphs referenced by the table. Where this further detail includes the provision of specific algorithms used, these algorithms are presented hierarchically. That is, if a student record satisfies more than one of the criteria described, they will be classified on the basis of meeting the first of those criteria to appear in the description of the algorithm.

HUSIDXX

18. This field contains the HUSID field value in the HESA student record for the year denoted by XX. For example, HUSID09 is the HUSID field value in the 2009-10 HESA student record.

NUMHUSXX

19. This field contains the NUMHUS field value in the HESA student record for the year denoted by XX. For example, NUHMUS09 is the NUMHUS field value in the 2009-10 HESA student record.

BIRTHDTEXX

20. This field contains the BIRTHDTE field value in the HESA student record for the year denoted by XX. For example, BIRTHDTE09 is the BIRTHDTE field value in the 2009-10 HESA student record.

CAMPIDXX

21. This field contains the CAMPID field value in the HESA student record for the year denoted by XX. For example, CAMPID09 is the CAMPID field value in the 2009-10 HESA student record.

COMDATEXX

22. This field contains the COMDATE field value in the HESA student record for the year denoted by XX. For example, COMDATE09 is the COMDATE field value in the 2009-10 HESA student record.

DOMXX

23. This field contains the XDOMHM01 field value in the HESA student record for the year denoted by XX. For example, DOM09 is the XDOMHM01 field value in the 2009-10 HESA student record.

EXCHANGEXX

24. This field contains the EXCHANGE field value in the HESA student record for the year denoted by XX. For example, EXCHANGE09 is the EXCHANGE field value in the 2009-10 HESA student record.

FPEXX

25. This field contains the XFPE01 field value in the HESA student record for the year denoted by XX. For example, FPE09 is the XFPE01 field value in the 2009-10 HESA student record. The XFPE01 field apportions the subject areas of study to a full-person equivalence, according to the number of subjects submitted per instance and the weighting allocated to an instance's course subject. It is expressed as a percentage.

INSTCAMPXX

26. This field contains the INSTCAMP field value in the HESA student record for the year denoted by XX. For example, INSTCAMP09 is the INSTCAMP field value in the 2009-10 HESA student record.

JACSXX

27. This field contains the XJACS01 field value in the HESA student record for the year denoted by XX. For example, JACS09 is the XJACS01 field value in the 2009-10 HESA student record.

LEFTXX

28. This field contains the ENDDATE field value in the HESA student record for the year denoted by XX, where XX is one of 07, 08, 09 or 10 and DATELEFT if XX is 06. For example, LEFT09 is the ENDDATE field value in the 2009-10 HESA student record.

LEVXX

29. This field contains the XLEV501 field value in the HESA student record for the year denoted by XX. For example, LEV09 is the XLEV501 field value in the 2009-10 HESA student record.

LOCSDYXX

30. This field contains the LOCSDY field value in the HESA student record for the year denoted by XX. For example, LOCSDY09 is the LOCSDY field value in the 2009-10 HESA student record.

MODEXX

31. This field contains the XMODE01 field value in the HESA student record for the year denoted by XX. For example, MODE09 is the XMODE01 field value in the 2009-10 HESA student record.

MSTUFEEXX

32. This field contains the MSTUFEE field value in the HESA student record for the year denoted by XX. For example, MSTUFEE09 is the MSTUFEE field value in the 2009-10 HESA student record.

QAIMXX

33. This field contains the COURSEAIM field value in the HESA student record for the year denoted by XX, where XX is one of 07, 08 and 09 and QUALAIM if XX is 06. For example, QAIM09 is the COURSEAIM field value in the 2009-10 HESA student record.

QENTXX

34. This field contains the QUALENT2 field value in the HESA student record for the year denoted by XX. For example, QENT09 is the QUALENT2 field value in the 2009-10 HESA student record.

QLEVXX

35. This field contains the XQLEV501 field value in the HESA student record for the year denoted by XX. For example, QLEV09 is the XQLEV501 field value in the 2009-10 HESA student record.

QOBTXX

36. This field contains the XQOBTN01 field value in the HESA student record for the year denoted by XX. For example, QOBT09 is the XQOBTN01 field value in the 2009-10 HESA student record.

SPCSTU06

37. This field contains the SPCSTU field value in the HESA student record for 2006-07.

STULOADXX

38. This field contains the STULOAD field value in the HESA student record for the year denoted by XX. For example, STULOAD09 is the STULOAD field value in the 2009-10 HESA student record.

TTCIDXX

39. This field contains the TTCID field value in the HESA student record for the year denoted by XX. For example, TTCID09 is the TTCID field value in the 2009-10 HESA student record.

TYPEYRXX

40. This field contains the TYPEYR field value in the HESA student record for the year denoted by XX. For example, TYPEYR09 is the TYPEYR field value in the 2009-10 HESA student record.

UKPRNXX

41. This field contains the UKPRN field value in the HESA student record for the year denoted by XX. For example, UKPRN09 is the UKPRN field value in the 2009-10 HESA student record.

YEARSTUXX

42. This field contains the YEARSTU field value in the HESA student record for the year denoted by XX. For example, YEARSTU09 is the YEARSTU field value in the 2009-10 HESA student record.

POPYEAR

43. Paragraphs 15 to 25 of Annex A provide the definition of the transition population, and describe its four component parts. Analysis of transition considers students' change in state from one academic year to the next. This field identifies the initial year for the transition analysis.

