

5 February 2010

Mr Jason Coutts
Branch Manager
Equity, Performance and Indigenous Branch
Department of Education, Employment and Workplace Relations
GPO Box 9880
Canberra City ACT 2601

Dear Jason

Go8 Response to the *Measuring the Socio-Economic Status of Higher Education Students Discussion Paper, 2009*

The Group of Eight's (Go8) detailed comments the *Measuring the Socio-Economic Status of Higher Education Students Discussion Paper* are attached. Thank you for the opportunity for input. Comments here overlap to some extent with our comments on related issues in response to the Higher Education Participation and Partnership Program Guidelines consultations and the Performance Funding Indicator Framework Discussion Paper. The Go8 would be happy to discuss these issues further with you.

The Go8 welcomes initiatives which provide increased access to higher education for people from less advantaged backgrounds, which include support mechanisms to assist them to succeed in their courses.

In summary, the Go8 supports a move to measure the real socioeconomic characteristics and background of students by exploring additional, relevant, directly measured variables (subject to validation of quality), rather than relying on the present use of imputed values, based on postcode areas. Go8 encourages the development of improved data sets, to better understand the interaction of factors which lead to success in higher education, and to support data-based evaluation of policy initiatives and programs.

The Go8 supports analysis at Population Census Collector's District (CD) level, rather than postcode (recognising that CDs are generally more homogeneous and smaller, and more stable). However, the fundamental problem remains that relatively advantaged students coming from particular areas will be counted as low SES because of the disadvantages of others who live in the same geographical area. The government's low SES initiatives will be measured in terms of whether more students come from those and similar areas, and not in terms of whether more truly disadvantaged students enrol.

The Go8 also supports the inclusion of relevant Centrelink data related to students and their parents, subject to data validation and timeliness, on the basis that those students have an evident, data-verified disadvantage in having to meet eligibility criteria.

However, the Go8 calls for a more thorough approach in the medium term. Australia can and must establish more accurate data sets, to better inform policy development and to enable initiatives to be evaluated with some degree of confidence. While establishing such a data set will cost universities and government, the issues are of such immediate and ongoing importance that the Government should commit to such investment.

The Go8 would be very happy to meet with the Department, to discuss these comments in more detail, and to work towards better data sets for analysis and as a basis for funding allocations.

Yours sincerely

[signed]

Michael Gallagher
Executive Director

Go8 Position Paper

Response to the DEEWR discussion paper: *Measuring the Socio-economic Status of Higher Education Students,* December 2009

1. The Need for Reliable Data about Socio-economic Status and Higher Education

In terms of the conceptual and measurement issues addressed in the discussion paper, it is important to recognise that there are (at least) two separate but related purposes for a data set which contains information on the socioeconomic background of higher education students.

Firstly, there is a need for a student data set which reliably identifies the actual, personal values of socioeconomic and other variables related to each higher education student, and also contains information on education status – whether each student completed successfully, measures of academic progress and outcomes, or reasons why they did not progress. (Information on education prior to entry to higher education is relevant too.) Such a data set is critical to the evidence-informed consideration of what factors facilitate or mitigate against successful entry to and completion of higher education. It is only from the analysis of such an individual student data set that barriers and successful pathways will be identified, based on outcome data.

Second, the Government has a funding allocation formula use for a data set. Socioeconomic status (SES) measurement is currently imputed from averages such as geographical (postcode) socioeconomic index values. It doesn't record the actual socioeconomic situation of actual students, but imputes statistically derived averages based on address. (This means that all students whose personal situations are at variance - in either direction - from the postcode average are not well represented in the funding formula, and the funding outcomes are less than accurate. There is no ongoing analysis of whether students purported to be low SES are in fact from low personal SES circumstances. It is arguable that persons with a personal socioeconomic advantage from lower SES areas are more likely to enter higher education than persons with personal socioeconomic disadvantage from the same area – but both are treated equivalently in the DEEWR analysis.)

These separate uses need to be acknowledged, as data deficiencies potentially impact in different ways, depending on policy options being considered.

