UNSW THE UNIVERSITY OF NEW SOUTH WALES

Social Policy Research Centre Report Series

Never Stand Still | Faculty of Arts and Social Sciences

Report for the NSW Child Death Review Team on Measuring Socioeconomic Status

Peter Saunders

Social Policy Research Centre,

May 2012

SPRC Report 08/12

SPRC
Social Policy Research Centre

For a full list of SPRC Publications visit:

www.sprc.unsw.edu.au

or contact:

Publications, SPRC,
Level 2, John Goodsell Building
University of New South Wales,
Sydney, NSW, 2052, Australia.

Telephone: +61 (2) 9385 7800
Fax: +61 (2) 9385 7838
Email: sprc@unsw.edu.au

ISSN: 1446-4179

ISBN: 978-0-7334-3179-1

Published: May 2012

Social Policy Research Centre, UNSW

Research team

Peter Saunders and Melissa Wong

Other agencies

Child Death Review Team, New South Wales Ombudsman

Author

Peter Saunders

Contact for follow up

Professor Peter Saunders, Social Policy Research Centre, University of New South Wales, Sydney NSW 2052, Ph: (02) 9385 7806, p.saunders@unsw.edu.au

Suggested Citation

Saunders, P (2012), 'Report for the NSW Child Death Review Team on Measuring Socioeconomic Status', Social Policy Research Centre, University of New South Wales, Sydney.

Contents

Background[*]

The NSW Child Death Review Team (CDRT) is seeking advice from a recognised expert in the field on appropriate options for measuring and reporting socioeconomic status (SES) in relation to children whose deaths are subject to review by the Team. It is also seeking advice on appropriate options for analysing and reporting child deaths in NSW on a geographic basis.

The primary purpose of considering the SES of the families of children who have died is to identify any links between child deaths and SES, or particular causes of death, so that such information can inform the development of preventative actions and strategies.

In the past, a number of methods have been used to measure and report on the SES of families whose children have died, although concern has been expressed about the utility and reliability of the different methods used, the number of indices, some aspects of their variability within areas, and their stability over time.

Problems with the approach used in the past on the geographic reporting of child deaths include the relatively small number of cases, and issues related to children who normally reside in NSW but die interstate. The latter is a particular issue for those residing in border regions, including the far North coast.

The latest (2010) NSW CDRT *Annual Report* noted (p.117) that the method used there to report on SES was subject to 'substantial uncertainty' about its ability to capture the actual socioeconomic status of families. In light of this uncertainty, the report noted that: 'Improving the measurement of socioeconomic status is a priority for future reports'.

Reporting on the geographic spread of child deaths last appeared in the 2007 *Annual Report,* and the following year's Annual Report noted (p. xxviii) that 'over the next 12 months the Team will continue to investigate currently available methods to inform the analysis of child mortality across the geographic regions of NSW.' However, this work was not progressed in the period prior to the transfer of responsibility of the CDRT from the NSW Commission for Children and Young People to the Office of the NSW Ombudsman in February 2011.

This report is in two parts. Part I considers issues associated with the measurement of socioeconomic status – in general and in relation to the specific approaches used by, and available to, the CDRT. Part II addresses issues associated with the geographic reporting of child deaths. Options for addressing both issues are presented and analysed.

PART I: MEASURING SOCIOECONOMIC STATUS

Introduction

Measures of socioeconomic status (SES) are commonly used in social science studies that seek to link the circumstances and conditions of individuals and families to the incidence of a variety of social problems. The degree of sophistication employed in these studies varies

[*] The author acknowledges the statistical support provide by Melissa Wong.

greatly, although they generally aim to establish whether the *level* of SES is closely associated with the extent of a specific problem, whether and how its incidence varies with SES (the social gradient), and whether the degree of *inequality* in SES is associated with differences in the observed outcomes of a specific variable. Particular attention often focuses on whether those identified as low SES are particularly prone to the issue under examination.

A common implication of all approaches is that the social problem under examination will only be effectively addressed if the variation in SES is reduced – either by raising the levels experienced by those most disadvantaged, or by reducing the differentials that exist between disadvantaged and advantaged in the community.

In recent years, there has been an increasing interest in the role of SES in influencing patterns of participation in a broad range of social programs, and hence as a factor affecting the distribution of social outcomes. One example is the Index of Community Socio-Educational Advantage (ICSEA) that is provided on the Australian Government's My School website (see Barnes, 2011). Another is the work being done to develop an indicator of the SES background of tertiary students as part of efforts to increase the proportion of students from a low SES background to 20 per cent by 2020 following the recommendations of the Bradley Review (Commonwealth of Australia, 2008).

In both cases, the increased emphasis given to identifying the SES backgrounds of students reflects a broader acknowledgment of the role that SES plays in influencing educational outcomes. It is argued that unless the impact of SES on educational participation, performance and outcomes is weakened, the ability of the education system to contribute to greater equality in opportunities and outcomes will be constrained. By concentrating more education resources on those at the lower end of the SES scale, the aim is thus to increase the effectiveness of public education spending in reducing overall inequalities in life chances and economic outcomes.

The Concept of Socioeconomic Status

Although the role of SES in influencing educational (and other) outcomes (e.g. in relation to health status and health inequalities) is supported by research studies, the approach used to identify SES varies. This reflects the nature of the concept of socioeconomic status, which has been described as 'an abstract concept for which there is no agreed international method of measurement' (Department of Education, Employment and Workplace Relations, DEEWR, 2009, p. 3).

Although there is no agreement on the precise measurement of SES, there is a broad consensus about what constitutes its main elements. Thus, the DEEWR report cited above argues that:

> 'Most measures of SES use one or more of the following key dimensions of SES – educational attainment, occupation, economic resources and other social and cultural resources. Some measures also include indicators of area and context related aspects of socio-economic status such as geographic location or community' (DEEWR, 2009, p. 3)

In a similar vein, the Australian Institute of Health and Welfare (AIHW) argues in relation to examining the impact of SES on health status, that:

> ‘Education, employment and income are the most commonly used measures of socioeconomic status. However, many other factors can be used, such as housing, family structure and occupation as well as access to resources. Some measures can be based on a single characteristic and others may be composite’ (AIHW, 2010, p. 252)

In an earlier contribution, a report by the Australian Council for Educational Research (ACER) argued that the single most important component of SES was occupation, and that:

> ‘Socioeconomic status can be derived from a single measure or calculated from several variables relating to occupational status. Most often single variable based measures are derived from responses to questions on an individual’s occupation. In contrast, multiple measures can be derived from a range of variables such as father’s and mother’s occupation and educational attainment, income, possessions (such as video recorders, television, cars, size of home etc.), the number of books in the home, and home ownership’ (Marks, 1998, p. 6)

The important role that social status, broadly defined, plays in the emergence and perpetuation of a range of social problems has long been recognised in the academic literature. The authors of the highly cited book, *The Spirit Level. Why More Equal Societies Almost Always Do Better,* argue that:

> ‘It’s hard to disregard social status because it comes so close to defining our worth and how much we are valued. To do well for yourself or to be successful is almost synonymous with moving up the social ladder. Higher status almost always carries connotations of being better, superior, more successful and more able. If you don’t want to feel small, incapable, looked down on or inferior, it is not quite essential to avoid low social status but the further up the social ladder you are, the easier it becomes to feel a sense of pride, dignity and self-confidence. Social comparisons increasingly show you in a positive light – whether they are comparisons of wealth, education, job status, where you live, holidays, or any other markers of success’ (Wilkinson and Pickett, 2009, p. 40)

Here, it is clear that social status affects not only how people are perceived by others, but also how people perceive themselves. A low position on the social status ladder can thus become entrenched as it affects how people are judged by others and the actions and motivations that they themselves take.

The most detailed analysis of SES in the Australian context has been undertaken by the Australian Bureau of Statistics (ABS). In a recent report on the topic, it is argued that:

> ‘Socioeconomic status refers to the social and economic position of a given individual, or group of individuals, within the larger society. Socioeconomic status is usually, but not always, conceived of as a relative concept and can be measured for the individual, family, household, or community area. *The ABS defines relative socioeconomic advantage and disadvantage in terms of people’s access to material and social resources, and their ability to participate in society*’ (ABS, 2011a, p. 1: italics added)

The SES Status of Children

It is important to note at this stage that since the focus of interest in the current context is on children and young people, it is their socioeconomic background that is relevant, and this is normally equated with the SES status of their immediate family. This perspective, although often implicit, underlies most studies of the economic status of children (e.g. studies of child poverty) although it is based on the assumption that the economic resources available to the family are pooled and shared amongst all members so that each individual attains the same standard of living.[1]

The assumption of equal sharing of resources within the family has been challenged by a number of critics of the standard approach, including by feminists, who argue that it ignores gender-based power relations within the family. On this argument, men may end up with a disproportionate share of family resources (income) at the expense of women (and possibly children) who are forced to endure a lower living standard.

It has also been argued that the assumption that all family members are treated equally is inappropriate, and there is evidence that parents (particularly mothers) often put their children's material needs ahead of their own. If this is true, it implies that the standard of living of children is under-estimated in the standard approach (and that of parents over-estimated).

