Technical algorithms for institutional performance measures

Changes 2019 to 2020

Enquiries to providermetrics@officeforstudents.org.uk

Publication date 7 May 2020

Contents

Introduction

Purpose

1. This is one of a series of technical documents that provide detail of the definitions and methodology used by the Office for Students (OfS) in constructing institutional performance measures[1]. Wherever possible we have used consistent definitions and approaches, to minimise the burden on higher education providers of understanding these. This document provides a technical description of the categorisations applied to individualised student data records and currently used by the OfS to generate the indicators underpinning the OfS's functions related to:

    a. The assessment of registration condition B3[2] for the purposes of initial registration and ongoing monitoring.

    b. The regulation of access and participation.

2. This document relates to the institutional performance measures produced in March 2020 in support of these two functions. It is also relevant to the indicators constructed by the OfS as our Key Performance Measures, and used more widely in reporting sector-level analysis of patterns and trends in student outcomes[3]. It supplements, and should be read alongside, the following documents:

    - 'Technical algorithms for institutional performance measures: 2020 core algorithms'[4]
    - 'Technical algorithms for institutional performance measures: Regulatory indicators, methodology and rebuild descriptions'[5]
    - (for registration condition B3 purposes) 'Condition B3: baselines for student outcomes indicators'[6]
    - (for access and participation purposes) 'Access and participation data resources: Dashboard user guide' and 'Access and participation data resources: Supporting data user guide'.[7]

[1] See www.officeforstudents.org.uk/data-and-analysis/institutional-performance-measures/technical-documentation/

[2] See www.officeforstudents.org.uk/advice-and-guidance/regulation/conditions-of-registration/.

[3] See www.officeforstudents.org.uk/about/measures-of-our-success/.

[4] See www.officeforstudents.org.uk/data-and-analysis/institutional-performance-measures/technical-documentation/

[5] See www.officeforstudents.org.uk/data-and-analysis/institutional-performance-measures/technical-documentation/

[6] See www.officeforstudents.org.uk/publications/registration-key-themes-and-analysis/

[7] See www.officeforstudents.org.uk/data-and-analysis/access-and-participation-data-dashboard/

3. The algorithms used to define the institutional performance measures underpinning a number of OfS functions have been refined for the outputs produced in March 2020. This document explains the nature and impact of those changes, by comparison with the algorithms previously published by the OfS (in the 2019 publication of 'Technical algorithms for institutional performance measures: Core algorithms')[8].

Improving alignment within OfS institutional performance measures

4. Readers may be aware that a number of OfS functions make use of similar institutional performance measures that are constructed from individualised student data returns. These functions currently include our regulation of access and participation, and of quality and standards through the initial registration and ongoing monitoring of registration condition B3. We anticipate the continued use of some or all of these measures within the evidence base supporting a revised Teaching Excellence and Student Outcomes Framework (TEF).

5. To support the coherence of the evidence base that underpins the OfS's regulatory functions, and an improved user understanding of that evidence base, we have been reviewing the definitions and algorithms that are applied in the production of institutional performance measures. This work has sought to identify and eliminate any unnecessary definitional differences across these measures. To date, the review has highlighted a number of areas where alignment of the definitions and algorithms could be improved. It has also highlighted opportunities to establish principles that may improve the consistency of our approach to some aspects of our data outputs, and a number of refinements to algorithms that are desirable based on changes in the wider data landscape.

6. The OfS is responding to these review findings by taking forward a series of refinements to the definitions used in institutional performance measures. A first phase of this work is reflected in this document. It refers to definitional changes reflected in the document 'Technical algorithms for institutional performance measures: 2020 core algorithms', which represent an evolution of those previously published by the OfS (in the 2019 publication of 'Technical algorithms for institutional performance measures: Core algorithms'). We anticipate that further development will follow – in particular, to accommodate the outcomes of forthcoming consultation on future approaches to the TEF and to ongoing monitoring of registration condition B3 later in 2020.

7. The first phase of refinements reflected in the 2020 core algorithms document includes some that are primarily cosmetic (and intended to simplify user understanding and navigation) and a small number that respond to changes in supporting data infrastructure (and have been included in the scope of previous consultation activity). Modelling and regression testing of these algorithm refinements have determined that their impact is marginal throughout, both singularly and collectively. To the extent that the impact of these refinements is non-negligible in some individual circumstances, the OfS considers that this impact is outweighed by the added value they afford and will consider their regulatory consequences on a case-by-case basis.

