

FRAME DEVELOPMENT FOR SURVEYS OF FINANCIAL INSTITUTIONS: USING ADMINISTRATIVE DATA FROM REGULATORY AGENCIES

Carl Ramirez
U.S. General Accounting Office

Abstract

This paper reviews the advantages and disadvantages of using administrative data from governmental regulatory reporting systems to create frames and draw samples for establishment surveys. Using the closely supervised U.S. banking industry as a case study, the nature and quality of frames resulting from the records of federal regulators is evaluated. This paper also describes the characteristics of the U.S. banking industry and its regulatory oversight, and makes observations about the impact of rapid change in the industry on the administrative data systems maintained by the regulators.

Introduction

Administrative records systems can be particularly useful for establishment survey sampling frame development and sample design, in addition to providing key analysis items through linkage of those records to survey responses. Many of the list frames that are widely used in establishment surveys conducted by government are based on administrative records on businesses compiled for tax, registration, or regulatory purposes mandated by government statutes. (FCSM, 1988; Swaim, 1993).

In addition to the tax filings that businesses are required to make to the IRS, certain U.S. industries are closely regulated at the federal level, and enterprises in those industries are often required to submit periodic reports to their regulators on a wide range of business practice and financial subjects.

This paper focuses on one such population -- the "banking industry," whose members are depository financial institutions -- for which statements of financial condition ("Call Reports") are required quarterly by federal banking regulators. The administrative data files resulting from this reporting system can be used to create sampling frames and to provide auxiliary data for sample design.

This general approach may be of interest to those researchers who wish to conduct surveys of a specific industry but who do not have access to a proprietary frame or master business register for producing such a frame, such as the Standard Statistical Establishment List maintained by the U.S. Bureau of the Census, or the Business Establishment List used by the Bureau of Labor Statistics. The cost of creating and maintaining a comprehensive in-house frame for even one industry may be prohibitive for some research organizations.

Of course, survey researchers might craft a multiple-source frame from any combination of a number of directories, databases, registers and lists of banking organizations available from public and commercial sources that vary in quality and coverage, but this paper focuses on this type of official data because of its potential for high coverage and cost effectiveness (Perry, 1993).

After a profile of the banking industry and the regulatory data system is made, the application of this administrative data to the task of the sample survey will be described and an evaluation of this records system in its capacity as a sampling frame will be undertaken. The characteristics and requirements of establishment sample frames have been well described elsewhere (Colledge, 1995; FCSM, 1988; Deming, 1960), and will be covered here only in passing.

The U.S. banking industry

There were approximately 10,400 chartered commercial banks and savings institutions in the U.S. in March 1999, down from nearly 12,700 that existed ten years before (Federal Deposit Insurance Corporation, 1999). This consolidation is a function of several historical factors. In the 1990's, some banks have been closed and relatively few new banks have entered the industry, but more importantly, the rate of mergers and acquisitions among existing institutions has accelerated in recent years (Holland, et. al, 1996). These mergers typically take the form of a large multi-bank holding

company acquiring one or more independent banks or single-bank holding companies. While some acquired institutions may retain their names, charters (licenses that determine a bank's regulator, the insurance fund it is covered by, and the business practices it is allowed to engage in), or both, others may continue to operate in the same business locations as nothing more than branch offices of another bank. This rapid consolidation has been furthered by the relaxation of prohibitions on interstate banking, making it easier for large banking organizations to expand through acquisition of other banks.

As in other populations of establishments, a disproportionately small number of firms account for most of the business activity of the industry – in the banking industry, the eight largest banking institutions accounted for almost 36% of the industry's assets in 1997 (Berger, et. al, 1998). This concentration of the banking industry is increasing, while the overall size of this industry, in terms of assets under management, is also increasing.

 Source: (Federal Deposit Insurance Corporation, 1999)

1. Change in the Structure and Classification of depository institutions

In addition to births and deaths in an establishment population, other changes can occur to individual statistical units, through consolidation (as in the mergers described above) or change in identity or organizational classification (Struijs and Willeboordse, 1995).

In the banking industry, there is a significant amount of this last type of change as well. For example, institutions can change their charters. A large number of savings associations supervised by the Office of Thrift Supervision (OTS) have become either state savings banks regulated by the Federal Deposit Insurance Corporation (FDIC) or commercial banks regulated by FDIC or another federal regulators.

In addition, holding companies (particularly savings and loan parent companies) are now allowed to engage in nonbank activities and own nonbank subsidiaries, further blurring the distinction between the banking industry and other financial services. Some of these parent institutions may be federally regulated as banking firms, but might be considered by researchers as members of other industries, too.