All of these approaches effectively treat children and young people as passive recipients of family resources, with no say in how those resources are divided up, nor make any contribution to how they are generated in the first place. The assumptions underlying this approach have been challenged in recent literature on the sociology of childhood, which argues that the conventional economic approach is flawed because it implicitly treats children not as individuals in their own right (as "beings", that have their own needs and can exert agency and act to supplement or reduce their demand on family income) but as future adults ("becomings") that simply have material needs that impose a drain on family resources.

This is not the place to delve into these arguments in depth, although it is important to emphasise that a choice does have to be made about the relevant *unit of analysis* when estimating the SES of children (e.g. the nuclear family or the household) and this can affect the classifications that are produced.

For most purposes, the unit of analysis would be the (immediate) family comprised of children and parents living together, although the broader concept of the household may be relevant when older children (those aged over 17) are also residing in the parental home.[2] The two approaches will give the same result in the majority of cases, although it is worth noting that a decision has to be made about the ages used to define children and young people and hence to distinguish between them and adults.

[1] When income is used to measure the economic status of families or households, an adjustment must be made to allow for differences in family size and composition and hence in the needs that have to be met out of that income. This called the equivalence adjustment and involves adjusting income by an equivalence scale that captures differences in the relative needs of different families.

[2] ABS income distribution statistics are now based on the household as the basic unit of analysis – see ABS (2011b).

Children are defined for CDRT purposes as under 16 years of age (in line with child protection legislation), with young people defined as between 16 and 17 years of age. In its studies of income distribution, the ABS defines children as being under 15 years of age, while dependent children also includes persons aged 15–24 years who are full-time students, have a parent in the household and do not have a partner or child of their own in the household.

Most economic studies of income distribution and poverty define children as being under 15 years of age, with young people (aged 15-17) considered to have a degree of financial independence which suggests that they should be treated as adults rather than as children.

Definitional and Measurement Issues

An important feature to emerge from the above discussion is that the concept of SES should embrace both economic and social dimensions. This implies that one must either find a single variable that captures these two dimensions, or measure SES using a multi-dimensional approach (or composite variable).

One of the few variables that appears to fit the former (uni-dimensional) approach is *occupation*, since this is related both to the economic returns generated in the labour market, but also captures the social prestige that is attached to different job classifications. Another is *educational attainment*, although the links between both education and income, and between education and social status are weaker (and changing as education practices evolve), making it less likely to provide the basis for an accurate measure of SES.

Both occupation and educational attainment contain strong cohort effects which can influence their interpretation: thus, for example, older people tend to be more likely to have been employed in manufacturing rather than services, and often have lower levels of educational attainment than younger people. This suggests that a group (or area) may have a low average occupational status or educational attainment simply because it contains a higher percentage of older people.

In addition, in both cases, decisions would have to be made about which parent is relevant when there is a two-earner family, and about how to determine the occupational status of jobless families.

A third variable that is often used as a summary measure of economic status (with implications for social status) is *income*. When used as an indicator of the SES of children, the relevant variable is the adjusted (or equivalised) disposable (after-tax) income of the family, which is subject to the qualifications discussed earlier about income-sharing assumptions and is also dependent on the choice of equivalence scale - about which there is also no universal agreement.[3]

[3] Most studies of poverty and economic inequality use the modified OECD equivalence scale, including the ABS in its income distribution reports (ABS, 2011b) and academic research studies of poverty (Wilkins, 2008; Saunders and Hill, 2008). That scale assigns a score of 1.0 to the first adult in the family, 0.5 to each subsequent adult and 0.3 to each child. The total equivalence score for a family if two adults and two children is thus equal to 1.0 + 0.7 + (2 x 0.3) = 2.3. This is less than the number of family members (4) because of the economies of scale in consumption of some items (all four members can use the same furniture and be covered by the same insurance policy) and because children's items tend to cost less than adult items and larger families have more children.

Another problem facing the use of income relates to the difficulties involved in collecting accurate information. This has been acknowledged by the ABS (2003) to be a particular problem for those with incomes at the lower end of the income distribution. There is the additional problem of non-disclosure: many people are unwilling to provide details of their income when asked in surveys, while others forget to include income that is only received infrequently (e.g. interest income).

All three of the variables mentioned above – occupation, educational attainment and income – are measured at the parental or family level and provide an indication of the SES status of children living in families. However, whereas both occupation and education are relatively stable over time, family income is likely to be more variable, either in itself (as employment or pay levels change) or because changes in family size or composition result in a change in equivalent income.

In all three cases, it is generally necessary to devise a small number of categories when presenting the data so that it is more easily presented and assimilated. In the case of occupation, there are a small number of classifications that have been developed by statistical agencies or academic researchers (blue collar, professional, managerial, etc.).

Educational attainment can also be summarised in a small number of categories (high school only, trade certificate, undergraduate degree, postgraduate degree, etc.). In contrast, because income is a continuous variable, it is necessary to compress it into a small number of groups (e.g. deciles of quintiles) that depend on where one fits in the overall ranking of incomes. The use of these summary categories is widespread and it makes sense to adopt them since this allows the results to be more easily compared with other available data.

The choice of summary measure applies to all single variable measures and to composite indicators (see below) and can have an important impact on the implications that are drawn from the data presented. For example, if interest focuses mainly on the lower end of the SES spectrum, then presenting results for the lowest category (in the case of occupation or education) or decile or quintile (in the case of income or a constructed index) will be appropriate. If, in contrast, there is greater interest in the impact of inequality, then a comparison of the lowest and highest categories (or the lowest and highest deciles or quintiles) will provide a better picture.

There is no single method for deciding which approach is preferable, since this depends on the underlying purpose. It is, however, important to bear in mind that the choice of summary measure is important, and that *the use of an inappropriate summary measure can undermine the benefits of selecting a good variable or index in the first place*.

An example is the use in the CDRT *Annual Report* for 2010 of a two-way classification of SES based on the Index of relative Socio-economic Disadvantage (IRSD) into below and above the median (see CDRT, 2011, p. 29). This approach is problematic because the distribution of IRSD scores is closely bunched together around the median (see CDRT, 2010, Figure 25). This implies that an area ranked just below the median is classified into the low SES category, while one ranked just above the median (and thus only marginally higher in the overall ranking than the first area) is classified into the high SES category. The use of a small number of (ranked) categories in such cases needs to take account of the nature of the distribution on which they are based.

It is also important to distinguish between measures that apply across the entire SES spectrum (like IRSD the other SEIFA indexes discussed later) and those that are specifically designed to identify those with low SES. Examples of the latter include whether or not someone is poor which, despite all of its acknowledged limitations, remains an important indicator of low SES. The use of these latter indicators means that the population is separated into those who satisfy the definition (i.e. are below the poverty line) and everyone else, with the variations within each of these two broad groups suppressed.

The choice of whether to apply an approach that extends across the entire spectrum (like SEIFA) or one that only identifies those with low SES (like poverty status) depends on the purpose at hand. The advantage of the former is that it provides the basis for examining whether the variable under investigation is randomly distributed across all quintiles, varies systematically across the quintiles (as would be the case if there is a social gradient), or is concentrated in particular quintiles.

The advantage of the latter is that it focuses attention directly on that part of the distribution – those at the bottom with low SES – that is thought to be a causal factor in the observed incidence patterns.

In the CDRT context, it can be argued that the main focus of interest is not so much on how the incidence of child deaths varies across all SES categories, but rather whether or not children and young people from a low SES background are more susceptible, and if so, how much more than the average (or median).

If the latter perspective is seen as most important (albeit not at the total neglect of the former), then indicators that seek to identify those from low SES backgrounds should be given greater prominence than those that seek to represent SES more generally.

If a multi-dimensional approach to the specification of SES is adopted, it is necessary to determine which dimensions (or variables) are important, and how they should be combined. Both issues raise a number of new challenges and there is no general agreement about how these should be resolved.

The two approaches that have been most commonly used in the literature are the use of factor analysis to identify the underlying (latent) factor (or factors) that best captures the variation in a set of variables (or dimensions), or to derive a single summary variable that is equal to a weighted sum of a set of variables/dimensions.[4] The main difference between the two approaches is that whereas factor analysis allows the variation in the data itself to identify the latent variables and the weighting attributed to each dimension, judgment must be used to decide which variables to include as components of the summary measure and which weights to use when combining them.[5]

[4] The weights might be set equal to each other, in which case the summary measure is equal to the average of the values of the component variables.

[5] The use of factor analysis in social research has come under criticism because it is 'entirely data driven' (McKay and Collard, 2003, p. 45) and because 'no strong theoretical justification is required in deciding which variables to include or exclude from the analysis' (Tomlinson, Walker and Williams, 2008, p. 601). Reflecting these concerns, those who use factor analysis often apply judgment to the statistical output when deciding which variables to maintain, and assign to each factor.

