

International Census Coverage Discussion

Howard Hogan

US Census Bureau, Washington, DC 20233¹

These have been three interesting and informative presentations. More importantly, the results that they present are the result of three well designed and well executed coverage measurement studies. For those who know the history of census coverage measurement, the sophistication of the designs and the quality of execution of each of these programs is something of which each organization can be proud. It is also something that the statistical community can be proud of, as these three programs are a result of a long history of collaboration.

At the time for 1980 Census rounds, census evaluation based on case-by-case matching was methodologically in poor condition. The 1970 US study, based on a match between the Current Population Survey and the Census, was so badly flawed that the results were never released. The 1980 (US) Post-Enumeration Program (PEP) produced a set of 12 different estimates, based on different data sets and assumptions. These ranged from 1.0 percent overcount to a 2.1 percent undercount, a range too wide to be of much value. The 1990 (US) PES produced an initial estimate of 2.1 percent, but after a processing error was discovered, this was lowered to 1.6 percent. The 2000 US study initially badly underestimated the level of duplication in the census, which led it to greatly overestimate the net error. Only additional study and statistical modeling allowed the Census Bureau to produce this final estimate of an overcount of approximately one half of a percent.

The United Kingdom's program in the 1980s was also flawed. Its evaluation of the 1981 (UK) Census produced an estimate of literally unbelievable census accuracy. Except in a few large cities, it estimated almost no net error, even at the local level. The Canadians have had the most stable and most successful program. They have used the same methodology, the Reverse Record Check (RRC), in every census since 1961, so that 2011 was their eleventh time using essentially the same approach and using it successfully. Even so, the early evaluations measured only census omissions, with no measurement of erroneous enumerations.

I bring up this history not to criticize those who designed and directed these earlier efforts; indeed I was one of the offenders. Rather, I review the history so the reader can comprehend what an outstanding accomplishment each individual study discussed in this session has been, and that all three seem to have been successful.

¹ This discussion is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The views expressed are those of the author and not necessarily those of the U.S.Census Bureau.

I would like to turn in the rest of my discussion to a comparison of the three approaches, first simply because it is interesting, but also it may lead to future collaboration. The authors did not cover exactly the same topics, so there is occasionally “missing data” in my analysis.

	US	UK	CANADA
Method	PES	PES	RRC

As is well known, the US and the UK rely on Post-Enumeration Surveys, while Canada uses the Reverse Record Check. I think that there are three good reasons for these choices. First, the 5-year census cycle makes the RRC possible in Canada, while a 10 year gap may be too large to bridge. Secondly, the US has too many undocumented immigrants to allow for a reverse record check to be successful. Thirdly, the Canadian national administrative record system greatly helps with the successful tracing.

	US	UK	CANADA
Universe	HH Pop	Almost All	All
Samples Size (persons)	390k	500k	70k
Typical Cluster size	30 Housing Units	20 Housing Units	1 person

The universe for the Canadian programs comprises the whole population that was to be enumerated. The UK program covers the household population plus small and median group quarters. Very large GQs (prisons big student halls) were excluded. In a sense, this is a necessary outcome of the RRC as one doesn't know where a sample person will be in five years. It was a design choice for the UK. The 2000 and 2010 US programs were restricted to the household population only. Interviewing, matching and follow-up was considered too difficult in dormitories, jails and nursing homes, approximately 8 million (2.6 percent) of the resident population. The approximately one million over-seas enumerations included in the apportionment counts are also excluded. The 1990 program had included non-institutional group quarters.

Interestingly, although the US and the UK populations are quite different in size, the samples of their evaluation programs were similar. This result is not due the well-known fact that the necessary sample size is largely independent of the population size. Rather, the sample sizes are driven by the need for sub-national estimates. Perhaps this says something about the maximum samples sized that can be controlled during interviewing, matching and follow-up. Canada gets by with a much smaller sample size, although roughly proportional, with respect to the US, to their total population. However, the main difference is that Canada's sample is unclustered. Therefore, they get great precision from the smaller sample.

	US	UK	CANADA
Timing of initial interview	17 weeks	6 weeks	Not Applicable
Follow up interview	9+ months	None	8+ months
Extent of Search	Nationwide	One of 10 Region	Nationwide

The timing of interviewing has traditionally been of great concern to the designer of evaluation surveys. To the extent that the respondent must recall the situation on Census Day, any delay can introduce errors. The four month delay for the initial interview in the US looks especially problematic. However, two methodological advances seem to have lessened the impact.

The 2010 program returned to the approach of asking “who lives here now” rather than “who lived here on Census Day.” This “now” approach was also used in 1990, but not 2000. The problem with the “who lived here on Census day” is that the people most likely to have been missed by the census are also the people most likely to be forgotten by the current residents. The advantage of the “now” approach is that the respondents do not need to recall who lived here long ago, perhaps even the names of the previous residents. The disadvantage has been that matching the current residents who have moved to their census records was time consuming and error prone. However, (1) the fact that the US Census now captures names and (2) advances in computer and computer assisted clerical matching seem to have overcome this weakness. This also allows nation-wide search.

The nine month gap on the follow-up interview raises separate issues. This re-interview is used to verify information gained in the initial interview. For example, one purpose is to identify and exclude fictitious “people,” incorrect names, incorrect dates of birth, etc. gathered in the initial evaluation interview. The gap makes this much harder. Failure to exclude these erroneous evaluation interviews can badly bias the evaluation results up (estimated population too high), but excluding true census misses as “problematic” can badly bias the evaluation results down (estimated population too low).