2. Regulatory structure

A depository institution can be granted one of a number of different federal or state charters and can be supervised by one of four federal regulators. The core components in the definition of the "depository institution" population is that an institution must be individually chartered, supervised by a federal regulator and insured by the FDIC. Such institutions include commercial banks, savings and loans (thrifts), mutual savings banks, industrial banks, some trust companies, and other special types of institutions. Nonbank subsidiaries owned or controlled by bank holding companies and other special purpose entities are supervised by the Federal Reserve Board. It is possible that individual institutions that are part of one bank family may be chartered and supervised by different regulators.

Table 1: Current number of FDIC-insured depository institutions, by federal regulator, as of 3/99

	Federal Regulator Responsible for Supervision			
	Federal Deposit Insurance Corporation (FDIC)	Federal Reserve (FRB)	Office of Comptroller of the Currency (OCC)	Office of Thrift Supervision (OTS)
Commercial Banks	5,294	994	2,433	
Savings Institutions	540			1,129
US Branches of Foreign Banks	25			

Source: (FDIC, 1999)

The Federal Financial Institutions Examination Council (FFIEC) is an umbrella organization that coordinates examination, financial data collection, and supervision between the four federal regulators. The FFIEC coordinates the reporting system that forms the basis for the administrative frame that is the subject of this paper.

It is important to note that the regulatory framework uses the possession of a legal charter to define the unit of analysis, while in practice, the distribution of business activity may not be cleanly divided between chartered entities. Instead, control over programs or policies may reside in holding companies or other parent entities, rather than in individual chartered entities. The administrative fact that a bank family is composed of 5 individually chartered members may be meaningful only to bank examiners, not the management or customers of the banks.

The Statements of Financial Condition – "Call Reports"

All federally chartered and insured banks and thrifts are required by law to file "Call Reports" -- Statements of Financial Condition, or Thrift Financial Reports (for thrifts, which are also known as savings and loan institutions, and overseen by OTS) quarterly with their regulators. Information on branch offices operated by the institutions is also collected and maintained separately. In addition, holding companies of individually-chartered institutions also have to report on a separate form. Individual institutions that are the parents of other institutions have to consolidate the financial information for those subsidiaries in their report.

The reports may be filed from the chartered institution's main office, but may contain data on the assets and liabilities of subsidiaries and branches located outside the institution's home state. Therefore, instead of an establishment-based listing where the elements correspond to separate geographical business locations, this database is organized to reflect the regulatory structure and also the hierarchical patterns of control in the industry.

The Call Report record contain numerous fields of financial and operational data on banks. In addition to balance sheet items, a variety of data is collected on practices ranging from the sale of mutual funds to the securities trading activity of banks. The information collected can change over time, with new items being added and others changed or removed from the data collection schedule. Detailed classification information is available on the type, charter, and deposit insurance fund membership of each institution. The data collection process is coordinated across the regulators by the FFIEC, and the FDIC is primarily responsible for processing and maintaining the resulting database.

In addition to the items collected quarterly, banks are required to obtain regulatory permission for certain organizational actions such as mergers, name changes, and changes in charter. Some information on the nature and date of recent changes is continually added to the current administrative bank records, which are based on Call Report data, and are linked to a separate database of "changes in structure" that the FDIC maintains.

Each chartered institution is assigned a unique "Certificate ID" number at birth, and these identifiers are not reused. In fact, Some Call Report database systems maintain records on "dead" institutions that are no longer in operation.

Currently, a limited number of fields from these reports are publicly available online through the FDIC or the FFIEC websites. The Federal Reserve System also maintains an online database covering those institutions it supervises. Much of the data collected are proprietary and sensitive; financial regulators limit dissemination of any information that could allow researchers or the general public to determine the financial condition of open banks. Databases constructed largely from Call Report data but with added search and reporting capabilities are also sold by commercial vendors.

How Call Report data can be used in sample surveys

The data from the Call Reports, when combined with other administrative data that the regulatory agencies maintain, has many research and analysis applications. Researchers at the U.S. General Accounting Office rely on Call Report data for drawing samples of banks for examination and survey purposes. Because the population is skewed to a great extent, with a relatively few large institutions accounting for most of the business activity in this industry, stratified sample designs with widely different rates of selection often result. The financial data in the Call Reports provides a

number of measures of size to choose from for this purpose. In addition to total assets, samples may be stratified based on deposits, loans, or number of branches.

In a survey of Automated Teller Machine (ATM) operations, for example, retail deposits or branches may be more correlated with ATM activity than total assets, whereas in a survey concerning farm lending, the detailed loan type breakdowns reported allow the specification of very efficient samples. GAO has recently conducted surveys of bank operations of ATMs, their mutual fund sales practices, and their use of derivative financial products.

Call Report data (and other regulatory databases created from this data) also serves as a sampling frame for surveys conducted by the regulators themselves. The Federal Reserve's "Survey of Terms of Bank Lending to Farmers" uses auxiliary data on loan volumes outstanding from the Call Reports, and uses commercial & industrial loans and farm loans to stratify and to create stratum weights (Walraven, 1993).

