

A Word of Caution on ELSI
By Stuart Birks
Centre for Public Policy Evaluation
Massey University, Palmerston North
26 July 2006

Abstract

The Economic Living Standard Index (ELSI) is an example of a class of measures which are increasingly being developed and used by researchers. The short form of the index is described in detail in (Jensen, Spittal, & Krishnan, 2005). While it is tempting to draw parallels with price indices, care must be taken in so doing. The legitimacy of the index construction process and the interpretation of the resulting values may be specific to the particular index being considered. Here I discuss aspects of this from first principles. Clear grounds for caution are indicated.

The construction of an index requires three steps: i) the constituent items have to be selected; ii) weights must be assigned to these items; and iii) values for each of the items must be determined. Unless there is a specific basis for each step, it may be possible to construct numerous indices meeting the same criteria, but giving different results.

Consider a price index such as the consumer price index. A “representative” bundle of goods and services is selected based on the spending of consumers. People may have their own unique patterns of expenditure, but we can nevertheless identify some average or representative individual or household. For that unit we can then determine, at a particular time, what goods and services are purchased, and in what quantities. The choice of bundle for calculation of the index can be challenged on the basis that it is not representative, and it can be verified as suitable by observing spending behaviour. The items are the goods and services selected, the weights are the quantities or volumes of these goods and services, and the values are the prices that are observed. These values are cardinal measures, thus it is legitimate to make statements such as price X is twice price Y.

The three steps, when applied to construct a price index, yield index values which have a clear meaning. Namely, the values give the relative cost of purchasing a specified bundle of goods and services, a bundle that has been determined to be representative of spending patterns, compared to the cost in the base year.

In recent years there has been a proliferation of indices that have been constructed in an attempt to make comparisons over time and space. The Economic Living Standard Index (ELSI) is one such index. The manual for the short form of the index describes it as follows:

The Economic Living Standard Index Short Form (ELSI_{SF}) is a survey tool for measuring people's economic standard of living. Economic standard of living refers to the material aspect of wellbeing that is reflected in a person's consumption and personal possessions – their household durables, clothing, recreations, access to medical services, and so on. (The terms “living standards” and “standard of living” are used here interchangeably to refer to the same construct.) The ELSI_{SF} tool yields a score from combining information from a set of items that require 4–6 minutes to administer.(Jensen et al., 2005 p. 1)

The pen and paper version starts with fourteen questions about items or activities each with the same four options for response. Hence the first item is:

Telephone
a <input type="checkbox"/> Yes – have it b <input type="checkbox"/> No – because I don't want it c <input type="checkbox"/> No – because of the cost d <input type="checkbox"/> No – for some other reason

There are six other items, including a washing machine, a good pair of shoes and a personal computer, and seven activities such as visiting the hairdresser, having holidays away from home and having family or friends over for a meal.

These are followed by eight economizing measures with three options for each. Hence:

Gone without fresh fruit and vegetables to help keep down costs
A <input type="checkbox"/> Not at all B <input type="checkbox"/> A little C <input type="checkbox"/> A lot

Other measures include “Spent less time on hobbies than you would like to help keep down costs” and “Done without or cut back on trips to the shops or other local places to help keep down costs”.

The remaining three questions are:

Generally, how would you rate your material standard of living?
a <input type="checkbox"/> High b <input type="checkbox"/> Fairly high c <input type="checkbox"/> Medium d <input type="checkbox"/> Fairly low e <input type="checkbox"/> Low

Generally, how satisfied are you with your current material standard of living?

- a** ☐ Very satisfied
- b** ☐ Satisfied
- c** ☐ Neither satisfied nor dissatisfied
- d** ☐ Dissatisfied
- e** ☐ Very dissatisfied

How well does your (and your partner's combined) total income meet your everyday needs for such things as accommodation, food, clothing and other necessities? Would you say you have not enough money, just enough money, enough money, or more than enough money?