The Canadian RRC doesn’t have an initial interview, so recall issues are not a problem in constructing the sample. Still, for most of the sample they face a 5 year gap between the reference date of the initial information and census day. Using the many administrative data sources to which they have access and with a variety of searching tools, they are able to match the large majority of sample cases to the Census without field work. Further, the follow up interview is not used to correct the initial sample, merely to gain more information, so the effect is not likely to be as troublesome.

	US	UK	CANADA
Required Information for Matching	Name & 2 characteristics	Name or Dob & One additional item	Name (first or last) & birth month & year
Estimator	DSE	DSE	HT Estimator
Additional Statistical Modeling	Logistic “post-stratification”	Calibration	Non-response adjustments
Correlation bias Correction	DA Sex Ratios	AHE	Not Applicable

One of the problems that bedeviled early evaluation surveys was how to handle incomplete census and evaluation interviews. The tendency was to try to match everything, and then carefully study the unmatched to see which should be excluded as having insufficient information for matching and follow-up. The problem was that only cases that did not match were subjected to this subjective evaluation, creating a strong bias. Again, this problem seems to have been successfully dealt with. Each program has explicit rules for insufficient information for matching.

Both the US and UK relied on dual-systems estimation, with its assumption of independence between the probabilities of being observed in the census and the probability of being included in the evaluation interview. This approach is necessary because the evaluation interviews include coverage issues of their own. However, by carefully building up the RRC frame over several decades, the Canadians can logically argue that they have a complete frame, and need not resort to the independence assumption.

Both the US and the UK have extensive statistical modeling of their DSE results. The US used, for the first time, a method of logistic modeling in lieu of traditional post-stratification. The UK office used calibration. Both the US and the UK include steps to address response correlation bias. In the US, this is done through an assumption based on the sex ratios estimated by its Demographic Analysis program. In the UK, they construct an Alternative Household estimate (AHE) based on estimates from the address register and counts for the Enumeration Area by “Hard-to-Count Group.” Sex ratios were also used nationally to correct for any remaining residual.

In Canada, in order to create homogeneous non-response adjustment groups, they do use some administrative information. They also use a model that assumes that $\Pr(\text{obtaining a RRC interview given the sampled person has not moved and was enumerated in the census}) = \Pr(\text{obtaining a RRC interview given the sampled person has not moved and was not enumerated in the census})$. However, this modeling is much less extensive than that of the UK or US. This is not a criticism. The completeness of the RRC frame lets them avoid much of the complexity introduced by the DSE.

	US	UK	CANADA
E sample	Yes	Yes	Yes
Components of Error Estimated and reported	Yes	Yes	Yes
Housing Unit Coverage Measured	Yes	No	No

All countries now include a program to verify whether census enumerations were correct, a so-called Enumeration or E sample, and all report information on the components of census error rather than just the net error. This is a great advance from early evaluations. The Canadian approach is independent of the RRC and involves matching the census database to itself to identify duplicate enumerations. This can lead to perhaps a small bias, as it cannot identify completely fictitious enumeration or other “out of scope” enumerations (e.g. of non-resident Americans).

The US evaluation also provides estimates of the coverage of housing units. This is important not only for those interested in housing issues, but also for statisticians wanting to draw an address sample from the census or demographers wishing to use a “housing unit method” to construct post-censal estimates. Being completely person-based, the RRC cannot provide this data.

	US	UK	CANADA
Carrying down	Synthetic	Model based	No
Geographic Domain	State, Counties, Places 100k+	Local area	Province/ Territory Large cities
Related Stat Processes	None	Dynamic allocation of Resources	Sample Dwelling Classification Survey

Both the US and the UK use models to carry down the “direct” DSE estimates to the local areas. The RRC is designed to provide direct estimates for the provinces and territories plus large cities. However, their population estimates program, to estimate small domains, does use the coverage studies estimates plus a considerable amount of smoothing, modeling (DA and sex ratios), calibration, and small-domain methods in preparing their estimates.

The Canadian Census uses a coverage survey – the Dwelling Classification Survey (DCS) to provide estimates of the number of occupied dwellings, with the household size distribution, in both the unoccupied and non-response dwelling universes. These

estimates are used to derive adjustments to be incorporated on the Census database via imputing whole households into appropriate proportions of the dwellings with no census response.

	US	UK	CANADA
So What?	May be used to plan next census	One number census	Post/Intercensal estimates correction

How are the results of the census coverage evaluations used? The UK has an explicit “one-number census” where the coverage estimates are immediately and directly incorporated into the official census results. In Canada, the results are incorporated into the official population estimates program, which is used for fund allocation. The US program is strictly to “inform users and help plan the next census,” but the estimates are not incorporated into any official statistics.

With three successful evaluation programs, it is possible to line up the results in percentage terms.

	2010 US	2011 UK	2006 Canada
Net National	0.0	0.0	2.7
Imputed	2.0	5.5	2.9
Erroneous	3.3	0.6	1.6
Not Observed	5.3	6.1	7.2

I borrow the term “Not Observed” from the Dolson paper. I think that it may be less ambiguous than “missed.” Some argue that people who were statistically accounted for through whole-person imputations were not “really” missed. However, the fact that all their data had to be imputed indicates that the observation was highly incomplete if at all.

It is remarkable that the percent “not observed” varies far less than does the net result. For the UK, this is by design. The “one-number census” is designed to account through imputation for all measured errors. The US achieve the same net error by having a combination of imputations, duplicates, and other erroneous inclusions. With higher “not observed” and lower imputations and erroneous inclusions, Canada has a higher net error.

It was not the purposes of these three excellent papers but one cannot help but asking, if one considered the costs versus the results, which census approach is preferable?