Discussion: Topics in Estimation and Modeling for National and International Surveys

Michael P. Cohen

Senior Consultant to NORC
1801 Pennsylvania Avenue NW
Washington DC 20006-3608
MPCohen@Juno.com

Introduction

In essence, this was a session about estimation in surveys. Before discussing the individual papers, I will first make some brief remarks about estimation in general. I will then comment on small area estimation, the topic of two of the three papers. Turning to the individual papers, I first discuss the paper on small area estimation for Navarra, then the one on small area estimation for the Current Population Survey. Next comes the paper on weighting for the American Community Survey followed by a brief conclusion.

Estimation

I want to take this opportunity to comment on estimation as a specialty within survey statistics. Often estimates are just weighted averages so what's the big deal?

In fact, there are many challenging areas for the estimation specialist:
(1) Price indexes
(2) Seasonal adjustment of estimates
(3) Composite estimation and small area estimation more generally
(4) Principal person weighting and other subtle weighting issues

Estimation is a topic of international scope. As this session amply demonstrates, it behooves practitioners and researchers in estimation to pay attention to what is happening throughout the world.

Small Area Estimation (SAE)

SAE is an important topic in estimation and its importance is increasing. The underlying principle of SAE is to "borrow strength" from a domain for which there is more data. This borrowing strength can introduce some bias in the estimates but the substantial variance reduction makes the technique very worthwhile.

One can borrow strength from a larger geographic area or from a longer time period (see diagram). The two SAE papers in this session borrow strength geographically.

In designing a survey to make estimates for small areas, one must consider whether the primary purpose is to make estimates of *level* or of *change*. All too often, surveys are designed as if level were of paramount importance whereas the analytical reports derived from the survey focus on change.

Despite the intense activity in SAE, there is much room for further research in developing estimates, variance estimates, and confidence intervals that are efficient yet relatively insensitive to modeling assumptions.

Direct Estimator	Borrow Strength Across Geography
Borrow Strength Across Time	Borrow Strength Across Geography and Time

Ways to Borrow Strength in Small Area Estimation

Navarra SAE

Let us now turn to the very interesting paper by Maria Ugarte, Ana Militino, and Tomas Goicoa, all with the Public University of Navarra. Navarra is a subnational area in Spain. I am not aware of any work of this caliber by any state in the United States.

A composite estimation method is employed that "shrinks" the estimate toward a synthetic estimate and away from the direct one. The amount of shrinkage is controlled by a parameter called α. The largest value of α considered turned out to be the best. I wondered why the authors did not investigate still larger α values. Also worth considering are composite estimation methods that base the amount of shrinkage on variance estimates. Such methods offer the potential of great gains in efficiency but can be sensitive to outliers or skewness in the data.

I was impressed by the careful and thorough testing of the estimation methods on simulated data. With regard to variance estimation, the authors state: "All the [variance estimation] methods proposed here tend to overestimate the MSE…" For most official statistical purposes, some overestimation of MSE is usually very acceptable. If the MSE estimate is to be used as a component of some other estimate (as in, e.g., components of variance models), overestimation may be as serious as underestimation.

Current Population Survey SAE

The paper by Swamy Paravastu (U.S. Bureau of Labor Statistics), Tamara Zimmerman (U.S. Bureau of Labor Statistics), and Jatinder Mehta (Temple University) on SAE for the U.S. Current Population Survey is especially noteworthy for the clever way that it addresses nonsampling error. It employs a simple time series model to get at the error structure.

A caution: components of variance techniques may lead to delicate estimation problems. This caution notwithstanding, this research is a real tour de force and opens up an extremely promising path. Empirical testing should be the next step

American Community Survey Weighting

Turning from SAE, the last paper by Mark Asiala of the U.S. Census Bureau addresses weighting for the U.S American Community Survey. The challenge is to develop weights that are consistent at the both the household and person levels. This is a fascinating estimation problem. The method Asiala develops seems very reasonable.

Weighting of the U.S Consumer Expenditure Surveys was a major topic of the early 1980s. Zieschang (1986, 1990) and Alexander and Roebuck (1986) are recommended reading.

I would like to encourage more subject matter based analysis. As just one brief example, one would expect weighting up of young adults to increase estimates of the number of housing units.

Conclusion

This session well illustrates the richness of the challenges in survey estimation and the exciting work being done in the area both nationally and internationally.

Acknowledgements: I sincerely thank the authors/presenters not only for their excellent presentations but also for their well prepared papers given to me in advance. Tamara Rib ably chaired the session and kept everything running smoothly. Very special thanks to Kevin Cecco for organizing the session and inviting me to discuss.

References

Kimberly D. Zieschang, Sample Weighting Methods and Estimation of Totals in the Consumer Expenditure Survey, *Journal of the American Statistical Association*, Vol. 85, No. 412. (Dec., 1990), pp. 986-1001

Kimberly D. Zieschang, A Generalized Least Squares Weighting System for the Consumer Expenditure Survey, *American Statistical Association Proceedings of the Section on Survey Research Methods*, 1986, 64-71

Chip Alexander and Michael Roebuck, Comparison of Alternative Methods for Household Estimation, *American Statistical Association Proceedings of the Section on Survey Research Methods*, 1986, 54-63