

Annex C

Projected outcomes and efficiencies—technical notes and detailed information

1. The projected outcomes and learning efficiencies have been produced using a census cohort approach. The method involves identifying a group of starting students known as a cohort, consisting of all students who started at the institution on a full-time first degree course of study in a particular year. Their future progression outcomes—whether they qualify, transfer to another institution, or discontinue—are then projected based on current progression patterns at their institution.
2. This annex provides details of the method used, and shows the complete transition matrix for these factors, of the type used to produce Table T5. A similar matrix was calculated for each institution.

Technical details

General information

3. The progression behaviour of all the students at an institution can be summarised in a transition matrix. The rows of this matrix are possible starting states, the columns possible end states. The values in these cells of each row are the proportions of students from that starting state who make the transition to the end state. One starting state (row) might be 'full-time first degree year of programme 1', a possible end state (column) could be 'part-time first degree at the same institution'; the value at their intersection would be the proportion of students who are on year of programme 1 of a full-time first degree who then move to a part-time first degree. By defining states such as recently absent students and those moving to other modes or levels of study, we can allow for students returning to study or being awarded a degree from part-time study, for example.
4. It is assumed that the probability of moving from one state to another is the same, regardless of how the state has been reached. So, for example, this means that a student who has gone from first year to second year to third year, and is now in the third year of a degree has the same probability of qualifying at the end of the year as a student who transferred into the second year of the course from another institution, or one who repeated a previous year. It also means that we allow the possibility of a student repeating a year more than once (which in practice is usually not acceptable). Although the assumption is not completely accurate, any errors it introduces are generally small. The effect of the errors will usually be to slightly over-estimate the time a student will spend in the system, and under-estimate the proportion of students leaving with no qualification.
5. Some small institutions, and institutions which have made major changes either to the format of the degree programmes or to the way they record that format, may find that the projections lead to some difficulties. For example, if an institution had no record of students reaching an 'unknown' state. This can occur if there were no students in a particular stated year, such as 1999-2000, but there was movement into that state in 2000-01. For example, if an institution had no record of students moving into year of programme 5, but no record of students moving into year of programme 5, but no record of being of unknown outcomes.

Efficiency

6. The efficiency measure is defined as the ratio of the total efficient time for all starting students to the actual time taken by these students. The efficient time is the time that they would be expected to take to obtain a qualification if they did not repeat a year, change course and soon. To produce the measure of efficiency, we need to know from which year of programme the student has qualified, or between which programme years they are transferring. A student obtaining a degree from year of programme 4, for example, is assumed to be on a degree course that should take four years, and is allowed four notional years as his or her assumed time to complete the course.

her 'efficient time'. Similarly, a student who transfers from year of programme 1 at one institution to year of programme 2 at another institution is allowed one optional year for efficient time for the first institution; but a student transferring from year 1 at the first institution to year 1 at another institution gets no optional years. Students who leave higher education with no qualification have an efficient time of zero.

Projected outcomes

7. To ensure that the data used in the calculations are robust, it is important to define the stages that a student has reached by reference to four years of data. These are: the 'reference year', which is the year to which all calculations relate; the two years prior to the reference year; and the year following the reference year. For this publication, the reference year is 1999-2000.
8. At present, the method used is restricted to students who are residents of the United Kingdom, excluding the Channel Islands and the Isle of Man, who are studying full-time for a first degree. In what follows, the term 'student' is taken to refer to someone with these characteristics, unless it is further qualified.
9. There are three relevant states in which a student can be in any one year. These are:
- State **A**: Active at the institution as a full-time student with first degree qualification aim
 - State **a**: Active at the institution as an undergraduate student not in state **A** (could be full-time sub-degree, part-time degree and soon)
 - State **D**: Not active at the institution as an undergraduate student (could be at another institution)
10. The starting population consists of students who studied full-time for a first degree in the reference year (in other words, were in state **A**) and were not in that state in either of the previous two years (in other words, were in one of the states **a** or **D** in the two years preceding the reference year). Note that a house has been made of the date of commencement of studies in defining the starting population. The term 'entrants' is reserved for students reported as commencing studies in a given year, and 'starters' for members of the starting population as defined here.
11. As well as the starting population, we need to define a 'transition population', which will be used to determine the pattern of progression at the institution. This transition population is defined as all students who are in state **A** at the institution in the reference year, plus all those students who were in state **A** in the previous year and are in state **a** or state **D** in the reference year but have not obtained a degree. This is essentially all students who are currently, or were in the previous year, full-time first degree students. Students who are on a sub-degree programme are only included if they have previously transferred from a degree programme.
12. For each student in the transition population, we need to know what year of programme they were on in the reference year, or what year of programme they were on previously if they are currently in state **D**. For the year following the reference year, we need to know:
- if they have qualified with a degree or a sub-degree following study in the reference year
 - which year of programme they are on if they are still at the institution, and whether they are still studying full-time for a first degree
 - if they have transferred to a sub-degree or part-time programme
 - if they have transferred to another institution.
13. A student is assumed to have discontinued with 'no award or transfer' if there is no record of the student as a full-time first degree student in the reference year, and no record as an undergraduate student in the

next year. A student who is in state **A** in the reference year, but for whom no record can be found the following year, is initially deemed to become 'inactive', with 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 'neither award nor transfer'.

