

## Regional Socio-Economic Indicators – External Review

As part of an evaluation of the methodology and results of the Regional Socio-Economic Indicator Project, advice was sought from various researchers in academic institutions across Canada as well as federal government departments. The external reviewers were asked to comment on:

- The variables used in the construction of the indices
- Validity of the data
- Reliability of the data
- Weights used in estimating the indicators
- General views on the use of indicators for public policy.

Of the ten external reviewers who were contacted in early January 1999, four responded with written comments at the time this report was prepared. They were:

- John Gilbert, Ph.D., Professor, Coordinator of Health Sciences, University of British Columbia
- Ian McKinnon, Pacific Issues Partners, Victoria
- Garnet Picot, Director, Business and Labour Market Analysis, Statistics Canada, Ottawa
- Michael Wolfson, Director General, Institutions and Social Statistics Division, Statistics Canada, Ottawa

A summary of the comments and responses to these comments follows.

<b>Comment</b>	<b>Response</b>
The population estimation methodology used to generate current population estimates by sub-provincial regions is statistically sound and applicable to normalize the various data variables.	No response necessary.
It is not always clear how the weight judgements were made.	This critique was expressed by a number of the reviewers and is addressed later in this report.
Why was the prevalence of smoking not included in the Index of Health given that we know more about this than we do about suicides?	The data was not available by Regional District (only Health Unit which is an aggregation of RD's) and there are currently no plans to repeat the survey. If the survey were ever repeated, we would try to get RD data; and definitely would use it in the Index.

<p>The data contained in the Index of Crime may be subject to reporting bias.</p>	<p>The position of the Attorney General Ministry is that there are no variations in the “degree of policing” across the province. To ensure we minimize reporting bias, we really need to get the serious crimes broken out and either drop the minor crimes from the index altogether, or at least give them a minimum weight.</p>
<p>Why is spousal assault rate not included in the Index of Crime?</p>	<p>First, spousal assault is included in the “violent crime” measure, which is included in the Index. Second, we do not split it out and give it its own weight because this is one crime that we know is subject to reporting bias (as opposed to a policing bias) among age groups, different cultures, etc. If we were able to separate out the more serious spousal assaults and observed that this indicator was not highly correlated with violent crime, then it would be appropriate to give it its own weight -- presumably a heavier weight than random violent crime.</p>
<p>As the indices are not designed for temporal analysis, concern was expressed with regard to the potential misuse of the index by those who would attempt to apply a time dimension to the results.</p>	<p>Good point and it should be footnoted on each table. This concern was also expressed by one of the Statistics Canada reviewers as well.</p>
<p>Subject to concerns with respect to why certain weights have been applied, the use of these particular indicators for public policy decisions is a sound idea.</p>	<p>No response.</p>
<p>The initiative is an important one and useful to policy makers within the government.</p>	<p>No response.</p>
<p>While it is reasonable to look at a summary measure for specific issue areas, I would be very leery of trying to summarize to a single figure. The decision about variables to include, weighting and scaling are all fundamentally arbitrary.</p>	<p>True, the weights reflect a value judgement on the part of the creator, which is a potential problem that we have known from the beginning. Our hope is that the policy makers who use the index also share this value statement.</p> <p>The alternative is to give decision-makers large volumes of data covering many variables across many regions and expect them to effectively synthesize the results into a coherent response. In the end, a final single summary of the data will be produced, either formally through this process or informally in the mind of the decision-maker.</p>
<p>The current geographic base for the data (i.e. regional districts) probably restricts its policy relevance. The disparity in population size between the smallest and largest region is too great. Also, the larger population centers such as the GVRD are not homogenous in terms of their socio-economic make-up. A better set of regions would be the new Health Area geography.</p>	<p>Disparity in population size is not a problem if the data variables are measured in terms of rates, AND the “sample size” is such that the rate is stable. The critique holds true only for regions with unstable rates due to a small number of observations.</p> <p>The lack of homogeneity in some regions is a valid concern. Examples include the GVRD and Squamish-Lillooet.</p>

