

## DISCUSSION

Michael Lee Cohen, National Academy of Sciences  
National Academy of Sciences, CNSTAT, Room HA-192, 2101 Constitution Ave., Wash, DC 20418

The five presentations represent a great variety of interesting estimation problems of real importance in federal statistics. Areas of statistical methodology that are made use of include sampling, ANOVA, time-series, and regression. The breadth of problems makes me somewhat unsure of my comments and questions and so they should be considered to be from someone not totally informed with respect to each area and problem.

The first paper was "Some Basic Technical Information About the American Community Survey". In my discussion I will not raise questions concerning the impact of the introduction of this survey on the federal statistical community and other existing surveys, though such questions are important to raise. The paper represents enormous progress since the last presentation I heard on this topic. Much of this is due to the open process that the Census Bureau has had of asking for and seriously considering comments on their plans. Clearly, the current plans indicate successful resolution of many complications raised internally and externally.

I have several questions about the American Community Survey. 1) What will happen to the American Community Survey during the decennial census? Won't there be confusion with a resulting deleterious impact on the response? 2) Has any thought been given to having some of the sample retained from one year to the next to help estimate yearly change? 3) If modeling seems necessary for benchmarking the ACS to the sample (long) form, wouldn't one also need a sample form (possibly reduced in size) in 2010? Some remaining points and questions are: (1) Mail self-administered forms are limited with respect to their ability to tease out subtle concepts, and so ACS is not a panacea, (2) Controlling to intercensal demographic estimates is not clearly the right thing to do. Moving the marginal totals in that direction seems fine, but both the ACS and the demographic marginal totals will have error, possibly comparable -- especially late in the decade, and therefore the margins should be made up of some sort of weighted combination, (3) Users might enjoy a single field that had an arrow that indicated whether the average over years is "hiding" a three or five-year up or down trend (that was statistically significant accounting for sample variances), (4) Has there been any thought to augmenting MAF using DES or Census Plus addresses or weighting factors to help in dealing with either whole household undercoverage or within household unit

nonresponse? and (5) What about use of administrative records? Earlier presentations indicated a broad role for the use of administrative records in ACS.

The second paper was "Estimation of Agricultural Commodities Using Multiyear Area Frame Survey Data." The paper presents an excellent idea of making use of the added information for area sampled farms that appear for consecutive years, being careful to consider the amount of rotation between years, year-to-year correlations, and the optimal number of years to look back. This is a very hard problem.

While the approach used has great promise, it does seem as if there is a wider class of models that could also be examined, at least for comparison purposes. This would include multiplicative models. Graphical methods might help convince us that the linear model was appropriate.

One question is whether one could examine a list approach for the large farms, an area frame for the little farms, and then use a combination of cut-off sampling or certainty sampling for the large farms plus an area frame approach for the small farms? Of course, rare commodities would still be a problem. Then one would use the ANOVA model only on the area frame information to make use of the multi-year information. This might provide cleaner modeling.

Another question is how do you define outliers? How are they treated? Have you considered using robust ANOVA as a way of dealing with them?

Finally, bootstrapping residuals to investigate the validity of model assumptions is good for examining some departures from model assumptions, but not those that affect the residuals strongly. An obvious example is outliers. Instead, in that case one should bootstrap robust residuals. Further, there are more straightforward diagnostics for simpler models that might be adapted for this complicated situation. However, the correspondence of the relative efficiencies between the bootstrap and the model-based approach is extremely impressive and comforting for the validity of the chosen model.

With respect to the third paper, "Estimating Hartley's Dual Frame Design, Case IV," the only question I have

With respect to the third paper, "Estimating Hartley's Dual Frame Design, Case IV," the only question I have is what would you do to estimate parameters other than the population mean? The paper produces a final answer to a hard, important question. I can only offer congratulations for the excellent work.

The fourth paper, "Adjusting for a Calendar Effect in Employment Time Series," represents a very careful job of addressing this hard problem. In fact, it would be difficult to even discover whether there was a problem. I like the idea of looking at three different time spans to provide some model robustness but could one possibly have less overlap of the windows using shorter time spans?

Everyone likes to bounce back and forth between modeling and subject-matter considerations. Some questions related to this are: (1) Do the implications of the models make subject-matter sense? There was some discussion of this but I would have enjoyed seeing more. For example, did the estimated alpha's have a sensible structure, e.g., across outputs? (2) Did it make sense that the industries that were removed from the analysis would not be well-modeled? Shouldn't all industries be subject to calendar effects?

Some final points and questions are: (1) Identifying outliers as those points  $3.5\sigma$  away from the center is not a great rule unless  $\sigma$  is estimated robustly; (2) Are there any true values that could be used to help us evaluate this model? Are there any censuses? A related question is that when adjustment did "harm" by reducing smoothness, is it possible that smoothing eliminated a real phenomenon?; (3) Finally what did the estimated covariance matrix look like? Did the correlations make intuitive sense?

The paper finished with a great list of subtle points that concluded an excellent answer to a hard and important problem.

The final paper was "Forecasting Wholesale Gasoline Prices". I am comforted that here we have an evaluation based on true values. The transfer function model seems very appropriate for this problem and is obviously doing an excellent job. If users want to do better they might help EIA buy DTW data.

It would be fun to examine some competing approaches, for comparison purposes. One possibility would be to model the ratio of the price to some weighted average price, rather than the price itself. This leads to the first question: What is the correlation

of the three prices of interest?

Another question that occurred was whether you could use early returns of the 782 in your model? You don't seem to in this model and yet some of these might be available.

Being closer than \$.02 is a good loss function, since it was motivated by user concerns, but it is useful to mention that it has a minor disadvantage in that when a forecast isn't within \$.02, it could either be quite close to .02 or it could be quite large. Therefore, I was glad to see use of the mean absolute error as well.

I don't understand why different PADD's need to be modeled differently. Is there a subject-matter reason? In this situation also the examination of the model's assumptions using subject-matter considerations might pay-off.

Again, all five papers are great examples of the careful and hard work that is going on in federal statistics to solve important and challenging problems. It was a pleasure being asked to discuss them.