

APPENDIX A

TECHNICAL NOTES

The primary data source for this report is the Federal Deposit Insurance Corporation’s (FDIC) 2022 Small Business Lending Survey (SBLs), a nationally representative survey of banks about their small business lending practices. The 2022 survey follows and updates the first Small Business Lending Survey conducted in 2016. The FDIC undertook the 2022 survey to highlight many aspects of small business lending for which detailed information is not generally available or is not representative of all parts of the banking industry. This report provides insights on how banks engage in small business lending, including how they interact with their small business customers, define their market for small business lending, develop and use financial technology, perceive their competition and competitive advantages, and underwrite loans, including loans to start-ups. The SBLs design allows for the report to make statements about the industry at large and to highlight significant differences in the industry.

The FDIC developed the SBLs in consultation with survey experts at the U.S. Census Bureau. During the survey development phase in 2021, the FDIC and the Census Bureau conducted three rounds of interviews with approximately 50 banks of various sizes. The interviews were conducted to ensure that bank personnel understood the survey questions, that the information gleaned from the questions matched the intent of the researchers, and that respondents could provide the requested information using a reasonable amount of bank resources. The survey was estimated to take bank staff 3.5 hours to complete, depending on the size and complexity of the bank. In the first half of 2022, the Census Bureau tested the usability of the survey’s web interface to ensure it functioned as intended.

The Census Bureau collected data for the SBLs between June 2022 and December 2022. The data collection was conducted online through the Census

Bureau’s Centurion data collection website. An informational copy of the survey instrument can be found in Appendix B. The findings in this report are a result of an analysis of Sections I, II, and III.

A.1 Sample Selection

The universe of banks covered by the SBLs was constructed using data from the September 2021 Consolidated Reports of Condition and Income (Call Reports) and the June 2021 Summary of Deposits (SOD). A bank was included in the universe if it:

1. Was present in both the September 2021 Call Reports and the June 2021 SOD data
2. Was insured by the FDIC
3. Had positive net loans and leases in its September 2021 Call Report
4. Had at least one full-service office in a U.S. state or in the District of Columbia
5. Was not an International Banking Act institution

The sampling methodology resulted in a universe of 4,868 banks out of the 5,310 banks that submitted a September 2021 Call Report. The universe was then stratified along two dimensions: total asset size and whether the bank responded to the 2016 SBLs. Combining these two dimensions created 12 strata, the characteristics of which are summarized in the first three columns of Table A.1. The fourth column is the number of banks in each stratum.

The sampling methodology drew disproportionate stratified random samples for Strata 1 through 6 (covering all banks with less than \$3 billion in assets) and “certainty” samples for Strata 7 through 12, which contain all of the largest banks.⁹³ This methodology was used to ensure that the SBLs could produce accurate estimates for both small and large banks. The sample sizes for Strata 1 through 6 were assigned based on power calculations

⁹³ “Disproportionate” indicates that the probability of selection into the sample is different for banks in different strata.

Table A.1: Sample by Stratum

Stratum	Assets	Responded to SBLs 2016	Approximate Initial Sample Size	Approximate Adjusted Sample Size	Approximate Number of Respondents	Approximate Stratum Response Rate
1	Less Than \$500M	No	500	500	350	70%
2		Yes	250	200	150	77%
3	\$500M to Less Than \$1B	No	250	250	150	64%
4		Yes	150	150	100	73%
5	\$1B to Less Than \$3B	No	350	350	200	59%
6		Yes	150	150	100	76%
7	\$3B to Less Than \$10B	No	150	150	80	59%
8		Yes	80	80	60	80%
9	\$10B to Less Than \$50B	No	60	50	40	70%
10		Yes	50	50	40	70%
11	Greater Than \$50B	No	30	30	20	70%
12		Yes	20	20	20	70%
Approximate Total			2,000	2,000	1,300	68%

Note: In line with Census Bureau guidelines on preventing inadvertent disclosure, only approximate counts can be provided publicly.

conducted by the Census Bureau.⁹⁴ The approximate distribution of the initial sample of 2,000 banks is in the fourth column of Table A.1; in line with Census Bureau guidelines on preventing inadvertent disclosure, only approximate counts can be provided publicly. Between the time that the sample was drawn and the survey was conducted, about 50 banks initially included in the sample had closed or could not be contacted. These institutions were removed from the sample.