<p>I think that the computational technique used for calculating the scores is a sound one. However, the results are dependent on the geographic divisions used to analyze the data.</p>	<p>True, but only to the extent that there does not exist homogeneity within the regions selected. A different ranking would likely exist if health region geography were used. However, with the exception of flagging some “hot spots” in downtown Vancouver, the overall regional picture would likely remain similar.</p>
<p>The heavy reliance on income assistance data within the Economic Hardship Index implies that it cannot be used for analysis across time with any confidence. The principal reason is that increases or decreases in the rates can be attributed to administrative changes such as eligibility rules, and hence, may not reflect social or economic conditions.</p>	<p>The cross-sectional nature of the index was stated in the introduction to the report. The measure as designed is a cross-sectional one only, and cannot be used for time-series analysis.</p>
<p>Economic hardship is a multi-faceted phenomenon that includes employment, income and transfers and we should seek to measure more than IA.</p>	<p>We now have the MSP data for 1998. We will add the per cent of population requiring an “MSP subsidy but not on IA” to the June version of the indicators as a measure of the “near poor”. This indicator likely will be included as part of the index at a lesser weight than the IA.</p>
<p>The data used in the construction of the Impending Change in Economic Hardship Indicator are very likely to be too variable month-to-month to give robust indicators.</p>	<p>This is true and is also of concern to us. Our initial sensitivity tests indicated that this “change” variable had the potential to become unstable. Hence, it was modified in later versions, and was eventually given a low weight in the overall index of only 5%. The evaluation contained in our report notes that this indicator will be dropped from the overall Stress Index if the swings cannot be tempered.</p>
<p>The use of dependence on primary industries in the Impending Change in Economic Hardship Indicator may not be appropriate as the data is updated infrequently. In addition, the decline in forestry has been recent and may not be a negative factor in future years.</p>	<p>All true. We had noted this in our report and it was discussed at our last meeting. We may want to include an “economic diversity measure” as background (i.e. unweighted) information and remove the dependency on resource industries variable.</p>

<p>The Impending Change in Economic Hardship Indicator is very narrowly based. Possible improvements may be to:</p> <ul style="list-style-type: none"> <li>• use new IA and EI claims,</li> <li>• capital investment intentions, or</li> <li>• population change.</li> </ul>	<p>We tried change in IA Claims of less than one year but it seemed to be less reliable as a predictor than change in total IA recipients. New EI claims might be worth pursuing and we have the data to do it. However, EI data in the short-run can be misleading, as one cannot distinguish temporary layoffs from the real thing. As we all agreed, this Index is weak and needs work and further monitoring.</p> <p>Regarding the population change, if we could get some indication on a more up-to-date six-month time frame, it would be worth including it. For example, we did not understand why the IA count fell in Skeena/QC, when other indicators were showing no improvement. When the population numbers came out in November it became obvious that an out-flow of population may have been the cause for the improvement in the IA rate. Being able to capture population changes early would be worthwhile as a predictor of impending change; however, considerable additional resources would be required to generate semi-annual or quarterly regional population estimates.</p> <p>Major Project Capital investment intentions are possible. If we do that, we should also consider the phasing out of existing Major Projects.</p>
<p>In general, I like the approach that has been taken to deal with crime statistics. I share BC Stats' desire to break out property crime into categories reflecting the severity of the offence.</p>	<p>We agree.</p>
<p>Due to possible reporting bias for minor offences, I would strongly recommend weighting the index even more heavily towards the rates of the more serious crimes.</p>	<p>See comments above.</p>
<p>I generally agree with the health indicators chosen and particularly like the use of the PYLL.</p>	<p>No comment necessary</p>
<p>Measures of neonatal morbidity and mortality may also be very good summary measures of health.</p>	<p>Infant mortality was discussed as an indicator but because the PYLL is so sensitive to neonatal death we decided the PYLL would suffice. Presumably, neonatal morbidity will be highly correlated to mortality.</p>