Call Report data is also the foundation of sample frames used for some surveys conducted by professional or trade associations in the banking industry. This database is often used to create subsets of the bank industry, such as the largest banks, or community banks (under a certain asset size and doing business only in the geographical area where the lead bank is located) for further study.

This list of bank organizations can also be used as a first-stage sample frame, where the ultimate target population is the customers or employees of the selected institutions.

Frame evaluation

Bank sample frames derived from regulatory records can be evaluated with criteria that might be grouped under the following six categories:

1. Data Quality

The quality of information in the administrative records that make up the frame is critical. The accuracy of the data is a function of the inherent *reliability* of the data coming into the administrative system, and the *timeliness* of updates reflecting a rapidly changing industry.

Call Report data collected by regulators is generally viewed as being very reliable. A mandatory reporting requirement exists, and reporting institutions are motivated by the control that regulators exercise in this industry. Much of the data requested conforms to the specifications of accounting and recordkeeping systems required for other official and tax reporting requirements. Regulator staff are assigned to review and edit database records to some extent.

Timeliness matters because changes in this volatile industry can result in births, deaths, and changes in structure. The frame has to be maintained by keeping administrative records up to date.

Information systems based on the Call Reports should be more up-to-date than others developed to cover the banking industry, because regulatory needs require real-time change information. Banks and thrifts must apply to regulators before making major organizational changes such as mergers or acquisitions, eliminating the need to conduct regular post-hoc surveys to detect births or deaths in the population. The FDIC maintains a publicly-available "change in structure" file with the most recent five events coded – changes to title, charter, and insurance fund affiliation are recorded here after mergers resulting from acquisitions, consolidation of family members into one chartered entity, name changes, switching regulators, incorporation and failure. Theoretically, the "change in structure" data could be used to continually update the larger body of financial records derived from the quarterly Call Reports.

2. Coverage

Errors of undercoverage (failure to include all units in the defined population) and overcoverage (inclusion of ineligible or out-of-scope elements, either from misclassification or an overly broad definition of the population)

can result from inadequate frames.

The population is well defined. In the Call Report database, classification and coding of elements is very precise. While Standard Industrial Classification or North American Industrial Classification (NAICS) coding schemes achieve some precision in coding financial institutions, bank class, charter, insurance membership information in the Call Reports allow for much finer categorizations.

Coverage of the banking industry by the regulatory information system is thorough. Even relatively recent organizational forms are covered, such as internet-based banks without a physical presence in communities they service. On the other hand, the unique definition of "depository institution" as operationalized by the database may exclude some institutions of interest to the researcher, and include others not generally associated with the banking industry. Because the institution must be regulated and insured by the FDIC, credit unions are not covered in Call Report datasets. Yet some trust institutions, industrial banks which are closer to investment banks and offer little if any retail banking services, domestic offices of foreign banks, and some nonfinancial institutions might be included in a frame derived from Call Report data.

The possibility of duplicate listings in the frame can lead to a form of over-coverage, because a duplicate listing gives that element an undeservedly higher probability of selection. While infrequent processing errors may lead to this condition, duplicate elements in the Call Report-based frame are more likely to exist because parent institutions are required to file consolidated reports of all subsidiaries, while those subsidiaries that are chartered depository institutions themselves are also required to file. Therefore, steps must be taken to account for this consolidation if sampling or analysis is performed that uses all holding company and individual institution records. Some relationship information on bank families, their branches, parents and holding companies exists in administrative data records incorporated with call report data, and must be used ensure that each entity initially has an equal probability of selection.

3. *Concordance of sample unit on administrative frame and reporting unit*

Ideally, the organizational units specified for reporting by the administrative records system would be consistent with the reporting units required by the survey (Perry, 1993). Unfortunately, the locus of control of a complex enterprise may not be neatly defined as the regulated entity for which administrative reports are made. Often, the researcher asks for data in an establishment survey corresponding to some organizational unit (an "independently chartered bank," for example) that the organization does not recognize as a meaningful unit, and does not keep data on. In a bank family, the locus of control may not reside in individually chartered entities, which exist only "on paper," but at the parent bank or holding company level. Requesting data from each member of such a family may result only in one aggregated report from a higher level entity. Sampling individual institutions listed in a regulatory frame may produce survey reports based on different units.

4. *Information on linkages between organizations*

The frame should document the linkages between bank family members and reflect the hierarchical relationships between holding companies, parent or "lead" banks, other family members, branch offices, and nonbank subsidiaries. This information may be crucial to selecting the appropriate reporting unit, or preventing unwanted duplication effects from consolidated reporting. This type of information is available in some versions of Call Report-based bank databases, although it may require significant processing to make use of it, because this data was not necessarily intended to serve as a means for selecting only one sample unit per firm. Branch office data produced by regulators may serve as a link between enterprises (bank families) and individual establishments.

