

SUPPORTING STATEMENT B

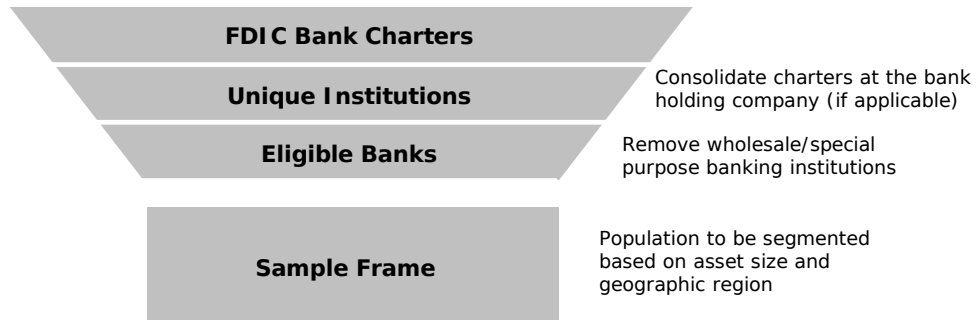
National Survey on Banks’ Efforts to Serve  
The Unbanked and Underbanked  
(new collection)

B. STATISTICAL METHODS

1. Potential respondent universe (including a numerical estimate) and any sampling or other respondent selection methods to be used.

**Sample Frame**

For this study, the sample frame will be bank holding companies (or ‘unique institutions’ in the FDIC database if they do not have a bank holding company). The sample frame will include FDIC insured financial institutions that operate retail banking operations; as such, wholesale banks and special purpose banks are not relevant for this survey, and will be excluded from the survey sample based on specialty charter and branch characteristics (i.e., assets per branch indicative of a money-center.) These two screening criteria will be used to determine the sample frame of U.S. banks. (In subsequent discussions of the sampling frame, for simplicity the sampling units will be referred to as banks or holding companies.)



Based on the application of these two screens, the sample frame includes 7,487 banks, as shown in the table below by asset size groupings.

	# of Bank Charters	# of Banks
<b>Tier 1</b> Top 25 banks by assets	91	25
<b>Tier 2</b> Banks with \$1 Billion or more in assets outside the Top 25	1,115	564
<b>Tier 3</b> Banks with less than \$1 Billion in assets	7,435	6,898
<b>Sample Frame</b>	<b>8,641</b>	<b>7,487</b>

<i>Ineligible*</i>	9	5
Total	8,650	7,492

\*Tier 1, as currently defined, already excludes a number of wholesale and special purpose banks.

### **Sampling Unit**

The sampling unit will be at the bank holding company level (if one exists), as reflective of legacy arrangements. The respondent sought will be the bank officer responsible for retail bank operations. For U.S. owned holding companies, we will treat each of the top 25 holding companies as a single-unit and ask the bank leadership to specify contacts for the study.

#### **2. Procedures for the collection of information.**

A paper survey will be mailed to the designated officer at each bank included in the sample. Additional copies will be made available electronically as a MS-Word file or in paper as requested. Each bank receiving the survey will be tracked to compute response rates and for follow-up requests for participation. Survey data will be collected either from the paper form or electronically and consolidated into a data set. Quantitative responses will be entered into SPSS for statistical analysis and qualitative responses will be grouped in MS-Word files for thematic analyses.

- a) Statistical methodology for stratification and sample selection

### **Stratified Sample Design**

The top 25 banks, in terms of asset size, will constitute the certainty stratum, and as such will be automatically included in the sample. For the remaining banks, random selection procedures will be used to draw a stratified random sample, where the strata are defined by the two remaining size classes and the nine Census Divisions. The primary strata will be defined by the three bank size groups: top 25 (certainties), banks with total assets of \$1 Billion or more, and banks with under \$1 Billion in assets. The next level of stratification will be on the nine Census Divisions. The sample design will allow us to examine differences among bank asset size groups and geographic areas. As shown in the table below, this sampling approach will create 19 strata for analysis: comparisons by the three size groups or the nine Census Divisions.

#### **Sample Stratification Scheme**

	<b>Certainty</b>	<b>Mid-Sized</b>	<b>Under \$1B</b>
<b>New England</b>			
<b>Middle Atlantic</b>			
<b>South Atlantic</b>			
<b>East South Central</b>			
<b>West South Central</b>			
<b>East North Central</b>			
<b>West North Central</b>			
<b>Mountain</b>			
<b>Pacific</b>			

### Stratified Sample Selection

Because of the disproportionate influence of the largest banks on availability of retail banking services, a stratified design with an oversampling of larger banks will improve sampling efficiency. Specifically, a certainty stratum of the Top 25 banks (goal = 100% participation) will be defined. In addition, there will be an oversampling of the Tier 2 banks (total assets of \$1 Billion or more).

