

Iris Shimizu, National Center for Health Statistics

### 1. Introduction

A list of sampling units is a prerequisite for any survey. For surveys of firms that seek business from the public, the yellow pages phone directories can be a possible source for identifying and listing many, if not all, businesses of interest.

Yellow pages phone directories should not, however, be taken as the sole source for a survey's sampling frame until the completeness of the yellow pages phone directories coverage of survey eligible firms is determined to be adequate. The best use of the yellow pages in frame construction is probably as a supplement to other sources.

The National Center for Health Statistics (NCHS) used yellow pages phone directories to construct the sampling frame for a survey of nursing and related-care homes in a probability area sample. The methods used to identify survey eligible firms in that survey can be applied to any area sample survey of businesses that tend to advertise in yellow pages phone directories.

This paper discusses the methods used to identify and obtain the yellow pages phone directories for that survey. Space does not permit including detailed discussion about the search of directory pages for survey eligible firms, but information about those procedures may be obtained from the author.

The next section describes the survey in which yellow pages phone directories were used as a source for constructing the sampling frame for an area sample. It also describes plans for evaluating the yellow pages coverage of survey eligible firms. Section three discusses the methods used to identify yellow pages phone directories serving the sampled areas. Section four discusses the available sources from which yellow pages phone directories may be obtained. Section five summarizes recommendations based on our experiences for identifying and obtaining yellow pages phone directories for an area sample.

In the remainder of this paper, "yellow pages phone directories", "yellow pages," and "directories" will be used interchangeably.

### 2. National Master Facility Inventory Complement Survey

The yellow pages were used in constructing the sampling frame for the 1982 Complement Survey to the 1982 National Master Facility Inventory (NMFI). The NMFI is the most complete list of nursing and related-care homes in the U.S. In addition to providing data about the nation's supply of nursing and related homes, the inventory serves as the sampling frame for the National Nursing Home Survey.

Through 1982, the census of the NMFI was conducted in even numbered years. An agency reporting system is used to update the NMFI in other years. A complement survey based on probability area samples is used periodically to measure the completeness of the NMFI. Facilities eligible for the inventory in the area sample are identi-

fied, surveyed, and matched against those in the corresponding NMFI. Unbiased estimates are then made of the proportion of eligible homes which are not in the inventory.

The yellow pages current in 1981 were used to compile the mailing list for the 1982 NMFI Complement Survey. The reference time 1981 was used since the mailing list for the 1982 NMFI census was compiled during 1981. The decisions to conduct this Complement Survey by mail and to use the yellow pages to construct its survey frame were based on research in the 1980 Complement Survey that evaluated the feasibility of using yellow pages to construct sampling frames of certain types of health facilities.

The Complement Survey for the 1982 NMFI was conducted in 1982 in a probability subsample of 75 primary sampling units (PSU's) selected from the 383 PSU's in the sample for the National Health Interview Survey (NHIS), which is also conducted by the NCHS. Each PSU consists of a county or a group of conterminous counties. For the Complement Survey, the NHIS PSU's were stratified on the basis of their 1970 population: the first stratum consisted of PSU's having more than 2.3 million persons in 1970, the second consisted of the other PSU's which were self-representing in the NHIS sample, and the third consisted of the remaining PSU's.

The coverage of nursing and related-care homes in the yellow pages is not now unknown. However, when the results of the 1982 NMFI census are available, we plan to match the homes in that NMFI against those found in the yellow pages for the sampled areas in the 1982 Complement Survey. An estimate will be produced for the proportion of eligible homes covered in the yellow pages, probably by using a dual frame estimator since we will have data from two frames--the NMFI and the frame constructed from yellow pages.

The use of yellow pages as a source for the NMFI Complement Survey sampling frame is appealing for several reasons. First the yellow pages will not be the sole source for other surveys of nursing and related-care homes. Second the use of yellow pages saves the expense of having interviewers travel through the sampled areas to identify of facilities eligible for the NMFI. Also, because the facilities eligible for the NMFI seek business from the public they are likely to be listed in the yellow pages serving their respective areas.