<p>Additional data sets that would illuminate important regional differences:</p> <ul style="list-style-type: none"> <li>• Results of the Provincial Learning Assessment Program for Reading, Writing and Math.</li> <li>• Grade 12 completion rates as well as post-sec transition rates.</li> <li>• Census information on English language proficiency or percent foreign born.</li> <li>• If available, cigarette consumption by region, or alternatively alcohol sales by region.</li> </ul>	<p>The Reading Assessment scores from Provincial Learning Assessment Program were included in the Children at Risk Index. These are excellent indicators but they are not updated very often (Note: we were lucky that the reading and writing scores were just updated, but the math scores are as of 1995). Timely data on series that could change from one time period to the next was one of our criteria. We chose the graduation rate overall and the Math 12 and English 12 pass rates for all students who started Grade 12 because that data is updated annually.</p> <p>We already use the number of High School Grads as a per cent of 18-year-olds, which is considered a better indicator of total completion than the Grade 12-completion rate. The graduation rate only measures the completion of those enrolled in Grade 12 and therefore does not capture high levels of Pre-Grade 12 dropouts.</p> <p>We can provide a foreign born variable on the demographic page as background information.</p> <p>We have alcohol sales by RD and we could easily add this indicator. We discussed this at one of our meetings and the general feeling was that the growth in self-brewing/wine making distorted the numbers.</p>
<p>The sensitivity of the results to the selection of the variables and the weights would be useful to get a feel for how the indices are working. You want to ensure that the results are robust and not driven by changes in one or two variables.</p>	<p>True. We have done sensitivity on some of the variables and weights, but the results haven't been formally documented. More work in this area would be beneficial.</p>
<p>A correlation coefficient among the weighted variables could be calculated to support the assertion of independence.</p>	<p>Yes, this can be done although we need to ensure that any correlation is causal and not spurious.</p> <p>It should be noted that one of the reviewers felt that the criteria that the variables chosen should measure different concepts that are not cause or effect seems too stringent, nor was it felt to be a necessary criterion</p>

<p>The fact that the index cannot be used as a temporal measure is a drawback and may result in a potential for misuse of the results. Is it possible to construct an index that is both cross-sectional and temporal?</p>	<p>This is a valid concern. Statistics Canada has had first hand experience with this problem in relation to the LICO being misused as a measure of poverty. No matter what Statistics Canada has done to dissuade this notion, politicians, the media, some analysts and various special interest groups continue to use the words LICO and poverty interchangeably.</p> <p>It may not be possible to convert these Indexes (at least in their current format) into temporal indicators to show change in absolute well being. Because so many of the Indices rely on administrative data, which change according to program requirements, consistent time series would be difficult. We probably should try to look at changes in relative standings -- those regions improving or faltering relative to other RD's.</p>
<p>In the case of the Education Index, the educational attainment of a region reflects, to some extent, the industrial structure of the community. The real issue is the extent to which the educational attainment of the region relates to that required by the jobs in that region (or the jobs that one wants in the region). Hence, the educational level of the "best" region is perhaps not the correct benchmark.</p>	<p>The critique that the quantity of education does not necessarily measure the relevance to the labour market of the education justifies combining the indices of education plus economic hardship together, as we now do in our "Stress Index". We hope the two together will offset divergences in the level of education and the appropriateness of that education for the jobs in the region.</p> <p>In other words, when a region is highly educated (hopefully, we all agree that education, in itself, is a "good", regardless of labour market outcomes) but suffers economic hardship, the combining of the two indices puts a qualifier on the appropriateness of the education. When there is economic hardship and low levels of education combined, we can assume there is stress in the region.</p>
<p>In the overall Stress Index, the Economic Hardship component is probably the most important as it influences the other components of Crime, Education, Health, Children and Youth. Should it not have a larger weight?</p>	<p>Some would say it should have a lower weight because of the high correlation with the other variables. In other words, we are multi-counting the economics of the region through each topic we add.</p>
<p>The Economic Hardship Index seems to focus on those on the bottom end of the income scale, which is reasonable. Is it also reasonable to have an "economic conditions" factor in the index?</p>	<p>Not sure we understand this point. We now have an income inequality indicator in the set of series, but this indicator is not timely so it is not reasonable to use it in the index. We know that the best indicator of current economic hardship would be an unemployment rate, which we don't have by regional district.</p>
<p>Concern was expressed about the selection of the variables and weights reflecting the values of the originator, and whether the population or politicians at a given point share these values in time.</p>	<p>We're hoping that it does. If not, those who object to these values need to give us direction.</p>

