SAMPLING AND ESTIMATION FOR THE FDIC'S ASSET LOSS RESERVE

David W. Chapman, Terry L. Kissinger, and Henry S. Foster, FDIC
David W. Chapman, FDIC, Rm. 2074, 550 17th St., N.W., Washington, DC 20429

KEY WORDS: Optimum Allocation, Synthetic Estimator, Bootstrap Variances

1. Introduction

As part of the preparation of its annual financial statement, the Federal Deposit Insurance Corporation (FDIC) conducts an annual assessment of assets it has acquired from failed financial institutions. Because of the effort involved with valuing assets, these assessments are based on the review of a probability sample of these acquired assets. Estimated asset recovery values, which are based on sample valuations, are compared with associated liabilities to estimate the total liability to the FDIC insurance funds. This process is referred to as the Asset Loss Reserve (ALR).

The sample of assets selected for the ALR is a stratified random sample, where the strata are defined by (1) the insurance fund (two types), (2) legal classification (two categories), (3) asset type (several categories), and (4) asset size (three categories). The sample design includes certainty strata for assets with the largest book values, and for assets for which the book value is unknown. The sample size for the non-certainty portion of the sample is allocated to strata using Neyman allocation.

To obtain an estimate of the liability to the insurance fund, estimated recoveries must be made for each asset in every receivership (failed financial institution). This is because the calculation of liabilities depends on the recovery amounts at the receivership level. However, the sample is not designed to allow precise estimates at the receivership level. Except for a few receiverships, the sample is not large enough to support direct receivership estimates. Therefore, a type of synthetic estimator is used to estimate asset recoveries for a receivership, by applying the sample recovery rate for a stratum to each asset in the corresponding stratum in the receivership. The estimated liabilities for receiverships are summed to obtain the fund-level liability estimate. Because of the complexity of this estimator, the variance is estimated using a bootstrap (resampling) procedure.

This paper describes the design and allocation of the ALR sample, and the methods used for selecting the sample, estimation of the liability to the insurance fund, and estimation of the precision of the liability estimator. It concludes with a discussion of future research.

2. Sample Design and Allocation for the ALR

The universe for the ALR sample is all FDIC assets obtained from failed institutions, as of a specific date, except for those few assets that were acquired from financial institution failings covered by the Federal Savings and Loan Institution Corporation Resolution Fund (FRF). (For the few assets remaining that are covered by FRF, which is no longer an active insurance fund, the assets are valued using a non-statistical methodology.) There are several sources of the ALR asset universe that are merged to create the sampling frame. The 2000 ALR sampling frame contained 731 assets.

To try to improve sampling efficiency, the assets are stratified prior to sampling into relatively homogeneous strata, defined by the following variables:

- Insurance fund type: BIF or SAIF. These are the active funds that cover deposits in financial institutions: the Bank Insurance Fund (BIF) or the Savings Association Insurance Fund (SAIF).

- Two legal classifications: Receivership or Corporate Purchased. After a financial institution fails, all of the assets acquired by the FDIC are receivership assets. However, over time the FDIC will sell these assets. At some point, there may be only a few assets left in a receivership. To allow the receivership to be closed out, the FDIC may purchase the few remaining assets, which then become corporate purchased.

- Type of asset: Several asset type categories are defined, depending on the fund type and legal classification. Examples of asset type categories are consumer loans, commercial loans, mortgage loans, and real estate loans. The maximum number of asset types for a fund type/legal cross-classification is seven.

- Asset size (book value): Within each cross-classification of the first three stratification variables defined above, two or three size categories are defined in terms of book values: small, medium, and large/certainty. The size cutoffs defining these categories vary, depending on the size distribution of the assets in the various cross-classifications of the first three stratification variables.

For the 2000 ALR, there were 23 noncertainty strata. Since the asset population can change substantially in a given year, the strata definitions are reviewed each year, and revised as needed.

The total sample size is determined as the minimum sample size needed to estimate the total recovery amount for the assets in each of four groups, defined by the two insurance funds (BIF and SAIF) and the two legal classifications (receivership and corporate purchased), to within ± 10% with 95% confidence. It would be preferable to design the sample to meet a precision target for estimating the insurance fund liability amount. However, because of the complexity of the estimator of the liability amount, this would not be feasible. Therefore, the precision target is based on the estimator of the recovery amount, which is also entered on the FDIC's financial statement.

The procedure used to identify the minimum sample size needed to achieve the target precision for estimating each of the four recovery amounts defined above is an iterative process. First, the basic estimator of each the four recovery amounts is a separate ratio estimator, $\hat{Y}_{Rs}$. This estimator is given in Equation 6.44 in W.G. Cochran's sampling text (1977), page 164, as follows:

$$\hat{Y}_{Rs} = \sum_{h=1}^{L} \frac{y_h}{x_h} X_h \,, \qquad (1)$$

where

- L = the number of strata,
- y_h = the sum of the sample recovery values in stratum h,
- x_h = the sum of the sample book values in stratum h,
- X_h = the sum of the book values of all assets in stratum h.