The discussion paper appropriately recognises some of the shortcomings of the current data set based on ABS social economic indexes for areas (SEIFA), but does not fully explore the impacts and results of those shortcomings. In particular, it proposes some marginal adjustments with its proposed interim measure, which it hopes, and we agree, may improve the potential distortions in the allocation of funding. We do not believe the paper adequately explores the range of possible options for the development of an analytical data set which would support real policy analysis on these issues.

On its face we strongly support the development of an index of SES which includes multiple measures relevant to each dimension of disadvantage weighted according to their differential impact on dependent variables such as propensity for access and success at higher education. At an institutional level, if such index scores were to be reported at the individual student level, we can see great value in terms of being able to identify potential students at risk and in targeting scarce resources to improve outcomes.

2. Policy Context and Perspective

The key policy issue and purpose that underpins this discussion paper is the fostering of greater access to, and achieving of greater levels of participation and success in higher education, from people in those population subgroups in society who have not proportionately proceeded to such study. To achieve this policy objective will require change to past patterns and practices.

However, success is not simply a matter of encouraging higher education institutions to enrol more students from disadvantaged backgrounds. In part, there needs to be greater understanding of the reasons why some secondary students do not proceed to higher education.

There is, firstly, aspiration and motivation to proceed to higher education. This is an issue for schools (teachers and career advisers), for parents, and for the students themselves. The fact that no-one or few others in the family have proceeded to higher education must not be allowed to be a barrier to capable students. This issue is beyond the scope of this discussion paper, but needs to be addressed – and it is relevant to the discussion paper, given the recurring intergenerational patterns of relatively low academic attainment and socioeconomic disadvantage among some population subgroups.

Second, there is the requirement that these students receive secondary education of a quality that enables them to be admitted to and equips them to be successful in courses at the higher education level. This issue is out of scope of this paper, but needs to be recognised explicitly as a major factor in some population subgroups being under-represented in the higher education student population.

There is, third, the financial cost of proceeding to higher education. While income contingent loan (HECS HELP and FEE HELP) arrangements address some direct costs, and Austudy assists those who meet the eligibility criteria, there may be other costs of possible relocation, travelling and living costs, which fall more heavily on some types of students. As well, there may be opportunity costs of earnings foregone, or contribution to a small family business because of study rather than entry to the labour force, which may be factors which act as barriers to entry to and success in higher education. These factors may apply more strongly to persons from low socioeconomic status backgrounds.

3. Impact of Socioeconomic Status on Entry to and Success in Higher Education

Before coming to the issues directly addressed by the discussion paper, it is important to note that there are two other pieces of analytical work required related to socioeconomic status, to inform government and the sector and to enable effective responses to be developed to policy objectives and recommendations.

These are:

- I. An analysis of attrition – to answer the question whether socioeconomic status is a contributing factor to failure or withdrawal of students who commence but do not complete higher education. This analysis is needed in respect of true socioeconomic status, and not an imputed area average. (Better analysis of transition generally is needed, to know whether students now counted as attrition are in fact moving to other tertiary institutions, and whether a proportion of them are returning to higher education over time, and completing.)
- II. An analysis of the socioeconomic and other characteristics and reasons of those completing secondary students who do not proceed to higher education – particularly

those who would have qualified (or almost qualified) to enter. (Again, the analysis is needed in respect of true socioeconomic status, although an imputed area average SES would begin to identify potential problem areas needing closer analysis of the impact of specific socioeconomic, educational and other factors.)

4. *Analytical data*

We return to the issue of the need for a high quality analytical data set, which contains a wide range of independent variables, or factors considered to impact on success in higher education. It needs also to have a dependent variable, such as:

'Propensity to be admitted to and to successfully complete a higher education course/qualification'.

A statistical relationship would then be derived from the data, describing in quantitative terms the relative strength or contribution of those various factors in leading to success in higher education: in technical terms, a multivariate functional equation with independent variables, and weights, derived from a systematic and thorough analysis of relevant data, of acceptable quality. There are standard statistical methodologies for this analysis. Such information would include the various factors which may increase or inhibit success in higher education (and some of these are likely to be correlated).

We need a contemporary analysis of the factors which influence participation and success in higher education.