<p>In general, these look to be a very thoughtfully compiled set of indicators. The five plus two major domains appear reasonable.</p>	<p>No comment necessary</p>
<p>Whether or not these are the “best” or “most useful” seven domains are probably best determined by publishing them, and then following up on the ways in which stakeholders use them.</p>	<p>Agreed.</p>
<p>The fact that the index cannot be used as a temporal measure is a drawback. It would be a clear advantage if the index could be constructed for both cross-sectional and time trend analyses.</p>	<p>True, the lack of time series comparability is a limiting factor in the use of this index.</p>
<p>Users of the index can draw two kinds of conclusions from the composite index:</p> <ol style="list-style-type: none"> <li>1. my region is K<sup>th</sup> out of the 26, and</li> <li>2. my region is j points behind the leading region.</li> </ol> <p>There is no problem with the first type of conclusion. However, the second is problematic as it is not immediately clear what this means.</p>	<p>Good point. When we sum the weighted index scored to form a composite index, we are assuming that the index scores have a meaning. However, a regional outlier (i.e. the “best” or “worst” value) may overly influence the gap between scores.</p>
<p>An alternative approach would be to scale the raw values using a formula that measured the dispersion from the median and not the maximum value, and normalize by the inter-quartile range and not the overall range.</p> $Index R_j = (R_j - median)/(inter-quartile range)$ <p>This approach is more robust, statistically and preserves the possibility of “cardinal” interpretations.</p>	<p>This is a really good idea. The influence of outliers on the index had concerned us but we couldn’t think of a way to fix it. For the next meeting, we will convert one of the indices to this method and evaluate it.</p>
<p>The median/inter-quartile range approach also has a potential to be used for temporal analysis.</p> <p>At the provincial level, one would show how medians and inter-quartile ranges have changed over time (e.g. is the entire province improving and is the “pack” getting more tightly clustered).</p> <p>On a region level if a score went from 107 to 86 it would be known that it was still above the middle of the pack but falling relative to the degree of dispersion.</p>	<p>The potential to use these data for temporal analysis to monitor whether conditions are improving <u>absolutely</u> in a region may be limited. Changes to programs and processes within government create real problems in the consistency of administrative files. The IA/EI data, which are such an important part of the indices, are a case in point.</p> <p>We agree that this revised index would be a more meaningful way to express changes in the index but we are still looking at relative changes or changes in dispersion as opposed to absolute changes.</p> <p>The example used is difficult to interpret let alone explain to lay persons. It could mean that a region was improving relative to the rest of the group, or the dispersion for the group was diminishing with no idea of how the region is faring relatively or absolutely.</p> <p>Our suggestion is that when we get two solid sets of observations to work with, we try to come up with some kind of a temporal measure that could be readily understood.</p>