There is no agreement on which of these two approaches is superior, although the fact that there are many viable (and valid) alternative approaches from which to choose when developing a multidimensional (composite) measure means that it is best to think of what emerges as *indicators* rather than *measures.*

As the author has recently noted in the context of developing indicators of social exclusion:

> 'The use of indicators is inevitable where the underlying concept is both multi-dimensional and elusive … Indicators are signposts or pointers that represent things that are not amenable to measurement, but they do not represent the thing itself … [However], the more agreement there is among the indicators about the nature of the issue being examined, the more confidence one can have that the indicators are capturing what is going on, but such agreement will not always exist and there will be instances where different indicators suggest different, perhaps contradictory, conclusions' (Saunders, 2011, p. 196)

These remarks are as relevant to the concept of SES as they are to the concept of social exclusion, and much of the literature on social exclusion has parallels in earlier studies of SES. Concepts like occupation, educational attainment and income are measured on a single dimension, have an agreed (and precise) definition, embody international standards for statistical measurement exist, and reflect agreement over the appropriate measurement metric. In contrast, none of these conditions apply to a concept like SES and it thus becomes necessary to rely on indicators, raising questions about which ones to use that depend in part on the properties of the many alternatives that are available.

Indicator Properties

There is an extensive literature on the properties of indicators, much of which has been revived in recent years as part of the policy interest in complex multi-dimensional issues like wellbeing, social inclusion and (in the educational context) the SES background of students. The ABS has developed a suite of indicators for inclusion in its reports on *Measures of Australia's Progress* (MAP) (e.g. ABS, 2010) that are designed to inform decisions about Australia's progress in three broad dimensions – economic, social and environmental.

It identifies the following criteria as defining what constitutes a 'good' headline indicator (ABS, 2002, Appendix 1):

- Relevance (to a particular aspect of progress);
- Outcome-focused (as opposed to input- or process-focused);
- Unambiguous in interpretation (in relation to progress);
- Supported by timely and good quality data;
- Availability as a time series;
- Sensitive (to changes in underlying conditions);
- Be summary in nature;
- Be capable of disaggregation (by population groups or regions); and
- Be intelligible and interpretable by the general reader.

These properties overlap with those identified in a report that has been influential in the European Union, where a suite of indicators (the "Laeken indicators") have been used by to monitor progress within member countries towards the EU social agenda (Atkinson, Cantillon, Marlier and Nolan, 2002).

The principles of good indicators identified in that report are:

- Clarity and lack of ambiguity;
- Robustness and validation;
- Policy responsiveness (and lack of manipulation);
- Comparability (within and between countries) and consistency (with established national and international standards);
- Timeliness (but subject to revision); and
- Avoidance of unnecessary informational burden on states, enterprises and citizens.

Finally, in the context of its work on identifying the socioeconomic status of Australian tertiary students, DEEWR (2009, p. 2) identifies the following characteristics of a good measure/indicator:

- Construct and predictive validity;
- Transparency;
- Reliability;
- Data efficient;
- Cost effective in collection and timely; and
- Minimal intrusion for respondents.

Differences in the details of the three list of indicator properties reflect different underlying purposes, but there is general agreement that the indicators should:

1. Measure what they are intended to in an unambiguous manner
2. Be transparent and statistically robust
3. Be timely and comparable over time
4. Make efficient use of existing data and/or not be burdensome or intrusive to collect

In other words, the indicators should do the job they are intended for, draw as far as possible on existing data, and not impose an undue burden on any new data providers.

It is important that the requirements of the CDRT to provide information on SES should be considered in the light of these indicator properties.

Particular attention should be paid to the final property if the intention were to collect any new data (as opposed to accessing existing data), and any move in this direction would have to be within the provisions of the existing Act or draw on data that is otherwise freely available. A clear case would have to be made that the benefits of any new data would outweigh the costs of collecting new data and/or of accessing existing data.

There are strong risks in any attempt to collect new data on dimensions of SES and this path should be avoided. The alternative involves exploiting existing available data more effectively, and this latter approach informs the remainder of this report.

Area-based Indexes (SEIFA)

One of the most commonly used multidimensional indicators of SES is the Socio-economic Index for Areas (SEIFA) derived by the ABS (2008a). The SEIFA data have been used in earlier CDRT reports in NSW other states and are still used in Queensland and Sotuth Australia (see below).

When describing and presenting the SEIFA data, the ABS notes that in most circumstances, although SES is normally *conceptualised* for the individual (or family), it is often *measured* (including by ABS) at an area level.

This can be justified on the basis that the characteristics of an area will often influence the status of the individuals who reside in it. The availability and quality of services such as public transport, schools, health and community care, the degree of exposure to crime or environmental risks and access to recreational facilities such as community clubs, sports facilities and restaurants all vary across locations in ways that will influence the wellbeing and living standards (and hence the SES) of local residents.

Some of these factors will have a direct bearing on people's economic status, either because the cost of accessing them will be lower when they are more readily available, or because they allow residents to better maximise their economic potential. Others will affect people's social status, either because they are employed in the provision of local services, or because of the prestige (or lack of it) that attaches to particular localities or addresses.

However, the link between the quality of the area in which one lives and one's overall socioeconomic status is more tenuous and by no means exact: high SES individuals may live in low SES areas, and *vice versa*. This is illustrated by the qualifications attached to the SEIFA, where it is noted that:

> 'While an individual's own socioeconomic status is important, in some contexts it is the socioeconomic status of one's community or the area in which they reside, work or go to school that is either of greater interest *or the only measure available for analysis* … Area based measures of socioeconomic status [do not] provide information about individuals … As with any area or community, there is variation in the characteristics of the overall population in any one place. Judgements about individuals *based solely on the area in which they live* have a high potential for error in the conclusions, due to this variation in the characteristics of the individuals' (ABS, 2011a, p. 12: italics added)

The need for caution when using the SEIFA information to draw conclusions about individuals has been highlighted in two recent ABS studies that have modified the methodology used to develop SEIFA (which utilises principal components analysis, PCA) to produce a new set of Socio-economic Indexes for Individuals (SEIFI), initially for Western Australia (Baker and Adhikari, 2007) and more recently for Australia (Wise and Mathews, 2011).

The latter paper examines the degree of diversity of socioeconomic advantage and disadvantage within census collector districts (CDs) – one of the basic units of analysis used

to construct SEIFA. The results, while only illustrative at this stage, reveal that there is considerable individual diversity of the SEIFI scores within areas sorted into deciles on the basis of their SEIFA scores.

Thus, for example, when CDs are ranked by their score on the SEIFA Index of Relative Socio-economic Disadvantage (IRSD) only 45.4 per cent of individuals who reside in the lowest decile of areas score lowest (on a 4-point scale) in terms of their SEIFI score. Furthermore, almost one-third (31.0 per cent) of individuals with the lowest SEIFI score reside in areas that fall in the top 5 deciles of the SEIFA ranking (Wise and Mathews, 2011, Table 5.10).[6]

The degree of diversity is similar across all states, although ACT and the Northern Territory were consistently identified as having the greatest proportion of areas with a high incidence of (individual) diversity. In contrast, New South Wales (and Western Australia) had the highest proportions of most relatively advantaged people living in the more advantaged areas.

The mis-match between the socioeconomic status of an area (as indicated by SEIFA) and the SES of the individuals who live in the area can be illustrated with data from the *Poverty and Exclusion in Modern Australia* (PEMA) survey conducted by the author in 2010 (further details can be found in Saunders and Wong, 2012). The survey was conducted on a random sample of 6,000 adult Australians draw from the electoral rolls and generated 2,644 responses – equivalent to an effective response rate of 46.1 per cent.

The survey collected information on income (in ranges) and this can been used to estimate the poverty status of respondents using a poverty line set at 50 per cent of median income – the most common poverty line now used in poverty research in Australia and other developed countries. Information is also available on the postcode of (most) respondents and this can be matched to the ABS IRSD deciles, thus making it possible to compare the poverty rates of households living in each IRSD decile.

This is done in Figure 1, while Figure 2 shows the composition of the poor across the IRSD deciles, i.e. allocates all poor households to the IRSD decile in which they reside. The poverty rate does decline with SES as indicated by the IRSD decile, with the poverty rate in the lowest decile (19.9 per cent) more than three times that in the highest decile (6.4 per cent). The results imply that a household selected at random from the most disadvantage decile of areas would have a one-in-five chance of being poor, while a household chosen randomly from the least disadvantaged decile of areas would have around a one-in-sixteen chance of being poor.

The calculations illustrate the fact that although many poor households live in poor areas, so too do many non-poor households and conversely, that many poor households live in non-poor areas.

[6] These results are restricted to those aged between 15 and 64 years and exclude some observations. Further details are provided in Section 2.4 of Wise and Mathews (2011).

