An Institutional Research Approach to Access and Equity Analysis

Ian R Dobson, Monash University

Introduction

This paper considers aspects of the use of an 'institutional research' methodology which could be used for enquiry into Australia's designated equity groups. To some extent, it provides a recipe for using data files dispatched by universities to the Australian government, focussing on the important subject of planning for access and equity. It also looks at aspects of the background to current equity policy and some of the issues institutional researchers need to consider when analysing and interpreting equity policy.

As noted elsewhere in this *Primer*, institutional research has been characterised as incorporating several basic supportive activities, which, as summarised by Maasen and Sharma, include collecting data on institutional performance and the environment, analysing and interpreting the data collected, and transforming the results of the data into information for decision support (cited in Maasen & Sharma, 1991). The first stages of quantitative research into access and equity issues can be undertaken by using such a methodology.

Background on higher education equity policy in Australia and institutional research

The objectives of the Government of the late 1980s in expanding the Australian higher education system included not only a desire to create a better educated and more highly skilled population, but also to improve equity in Australian higher education. The Government's White Paper on higher education (Dawkins, 1988) considered improved access to be vital, but it was noted that the benefits of higher education had been enjoyed disproportionately by the more privileged members of society in the past. The White Paper said that the Government was committed to improving access to, and success in, the higher education system, as goals critical to Australia's ability to realise the full potential of all its citizens.

An improvement in access for those groups previously excluded from higher education was seen as being heavily dependent on growth in the system. However, growth alone was not seen as providing the complete answer. Strategies it was believed would achieve specific access and participation goals were identified and set in place. The White Paper, in which the Government stated that it '....is committed to the development of a more equitable higher education system with improved opportunities and outcomes for all Australians' (Dawkins, 1988:53), led to the National Board of Employment, Education and Training being asked to prepare a national 'overview' of equity issues in education. The outcome of this overview was a discussion paper *A Fair Chance for All — higher education that's in everyone's reach* (DEET, 1990).

The discussion paper was thorough. It defined the overall national equity objective for higher education; it set national equity objectives and targets for specific groups; it presented a range of strategies for each allegedly disadvantaged group to assist institutional planning; and it set out the responsibilities of both the Commonwealth and the universities in achieving national equity objectives. (DEET, 1990:1).

Paraphrasing the White Paper, *A Fair Chance for All* noted that '[T]he overall objective for equity in higher education is to ensure that Australians from all groups in society have the opportunity to participate successfully in higher education. This will be achieved by changing the balance of the student population to reflect more closely the composition of society as a whole'. (DEET, 1990, p8). *A Fair Chance for All* went on to identify six 'disadvantaged' groups:

- Aboriginal & Torres Strait Islander students
- Students from low socio-economic status backgrounds
- Students from rural and isolated areas
- Students with disabilities
- Students from non-English speaking backgrounds
- Women, particularly in non-traditional courses and postgraduate study. (DEET, 1990:10)

These categories are considered later, in the context of the quantitative data available from university statistics collections. It should be noted that 'equity and access' policies relate only to domestic students; overseas students are not included in calculations. All universities are now required to have specific policies and procedures about the admission of students from the equity groups, and many had such policies in place before officially being required to do so.

University statistics — an introduction and summary

In the late 1980s, the Commonwealth Tertiary Education Commission (CTEC) introduced a uniform data collection methodology for both halves of the then binary system of universities and colleges of advanced education. To the current day, universities continue to be required to use the methodology to supply information to the government by way of a series of unit record files. The system has been amended, and the collection software upgraded several times since the first collections were made. CTEC was in effect abolished in 1987, and its role so far as statistics collecting was concerned was absorbed into the Department of Employment, Education and Training (DEET).[1] Currently it is known as the Department of Education, Science and Training, and hereafter the current acronym DEST has been used whenever it was necessary to identify the Department. Data collection and provision is compulsory, and its requirement is supported by provisions in the Higher Education Funding Act (HEFA). Although there have been many changes to the collection, in essence it is little different to the system set in place in the late 1980s.