8. The primary change reflected in the 2020 core algorithms document relates to naming conventions for derived variables. Our previously published algorithms used a naming

[8] See www.officeforstudents.org.uk/data-and-analysis/institutional-performance-measures/technical-documentation/

convention in which all derived variable names were prefixed with ‘B3MON’. Readers should note that all derived variable names in this document and the supporting individualised files now use ‘IP’ as the prefix. This change signals our intention that derived variables are aligned to a single definition that serves multiple uses across different OfS institutional performance measures. For a large number of the derived variables in the 2020 core algorithms document, the change of prefix in the variable name represents the full extent of the change between these algorithms and those previously published.

9. However, a number of changes reflected in the 2020 core algorithms document are more substantial change than naming conventions. To support users in identifying and understanding the nature of the changes, each algorithm’s description in that document notes the type of change (or lack thereof) from the previously published algorithms. The changes are described more comprehensively in this document and fall into four broad categories:

    a. Variables that have been renamed, more substantially than a change of prefix from ‘B3MON’ to ‘IP’. The definition of these variables remains unchanged, and the change of name results from efforts to simplify algorithms for user understanding and navigation of supporting individualised files.

    b. Addition or removal of variables, which results from efforts to streamline and simplify algorithms for institutional performance measures wherever possible.

    c. Variable definitions that have been refined compared with the equivalent version previously published, and where the definitions included in this document reflect minor improvements or clarifications to the categorisations we are applying.

    d. Variable definitions that have changed more substantially by comparison with the equivalent version previously published, to reflect changes in OfS policy or underlying data infrastructure.

10. This document further describes, and where appropriate quantifies, the impact of these refinements on the institutional performance measures generated at provider level.

Guidance for using this document

11. The algorithms referenced in this document are applied to the 2010-11 to 2017-18 individualised student records collected annually by the Higher Education Statistics Agency (HESA) or the Education and Skills Funding Agency (ESFA). This document is aimed at readers with in-depth knowledge of the HESA, HESA alternative provider or Individualised Learner Record (ILR) student data.

12. Individualised student data files are supplied to higher education providers by the OfS to support their understanding of our approach to calculating institutional performance measures. These files contain data relating to a provider’s own students and show how they have been categorised according to the algorithms referenced in this document. Similar files were made available to providers in 2019, containing the categorisations applied on the basis of our previously published algorithms.

Enquiries and feedback

13. Enquiries regarding the algorithms and changes described within this document, and any other questions about the OfS's approach to the construction and use of institutional performance measures, should be raised with providermetrics@officeforstudents.org.uk, 0117 931 7230.

Description of changes

Changes to data infrastructure

14. For the first time, the algorithms detailed in the 2020 core algorithms document incorporate the specifications of the 2018-19 individualised student data returns as submitted to HESA or the ESFA.

Changes to accommodate missing FTE (STULOAD) information in ILR returns

15. The full-time equivalence (FTE)[9] is calculated for each instance of study and is primarily used in institutional performance measures to provide student number counts that act as contextual data. In the measures themselves, FTE only plays a role in calculating continuation outcomes (where it helps indicate whether a student is in active study in the years assessed).

16. The STULOAD variable may sometimes be absent for a small number of ILR student records. Where this has occurred, rather than treating the FTE information as missing, it is now derived using a method also used in the OfS student numbers data (for the purposes of calculating OfS registration fees). Information on this approach, including a technical description of the algorithms used, can be found at www.officeforstudents.org.uk/data-and-analysis/student-number-data/.

17. The adoption of the OfS student numbers approach to missing FTE information for ILR student records has a minimal impact in a vast majority of cases. This is because missing FTE information tends to be concentrated in provision that is largely outside of the scope of existing institutional performance measures. Where it does impact, the change represents an improvement to the accuracy of continuation outcomes being reported.

Changes to classifying learning aims and levels of study

18. For student records drawn from the ILR, the HEFQAIM variable was previously used to allocate learning aims to broad types of higher education formerly recognised for funding by the Higher Education Funding Council for England. The HEFQAIM variable has been replaced within the 2020 core algorithms by the IPOFSQAIM variable, which allocates ILR learning aims to broad types of higher education, irrespective of whether they are recognised for OfS funding purposes. This follows the replacement of HEFQAIM with OFSQAIM in OfS algorithms used for funding purposes, as detailed in the ‘Classifying learning aims technical document’ for 2018-19

[9] The concept of full-time equivalent student numbers is defined in full at www.hesa.ac.uk/support/definitions/students.

ILR data[10]. For the avoidance of doubt, the definition of IPOFSQAIM is identical to the definition of OFSQAIM for 2018-19 ILR student records[11].

19. To improve consistency in allocating learning aims to broad types of higher education, the 2020 core algorithms also define the IPOFSQAIM variable for student records drawn from the HESA student and HESA AP student returns. This variable now provides a consistent level of granularity underpinning the allocation of students to various level of study aggregations, allowing for the elimination of the two variables, AIMTYPE and LEVEL_DETAIL, that previously sought to capture this detail for student records returned to HESA.

