Combining Years of Data from a Rolling Sample Survey

Alan Tupek
United States Census Bureau, Washington DC 20233 USA

1. In Memory of Charles (Chip) Alexander

Charles (Chip) Alexander was in the process of preparing this paper when he died tragically in a drowning accident in September 2002. He was 55. I have attempted to provide context for Chip's latest thinking on multi-year averages and have adapted his paper for presentation at this ISI session honoring Leslie Kish. Leslie and Chip collaborated for over a decade on the development of the American Community Survey. Chip was a dear colleague of mine who carried on the proud tradition of Census Bureau statisticians such as Morris Hansen and Joe Waksberg who have contributed much to sample survey methods. Chip is fondly remembered by his colleagues at the US Census Bureau as the "father of the American Community Survey." The Census Bureau is gearing up to fully implement the American Community Survey beginning in the second half of 2004. This paper is about the plans to use the American Community Survey methods as a replacement for the decennial census long form and, in particular, the use of multi-year averages from this annual survey for producing estimates for small domains such as towns, city neighborhoods, and American Indian reservations.

Chip wrote of Leslie Kish's contributions to the development of the American Community Survey*: "The long list of articles by Leslie Kish on the subject of rolling samples clearly demonstrates the intensity and tenacity of his campaign for what he understood as an important idea. The evolution of the idea over the course of these papers also illustrates the depth of his attention to "philosophical" questions about the fundamental quality objectives for a survey: What are we trying to do? How does the choice of survey design relate to what we are trying to do, and why? This kind of guidance is crucial at the start of a survey program, when the "big questions" are being addressed, and makes the difference between ideas that quickly fall by the wayside and those that are "still rolling".*

Leslie's personal support of other statisticians went far beyond his papers. Though I was by no means one of his closest colleagues, he regularly provided personal advice or encouragement when he sensed it was needed. The "still rolling" in this paper's title was the title I used in e-mail messages to him when I had news about the ACS's perilous passage through the annual budget cycle, which was most of the time. He would respond briefly by e-mail, but important messages always came in the form of handwritten letters." Alexander (2001)

2. INTRODUCTION

At the 1999 session of the International Statistical Institute, Leslie Kish urged his colleagues to explore a "new field" of combining surveys from different populations. He saw common aspects to a diverse set of problems. The problems include combining statistics from surveys in different nations, as well as combining data from surveys conducted at different points in time for the same geographic domain. Kish saw similarities to work on methods for combining separate experiments, by Cochran (1937, 1954).

This paper discusses the quality and usefulness of estimates from the planned American Community Survey (ACS), for very small populations groups. The ACS[1] is intended to replace the long form survey in the 2010 census. Since the long form is unique as a source of information about smaller population groups, a priority objective of the ACS design has been to provide good information about smaller groups.

[1] For background information on the American Community Survey, see the U.S. Census Bureau website at www.census.gov/acs

The general premise of the ACS design is that by spreading the "long form" sample across the decade, it is possible to provide updated information for all sizes of population groups. In principle, this should be especially advantageous for small population groups, because there is currently very little information about how these populations change over time. Also, the ACS is expected to have more consistent quality because of the advantages of a continuous operation, which is especially important for those small groups that have traditionally been difficult to include in surveys and collect information about characteristics.

This paper applies these ideas to the problem of combining data over time from the new American Community Survey. This survey will use Kish's "rolling sample" design to facilitate cumulating data over time, to make direct estimates of demographic and housing characteristics for small domains across the U.S. The problem of combining data across multiple years is viewed as a problem of combining information from multiple "experiments." Looking at the problem this way suggests that the appropriate methods for combining the multiple years depend on the purpose of the "experiments" and on the model assumed for the data in the different years.

3. 5-Year Moving Averages for Small Domain Estimates

The ACS was designed to meet the minimum requirements of all the major groups of census data users. Since the early 1990s, we worked with data users to determine minimum requirements for the ACS to replace their need for census long form data.

A critical design decision was to offer a 5-year average as the standard replacement for the full array of long form estimates. Based on discussions over the past decade, we concluded that five years was the longest time period that most data users would accept. Another critical design decision was to follow-up only some of the nonrespondents using personal visits. Since personal visits are many times more expensive than mail and CATI cases, we only followup on 1/3 of the nonrespondents to achieve the most cost-effective operation.