It is disappointing that the discussion paper presents no analysis from data held by the Department, or from other sources, of those social and economic factors which contribute to success or failure in pursuing higher education. The analytical reference list is relatively thin and dated, and contains no international literature. There is no discussion of success (completions) and attrition rates, based on the data available to the Department or other studies.

For example, a Canadian Study published in 2005 on impact of family background on access to postsecondary education quantified the post-secondary education influence of education level of parents, and also family type (comparing two parent and single-mother families), living in a rural area, and academic performance in high school. See: <http://www.statcan.gc.ca/daily-quotidien/050118/dq050118c-eng.htm>

It would be most helpful to developing incentive programs and funding models related to this issue if DEEWR were both to look at other countries' data, and to analyse Australian data to evaluate whether similar experience applies.

As a priority, the Department needs to use and add to existing data to explore these issues, and to strengthen the data set to support more extensive analysis than is likely to be possible at present.

There is considerable capacity for understanding the motivations and success factors impacting on students in the Australian context by our universities. The Go8 would be happy to facilitate greater sharing of information with the Department and to provide a link to the personnel within our universities that can assist.

5. Improving Data to Support Funding

The Discussion Paper proposes geocoding of the home address of students, to be able to impute averages from smaller areas, such as Census Collectors Districts (CDs). Subject to cost-effectiveness, this is supported, on the basis that CDs are more homogeneous and stable than postcodes. However, this does not remove the essential deficiency of imputing an average which may or may not reflect the circumstances of the individual student. The discussion paper acknowledges this.

The Department also proposes to include Centrelink data on numbers of students receiving means tested study related income support allowances and supplements, as a proxy for numbers of students of low income backgrounds at each institution. We support this. There is no question that such students are appropriately counted as from low SES background – but there are many others of low SES who are not directly counted by this mechanism.

DEEWR needs to continue to consider ways of counting such persons, directly.

Another potential source referred to is ATO parental income data. Setting aside the questions of privacy and of access, definitional and data quality issues would need to be explored, including the aggregation of individual data to household income, and developing a formula to take account of, for example, numbers of dependents.

6. The Present Situation – Data Deficiencies

At present, there is a student data set with significant conceptual, definitional and measurement deficiencies, which nonetheless contains some limited information on student backgrounds. In particular, it has student home address postcode. The priority should be to establish a data set of higher quality and integrity, both to better understand the factors which lead to higher education success, and to allocate funds on a demonstrated evidentiary basis which will encourage the highest levels of success

Using Australian Bureau of Statistics Census of Population and Housing data and multivariate statistical analytical techniques, various summary statistical measures are derived. From the most recent (2006) Census, the following four Socio-Economic Indexes for Areas (SEIFA) have been produced:

- **Index of Relative Socio-economic Disadvantage:** is derived from Census variables related to disadvantage, such as low income, low educational attainment, unemployment, and dwellings without motor vehicles.
- **Index of Relative Socio-economic Advantage and Disadvantage:** a continuum of advantage (high values) to disadvantage (low values) which is derived from Census variables related to both advantage and disadvantage, like household with low income and people with a tertiary education.
- **Index of Economic Resources:** focuses on Census variables like the income, housing expenditure and assets of households.
- **Index of Education and Occupation:** includes Census variables relating to the educational and occupational characteristics of communities, like the proportion of people with a higher qualification or those employed in a skilled occupation.

The ABS notes that:

*“The concept of relative socio-economic disadvantage is neither simple, nor well defined. SEIFA uses a broad definition of relative socio-economic disadvantage in terms people's access to material and social resources, and their ability to participate in society. While SEIFA represents an average of all people living in an area, **SEIFA does not represent the individual situation of each person.** Larger areas are more likely to have greater diversity of people and households.”*

(See: ABS Cat No 2039.0 - *Information Paper: An Introduction to Socio-Economic Indexes for Areas (SEIFA)*, 2006 (emphasis added))

DEEWR attributes the value of the fourth of these indexes, that is, average community educational attainment and average proportions in skilled occupations, for the student's postcode, irrespective of the actual parental educational attainment background, parental occupation, income, housing, country of birth, languages spoken at home, availability of education enablers like household internet connection, etc.