20. In reviewing the differences between the HEFQAIM and IPOFSQAIM mappings, we have determined that some edge-case qualifications may now feature in different categories at this level of granularity:

    a. For student records drawn from the ILR, this reclassification is wholly consistent with that which impacts on OfS funding calculations.

    b. For student records returned to HESA, the low level of reclassification results from refinements implemented following a detailed review of the placement of each COURSEAIM value within this definition. The most noteworthy reclassification affects the COURSEAIM value I16, which would previously map to a first degree (FIRST) categorisation in all cases and may now alternatively map to a medicine, veterinary and dentistry (MEDVETDENT) categorisation if the provision is regulated by one of the relevant bodies.

21. Institutional performance measures produced by the OfS normally report on levels of study at a broader level of aggregation than the full range of IPOFSQAIM categories. As a result, the nesting of IPOFSQAIM values into this more aggregated level of study categories negates the impact of some of the reclassifications at the most granular level.

22. However, the IPOFSQAIM variable allocates learning aims to broad types of higher education irrespective of whether they are recognised for OfS funding purposes, to better reflect the full scope of OfS regulation. This means that some student records pertaining to higher education study drawn from the ILR will now be reclassified, having previously been categorised as 'OTHER' by the HEFQAIM algorithm. While all HEFQAIM values of 'OTHER' were previously excluded from the coverage of institutional performance measures, all values of IPOFSQAIM other than 'Further education course' have the potential to be included in this coverage. Because the coverage of the 'HE learning delivery' in the ILR specification is currently incomplete across all higher education aims, these aims remain outside the coverage of institutional performance measures for the time being (and the variable IPHEENTITYFLAG has been introduced to facilitate this). Readers should note the OfS's intention that – in the fullness of time, and pending both consultation and extension of the coverage of the 'HE learning

[10] See www.officeforstudents.org.uk/data-and-analysis/data-checking-tools/2018-19-ilr-data-checking-tool/.

[11] The IPOFSQAIM algorithm has been applied to earlier years of data retrospectively, and the algorithm definition provided in the 2020 core algorithms document includes clauses accommodating these earlier years.

delivery' entity of the ILR – all higher education students should be included in all institutional performance measures.

23. Readers should also note that level of study is used throughout the indicator and algorithm definitions used by the OfS. Minor differences may result in terms of classification at individual student level compared with other variables whose definitions depend on level of study (or level of qualification awarded). Examples include mode of study and other variables used to calculate continuation outcomes.

24. A more minor refinement affecting level of study concerns the level of qualification obtained, and occasions where there is a mismatch between this and the level of the qualification aim. The OfS is aware of a number of such mismatches, in particular undergraduate qualifications awarded from a postgraduate aim, which previously resulted in the students being reported as an undergraduate qualifier based solely on the level of qualification awarded. We consider that the student experience and fee level will both reflect the postgraduate course that the student had been following (rather than the undergraduate award they received), and the algorithms for institutional performance measures have been refined to facilitate this assignment where it is most appropriate. The OfS will monitor the impact of this approach, in particular with regard to more flexible forms of study where student intentions may be less clear-cut, and providers are invited to submit any feedback to providermetrics@officeforstudents.org.uk.

25. The impact of the changes described in paragraphs 18 to 23 is illustrated in Tables 1 and 2. Table 1 shows a comparison of student level of study categorisations before and after changes to algorithms, where direct matches between the new and previous categorisations are shaded blue, and any mismatches are highlighted in yellow. Table 2 focuses on the number of student instances previously categorised as other (OTH) and shows the extent to which these students will remain outside the current scope of institutional performance measures based on the status of the 'HE learning delivery' entity within ILR records (as determined with the variable IPHEENTITYFLAG).

Table 1: Difference in level of study categorisations, before and after changes to algorithms

Previous LEVEL categorisations (combining LEVEL and AIMTYPE)	New IPLEVEL categorisation: PHD	OPGR	PGTM	PGCE	OPGT	PG CREDIT	PG UNSPEC	DEG	PUGD	PUGO	OUG	UG CREDIT	UG UNSPEC	Other (FE and NA)
PHD	148,400													
OPGR		900												
PGTM			445,000											
PGCE				26,000										<100
OPGT			7,400		88,600	600								
PGCREDIT						36,800								
PGUNSPEC							1,200							
DEG								139,100	2,000					<100
PUGD									131,200	300				
PUGO										13,700				
OUG											166,400	2,100		
UGCREDIT												84,900		
UGUNSPEC													2,200	
OTH					400	<100					43,300	12,000		7,200

Note: Student numbers rounded to the nearest 100.