<p>The criteria that the variables chosen should measure different concepts that are not cause or effect seems too stringent, nor is it a necessary criterion. In any case, this criterion seems not to have been applied to the actual choices of basic variables.</p>	<p>OK. What we meant to say is that the variables chosen shouldn't be <b>perfectly</b> collinear. Although, we still feel that independence among the data variables is a desirable criteria.</p> <p>As a hypothetical example, if the two variables of property crime rates and violent crime rates were perfectly collinear, there would be no need to go through the expense to collect information on both as one would simply be a linear transformation of the other. Hence, if a summary index on personal safety was to be constructed, only one data variable would be required. In the real world, perfect collinearity rarely exists; however, we do frequently have situations where variables are highly correlated as a result of a causal relationship. In these situations the same logic applies. Why include data variables in the construction of a composite indicator if they are not adding any new information? In the language of Regression Analysis, this would result in the condition of multicollinearity.</p>
<p>I have no comments on the weights. The report is quite clear that these are subjective. Weights are essential to the construction of composite and overall indices; but there are no simple, objective methods to derive them; so what has been done is reasonable under the circumstances.</p>	<p>No comment necessary.</p>
<p>Is the administration of income assistance rates sufficiently uniform across regions such that the indicator primarily reflects low income rather than administrative practice?</p>	<p>Yes, we believe this to be true.</p>
<p>To what extent would regional differences in IA be affected by regional differences in unemployment rates, hence different eligibility duration's for EI.</p>	<p>Because we are measuring economic hardship and not unemployment rates, we don't know whether this comment is relevant. Definitely, the eligibility criteria for EI will vary between regions meaning those living in regions with low unemployment, and hence more stringent criteria for collecting EI, will be forced more readily onto IA. But if we agree that being on EI is not nearly as "economically hard" as requiring IA, then we are measuring the correct concept.</p>
<p>Could you not use income tax administrative data as well?</p>	<p>We could, but it is not as timely. The income tax data from Revenue Canada is two to three years out of date when we receive it. The GIS data, on the other hand, has a three month time delay.</p>
<p>In the case of the Impending Economic Change Index, there may be a potential to draw on Statistics Canada's SEPH database. There are plans to expand this data to the sub-provincial level.</p>	<p>All true. B.C., specifically BC Stats, has agreed to participate in the pilot. However, given existing difficulties with SEPH at the provincial level, we feel that it will be a few years before we see usable data from this system.</p>

<p>I agree with trying to remove “minor” crime from the Crime Index.</p>	<p>Yes.</p>
<p>It would be very useful to include other perspectives on health such as disability. Data from the 1991 Health and Activity Limitations Survey (HALS) may be useful.</p>	<p>Yes, but the data is not updated very frequently. Also, PYLL would likely mirror the disability rates of the region. Premature deaths due to accidents would likely be highly correlated with disabling accidents, high neonatal deaths correlated with genetic birth irregularities, and early death due to disease, correlated with adult illness.</p>
<p>There are well-known concerns about the quality in the coding of suicides on death certificates.</p>	<p>Agreed, they tend to be under reported. Lack of mental illness data is a major shortcoming of the Health Index.</p>
<p>It is not clear why there is a distinction between “natural” and “accidental” PYLL’s, or if it can be done reliably. “Premature” death (e.g. before age 70) is much less ambiguous.</p> <p>If there is an interest in accidents, why not go after it directly with Workers Compensation data?</p>	<p>Much of the “natural PYLL” is defined by infant and elderly deaths while the “accidental PYLL” is more heavily weighted towards the prime age population. Hence, we have made a value judgement that the costs to society associated with the latter type death are higher, and therefore are weighted more heavily.</p> <p>Accidents were desegregated to try and filter out the effects of “non-medical” losses (i.e. things that the Health Ministry can’t do much about but rather are the responsibility of WCB, AG’s, Highways, etc.)</p>
<p>In the case of the Education Index, a more exact measure of the home learning environment is to use the census data on educational attainment of parents of school age and per-school age children as opposed to the population 25 to 44.</p> <p>Another approach to measuring the capabilities of the local labour force would be to measure vulnerability by using the proportion of individuals or parents with low educational attainment. Or counting the total or average years of education completed beyond high school.</p>	<p>We were trying to get a measure of the societal values of the region as a whole, as well as the home environment of the school age population. We don’t think this warrants consideration because of the additional costs involved to buy the series and the likelihood that the two series would be very highly correlated.</p> <p>The point regarding using the “low education” instead of the “high education” may be worth considering. The per cent of adults without completed high school is already an indicator, but is not included in the index. To use the highly educated as the more important measure between the two was an arbitrary decision. While these two indicators are fairly highly correlated, there are some differences in some regions. A good examples of this is Fort Nelson-Liard, which shows a more positive PS completion than the relatively high per cent without HS would indicate (likely reflects a more polarized society). Another example is Kitimat Stikine, with the opposite situation (probably due to Alcan that provides so many high paid low skilled jobs that do not require PS certification).</p>

