Notes on the ecological fallacy when area-based indexes of disadvantage/advantage are applied to schooling in Australia[4]

Barbara Preston

28 March 2010

The Australian Bureau of Statistics has strongly cautioned against the inappropriate use of area-based indexes of disadvantage (and/or advantage) such as the ABS's own four Socio-Economic Indexes for Areas (SEIFAs) to indicate the disadvantage (and/or advantage) of individuals living in the area. This is in recognition of the heterogeneous nature of areas as small as the ABS Census Collection District (CD) of around 250 households. Such inappropriate use involves the 'ecological fallacy'. The appendix provides detail of ABS's concerns.

Area-based indexes are used as an indicator of the relative disadvantage or advantage of Australian schools for the purposes of providing funding to nongovernment schools (the DEEWR SES measure) or for providing the basis for judging the relative performance of every Australian school (the My School Index of Community Socio-Educational Advantage, ICSEA). These indexes use the most recent ABS Census data and are similar to the SEIFAs, but with some different components and weightings judged more appropriate to education-specific disadvantage/advantage. The indexes are used to obtain a score for each CD, in the same way as SEIFA is used to obtain a score for each CD. A school is then scored according to the average score of the home CDs of every student (or a representative sample of students). Thus a very small school that drew students from just one CD would have a score directly based on the score of that CD, and one that drew equally from ten CDs would have a score directly based on the average scores of those ten CDs. While these indexes are not making assessments of the disadvantage/advantage of individuals as their final outcome, their *central mechanism* involves making such assessments of the disadvantage/advantage of individuals. Thus the cautions of ABS are relevant.

Not only is there great diversity of family SES levels within CDs – the heterogeneity that ABS refers to - but, most importantly, *the different schools attended by children in the same CD is not randomly determined.* In discussions around the Commonwealth Government's SES funding model for nongovernment schools some very obvious cases were raised. In particular, the situation where high fee metropolitan boarding schools were assessed as not very high SES because of the large number of students who were the children of high income/high wealth broad acre farmers who in live in CDs classified as relatively low SES because of the many low SES station hands and rural village residents in those CDs. The children in those lower SES families attend local schools, not high fee boarding schools. The corollary is that those local schools, not attended by the farm-owners' children, would be classified as higher SES because of those high SES farm-owners in the same CDs as the local students. Such instances, though stark and easy to explain, are not anomalies. The same pattern operates throughout Australia - to a greater or lesser degree.

---

[4] The ABS 2001 Census data used in these notes was obtained as custom tables by the author for another purpose six years ago. Similar analyses could be carried out using 2006 Census data, which is the basis of the current DEEWR SES and ICSEA indexes.

The Census does not (of course) provide data on the actual schools attended by school age students, but it does provide data on level of schooling (primary or secondary) and type of school (government, Catholic and other nongovernment/independent) in a classification, 'Type of Educational Institution Attending' (TYPP). The Census also provides data on Family Income (FINF) which can be applied to individual school students[5]. It is thus possible to analyse school students in every CD (scored by an appropriate index according to advantage/disadvantage) by FINF[6] and TYPP (the type of school attended by level and sector).

Custom 2001 Census data for the population of all secondary school students in Statistical Local Area (SLA) of Penrith in NSW was obtained (Preston 2004), with variables of FINF (divided into approximate thirds – LOW, MEDIUM and HIGH), TYPP (government, Catholic or independent secondary school), and CD. The CDs were classified according to the SEIFA index of advantage/disadvantage. Some of the findings from this dataset include:

- Across the whole Penrith SLA, a public school secondary student is *equally likely* to have a HIGH or LOW family income, while a private (Catholic or independent) school student is *more than twice as likely* to have a HIGH as a LOW family income.
- In the ten most *advantaged* CDs, a public school student is *four* times as likely to have a HIGH as a LOW family income, while a private school student is *ten* times more likely to have a HIGH as a LOW family income.
- In the ten most *disadvantaged* CDs, a public school student is *sixteen* times more likely to have a LOW as a HIGH family income, while a private school student is only *three* times more likely to have a LOW as a HIGH family income (and an independent school student is *equally likely* to have a HIGH as a LOW family income).

Thus, if drawing from just the ten most disadvantaged CDs, a representative public school would have 16 disadvantaged (LOW family income) students for every one advantaged (HIGH family income) student, while a representative independent school drawing only from the same CDs would have equal numbers of disadvantaged and advantaged students. Yet if an index that was similar to the SEIFA index of advantage/disadvantage (such as the DEEWR SES index or the ICSEA) was applied, both schools would be classified as equally disadvantaged.