Table 2: Students previously categorised as other (OTH) by status of the ILR 'HE learning delivery' entity

	New IPLEVEL categorisation: OPGT	PG CREDIT	OUG	UG CREDIT	Other (FE and NA)
Entity not required, not included in measures (IPHEENTITYFLAG = 0)	400	<100	41,700	11,900	<100
Entity required, included in measures (IPHEENTITYFLAG = 1)	0	0	1,600	100	7,200

Note: Student numbers rounded to the nearest 100.

Changes to continuation algorithms

26. In addition to the work to review the definitions and algorithms applied in the production of institutional performance measures, the OfS is also undertaking a programme of development work looking at the package of continuation measures. This work has the aim of:

    a. Improving alignment with emerging work to measure student transfers in support of the OfS's obligations in relation to section 38 of the Higher Education and Research Act 2017[12], and to measure completion within the transparency information condition[13].

    b. Developing new measures that can supplement the understanding of continuation outcomes afforded by the existing OfS continuation and completion indicators used as institutional performance measures.

27. Subsequent phases of the development work will seek to introduce (via consultation) a range of measures able to take a nuanced view of issues related to continuation before the end of the year of entry, as well as those related to resumptions of study.

28. However, a first phase of this work has involved refining the algorithms for our existing continuation measures to retain further granularity from the building blocks that contribute to that indicator. This work allows us to retain additional information – including whether the student continues at the same level and subject of study, and continuation outcomes after each of one, two and three years following entry – while still being able to aggregate it to create our established measure. As a consequence, the 2020 core algorithms describe a series of continuation algorithms that have been restructured to facilitate these additions, while also maintaining the existing coverage and outcomes.

29. The previously published algorithm variables of IPCONINDFULL and IPCONEXCL report the established continuation outcomes and exclusions, to show outcomes evaluated one year after entry for full-time students and two years after entry for part-time students. A number of additional variables have also been created. While we have no reason to doubt that the algorithms used to create these additional variables are robust, we note that they should be considered experimental. This experimental classification signals only that the algorithms are newly developed and undergoing evaluation. The OfS will continue to evaluate the new algorithms and providers are invited to submit any feedback to providermetrics@officeforstudents.org.uk.

30. The continuation algorithms have been updated to identify the continuation outcomes one, two and three years after entry with the variables IPCONINDFULL_Y1, IPCONINDFULL_Y2, and IPCONINDFULL_Y3. Additionally, continuation outcomes now include information on subject-level continuation, identifying cases where the activity in the later year is in the same subject as

[12] See www.legislation.gov.uk/ukpga/2017/29/section/38/enacted.

[13] See www.officeforstudents.org.uk/publications/regulatory-advice-8-guidance-for-providers-about-condition-of-registration-f1-transparency-information/.

the subject studied in the year of entry[14]. These outcomes are identified in the variables IPCONINDFULLSBJ_Y1, IPCONINDFULLSBJ_Y2, and IPCONINDFULLSBJ_Y3.

31. In addition, the continuation outcomes reported via the 2020 core algorithms describe outcomes at a more granular level within IPCONINDFULL variables. The outcome is assigned according to the following hierarchy:

    - 'QUALIFIED': The student qualified, at the same provider, at an equivalent level or higher than their study in the entrant year.
    - 'CONTINUING_H': The student was in active study, at the same provider, at a higher level than their study in the entrant year.
    - 'CONTINUING': The student was in active study, at the same provider, at an equivalent level to their study in the entrant year.
    - 'TRANSFER': The student was in active study (or qualified), at a different provider, at an equivalent level or higher than their study in the entrant year.
    - 'QUALIFIED_L': The student qualified, at the same provider, at a lower level than their study in the entrant year.
    - 'CONTINUING_L': The student was in active study, at the same provider, at a lower level than their study in the entrant year.
    - 'TRANSFER_L': The student was in active study (or qualified), at a different provider, at a lower level than their study in the entrant year.
    - 'INACTIVE': The student was not found in active UK study (at an appropriate level).

32. In addition, the 2020 core algorithms have implemented some minor improvements identified in the review of our existing algorithms. These improvements have a minimal impact on the continuation outcomes calculated within institutional performance measures:

    a. Following changes to data reporting guidance some years ago, whether a student has been studying on a standard or non-standard academic year basis is now redundant for all but Scottish providers in establishing whether or not the student is active at the relevant census point. All cohorts who were impacted by the previous reporting guidance are no longer included in the time series for any institutional performance measure. The IPCONACTIVE algorithms have been simplified for English, Welsh and Northern Irish providers by removing this parameter, and only retaining it for Scottish providers.

    b. Within the previously published algorithms, there was an underlying assumption that any postgraduate student who progressed to an undergraduate qualification on or before the relevant census date was counted as 'inactive' in postgraduate study. This assumption did not apply if the progression was to an 'undergraduate programme with postgraduate components', despite this category being otherwise treated as an undergraduate aim