The banks selected in the sampling process were sent an advance contact letter and an initial survey request by mail, which introduced the study and outlined the procedures for logging on to the Census Bureau's data collection website to respond to the survey. Sampled banks that did not respond within 45 days of the beginning of the response period received follow-up letters and telephone calls.

A.2 Survey Response Rate

As shown in Table A.1, of the approximately 2,000 banks in the adjusted sample, about 1,300 responded to the survey, for a response rate of about 68 percent.⁹⁵ This response rate exceeds the Census Bureau's

standards for quality control.⁹⁶ The response rates varied somewhat between strata, ranging from 59 percent for Strata 5 and 7 to 80 percent for Stratum 8.

While the response rate met the Census Bureau's standards, nonresponse bias remained a concern. To address this concern, banks in the adjusted sample (excluding banks in the initial sample that had closed or could not be contacted) were compared to the subset of banks that responded to the survey, based on information in the Call Report and the Summary of Deposits. Table A.2 shows how sample banks differed from respondent banks within each stratum based on ten relevant characteristics. The "Resp" columns give the values for respondent banks within each stratum, the "Samp" columns give the values for the sample banks, and the "Sig" columns indicate whether the difference is significant at the 10 percent, 5 percent, or 1 percent level.

Overall, testing for significant differences between respondent and nonrespondent banks in each stratum finds that only 14 out of 120 tests (11.7 percent) were significant at the 10 percent confidence level, only slightly greater than random chance. The most notable difference is that in Stratum 8 (panel banks with \$3 to \$10 billion in assets), respondent

⁹⁴ Power calculations estimate how many banks must be sampled in each stratum to yield a desired level of precision.

⁹⁵ A bank was considered to have responded to the survey if it gave a response to question IA.1, the first non-screener question in the survey.

See Section A.3 below for an analysis of question-by-question response rates.

⁹⁶ See Sub-Requirement F1-6.2b of U.S. Census Bureau Statistical Quality Standards

Table A.2: Mean Values for Respondent and Sample Banks, by Stratum

Strata Number	1			2			3			4		
Assets	<\$500M						\$500M to \$1B					
Panel?	No			Yes			No			Yes		
Group	Resp	Samp	Sig	Resp	Samp	Sig	Resp	Samp	Sig	Resp	Samp	Sig
Natural Log of Assets	12	12		12	12		13.5	13.5		13.5	13.5	
Employees per \$100M Assets	17.8	18.3		16.9	17.2		15.6	16.2		18.9	18	
Branches/Offices per \$100M Assets	2	2		1.9	1.9		1.2	1.3	+	1.2	1.2	
Percent Assets in C&I Loans	7	6.9		6.5	6.7		8.7	8.3		8.4	8.5	
Percent Assets in CRE Loans	11.8	11.9		12	12		20.7	20.1		19.4	20.1	
Percent Commercial Specialist Bank	36.3	37.4		35.7	38.3		69	67		65.7	67.6	
Percent Headquartered in South (or Puerto Rico)	32.1	35.1	*	30.4	33.8	+	31	35.8	*	44.1	43.2	
Percent Headquartered in Northeast	4.5	4.6		7.6	7.2		21.6	18.7	+	10.8	10.8	
Percent Headquartered in Midwest	56	52.9	*	56.7	53.6	+	34.5	34.3		35.3	36	
Percent Headquartered in West	7.4	7.3		5.3	5.4		12.9	10.8		9.8	10.1	

Note: A “+” denotes statistical significance at 10 percent, “*” at 5 percent, and “***” at 1 percent. Hypothesis tests compare differences between respondents and nonrespondents. C&I is commercial and industrial. CRE is commercial real estate.