14. A student who is awarded a sub-degree qualification from state **A** is also initially placed in the inactive state. This allows for students whomay return after a year to complete their degree programme. Students in state **a** in the reference year are categorised as failing to get an award if they cannot be traced in the following year.

15. To calculate the transition matrix, the states students are in need to be specified more fully. The specification includes mode of study, year of programme, and whether or not the student has obtained a qualification. The values in the matrix are then found as:

$$t_{ij} = n_{ij} / n_i$$

where n_{ij} is the number of students at the institution in state i in the reference year and state j in the following year, and n_i is the number of students at the institution in state i in the reference year.

16. An explanation of the states used is given in Table C3, and Table C5 provides a simplified form of the transition matrix for the sector. The actual computations used a more detailed matrix than that shown here.

17. The matrix used in calculating the adjusted sector benchmark is produced in a similar way, using all the students in the sector. However, each student is given a weight which reflects the relative importance of their subject of study and entry qualifications at the institution, and it is the sum of these weights, rather than the actual number of students, that are used to calculate the values in the matrix. The weights are defined as:

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

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 number of students in the population in the institution and the sector respectively.

18. Several of the states are what are known as 'sink states'. Such states are characterised by the fact that once a student enters such a state, that is the end of their progress through the system as far as the institution is concerned. Four of the 15 states shown are sink states, namely 'qualifying with a degree', 'qualifying with a sub-degree without subsequent degree-level study', 'transfer', and 'absent'. Students cannot return to the institution, as a continuous part of their progress, from any of these states.

19. The data are then used to create a transition matrix **T** from the students in the transition population. Each row represents a state in the reference year, with each column representing a state in the following year. Each row of the matrix contains details of the percentage of students who were in the row state in the reference year, who are in the column state in the following year. Rows representing sink states contain only a non-zero value, on the diagonal, as all students in such a state will remain there.

20. The starters are described by the vector **n** which gives the number of starters in each of the states. All starters will be on a first degree programme, and most will be in year of programme 1, but at some institution there may be considerable numbers of starters in later years.

21. Each state has an associated FTE, described by vector **f**, giving the amount of study time incurred by being in that state for a year. The value of FTE is either 0 or 1, with 1 taken for all full-time years of study, and 0 for all part-time study, as well as for inactive and unknown states.

22. The progression of the starters can be computed by applying the transition matrix to the vector of starters. After i years the cohort will be in state $\mathbf{n} \cdot \mathbf{T}^{i-1}$. A total of $\sum_{i=1}^{\infty} (\mathbf{n} \cdot \mathbf{T}^{i-1}) \cdot \mathbf{f}^T$ study time will have been used in year i , and the 'actual' time used is given by the sum of these values for large i . The value of i is considered 'large' when the values in $\mathbf{n} \cdot \mathbf{T}^{i-1}$ of the numbers of students qualifying or transferring no longer change as i increases. At that stage, the value of $(\mathbf{n} \cdot \mathbf{T}^{i-1}) \cdot \mathbf{f}^T$ is also effectively zero.

23. In some circumstances—for example, when an institution is starting a new course, or changing the pattern of an existing course—there are values in the state vector $\mathbf{q} = \mathbf{n} \cdot \mathbf{T}^{i-1}$ for large i which are not in sink states. In such cases the outcome is recorded as unknown, and the projected outcomes should be treated with caution.

Efficiency

24. The efficiency of an institution has been defined in terms of how long students should take before reaching a valid endpoint compared to how long they actually do take. We have defined a number of outcomes as 'valid ends' for this purpose, namely qualifying with either a degree or a sub-degree, and transferring to another HEI.

25. Each state has a start and end value, described by vectors \mathbf{s} and \mathbf{e} . These are used to calculate the efficient time taken. Any state which is not a 'valid end' has an end value of 0. Any state which is a 'valid end' has an end value which reflects the time that should be used in reaching it, its efficient time. These start values are only relevant to the year of degree programme start, while each has a start value one less than their year of programme value.