Universities provide DEST with a range of student-based unit record files. These are:

The *student enrolment file*, in which is recorded (for each student in a given course), information on course of study and a range of personal attributes, some provided by the student (such as sex, date of birth, etc.) and others provided by the institution about each student (such as 'basis for admission').

The *student load file*, in which is recorded information on the subjects the student is enrolled in, including the discipline of those subjects, and the relative proportion of a year's work the subject represents.

The *HECS liability status file* contains many elements included in the enrolment and student load files, but also information on the student's 'fee' status (eg HECS liability or exemption from HECS, for instance, by virtue of being a fee-paying student, or holding a HECS exemption scholarship).

The *past course completions file* also contains elements contained in the student enrolment file, but a student will be recorded on this file only when she or he has completed their course of study.

One data element reported on each student recorded on all of these files is a unique identifying number (ie Student ID), which means that institutional researchers within an organisation with access to the files can match students between files.

Universities also provide DEST with two files which provide information on courses and teaching departments. The *course file* provides the link between courses and fields of education. The *academic organisational unit (AOU) file* provides information on teaching departments, for aggregation purposes.

These files provide researchers with a rich source of highly accurate data on their own institution, but an institutional researcher interested in conducting system-wide research can utilise aggregated data sets,

[1] The Australian 'Department of Education' has had a variety of names: Department of Employment, Education and Training (DEET); Department of Employment, Education, Training and Youth Affairs (DEETYA); Department of Education, Training and Youth Affairs (DETYA); and Department of Education, Science and Training (DEST).

available from DEST. Several 'standard' files can be down loaded from the DEST web site, but much finer analysis can be undertaken by ordering (for a price) user-specified data sets. These aggregated files of higher education data provide the opportunity for analysts to permutate and combine a range of data elements to describe many aspects of the student body and as well as to test hypotheses about the student body.

The higher education system's data integrity relies on universities adhering strictly to the definitions contained in a set of data element dictionaries, which provide assistance in understanding the scope of what has to be collected. Data elements defined for the student collection include student-related information collected from students themselves as part of the enrolment process. These include sex, date of birth, permanent and semester residence information (collected in the form of postcodes or overseas country codes), previous scholastic background information, and background information on country of birth, year of arrival in Australia, language spoken at home, and/or indigeneity.

Another set of data elements on students is generated by universities. These data elements include students' basis of admission to their course, mode of attendance (internal, external or multi-modal), attendance type (full time or part time), and their liability for, or exemption from, paying Higher Education Contribution Scheme (HECS) fees.

In addition, researchers can derive further information from the material universities supply, for instance, by linking postcodes into indicators of location (Rural, Isolated or Urban), and socioeconomic status (High, Middle or Low). Although postcode-related information has been a part of the standard material provided by DEST in recent years, it must be used with care: Australia Post, the 'owner' of postcodes, adds and removes postcodes on a regular basis, and DEST concordance tables between postcodes and SES or rural/urban values are often out of date. Linking these characteristics to postcodes is based on information collected in the Australian Bureau of Statistics' quinquennial Census of Population & Housing.

A simple but fundamental change to aspects of the system was introduced for the collection in 2001. Prior to 2001, 'courses' (such as B.A. or B.Sc.) were classified into 'fields of study', a six-digit field which allowed for specific classification. 'Subjects', being components of 'courses' (eg French 1 or Chemistry 2A), were classified according to their discipline. The field of study classification was not the same as the discipline classification. The set of changes from 2001 involved the creation of a classification of 'fields of education' and 'disciplines' which were drawn from a common classification. The new classification allowed for much more specific classification of courses and subjects than had previously been possible.