Value	Description
2009-10	Record falls within 2009-10 HESA student records and consideration of students' change in states between 2009-10 and 2010-11 is required
2008-09	Record falls within 2008-09 HESA student records and consideration of students' change in states between 2008-09 and 2009-10 is required
2007-08	Record falls within 2007-08 HESA student records and consideration of students' change in states between 2007-08 and 2008-09 is required

REFYEAR

44. Paragraphs 15 to 25 of Annex A provide the definition of the transition population, and describe its four component parts. For the purposes of this analysis we require information about changes in state from one of two reference years to the following year. This field identifies the reference year for the transition analysis.

Value	Description
2009-10	Record contributes to analysis of transitions between 2009-10 and 2010-11
2008-09	Record contributes to analysis of transitions between 2008-09 and 2009-10

TRANSITION

45. Paragraphs 15 to 25 of Annex A provide the definition of the transition population, and describe its four component parts. This field identifies which of the four parts of the transition population a given record contributes to.

Value	Description	Definition
(i)	Record falls within component (i) of the definition of the transition population:	POPYEAR = 2009-10 and REFYEAR = 2009-10
(ii)	Record falls within component (ii) of the definition of the transition population:	POPYEAR = 2008-09 and REFYEAR = 2009-10
(iii)	Record falls within component (iii) of the definition of the transition population	POPYEAR = 2008-09 and REFYEAR = 2008-09
(iv)	Record falls within component (iv) of the definition of the transition population	POPYEAR = 2007-08 and REFYEAR = 2008-09

DURATION

46. This field considers whether the ENDDATE/DATELEFT field is completed in the academic year identified by POPYEAR. Where an end date of the instance of study is provided, the duration of that instance is calculated as a number of days.

Value	Description	Definition (where XX relates to the academic year identified by POPYEAR)
0 – 99,999	Where the ENDDATE/ DATELEFT is completed, duration of study is calculated as the difference between the COMDATE and the ENDDATE/DATELEFT in the academic year identified by POPYEAR	DURATION = LEFTXX – COMDATEXX where LEFTXX ≠ NULL
(Blank)	ENDDATE/DATELEFT is not	LEFTXX = (Blank)

	completed in the academic year identified by POPYEAR

NEWCAT3

47. This field identifies the grouping of the student's highest qualification on entry that has been derived from data recorded in the academic year identified by POPYEAR.

Value	Description	Definition (where XX relates to the academic year identified by POPYEAR)
PG	Higher education qualifications held on entry: Postgraduate qualifications	QENTXX = 01, 02, 03, 04, 05
FD	Higher education qualifications held on entry: First degree of UK institution and graduate equivalent qualifications	QENTXX = 10, 11, 12, 13, 14, 15, 16
OT	Entry qualifications at neither postgraduate nor first degree level	Otherwise

PIDOM

48. This field identifies the grouping of the student's pre-course domicile that has been derived from data recorded in the academic year identified by POPYEAR.

Value	Description	Definition (where XX relates to the academic year identified by POPYEAR)
H	UK-domiciled	DOMXX = 1, 2, 3, 4
E	EU, Guernsey, Jersey or Isle of Man domiciled	DOMXX = 5, 6
O	Otherwise	Otherwise

PIFUND

49. This field identifies the grouping of the student's source of funding that has been derived from data recorded in the academic year identified by POPYEAR.

Value	Description	Definition (where XX relates to the academic year identified by POPYEAR)
RC	Students whose major source of tuition	MSTUFEEXX = 11, 12, 13, 14,

	fees was recorded as being the UK Research Councils	15, 16, 17, 18, 19
OT	Otherwise	Otherwise

PIMOD

50. This field identifies the grouping of the student's mode of study that has been derived from data recorded in the academic year identified by POPYEAR.

Value	Description	Definition (where XX relates to the academic year identified by POPYEAR)
1	Full-time	MODEXX = 1
2	Sandwich	MODEXX = 2
3	Part-time	MODEXX = 3
4	Writing-up	MODEXX = 4
5	Sabbatical	MODEXX = 5
6	Dormant	MODEXX = 6

PISUB

51. This field identifies the grouping of the student's principle subject of course aim that has been derived from data recorded in the academic year identified by POPYEAR.

Value	Description	Definition (where XX relates to the academic year identified by POPYEAR)
SCI	Science-based subject area of study	JACSXX = 1, 2, 3, 4, 5, 6, 7, 9
OTH	Not a science-based subject area of study	Otherwise

COMAGE

52. This field identifies the grouping of the student's age as at 30 September in the year of commencement of their programme of study.

Value	Description	Definition (where XX relates to the academic year identified by POPYEAR)
U	Age at 30 September in the year of commencement is unknown	Age on 30 September in academic year of COMDATEXX is unknown on account of missing or invalid

		BIRTHDTEXX
M	Mature student aged 25 or over	Age on 30 September in academic year of COMDATEXX is 25 or over
Y	Young student aged under 25	Age on 30 September in academic year of COMDATEXX is less than 25

SUBWT

53. This field contains a derivation of a student's full-person equivalence across the principle subject(s) of course aim in the academic year following the one identified by POPYEAR. The full-person equivalence is given as a headcount (as opposed to a percentage as given by FPEXX).

Value	Description	Definition (where XX relates to the academic year identified by POPYEAR)
0 – 1	Full-person equivalence across the principle subject(s) of a student's course aim, expressed as a headcount	FPEXX / 100

YEAR0

54. This field identifies the grouping of the student's year of student on this instance that has been derived from data recorded in the academic year identified by REFYEAR.