[14] Where subjects are defined at level 2 of the Common Aggregation Hierarchy, as at www.hesa.ac.uk/innovation/hecos.

throughout the institutional performance measures. Removing progression to an 'undergraduate programme with postgraduate components' is intended to improve alignment of coverage and user understanding.

c. In the previously published algorithms, there was an underlying assumption that any undergraduate student who progressed to a postgraduate qualification on or before the relevant census date was counted as an undergraduate qualifier, even if no undergraduate qualification award had been returned in the HESA or ILR student data. This assumption applied, and resulted in a positive continuation outcome, even if the student then became inactive in postgraduate study before the relevant census date. The 2020 core algorithms no longer make this assumption, so this scenario will now result in a continuation outcome of 'inactive'.

33. The collective impact of the changes described in this document on full- and part-time continuation outcomes is illustrated in Figures 1 and 2, where each blue dot along the x-axis represents a provider.

Figure 1: Difference in full-time continuation outcomes, after changes to algorithms

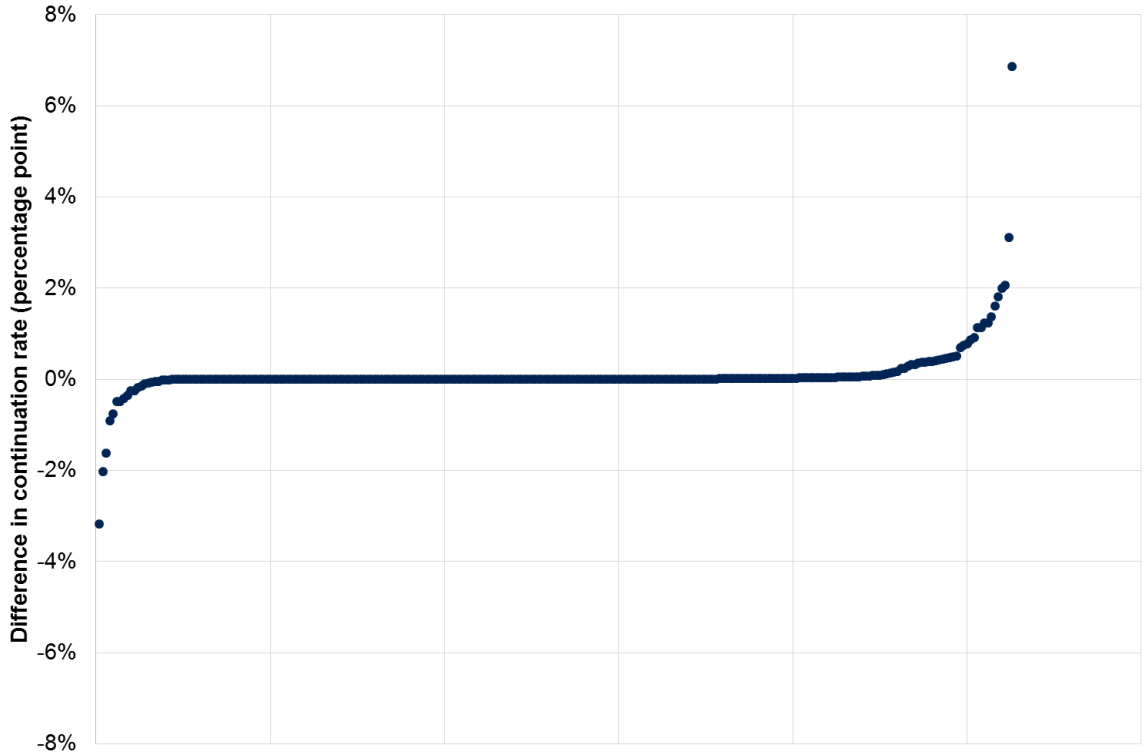
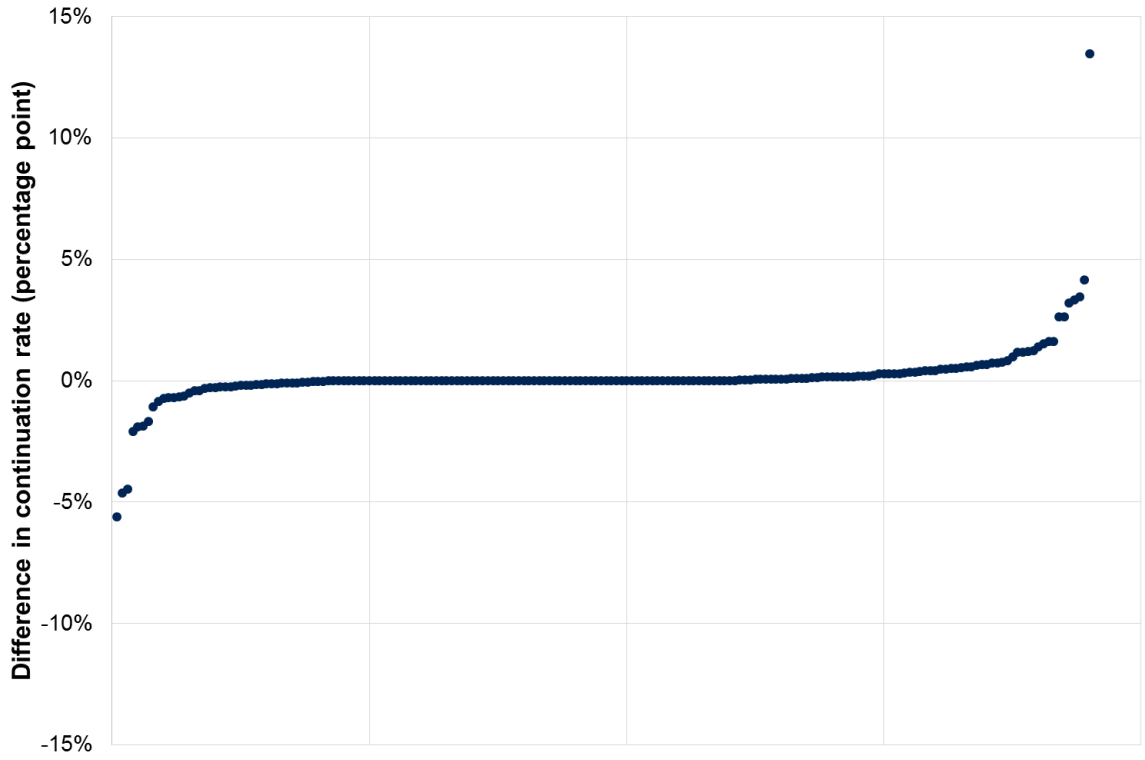