### 3. Identifying Directories

The task of identifying the yellow pages serving sampled areas is difficult because there is no national list of phone directory titles sorted by county. The lists of directory titles that do exist do not include a listing for every town/city in the U.S.

#### 3.1. Procedures

Procedures for identifying and acquiring the yellow pages phone directories evolved during efforts on these tasks for the 1982 Complement Survey. Several references were used to identify

directories for the final list of directories for the survey. Among these were two catalogs of price list). The directories price list contains in alphabetical order the titles of directories published for 1981 by the American Telephone and Telegraph Company (AT&T) and most other telephone companies serving the U.S. The names of major communities are also included with a reference to the primary directory serving each if the community name is not included in a directory title.

The second catalog is the 1981 Phonefiche [4] catalog that lists titles of over 1400 directories (white and yellow pages) available on microfiche. These directories cover about 25,000 U.S. communities. The catalogs for later years may include additional directories, since continuous efforts are reportedly made to obtain permission from directory publishers to issue their directories on microfiche.

Another list of directory titles used was the Phonefiche 1981 Community Cross-Reference Guide to Phonefiche [3] (referred to hereafter as the cross-reference). The cross-reference lists nearly 17,000 communities alphabetically by State and cites for each community the directories serving it that are also available on microfiche.

Also used were detailed maps showing county boundaries and location code books that list towns/cities alphabetically with county code or name within State. The official code book for the 1982 Complement Survey was the latest edition of the Worldwide Geographic Location Codes [6].

Due to concern about the completeness of the final list of directories serving sampled areas, three methods were used to identify directories for the 1982 Complement Survey. The first method consisted primarily of calling directory assistance (area code plus 555-1212) for each of the sampled areas. The operator was asked about directories serving a particular area. If an operator was unable to provide the requested assistance, then the information was requested in calls to the business office of the telephone company serving the area, the Chamber of Commerce, and/or county offices in the county seat.

In the second method, all places included in the titles of directories listed in a directory catalog were sought on detailed maps showing county boundaries or in a location code book. If one of these places lays inside the boundaries of a sampled area, the directory having that place in its title was added to the Complement Survey list of directories.

In the third method, the names of all towns/cities in the sampled areas were first determined from the location code book. Each town/city thus found was then sought in the directories price list and/or the community cross-reference. If the town/city was listed in either of these references, all the directories indicated as serving the town/city were added to the Complement Survey list of directories.

For convenience, descriptive labels are used in the following discussion for the three methods. These are displayed by method in Table A. As can be seen in Table A, two of the methods were only partially applied in the 1982 Complement Survey. The need to meet mailing

deadlines for the Complement Survey prevented completing the search of directory title places on PSU maps. The search was completed for directories on microfiche since those directories were already on hand whereas there was no assurance, at the time, that the remaining directories could be obtained in time to meet the survey schedule.

Resource limitations also prevented completing the search among directory served places in all 75 PSU's. The procedure, however, was completed in a random subset of the Complement Survey sample PSU's so that the PSU's where the method was applied represents those in the full sample.

After directories were identified as potentially serving a sampled PSU, some were determined to be out-of-scope for the survey. Business-to-Business directories and Consumer directories were declared out-of-scope after it was determined that the publishers of such directories believe the listings in those directories are subsets of the listings contained in other directories for the same area. Most of the out-of-scope directories, however, bore titles that were variations of names for towns/cities in sampled PSU's. A search of PSU maps revealed that the town/city named in the title for each of these directories and the town/city with a similar name in a sampled PSU were two distinct towns/cities and, frequently, they were such a distance apart that it was unlikely that the same phone directory would serve both towns/cities.

### 3.2 Comparison of Methods

Since different methods were used to produce the final list of directories for the Complement Survey there is interest in evaluating these methods for future reference. Of special interest is evaluation of the completeness of the directories lists yielded by the different methods. The interest in evaluation did not arise, however, until after the different methods were applied. Hence, data useful for a full evaluation were not always recorded. For example, notes were not kept about which directories identified by calling directory assistance were also identified by searching PSU maps when that method was used to supplement the list produced by calling directory assistance. Hence, no total count is available for directories found by searching PSU maps. Comparisons are also hampered by incomplete applications described earlier for two methods.