Table A.2 (cont.): Mean Values for Respondent and Sample Banks, by Stratum

Strata Number	5			6			7			8		
Assets	\$1B to \$3B						\$3B to \$10B					
Panel?	No			Yes			No			Yes		
Group	Resp	Samp	Sig	Resp	Samp	Sig	Resp	Samp	Sig	Resp	Samp	Sig
Natural Log of Assets	14.3	14.3		14.4	14.4		15.5	15.5	*	15.5	15.5	+
Employees per \$100M Assets	13.8	14.6		13.6	13.3		10.9	11.2		11.7	11.6	
Branches/Offices per \$100M Assets	0.9	0.9		1	1		0.8	0.7		0.8	0.8	
Percent Assets in C&I Loans	8.9	9.7	*	9.3	9.2		10.5	11.1		10.7	10.8	
Percent Assets in CRE Loans	21.4	21.1		21.8	21.8		17.6	17.7		19.9	19.8	
Percent Commercial Specialist bank	80.3	83.1	+	79.5	78.9		76.9	78.8		88.7	85.4	
Percent Headquartered in South (or Puerto Rico)	36.4	36.4		23.2	23.1		37.2	36.4		17.7	18.3	
Percent Headquartered in Northeast	20.7	20.4		30.4	31.3		19.2	20.5		37.1	28	**
Percent Headquartered in Midwest	27.3	29		33.9	31.3		26.9	28		25.8	31.7	*
Percent Headquartered in West	15.7	14.2		12.5	14.3		16.7	15.2		19.4	22	

Note: A “+” denotes statistical significance at 10 percent, “*” at 5 percent, and “***” at 1 percent. Hypothesis tests compare differences between respondents and nonrespondents. C&I is commercial and industrial. CRE is commercial real estate.

Table A.2 (cont.): Mean Values for Respondent and Sample Banks, by Stratum

Strata Number	9			10			11			12		
Assets	\$10B to \$50B						>\$50B					
Panel?	No			Yes			No			Yes		
Group	Resp	Samp	Sig	Resp	Samp	Sig	Resp	Samp	Sig	Resp	Samp	Sig
Natural Log of Assets	16.7	16.8		16.9	16.9		19.1	19.1		19.1	19	
Employees per \$100M Assets	8.9	8.3		10.3	10.4		6.5	6.7		9	8.5	
Branches/Offices per \$100M Assets	0.5	0.4		0.6	0.6		0.2	0.2		0.4	0.4	
Percent Assets in C&I Loans	8.4	9.2		11.6	12.2		9.8	9.5		12.9	13.4	
Percent Assets in CRE Loans	16.3	14.6	+	19.2	18.5		5.2	4.4		8.1	8.5	
Percent Commercial Specialist Bank	73	68.5		85	87		33.3	28		64.7	65.2	
Percent Headquartered in South (or Puerto Rico)	32.4	35.2		35	37		27.8	28		35.3	43.5	
Percent Headquartered in Northeast	24.3	22.2		20	22.2		22.2	24		17.6	13	
Percent Headquartered in Midwest	10.8	13		27.5	24.1		22.2	20		23.5	17.4	
Percent Headquartered in West	32.4	29.6		17.5	16.7		27.8	28		23.5	26.1	

Note: A “+” denotes statistical significance at 10 percent, “*” at 5 percent, and “***” at 1 percent. Hypothesis tests compare differences between respondents and nonrespondents. C&I is commercial and industrial. CRE is commercial real estate.

banks were about 9 percentage points more likely to be headquartered in the Northeast than were banks in the overall sample.

A.3 Item Response Rate

In addition to the high survey response rate, most questions also exhibited a high item response rate, defined as the proportion of banks that answered a question conditional on having seen the question. Not all banks saw all the questions in the survey, and whether a bank saw a particular question could be a function of their previous answers and the bank’s stratum.⁹⁷ Item response rates were high across the board, with most questions having a response rate of above 85 percent, exceeding Census Bureau standards for quality control.⁹⁸

The notable exception was Question IB.10, which had only a 43 percent response rate. This question was a

follow-up to Question 1B.9, which was asked only if the respondent indicated that the first level of loan approval for a small business loan was conducted by a group of decision-makers. The question asked the respondent to indicate the types of personnel who comprise the decision-making group and the list was nearly identical to those that were provided in the prior question, excluding the group option. It is possible that respondents clicked through this question at a high rate due to a perception of having already just answered it. Response rates also declined somewhat towards the end of Part B of Section 1, in the 80-90 percent range for IB.20-IB.27, and 70 percent for IB.28 (the final question in the section).⁹⁹

A.4 Analysis and Statistical Precision

In line with standard Census Bureau methodology for stratified random sampling, banks were assigned weights based on the inverse probability of selection

⁹⁷ See the informational copy of the survey instrument in Appendix B for the skip patterns.

