Sample Design for the FDIC's Asset Loss Reserve Project

David W. Chapman
Federal Deposit Insurance Corporation

Abstract

Each year the FDIC must include in its financial statement the value of assets it has acquired as the receiver for failed institutions. Since it is not practical to value all these assets, a sample of assets is selected for estimating total recovery values. The sampling and estimation process is referred to as the Asset Loss Reserve (ALR) project. There are some special constraints that make it difficult to optimize the design of the ALR sample. As an example, the ALR asset universe is a "moving target" because it is continually changing as assets are sold and banks fail. As a result, the ALR asset universe is defined as of June 30, though the financial statement pertains to December 31, which creates estimation problems. As a second example, the sample design is a stratified random sample, but many of the current strata contain too few assets. This introduces the possibility of bias in the separate ratio estimator that traditionally has been used. There are barriers to making improvements in the sample design or estimation methodology used. Assessments of the impact of three special constraints on the ALR project, and approaches used to address these constraints, will be discussed.

Keywords: Optimum Allocation, Ratio Estimation, Stratification

1. Introduction

As part of the preparation of its annual financial statement, the Federal Deposit Insurance Corporation (FDIC) conducts an annual assessment of all assets it has acquired from failed financial institutions that become receiverships. Because of the effort involved with valuing assets, these annual assessments are based on the review of a sample of acquired assets. Estimated asset recovery values, which are estimated from sample valuations, are compared with associated liabilities at the receivership level to estimate the loss liability to the FDIC insurance funds (i.e., the Bank Insurance Fund, BIF, and Savings Association Insurance Fund, SAIF) associated with these acquired assets. The sampling of assets, valuation of assets, and estimation of the loss liability to the insurance funds are the main components of the Asset Loss Reserve (ALR) project.

The sample of assets selected for the ALR project is a stratified random sample, where the strata are defined by seven types of assets. The sample design includes certainty selections for assets with the largest book values (generally over $1,000,000). The noncertainty portion of the sample is allocated to strata using an iterative optimization approach.

There are some special problems or constraints associated with the design of the asset sample and the development of optimum estimators. Three of these special problems are the following:

(1) The FDIC asset universe is a "moving target" because it is continually changing as assets are sold off and banks fail. This makes it difficult to derive appropriate estimates of the liability to the insurance funds that apply to the ALR asset universe at the end of the year.

(2) The FDIC asset universe has been shrinking over time from about 40,000 assets in the mid 1990s to 765 for the 2004 ALR sample design. Because of the practical difficulties of modifying strata definitions, some strata contain a small number of assets, which introduces the possibility of bias in the separate ratio estimator that has been traditionally used to estimate the total recovery value for the asset universe.

(3) The final estimates of liability to the insurance funds are calculated at the closed bank (receivership) level. However, there are too many receiverships to use them as strata. As a result, the liability calculations require the use of a synthetic-type estimator.

The next section provides a summary of the ALR sample design. Sections 3, 4, and 5 provide detailed discussions of the impact of each of the three special problems listed above. The discussions will include implications for sample design and estimation, and approaches used to address these issues. The final section, Section 6, covers topics for future research.

2. Sample Design and Allocation for the ALR Project

The universe for the annual ALR sample is all FDIC assets acquired from failed institutions (and not yet sold), as of June 30. The universe for the 2004 ALR sample contained 765 assets. In order to enhance the estimates of asset recoveries, the universe is stratified by the following seven asset types:

1. Consumer Loans
2. Commercial Loans (Certainty cutoff: $2,000,000)
3. Securities (Certainty cutoff: $1,000,000)
4. Real Estate Mortgages (Certainty cutoff: $1,000,000)
5. Other Owned Real Estate (Certainty cutoff: $1,000,000)
6. Other Assets—Judgments (Certainty cutoff: $1,000,000)
7. Net Investments in Subsidiaries (Certainty cutoff: $1,000,000)

Each of these strata has a certainty and noncertainty component, except for the first stratum (consumer loans), which contains assets with relatively low book values. The certainty cutoff in terms of book value, which is generally $1,000,000, is given in parentheses in the list above.

The stratum sample sizes are derived to meet a specific (relative) precision target. Ideally this criterion would be based on estimating total liability to the insurance funds for the FDIC's financial statement. However, that is a very complex estimator (discussed in more detail in Section 5), and therefore cannot be used to determine optimum sample sizes. Instead, the target estimator for designing the sample is the estimator of the total recovery value of the asset population, based on a separate ratio estimator, where the covariate is the book value of the asset. This estimator is given in Equation 6.44 in Cochran (1977, p. 164).