Figure 2: Difference in part-time continuation outcomes, after changes to algorithms



Other minor refinements

Population identifiers

34. The 2020 core algorithms contain a number of population identifier variables that underpin the identification of students who are and are not included in the calculation of institutional performance measures. The OfS's work to review the definitions and algorithms applied in the production of institutional performance measures has identified opportunities to rationalise and restructure these identifier variables. While it is anticipated that this rationalisation will simplify use, and improve transparency and quality assurance, of the data, we note that the changes are primarily in presentation: the changes have no impact on the populations identified.

35. In support of the aforementioned rationalisation, variables that only serve to exclude students from the cohorts measured in institutional performance measures have been discontinued in the 2020 core algorithms. These include CONEXCL512 and EMPEXCL16, both of which excluded students who did not count in the relevant indicators because they did not fall within a relevant base year for the given set of published outputs.

Travel to work area of study location

36. The definition of the variable IPSTUDYTTWA has been refined so that the travel to work area of the student's study location is calculated regardless of the student's domicile. Previously it had only been calculated for UK-domiciled students.

Ethnicity

37. The definition of the variable IPETHNIC has been revised with regard to the assignment of the Gypsy or Traveller group. In incorporating an additional variable providing a more detailed 18-category breakdown of ethnicity, we have reviewed the assignment of this group with reference

to both best practice for reporting of ethnicity categories (including consideration of best practice from Advance HE's equality, diversity and inclusion work, and harmonisation guidance from the Government Statistical Service) and the OfS policy objectives with regard to access and participation.

38. Best practice with the harmonisation principles tell us that the Roma group should ideally be reported distinctly as a separate category, and the Gypsy and Traveller group should be included as a subset of the 'White' group. This is not possible without changes to the HESA and ILR data collections, which currently collect a single category of 'Gypsy, Traveller', meaning that appropriate attribution to either 'White' or 'Other' as the broad ethnicity category is not clear-cut. It is not clear where Roma students are identifying themselves in the data returned in the HESA and ILR categories (because no 'Roma' category is currently available to them), but in any eventuality we anticipate it to be very small given the overall size of the population and the size of the 'Gypsy, Traveller' population in UK higher education.

39. As an access and participation target group, it is considered that the inclusion of any of the Roma, Gypsy and Traveller groups within the 'White' category is unhelpful as it entirely masks them and their outcomes from view. Notwithstanding the best practice, it has been considered important to provide some differentiation from the 'White' group to help with access and participation priorities. Given the size of the group, especially within higher education, the movement of this group from the White to the Other broad category has only a marginal impact on institutional performance measures.

Addition of experimental variables

40. A number of additional variables are defined within the 2020 core algorithms. In each of the following cases, the variables should be considered experimental with respect to the role that they play in institutional performance measures produced by the OfS. We have no reason to doubt that the algorithms and mappings used to create these variables are robust. Rather, we note that they will be being used for the first time in 2020 institutional performance measures. This experimental classification signals only that the algorithms are newly developed and undergoing evaluation. The OfS will continue to evaluate the new algorithms and providers are invited to submit any feedback to providermetrics@officeforstudents.org.uk.

