NATIONAL BUREAU OF STANDARDS

NBS TECHNICAL NOTE 921

U.S. DEPARTMENT OF COMMERCE / National Bureau of Standards

Survey Manual for Estimating the Incidence of Lead Paint in Housing

NATIONAL BUREAU OF STANDARDS

The National Bureau of Standards[1] was established by an act of Congress March 3, 1901. The Bureau's overall goal is to strengthen and advance the Nation's science and technology and facilitate their effective application for public benefit. To this end, the Bureau conducts research and provides: (1) a basis for the Nation's physical measurement system, (2) scientific and technological services for industry and government, (3) a technical basis for equity in trade, and (4) technical services to promote public safety. The Bureau consists of the Institute for Basic Standards, the Institute for Materials Research, the Institute for Applied Technology, the Institute for Computer Sciences and Technology, and the Office for Information Programs.

THE INSTITUTE FOR BASIC STANDARDS provides the central basis within the United States of a complete and consistent system of physical measurement; coordinates that system with measurement systems of other nations; and furnishes essential services leading to accurate and uniform physical measurements throughout the Nation's scientific community, industry, and commerce. The Institute consists of the Office of Measurement Services, the Office of Radiation Measurement and the following Center and divisions:

Applied Mathematics — Electricity — Mechanics — Heat — Optical Physics — Center for Radiation Research: Nuclear Sciences; Applied Radiation — Laboratory Astrophysics[2] — Cryogenics[2] — Electromagnetics[2] — Time and Frequency[2].

THE INSTITUTE FOR MATERIALS RESEARCH conducts materials research leading to improved methods of measurement, standards, and data on the properties of well-characterized materials needed by industry, commerce, educational institutions, and Government; provides advisory and research services to other Government agencies; and develops, produces, and distributes standard reference materials. The Institute consists of the Office of Standard Reference Materials, the Office of Air and Water Measurement, and the following divisions:

Analytical Chemistry — Polymers — Metallurgy — Inorganic Materials — Reactor Radiation — Physical Chemistry.

THE INSTITUTE FOR APPLIED TECHNOLOGY provides technical services to promote the use of available technology and to facilitate technological innovation in industry and Government; cooperates with public and private organizations leading to the development of technological standards (including mandatory safety standards), codes and methods of test; and provides technical advice and services to Government agencies upon request. The Institute consists of the following divisions and Centers:

Standards Application and Analysis — Electronic Technology — Center for Consumer Product Technology: Product Systems Analysis; Product Engineering — Center for Building Technology: Structures, Materials, and Life Safety; Building Environment; Technical Evaluation and Application — Center for Fire Research: Fire Science; Fire Safety Engineering.

THE INSTITUTE FOR COMPUTER SCIENCES AND TECHNOLOGY conducts research and provides technical services designed to aid Government agencies in improving cost effectiveness in the conduct of their programs through the selection, acquisition, and effective utilization of automatic data processing equipment; and serves as the principal focus within the executive branch for the development of Federal standards for automatic data processing equipment, techniques, and computer languages. The Institute consists of the following divisions:

Computer Services — Systems and Software — Computer Systems Engineering — Information Technology.

THE OFFICE FOR INFORMATION PROGRAMS promotes optimum dissemination and accessibility of scientific information generated within NBS and other agencies of the Federal Government; promotes the development of the National Standard Reference Data System and a system of information analysis centers dealing with the broader aspects of the National Measurement System; provides appropriate services to ensure that the NBS staff has optimum accessibility to the scientific information of the world. The Office consists of the following organizational units:

Office of Standard Reference Data — Office of Information Activities — Office of Technical Publications — Library — Office of International Relations — Office of International Standards.

---

[1] Headquarters and Laboratories at Gaithersburg, Maryland, unless otherwise noted; mailing address Washington, D.C. 20234.

[2] Located at Boulder, Colorado 80302.

Survey Manual for Estimating the Incidence of Lead Paint in Housing

William G. Hall

Institute for Basic Standards
National Bureau of Standards
Washington, D.C. 20234

and

Lillian T. Slovic

Institute for Applied Technology
National Bureau of Standards
Washington, D.C. 20234

Prepared for:

Office of Policy Development and Research
The Department of Housing and Urban Development
Washington, D.C. 20410

U.S. DEPARTMENT OF COMMERCE, Elliot L. Richardson, *Secretary*

Edward O. Vetter, *Under Secretary*

Dr. Betsy Ancker-Johnson, *Assistant Secretary for Science and Technology*

NATIONAL BUREAU OF STANDARDS, Ernest Ambler, *Acting Director*

Issued September 1976

National Bureau of Standards Technical Note 921
Nat. Bur. Stand. (U.S.), Tech. Note 921, 111 pages (Sept. 1976)
CODEN: NBTNAE

U.S. GOVERNMENT PRINTING OFFICE
WASHINGTON: 1976

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D. C. 20402
(Order by SD Catalog No. C13.46:921). Stock No. 003-003-01660-8 Price: $2.20.
(Add 25 percent additional for other than U.S. mailing).