Unfortunately, due to small stratum sample sizes and some changes of stratum boundaries from year to year, there has not been sufficient data available to derive optimum stratum sample sizes for the ratio estimator of total recoveries either. Therefore, book value has been used as a proxy for recovery value in the optimization process. The criterion is to derive the minimum sample size for estimating total book value of the BIF/SAIF receivership assets to within ± 10% with 95% confidence, subject to the constraint that at least two assets are selected from each noncertainty stratum.

If x_{hi} represents the book value of the ith asset in noncertainty stratum h, the estimator of the total book value, X, and the corresponding variance of the estimator, $V(\hat{X})$, are based on Equations 5.4 and 5.10 in Cochran (1977, p. 92-93), as follows:

$$\hat{X} = \sum_{h=1}^{7} N_h \bar{x}_h \, , \qquad (1)$$

$$V(\hat{X}) = \sum_{h=1}^{7} N_h (N_h - n_h) \frac{S_h^2}{n_h} \, , \qquad (2)$$

where $\bar{x}_h \; and \; S_h^2$ represent the sample mean and universe variance for the asset book values for stratum h.

The relative precision of the estimator, based on a 95% confidence interval, is the following:

$$RP(\hat{X}) = 1.96 \frac{\sqrt{V(\hat{X})}}{X} \, . \qquad (3)$$

Because of the constraint that there must be at least two assets selected from each stratum, standard textbook formulas for deriving the minimum sample size needed to meet the precision criterion cannot be used. Therefore, an iterative process is used to obtain the optimum sample size to meet the precision criterion (± 10% at 95% confidence). The initial step in the process is to allocate two sample units to each noncertainty stratum. For this initial allocation, the relative precision for estimating total book value is derived, based on equation (3). If the precision target is not reached with this initial sample size, one sample unit is added to the stratum that improves the estimated precision the most (i.e., that reduces the estimated standard error of the estimated total book value the most). If the precision target is not met for the new sample size, another sample asset is added to the stratum that improves the estimated precision the most. This iterative process is repeated until the precision target is achieved.

Estimated recovery rates at the stratum level are critical in determining the loss reserve estimates for the FDIC's financial statement. Therefore, some modest increases are made to the optimum sample sizes for a few strata to avoid very small stratum sample sizes. Because of the complexity of the estimated liability to the insurance funds associated with these acquired assets, it is difficult to determine what the appropriate increases in stratum sample sizes should be. It is clear that there should be more than two assets selected from most strata, but the derivation of an optimum increase in stratum sample sizes is an open question for research. The designated sample for each stratum is selected randomly from the noncertainty assets in the stratum.

3. The Problem of the Universe as a Moving Target

The ALR asset universe is constantly changing as assets are sold and bank failures occur. This creates an estimation problem because the FDIC's annual financial statement is prepared as of December 31. Some type of allowance has to be made for the changes that occur near the end of the year.

One approach that has been considered to address this problem is to select an initial sample from the universe as it exists at a specific point in time (like July 1 or Sept. 1) and then update the sample as the universe changes between then and the end of the year. This would involve the possibility of modifying stratum sample sizes as the strata either shrink or expand during the latter months of the year. It was determined that this approach would not be feasible for two main reasons. First, there could be assets for which valuations were made (which take a considerable amount of FDIC staff time and effort) but not used because the assets are sold before the end of the year. This would be an inefficient use of FDIC staff resources.

Second, if additions of assets were made to the universe late in the year, additional sample selections may be needed. Depending on how many additional assets would have to be valued, this could create a resource problem, especially during the yearend holiday season.

Therefore, it was determined that the approach of continuous updating of the universe and sample until late in the year would not be practical. Instead, the universe is "frozen" at a specific time (currently July 1). The sample is selected at that time, the selected assets are valued, and the recovery and liability estimates are made. Then, FDIC accountants apply a "roll forward" process to provide acceptable estimates for the December 31 universe. This "roll forward" process is an accounting procedure, and does not involve any statistical inference. (For a description of this "roll forward" process, see Federal Deposit Insurance Corporation, 2004a.)

4. The Problem of Shrinking Stratum Sizes

Over the years the number of assets in the ALR universe has gone from about 40,000 in the mid 1990s to 765 in 2004. As the number of assets has decreased, the number of strata has also decreased. However, the most recent reduction in the number of strata, which took some time to develop and required the approval of the ALR Board of Directors, occurred about four years ago when the universe still contained about 1,500 assets. At that point, there were 20 strata. Some of the strata contained rather small numbers of assets then, but the numbers are even smaller now.