Data collection will be performed at the highest organizational level of the bank within the United States. The sample will be selected using random selection procedures within each sampling stratum.

### Sample Size and Sample Allocation

The target number of completed interviews for the survey is 865. To achieve this target, we plan to invite about 1200 banks to participate in the survey, with a goal of achieving a 65% or better response rate.

The following table shows the target number of completed interviews and the corresponding sampling rates (percentage of banks covered) for each of the three asset size strata. Overall, the sample goal for completed interviews represents more than 10% of all federally-insured banks in the United States.

**Sample Strata Universe Counts and Sample Goals**

<b>Strata</b>	<b># of Unique Institutions</b>	<b>Sample Goal (Completes)</b>	<b>Percentage of Banks Covered</b>
<b>Tier 1 (Certainty)</b> Top 25 banks by assets	25	25	<b>100.0%</b>

<b>Tier 2 (Mid-Sized)</b> Banks with more than \$1 Billion in assets outside the Top 25	564	180	<b>31.9%</b>
<b>Tier 3 (Under \$1 Billion)</b> Banks with less than \$1 Billion in assets	6,898	660	<b>9.6%</b>
<b>Total</b>	<b>7,487</b>	<b>865</b>	<b>11.6%</b>

In addition to asset size stratification, the sample of banks outside the top 25 will be allocated proportionally to the nine Census divisions within each of the other two size strata, to ensure geographic representation.

The nine divisions are:

1. New England Division: Maine, Vermont, New Hampshire, Massachusetts, Connecticut, Rhode Island
2. Middle Atlantic Division: New Jersey, New York, Pennsylvania
3. South Atlantic Division: Maryland, Delaware, West Virginia, Virginia, North Carolina, South Carolina, Georgia, Florida
4. East South Central Division: Kentucky, Tennessee, Alabama, Mississippi
5. West South Central Division: Oklahoma, Arkansas, Texas, Louisiana
6. East North Central Division: Wisconsin, Michigan, Ohio, Indiana, Illinois
7. West North Central Division: North Dakota, South Dakota, Nebraska, Kansas, Missouri, Iowa, Minnesota
8. Mountain Division: Idaho, Montana, Wyoming, Colorado, Utah, Nevada, Arizona, New Mexico
9. Pacific Division: Washington, Oregon, California, Alaska, Hawaii

b) Estimation procedure

### **Selection Probabilities**

The sample design for the survey of banks will be a single stage stratified design with random sampling of banks from within strata. The selection probabilities for all units within a stratum will be equal to the sampling rate for the stratum. The selection probabilities will be equal to 1 for all banks in the ‘top 25’ certainty strata.

Let  $h$  denote the stratum and the index  $i$  denote a bank (or bank holding company) within the stratum. If we select  $n_h$  (the sample size for stratum  $h$ ) out of the  $N_h$  (the population size for stratum  $h$ ) banks in stratum  $h$ , then the selection probability for stratum  $h$ ,  $\pi_h$  is given by

$$\pi_h = n_h / N_h .$$

The selection probabilities  $\pi_h$  define the base weights as discussed below.

### **Base Weights**

Survey responses from the respondents will need to be inflated to obtain estimates for the entire population using sampling weights. These weights will (1) compensate for unequal selection probabilities, (2) attempt to compensate for non-response by some banks, and (3) attain greater precision for the survey estimates through post-stratification (or ratio estimation) if useful auxiliary data for post-stratification can be attained. The selection probabilities of the banks will vary by stratum due to different sampling rates for the size strata. The base weight for a sampled bank is defined to be the reciprocal of the selection probability of the bank. Such base weights produce unbiased estimates of totals and percentages in the absence of any nonresponse in the survey.

If  $\pi_h$  is the selection probability for the sampled banks from stratum  $h$  then the base weight (or design weight) assigned to the sample banks in stratum  $h$  is defined as

$$w_h = 1 / \pi_h .$$

Properly weighted estimates using the base weights (as given above) would be approximately unbiased if every sampled unit agreed to participate in the survey. However, we anticipate some nonresponse in this survey. Therefore, a nonresponse weight adjustment will be used to attempt to minimize the potential for nonresponse bias.

A ratio adjustment of sample weights at the individual stratum level may be performed using the auxiliary variable (assets) to reduce the variability of sample estimates. The improvement in the reliability of the estimates due to ratio adjustment depends on the correlation between the study variable and the auxiliary variable used for ratio adjustment.

