Business Trends and Forecasts: Using Securities and Exchange Commission's Electronic Records from the National Archives and Records Administration

Albert C. Minnick*

National Archives and Records Administration

The National Archives and Records Administration (NARA) is the official repository of United States Government records, including the records of businesses. These business records may relate to federal agencies in their regulatory capacities, or in their analytical and data gathering functions. In the past, the records of archival value were transferred to NARA almost exclusively in paper format. Today such records are increasingly created, stored, and transferred to NARA in electronic form. It is important for the researcher to be informed of the specific format of the records in order to conduct research in that format. This paper presents an overview of records transferred to NARA from the Securities and Exchange Commission (SEC).

The record keeping responsibilities of the SEC is the primary reason that these records are permanently kept at NARA. The SEC was created by the Securities Act of 1933, amended the following year by Congress, to remedy some of the abuses which led to the Great Wall Street Crash of 1929. The Securities Exchange Act of 1934 called for the regulation of the activities of the securities exchanges as well as broker dealers who operated in those trading arenas. The 1934 Act also requires periodic disclosures by companies that are publicly owned. These disclosures are rich sources of historical evidence. In order to understand which disclosures are available at NARA it is helpful to have a brief description of the SEC's organizational functions.

The SEC's main regulatory functions are enforcement, disclosure, and regulation. Regulation may be further divided by the regulation of public utility holding companies (1935), regulation of investment companies and advisers (1940), and regulation of the securities markets, e.g., the New York Stock Exchange. All of these functions are documented in several records series and available as magnetically stored files which may be accessed using automated data processing equipment. These types of records are commonly called electronic records. These are extremely useful records which appear to be largely unknown to business historians. Reference usage statistics suggest that

* I would like to thank the staff at the Center for Electronic Records, Reference Unit for encouragement, statistics, and assistance from Margaret Adams, Lee Gladwin and Ted Hull.

BUSINESS AND ECONOMIC HISTORY, Volume Twenty-seven, no. 1, Fall 1998.
 ISSN 0894-6825.

the data files are mainly requested (Table 1) and purchased by academic institutions. Published citations to the records, however, are scarce.

Table 1: *SEC Information Requests by Requestor Type, 1986-1998*

Private, Not-Profit	5
Private, for profit	625
Academia	1,813
Government	0

Source: Center for Electronic Records, 1998.

One reason that information requests are used primarily by academic institutions is that quantitative research techniques have matured within the past ten years. The hardware, software, and storage media necessary for processing these often large data files are now currently available on the desktop computer. As recently as 1983, when the Compact Disc Read Only Memory (CD-ROM) became widely available in libraries, there were no storage disks large enough on personal computers. The exception was at data processing centers in large universities. The research usually had to be conducted offline, often overnight on mainframe and minicomputers by data processing programmers and other non-historians. Now a historian who has access to a current desktop computer equipped with a CD-ROM drive, a database management system (DBMS), and a statistical analysis package can perform the type of quantitative analysis once accomplished only at great cost by programmers.

Although quantitative research of electronic records has grown easier and more generally available, this has still not increased their usage for business history. One reason may be that regulatory records are commonly entered into database applications, and as such they create a data pool too large to handle effectively. Another way of looking at this phenomenon is that the data create a useful model of the agency's organizational activities, but the model is unwieldy and difficult to grasp theoretically.

This difficulty to encapsulate the agency's mission is not evident in the SEC's electronic records. This is true partly because the SEC has maintained a continuity of the records by rarely changing the data structure, and partly because the SEC has had a relatively long cooperative relationship with NARA. Although the relationship has had lapses in which no records were transferred for several years, the SEC has always corrected the situation. The SEC has consulted with NARA in designing electronic submission systems since the 1970s. It is in this long research and development phase that NARA will soon reap further benefits when transfers from the Electronic Data Gathering and Retrieval (EDGAR) system arrive.

While the SEC and NARA have cooperated to identify, schedule, and permanently preserve the SEC's magnetically recorded records, the cost of copying this data has remained fairly constant since the early 1980s. This is not the case with the escalation of costs driven in the value-added private sector vendor industry. On any given day an Internet researcher might locate more than a dozen such vendors which offer data from the Ownership Reporting System (ORS) files in several permutations, either as single one line searches at

several hundred dollars a line downloaded or in time-driven contracts which cost several thousands of dollars. These same files are available at NARA at a fraction of the cost. Perhaps the only downside to accessing the data files from NARA is that they are indeed raw; there are no value-added bells and whistles. There are several benefits however, that may not be readily visible to the business researcher who is unfamiliar with NARA as a data resource.

One of the benefits of obtaining the SEC records from NARA is that it is possible to combine the data in relational databases and to research entirely new relationships. Many of these relationships have already been identified for the customer through painstaking documentation by NARA staff, e.g., the Investment Company Datafile (IVT), 1979-1988, which has more than 11,000 investment companies complete with names, addresses, and assets. These data fields may be combined with other files, for example, with files created by the Census Bureau. This combination could reveal a comprehensive market and investment analysis of companies in a particular field. One of the primary keys are the form numbers. Data from these particular forms will eventually be captured in the SEC's EDGAR system.