⁹⁸ See Sub-Requirement F1-6.2b of U.S. Census Bureau Statistical Quality Standards.

⁹⁹ Due to the length of the survey, item response rates are not reported in this appendix.

into the sample (with a slight correction due to the banks excluded from the sample due to closure or unavailability), which were then adjusted for nonresponse. All results in this report use these weights to produce accurate estimates that reflect the universe of banks in each stratum. Estimated differences discussed in this report are significant at the 10 percent level or greater. This indicates that, if, for example, small and large banks were equally likely to respond in the same way to a question, the probability of obtaining estimates with a difference equal to that observed or larger would be no more than 10 percent. In many cases, the observed differences were statistically significant at much lower thresholds than 10 percent.

A.5 Focal Product

Three parts of the survey prompt banks to answer questions in reference to a specific lending product. This specific product can differ from bank to bank and is determined for any one bank based on the bank's response to screener questions SN4 and SN5. Specifically, for questions analyzed in survey parts IB (Loan Approval Process), IC (Underwriting), and IIA (Lending Markets and Practices), banks are asked to consider their top lending product to small businesses by total dollar volume other than credit cards or government-guaranteed lending (their "focal product"). This approach reflects a decision to maintain comparability of responses across banks by excluding certain products (such as credit cards and government-guaranteed lending) that typically differ systematically from traditional credit products. As shown in Figure 2.7, the majority of banks answered questions either for term loans or lines of credit. There were significant differences in the distribution of the top product between small and large banks. To ensure that the differences between banks (such as between small and large banks) discussed in this report are in fact due to those characteristics and not the specific focal product, regressions were run exploring each difference that included controls for the focal product of the bank. In all cases, including controls for bank focal product did not qualitatively change the results.

A.6 Construction of Information Indices

This sub-section describes the construction of the information indices discussed in Section 3 and shown in Figures 3.16, 3.17, 3.18, and 3.19.

The indices are based on survey questions IC.4, IC.10, IC.16, and IC.22, Subparts a-l, which ask about the information that banks evaluate when underwriting loans of three sizes. The questions contain 12 categories of information, referred to here as subparts (such as business plan and loan officer's assessment), and ask banks how commonly they evaluate the subparts for each loan size.¹⁰⁰ For example, a bank may evaluate the personal credit score or history of an applicant most of the time for a loan of about \$25,000. The overall index for a given loan size was calculated by assigning 0 to 3 points as described below to the answers for each subpart, summing the points across the responses for each bank, dividing by the maximum number of possible points, and multiplying by 100. This process yielded an information score ranging from 0 to 100 for each bank and loan size.

Points were assigned as follows:

- 0 points for "Evaluates for no or almost no loans of this size," "Don't know," or a missing response
- 1 point for "Evaluates for some loans of this size"
- 2 points for "Evaluates for most loans of this size"
- 3 points for "Evaluates for all or almost all loans of this size"

Since there were 12 subparts, the maximum score for a bank that responded to all subparts with an answer other than "Don't know" was 36. As another example, a bank that answered "Evaluates for no or almost no loans of this size" to three subparts, "Evaluates for some loans of this size" to four subparts, and "Evaluates for all or almost all loans of this size" to five subparts would have a score of 52.8: $100 \times ((3 \times 0) + (4 \times 1) + (5 \times 3)) / 36 = 52.8$.

¹⁰⁰ To reduce survey burden, banks answered only the information questions for either \$1 million or \$3 million loans. Banks that made \$3 million loans to small businesses were asked about their \$3 million loans, and banks that made \$1 million loans but not \$3 million loans to small businesses were asked about their \$1 million loans. Responses for \$1 million and \$3 million loans were aggregated as "large loans." In addition, the information questions were specifically keyed to the respondent bank's largest volume loan product other than government-guaranteed products or credit cards. For simplicity of presentation in this report, responses are generally aggregated across loan products.