<p>My preference is to put less weight on process measures like graduation rates and more on outcomes like test scores.</p>	<p>We partially addressed this issue above when the lack of timeliness of the test scores was noted. Further to this, we had hoped that the Grade 12 English and Math pass rates would provide a quality of education dimension to the index as well. These two subjects do not have to be completed in order to graduate. Presumably, the pass rate for these two subjects would be a measure of the proportion of the HS population who have their sights set on University.</p>
<p>What not use low birth weight for infants and hospital admissions for respiratory problems and for injuries within the Children at Risk Index?</p>	<p>All are good points. Variability is a problem in infant birth rates, even though we do a 5-year average. Respiratory problems may be caused by the physical environment and unrelated to infant deaths. Child injury also may be independent of deaths.</p> <p>We should have further discussion with Health on the feasibility of these three indicators.</p>
<p>In the case of the Youth at Risk Index, an alternative to the ratio of high school graduates to 18 year olds is the proportion that has completed high school by age 20. However, the preference is to use performance on standardized test.</p>	<p>We do not agree with the first point. We should be capturing the regional differential in the length of time it takes to get through high school. Our numerator includes all graduates, regardless of age, while the denominator is just a single year age group used to standardize the numerator. The outcome should be similar no matter what age is used.</p> <p>The timeliness of the standardized test scores is an issue.</p>

There were many very good comments made by the various reviewers. Although it is difficult to summarize all the comments into a few statements, we have attempted just that. In general, we feel that the most substantive critiques include:

1. Weights appear to be subjectively determined.
2. An overall composite indicator (i.e. the Stress Index) should not be attempted.
3. Regions should have a greater degree of homogeneity.
4. Make a distinction between severe and less severe crime in the index.
5. Sensitivity analysis of the weights should be carried out.
6. Lack of temporal comparisons could be a problem with users.
7. The construction of the index should be changed to support magnitude interpretations.

1,2 Yes, the weights are subjectively determined. However, we don't feel that this is a sufficient reason not to construct composite and overall indices. This point was noted by one of the reviewers where he states, "Weights are essential to the construction of composite and overall indices; but there are no simple, objective methods to derive them; so what has been done is reasonable under the circumstances."

3. Moving to a geography that reflects a smaller population size for some regions such as the GVRD would likely improve the policy relevance. However, this would come at a cost in terms of time to gather the data, and additional dollars. The improvement in policy relevance is probably best judged by the users of the indices and not by those creating the indices.
4. We would like to have the support of the Ministry of Attorney General in obtaining offence data by severity.
5. We feel that additional analysis and documentation of the sensitivity of the results to the selection of the weights is appropriate.
6. We also agree that the lack of temporal comparability of the index values is a limitation of the current index, and could lead to misuse of the measures.
7. We feel that the index recommended by one of the reviewers has considerable merit. The influence of outliers on the index is of concern. However, the potential to use these data for temporal analysis to monitor whether conditions are improving absolutely in a region may be limited. Changes to programs and processes within government create real problems in the consistency of administrative files. The IA/EI data, which are such an important part of the indices, are a case in point.