For this estimator, the Neyman optimum allocation of the sample to stratum h, is based on the sample size formula given by Cochran (1977), page 172, as follows:

$$n_h = n \frac{N_h S_{dh}}{\sum_{h=1}^{L} N_h S_{dh}} \,, \qquad (2)$$

where

- n = the total fund-level sample size,
- N_h = the total number of assets in stratum h,
- S_{dh} = equals the standard deviation for stratum h of the variable d_{hi}, which is defined as $(y_{hi}-R_h x_{hi})$, where
- y_{hi} = the recovery value of the i^{th} asset in stratum h,
- x_{hi} = the book value of the i^{th} asset in stratum h, and
- R_h = Y_h / X_h, where
- Y_h = the sum of the recovery values of all assets in stratum h.

The ratio estimation standard deviation, S_{dh}, in Equation (2) is estimated from the previous year's ALR data. Then, an initial fund-level sample size, n, is chosen, which may be the sample size used in the previous year. Next, the sample size is allocated optimally to the strata, based on Equation (2). Then, to see if the precision target is met, the variance for the estimated recovery amount is computed, using the following equation (see Cochran, 1977, page 172, Equation 6.71):

$$V(\hat{Y}_{Rs}) = \sum_{h=1}^{L} \frac{N_h (N_n - n_h)}{n_h} S_{dh}^2 \,. \qquad (3)$$

The same estimates used for S_{dh}^2 in Equation (2) (from the previous year's ALR data) are used to compute the variance in Equation (3). The relative precision of $\hat{Y}_{Rs}$ is derived from the estimated standard error of $\hat{Y}_{Rs}$ (square root of the variance from Equation 3) and the estimated recovery amount derived from Equation (1), which is also based on the previous year's ALR data.

The process of sample allocation and checking to see if the precision target is met is repeated, using different sample sizes, until the minimum sample size is identified that meets the precision requirement. Each of the four sample sizes is computed in the same way, and then summed to obtain the total sample size.

If the strata are revised for a particular year's ALR sample, the parameter estimates in Equation (2) needed to allocate the sample to strata cannot be obtained from the previous year's ALR data. Accordingly, the optimum allocation and sample size derivations are computed as though total book value (rather than total recovery value) were being estimated. In that case, the stratum standard deviations used in the allocation process in Equation (2) are book value standard deviations, rather than ratio estimation . Also, the estimated total is a simple expansion estimator, rather than a ratio estimator.

Because estimated recovery rates at the stratum level are critical in estimating the liability to the insurance fund, the "optimum" sample sizes derived as described above are reviewed to be sure that none of the strata have very small sample sizes. As a result of this review, some modest adjustments (mostly increases) are made to the Neyman allocation sample sizes to improve stratum estimates of asset recovery rates.

3. Sample Selection

Once the stratum sample sizes are determined as described in Section 2, the sample is selected using a permanent random number procedure. This approach makes it easy to select the sample each year, incorporating new assets into the universe, and taking into account the switches that some assets make to a new stratum. With this approach, many of the same assets are valued in subsequent years, which has some cost advantages.[1]

The permanent random number method of selecting the sample is carried out in the following three steps:

(1) Each asset is assigned a unique (permanent) random number when it enters the ALR universe, which it retains as long as it remains in the universe.
(2) Assets within each stratum are ordered by their random numbers, from smallest to largest.
(3) The designated number of sample assets for each stratum (as determined by the methods described in the previous section) are selected consecutively, beginning at the top of the ordered list.

In subsequent survey periods old assets will drop out of the population and new assets will come into the population. Each new asset will receive its own permanent random number. Assets which move from one stratum to another will retain their original random number. Neyman allocation will continue to be used to estimate optimum strata sample sizes. The assets in each stratum will then be sorted by random number so that new or moved assets will assume their appropriate position in the stratum list. As in previous years, the designated number of sample assets for each stratum will be selected from the top of the ordered list.

With this method, it is possible that an asset selected in a previous survey period may be "displaced" from the sample if new/moved assets with lower random numbers come into the stratum, or if the stratum sample size is decreased. Such an asset may also return to the sample in a subsequent survey period. However, for the selection of any ALR sample, the assets in each stratum each have the same sample inclusion probability.

4. Estimating the Liability to the Insurance Fund

The FDIC financial statement includes estimated liabilities for the BIF receivership assets and the SAIF receivership assets. These liability estimates are complex because of the need to make net liability calculations at the receivership level. The estimates of net liabilities for all receiverships are summed to get each fund-level liability estimate. (These estimates are not required for corporate purchased assets because all liabilities belong to the corporation rather than individual receiverships.)