41. The variables introduced in the 2020 core algorithms on an experimental basis are as follows:

    a. IPETHNICDETAIL provides a more detailed ethnicity breakdown, and supplements the variable IPETHNIC. Information at the lower level of granularity included within IPETHNICDETAIL is used by the OfS in the definition of the associations between characteristics of students outcome groups[15].

    b. IPIMDEXPERIMENTAL provides the Index of Multiple Deprivation (IMD) quintile for each student based on the 2019 version of the English and Welsh IMDs.

    c. IPAWARDBOD defines the awarding body for the qualification.

[15] See www.officeforstudents.org.uk/data-and-analysis/associations-between-characteristics-of-students/.

d. A new series of variables is used to derive a student's qualifications held on entry to higher education.

Summary lists

42. As a reference guide, the following lists summarise all of the variable changes included in the 2020 core algorithms. This list is also available in tabular form[16].

Renamed variables

- Any variables suffixed with a year reference and relevant to calculation of continuation outcomes have been renamed to end with either _Y1, _Y2 or _Y3. These variables are defined consistently with the un-suffixed version of the variable, but calculated on the basis of data from either one year (_Y1), two years (_Y2) or three years (_Y3) after the base year. For example, HUSID_Y1 indicates the HUSID of that student in the year following the base year.

- AGEBAND31AUG has been renamed to IPAGEBAND.

- EMPINDHIGHPGNUM has been renamed to IPHSEMPPGSTUDY.

Additional variables

- IPAGE18 (identifies whether a student is aged 18 in the year they commence their studies)

- IPHEENTITYFLAG (identifies ILR records for which we expect completion of the 'HE learner delivery' entity).

- IPBASEQUALPOP (identifies whether the record is in scope for institutional performance measures focussed on qualifying cohorts).

- IPUGBASEQUALPOP (identifies whether the record is in scope for institutional performance measures focused on qualifying cohorts of undergraduates only).

- IPLEVELNUM (for student records drawn from the HESA student and HESA AP student returns, assigns the academic level of the qualification aim, based on the Framework for Higher Education Qualifications of UK degree-awarding bodies[17]).

- IPAWARDLEVELNUM (analogous to IPLEVELNUM with respect to the qualification awarded).

- IPAWARDBOD (identifies the awarding body for the qualification).

- IPETHNICDETAIL (provides a more granular breakdown of a student's ethnicity).

[16] See 'Technical algorithms for IP measures: 2020 algorithm status' available as an Excel download at www.officeforstudents.org.uk/data-and-analysis/institutional-performance-measures/technical-documentation/

[17] See www.qaa.ac.uk/quality-code/qualifications-and-credit-frameworks.

- IPIMDEXPERIMENTAL (provides the Index of Multiple Deprivation quintile for each student based on the 2019 version of the English and Welsh IMDs).

- IPCONINDFULL_Y1, IPCONINDFULL_Y2 and IPCONINDFULL_Y3 (describe continuation outcomes at census points one, two and three years following entry respectively).

- IPCONEXCL_Y1, IPCONEXCL_Y2 and IPCONEXCL_Y3 (identifies whether the record is excluded from continuation indicators when calculating continuation outcomes one, two and three years following entry respectively).

- IPCONEXCL256_Y1, IPCONEXCL256_Y2 and IPCONEXCL256_Y3 (identifies one record to be retained in the calculation of continuation outcomes for each student per provider, mode and level of study, one, two and three years following entry respectively).

- IPCONLEVEL_E, IPCONLEVEL_H and IPCONLEVEL_L (identifies which levels of study being pursued at the census date for continuation outcomes, or qualifications obtained by that census date, would be considered as at an equivalent (_E), higher (_H) or lower (_L) level than that studied in the year of entry to higher education respectively).

- IPCONACTIVE_Y1, IPCONACTIVE_Y2 and IPCONACTIVE_Y3 (identifies whether the student was categorised as active, when calculating continuation outcomes one, two and three years following entry respectively).

- IPCONCENSUS_Y1, IPCONCENSUS_Y2 and IPCONCENSUS_Y3 (identifies the date X years and 14 days after the start of the student's instance, when calculating continuation outcomes one, two and three years following entry respectively).

- IPCONVALIDMODE_Y1, IPCONVALIDMODE_Y2 and IPCONVALIDMODE_Y3 (identifies the modes of study considered as an 'active' mode based on a student's level of study, when calculating continuation outcomes one, two and three years following entry respectively).

