



# ROLLING CENSUS

## QUESTIONABLE SCIENCE

by Michael Strube, Ph.D.

A rolling census spreads the measurement over a fixed time interval (e.g., a year with measurement every quarter). This has the advantage of spreading costs and (perhaps) disruption of ongoing practices. But, this approach also has some important disadvantages that need to be considered.

In a rolling census, two kinds of comparisons are usually of interest: (1) comparisons among respondents who took the survey at the same time (e.g., during the first quarter), and, (2) comparisons across time (comparisons of first and second quarter results). The first type of comparison is, essentially, an “apples to apples” comparison because survey responses are gathered under the same conditions. Comparisons across workgroups, for example, are meaningful and may reveal important differences that can be translated into action plans. Unfortunately, because a reduced sample of participants is assessed at any given time point, these within-time-point comparisons can be imprecise. The confidence intervals around the mean responses are necessarily wider than would be true for a one-time full census, and consequently, it is harder to identify important and subtle differences that could need attention.

The across-time comparisons are likewise hampered by the smaller sample sizes at each measurement period. This makes it harder to detect important differences that may be occurring over time. In addition, unless the quarterly samples have been randomly selected, the comparability of the samples is questionable making any differences open to several explanations. For example, differences between the two quarters could be due to important changes that have taken place, or, it could be that the two samples assessed at different time points are different from each other in important ways that have nothing to do with changes over time. In short, without very careful sampling, the across-time comparisons can be an “apples to oranges” comparison.

The advantage of a full, one-time, census is that comparisons among workgroups have maximum sensitivity. These can be used to identify the issues in need of attention, which can then be used to craft selective and targeted pulse-taker surveys that are given at later times to assess change in response to action plans.

If those targeted pulse-taker assessments are given to, for example, all members of a workgroup (i.e., a repeat census of the workgroup) or a random sample of workgroup members, then the comparison over time is valid and speaks to changes that have occurred since the original assessment.

The bottom line is that for a fixed number of responses in a survey there is a particular amount of sensitivity for detecting differences. When the survey assessment is spread out over time, the sensitivity is diluted as well.

Linking internal performance data to survey responses involves building prediction models that identify the key drivers of performance measures. This can be done in several ways, each having distinct advantages. The most convincing approach uses survey responses collected at a particular time (Time 1) to predict performance data at a later time (Time 2), while controlling for performance data collected at Time 1. This produces an assessment of change that correctly takes into account the pre-existing levels of performance when the survey was given and factors out the extent to which the performance data would be expected to change anyway. Structural equation modeling provides a powerful approach to this kind of analysis that takes into account the reliability of the survey responses.

A hidden danger in the usual structural equation approach, however, is the failure to take into account the multiple levels of data in an organization. That is, the respondents exist within workgroups, which in turn may exist within departments, which in turn may exist within locations. Other levels beyond these may exist as well. This “hierarchical” or “nested” nature to the data can produce severe bias in the construction of prediction models if it is not explicitly taken into account in the statistical analysis. The correct approach—called hierarchical linear modeling—correctly identifies the sources of error in prediction that arise at the different levels of analysis, corrects for them in statistically appropriate ways, and ensures that the resulting prediction models are free from bias and bounded by the correct confidence intervals.