Value	Description	Definition (where XX relates to the academic year identified by REFYEAR)
7	Year of student on this instance was recorded as 7 or higher	YEARSTUXX >= 7
2, 3, 4, 5, 6	Year of student on this instance was recorded as one in the range from 2 to 6 inclusive	YEARSTUXX = 2, 3, 4, 5, 6 respectively
1	Otherwise	Otherwise

LINK

55. This field contains a flag that indicates whether or not a student who was recorded in the academic year identified by POPYEAR has been identified in the HESA student records in the following academic year.

Value	Description
0	No record is found for the student in the academic year following POPYEAR
1	The student has an entry in the HESA student records in the academic year following POPYEAR

LINK2

56. This field contains a flag that indicates whether or not a student who was recorded in the academic year identified by POPYEAR and falls within the TRANSITION = (ii), (iv) population has been identified in the HESA student records in the academic year two years later.

Value	Description
0	No record is found for the student in the academic year two years after POPYEAR or TRANSITION ≠ (ii), (iv)
1	The student has TRANSITION = (ii), (iv) and an entry in the HESA student records in the academic year two years after POPYEAR

YEAR1

57. This field identifies the grouping of the student's year of student on this instance that has been derived from data recorded in the academic year following the one identified by REFYEAR.

Value	Description	Definition (where XX relates to the academic year following the one identified by REFYEAR)
7	Year of student on this instance was recorded as 7 or higher	YEARSTUXX >= 7
2, 3, 4, 5, 6	Year of student on this instance was recorded as one in the range from 2 to 6 inclusive	YEARSTUXX = 2, 3, 4, 5, 6 respectively
1	Otherwise	Otherwise

PROGXX

58. This field contains a classification of a student's progression into an academic year identified by XX from the preceding academic year.

59. For TRANSITION = (i), (iii) we derive PROGXX where XX relates to the academic year following the one identified by REFYEAR:

Value	Description	Definition
Qualify RDeg	Qualified with research degree award	QOBT(XX-1) = 02, D00, D01 **and** QLEV(XX-1) = 1
Qualify MPhil	Qualified with MPhil award	QOBT(XX-1) = 04, 06, L00, L80 **and** QLEV(XX-1) = 1
RDeg	Full-time research degree study	LINK = 1 **and** QOBT(XX-1) ≠ 02, 04, 06, D00, D01, L00, L80 **and** MODEXX = 1, 2 **and** LEVXX ≠ 3, 4, 5 **and** STULOADXX >= 0 **and** INSTXX = INST(XX-1) **and** QAIMXX = 02, D00, D01
PT RDeg	Part-time research degree study	LINK = 1 **and** QOBT(XX-1) ≠ 02, 04, 06, D00, D01, L00, L80 **and** MODEXX = 3 **and** LEVXX ≠ 3, 4, 5 **and** STULOADXX >= 0 **and** INSTXX = INST(XX-1) **and** QAIMXX = 02, D00, D01
Writing up	Writing up research degree study	**(**LINK = 1 **and** QOBT(XX-1) ≠ 02, 04, 06, D00, D01, L00, L80 **and** MODEXX = 4, 5 **and (**LEVXX = 3, 4, 5 **or** STULOADXX <= 0**) and** INSTXX = INST(XX-1) **and** QAIMXX = 02, 04, 06, D00, D01, L00, L80, L99**) or** **(**LINK = 1 **and** QOBT(XX-1) ≠ 02, 04, 06, D00, D01, L00, L80 **and** MODEXX = 4, 5 **and** LEVXX ≠ 3, 4, 5 **and** STULOADXX >= 0 **and** INSTXX = INST(XX-1) **and** QAIMXX = 02, 04, 06, D00, D01, L00, L80, L99**) or** **(**LINK ≠ 1 **and** QOBT(XX-1) ≠ 02, 04, 06, D00, D01, L00, L80 **and** MODEXX = 4, 5 **and** INSTXX = INST(XX-1) **and** QAIMXX = 02, 04, 06, D00, D01, L00, L80, L99**)**
MPhil	Full- or part-time MPhil study	LINK = 1 **and** QOBT(XX-1) ≠ 02, 04, 06, D00, D01, L00, L80 **and** MODEXX = 1, 2, 3 **and** LEVXX ≠ 3, 4, 5 **and** STULOADXX >= 0 **and** INSTXX = INST(XX-1) **and** QAIMXX = 04, 06, L00, L80, L99
Transfer	Transferred to another UK HEI	LINK = 1 **and** QOBT(XX-1) ≠ 02, 04, 06, D00, D01, L00, L80 **and (**MODEXX = 1, 2, 3, 4, 5 **and** LEVXX ≠ 3, 4, 5 **and** STULOADXX >= 0 **and** INSTXX ≠ INST(XX-1)
Inactive	Not currently active in postgraduate research HE	Otherwise

60. For TRANSITION = (ii), (iv) we derive PROGXX where XX relates to the academic year identified by REFYEAR. This derivation is consistent with that described at paragraph 59.