It has long been recognised that attributing an average derived from a geographical area developed for delivery of postal articles, and not based on statistical homogeneity, has serious flaws. Some students are presently being attributed to low SES, who do not come from a low SES family background; others who are from a relatively disadvantaged background are being rated as higher SES.

The discussion paper discusses at length the deficiencies of a statistical product which was not designed specifically for the purposes for which DEEWR has used it, without addressing the biases that result from that use, or fundamental policy or statistical issues, or the 'policy levers' to increase the participation and success in higher education.

The proposal in the discussion paper to collect information on each parent's educational attainment directly from students in the 2010 higher education student's collection represents an attempt to have actual student data, rather than postcode averages. However, there are practical concerns about discussion in the DEEWR paper about the potential for universities to collect data at enrolment or through surveys such as parental income and schools attended. Students do not necessarily have access to this information, parents may be reluctant to provide it and there are privacy issues for the University in collecting such information. Our preference would be for these data to be accessed through Centrelink or the ATO. On the other hand, as a first step, we support the collection at enrolment of information about parental educational levels from 2010, but this should be done by category choice rather than by free text.

The discussion paper raises the possibility that parental occupation and income be collected. These would each require evaluation as to precise definition, and ability to capture them accurately. However, Go8 would support exploring availability of such data items.

There are other variables which might be considered for direct measurement, but it is questionable whether adding them to the existing, postcode-dominated data set would be of analytical value. More fundamental changes to the data collection are required.

It might have been expected that the discussion paper would have given more attention to the possible analysis of data from the Census of Population and Housing in different ways than at present, to explore relationships between some of the education, occupation, income and other variables, free from the distortions introduced by the SEIFA postcode averages.

Additionally there is the possibility of building on the Census data set. The ABS is working on an enhanced Census data set whereby it combines administrative by-product data with census data

from the 2006, 2011 and subsequent censuses. Under the heading **Benefits of the Statistical Longitudinal Census Dataset**, the ABS states

“The richness of the census data lies in the information it provides on topics such as

- *family structure;*
- *education and qualifications;*
- *work, including hours worked, occupation and industry;*
- *income; and*
- *housing.*

Patterns in individual experiences over time in these areas, and factors that might influence these experiences, would be those where the Statistical Longitudinal Census Dataset would provide most insight. Examples include studies of the impact of parents’ education and labour force participation on their children’s subsequent participation, and the impact of factors such as family structures and household income on these outcomes.”

(See: ABS Cat No 2062.0 *Census Data Enhancement Project: An Update*)

The ABS intends to augment successive Census data records with administrative data – school achievement, higher education ENTER scores, and outcomes of higher education studies might well be considered suitable and of analytical value. It is, perhaps, surprising that the Discussion Paper does not even refer to this possibility.

7. Potential Impacts on VET

Few policy changes are totally isolated in their impacts. Both school leavers and others entering tertiary education make choices about the types of courses most suitable for their skill set and aptitudes, against their aspirations and competencies. While it is in no way an argument against change or changing the balance of socioeconomic status of students entering different tertiary options, it will be important for the government to undertake some thorough analysis of the potential impact on the VET sector. A significant shift to higher education of those types of students currently entering further study and training in VET might lead to critical shortages in skilled labour inputs in a range of occupations and industries.

Additionally, it would be useful to ensure consistency in low SES measurement between different levels of education. We understand that SEIFA-based measures of SES are used to varying degrees by both the Commonwealth and State Governments to determine funding allocations for programs covering childcare centres, pre-schools, primary and secondary schools (notably public and private), and TAFES. While we recognise that it may be difficult and will take time, the ideal outcome of this project would be to reach agreement across all sectors about an index of SES that could be applied uniformly, with consistency in the type of data collected, and the way they are collected and reported.

8. Questions Raised in Discussion Paper

Rather than respond to each of the questions posed in the Discussion Paper, Go8 notes that the fundamental issues to be addressed are specific measures related to individual students, rather than imputing averages, and improved data quality. When these are addressed, there is a need for rigorous analysis. Go8 encourages DEEWR to develop and undertake a work program, with ABS and other stakeholders, to address the data deficiencies and quality issues. Go8 would welcome the opportunity to work with DEEWR on these issues.