Much as this new arrangement allows for new fields of education/disciplines to be included in classifications, it also means that temporal comparisons are much more difficult. For example, while reasonably accurate concordance tables have been constructed to map courses from the old 'field of study' classification to the new 'field of education', it is nearly impossible to map subjects from 'discipline' to 'field of education' in historical analyses. Moreover, the new classification really did nothing to alleviate the problem of classifying generic courses (eg BA, BSc, MBA) which subsume a large number of (untagged) majors. It is still difficult for DEST and institutional researchers using national data sets to identify precisely what people are studying.

University statistics and institutional research into equity categories

Perhaps the simplest analysis which can be undertaken from DEST university statistics is an enumeration of students in each equity category, and often this will be all that an institutional researcher will be seeking to do. However, one ought also be mindful of other issues driven by equity policy and some of the issues are considered in the section below. The designated equity categories can be identified in DEST data sets according to one or more data elements. Although all these categories of students self identify to a point, in only three of the cases it is via a direct question in the statistical questionnaire which forms part of the university enrolment process:

- Aboriginal and Torres Strait Islander (ATSI) students (the information is recorded in Data Element 316 "Aboriginal & Torres Strait Islander code"). In years prior to 2000, each student could indicate that she

or he was (or was not) an ATSI student. Now students can indicate in more detail whether they identify as being Aboriginal, OR a Torres Strait Islander, OR both (OR none of the above).

- Students with disabilities (the information is recorded in Data Element 386 "disability code": This information was not formally collected until the late 1990s). This data element is an eight-character field, and students identify that they have a disability by recording '1' as the first character. Characters two to seven are used to identify particular disabilities, and the eighth character is used by students if they wish to receive advice on support services.

- Women (the information is recorded in Data Element 315, "sex code"). Although the original purview of this category was all women, in practice the equity group so far as it relates to undergraduates is restricted to certain fields of study/education. Policy makers noted women's under representation in various 'non-traditional' fields of education. As noted above, universities link courses to fields of education via the Course File. From 2001, 'non-traditional' areas include any courses coded to broad fields of education Natural & Physical Sciences; Information Technology; Engineering & Related Technologies; Architecture & Building; Agriculture; Management & Commerce, and the narrow field of education Economics & Econometrics (DEST 2002: 233)

In the case of the other three categories, the information is *derived* from student responses to the enrolment questionnaire, but students do not specifically allocate themselves to these equity categories:

- Non-English speaking background students identify that they normally speak a language other than English at home (information recorded in Data Element 348, "language spoken at home"), but this alone is insufficient for a student to be recognised by official policy as 'non-English speaking': The student must also have been born outside Australia (Data Element 346, "country of birth"), AND must have arrived in Australia within the past ten years (Data Element 347 "year of arrival"). Second generation students from non-English speaking backgrounds are therefore excluded definitionally from this equity category.

- Students from low socio-economic status backgrounds and students from rural and isolated areas are identified according to the postcode of their permanent residence (Data Element 320 "permanent home residence"). So far as students from low SES backgrounds are concerned, the EdOcc index of postcodes is used. Following the Australian Bureau of Statistics' analysis of education levels and occupation categories, each postcode is designated as 'high', 'middle' or 'low' socio-economic status, with these groups containing 25%, 50% and 25% of the population, respectively (Martin, 1994:132). For identifying rural or isolated students, postcodes are split into urban, rural or isolated, based on population and/or proximity attributes of each postcode. This categorisation was originally based one devised by the Department of Primary Industry & Energy (Martin, 1994:101).

Using the data elements outline above permits analysis of overall numbers of students in each equity group, but more information may be required. It will often be necessary to analyse students in each of these equity categories against other variables. It is interesting to compare the members of designated equity groups according to attributes such as gender, age, country of birth, language spoken at home, basis of entry to university or field of education. For example, such an analysis will reveal many more female than male ATSI students, and also a preponderance of ATSI students in education, humanities or health courses.