61. In addition, for TRANSITION = (ii), (iv) we also derive PROGXX where XX relates to the academic year following the one identified by REFYEAR:

Value	Description	Definition
Qualify RDeg	Qualified with research degree award	QOBT(XX-1) = 02, D00, D01 **and** QLEV(XX-1) = 1 **and** INST(XX-1) = INST(XX-2)
Qualify MPhil	Qualified with MPhil award	QOBT(XX-1) = 04, 06, L00, L80 **and** QLEV(XX-1) = 1 **and** INST(XX-1) = INST(XX-2)
RDeg	Full-time research degree study	LINK2 = 1 **and** QOBT(XX-1) ≠ 02, 04, 06, D00, D01, L00, L80 **and** MODEXX = 1, 2 **and** LEVXX ≠ 3, 4, 5 **and** STULOADXX >= 0 **and** INSTXX = INST(XX-2) **and** QAIMXX = 02, D00, D01
PT RDeg	Part-time research degree study	LINK2 = 1 **and** QOBT(XX-1) ≠ 02, 04, 06, D00, D01, L00, L80 **and** MODEXX = 3 **and** LEVXX ≠ 3, 4, 5 **and** STULOADXX >= 0 **and** INSTXX = INST(XX-2) **and** QAIMXX = 02, D00, D01
Writing up	Writing up research degree study	**(**LINK2 = 1 **and** QOBT(XX-1) ≠ 02, 04, 06, D00, D01, L00, L80 **and** MODEXX = 4, 5 **and (**LEVXX ≠ 1 **or** STULOADXX <= 0**) and** INSTXX = INST(XX-2) **and** QAIMXX = 02, 04, 06, D00, D01, L00, L80, L99**) or** **(**LINK2 = 1 **and** QOBT(XX-1) ≠ 02, 04, 06, D00, D01, L00, L80 **and** MODEXX = 4, 5 **and** LEVXX = 1 **and** STULOADXX >= 0 **and** INSTXX = INST(XX-2) **and** QAIMXX = 02, 04, 06, D00, D01, L00, L80, L99**) or** **(**LINK2 ≠ 1 **and** QOBT(XX-1) ≠ 02, 04, 06, D00, D01, L00, L80 **and** MODEXX = 4, 5 **and** QAIMXX = 02, 04, 06, D00, D01, L00, L80, L99**)**
MPhil	Full- or part-time MPhil study	LINK2 = 1 **and** QOBT(XX-1) ≠ 02, 04, 06, D00, D01, L00, L80 **and** MODEXX = 1, 2, 3 **and** LEVXX = 1 **and** STULOADXX >= 0 **and** INSTXX = INST(XX-2) **and** QAIMXX = 04, 06, L00, L80, L99
Transfer	Transferred to another UK HEI	LINK2 = 1 **and** QOBT(XX-1) ≠ 02, 04, 06, D00, D01, L00, L80 **and** MODEXX = 1, 2, 3, 4, 5 **and** LEVXX = 1 **and** STULOADXX >= 0 **and** INSTXX ≠ INST(XX-1)
Absent	Absent from postgraduate research HE	Otherwise

STATE0

62. This field contains a classification of a student's activity in the academic year identified by REFYEAR. Details of the values of state 0 are described at paragraphs 28 to 31 of Annex A.

Value	Description	Definition (where XX relates to the academic year identified by REFYEAR)
RDeg 1, 2, 3, 4, 5, 6, 7+	Full-time research degree, year of study 1, 2, 3, 4, 5, 6 or 7 or above respectively	**(**PIMOD = 1, 2 **and** QAIMXX = D00, D01 **and** YEAR0 = 1, 2, 3, 4, 5, 6, 7 **and** TRANSITION = (i), (iii)**) or (**PROGXX = 'RDeg' **and** TRANSITION = (ii), (iv)**)**
MPhil	Full- or part-time MPhil	**(**PIMOD = 1, 2, 3 **and** QAIMXX = L00, L80, L99 **and** TRANSITION = (i), (iii)**) or (**PROGXX = 'MPhil' **and** TRANSITION = (ii), (iv)**)**
PT RDeg	Part-time research degree	**(**PIMOD = 3 **and** QAIMXX = D00, D01 **and** TRANSITION = (i), (iii)**) or (**PROGXX = 'PT RDeg' **and** TRANSITION = (ii), (iv)**)**
WU 0	Writing up research degree, year of study 0, 1 or 2	**(**PIMOD = 4, 5 **and** YEAR0 = 1, 2 **and** TRANSITION = (i), (iii)**) or (**PROGXX = 'Writing up' **and** YEAR0 = 1, 2 **and** TRANSITION = (ii), (iv)**)**
WU 3+	Writing up research degree, year of study 3 or above	**(**PIMOD = 4, 5 **and** YEAR0 = 3, 4, 5, 6, 7 **and** TRANSITION = (i), (iii)**) or (**PROGXX = 'Writing up' **and** YEAR0 = 3, 4, 5, 6, 7 **and** TRANSITION = (ii), (iv)**)**
Transfer	Transferred to another UK HEI	**(**PROGXX = 'Transfer' **and** TRANSITION = (ii), (iv)**)**
Inactive	Not currently active in postgraduate research HE	**((**PIMOD ≠ 1, 2, 3, 4, 5 **or** QAIMXX ≠ D00, D01, L00, L80, L99**) and** TRANSITION = (i), (iii)**) or (**PROGXX = 'Qualify RDeg', 'Qualify MPhil', 'Inactive' **and** TRANSITION = (ii), (iv)**)**

STATE1

63. This field contains a classification of a student's activity in the academic year following the one identified by REFYEAR. Details of the values of state 1 are described at paragraphs 32 to 36 of Annex A.