There are currently seven series of SEC records available at NARA. Besides the ORS and IVT previously mentioned. These series include the Broker Dealer Directory (BDD), 1935-1993; the Corporation Index System (CIN) which has data on which companies have filed annual reports from 1979 to 1993; the Investment Adviser Directory (IA) which has profile and statistical information from 1979 to 1993; the Proposed Sale of Securities (PSS) which has information from form 144 on proposed sales and blocks of securities not offered publicly; and the Registration Offering Statistics (ROS) System which among other phenomena includes information on shares purchased by foreigners. There is complete documentation available on each of these series along with thorough user support offered by the Center for Electronic Records Reference Staff.

Reference statistics indicate regular usage of only two of these series, that is, the Ownership Reporting System (ORS) and the Proposed Sale of Securities (PSS). Legal insider trading as documented in the ORS is probably the most prominent and widely recognized historical phenomena. Foreign financial organizations are also counted in the reference statistics because of their interest in how the United States monitors, documents, and punishes criminal activity related to insider trading. Several inside traders are also extremely well known by the popular press and recent articles have been written about the SEC's current chairman's actions to crack down on insider trading [Crock, 1997, p. 50].

In order to understand the research possibilities is it necessary to briefly describe what the ORS contains and how it may be used. First, it is useful for the researcher to understand that the ORS, while stable since 1978, has actually been modified several times. The Committee on Uniform Security Identification Procedures Stock Number (CUSIP) still remains the primary key. This is essentially a Social Security-type number for all of the corporations which the SEC regulates. The files contain monthly and in some cases yearly filings of security transactions. These filings are organized first by CUSIP and thereafter

by the name of the reporting person who has an insider relationship or beneficial ownership as defined by the SEC Act of 1934. Other data fields include the security name and stock symbol in later years, the relationship to the issuer, and nature of the ownership as well as date and type of transaction. Financial information includes the number of shares, value per share, and holdings at the end of the reporting period. Changes in holdings are reported when involving either more than 5,000 shares or more than $5,000 exchanged. These transactions are reported on SEC forms 3, 4, and since 1991, form 5.

The ORS may be used in several different ways. Probably one of the most effective is to do several tabulations or frequencies on either the names of the holding companies or the individual insiders, e.g., A.G. Edwards and Sons (company). There are exempted individuals whose names do not appear in the ORS. However, their names may be located in the financial sections of newspapers. It may be possible to generally infer which stocks an important investor is selling. The financial accounting data itself will then reveal monetary trends of which securities are most popular with inside traders and when are they most likely to be traded. The research possibilities are even more positive when sitting at a powerful DBMS in which the relationships to issuers are revealed in the SEC records. It is these relationships which might then be explored in comparative analysis with the records available at NARA and other social science repositories which should reasonably interest the business history researcher.

Much has been written in the past few years in the computing trade journals about legacy systems and the year 2000 problem. While it is true that the ORS is a legacy system in the sense that it is destined to be migrated into EDGAR in the future, because of its value it will continue either on the World Wide Web or in near online repositories such as data warehouses. One of the more promising techniques also written about in these journals is that of data mining. Combine these new information analysis applications with older but still trustworthy statistical analysis and the researcher has a valuable toolbox for effective historical writing.

There are definitely good records available at NARA from the SEC which are largely untapped [Securities and Exchange Commission, 1984]. These records may form the quantitative structure for solid business history. These records include data on topics such as insider trading, broker dealer information, and the amounts of securities exchanged since 1978. The regulatory function of the SEC has one of the most crucial missions, that is, to maintain U.S. business integrity. This mission is fairly, competently, and comprehensively completed as evidenced by these records. Business history visualization of the information from these records may help bring this ongoing mission to the research forefront.

Finally, it is important for the federal records researcher to know that the SEC records are generally all contained within an intellectual universe which is known as Record Group 266. Experienced staff will help the researcher in both the textual and electronic records units within NARA. Written guides, finding aids, and indexes are available to aid this research. When electronic records are transferred to NARA there are rigorous procedures instituted to

assure the quality, reliability, and verified readability. The primary procedure is called (interchangeably) validation or verification. Software assisted printouts and photocopies are also available at cost to convey this verification information to the researcher. In most cases when agency manuals, code books, or user guides are available, again these are identified and provided by reference staff. To contact the Center for Electronic Records, please either write, phone, or electronically mail a message.

Electronic and Special Media Records Services Division (NWME)
8601 Adelphi Road
College Park, MD 20740-6001
301-713-6630, fax 301-713-6911
cer@nara.gov

To use some of the online resources over the World Wide Web, point your browser to www.nara.gov/nara/electronic/. There are several different online tools available, especially the comprehensive Title List which currently lists more than 10,000 data files from more than 300 record groups. The Title List is also available from the Center's homepage at www.nara.gov/nara/electronic/tlintro.htm.

References

Center for Electronic Records, [Series Descriptions], January 22, 1998.

Crock, S., "Insiders Watch Your Back," *Business Week*, September 29, 1997, 50-52.

United States, National Archives and Records Administration, Center for Electronic Records, Reproductions For Fiscal Year, 1986-1996, March 9, 1998.

United States, Securities and Exchange Commission, *Good People, Important Problems, and Workable Laws* (Washington, DC, 1984).