Figure 1: Poverty Rates by IRSD Deciles

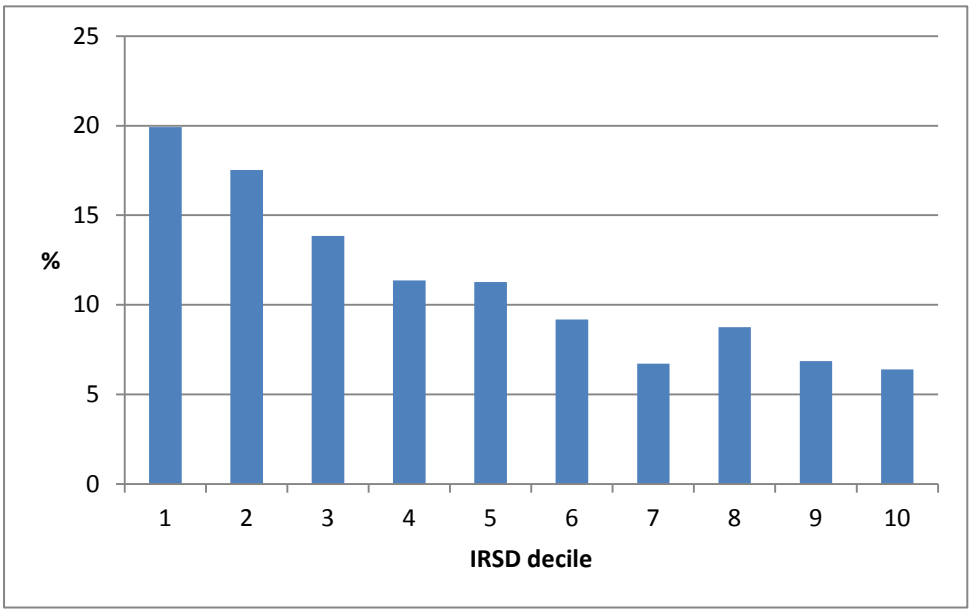
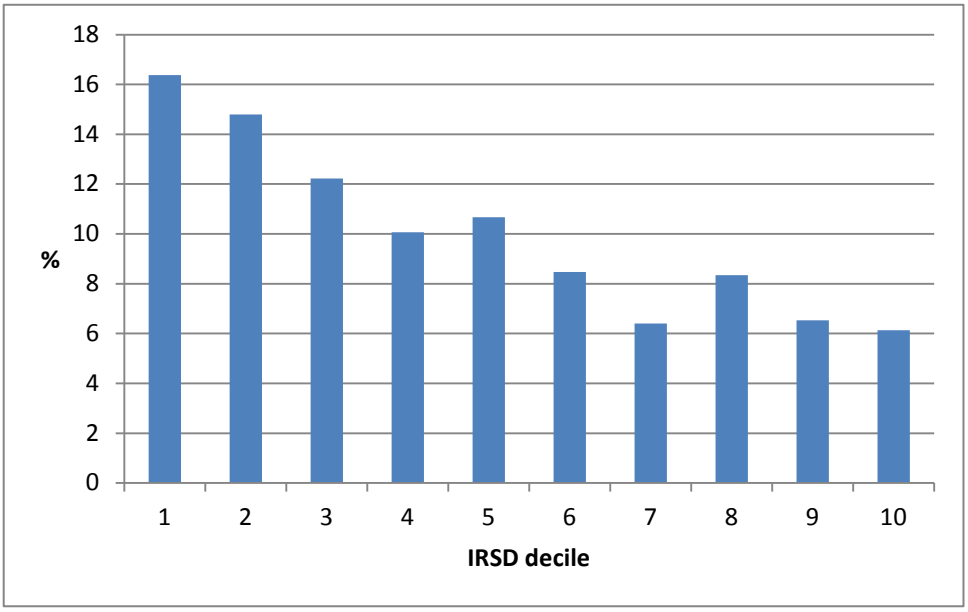


Figure 2 indicates that less than one-third (31.2 per cent) of households with incomes below the poverty line live in the lowest quintile of IRSD areas and just over one-half (53.5 per cent) live in the lowest two quintiles. In contrast, around one-in eight (12.6 per cent) of poor households live in the least disadvantaged quintile of areas, and over one-quarter (27.3 per cent) live in the two least disadvantaged quintiles. Clearly, it is important to distinguish between *households who are themselves poor*, and *households that live in the poorest areas.*

Figure 2: The Composition of Poverty by IRSD Deciles



Returning to a more general discussion of SEIFA, there are four versions of SEIFA, which differ according to which aspect of SES they embody, and the variables used to construct them. The four are identified below, along with a brief description of what each index seeks to capture:

- *The Index of Relative Socio-economic Disadvantage* (IRSD) is a general index that summarises a range of information about the economic and social resources of people and households within an area, restricted to measures of relative disadvantage;
- *The Index of Relative Socio-economic Advantage and disadvantage* (IRSAD) is similar to IRSD but includes measures of both relative advantage and disadvantage;
- *The Index of Economic Resources* (IER) focuses on the general level of access to economic resources of people and households within an area; and
- *The Index of Education and Occupation* (IEO) focuses on the general level of education and occupation-related skills of people within an area.

The four indexes capture different aspects of SES although they are fairly highly correlated, with the correlations between IRSD and IRSAD, IER and IEO (in 2006) being 0.94, 0.92 and 0.80, respectively (ABS, 2008b, Table 5.5).

Which of the four indexes should be used? This depends on the purpose at hand, but those studies that focus on social disadvantage normally use IRSD and, as indicated earlier, this is the case with previous CDRT practice, although Queensland currently uses IRSAD. The former index (IRSD) is derived from the 17 variables listed in Table 1.

Based on the results of the analysis used to derive the indexes (see below), the following four of these variables are not used to calculate IRSAD (see ABS, 2008a, Appendix 4):

- % People who identified themselves as being of Aboriginal and/or Torres Strait Islander origin
- % People aged 15 years and over who are separated or divorced
- % People aged 15 years and over who did not go to school
- % People who do not speak English well

At the same time, the following seven new variables are included in the estimation of IRSAD:

- % Occupied dwellings with four or more bedrooms
- % People aged 15 years or over at university or other tertiary institution
- % Households paying mortgage greater than $2,120 per month
- % People aged 15 years and over with an advanced diploma or diploma qualification
- % Employed people classified as Professionals
- % Occupied private dwellings with broadband internet connection
- % People with stated annual household equivalised income greater than $52,000 (approximates the 9^{th} and 10^{th} deciles)

It is clear that the variables that are dropped when constructing IRSAD and the new variables introduced reflect the shift in emphasis away from the notion of relative disadvantage to one designed to capture both disadvantage and advantage. However, it is not clear that all of these new variables adequately capture the notion of advantage, or do so in a way that is likely to be enduring.

This is the case for the broadband internet access variable, for example, which can be criticised as an indicator of advantage in today's environment and is in any case likely to become increasingly redundant as the NBN scheme is rolled out. If this is the case, the IRSAD will have to be revised, undermining its comparability.

These factors, combined with the emphasis on low SES that was argued earlier to be of more relevance in the CDRT context, suggest that the IRSD is preferable to the IRSAD index for current purposes.

The method used to derive the IRSD is principal components analysis (PCA) which the ABS describes as follows:

> 'The aim of PCA is to summarise a large number of correlated variables into a smaller set of transformed variables, called 'principal components'. Each component is a weighted linear combination of the original variables. … The first principal component is the weighted linear combination designed to capture the maximum amount of variation present in the original dataset … The first principal component is used for SEIFA, as it is the single transformed variable that best summarises the common trend underlying the original set of variables' (ABS, 2008b, p. 17)

Once the first principal component has been identified, variable loadings can then be estimated as the correlation coefficients between each variable and the component, and variable weights which are the coefficients used in the linear transformation that produces the component. Those variables that have the highest loadings or greatest weight are given the greatest (statistical) emphasis in the summary measure or principal component.

Table 1 lists the 17 variables used to construct the IRSD from the 2006 census data and shows the loading and weight for each variable.[7] The variables included in the analysis are all defined in binary terms, and were given a value of one if the characteristic is present for the individual and zero otherwise.

[7] As ABS notes, variables with a loading of less than 0.3 were excluded from the analysis.

Table 1: Index of Relative Disadvantage: Variable Loadings and Weights in 2006

	Variable description	Variable loading	Variable weight
1	% Occupied private dwellings with no Internet connection	-0.85	-0.33
2	% Employed people classified as Labourers	-0.76	-0.30
3	% People aged 15 years and over with no post-school qualifications	-0.76	-0.30
4	% People with stated annual household equivalised income between $13,000 and $20,799 [(a)]	-0.76	-0.30
5	% Households renting from a Government or Community Organisation	-0.70	-0.27
6	% People (in the labour force) unemployed	-0.70	-0.27
7	% Families that are one parent families with dependent offspring only	-0.67	-0.26
8	% Households paying rent who pay less than $120 per week [(b)]	-0.67	-0.26
9	% People aged under 70 who have a long-term health condition or disability and need assistance with core activities	-0.61	-0.24
10	% Occupied private dwellings with no car	-0.57	-0.22
11	% People who identified themselves as being of Aboriginal and/or Torres Strait Islander origin	-0.52	-0.20
12	% Occupied private dwellings requiring one or more extra bedrooms [(c)]	-0.52	-0.20
13	% People aged 15 years and over who are separated or divorced	-0.51	-0.20
14	% Employed people classified as Machinery Operators and Drivers	-0.51	-0.20
15	% People aged 15 years and over who did not go to school	-0.44	-0.17
16	% Employed people classified as Low Skill Community and Personal Service Workers	-0.44	-0.17
17	% People who do not speak English well	-0.33	-0.13

Notes: (a) This range approximates the 2nd and 3rd deciles of the income distribution; (b) Excludes zero values; (c) This variable is based on Canadian National Occupancy Standard.
Source: ABS (2008b, Table 4.2)

It can be seen that the IRSD captures five broad characteristics of the household and its members: occupation (variables 2, 14 & 16), educational attainment (variables 3 & 15), housing status (variables 5, 8 & 12), household type (7 &13) and a range of household characteristics (variables 1, 4, 6, 9, 10, 11 & 17). The first four variables listed in Table 1 all have loadings in excess of 0.75 and weights of 0.3 or more. They capture the household's access to the internet, and the occupation, educational attainment and income of its members.