Value	Description	Definition (where XX relates to the academic year following the one identified by REFYEAR)
RDeg 1, 2, 3, 4, 5, 6, 7+	Full-time research degree, year of study 1, 2, 3, 4, 5, 6 or 7 or above respectively	**(**MODEXX = 1, 2 **and** QAIMXX = D00, D01 **and** YEAR0 = 1, 2, 3, 4, 5, 6, 7**) or** PROGXX = 'RDeg'

MPhil	Full- or part-time MPhil	PROGXX = 'MPhil'
PT RDeg	Part-time research degree	**(**MODEXX = 3 **and** QAIMXX = D00, D01**) or** PROGXX = 'PT RDeg'
WU 0	Writing up research degree, year of study 0, 1 or 2	**(**MODEXX = 4, 5 **and** YEAR1 = 1, 2**) or (**PROGXX = 'Writing up' **and** YEAR1 = 1, 2**)**
WU 3+	Writing up research degree, year of study 3 or above	**(**MODEXX = 4, 5 **and** YEAR1 = 3, 4, 5, 6, 7**) or (**PROGXX = 'Writing up' **and** YEAR1 = 3, 4, 5, 6, 7**)**
Transfer	Transferred to another UK HEI	PROGXX = 'Transfer'
Inactive	Not currently active in postgraduate research HE	**(**PROGXX = 'Inactive' **and** PROG(XX-1) ≠ 'Inactive'**) or (**PROGXX ≠ 'Qualify RDeg', 'Qualify MPhil', 'Transfer' **and** QAIMXX = D00, D01 **and** MODEXX ≠ 1, 2, 3, 4, 5**) or (**PROGXX = 'Absent' **and** PROG(XX-1) = 'Writing up', 'Qualify MPhil', 'PT RDeg', 'MPhil'**)**
Qualify RDeg	Qualified with research degree award	PROGXX = 'Qualify RDeg'
Qualify MPhil	Qualified with MPhil award	**(**PROGXX = 'Qualify MPhil'**) or (**PROGXX = 'Absent' **and** PROG(XX-1) = 'Qualify MPhil'**)**
Absent	Absent from postgraduate research HE	**(**PROGXX = 'Absent' **and** TRANSITION = (ii), (iv)**) or (**PROGXX **=** 'Inactive' **and** PROG(XX-1) **=** 'Inactive'**)**

MTOPXX

64. The rates of qualification from postgraduate research degree study are calculated with respect to a population of research degree starters in academic year 2009-10. The intention is to include within the starting population those students who convert a full-time MPhil qualification to a postgraduate research degree at the earliest time that doctoral study can be identified[35]. This field contains a flag that indicates whether or not a student who was recorded in the 2009-10 academic year has shown evidence of such a conversion. That is, whether a student was registered as a full-time MPhil student in the 2009-10 academic year, and showed evidence of studying for a doctoral qualification in 2010-11 while having had no MPhil qualification awarded.

[35] This intention is explained further at paragraphs 11 to 14 at Annex A.

Value	Description	Definition (where XX relates to the academic year identified by REFYEAR)
1	Student is found to show evidence of an appropriate conversion from an MPhil qualification to a postgraduate research degree	REFYEAR = 2009-10 **and** STATE0 = MPhil **and** PROG(XX+1) ≠ 'Qualify RDeg', 'Qualify MPhil', 'Transfer' **and** QAIM(XX+1) = D00, D01 **and** MODEXX = 1, 2
0	Student is not found to show evidence of an appropriate conversion from an MPhil qualification to a postgraduate research degree	REFYEAR = 2009-10 **and (**STATE0 ≠ MPhil **or** PROG(XX+1) = 'Qualify RDeg', 'Qualify MPhil', 'Transfer' **or** QAIM(XX+1) ≠ D00, D01 **or** MODEXX ≠ 1, 2**)**
NULL	Student is not considered in respect of conversions from MPhil to postgraduate research degree because the reference year for the transition analysis is not 2009-10	REFYEAR = 2008-09

STARTER

65. The rates of qualification from postgraduate research degree study are calculated with respect to a population of research degree starters in academic year 2009-10. The starting population includes entrants to a full-time research degree programme of study (those whose commencement date for their programme of study fell within the 2009-10 academic year). It further includes those who converted a full-time MPhil qualification to a research degree in the manner described at paragraph 64.

66. This field contains a flag that indicates whether or not the student was grouped within the starting population in the 2009-10 academic year, where the definition of a starter is that described at paragraph 65.

Value	Description	Definition (where XX relates to the academic year identified by REFYEAR)
1	Student is grouped within the 2009-10 starting population	REFYEAR = 2009-10 **and (**MtoPXX = 1 **or (**COMXX >= 1 August 20XX **and** COMXX <= 31 July 20XX **and** QAIMXX = D00, D01 **and** PIMOD = 1, 2**))**
(Blank)	Student is not grouped within the 2009-10 starting population because the reference year for the transition analysis is not 2009-10	REFYEAR ≠ 2009-10

0	Student is not grouped within the 2009-10 starting population	Otherwise

YEARHIN

67. Considering the five-year period prior to the 2009-10 academic year for which the study outcomes are projected, the HIN is used to identify the academic year in which an individual student was most recently found to be active on an instance of full-time postgraduate research study, not having reached one of the end states prior to the academic year for which the outcomes are projected. Minimal information has been retained from the HESA student records for each of these five academic years: namely, the HIN information, and the student's commencement date of study, mode of study, qualification aim and any qualification obtained.