Table A: PSU's and Directories to Which Directory Identification Methods Were Applied by Method

Method and Descriptive Label	Number of PSU's	Type of Directory	
		All	On Microfiche
1. Call Directory Assistance	75	x	
2. Search PSU Maps	75		x
3. Search among Directory Served Towns	38	x	

To evaluate for completeness in results from the different methods of identification, we assume that the procedure or combination of procedures producing the most directory titles for the sampled areas is the best. Table B displays available counts of directories identified in the 1982 Complement Survey by identification method and by PSU size strata. The table includes separate counts for the random subsample of 38 PSU's in which all three methods were performed as well as counts for the full sample of 75 PSU's.

Some observations about the effect of the different techniques are possible from the counts in table B. It can be seen that a total of 634 directories were identified by one or more of the three methods. The call directory assistance method yielded 444 or most directories, but a substantial number of additional directories were identified through the other methods. The call directory assistance method is inadequate by itself for producing lists of directories for sample areas. Directory assistance operators and the others contacted are apparently not well versed about the yellow pages containing listings for their respective areas.

To compare the success of the three methods, one can restrict attention to the 38 PSU's where the search among directory served places method was used. The search among directory served places yielded all but one of the directories yielded by the other two methods and it yielded 61 additional directories. Most of these 61, however, were not available on microfiche and those few that were on microfiche primarily

served boundary areas of sampled PSU's and bore titles void of names for towns/cities they served in the sampled PSU's. Many of the 61 directories, hence, would probably be identified in a search of PSU maps if that method were applied to directories unavailable on microfiche. One can, however, only speculate that a search of PSU maps would yield about the same number of directories as would a search among directory served towns if the search PSU maps was extended to directory served places as well as those places included in directory titles.

With respect to efforts required for the three methods, figures for person hours spent identifying directories were not kept except for the search among directory served places. For that method, an average of about 9.5 person hours were spent per PSU with more than 40 person hours spent on the PSU that had almost 100 yellow pages phone directories serving it, the most of any PSU among the 75 sampled. The efforts expended in the search among directory served towns are believed, however, to be no greater than those used for the other two methods of identifying yellow pages.

### 3.3 Other Observations

The reference materials about phone directories may be incomplete and thus affect the completeness of any list of directories produced for sample areas. The directories price list, which was used as the primary source list for the Complement Survey, is only claimed by its publisher to contain directories published for most phone companies. Some phone company serving a

Table B: Numbers of Yellow Pages Telephone Directories Identified for Sampled Areas by Individual or Combination of Identification Methods and PSU Population Size

Directories Counted and Identification Methods Used	PSU Size			
	Total	Largest	Medium	Smallest
	<u>Counts for 75 PSU's</u>			
Directories From:				
All Three Methods	634	273	273	88
Call Directory Assistance and Search PSU Maps Methods	573	254	238	81
Call Directory Assistance Method	444	203	164	77
Search PSU Maps Method Minus Those from Call Directory Assistance Method	129	51	74	4
	<u>Counts for 38 Subsample PSU's</u>			
All Three Methods	383	165	182	36
Call Directory Assistance Method	235	112	95	28
Search PSU Maps Method Minus Those from Call Directory Assistance Method	87	34	52	1
Search among Directory Served Towns Method Minus Those from Other Two Methods	61	19	35	7
Call Directory Assistance Method Minus Those from Search among Directory Served Towns Method	0	0	0	0
Search PSU Maps Method Minus Those from Search among Directory Served Towns Method	1	0	1	0

sampled PSU may be omitted from the price list and, hence, the yellow pages for that company would be missed in the survey. However, because every directory identified through contacts with persons in the sampled PSU's was located in the directory price list, it is believed that if there are companies whose yellow pages are omitted from the directories price list, they are probably small and have few, if any, directories serving the sampled PSU's.