Once the PCA has been conducted, a score is derived by applying the weights (shown in Table 1) to the average values of each variable for each CD, with the more disadvantaged areas receiving a lower score. The scores for all CDs are then standardised so that the mean score is equal to 1,000 and the standard deviation is equal to 100. These scores are then ranked from lowest to highest and can be split into deciles or quintiles.

It is important to emphasise that the derivation of the scores themselves and their ranking mean that they are *relative* measures only. This implies, for example, that a CD with a score of 500 is not twice as disadvantaged as a CD with a score of 250, nor that the difference (in disadvantage) between CDs with scores of 800 and 700 is the same as the difference between CDs with scores of 400 and 300.

The distribution of IRSD scores is skewed to the right (for Australia, see ABS, 2008a, Figure 4.1; and for NSW, NSW Child Death Review Team, 2011; Figure 25) with a long left-hand tail

and with many CDs bunched closely together in the middle deciles (4-8). The ABS (2008a, p. 8) notes that the IRSD scores can be used in 'more sophisticated analysis' but argues that the deciles 'should be used for most analyses'.

It also notes that changes introduced in 2006 mean that: 'Comparison with previous indexes is not recommended'. These changes affected the income variable used in the PCA, which included an equivalence scale adjustment for the first time (using the modified OECD scale described earlier) and the inclusion of new variables relating to Internet access and the need for assistance with core activities. All three variables entered into the derivation of IRSD for the first time in 2006 (see Table 1), making the estimates for that year not comparable with those for earlier years.

However, the extent to which the non-comparability of the IRSD estimates gives rise to cause for concern is ultimately a matter for judgment and is dependent on the specific purpose for which the index is being used. It would clearly not be appropriate to compare IRSD scores for a particular area in 2001 and 2006 and draw conclusions about how it changed over the period. However, there may be less of a problem when ranking areas by the decile (or quintile, or quantile) of the distribution into which they fall in different census years.

Use of the more sophisticated income variable and the inclusion of the Internet access and core assistance requirements variables in 2006 provides a better estimate for that year but is less likely to cause major shifts in an area's position in the overall distribution of IRSD scores (particularly if the new variables are closely correlated with those they replaced and with other variables used to construct IRSD).

It seems likely that the ABS will continue to improve the methods used to construct its SEIFA indexes and so the issue of non-comparability will be on-going. This will make long-term comparisons particularly hazardous, but since the focus of CDRT is on better understanding the *cross-sectional* relationship between SES and the incidence of child deaths, the issue of inter-temporal non-comparability is not a major issue.

One great advantage of the IRSD (and other SEIFA indexes) is that the information is available for areas classified in a number of different ways. One of these is Postal Areas (POAs), which are formed by aggregating CDS in order to reflect the postcodes used by Australia Post to deliver mail. The matches are not always precise (because a CD may span two separate postcodes but can only be assigned to one of them, for example) and the ABS notes that: 'SEIFA users need to be aware that Postal Areas and postcodes are not always good matches, and should use POAs with caution' (ABS, 2008a, p. 48).[8] This caution should be kept in mind, but is again unlikely to adversely affect comparisons that are conducted at a high level of aggregation such as for IRSD deciles or quintiles.

By making it possible to link information on an individual's postcode (which is information readily available and easily collected) to the IRSD score, rank or decile, it is straightforward to identify the relative disadvantage of a person's location on a national or state-wide basis. There will, of course, always be some individuals whose characteristics do not align with those of the area in which they live and they will be falsely assigned (the ecological fallacy),

[8] There were 2,515 postal areas used in 2006, of which 2,474 were included in SEIFA.

but the overall power and minimal data requirements of the general approach explains the popularity and widespread use of the SEIFA indexes.

It was noted earlier that the ABS has been exploring the development of a new set of indexes (SEIFI) that are derived directly from the *specific* characteristics of individuals and households rather than the *average* characteristics of the areas in which people live. This raises the obvious question of the degree to which the variables identified as important in the PCA used to construct these indexes overlap.

This can be judged from Table 2, which compares the factor loadings and weights used to construct the SEIFI and SEIFA IRSDs using census data for 2006. The number of variables included in the SEIFA analysis summarised in Table 2 is lower than the 17 shown in Table 1 because of differences in the method used to identify the final model.[9]

Table 2 shows that the move from the area-based to an individual-based measure of disadvantage results in a general reduction in the variable loadings (and in the degree of variation explained by the first component; see ABS, 2011a, p. 17).[10] This is the case for all but the first two variables (renting from government or social housing, and low weekly rent) and these two variables now have the highest loadings and weights.

Table 2: Comparing Variable Weights and Loadings on the SEIFA IRSD and SEIFI IRSD in 2006

	SEIFA IRSD		SEIFI IRSD	
Variable	Loading	Weight	Loading	Weight
Person resides in a household renting from Government or Community Organisations	-0.70	-0.27	-0.75	-0.52
Person resides in an occupied dwelling paying less than $120 rent per week	-0.67	-0.26	-0.72	-0.50
Person resides in an occupied dwelling with no internet connection	-0.85	-0.33	-0.50	-0.34
Person resides in an occupied dwelling with no car	-0.57	-0.22	-0.49	-0.34
Person identifies themselves as being of Aboriginal and/or Torres Strait Islander origin	-0.52	-0.20	-0.44	-0.30
Persons aged 15 and over with no post-school qualifications	-0.76	-0.30	-0.35	-0.24
Persons aged 15 and over who left school after year 11 or lower	N/A	N/A	-0.33	-0.23
Person has stated annual household equivalised income between $13,000 and $20,799	-0.76	-0.30	-0.32	-0.22

Source: Wise and Mathews (2011, Tables 2.1 & 4.1)

The declines are greatest for the no internet connection, no post-school qualifications and low income variables, casting doubt on the ability of these variables to capture disadvantage at the individual level. All of the occupation variables that had high loadings in the original

[9] In the comparative analysis reported in Table 2, variables with a loading of less than 0.3 on the first principal component were dropped and the PCA analysis re-run until all remaining variables had a loading above 0.3 (see ABS, 2011a, p. 16).

[10] All variables were again defined in binary terms, depending on whether or not the characteristic was present in the individual

SEIFA analysis (see Table 1) no longer appear amongst the variables used to construct SEIFI, although (low) income and two of the education variables do remain in the list of SIEFI variables.

The derivation and presentation of SEIFI results for 2006 has highlighted the extent of individual level diversity within disadvantaged areas. As ABS notes these new results:

> ' … are important illustrations of how diversity of advantage and disadvantage within an area can exist, and the extent to which individuals with differing levels of socio-economic advantage and disadvantage reside in the same area' (Wise and Mathews, 2011, p. 38)

There are, however, some shortcomings with the construction of the SEIFI indexes that need to be noted. They include the far higher proportion of cases that had to be excluded when constructing the index from the raw census data, either in order to make the SEIFA and SEIFI comparable or because of applicability issues. The most important of these was the restriction of SEIFI to those individuals aged between 15 and 64, which together with a number of other exclusions led to the exclusion of over 6.5 million persons (33.1 per cent).

The decision to restrict the SEIFI to those aged between 15 and 64 years was based on the view that many of the variables used to construct the index are not relevant to those who fall outside of this age range. As ABS notes:

> 'Many of the census variables relating to advantage and disadvantage address factors such as employment, education and economic resources. These aspects of advantage and disadvantage are not necessarily relevant for all person in the population. For instance, young people under the age of 15 will most likely not have completed their education or be employed, and so will have socio-economic characteristics largely determined by their parent or guardian, whilst retirees over the age of 64 find their accumulated wealth is a better indicator of their economic standing than their income. Therefore, because of the different stages of the life cycle, it is not practical to have a unique individual level index for all demographics of the population' (Wise and Mathews, 2011, p. 8)

The ABS also argued that the SEIFA approach is 'more theoretically and conceptually sound' than the SEIFI because of the strict exclusion rules applied and because more effort was put into external validation of the index.[11] While this latter shortcoming reflects the experimental nature of the new SEIFI indexes, the large number of excluded cases is a more profound problem that will not be so easily resolved. It was argued that SEIFA 'remains an important robust product' that should continue to be used for socio-economic analyses as long as the caveats about it measuring the status of an area, not of the individuals who reside within it are not forgotten.

One final, perhaps critical, factor that restricts the use of SEIFI is that it is no longer possible to link it to readily available data as is the case with the SEIFA, which can be linked to postcode. This linkage is a very important property of SEIFA (even though it reflects its *area-*

[11] Exclusion from the SEIFI was not based solely on age, but also affected those who were not in the CDs that were included in the SEIFA analysis. However, this exclusion accounted for only 2.3 per cent of all exclusions, with age accounting for the remainder.

based focus, which is a problem if interest is on the SES of *individuals*) because it means that one can locate an individual into the SEIFA ranking merely by knowing where they live (their postcode).

There is no corresponding link from readily available individual characteristics to the SEIFI and those variables that are given greatest weight when constructing SEIFI (rent status, rent level and income) cannot be derived for individuals without what many would regard as an unacceptable degree of intrusion.