Issues in institutional research into equity categories

For institutional researchers, the major concerns include the adequacy of the groups in the first place, and the definitions associated with identifying students included in each group. Another set of concerns relate to the effects of overlap between categories, and therefore the nature of 'disadvantage' to those students represented in more than one category. Does a student with multiple equity group characteristics suffer multiple disadvantage, or do particular equity characteristics override others? For example, does a female engineering student defined as being non-English speaking background AND who comes from a low SES, rural area carry disadvantage on account of gender alone, or is the influence of rurality and/or low

socio-economic status also important? Are any of the characteristics irrelevant so far as disadvantage is concerned? This hypothetical student has four points of 'disadvantage', according to formal equity policy.

A study undertaken under the auspices of the Australian Vice-Chancellors' Committee (published in 1997) identified from among a total of nearly 32,000 commencing undergraduates enumerated in at least one equity category, 15 students with four equity characteristics. (Dobson *et al*, 1997:37). It is virtually impossible for any student to bear all six equity characteristics, because few (if any) students identifying as being non-English speakers of Aboriginal or Torres Strait Islander descent will have been born overseas, automatically ruling them out of the non-English speaking background category (which requires birth overseas and arrival in Australia within the past ten years).

DEST statistics can be used to identify the number of students in each equity category by filtering according to the data elements identified above. This produces useful statistics, because it allows institutional researchers and those responsible for observing equity policy at each university to test for under representation. Arguably, the enumeration of students with specific equity (or other) characteristics should only be of concern if there is an under representation relative to the population as a whole. This raises the issue of establishing an appropriate denominator. For example, a national denominator figure for ATSI people is not valid, because ATSI people are not equally present in each state/territory. Analysis therefore on representation of ATSI students must take State/Territory population into account.

Age is critical in establishing relative cohort presence. To establish whether or not a certain equity category is under represented is dependent on matching university students with the population overall of university age. In 2001, 53% of bachelor students in Australian universities were aged 19 or under, with another 24% being aged 20 to 24 years. Only about 4% of bachelor degree students were aged over 39 years. Unfortunately, Martin's work, which went along way to operationalising equity analysis from DEST statistics, used the 15-64 years age group when establishing denominators from which to measure representation in universities (For example, Martin, 1994:76). It is dearly not valid to use people aged 15-64 years of age as a denominator because these ages in no way reflect the normal age range of university students. However, institutional researchers can surmount this problem by obtaining appropriate age cohort figures from Australian Bureau of Statistics publications.

Other measurement problems are more persistent and difficult to solve, notably those concerning socioeconomic status and rurality. Martin (1994) and Western *et al* (1998) both have extensive discussion on difficulties with these measures. The most recent consideration of these matters was undertaken by Jones (2001).

Enumeration alone is not enough, and Martin (1994) recognised this in the development of *Equity and General Performance Indicators.* In her study, indicators of access, participation, success and apparent retention were operationalised for the first time.

Enrolments of students from under represented groups can be monitored over time, to see if the presence of that group is improving. In this way, the scope of which cohorts of students should be considered to be an 'equity group' could also change over time, although this has not yet occurred under formal equity policy. It is also important to ensure that once at university, all students 'succeed', where success can be defined as (ultimately) successfully completing a course, and successfully completing subjects, on the way to completing a course. Changes in the scope of information collected can influence apparent outcomes. For example, the expansion in the specificity in higher education statistics for Aboriginal or Torres Strait Islander students was noted above. Whereas students were simply asked whether they were ATSI students or not prior to 2000, when the range of options for indigenous students was increased, some universities dearly failed to incorporate this change into their internal systems. This is evident because in aggregated data files for 2000, the number of students with 'unknown' ATSI status reported by some universities increased markedly. One result of this was speculation in the press that the number of ATSI students had declined. Given the poor response by some universities, this assertion could not really be tested. Hopefully data quality will improve. In light of the relatively low proportion of ATSI students in the total university population, this 'glitch' by some universities was an important one which has reduced researchers' capacity to undertake proper time series analysis on ATSI student presence in the university population.