TABLE OF CONTENTS

Page

1. List of Tables v

2. List of Figures. vi

Acknowledgement. vii

Abstract 1

Introduction 2

Part I - Preliminary Considerations. 5

o Overview of the Survey Process. 6

o Cost Determinants 7

o Establishment of Objectives 9

o Required Resources. 10

- Supervisors 10

- Data Base Manager 11

- Inspection Staff. 11

- Statistical Consultant. 12

- Clerk-Telephonist 12

- Equipment 12

Part II - Preparing for and Performing the Survey. 14

o Selecting the Sample. 15

- Random Sampling and the Binomial Distribution . . . 16

- Sample Accuracy 17

- Sample Size 17

- Non-Sampling Errors 22

o Proper Timing for the Survey. 25

o Characteristics to Look for in Hiring Inspectors. . . . 26

o Supervisor Training 27

Page

o Inspector Training 30

o The Survey Questionnaire 33

o Publicity Campaign 34

o Pretest. 37

o Record Keeping 37

o Notification of Residents. 38

o Supervisor Follow-up 42

Part III - Data Reduction and Analysis. 43

o Data Base Organization 44

o Construction and Maintenance of the Date Base. 45

o Estimating the Incidence of Lead Paint on Surfaces . . .50

Recommended Reading for Further Information on Surveys and Statistics. 53

Appendix A - Determination of Sample Size 54

Appendix B - Estimating Incidence of Lead Paint 73

Appendix C - Instructions for Completing the Sample Survey Form. 76

Appendix D - Specifications for Sample Generation Program. 84

Program Listing. . . . .

Specimen of Output . . . .

LIST OF TABLES

Page

1. Relation Between Sample Size and Precision - 95 Percent Confidence Intervals 19

2. Relation Between Sample Size and Precision - 90 Percent Confidence Intervals 20

3. Relation of Sample Size (n) to Hazard Fraction (p) for Various Population Sizes (N) 57

4. Percentage of Normal Distribution Lying Within Multiples of σ 58

LIST OF FIGURES

Page

1. Inspection Questionnaire Form. 28

2. Sample Introductory Letter 36

3. Call-Back Card 41

ACKNOWLEDGEMENT

The authors express their thanks to Harvey W. Berger of the Office of Housing and Building Technology at the National Bureau of Standards for his assistance in preparing this manual.

Special thanks go to Janet Bonk of the Allegheny County Health Department's lead paint survey project, for her many suggestions on survey administration.

Survey Manual for Estimating the Incidence of Lead Paint in Housing

William G. Hall & Lillian T. Slovic

ABSTRACT

This Manual is intended as a guide for municipal managers in performing a survey to determine the prevalence of lead based paint in their community's dwelling units. There are four parts to the Manual, each is intended for a different audience.

Part I discusses the preliminary considerations for a survey. It is intended for the department head or executive who will initiate plans for the survey. It presents a managerial overview of the processes, the cost determinants, criteria for the establishment of objectives and the resources required.

Part II is intended for the survey manager and the inspector supervisors. It contains more detailed information on the planning, staffing, training, and execution of the data collection phase of the survey.

Part III is for the use of the person responsible for the control and management of the data collected and for the analysis of these data.

The Appendices contain quite detailed information about procedures we have used in previous surveys. These may be used as they are described or may be modified or adapted to meet specific objectives.

Key Words: Lead paint; lead paint detection; lead paint programs; lead poisoning; portable x-ray fluorescence; random sampling; x-ray fluorescence.

INTRODUCTION

Lead poisoning, especially of children, is a serious national health problem. The sources of lead to which children are exposed include: air (industrial and automobile exhausts); water (industrial waste, minerals, and leaded pipes); food (packaging or storage containers; the food product itself), or dust (particulate fallout). Another source of lead intake in children is the ingestion of lead based paint.

Before the 1940's, paints with as much as 50 percent lead were widely used for residential applications. According to the 1970 Housing Census 1/, approximately 1.8 million houses, which were built before 1940, are in a dilapidated condition. Dilapidation denotes chipping, flaking and peeling paint as well as decaying walls and wood surfaces. Under these conditions lead paint if present is readily available for inhalation and ingestion and constitutes a potential health hazard.

Accurate diagnosis and careful treatment of lead poisoned children will not eliminate the problem if they are returned to the same environment to be reexposed to the hazard. The rate of recurrence of lead poisoning is high among such children. Reexposure of survivors of acute poisoning, to an unabated lead source, results in permanent damage almost 100 percent of the time 2/

1/ Census Bureau, U. S. Department of Commerce, Plumbing Facilities and Estimates of Dilapidated Housing, HC (6), 1973

2/ Jane S. Lin-Fu, "Childhood Lead Poisoning--An Eradicable Disease" Children: An Interdisciplinary Journal for the Professions Serving Children, Vol. 17, No. 1, January - February 1970.

The only sure preventative for the ingestion of lead paint, on housing surfaces, by children, is elimination of the hazard, either by removing it or covering it up with barrier type building materials. This is an expensive procedure--costing from hundreds up to several thousands of dollars per dwelling--but perhaps not so expensive for society as the care of youngsters who have suffered permanent physical or brain damage from lead paint poisoning.