Current (and Recent) Practice in the CDRT Context

The discussion so far indicates that the possible measures of SES considered are single variables like occupation, education and income and a composite index approach based on SEIFA or SEIFI. These are two extremes of a spectrum of possibilities that stretch from reliance on a single variable to ever more complex indexes that combine a group of variables in some way (either by using judgment to assign a weight to each variable, or by relying on PCA to identify correlations within the data and derive factor weights and loadings accordingly).

It has also been argued that in the CDRT context, particular attention should be placed on indicators that capture low SES as opposed to those that reflect the full range of SES.

The IRSD was used in NSW to measure SES between 2004 and 2008 and child deaths were allocated to one of six categories identified by the percentile rankings of IRSD (see the Appendix for details). No justification was given for the choice of these categories (over the use of the more familiar quintiles, for example).

The use of the IRSD to measure SES was discontinued in 2008 because;

> '… the ABS advises that this is not a good measure to use with time series data, as it is revised regularly and its construction and interpretation can change with each revision' (NSW Child Death Review Team, 2008, p. 384).

It was noted that a problem with indexes like the IRSD is that there are many ways in which an area may end up with a particular disadvantage score. Thus, a high area score might reflect a high proportion of Aboriginal residents, a high proportion of sole parent families, or a high proportion of people whose occupation is a Labourer. The implication is that these differences undermine the comparability of the scores or make it difficult to interpret them. However, this criticism can be leveled at the use of any composite index and misses the point about the importance of using a multidimensional index to measure a complex concept like SES.

As a replacement, the CDRT focused on a small number of variables used to construct the IRSD that were argued to be 'more likely to have a durable meaning across time'. For each variable, census figures for 1996, 2001 and 2006 were used, with values for the intervening years derived using (unexplained) interpolation.

The analysis concentrated on the following four variables:

- The proportion of the population 15 years old and older employed

- The proportion of the population with a post-school qualification (above the Higher School Certificate or an Australian Qualifications Framework Certificate II qualification)
- The proportion of the population that are Aboriginal or Torres Strait Islander
- The proportion of the population aged 17 years or younger

It was argued that all four variables capture an element of disadvantage (or low SES) although it was acknowledged that the final variable better captures family or community need as opposed to disadvantage as such.

Aside from the third variable, the others are all affected to some degree by the life cycle considerations that led the ABS to apply their new SEIFI methodology only to those between the ages of 15 and 64. Thus, both the first and second variables will be affected by the proportion of the population that is over 64, most of whom are no longer employed, and many of whom were educated in earlier periods when far fewer people acquired a post-school qualification. An area might therefore score low on both variables because it has a high proportion of residents aged 65 and over, not because its current working-age members are not employed or have poor education.

In conjunction with the (acknowledged) limited ability of the fourth variable to capture disadvantage as such, it is not clear that this new approach is capable of generating evidence on the extent of disadvantage or SES status that has 'a durable meaning across time' as was claimed.

The latest (2010) CDRT Annual Report for NSW re-introduced the use of IRSD, but (as noted earlier) areas were classified into high or low SES areas depending on whether or not they fall above or below the median value (see CDRT, 2010, Tables 20 & 21). Use of this cruder two-way classification was justified on the grounds that the SES status of households may vary widely within SEIFA areas, although it was acknowledged (p. 29) that the IRSD 'includes measures such as education and household income that are generally regarded as core elements of the concept of socioeconomic status.'

The IRSD was also used in a comprehensive report on trends in child deaths in NSW over the decade to 2005 that was conducted by the CDRT (NSW CDRT, 2008). Figures on the distribution of child deaths were presented for low, middle and high SES locations, where these were identified to cover the bottom, middle and top thirds of the distribution of IRSD.[12] The report noted that the correlation between household socioeconomic status and community socioeconomic background 'is mostly found to be a moderate one' but justified the use of IRSD on the grounds that 'information about community socioeconomic background … independently measures other characteristics of local communities that are important in their own right' (p. 459)

Two other states currently use variants of the SEIFA indexes to measure the SES background of children who have died. As noted earlier, in Queensland, the 2009-10 report uses another version of SEIFA, the Index of Relative Socio-economic Advantage and Disadvantage (IRSAD).

[12] The report notes (p. 460) that the IRSD rankings are based 'only on data for populations aged 0-17' although what this implied in practice was not specified and is unclear.

Unlike the IRSD, the IRSAD includes variables that reflect the degree of advantage as well as those that reflect the degree of disadvantage that exists within an area.[13] [14]

For presentational purposes, child deaths in Queensland are separated into those from 'low to very low'. 'moderate' and 'high to very high' SES backgrounds using the IRSAD (see Commission for Children and Young People and Child Guardian, 2011, Figure 1.3 (incorrectly headed), p. 32). The report does not define how these three SES categories are themselves defined.

South Australia presents a breakdown of child deaths in that state by SES, where the IRSD is used to measure SES. Child deaths are located into the quintiles of IRSD into which their location belongs and results are presented for the number of deaths in each quintile (see Child Death and Serious Injury Review Committee, 2011, Table 2, p. 24).[15]

This brief review of current applications in three states illustrates how even the use of a common approach to measuring SES (based on the SEIFA indexes) still leaves decisions to be made about: (1) the choice of which of the 4 available SIEFA indexes to use; and (2) the choice of how to summarise the chosen index for presentational purposes.

In relation to the former, it has been argued here that the IRSD is preferable given the focus of the CDRT (but see below). In relation to the latter, current and past practice covers a broad range of summary measures, including the six-way classification based on percentile rankings used in NSW until 2008, a bottom-third/middle-third/top-third classification used in the 2008 report produced by the NSW CDRT, the two-way, above/below the median used in NSW in 2010, the use of a three-way, low/moderate/high breakdown used in NSW up until 2008, the similar but unexplained approach used in Queensland, or the quintile breakdown used in South Australia.

Although there are good conceptual and practical grounds for using a summary measure that is based on a small number of SES categories, this still leaves many potential options and it is clear that past practice has adopted many different approaches. The trend in NSW has been towards using a smaller number of options (from 6 down to 3 or 2), all based on IRSD, although Queensland still uses 3 categories and South Australia uses 5 quintiles.

A strong case can be made for the use of quintiles when seeking to summarise the SES indicators (whichever ones are used) in the CDRT context. This assessment is based on three

[13] The 2010 NSW Annual Report preferred to use IRSD rather than IRSAD because it does not include 'measures such as mortgage payments that are likely to depend upon the average time of residence in the area' (p. 117). It is not clear whether it is length of residence or age that is the important factor here and if it is the latter, the same criticism can be applied to some of the other factors (e.g. educational attainment) that are used to construct the IRSD.

[14] Unlike the IRSD, the IRSAD does not include Indigenous status as a variable, which means that it can be used to analyse information about Indigenous status while IRSD cannot because there will be a definitional relationship between the index and one of the variables used to construct it (see ABS, 2008a, p. 11-12).

[15] The latest SA report also notes that a number of social or environmental factors associated with the child or their family may affect its well-being, including: family and domestic violence, alcohol use, parental supervision, drug or substance abuse or homelessness. The report notes that: 'Reviews of investigable deaths often highlight the impact of these factors on the circumstances leading up to the child's death and, where this occurs, these factors are recorded to enable an analysis of patterns and trends to assist in considering ways to prevent or reduce future deaths'

consideration: the wide use of quintiles in a broad range or related studies of SES (e.g. the ABS income distribution studies); the fact that quintiles are not only widely used, but easy to explain; and the fact that the lowest quintile can itself be used as an indicator of low SES.

Towards a New Approach

As noted earlier, the range of options available to measure SES is restricted by the constraints facing the CDRT, specifically those relating to data availability. It is not a practical option to collect new data so any measure must be based on data that is already available and/or easily accessible by the CDRT. It is, however, clear that even within these limits, choices have to be made that will influence the relevance and usefulness of alternative measures.

Table 3 summarises the features of the available SES indicators using the criteria used generally to identify a good indicator – one that achieves the task at hand (of measuring SES), is readily understood, statistically robust and based on data that is readily available.

The first three variables score well in terms of clarity and transparency but low in terms of robustness because none of them individually correlates that well with SES. This problem is overcome by the two composite indicators (which include the first three variables in their construction) but at a cost in terms of either their relevance to individuals as opposed to areas (SEIFA) or in terms of their experimental nature and uncertain availability (SEIFI).

Table 3: Summary Criteria Ranking of Alternative Approaches

MEASURE/INDICATOR:	Clarity (lack of ambiguity)	Transparency	CRITERIA: Statistical Robustness	Timely/ comparability	Data availability
Single variable:					
Occupation	HIGH	HIGH	MEDIUM/LOW	HIGH	HIGH
Educational Attainment	MEDIUM	MEDIUM	LOW	MEDIUM	MEDIUM/HIGH
Income	HIGH/MEDIUM	HIGH	MEDIUM/LOW	MEDIUM/LOW	LOW
Composite indicator:					
SEIFA/IRSD	MEDIUM/LOW	MEDIUM	HIGH	MEDIUM	HIGH
SEIFI	HIGH/MEDIUM	MEDIUM	HIGH/MEDIUM	???	???

Although the use of SEIFA has been subject to much criticism, it remains the single most widely used indicator of SES and has the enormous advantage that a household can be assigned its place in the national (or state) ranking on the basis of its address (and hence postcode).