68. This field identifies the academic year in which an individual student was most recently found to be active in full-time postgraduate research study, not having reached the end states of having 'qualified' with a postgraduate research qualification.

Value	Description	Definition
		(where YY relates to the most recent of the five academic years prior to the reference year – 2009-10 – in which the criteria are satisfied, and XX relates to the reference year and ZZ relates to YY and all subsequent years up to 2008-09)
20YY	Academic year in which the student was last found to be active on a full-time postgraduate research qualification, not having gained a postgraduate research qualification	HINYY = HINXX **and** QAIMYY = 02, 04, 06, D00, D01, L00, L80, L99 **and** MODEYY = 1, 2 **and** any QOBTZZ ≠ 02, 04, 06, D00, D01, L00, L80, L99
(blank)	Otherwise	Otherwise

MULTIPLE

69. Considering the five-year period prior to the 2009-10 academic year for which the study outcomes are projected, the HIN is used to identify the academic year in which an individual student was most recently found to be active on an instance of full-time postgraduate research study, not having reached one of the end states prior to the academic year for which the outcomes are projected. Minimal information has been retained from the HESA student records for each of these five academic years: namely, the HIN information, and the student's commencement date of study, mode of study, qualification aim and any qualification obtained.

70. This field identifies whether an individual student was found to have reached the end state of having 'become absent from HE': they have been identified as being inactive on the programme of study for two or more consecutive academic years.

Value	Description	Definition
		(Where XX relates to the reference year, and YY relates to the academic year identified by YEARHIN. Let ZZ relate to the academic year identified by COMDATEYY, or the academic year five years prior to the reference year, whichever is later. Let WW relate to ZZ and all subsequent years up to 2008-09.)
1	Found to have reached the 'become absent from HE' end state prior to 2009-10	MODEWW ≠ 1, 2, 3, 4, 5 and MODE(WW+1) ≠ 1, 2, 3, 4, 5
0	Otherwise	Otherwise

EXCL1XX

71. This field contains a flag that indicates whether or not a student recorded in the academic year identified by POPYEAR has been excluded from the starting and/or transition populations on account of exclusion criteria relating to their location of study. Non-UK based students studying by distance learning are excluded from the populations considered.

Value	Description	Definition
		(where XX relates to the academic year identified by POPYEAR)
1	Non-UK based studying by distance learning	LOCSDYXX = 7, S
0	Otherwise	Otherwise

EXCL2XX

72. This field contains a flag that indicates whether or not a student recorded in the academic year identified by POPYEAR has been excluded from the starting and/or transition populations on account of exclusion criteria relating to their pre-course domicile. Students recorded as having been domiciled outside of the EU member countries (and outside Guernsey, Jersey or the Isle of Man) are excluded from the populations considered.

Value	Description	Definition
1	Domiciled outside of the EU, Guernsey, Jersey or the Isle of Man	PIDOM = O
0	Otherwise	Otherwise

EXCL4XX

73. This field contains a flag that indicates whether or not a student recorded in the academic year identified by POPYEAR has been excluded from the starting and/or transition populations as they are an incoming or visiting exchange student.

Value	Description	Definition (where XX relates to the academic year identified by POPYEAR)
1	Incoming or visiting exchange student	**(**SPCSTUXX = 3, 4, 5, 6, 8 **and** POPYEAR <= 2006-07**) or ((**EXCHANGEXX ≠ 0, 5, 7, B, C, D, E, Y, NULL **or** TTCIDXX = E, F**) and** POPYEAR > 2006-07**)**
0	Otherwise	Otherwise

EXCL8XX

74. This field contains a flag that indicates whether or not a student recorded in the academic year identified by POPYEAR has been excluded from the starting and/or transition populations on account of exclusion criteria relating to their mode of study. Students recorded as not having been studying full-time either in the academic year identified by POPYEAR or in any of the five years prior to 2009-10 are excluded from the populations considered. Students recorded as having been studying full-time in the academic year identified by POPYEAR but found to have reached one of the end states prior to 2009-10 are also excluded.

Value	Description	Definition
1	Not studying on a full-time mode of study in the academic year identified by POPYEAR, and no evidence from HIN information that the student studied full-time on a postgraduate research qualification in any of the five academic years prior to 2009-10. Or, not studying on a full-time mode of study in the academic year identified by POPYEAR when TRANSITION is (i) or (iii), which by definition require a full-time mode of study in the academic year identified by REFYEAR. Or, studying on a full-time mode of study in the academic year identified by POPYEAR but evidence from HIN information that they reached one of the end states prior to 2009-10.	**(**PIMOD = 3, 4, 5, 6 **and** YEARHIN = POPYEAR, (blank)**) or (**PIMODE = 3, 4, 5, 6 **and** POPYEAR = REFYEAR **and** TRANSITION = (i), (iii)**) or (**PIMOD = 1, 2 **and** MULTIPLE = 1**)**
0	Otherwise	Otherwise

EXCL16XX

75. This field contains a flag that indicates whether or not a student recorded in the academic year identified by POPYEAR has been excluded from the starting and/or transition populations on account of exclusion criteria relating to the general qualification aim of their course. Students recorded as not having been studying for an appropriate postgraduate research qualification are excluded from the populations considered.