- a** ☐ Not enough
- b** ☐ Just enough
- c** ☐ Enough
- d** ☐ More than enough

The responses are then scored. For the first fourteen questions, responses “No, because of the cost” score zero, all other answers score one. For the next eight questions, “Not at all” scores 2, “A little” scores 1, and “A lot” scores 0. Finally, of the last three questions, the first two score a-e from 4 down to 0, and the third scores a-d from 0 up to 3.

These give a maximum score of 41. Somewhat surprisingly, anyone scoring below 10 is then assigned a score of 10 (“to truncate the outliers”), after which 10 is deducted from all scores, so that everyone then will score somewhere between 0 and 31. The final step is to classify into intervals as follows:

Score ranges for the ELSI_{SF}		
ELSI_{SF} score	Living standard level	Label
0 – 8	1	Severe hardship
9 – 12	2	Significant hardship
13 – 16	3	Some hardship
17 – 20	4	Fairly comfortable
21 – 24	5	Comfortable
25 – 28	6	Good
29 – 31	7	Very good

We can now consider this index in terms of the three steps listed at the start of this paper, asking if there is a specific basis for the results of each step:

- i) the constituent items have to be selected

Jensen et al state:

Over time, the item set will age and some items will change their properties and cease to contribute to the measurement of living standards...it is desirable – as with the Consumer Price Index – for the scale's content to be reviewed from time to time to ensure that its validity and discriminating power are preserved. (Jensen et al., 2005 p. 30)

This is true. However, the selection of items for the CPI is based on people's spending habits. It is hard to see an equivalent justification for the particular items selected for ELSI_{SF}. This is important, because we could imagine two or more alternative selections which may yield different results, but over which we would have no basis for claiming that one selection is superior to the other(s). Similarly, we could have two selections, where one contains all the items in the other, plus some more. It may be possible to change the results (in terms of people's relative scores, for example) simply by adding more items.

- ii) weights must be assigned to these items

In the CPI, the weights are expressed in terms of the quantities of the various items included in the bundle. As above, this is based on people's spending patterns, and there is a clear basis for judging whether the chosen bundle is representative. There is no equivalent basis for weighting the different questions in the ELSI_{SF} questionnaire. For example, is there any special reason why the first 14 questions all have equal weight? Moreover, the relative weights of the last three questions could be reduced by expanding the initial group of items by a further 10. The last three questions could be omitted completely, or an additional five questions with scores between zero and four could be added. There is no a priori reason why any one selection should be preferred over another.

- iii) values for each of the items must be determined

For the CPI, the values assigned are the prices that are observed. Once again, there is a clear basis for these values, and a precise interpretation of the results of the calculations (the cost of purchasing the specified bundle, or, more precisely, the cost in comparison to the cost of that bundle in the base year). There is no such interpretation for an ELSI_{SF} score, either with the short or the long form of the index. Is there any particular reason why each of the first 14 questions should have values of zero or one? We could even ask about the interpretation of the questions. Someone deciding an overseas holiday is not feasible, and so opting for a holiday in a local bach could state that they did not want an overseas holiday (score of 1), or see it as an option that they cannot afford (score of 0). Those who have come to terms with their current circumstances would be recorded as

having higher living standards than those in identical circumstances who have expectations that are higher (and presumably unrealistic if they are not affordable). Similarly, there is no a priori reason why the answers of the last questions should be scored 0-4 or 0-3 in increments of one, rather than changing in twos or fives, say.

In other words, at each of these steps, if we are to ask, why has the questionnaire been constructed in this particular way, there is no firm answer that can be given, and no unique form that can be determined. It may therefore be possible to construct numerous indices meeting the same criteria, but giving different results.

Jensen et al state “One of the goals in the development of the ELSI_{SF} was to develop a scale that closely replicated results that would be obtained if the full ELSI scale were used”.(Jensen et al., 2005 p. 29) While this is desirable in that it provides a more economical measure than the full scale, it does not overcome the legitimacy problem, as this applies to both measures. We do not know whether the results for particular individuals would be similar on both measures, as all that was found was a similarity in “mean scores and spreads for the population and for important population subgroups”.(Jensen et al., 2005 p. 29)

One reason for similarity in results could be that the individual components are correlated. If so, then it would be more transparent, and no less meaningful, to consider one or two indicator variables.