One particular concern with definition relates to students from non-English speaking backgrounds. Formal policy requires that students must have been born overseas, and have arrived in Australia within the past ten years. This definition therefore overlooks second generation Australians from non-English speaking homes, and might therefore fail to include students with a specific language disadvantage. A study by Dobson *et al* (1996) showed that several second generation language groups were under represented in Australian undergraduate higher education.

Relative success by students as they progress through their course is important, and measuring 'student progress units' for equity groups and reciprocal non-equity groups is readily undertaken. Analysis by Dobson *et al* (1996: 40 ff) revealed that all equity groups but Rural & Isolated students were significantly out-performed by students NOT in the equity category, but the relative performance WITHIN equity groups was interesting. In particular, female students significantly out performed males in all categories containing males (Dobson *et al*, 1996: 53). There was also a significant difference in performance by internal (on campus) students when compared with external students (Dobson *et al*, 1996: 53).

Of course, the ultimate test of success at university is the successful completion of a degree (preferably in minimum time). If students from each designated equity group complete their courses less often or more slowly than students from the reciprocal population, it ought to be of concern to universities. System-wide analysis is difficult on this point because of the range material made available by DEST. However, institutional researchers with access to their own institution's files over a number of years should be able to test the validity of the proposition that equity group students under-perform other students. In a project funded by DEST a number of years ago, software ('Cohort') was produced to simplify the process of measuring success in terms of outcomes (Bardsley, 1991). Equity professionals and institutional researchers alike would benefit by having the type of results promised by the Cohort software.

Conclusion

The richness of the DEST statistical collections can be utilised by institutional researchers to undertake deep and varied analysis into equity groups. Both the files supplied individually by universities, and aggregated data for the system can be used for this purpose. Institutional researchers ought to be analysing data beyond simple enumeration of equity groups. Only by considered analysis can universities establish the 'drivers' in terms of disadvantage. The commentary above suggested that official policy fails to consider important aspects of 'disadvantage' but this ought not prevent institutional researchers moving beyond superficial analysis.

DEST itself now publishes tables on equity characteristics, but these are of limited use, because they include information on ALL domestic students, whereas concerns about equity and access to university properly relate to students enrolling in a first undergraduate degree. Institutional researchers using DEST data files can go much further in their analysis.

Acknowledgements

The author wishes to acknowledge the considerable assistance received from the following people: Wayne Shippley and Jan Armodt of DEST; Eric Skuja and Bob Birrell of Monash University; Raj Shanna of Swinburne University and Angel J Calderon of RMIT.

References

Bardsley, WN (1991) *Cohort: a program for monitoring student completion rates in higher education.* DEET. EIP. AGPS, Conberra.

Dawkins, JS (1988). *Higher Education — a policy statement.* Canberra. AGPS.

Department of Employment, Education & Training (DEET), 1990, *A Fair Chance for All: Higher education that's within everyone's reach.* AGPS Canberra

Department of Education, Science & Training (DEST), 2002. *Students 2001. Selected Higher Education Statistics.* DEST. Canberra.

Dobson, Ian R, Birrell, Bob, and Rapson, Virginia (1996). The Participation of Non-English Speaking Background Persons in Higher Education. *People & Place* 4:1 46-54.

Dobson, Ian R, Sharma, Raj, & Ramsay, Eleanor (1997). *Designated Equity Groups in Australian Universities: Performance of Commencing Students 1996 — Undergraduate Award Courses.* AVCC. Canberra.

Jones, Roger, (2001) *Identifying Higher Education Students from Low Socio-Economic Status Backgrounds and Regional and Remote Areas.* DEET AGPS Canberra.

Maasen, Peter & Sharma, Raj (1991) What is institutional research? In Cameron, Barry, Sharma, Raj and Ng Gan Che (Eds.) *A Primer on Institutional Research in Australasia.* AAIR.

Martin, Lin (1994). *Equity & General Performance Indicators in Higher Education.* DEET. AGPS Canberra.

Western, John, McMillan, & Durrington (1998) *Differential access to higher education: the measurement of socioeconomic status, rurality and isolation.* DEET AGPS Canberra.