26. The efficient time is obtained by looking at the year of programme in which the end is reached. In general, each (non-repeated) year of study leads to one year for efficient time. So a student qualifying with a degree at the end of year of programme k is allotted k years for efficient time. A student obtaining a sub-degree qualification, from any year of programme after the first, is allotted two years for efficient time; a student obtaining a degree qualification from part-time study is allotted three years for efficient time. A student transferring from year of programme x at the institution to year of programme $(x+1)$ at another institution is allotted x years, while a student transferring from year of programme x to the same year of programme at another HEI is allotted $(x-1)$ years for efficient time. A student who is found on postgraduate study at the same institution but without a qualification is assumed to have obtained a degree, giving three years for efficient time. A student who is found on postgraduate study elsewhere is assumed to have obtained a sub-degree qualification, giving two years for efficient time.

27. The computations, using the above notation, are fairly straightforward. The actual time taken by a cohort of students at an institution is the sum of all the values $(\mathbf{n} \cdot \mathbf{T}^{i-1}) \cdot \mathbf{f}^T$ with summation over all values of i .

28. The efficient time taken by the cohort is given by:

$$\mathbf{q} \cdot \mathbf{e}^T - (\mathbf{q} \cdot \mathbf{c}^T) / \left(\sum \mathbf{n} \right) \cdot (\mathbf{n} \cdot \mathbf{s}^T)$$

where $c_i = 1$ if $e_i > 0$

$c_i = 0$ otherwise

Summation is over all values in the vector. The first term would be the efficient time if all starters were on year of programme 1; the second term is an adjustment to allow for starters on other years of programme.

29. Until the 1998-99 HESA return, there was no way of recording foundation year students. There were differences in the way foundation year students were returned by different institutions, and this had a slight

effect on the values taken for efficient time. For data from 1998-99, we have been able to obtain information from the HESA record about foundation year students, which we have used in the calculations.

30. Where an institution has returned foundation year students (YEARPRG=0), we have increased the efficient time for the institution by the number of such students. This errs on the generous side.

Standard deviations

Projected outcomes

31. The method of obtaining the adjusted sector outcomes is explained in paragraph 17 of this annex. Table C1 below shows the notation used to denote the proportions of students in each outcome category for both actual and adjusted outcomes.

32. We have only calculated the standard deviation for the 'neither transfer nor obtain any qualification' outcome category. If there are starters at an institution, and the benchmark for this outcome category is $(100 \times \pi_d) \%$, then the standard deviation of the proportion in category d has been calculated as

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

which assumes that the starters have been chosen randomly 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. This assumption is reasonably good unless the pattern of outcomes of the benchmark is quite different to the pattern of the actual values. In such cases, we have omitted the standard deviation altogether.

Table C1 Outcomes

Outcome category	Projected proportion	Adjusted sector proportion
Qualify with a degree	p_a	π_a
Qualify with a sub-degree	p_b	π_b
Transfer to another institution	p_c	π_c
Neither transfer nor obtain any qualification	p_d	π_d
Unknown	p_e	π_e

Standard deviation of efficiency

33. In defining efficiency, the length of time a student took in the system was compared to the efficient time for their outcome, that is, the time they would be expected to take to reach their particular outcome with no hitches on the way. Table C2 gives a few examples of factual and efficient times, and possible routes to these values. It is not a comprehensive list; it is provided to show the types of assumption we have made.

34. In producing the projected outcomes for the adjusted sector benchmarks, using the adjusted transition matrix, the number of students who actually took year x to reach an outcome when the efficient time as y years can be obtained. This is shown as n_{xy} in Table C2. The total number of students in the cohort (the starters) is

$$n = \sum_{\text{all } x \text{ and } y} n_{xy}$$

35. The average efficient time (\bar{x}) and the average actual time (\bar{y}) can be obtained as

$$\bar{x} = \frac{\sum_{\text{all } x \text{ and } y} x n_{xy}}{n}$$

and
$$\bar{y} = \frac{\sum_{\text{all } x \text{ and } y} y n_{xy}}{n}$$

The efficiency is then defined as $100 \times \bar{x} / \bar{y} \%$.

36. The standard deviations of both times (s_x and s_y), and their covariance (s_{xy}), are given by

$$s_x = \sqrt{\frac{\left(\sum_{\text{all } x \text{ and } y} x^2 n_{xy} \right) - n \bar{x}^2}{(n-1)}}$$

$$s_y = \sqrt{\frac{\left(\sum_{\text{all } x \text{ and } y} y^2 n_{xy} \right) - n \bar{y}^2}{(n-1)}}$$

and
$$s_{xy} = \frac{1}{(n-1)} \left(\sum_{\text{all } x \text{ and } y} x y n_{xy} - n \bar{x} \bar{y} \right)$$

37. The usual formula for the standard deviation of the efficiency as a ratio can then be applied to give the standard deviation

$$\text{s.d. (eff)} = \sqrt{\frac{1}{n\bar{x}^2} \left(s_y^2 + \left(\frac{\bar{y}}{\bar{x}} \right)^2 s_x^2 - 2 \left(\frac{\bar{y}}{\bar{x}} \right) s_{xy} \right)}$$

Transition matrix and starters

38. Table C4 shows the sector profile of starters on first degree courses by year of entry to the course. The majority of students (90 percent) start on the first year, but a few start in year 2 or later.