In addition, these linkages may be longitudinal in nature, showing the relationship between institutional predecessors and successors. Historical information in Call Report-based information systems on "dead" banks that have merged with or been acquired by other institutions may allow this type of linkage as well.

5. *Auxiliary information for use in sampling*

Having a variety of demographic and financial data on sample elements can be very useful for efficient sample

design. More appropriate measures of size can lead to more optimal stratification in a survey tasked with estimating a specific characteristic. This auxiliary information can also be used in data analysis in conjunction with survey data. The Call Report system offers much financial information on each institution.

6. *Contact information*

Respondent selection is a critical determinant of error in the establishment survey. Large banks can be complex, multidivisional organizations. Determining the best qualified reporter is not the only task facing the researcher – in some cases it is necessary to overcome gatekeepers. In other cases researchers must obtain the help of facilitators or reporting coordinators who have the authority to generate a response, and who can coordinate responses when multiple local data providers within a bank are necessary.

Unfortunately, the reporter of record on the Call Report form may not be the best respondent for the researcher's survey. In some cases, these are technical specialists such as accountants or finance executives, but in other cases those who sign off on these reports are merely reviewing and authorizing the final product.

In some large banks, there are government compliance offices dedicated to this regulatory task, but these offices may not be the best point of entry into the sampled bank. Therefore, the survey researcher will often have to augment the Call Report Frame with more targeted contact information. This information can be obtained from industry directories.

Even if the survey uses a panel sample design, a common problem in such surveys of establishments is losing previously recruited reporters who leave the firm or change roles within the firm. Thus, respondent selection and enrollment is a continuous process regardless of contact information available from the frame.

Discussion

While using regulatory data for surveys of the banking industry has several advantages, a number of troubling issues arise. First, is the regulatory data system adjusting to changes in the industry? For example, if the classification system of firm types no longer fits many of the firms that do business in the industry, the coverage of any frame derived from it is questionable. With the increasing deregulation of the financial service industries, more opportunities for firms outside the traditional banking industry (from the securities or insurance industry, for example) to provide depository services will arise, and vice versa.

Second, data quality in the frame is of paramount importance. When the researcher relies on an outside source – be it a vendor or regulator – for designing and maintaining the sample frame, care must be taken not to become too complacent on matters of data reliability and coverage. Also, administrative data collection programs may change – the scope of the population covered may change, or certain data items may be discontinued.

Third, the likely divergence between the regulatory definition of the units that make up the industry and the researcher's needs (or the prevailing definition used by population members themselves) can be a large impediment to using regulatory data for sample frames. In any case, the definitions of the populations that regulators use can be complex and hard to translate into easily understandable parameters for the user of the research. Reprocessing the regulatory dataset may be necessary to create a more useful snapshot of the industry.

References

- Berger, Allen N., Rebecca S. Demsetz, and Philip E. Strahan (1998) "The Consolidation of the Financial Services Industry – Causes, Consequences, and Implications for the Future." Working Paper Series, Social Science Research Network.
- Colledge, Michael J. (1995) "Frames and Business Registers: An Overview" in B.G. Cox, et al. (Eds.), *Business Survey Methods*. New York: Wiley, pp. 21-47.
- Deming, W. Edward (1960) *Sampling Design in Business Research*. New York: Wiley.

- Federal Committee on Statistical Methodology (1988) *Quality in Establishment Surveys (Statistical Working Paper 15)*. Washington, DC: U.S. Office of Management and Budget.
- Federal Deposit Insurance Corporation (1999) *FDIC Quarterly Banking Profile: 1st Quarter, 1999*. March, 1999, Washington DC: FDIC.
- Holland, David, Don Insoe, Ross Waldrop and William Kuta (1996) "Interstate Banking -- The Past, Present And Future." *FDIC Banking Review*. Vol. 9, No. 1, December 1996.
- Perry, John (1993) "The Development of A Business Survey Frame from Administrative Data," *Proceedings of the International Conference on Establishment Surveys*. Alexandria, VA: American Statistical Association, pp. 186-192.
- Struijs, Peter and Ad Willeboordse (1995) "Changes in Populations of Statistical Units," in B.G. Cox, et al. (Eds.), *Business Survey Methods*. New York: Wiley, pp. 65-84.
- Swaim, L, et al. (1993) "Frame Creation for Institutional Surveys at Statistics Canada," *Proceedings of the International Conference on Establishment Surveys*. Alexandria, VA: American Statistical Association, pp.771-774.
- Walraven, Nick (1993) "Surveying Banks About Agricultural Lending," *Proceedings of the International Conference on Establishment Surveys*. Alexandria, VA: American Statistical Association, pp.141-145.