The Census does not differentiate between selective/specialist and comprehensive public schools, or between high and low fee private schools. Thus the differences between, say, comprehensive public schools and high fee private schools are likely to be more pronounced than the general differences between public and private schools noted above. In addition, analysis of successive Censuses, including the 2006 Census, shows ncreasing divergences between the school sectors, with an increasing concentration of LOW family income students in public schools, and an increasing concentration of HIGH family income students in Catholic and independent schools (Preston 2007).

In conclusion, the use of area-based indexes of disadvantage/advantage are not appropriate for measuring the disadvantage/advantage of schools in Australia, especially for high stakes purposes and where there will be comparisons ('matching') between schools from different sectors or which have different formal or informal selection and exclusion practices.

---

[5] But not those at boarding school or otherwise living away from home on Census night, and some other categories.

[6] Or a family SES measure such as that developed by Baker & Adhikari (2007)

Appendix

The ABS cautions against the use of SEIFA indexes as a proxy for individual disadvantage/advantage.

The Australian Statistician, Brian Pink, points out that 'SEIFA says nothing about the socio-economic status of a particular person or household. …It is not correct to use SEIFA to say any particular person is relatively disadvantaged' (Pink 2006, p. 15).

In an ABS research paper, Pramod Adhikari develops the argument further:

> The relative disadvantage (or advantage) that SEIFA indexes summarise at CD level should … be used as contextual variables and not as indexes for individuals living in that area. If we attempt to explain individual level disadvantage from area level measure such as SEIFA then we are assuming that the relationships observed for areas hold for individuals. As we know people living in a CD are not homogeneous – everyone living in a disadvantaged CD will not all be equally disadvantaged. There will be some people who are less disadvantaged than others even in the most disadvantaged areas. The ecological fallacy is a result of the assumption that relationships observed for areas also hold for the individuals who live there. If we assign an area level index to an individual then there is a risk of an ecological fallacy. (Adhikari 2006, p.6)

The evidence for the substantial nature of the potential problem of the ecological fallacy in the use of SEIFA (or any similar area-based index) as a proxy for individual measures of disadvantage in Australia is provided in a 2007 ABS Methodological Advisory Committee research paper (Baker & Adhikari 2007) which uses Census data for Western Australia and 'explores the feasibility of creating individual and family level socio-economic indexes using the same conceptual and methodological basis as SEIFA' (p. 1). The authors conclude:

> We used the individual and family indexes to examine whether there is a high risk of an ecological fallacy if the IRSD [Index of Socio-Economic Disadvantage] is used as a proxy for individual or family level disadvantage. Our analysis found that individual and family relative socio-economic disadvantage was quite diverse within areas. This means that there is a high risk of an ecological fallacy if we use the SEIFA indexes as a measure of individual level disadvantage, rather than a measure of area level disadvantage. (p. 25)

References

Adhikari , Pramod 2006, 'Socio-Economic Indexes for Areas: Introduction, Use and Future Directions', Research Paper, Australian Bureau of Statistics, Canberra, viewed 2 February 2009, http://www.ausstats.abs.gov.au/Ausstats/subscriber.nsf/0/12AC236185F4B054CA2571F60017AE2C/$File/1351055015_sep%202006.pdf.

Baker, Joanna & Adhikari 2007, 'Socio-Economic Indexes for Individuals and Families', ABS Cat. No. 1352.0.55.086, Research Paper, ABS Methodological Advisory Committee, Australian Bureau of Statistics, Canberra, viewed 2 February 2009, http://www.ausstats.abs.gov.au/Ausstats/subscriber.nsf/0/87EA1D92F0BB437CCA25733F001F541D/$File/1352055086_jun%202007.pdf.

Pink, Brian 2006, 'An Introduction to Socio-Economic Indexes for Areas (SEIFA)', Information Paper, Australian Bureau of Statistics, Canberra, viewed 6 March 2009,

http://www.ausstats.abs.gov.au/Ausstats/subscriber.nsf/0/D729075E079F9FDECA2574170011B088/$File/20390_2006.pdf.

Preston, Barbara 2004, 'Choice and national schools policy', A paper presented at the annual conference of the Australian Association for Research in Education, Melbourne, 28 November to 2 December 2004

Preston, Barbara 2007, *The Social Make-up of Schools: Family Income, Religion, Indigenous Status, and Family Type in Government, Catholic and other Nongovernment Schools*, Australian Education Union, Melbourne, viewed 21 April 2008, http://www.aeufederal.org.au/Publications/Bprestonrep2007.pdf.