Against this, the SEIFA data are only available every five years (when the census is conducted), even then after a considerable delay and have been subject to revisions that have compromised the comparability of the data. Over time, however, the most telling criticism that has been leveled against the use of SEIFA in the current (and many other) contexts is that it captures the average SES status of an *area*, not that of the individual households that reside within it.

This is an important limitation, but not one that rules it out completely in the current context, where no single approach is both conceptually superior and practically achievable. Particularly when it comes to children and young people, the nature and quality of the local community in which they grow up will exert an important influence on them and it can be argued that an area-based indicator like IRSD captures key elements in the constellation of location-based factors that determine their well-being and SES.

Furthermore, the fact that not everyone living in a low SES would themselves be classified as low SES area does not mean that the area-based indicator is wrong for everyone: this clearly cannot be the case, since the IRSD is itself derived from the characteristics of local residents. And even if the use of IRSD produces an incorrect assessment for some households (the ecological fallacy) it may still do a better job *overall* than any of the alternative indicators that are available.

For all of these reasons, there are strong grounds for maintaining the use of SEIFA – at least as one element in what can be a multiple approach.

Which SEIFA should be used and should SEIFI be used if (or when) it becomes available? In relation to the first issue, it has been argued that there are strong grounds for using IRSD if the focus is primarily on identifying households (and children) that come from a low SES background. The main alternative, IRSAD, captures the degree of both advantage and disadvantage and it implies, for example, that if a large number of more advantaged households move into an area, its SES ranking will improve. In contrast, such a shift will have no (or very little) impact on the IRSD ranking of the area.

One can ask which of the two approaches better captures the impact of the change on the SES status of the children and young people who were already living in the area. *From their perspective*, the influx of better-off people may eventually raise the quality of the area but it can be argued that its immediate impact would more likely be to make the existing children more conscious of their low status and possibly increase their sense of inferiority. So the net effect is uncertain and it is probably wiser to adopt an index that does not change as a result of the shift, which favours the use of the IRSD.

The next issue that arises relates to the choice of summary indicator. This is an important consideration as explained earlier, where it was argued that the use of an inappropriate summary measure can undermine the value of the indicator itself. It has also been noted that while current practice is to disaggregate the indicator into a small number of sub-categories, there is no agreement on how these sub-categories should be defined. While it makes sense to use only a small number of categories some thought should be given to exactly how they are defined.

The earlier discussion suggests that the summary measure of SES should be based on the quintiles of IRSD.

There is, of course, no reason to use only a single indicator. The complex and contested nature of SES itself suggests that no one indicator can capture all of its dimensions and complexities. In addition, the concern over the use of the area-based IRSD to capture the circumstances of individual households suggests that *it may be wise to supplement IRSD with another variable that captures SES – or more specifically low SES – more directly.*

Three such variables proposed here are occupation, education and housing tenure. The use of both occupation and educational attainment as an indicator of SES has a long history and, if information on parental education and occupation is readily available either can serve to supplement the IRSD with an indicator that is based on the characteristics of the individual as well as those of the area of residence.

It is also worth noting that education (but not occupation) was included among the variables used to construct the experimental SEIFI index (see Table 2) and that both occupation and educational attainment enter into the calculation of IRSD (see Table 1). Against this, educational attainment, as was noted earlier, varies systematically over the generations and may in some circumstances act as a proxy for age (or cohort). There are also problems applying the education categories to those who were educated overseas. Occupation does not suffer from these problems although there are problems deciding how to treat two-earner families (now the norm) and those who are not employed, but these problems can be overcome.

It is also possible to address some of the concern about the ecological fallacy (that not all those living in a low SES area are themselves low SES) by combining IRSD with the fourth of the SEIFA indexes described earlier, the Index of Education and Occupations (IEO). This index includes two of the variables that are central to the concept of SES as it has evolved in the literature.

Thus it would be possible to identify those households who fall within the lowest quintiles of *both* IRSD and IEO to better represent those who are from a low SES background.

In relation to occupation, the Australian and New Zealand Standard Classification of Occupations (ANZSCO) has recently been revised (see ABS, 2009) but still maintains the following 8 major group categories:

Major Group 1: Managers

Major Group 2: Professionals

Major Group 3: Technicians and Trades Workers

Major Group 4: Community and Personal Service Workers

Major Group 5: Clerical and Administrative Workers

Major Group 6: Sales Workers

Major Group 7: Machinery Operators and Drivers

Major Group 8: Labourers

These 8 major groups could be collapsed into three broader occupational categories:

Occupational category 1: ANZSCO groups 1-3

Occupational category 2: ANZSCO groups 4-6

Occupational category 3: ANZSCO groups 7 & 8

Attention would then focus on those who belong to the third of these categories, as membership of this group would generally be expected to indicate low SES.

In relation to educational attainment, it would be possible to use a variable based on completed years of schooling (of the parent) if this is available. In its absence, the education variable would be captured in the IEO discussed above.

The third variable mentioned above is housing status, specifically whether or not one is residing in government or social housing. This variable has the greatest weight in the SEIFI index (see Table 2), has a high weight in IRSD (Table 1) and has been shown in a number of studies to be a strong predictor of low SES.[16]

Thus, for example, recent studies conducted by the author have estimated deprivation in Australia by identifying items that are regarded as essential ("things that no-one in Australian should have to go without today") and then establishing who does not have these items because they cannot afford them (Saunders, Naidoo and Griffiths, 2007; Saunders and Wong, 2012). A study conducted using the 2010 PEMA data referred to earlier indicates that the mean level of deprivation experienced by public housing tenants was 3.99 – more than three times the overall average 1.30 and higher than that among any other socio-demographic group (Saunders and Wong, 2012).

Information on this variable is presumably readily available if the child's residential address is known from state housing authorities.

The above information on occupation, educational attainment and housing tenure could then be used to refine (or validate) that on low SES based on the IRSD quintiles. It would be important not to be definitive about how the indictors should be combined but to experiment with some alternatives approaches. A useful starting point would be to present the IRSD quintiles as the base measure of SES and have the following three adjusted SES measures as additional refinements:

SES Adjustment 1: Belongs to the lowest quintile of IRSD *and* the lowest quintile of IEO

SES Adjustment 2: Belongs to the lowest quintile of IRSD *and* the lowest of the three broad occupational categories identified above.

SES Adjustment 3: Belongs to the lowest quintile of IRSD *and* is a resident of government/social housing.

SES Adjustment 4: Belongs to the lowest quintile of IRSD *and* the lowest occupational category *and* is a resident of government/social housing.

The use of these adjusted measures would provide a firmer basis for concluding that the children and young people concerned do indeed come from a low SES background. It is important to emphasise, however, that these adjusted measures are exploratory and their use should be considered experimental. They should also be seen as complements to the base indicator – membership of the lowest quintile of IRSD.

Their use should be accompanied by on-going search for existing readily available data that would allow the SES status of individual families/households to be established.

[16] The other variable that has a high factor loading in the construction of the SEIFI is the amount of rent paid (see Table 2). However, this information is unlikely to be readily available, thus ruling out its use as an indicator of SES in the current context.

PART TWO: REPORTING ON THE GEOGRAPHIC LOCATION OF CHILD DEATHS

Knowing where children and young people who have died resided is important information that can assist with a better understanding of the factors that led to the death and may also serve a preventative role. If, for example, there is a systematic relationship between geographic location and the availability of, or access to, services that provide support to children and their families then knowledge of where service provisions is currently most inadequate can direct attention to where action is most urgently needed.

The CDRT has in the past presented information on the geographic location of child deaths using the Accessibility/Remoteness Index of Australia (ARIA) originally developed by the National Centre for Social Applications of Geographical Information Systems(GISCA). ARIA measures the remoteness of a location based on the physical road distance to the nearest urban centre and classifies areas into five classes based on that measure.

This information is collated by the ABS into the following five main broad categories:[17] [18]

1. Major cities
2. Inner regional areas
3. Outer regional areas
4. Remote
5. Very remote.

The CDRT *2008 Annual Report* noted (p. Xxviii) that:

> 'Over the next 12 months the Team will continue to investigate currently available methods to inform the analysis of child mortality across the geographic regions of NSW'

However, no action was taken and the latest (2010) Annual Report provides (Tables 17-19) an analysis of the leading causes of death by remoteness that differentiates between children residing in major cities, in inner regional areas, and in more remote areas. Of the 530 deaths include in this analysis, 339 (64 per cent) were in major cities, 123 (23 per cent) were in inner regional areas and 68 (13 per cent) were in more remote areas.

As noted earlier, two concerns have arisen in relation to the CDRT geographic location analysis: first, there is the issue of small sample size, particularly for those living in remote areas; second, there is the issue of cross-border deaths (that are most prevalent among those living in the far North coast region, close to the NSW-Queensland border) where the required information is not available.

Although sample size is something of a concern for the number of deaths in more remote areas, this is an inevitable consequence of the disaggregation itself (which rightly focuses on the issue of remoteness). However, one consequence of this is that it restricts the possibilities for disaggregating the location-based data (at least for those in more remote areas) into other characteristics (e.g. age or SES) as a way of examining the independent

[17] There is a sixth (migratory) category but that is not relevant to this discussion.

[18] ABS notes that the rules for separating areas into these five classes can be relatively arbitrary, though it acknowledges that the members of each class are more like one another in terms of their distances to urban centres than they are to members of other classes

effects associated with each characteristic. This also applies to disaggregating by cause of death, except for the broadest categories.