Thus, data users were being asked to accept 5-year averages with less precision than long form estimates due to the 1/3 follow-up of nonrespondents. On top of that, we decided to reduce the size of the ACS initial mailout sample size from about 19,000,000 housing units (1 in 6 decennial census long form sample in Census 2000) to 15,000,000 housing units (ACS sample over five years. Using the effective sample size, the standard error for 5-year average ACS estimates will be 33% larger than long form estimates from Census 2000.

Of course, in turn, data users will get data every year rather than every ten years. Also, nonsampling errors are expected to be smaller for the ACS than for the decennial census long form, since the ACS workload is more manageable and conducted by experienced field representatives. We are currently assessing unit and item nonresponse and other issues affecting nonsampling errors with preliminary indications that we have met our expectations.

For a group of 400 people, the census long form would typically have a 90-percent confidence interval of roughly 280 - 520.[2] An ACS 5-year average would have a slightly larger interval, on the order of 240 - 560. In other words, a typical confidence interval for a hypothetical 2010 census long-form estimate of 400 would be ± 120. By comparison, a 2008-2012 ACS average estimate of 400 would have a 90-percent confidence interval of ± 160.

The basic premise of the ACS rolling sample is that this relatively moderate increase in the sampling error for one part of a decade is a reasonable tradeoff so as to profit from the ability to update the 5-year average every year and thereby gain a picture of the direction of change and relative differences among groups and areas. If the size and characteristics of the population change, such as from 400 to 480, the 5-year average gives a more accurate picture of current conditions than the out-of-date long form statistics.

The ultimate question for users of statistics for small groups is whether the long form's slightly greater precision for comparing groups is of such practical importance that it is worth

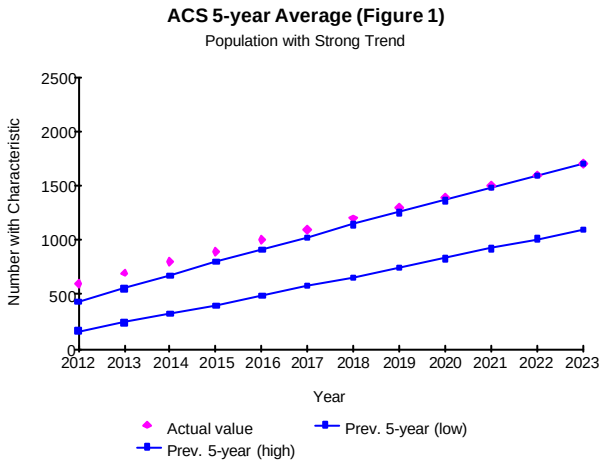
[2] This would be the confidence interval, centered on 400, if the estimate is 400. The actual estimate would not be exactly equal to the population value because of sampling error.

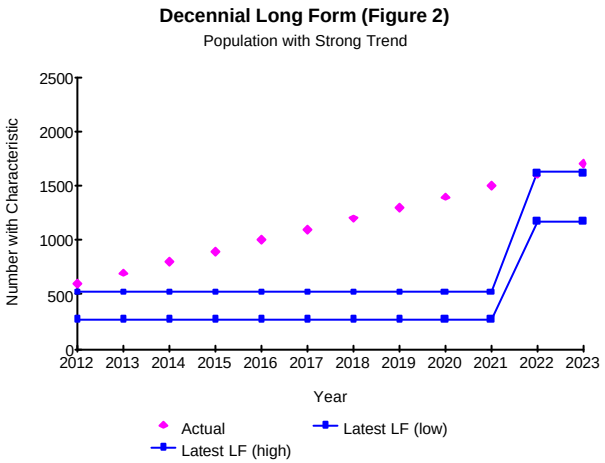
giving up the opportunity to learn about substantial changes in the size and characteristics of the small group over time. The premise of the ACS design is that, for small groups, the ability to learn about substantial changes over time is essential and worth a moderate loss of precision for any single point in time.

For example, consider the potential use of estimates of children under age 5 who speak a language other than English at home in helping school systems prepare and provide for appropriate educational opportunities in coming years. The series of ACS 5-year averages can monitor trends in the number of such children, and the 1- and 3-year averages can detect sudden large changes. By contrast, neither a single decennial estimate nor a single 5-year average, whether 400±120 or 400±160, has the precision or timeliness to be much help in planning. The 2010 long-form statistics will be available in late 2012, in time for planning for the 2013-2014 school year and preferable to a single 2007-2011 ACS average that will also be available in mid-2012. Rather, it is the series of updated ACS averages that would alert school planners more quickly when there are large changes in the needs of children who will be entering the school system and thus better inform their strategic planning.