Value	Description	Definition (where XX relates to the academic year identified by POPYEAR)
1	Not studying on an appropriate postgraduate research qualification aim. Or, not studying on an appropriate postgraduate research degree qualification aim when TRANSITION is (i) or (iii), which by definition require a postgraduate research degree qualification aim in the academic year identified by REFYEAR.	**(**QAIMXX ≠ 02, 04, 06 **and** POPYEAR <= 2006-07**) or (**QAIM XX ≠ D00, D01, L00, L80, L99 **and** POPYEAR > 2006-07**) or (**QAIM XX ≠ D00, D01, 02 **and** POPYEAR = REFYEAR **and** TRANSITION = (I), (III)**)**
0	Otherwise	Otherwise

EXCL32XX

76. This field contains a flag that indicates whether or not a student recorded in the academic year identified by POPYEAR has been excluded from the starting and/or transition populations on account of exclusion criteria relating to the duration of their study on the instance. Students recorded as having left the programme of study within 50 days are excluded from the populations considered.

Value	Description	Definition (where XX relates to the academic year identified by POPYEAR)
1	Student left the programme of study within 50 days	LEFTXX ≠ (Blank) **and** DURATION < 50
0	Otherwise	Otherwise

EXCL64XX

77. This field contains a flag that indicates whether or not a student recorded in the academic year identified by POPYEAR has been excluded from the starting and/or transition populations on account of exclusion criteria relating to the student instance FTE recorded. This exclusion criterion is applicable for 2006-07 student records only. Students recorded as not having been studying for an appropriate FTE in a standard academic year are excluded from the populations considered.

Value	Description	Definition
1	Student studied 0 FTE in a standard academic year	LOAD06 = 0 **and** TYPEYR06 = 1
0	Otherwise	Otherwise

EXCL128XX

78. This field contains a flag that indicates whether or not a student recorded in the academic year identified by POPYEAR has been excluded from the transition population defined in relation to the academic year identified by REFYEAR on account of progression patterns identified.

79. Students found to continue, qualify or transfer would be duplicated in the transition analysis if they are considered both where the transition analysis calls upon data from the academic year prior to the one identified by REFYEAR (that is, where TRANSITION = (ii) or (iv)) and where the transition analysis focuses on movements across the academic year identified by REFYEAR and the one that follows (that is, where TRANSITION = (i) or (iii)). Such students are excluded from the transition population considered in relation to the academic year identified by REFYEAR.

80. Students found to have been inactive on a postgraduate research qualification in the academic year prior to the one identified by REFYEAR (that is, where TRANSITION = (ii) or (iv)) who progressed to be inactive on a postgraduate research qualification in the following year are also excluded from the transition population considered in relation to the academic year identified by REFYEAR.

Value	Description	Definition (where XX relates to the academic year identified by REFYEAR)
1	Student found to continue, qualify or transfer in the transition analysis considering the academic year prior to the one identified by REFYEAR	**(**PROGXX = Qualify PhD, Qualify MPhil, PhD, Transfer**) or (**MODE(XX-1) ≠ 1, 2, 3, 4, 5 **and** PROGXX **=** 'Inactive**)** where TRANSITION = (ii), (iv)
0	Otherwise	Otherwise

EXCLUDEXX

81. This field considers whether information provided by the student record meets any of the exclusion criteria detailed above. As such it indicates whether the record will be included in the appropriate starting and/or transition populations, and hence will be included in the generation of rates of qualification from postgraduate research degree study. For students excluded from the starting and/or transition populations, the field contains the sum of all applicable values from the table below. Students included in the generation of rates of qualification from postgraduate research degree study have EXCLUDEXX = 0.

Value	Description	Definition
1	Non-UK based student studying by distance learning	EXCL1XX = 1
2	Domiciled outside of the EU, Guernsey, Jersey or the Isle of Man	EXCL2XX = 1
4	Incoming or visiting exchange student	EXCL4XX = 1
8	Not studying on a full-time mode of study	EXCL8XX = 1
16	Not studying on an appropriate postgraduate research qualification aim	EXCL16XX = 1
32	Student left the programme of study within 50 days	EXCL32XX = 1
64	Student studied 0 FTE in a standard academic year	EXCL64XX = 1
128	Student found to continue, qualify or transfer in the transition analysis considering the academic year prior to the one identified by REFYEAR	EXCL128XX = 1
0	None of the exclusion criteria are satisfied	EXCL1XX = 0 **and** EXCL2XX = 0 **and** EXCL4XX = 0 **and** EXCL8XX = 0 **and** EXCL16XX = 0 **and** EXCL32XX = 0 **and** EXCL64XX = 0 **and** EXCL128XX = 0.

Use of derived fields

82. The following paragraphs provide details of some key algorithms used in the generation of rates of qualification from postgraduate research degree study. In particular, we provide details of algorithms used to define populations used in the method described in the main body of this publication and at Annex A.

Population definitions

Starting population

83. To identify the 2009-10 starting population[36] from the individualised data that enables identification of progression patterns between the reference year and the year that follows select EXCLUDE09 = 0 and STARTER = 1. This is the population for whom the rates of qualification from postgraduate research degree study are calculated.

[36] As defined at paragraphs 11 to 14 at Annex A.

Transition population

84. To identify the transition population[37] both sets of the individualised data are required.

85. The individualised data that enables identification of progression patterns between the reference year and the year that follows has REFYEAR = 2009-10 for all records and contains student records that are considered in the definition of the part of the transition population that examines transitions between 2009-10 and 2010-11. To identify this population from this individualised data select **(**TRANSITION = (i) **and** EXCLUDE09 = 0**)** along with **(**TRANSITION = (ii) **and** EXCLUDE08 = 0**)**.