For the 2004 ALR sample design, three of the seven strata contained less than 30 assets. Since a separate ratio estimator is used to estimate recovery totals, there is concern about the size of the bias of the estimator due to small cell counts. A ratio estimator is used for ALR estimates of total recovery to take advantage of the correlation between book value and recovery value. Since it is believed that this correlation is stronger within strata (though this supposition has not been adequately verified), project staff continue to prefer to use the separate ratio estimator over the combined ratio estimator. Most project staff members have no understanding of the possible bias associated with the separate ratio estimator; project statisticians have the responsibility to investigate this concern and to advise project staff about it.

Investigation into the concern with bias of the separate ratio estimator has been rather limited, partly because the ALR universe is relatively small (especially in terms of total book value), and therefore the impact of poor recovery estimates on the accuracy of the FDIC's financial statement would be minor. We have reviewed the discussion of this issue given by Cochran (1977, p. 165). He gives an upper bound of the bias, relative to the standard error of the estimate, as follows:

Upper bound of the bias relative to the standard error is of order: $\sqrt{L}\ cv(\bar{x}_h)$ (4)

Since L=7, if $cv(\bar{x}_h)$ is about 0.1, the upper bound on the bias is about 0.26 times the standard error. Therefore, the bias could increase the mean squared error by about 7% (i.e., 0.26 squared). This does not appear to be a major concern, but something that needs to be checked each year the ALR Project is done.

5. The Problem of Estimates Being Required at the Receivership Level

The main entries in the FDIC financial statement that are derived from ALR sample valuations are estimated loss liabilities associated with acquired assets from failed institutions (receiverships). 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.

Since these liability calculations must be made at the receivership level, one option would be to use "receivership" as the major stratifying variable in the ALR sample design, and to be sure that receivership sample sizes were adequate to make the needed liability calculations with sufficient precision. However, this would require a substantially larger sample size than is practical. Therefore, the design of the ALR 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.

First, an estimated recovery rate, r_h, for each noncertainty stratum, is computed (for stratum h) 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 the stratum. 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 of these assets that were not selected for the sample, the estimated recovery amount is computed as the book value times the estimated (synthetic) stratum recovery rate.

The estimated loss liability for the receivership is calculated from the estimated recovery amounts, as defined above, and the liabilities for the receivership. The estimated loss liabilities for all the receiverships are summed to derive an estimated loss liability at the fund level. The same method is used for both insurance funds, BIF and SAIF. (For a description of the accounting methods used to derive the estimated liability to the insurance funds at the receivership level, see Federal Deposit Insurance Corporation, 2004b.)

Because of the use of a synthetic estimator to estimate the loss liability to the insurance funds, it is not possible to derive a closed-form estimator of the precision of the liability estimates, such as the variance. Instead, a bootstrap (resampling) estimator of the precision was developed about eight years ago, using 100 bootstrap estimates (see Cowan, 1997). Since then, project statisticians have been investigating possible improvements of the bootstrap methodology used, based on the work of Rao and Wu (1988).

However, since the size of the ALR asset universe has decreased so much in the past few years, the need for precision estimates for the estimated fund liabilities has become less important. Consequently, resources have not been allocated to this aspect of the ALR project, and currently precision estimates for estimated loss liabilities on the financial statement are not being made.

6. Future Research

Because of the trend over the past several years of continuing decreases in the size of the ALR universe, it is not likely that improving the ALR sample design will be a high priority. However, as long as a probability sample is used for the ALR Project (and that is not certain), some improvements will continue to be made, in spite of barriers associated with practical constraints.

In particular, efforts will be made to further reduce the number of strata so that stratum universe sizes will all exceed 25, though this will require ALR Board of Directors approval. As a related topic, additional investigation of the possible bias of the separate ratio estimator will be conducted. Another area for future research, referred to in Section 2, is the derivation of optimum increases in stratum sample sizes to take into account the critical role of estimated stratum recovery rates in the calculation of loss reserve estimates.

Finally, the need for precision estimates of the estimates of total net liability to the insurance funds will be revisited. If the need exists, and the resources are available, revised bootstrap methods will be developed.

References

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 memorandum, Jan. 2, 1997.

Federal Deposit Insurance Corporation (2004a). "Asset Loss Reserve Project '04 Process Memorandum." Internal FDIC memorandum available from David Chapman, FDIC Chief Statistician, Division of Insurance and Research, dachapman@fdic.gov.

Federal Deposit Insurance Corporation (2004b). "FDIC Fund Level Loan Loss Reserve Process Memo." Internal FDIC memorandum available from David Chapman, FDIC Chief Statistician, Division of Insurance and Research, dachapman@fdic.gov.

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.