Small sample size will also cause the numbers to fluctuate considerably from year to year, suggesting that not too much emphasis should be given to the numbers for any given year, or to changes from year-to-year. It is, however, possible to aggregate the figures over several years (e.g. by calculating 3- or 5-year moving averages) as a way of assessing whether there is any discernible longer-term trend.

Over time, this approach will produce a series that is likely to be more informative than the annual figures and this approach is recommended (with possible backdating if past data are still available).

The use of multi-year data will also make it possible to disaggregate by other factors in order to examine the extent to which location itself (as opposed to Indigeneity, for example) is systematically linked with the different types of child deaths, or with other characteristics of the child (e.g. age) or their family (e.g. their SES background).

In relation to the issue of cross-border deaths, this is a weakness of the CDRT data generally and is not specific to the breakdowns by geographic location (although the impact on those in more remote areas may have more practical significance here). Since the issue is largely related to the far North coast and presumably creates similar problems for Queensland, both states would have much to gain by cooperating with the other in order to provide the information needed to improve data quality in both jurisdictions.

Such effort seems an obvious place to begin to resolve an issue that is causing concern in both states.

References

Atkinson, A.B., Cantillon, B., Marlier, E. And Nolan, B. (2002), *Social Indicators. The EU and Social Inclusion,* Oxford: Oxford University Press.

Australian Bureau of Statistics (ABS) (2002), *Measures of Australia's Progress 2002,* Catalogue No. 1370.0, Canberra: ABS.

ABS (2003), 'Revised household income distribution statistics', in *Australian economic indicators, June,* Catalogue No. 1350.0, Canberra: ABS, pp 3-15.

ABS (2008a), *Information Paper. An Introduction to Socio-economic Indexes for Areas (SEIFA)*, Catalogue No. 2039.0, Canberra: ABS.

ABS (2008b), *Socio-economic Indexes for Areas (SEIFA) – Technical Paper,* Catalogue No. 2039.0.55.001, Canberra: ABS.

ABS (2009), *ANZSCO – Australian and New Zealand Standard Classification of Occupations. First Edition, Revision 1,* Catalogue No. 1220.0, Canberra: ABS.

ABS (2010), *Measures of Australia's Progress 2010,* Catalogue No. 1370.0, Canberra: ABS.

ABS (2011a), *Information Paper. Measures of Socioeconomic Status,* Catalogue No. 1244.0.55.001, Canberra: ABS.

ABS (2011b), *Household Income and Income Distribution, Australia 2009-10,* Catalogue No. 6523.0, Canberra: ABS.

Australian Institute of Health and Welfare (2010), *Australia's Health 2010,* Canberra, AIHW.

Baker, J. and Adhikari, P. (2007), 'Socio-economic indexes for individuals and families', Catalogue No. 1352.0.55.086, Canberra: ABS.

Barnes, G. (2011), *Report on the Generation of the 2010 Index of Community Socio-Educational Advantage (ICSEA)',* Sydney: Australian Curriculum, Assessment and Reporting Authority.

Child Death and Serious Injury Review Committee (2011), *Annual Report 2010-11*, Adelaide: Child Death and Serious Injury Review Committee.

Commission for Children and Young People and Child Guardian (2010), *Annual Report 2010. Deaths of Children and Young People,* Brisbane: Commission for Children and Young People and Child Guardian.

Commonwealth of Australia (2008), *Review of Australian Higher Education. Final Report*, Canberra, Commonwealth of Australia.

Department of Education, Employment and Workplace Relations (DEEWR) (2009), *Measuring the Socio-economic Status of Higher Education Students,* Canberra: DEEWR.

Marks, G. (1998), 'The measurement of socioeconomic status and social class in the LSAY project', *Technical Paper No. 14*, Melbourne: Australian Council for Educational Research.

McKay, S. and Collard, S. (2003), *Developing Deprivation Questions for the Family Resources Survey,* Bristol: Personal Finances Research Centre, University of Bristol.

New South Wales Child Death Review Team (2008), *Trends in Child deaths in New South Wales, 1996-2005,* Sydney: NSW Commission for Children and Young people.

New South Wales Child Death Review Team (2011), *Annual Report 2010,* Sydney: Ombudsman's Office. (And earlier years)

Saunders, P. (2011), *Down and Out: Poverty and Exclusion in Australia,* Bristol: Policy Press.

Saunders, P. and Hill, P. (2008) 'A consistent poverty approach to assessing the sensitivity of income poverty measures and trends', *Australian Economic Review*, vol 41, no 4, pp 371-88.

Saunders, P., Naidoo, Y. and Griffiths, M. (2007), *Towards New Indicators of Disadvantage: Deprivation and Social Exclusion in Australia*, Sydney: Social Policy Research Centre, University of New South Wales.

Saunders, P. And Wong, M. (2012, forthcoming), *Promoting Inclusion and Combating Deprivation: Recent Changes in Social Disadvantage in Australia. Final Report,* Sydney: Social Policy Research Centre, University of New South Wales.

Tomlinson, M., Walker, R. And Williams, G. (2008) 'Measuring poverty in Britain as a multi-dimensional concept, 1991 to 2003', *Journal of Social Policy,* Vol. 37, No. 4, pp. 597-620.

Wilkins, R. (2008) 'The changing socio-demographic composition of poverty in Australia: 1982 to 2004', *Australian Journal of Social Issues*, vol 42, no 4, pp 481-501.

Wilkinson, R. And Pickett, K. (2009), *The Spirit level. Why More Equal Societies Almost Always Do Better,* London: Allen lane.

Wise, P. and Mathews, R. (2011), 'Socio-economic indexes for areas: getting a handle on individual diversity within areas', Catalogue No. 1352.0.55.036, Canberra: ABS.

APPENDIX: Review of the Measurement of SES in Previous CDRT Reports

There is no explicit mention of socio-economic status (SES) in the CDRT Reports published prior to 2003. The 2000-01 Report includes a discussion of the ethnicity and Indigenous status of the household, and of risk factors present in the children's families. The identified factors include parental history of domestic violence, alcohol or drug abuse and criminal activity. The 2001-02 report extended this list to include 'social and economic difficulties, including those associated with relationships, major transitions and financial difficulties' (none of which were defined).

The 2003 report included for the first time an analysis of location and disadvantage over the three years 2000 to 2002. It was noted that:

> 'The Team will continue to monitor the geographical distribution of child deaths in future years to determine whether these patterns persist over time. Further examination of areas showing high death rates will be undertaken to understand the social and demographic factors that operate in these deaths if the current pattern is demonstrated in future years'

The 2003 and the following year's report (2004) used the Index of Relative Socioeconomic Disadvantage (IRSD) component of the ABS SEIFA index. For presentational purposes, areas were separated according to their IRSD ranking as follows:

- Lowest decile (<10%)
- 10^{th} to 25^{th} percentile
- 25^{th} to 50^{th} percentile
- 50^{th} to 75^{th} percentile
- 75^{th} to 90^{th} percentile
- Highest decile (>90%)

After presenting these figures, it was noted that the relationship between child deaths and socioeconomic disadvantage 'may not be a simple one' and that this issue would be explored in a separate study of trends in child deaths covering the period 1996 – 2005. (see NSW CDRT, 2008).

With some minor modification, this approach continued until 2008, when use of the IRSD was discontinued because of concern over its comparability over time and because it relates to the SES status of areas not of the individuals that live there.

Instead, information on the following four socioeconomic indicators was presented:

1. The proportion of the population 15 years old and older employed
2. The proportion of the population with a post-school qualification (above the Higher School Certificate or an Australian Qualifications Framework Certificate II qualification)
3. The proportion of the population that are Aboriginal or Torres Strait Islander
4. The proportion of the population aged 17 years or younger

Areas were classified according to their degree of socio-demographic disadvantage by separating the first, second and fourth of these variables into the lowest decile, the 10th to 90th percentiles and the highest decile, and separating the third variable into the lowest quartile (25%), the 25th to 50th percentile, 50th to 90th percentile and highest decile.

The 2010 Annual Report provides a breakdown in all tables that show the Demographic and Individual Characteristics of deaths of children into 'high' and 'low' socioeconomic status, defined on the basis of whether the area is above or below the median of the distribution of IRSD scores for NSW. The report noted (p. 117) that because the IRSD is defined for areas, 'there is substantial uncertainty about the actual socioeconomic status of the families of children who died' and that this uncertainty 'cannot be eliminated by using more or less specific indices defined by geographic area'.[19]

A comparison was presented (in Figure 25) of the distribution of IRSD scores for NSW with those of the 2010 CDR sample. It showed that the CDR sample was more heavily concentrated between scores of 950 and 1,000 (just below the mean, which by construction is equal to 1,000), with a greater proportion of scores at around 850 and 750. On this basis, it was noted that the CDR sample was not representative of the NSW population as a whole, and also (p. 118) that: ''Improving the measurement of socioeconomic status is a priority for future reports'.

[19] It was also noted that the use of IRSD rather than the Index of relative Socioeconomic Advantage and Disadvantage (IRSAD) was justified because the IRSAD includes a variable that reflects the level of mortgage payments, which is 'likely to depend upon the average time of residence in the area'.