39. Table C5 shows the transition matrix for these courses. As can be seen, most students in a particular state either move to the next year of programme, or qualify for programme 1 in 1999-2000, 77 percent moved on to year 2 of programme 2, 85 percent moved on to year 3 of programme 2 in 2000-01; and of those on year 3 of programme 2, 85 percent moved on to year 3 of programme 3.

40. Also worth noting is the fact that 74 percent of students who leave an institution in any year do not return to HE at the end of the following year. Five percent obtain some sort of qualification, 7 percent transfer to another institution, and the remainder return to their original institution.

41. Between 35 percent and 40 percent of students whom move to a part-time course of study return to full-time study after a year. In addition, over 20 percent achieve a qualification, at either degree or sub-degree level.

42. The transition matrices are only looking at patterns of movement for a particular institution. As a student ends their course, as far as that institution is concerned, they have left. However, many of these transferring students will go on to obtain a qualification at their new institution. This is taken into account in the sector projections given in paragraph 104 of the main document.

Table C2 Efficiency examples

Example	Efficient time (x)	Actual time (y)	Number of students	Comment
Student starts in first year, leaves after 1 December, and does not continue anywhere else.	0	1	n ₀₁	
Student completes the first year at institution A, then transfers to the second year of a course at another institution.	1	1	n ₁₁	Note that this is the time assigned to institution A
Student enters final year of course, having transferred from diploma course. Spends one year in the final year before qualifying.	1	1	n ₁₁	Shows another way of obtaining n ₁₁
Student starts on first year, completes first year, then transfers to second year of a diploma course. Completes this successfully.	2	2	n ₂₂	
Student starts in first year, has to repeat first year but does this part-time, then completes second and third years with no problem	3	3	n ₃₃	Part-time study counts as 0 years
Student starts in first year, progresses to second year, then repeats second year on a full-time basis, and moves to third year which is completed normally.	3	4	n ₃₄	
Student starts in first year, decides the course is not as expected and transfers after a year to the first year of another degree course which lasts four years. Completes this course with no problems.	4	5	n ₄₅	

Table C3 States used in the transition matrix

State	Description
QFD	Qualified with degree
QSD	Qualified with sub-degree
FTFD1	On full-time first degree, year 1
FTFD2	On full-time first degree, year 2
FTFD3	On full-time first degree, year 3
FTFD4	On full-time first degree, year 4
FTFD5	On full-time first degree, year 5
FTFD6+	On full-time first degree, year 6 or above
FTSD	On full-time sub-degree (having previously been FTFD)
PTFD	On part-time first degree (having previously been FTFD)
PTSD	On part-time sub-degree (having previously been FTFD)
Transfer	Transferred to another HEI
Inactive	Not currently in HE, but was FTFD in previous year
Unknown	Unknown (cannot be allocated to another state)
Absent	Absent from HE

Table C4 Starters on full-time first degree programmes by start year of programme

Year1	Year2	Year3	Year4	Year5	Year6	Number of starters
90%	5%	4%	1%	0%	0%	284,000

Table C5 All institution transition matrix for sector
 (Figures given are percentages)

	QFD	QSD	FTFD1	FTFD2	FTFD3	FTFD4	FTFD5	FTF	D6+	FTSD	PTFD	PTSD	Transfer	Inactive	Absent	Base numbers
QFD	100															
QSD		100														
FTFD1	1	0	6	77	0	0	0	0	0	1	0	3	10	0		272,500
FTFD2	1	0	0	4	85	1	0	0	0	1	0	1	6	0		236,300
FTFD3	62	0	0	0	4	27	0	0	0	1	0	0	4	0		230,400
FTFD4	85	0	0	0	0	2	9	0	0	1	0	0	3	0		68,300
FTFD5	87	0	0	0	0	0	3	7	0	1	0	0	1	0		6,600
FTFD6+	82	2	0	0	0	1	1	5	0	0	0	1	8	0		500
FTSD	1	18	5	3	4	1	0	0	41	1	1	4	0	22		2,100
PTFD	24	2	1	14	17	2	0	0	0	14	0	2	0	22		9,700
PTSD	0	11	1	8	19	3	0	0	2	4	9	3	0	38		500
Transfer													100			
Inactive	0	5	1	3	6	1	0	0	0	1	0	7	0	74		56,300
Absent															100	

The last column, 'Base numbers', shows the number of students on which each row is based.

The four rows for which only one value is given are 'sink states'; once a student enters one of these states, that is the end of their path through the HES system.

The matrices for individual institutions may include an extra state, 'Unknown', to allow for students moving into a state which is not currently represented at that institution.