For each subpart that a bank skipped or answered with “Don’t know,” the maximum score was reduced by 3 points. For example, a bank that answered 11 subparts with “Evaluates for all or almost all loans of this size” but skipped the 12th subpart would have a score of 100: $100 \times (11 \times 3) / 33 = 100$.

For indices that split the information types into two categories—soft information and hard information—the same procedure was applied but to a subset of the subparts of the question. The soft information index used five of the twelve subparts: (e) business plan, (f) experience in industry, (g) identity of business advisors, (h) loan officer’s assessment, and (i) market conditions. The hard information index used seven of the twelve subparts: (a) audited financial statements, (b) unaudited financial statements, (c) business credit score, (d) personal credit score, (j) purchase agreements, (k) willingness to offer collateral, and (l) willingness to offer guarantees. As with the overall information index, both were scaled from 0 to 100 so that the lowest score would be 0 and the highest 100.

A.7 Measure of Market Competitiveness

Figure 5.5 in Section 5.2 presents analysis showing how the market distance that banks reported through the survey varies with relative competitiveness of the markets in which they operate. The measure of competitiveness was constructed from non-survey data and was based on the Herfindahl-Hirschman Index (HHI) for deposits. The HHI is a common metric for characterizing the relative dominance of a firm or set of firms within a specific market and for a specific product. Typically, a low HHI indicates a relatively competitive market for the product with no particular firm holding a large share of the total quantity, while a high HHI suggests that a single or a small set of firms dominate the market. The five levels of competitiveness shown in the figure are based on the following procedure:¹⁰¹

1. Take the set of full-service branches from the 2021 Summary of Deposits data for branches in the 50 U.S. states and the District of Columbia. Note that this set of branches includes the branches of banks that did not necessarily meet all the conditions to be in the sample frame (see Appendix A.1).
2. For each county c , calculate the deposit market share s_i of each bank with a branch in that county. A county where only a single bank has a branch presence would have a single market share equal to 1. A county with two banks with the same total amount of deposits would have two market share statistics each equal to .5.
3. For the n banks in a county, calculate the HHI value for each county using the following formula:

$$HHI_c = s_1^2 + s_2^2 + s_3^2 + \dots + s_n^2$$

4. For each bank i , calculate the share of that bank’s deposits in county c , w_{ic} , where, for example, $w_{ic} = 0$ for counties where the bank holds no deposits and $w_{ic} = 1$ if bank i had all of its deposits in county c
5. For each bank i with deposits in C_i counties, calculate the weighted average HHI, where the weights reflect the share of that bank’s deposits in a given market such that

$$HHI_i = \sum_{c \in C_i} w_{ic} HHI_c$$

6. Take the set of bank-level measures of HHI_i and calculate the quintile values across all banks. Classify each bank into a quintile of competitiveness as in Figure 5.5 based on the quintile in which their HHI_i value lies. The quintile containing the smallest HHI values is labeled as the most competitive.

A bank with a high HHI does most of its business in markets dominated by a small number of banks, which is interpreted to mean that the bank faces a less-competitive landscape than a bank with a low HHI that does business in markets where market shares are more dispersed.

¹⁰¹ The procedure as described here uses a county-level market. Section 5 uses an alternative market definition where the market is the MSA when the county containing a bank’s branches is inside an MSA and the county itself for those counties not in an MSA. Therefore, the bank-level measure of HHI used in Section 5 is weighted across markets which may be a combination of MSAs and counties.

One limitation of the analysis is that using deposits to calculate an HHI value does not directly correspond to the competitiveness of the market for small business lending. Analysis not included in this report shows that the findings presented are qualitatively robust to alternative definitions of market and to using the number of branches instead of total deposits when calculating HHI values.

A.8 Cross-Year Survey Comparisons

Section 5.4 of this report compares results from the 2016 survey with results from the 2022 survey using the following methodology: The two surveys were treated as repeated cross-sections without adjusting standard errors and other statistics for overlap in respondents, so that observations in each survey year were treated as entirely independent. No adjustments were made for entries, mergers, closures, or re-sampling across the intervening period. This is analogous to a weighted least squares regression with the estimating equation

$$X_{it} = \alpha + \beta \text{Year2022}_{it} + \epsilon_{it}$$

where observations from both survey years are pooled and X is the outcome variable of interest. The estimate of β is therefore the estimated change in the variable of interest between 2016 and 2022, with the resulting standard errors used to test the hypothesis that $\beta = 0$, meaning there is no difference in the outcome X between the two surveys.