Directory titles listed in the price list were sometimes outdated, but this did not cause any directory to be omitted from the Complement Survey. Some directories listed in the 1981 Phonefiche catalog bear titles that combine two or more of the titles given in the directories price list.

Directories that were not the primary ones serving communities in the PSU may have been missed if they were not listed in the community cross-reference since the directories price list cited only one directory per community. It is unlikely, though, that this caused undercoverage of facilities in the Complement Survey since firms seeking the public's business tend to have listings in the primary directories serving the communities in which they are located.

#### 4. Obtaining Identified Directories

After identifying yellow pages telephone directories the next step is to obtain them. Individual telephone companies will generally not provide directories to individuals outside the company's service area. Orders for directories must be placed with the business office of the phone company serving the requester's area.

For the 1980 Complement Survey, it took over six months after ordering some directories to receive them and certain directories were never delivered because the current editions were out of stock and new editions were not yet published. To obtain copies of yellow pages not delivered by the phone company, the contractor had their field interviewers send copies of needed directories. There were still, however, some areas for which directories were never obtained.

Because of the difficulties in acquiring yellow pages for the 1980 Complement Survey, alternatives for getting directories on a timely basis were considered. Indeed a subset of the PSU's for the National Health Interview Survey was selected for the 1982 Complement Survey to assure that there would be interviewers located in the sampled PSU's who could obtain directories if the directories could not otherwise be obtained on a timely basis.

It was found impractical to copy pages from library copies of directories where large numbers of directories were involved. In the Washington, D.C., area, the only available means of copying was coin operated machines used by the public; placing a rental copier in the library was illegal. At the New York City Public Library, the only copying permitted was that done by their copying service and that required placing a completed form on each page to be copied.

Microfiche copies of directories, however, are available from the Phonefiche System. For 1981, the System has over 14,000 U.S. directories on microfiche, which cost a total of \$3,800 in 1982.

NCHS ordered the complete set of microfiche for 1981 U.S. telephone directories in mid-March 1982. These were received in two shipments within 4 weeks. Of the directories identified for the Complement Survey and listed in the "phonefiche" catalog, however, 14 were not among those delivered.

Directories not listed in the phonefiche catalog were sought at the Martin Luther King Library in Washington, D.C. (which was earlier found to possess the most complete set of current yellow pages phone directories in the D.C. area). The staff at the NCHS installation located in North Carolina provided a copy of the directory for the PSU near them.

Table C displays the numbers of identified directories obtained by source of acquisition. Overall, about 74 percent of the directories were on the microfiche received by NCHS but the percent decreases with PSU size. About 60 percent of the directories identified for the smallest PSU's were obtained through the business office for the telephone company serving NCHS.

A total of 143 directories were ordered at three different times. The first and third orders were received six weeks after the orders were placed, contrary to the experiences in the 1980 Complement Survey. The second order was, however, never delivered; the local phone company staff believe they misplaced the second order. Directories in that order were included in the third order.

Eight out of the 143 ordered directories were not delivered. Seven of these were out of stock

Table C: Numbers of Identified Yellow Pages Telephone Directories by Source of Acquisition and PSU Size

Method of acquisition	Total	PSU Size		
		Largest	Medium	Smallest
<u>Numbers of Directories</u>				
Total	634	273	273	88
Acquired	618	265	270	83
Microfiche	470	238	205	27
Phone Company	143	26	64	53
Libraries	4	1	1	2
NCHS Staff	1	0	0	1
Never acquired	16	8	3	5
<u>Percent of Directories</u>				
Total	100.0	100.0	100.0	100.0
Acquired	97.6	97.1	98.9	94.3
Microfiche	74.1	87.1	75.1	30.7
Phone Company	22.6	9.5	23.4	60.2
Libraries	0.6	0.4	0.4	2.3
NCHS Staff	0.2	-	-	1.1
Never acquired	2.5	2.9	1.1	5.7

and the corresponding 1982 editions were not yet available. The eighth was not delivered due to an error in the directory price list. The directory price list cited "Grand River Mutual, Mo." as the directory for Grand River, Iowa, instead of the correct title "Grand River Mutual, Iowa." Three of the nondelivered directories were found at the Library of Congress with 1981 publication dates. No copy was located there for a fourth directory, which served part of the Los Angeles PSU. No search was made at the Library to locate the other four directories.