### **Adjustments for Nonresponse**

The base weights will be adjusted to account for the nonresponding banks. As described above, base weights would produce approximate unbiased estimates only if there were no nonresponding banks. Due to the presence of nonresponders, a weight adjustment is used to account for nonresponding banks. This weight adjustment will be obtained by increasing the weights of the sampled banks for which data are collected. The weight adjustment will be applied within each stratum, based on the assumption that sample banks will be “missing at random” within strata, or at least that the survey variables for a nonresponding bank will be more like those of responding banks in the same stratum than those in other strata.

Suppose that  $n_h$  banks were sampled from stratum  $h$  and data were collected only for  $r_h$  banks. Then the adjustment is calculated as the ratio of the number of sampled banks (both respondents and nonrespondents) in the stratum to the number of banks that actually responded to the survey.

If  $A_h^{(nr)}$  denotes the weight adjustment due to the nonresponding banks within stratum  $h$ , then

$$A_h^{(nr)} = n_h / r_h.$$

In order to compensate for the nonresponse, the base weight for stratum  $h$ ,  $w_h$ , is multiplied by the nonresponse adjustment factor  $A_h^{(nr)}$  given above. The stratum level design weight adjusted for non-response then becomes

$$w_h^* = A_h^{(nr)} w_h.$$

The above weight is assigned to all banks belonging to stratum  $h$  for which survey data is obtained. Although the above weights can be used to produce survey estimates, we may also use the auxiliary information on the asset size of banks from the sampling frame for the purpose of ratio estimation. The benefit of ratio estimation is that it may help reduce the variability of survey estimates.

### Ratio Estimation

Ratio estimation is a procedure in which the weights of the respondents are further adjusted so that the survey estimates of an auxiliary variable are equal to the known population totals of the auxiliary variable for each stratum or some aggregate of strata. We would apply the ratio adjustment at the stratum level because the sample sizes at the stratum level should be adequate in terms of size.

Let  $X_h$  denote the sum of the assets for all the banks in the stratum  $h$ , as obtained from the sampling frame, and let  $\hat{X}_h$  be the corresponding survey estimate obtained by using the nonresponse adjusted base weight. Then the ratio  $X_h / \hat{X}_h$  is used as an adjustment to define the ratio estimation weight  $w_h^{(ratio)}$  as

$$w_h^{(ratio)} = \left( \frac{X_h}{\hat{X}_h} \right) w_h^*.$$

The superscript (ratio) denotes the ratio estimation weight, and the weight is applied to all banks in stratum  $h$ . It can easily be shown that the ratio estimation weight  $w_h^{(ratio)}$  for stratum  $h$  can also be expressed as

$$w_h^{(ratio)} = \left( \frac{N_h}{r_h} \right) \frac{\bar{X}_h}{\bar{x}_h},$$

where  $\bar{x}_h$  and  $\bar{X}_h$  are respectively the sample mean and the population mean of the auxiliary variable assets for stratum  $h$ , and  $r_h$  is the number of responding banks from stratum  $h$ . The ratio estimation weights  $w_h^{(ratio)}$  are the final survey weights and these will be used to tabulate the survey responses.

c) Degree of accuracy needed for the purpose described in the justification  
Academic levels of accuracy will be sought (i.e., 95% confidence level for estimating a universe

proportion to within  $\pm 5$  percentage points for a binomial or categorical question). An important aspect of the research is to gather qualitative information from open-ended questions which will identify what has been effective and may be best practices that could be used by banks to serve unbanked and underbanked individuals.

d) Unusual problems requiring specialized sampling procedures

For some bank holding companies it will be necessary to collect multiple responses if they have different or legacy procedures at retail banks within their bank holding company. Specifically, we will be requesting the largest banks to complete separate surveys if their policies differ by charter.

e) Any use of periodic (less frequent than annual) data collection cycles to reduce burden.

Survey may be repeated every two years at the FDIC's option

3. Methods to maximize response rates and to deal with issues of non-response.

The survey will be mailed to the bank holding company president or executive responsible for retail operations with a letter requesting participation by the Vice Chairman of the FDIC. In addition to the paper survey, an electronic version will be provided to make it easier for banks to participate. Extensive follow-up activity will be undertaken encompassing follow-up letter, emails and telephone calls to boost participation rates.

4. Description of all tests of procedures or methods to be undertaken.

Draft survey forms were reviewed internally by the FDIC and externally with compliance officers at four banks. Based on the feedback, forms were revised and simplified.

In the pilot test, the draft survey questionnaire will be reviewed with nine banks to gather their feedback about the questions and the process for gathering the data. Pilot participants will be asked to complete an evaluation and comment sheet, which will ask participants to identify any questions that were unclear or confusing. We will then speak with the participants to review their comments. Revisions will be made as required based on feedback from this pilot test

5. Name and telephone number of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze the information for the agency.

Ed Bachelder, Dove Consulting 617-753-9223

Attachments

1. Draft survey instrument
2. "First" Federal Register notice; draft "second" Federal Register notice