4. MORE DETAILED EXAMPLES AND DISCUSSION

Figure 1 shows the series of ACS 5-year averaged estimates. Figure 2 shows the long-form estimates for a population that changes from 400 in the year 2010, to 1400 in the year 2020, and the





corresponding probable actual values. . In each figure the “diamond” symbol represents the actual value compared with the estimates from each data set. The solid lines indicate the upper and lower bounds for the probable estimate averaged over the previous five years and then the updated survey estimate each year. The increasing spread between the upper and lower bounds of the ACS estimates in Figure 1 occurs because the number of people with the characteristics is increasing.

The 5-year averages in Figure 1 tend to lag slightly behind the actual population values, and the sampling errors are greater than those for the long form in Figure 2. Yet, the 5-year moving averages are obviously closer to the actual, current population value in most years than for the long form. Unlike the long form, the ACS 5-year averages reflect the direction of the actual trend.

In general, the 5-year averages more closely track the actual trend under most of the scenarios that we examined. With the continuously collected ACS data, it is possible to get considerable information about the magnitude and direction of change over time. Because of the sampling error, however, it will not be possible to be sure of picking up a slight trend, or whether a strong trend is steady or somewhat irregular. The long form, of course, provides no trend information except for two points ten years apart.

How well do the ACS 5-year averages track irregular, seemingly patternless change? This seems at first to be the most difficult situation for interpreting an average. Actually, it is the most natural situation for using an average. Averages would often be used in such situations, even if there were a census every year and there was no issue about sampling error. This is because the

average, over a period of time, provides a more stable description of the area. A decennial snapshot can give a very misleading picture in such situations. Similarly, a single-year spike, where members of a population move into an area for a single year and then move out, would not usually be picked up by a once-a-decade census. ACS information would be preferable, although far from perfect, for that type of scenario.

5. Conclusions

ACS estimates, using 5-year averages are almost always preferable to once-a-decade estimates for tracking changes in small population groups. We've examined a wide range of scenarios of trends and other changes in population groups, though not all were discussed in this paper. About the only situation that favors a decennial census is a population group that increases once a decade just before the decennial census. Even taking into consideration the somewhat higher standard error of ACS estimates, the 5-year averages are more likely to track actual variations under most scenarios.

As we build up a time series of data with the ACS, we may find that estimators other than the 5-year average do a better job of tracking actual variations. It is also possible that other sample design strategies can better measure change.

6. References

Alexander, Charles H. Jr. (2001), "Still Rolling: Leslie Kish's "Rolling Samples" and the American Community Survey," paper presented at the 2001 Statistics Canada Symposium.

Cochran W.G. (1937), "Problems Arising in the Analysis of a Series of Similar Experiments," Journal of the Royal Statistical Society (B), pp. 102-118.

Cochran W.G. (1954), "The Combination of Estimates from Different Experiments," Biometrics, 10, pp. 101-129.

Kish, Leslie, "Space/Time Variations and Rolling Samples," Journal of Official Statistics, Vol. 14, No.1, 1998, pp 31-36. Available at: http://www.jos.nu/Articles/abstract.asp?article=14131

RESUMÉ

À la session 1999 de l'institut statistique international, Leslie Kish avait invité ses collègues à explorer un « nouveau domaine » pour combiner des enquêtes de différentes populations. Il voyait des aspects communs à une variété de problèmes. Ces problèmes incluent la combinaison de statistiques d'enquêtes dans différentes nations, aussi bien que la combinaison des données d'enquêtes qui ont été faits à différents points à temps pour le même domaine géographique. Kish voyait les similitudes au travail sur des méthodes pour combiner des différents expériences, par Cochran (1937, 1954).

Cet article s'applique ces idées au problème de combinaison des données avec le temps de la nouvelle enquête américaine sur les communautés (the American Community Survey (ACS)). L'enquête emploiera de plans de sondage à « échantillon successifs » de Kish pour faciliter l'accumulation des données avec le temps, de faire des évaluations directes des caractéristiques démographiques et logements pour de petits domaines à travers les Etats-Unis. Le problème de combiner des données à travers des années multiples est considéré comme le problème de combiner l'information des « expériences » multiples. Dans ces conditions, l'on pourrait suggérer que les méthodes appropriées pour combiner les années multiples dépendent du but des expériences et du modèle qui est assumé pour les données en différentes années.