The completeness and promptness of order delivery may be due to several factors. First, information in National Yellow Pages Advertising Rates and Data, published by the National Yellow Pages Service Association [2] as well as the directory price list was used in placing the order. Second, instructions were given to deliver directories without delaying for volumes that were not readily available. Third, the 1982 volume was requested if the 1981 volume was no longer available since a more recent directory was believed better than none. Also, many of the directory volumes published in the first four months of 1982 were likely to have 1981 "corrected through" dates so that their contents would actually be current to the reference year for the Complement Survey.

Out of the 634 directories identified as serving the Complement Survey sample areas, all but 16 were finally obtained. Five of these were the directories that were ordered but not delivered or located at a library. Ten were listed in the 1981 Phonefiche catalog but the microfiche for them was not located among those received and they were never ordered.

##### 5. Summary of the Most Successful Procedures

The yellow pages phone directories were identified and acquired for a national probability area sample using several methods and sources. The best of the tried methods of identifying these directories requires several references: (a) the location code book published by the U.S. General Services Administration [6], (b) detailed maps showing PSU boundaries, and (c) cross-references listing communities and the directories serving them ([3], [4]). The procedure calls first for making a list from the location directory of all towns/cities in the sample areas. The towns/cities in that list are then sought in the cross-references between communities and directories. The directories indicated in the cross-references as serving that town/city are added to the list of directories for the sampled areas. In the event that a community served by a particular directory bears a name that is a variant of the name for some town/city in a sampled area, maps are used to determine whether the community served is the same as or located near that town/city in the sampled area or is a different community that is so geographically distant that its directory would probably not include listings from the PSU.

Two sources proved practical for obtaining large numbers of directories on a timely basis. One was the Phonefiche System [4] that offers microfiche copies of most U.S. directories, especially those for the more populated areas. These are usually available within four months after directory publication. Delivery of microfiche can be expected within two to four weeks after the order is placed. If sufficiently large numbers of directories are required for a survey, microfiche use may also save money; in our survey the directories received from the phone company were mostly small and averaged around \$3.20 when the whole set of microfiche for 1,400 1981 U.S. directories cost \$3,800 and that 1,400 tended to include the larger, more expensive directories.

One's own local phone company was the other source. To expedite delivery of directory volumes ordered through the business office of the local phone company, one should include in the order information from both the International Telephone Directories Price List [5] and the National Yellow Pages Advertising Rates and Data book [2] for each directory ordered. The order should also include express instructions to deliver ordered volumes without delaying for those volumes that are not immediately available. Delivery of most directories can then be expected about six weeks after the order is placed. However, a number of directories will likely be out of print and, thus, not delivered.

##### REFERENCES

- [1] "Complementing the Survey of Hospitals and Other Health Institutions (1980 Complement Survey)". Unpublished Project Report, RTI No. 255U-1913. Research Triangle Institute. 1981.
- [2] National Yellow Pages Advertising Rates and Data. National Yellow Pages Service Association, 888 W. Big Beaver Road, Troy, Michigan 48084. Telephone (313)362-3300. 1981.
- [3] 1981 Community Cross-Reference Guide to Phonefiche. Newspaper Indexing Center, Micro Photo Division, Bell and Howell Company, Wooster, Ohio. 1981.
- [4] 1981 Phonefiche. Newspaper Indexing Center, Micro Photo Division, Bell and Howell Company, Wooster, Ohio. 1981.
- [5] 1982 International Telephone Directories Price List. AT&T Long Lines, International Directory Group, Morris Plains, N. J. 1982.
- [6] Worldwide Geographic Location Codes. General Services Administration, Office of Finance. Superintendent of Documents. U.S. Government Printing Office, Washington, D.C. Stock No. 022-000-00130-2. June 1981.