Since the corporation balance sheet requires that net liability calculations be made at the receivership level, one option would be to select a sample from each receivership that is adequate to make the necessary calculations. However, this would require a substantially larger sample size than is practical. Therefore, the design of the asset sample, described in Section 2, ignores receivership groupings. Consequently, it is possible that, for a given receivership, very few, if any, assets are selected.

To make the necessary liability calculations, an estimated recovery amount is needed for each asset in every receivership. Since the asset sample for most receiverships is too small to use to make such estimates directly, a type of synthetic estimator is used which draws from the data collected from all receiverships.

The estimator starts with an estimated recovery rate, r_h, for each stratum, computed as the ratio of the sum of the estimated recovery amounts to the sum of the book values for the n_h sample assets in stratum h. For a given receivership, the estimated recovery rate for each asset in stratum h that was not selected for the sample is set equal to the estimated recovery rate for the stratum, r_h. For each asset, the estimated recovery amount is computed as the book value times the (synthetic) estimated recovery rate.

The net liability for the receivership is calculated from the estimated recovery rates, as defined above, and the liabilities for the receivership. The net

[1] In previous years the ALR process was carried out twice a year (Spring and Fall). With this format the cost savings were substantial when assets were retained in the sample from one period to the next, because a six-month "roll forward" process was used to value an asset that was reviewed in the previous ALR sample. This "roll forward" process cannot be used for periods over six months.

liabilities for all the receiverships are summed to derive an estimated liability at the fund level. The same method is used for both insurance funds, BIF and SAIF.

5. Estimating the Precision of the Fund Liability Estimator

Because of the use of a synthetic estimator to estimate the liability to the insurance fund, it is not possible to derive a closed-form estimator of the precision, such as the variance. Instead, a bootstrap (resampling) estimator of the precision is used.

With this method, the ALR sample is reselected from the original sample, with replacement, a large number of times. Each resample is referred to as a bootstrap sample. Since the ALR sample is a stratified random sample, the bootstrap sample for stratum h is obtained by selecting n_h assets with replacement from the original n_h assets selected for the sample in stratum h. The certainty selections are a part of each bootstrap sample. The number of bootstrap resamples that has been used for the ALR process is 100.

For each bootstrap sample the estimate of the liability to the insurance fund is recalculated, using the same estimation methodology as used to compute the original point estimate of the fund liability. This involves re-computing each of the stratum recovery rate estimates that are critical to the synthetic estimator. In general, there will be a different estimate of the fund liability for each of the 100 bootstrap samples.

From the 100 bootstrap samples a 95% confidence interval is computed for the net liability for each of the two funds, BIF and SAIF. This is done using the percentile method (see Efron and Tibshirani, 1993, pp. 168-169). With this method, the 95% confidence interval is the interval defined by the middle 95% of the distribution of the 100 bootstrap estimates. In general, this will not be a symmetric interval.

This method of constructing confidence intervals has desirable properties, as described by Efron and Tibshirani (1993, pp. 174-176). It is both transformation-respecting and range-preserving, which makes it preferable to some other methods based on bootstrapping. For example, a competitor known as the bootstrap *t* method does not have these two properties (Efron & Tibshirani, 1993, p. 163) and requires considerably more computations.

6. Further Research

Refinements to the existing methodology will be implemented in the ALR process in the future. It can be seen in the statistical literature pertaining to the bootstrap that the bias-corrected and accelerated percentile method (BC_a) generally yields better coverage than the percentile method described earlier. This method involves estimating a bias correction and an acceleration constant. A good discussion of the properties of the BC_a method in comparison to other methods is provided by Chernick (1999, pp. 54-57).

Another refinement that will be implemented is rescaling, as described for complex survey data by Rao and Wu (1988). This will involve using alternate bootstrap sample sizes in the various strata, as recommended in the statistical literature. Finally, more bootstrap samples will be generated to provide better confidence interval endpoints. Chernick (1999, p. 112-114) noted that "bootstrap folklore" suggests that 1,000 or more bootstrap samples be generated for constructing confidence intervals, and that for many applications many more bootstrap samples can be generated, with the computing power available today.

References

Chernick, M.R. (1999). *Bootstrap Methods: a Practitioner's Guide*, John Wiley, New York, NY.

Cochran, W.G. (1977). *Sampling Techniques*, 3rd ed., John Wiley, New York, NY.

Cowan, C.D. (1997), "Using the ALR to Obtain Fund Level Estimates of Asset Recovery and Liabilities." FDIC internal memo, Jan. 2, 1997.

Efron, B. and Tibshirani, R.J. (1993). *An Introduction to the Bootstrap*, Chapman & Hall, New York, NY.

Rao, J.N.K. and Wu, C.F.J. (1988), "Resampling Inference with Complex Survey Data," *Journal of the American Statistical Association*, 83, 231-241.