Particular problems arise if we try to treat these indices as cardinal measures, as in regression analyses. In linear regression, we would be estimating the relationship between a unit change in one variable and change in another variable. This change is assumed to be constant, irrespective of the absolute values of the variables. Hence, if an ELSI_{SF} measure is used as an independent variable, the impact of a change from 5 to 6 is assumed to have the same effect on the dependent variable as a change from 25 to 26. In multiple regression without interaction terms, the magnitude of the impact is also assumed to be constant irrespective of the values of all the other independent variables. This is a restrictive assumption in the best of circumstances, but the absolute values of ELSI_{SF} are a product of its construction, and there is no particular reason why one construction should be chosen over another. The significance of this cannot be overstated.

Consider a common research finding, namely that there is a positive, or an inverse, relationship between two variables. This is often taken to have policy significance. Now imagine a variable such as ELSI_{SF} where, for argument's sake, we shall assume that the ordering is correct. In other words, someone who really has a higher living standard also has a higher ELSI_{SF} score. By stretching sections of the scale and compressing others, it may be possible to reverse the sign of the relationship.

Here is a crude illustration with synthetic data and a small number of observations, as set out in Table 1. If we start with Y1, the points are distributed symmetrically around the X axis and, unsurprisingly, no relationship is found between X and Y. Y2 is a transformation of Y1, where positive Y values are divided by 5 and negative values

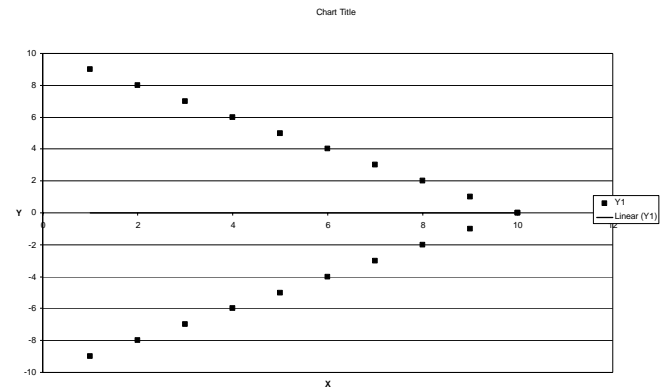
multiplied by 5. Note that this does nothing to change the ordering of the Y values. The result, shown in Figure 2, is a statistically significant, positively sloped trend line and an R^2 of 0.198. Y3 is also a transformation of Y1, but the reverse of that for Y2. Consequently the R^2 is the same, but this time the slope is negative, as seen in Figure 3.

Table 1

X	Y1	Y2	Y3
1	9	1.8	45
2	8	1.6	40
3	7	1.4	35
4	6	1.2	30
5	5	1	25
6	4	0.8	20
7	3	0.6	15
8	2	0.4	10
9	1	0.2	5
10	0	0	0
1	-9	-45	-1.8
2	-8	-40	-1.6
3	-7	-35	-1.4
4	-6	-30	-1.2
5	-5	-25	-1
6	-4	-20	-0.8
7	-3	-15	-0.6
8	-2	-10	-0.4
9	-1	-5	-0.2
10	0	0	0