86. The individualised file that enables identification of progression patterns between the year prior to the reference year and the reference year has REFYEAR = 2008-09 for all records and contains student records that are considered in the definition of the part of the transition population that examines transitions between 2008-09 and 2009-10. To identify this population from this individualised data select **(**TRANSITION = (iii) **and** EXCLUDE08 = 0**)** along with **(**TRANSITION = (iv) **and** EXCLUDE07 = 0**)**.

87. The populations extracted from the individualised data as described in paragraphs 85 and 86 are unified to form the transition population.

Classification of activity and the transition matrix

88. Students whose 2009-10 and 2010-11 data show evidence of an appropriate conversion from a full-time MPhil qualification to a postgraduate research degree are retrospectively added to the 2009-10 starting population[38]: in other words, at the earliest time that doctoral study can be identified. In the event that such a conversion is identified, the classification of a student's activity in 2009-10 is modified to enable appropriate counting of the record in the transition matrix and the associated calculations. That is, if MTOP09 = 1 then we have adjusted the STATE0 classification from STATE0 = MPhil to STATE0 = PhD 'YEAR0', where YEAR0 is one of the values 1 to 7.

89. In the generation of the rates of qualification from postgraduate research degree study we seek to project the proportions of a cohort of full-time research degree starters that are likely to be in each of three 'end states'. Paragraphs 43 to 48 of the main body of the report, and paragraphs 44 to 58 of Annex A, describe the use of a transition matrix to enable these projections of study outcomes. The transition matrix is constructed using the values of the STATE0 and STATE1 classifications for those student records forming the transition population, identified from the individualised files in the manner described at paragraphs 83 to 86.

Sector-adjusted averages

90. Paragraphs 51 to 53 of the main body of the report, and paragraphs 59 to 64 of Annex A, describe the provision of sector-adjusted averages to support the interpretation of the projected study outcomes. The sector-adjusted averages take into account the impact of variations in the subject area of study as well as students' domicile, highest qualification held

[37] As defined at paragraphs 15 to 43 at Annex A.

[38] The inclusion of these students in the 2009-10 starting population is explained further at paragraphs 11 to 14 at Annex A.

on entry and source of funding. The variables from the individualised data that are used in the benchmark calculations are detailed below.

Subject area of study

91. To identify students from a given subject area of study from the individualised data select PISUB = SCI and OTH for science-based and non-science-based students respectively. The algorithm for deriving PISUB is found at paragraph 51.

Students' domicile

92. To identify students from a given domicile from the individualised data select PIDOM = H, E and O for students domiciled: in the UK; in the EU, Guernsey, Jersey or the Isle of Man; and outside of EU member countries respectively. The algorithm for deriving PIDOM is found at paragraph 48.

Students' highest qualification on entry

93. To identify students with a given highest qualification on entry from the individualised data select NEWCAT3 = PG, FD and OT for students whose highest qualification on entry was: a postgraduate qualification; a first degree or graduate equivalent qualification; and other qualifications respectively. The algorithm for deriving NEWCAT3 is found at paragraph 47.

Students' source of funding

94. To identify students with a given source of funding from the individualised data select PIFUND = RC and OT for students whose source of funding was a Research Council or another source respectively. The algorithm for deriving PIFUND is found at paragraph 49.

Annex C: Linking processes

Outline of overall linking process

1. In order to link all available HESA records, a unique longitudinal identifier is created for each individual that appears at any point in the HESA record. This identifier is created as follows:

a. All students in a HESA individualised student record (year X) are matched to the following record (year X+1) using a number of match processes:

- records with matching HESA fields HUSID, HESAINST and NUMHUS (HIN linked)
- records matched on sex, birth date, first name and surname, with restriction for common names and an allowance for maiden name changes and spelling errors
- records matched on HUSID and either postcode, birth date, surname or first name
- records matched on HESAINST, HUSID, sex and surname with potential spelling errors or maiden name changes
- records matched on birth date, sex and first part of postcode. A combination of first name, HUSID and second part of postcode is further used to eliminate/select potential matches.

b. These five matching processes are also used to internally match up records belonging to the same student within a single academic year's HESA record. This internal matching is done for both year X and year X+1.

c. The identified matches are then resolved so that a single person identifier exists for year X and year X+1.

d. The process is repeated for matching between all pairs of years (X+1 and X+2, X and X+2, and so on).

e. The final step is to resolve all found links across all the years to produce a single HESA longitudinal identifier.

HIN linking

2. The HIN is a combination of three fields from the HESA individualised student records (the HESA unique student identifier, the UK provider reference number and the student instance identifier) which uniquely identify a student on a course (or instance of study). The HIN forms a year-on-year linking mechanism which can be used to track the student's progression on the instance from one year to the next. It follows that the HIN enables a student to be tracked from commencement of the instance through to completion.

3. Further information regarding the composition, usage and linking of the HIN is available from www.hesa.ac.uk under Data collection, Student and Year-on-year linkage.

Annex D: List of abbreviations

EU	European Union
HE	Higher education
HEFCE	Higher Education Funding Council for England
HEI	Higher education institution
HESA	Higher Education Statistics Agency
HIN	Husid, Institution, Numhus variables
HND/HNC	Higher National Diploma/Higher National Certificate
JACS	Joint Academic Coding System
PI	Performance Indicator
PISG	Performance Indicator Steering Group
QAA	Quality Assurance Agency for Higher Education
RDQR	Research degree qualification rate
TQI	Teaching Quality Information
UK	United Kingdom