- IPFPE_Y1, IPFPE_Y2 and IPFPE_Y3 (identifies the full-person equivalence (FPE) of a linked record when calculating continuation outcomes one, two and three years following entry respectively).

- IPCONINDFULLSBJ_Y1, IPCONINDFULLSBJ_Y2 and IPCONINDFULLSBJ_Y3 (describe continuation outcomes within the same subject of study at census points one, two and three years following entry respectively).

- IPCONEXCLSBJ_Y1, IPCONEXCLSBJ_Y2 and IPCONEXCLSBJ_Y3 (identifies whether the record is excluded from continuation indicators when calculating continuation outcomes at subject level one, two and three years following entry respectively).

- IPCONEXCL256SBJ_Y1, IPCONEXCL256SBJ_Y2 and IPCONEXCL256SBJ_Y3 (identifies one record to be retained in the calculation of continuation outcomes for each student per provider, mode, level and subject of study, one, two and three years following entry respectively).

- IPJACS_Y1, IPJACS_Y2 and IPJACS_Y3 (identifies the full four-digit JACS code of a linked record when calculating continuation outcomes one, two and three years following entry respectively).
- IPSBJ_CAH2_Y1, IPSBJ_CAH2_Y2 and IPSBJ_CAH2_Y3 (identifies the CAH2 code of a linked record when calculating continuation outcomes one, two and three years following entry respectively).
- IPSBJ_CAH2 and IPINTSBJ_CAH2 (respectively describe the subject area in the base year and the subject area from the previous year in which a student has intercalated, defined using the most recent version of CAH2).
- IPUKFLAG (simplifies user understanding and navigation of algorithm structures).
- IPRECID (simplifies user understanding and navigation of algorithm structures).
- IPLOCATION (simplifies user understanding and navigation of algorithm structures).
- Variables related to the classification of entry qualifications.
- Variables related to changes to accommodate missing FTE (STULOAD) information in ILR returns – for further information see paragraphs 15 to 17.

Removed variables

- AIMTYPE (superseded by new categories within IPLEVEL).
- LEVEL_DETAIL (replaced and enhanced by IPOFSQAIM).
- AWARDTYPE (superseded by new categories within IPAWARDLEVEL).
- CONEXCL512 (excluded students who did not count in the relevant indicators because they did not fall within a relevant base year for the given set of published outputs)
- EMPEXCL16 (excluded students who did not count in the relevant indicators because they did not fall within a relevant base year for the given set of published outputs).
- CONTEXTAGE (extraneous to IPAGEBAND).
- SBJ_OFSCAH2 (superseded by IPSBJ_CAH2).

Refined variables

- IPOFSQAIM (replaces HEFQAIM and extends to student records drawn from the HESA student and AP student returns as well as the ILR) – for further information see paragraphs 18 to 24.

- IPLEVEL (additional categories of PGCREDIT, PGUNSPEC, UGCREDIT, UGUNSPEC, FE and NA included, the OTH category has been removed)[18].

- IPAWARD_DETAIL (enhanced consistently with IPLEVEL_DETAIL and IPOFSQAIM).

- IPAWARDLEVEL (replaces B3MONEMPLEVEL and includes additional categories of PGCREDIT, UGCREDIT, FE and NA, analogous to IPLEVEL with respect to qualification awarded).

- IPINTFPE (now defined with reference to CAH2 subject areas).

- IPMODE, IPQUALIFIER, IPXPSR, IPCONTEXTPOP, IPACCEXCL8, IPCONEXCL16, IPEMPEXCL4, IPDODEGCLASS and IPHSEMPHLSTUDY (all refined to accommodate the modifications to variables categorising levels of qualification aims and qualifications awarded).

- IPETHNIC (now defined with reference to the additional IPETHNICDETAIL variable and refined with respect to the categorisation of students identified in the 'Gypsy, Traveller' group) – for further information see paragraphs 37 to 39.

- IPSTUDYTTWA (refined so that the travel to work area of the student's study location is calculated regardless of the student's domicile) – for further information see paragraph 36.

[18] The changes to IPLEVEL impact on a number of related and dependent variables, including those used to calculate continuation outcomes, IPMODE and IPXPSR (for student records drawn from the ILR), IPDODEGCLASS, DFAPAPPEXCL, IPCONQUAL, IPQUALIFIER, IPEMPEXCL4, IPHSEMPHLSTUDY.

List of abbreviations

AP	Alternative Provider
CAH	Common Aggregation Hierarchy
ESFA	Education and Skills Funding Agency
FPE	Full-person equivalence
FTE	Full-time equivalence
HESA	Higher Education Statistics Agency
ILR	Individualised Learner Record
IMD	Index of Multiple Deprivation
JACS	Joint Academic Coding System
OfS	Office for Students
PGCE	Post Graduate Certificate in Education
TEF	Teaching Excellence and Student Outcomes Framework