Question wording in the two surveys differed. For the full text of the 2016 survey, see Appendix C of FDIC (2018); for the full text of the 2022 survey, see Appendix B of this report. Sampling, weights, and non-response calculations for the 2016 survey are discussed in Appendix A of FDIC (2018) and for the 2022 survey in Appendix A.1-A.4 of this report.

A.9 Regression Analysis

This sub-section describes the multinomial regression analysis used in Section 7 and depicted in Figure 7.9. The results of the regression are shown in Table A.3.

The regression uses a linear probability model estimated using weighted least squares that incorporate the weighting scheme of the survey respondents. The estimating equation is

$$\begin{aligned} \text{StartUpLender}_i = & \alpha + \beta_1 \text{Large}_i + \beta_2 \text{Meet}_i + \\ & \beta_3 \text{SBA}_i + \beta_4 \text{PCS}_i + \beta_5 \text{HardTerc2}_i + \\ & \beta_6 \text{HardTerc3}_i + \beta_7 \text{CreditScore}_i \\ & + \beta_8 \text{TermLoan}_i + \beta_9 \text{Other}_i + \epsilon_i \end{aligned}$$

which specifies the probability that bank i lends to start-ups as a function of several characteristics of that bank. All variables take a value of zero or one depending on whether the bank has the characteristic or does the practice as described below.

- **StartUpLender** indicates that the bank typically lends to start-ups.
- **Large** indicates that the bank has at least \$10 billion in assets.
- **Meet** indicates that the bank typically has a loan decision-maker meet with small business applicants as part of its loan approval process.
- **SBA** indicates that the bank made at least one SBA loan in 2021.
- **PCS** indicates that the bank often or always requires loan applicants to provide a personal credit score to get small loans of about \$25,000.
- **HardTerc2** indicates that the bank is in the middle third of the distribution in terms of its usage of hard information for small loans of about \$25,000. See Appendix A.6 for details on the construction of the information indices.
- **HardTerc3** indicates that the bank is in the top third of the distribution in terms of its usage of hard information for small loans of about \$25,000. See Appendix A.6 for details on the construction of the information indices.
- **CreditScore** indicates that the bank lists credit scores and other items from the credit report as the most important factor for approval of small loans of about \$25,000.

- *TermLoan* indicates that the bank’s largest loan product type by dollar volume at origination (other than government-guaranteed loans or credit cards) was term loans.
- *Other* indicates that the bank’s largest loan product by dollar volume at origination (other than government-guaranteed loans or credit cards) was a product other than a term loan or line of credit.

The *HardTerc1* and *LineOfCredit* categories were omitted from the regression due to multicollinearity. The coefficients on *HardTerc2* and *HardTerc3* should be interpreted in reference to *HardTerc1*; thus, all else equal, a bank in the middle third of hard information usage (*HardTerc2*=1) is 10.6 percentage points less likely to lend to start-ups than a bank in the bottom third of hard information usage (*HardTerc1*=1).

Table A.3: Start-Up Lender Regression

Variable	Coefficient	Standard Error
Large	-0.032	(0.064)
Meet	0.192**	(0.047)
SBA	0.121**	(0.034)
PCS	-0.105*	(0.052)
HardTerc2	-0.106*	(0.043)
HardTerc3	-0.107*	(0.043)
CreditScore	-0.136*	(0.053)
TermLoan	0.002	(0.042)
Other	-0.229	(0.16)
Constant	0.638**	(0.072)
Approximate Observations	1,000	
Log Likelihood	-775.837	
Akaike Inf. Crit.	1,571.68	

Note: A “+” denotes statistical significance at 10 percent, “*” at 5 percent, and “**” at 1 percent. In line with Census Bureau guidelines on preventing inadvertent disclosure, only approximate counts can be provided publicly. “Inf. Crit.” is information criterion.