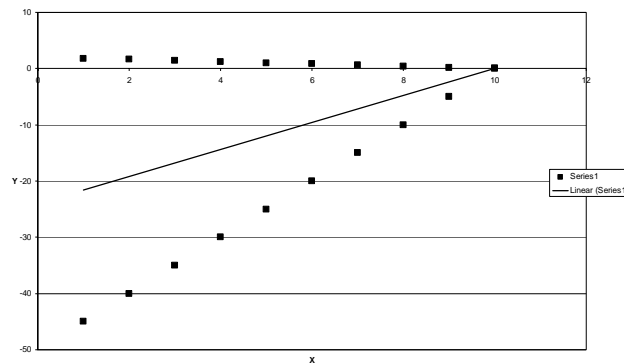
$$Y1 = 0 + 0X$$

$$R^2 = 0.0$$



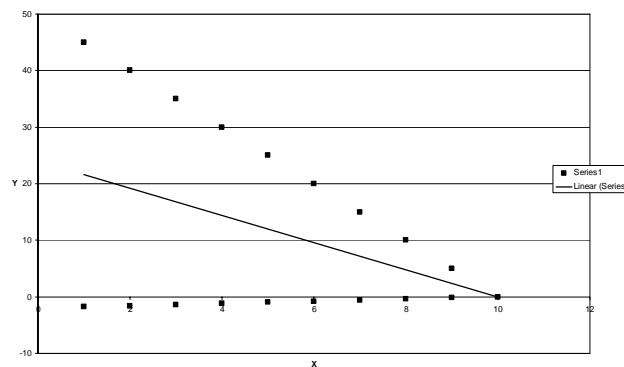
$$Y2 = -24 + 2.4X$$

$$R^2 = 0.198$$



$$Y3 = 24 - 2.4X$$

$$R^2 = 0.198$$



The fundamental conclusion to draw from this is that quantitative analyses that treat indices as cardinal measures should only be undertaken on indices that do actually have a cardinal basis. More generally, we must be very careful about interpretation of research findings depend on indices.

Nevertheless Jensen (2005 p. 2) suggests use of $ELSI_{SF}$ as a cardinal measure in at least two and possibly three of the four suggested broad research purposes, namely: examination of the effects of living standard differences on other outcomes; explanation of living standard variation; and possibly evaluation. Cardinality is not so important for description, the other suggested use.

In contrast to $ELSI_{SF}$, the New Zealand Index of Deprivation, NZDep, is presented in what is specifically described as an ordinal scale, with values 0-10, and as an interval score derived from a principal components analysis. This latter is also the basis for the ordinal scale. Consequently it can be argued that there is a statistical basis for the choice and weighting of components, and there is recognition of a lack of cardinality. (Salmond & Crampton, 2002) There may still be an issue about the specific values (as distinct from the weights) assigned to each component, however.

The NZDep2001 users manual includes the following warning about the ordinality of the measure:

- *If you are comparing two (or more) groups (eg fully immunised versus not fully immunised; or cot death cases versus control babies) compare the distributions of 10 scale values (or principal component scores) using a non-parametric test (since the scale values are ordinal, and the principal component scores are skewed, and may be more skewed in your dataset).*
- *If you are comparing rates of events with deprivation (eg mortality rates in a region compared across the ten deprivation scale values) you could calculate a rank correlation coefficient, or simply plot your results.*

(Salmond & Crampton, 2002 p. 10)

Nevertheless, while they caution, “Population weighted average scores and their decile scale values for census area units should be avoided where possible”, this is not because of ordinality, but rather, “as they disguise heterogeneity within census area units”. (Salmond & Crampton, 2002 p. 11)

Further caution is urged on p.13:

Can I compare NZDep scores between different censuses?

Area comparisons at the meshblock level, over time, should not be attempted.

Comparisons at a higher aggregation, such as Territorial Authorities, or perhaps Area Units, may be less fraught, but we would still urge great caution in the interpretation of changes from one area to another.

Comparing relationships between deprivation and another variable, over time, may be less fraught, but we would still urge caution

In conclusion, care should be taken when using any index. It may be that conclusions have been drawn from studies based on an assumption that indices can be used as cardinal measures.

References

- Jensen, J., Spittal, M., & Krishnan, V. (2005). *ELSI Short Form: User Manual for a Direct Measure of Living Standards*. Retrieved 21 July, 2006, from <http://www.msd.govt.nz/documents/work-areas/csre/elsi-short-form-manual.doc>
- Salmond, C., & Crampton, P. (2002). *NZDep2001 Index of Deprivation: User's Manual*. Retrieved 21 July, 2006, from [http://www.moh.govt.nz/moh.nsf/Files/phi-users-manual/\\$file/phi-users-manual.pdf](http://www.moh.govt.nz/moh.nsf/Files/phi-users-manual/$file/phi-users-manual.pdf)