If a serious lead paint poisoning problem has been identified in your community through discovery of a high lead poisoning incidence and/or by a child screening program, you may be looking for a strategy to combat it, either as an independent lead poisoning control program or as part of a broader housing improvement effort such as a code enforcement program. Either way, lead paint hazard abatement will affect your overall housing program in terms of funding and staff commitments. You may find that the addition of lead hazard elimination as part of your housing improvement program makes a raze-rebuild strategy more cost effective than a rehabilitation strategy, especially if the latter seemed only marginally feasible on economic grounds. Whatever course of action is anticipated, the first step is to estimate the number of contaminated housing units, the extent of contamination, and housing characteristics which may be associated with lead paint hazards (location, age, occupancy class, type, etc.). The sample survey techniques described in this manual can help you accomplish this much more conveniently and with less expense than the alternative--a unit by unit census.

Although this Manual is not intended to be used as a general guide for household surveys, some of the material is applicable to such activities. The text includes description of ways in which this survey differs from the usual opinion or household surveys; the focus on dwellings rather than people, the collection of physical measurements rather than interviews are distinguishing characteristics of this survey. A detailed procedure for drawing a sample is presented. Although it is by no means the only procedure which can be used it was designed to meet the specific objectives of estimating the hazard in a number of housing categories with a high statistical confidence level, with moderate expenditure of resources.

PART I

PRELIMINARY CONSIDERATIONS

OVERVIEW OF THE SURVEY PROCESS

There are many possible motivations for municipal officials to consider performing a lead paint survey in housing. The survey may be health-oriented or code enforcement oriented; it may be part of a large neighborhood or civic improvement program or an independent lead abatement program; it may be focused on the entire city or some particular target group or groups for which the city has special legal or social responsibilities. The findings of the survey may be required either within an agency or at a higher echelon of government; they may be used to allocate existing resources more efficiently in an operational sense, to permit cost effective long range planning for a lead program, or to estimate the benefits and costs of a lead abatement program relative to the benefits and costs of alternative action programs. Regardless of the conditions leading to the use of a sample survey, the following tasks should be accomplished sequentially to assure that the survey will be efficient and adequate.

1) Determine the objectives...Why is the information needed? What decisions will be based on the information gathered?
2) Define the housing units to be studied...Are all housing units to be included? Only owner-occupied units? Only rental units?
3) Determine the data to be collected and verify that they do not exist elsewhere...What data are needed? Can they be gathered by using data from preceding or current studies?
4) Choose the sample unit...Should it be the building? The dwelling unit? What size sample is needed? What sampling method is to be used? What degree of uncertainty can be tolerated?

5) Design and test the survey questionnaire or data collection form...Is it unambiguous? Is it clear? Does it minimize discretionary entries? Is it convenient for inspectors? Is it convenient for editing, transcription, and keypunching? Should abbreviations and/or conventions appear on the form or in a companion manual?

6) Determine the method of resident contact...Should appointments be made for inspectors or should they make unannounced calls? How should publicity be handled?

7) Acquire and train staff

8) Carry out the data collection

9) Edit the data for consistency and completeness.

10) Deal with nonresponses (units for which inspection cannot be performed.)

11) Analyze the data

12) Report the findings, including recommendations

Although these steps are in sequence, it should be understood that they are not independent.

Many skills and talents are required and they must be well coordinated if the survey is to be successful. A useful procedure is to have all who are expected to participate in the survey, or use the data collected engage in a series of give-and-take discussions during the planning stages. "What if" questions should be encouraged. These discussions should result in a useful, practical and economical survey design.

COST DETERMINANTS

The most critical determinant of survey cost is the size (number of dwelling units to be inspected) of the sample. The sample size is also the most important factor in the accuracy of the survey. There is therefore a direct trade-off of cost vs. accuracy. The accuracy requirements must be understood by all involved; insistence on more accuracy than that required will result in unnecessary expense; less accuracy than is required will result in an inadequate or even useless survey.

The following is a list of functions which may be considered as distinct elements for costing:

1. Planning and design
2. Management
3. Development of population list
4. Selection of sample
5. Development of data collection form.
6. Acquisition of lead measurement devices
7. Pretest
8. Printing
    a) Data collection forms
    b) Training materials
    c) Manuals
    d) Contact letters
    e) Call back forms
9) Inspection
    a) Recruitment
    b) Training
    c) Dwelling inspection
    d) Inspector supervision
    e) Inter-unit travel

10) Development of editing, coding and analysis procedures (manual or computerized)

11) Editing, coding, and keypunching

12) Telephone

13) Mailing

14) Analysis of data including programming and computer costs.

15) Report writing

16) Reproduction, dissemination, and presentation of report

As with the steps of the survey process, these functions are distinct but interdependent. The size of the sample will influence the methods by which the functions are to be accomplished. The sample size controls the scale of the entire operation. If the sample size is small, for example, editing will be manual rather than computerized, documents can be reproduced by office equipment rather than printing, a smaller organization is required.

ESTABLISHMENT OF OBJECTIVES

The first, most important, and most difficult step of a survey is the establishment of objectives. The objectives must be defined in terms of the purpose or purposes for which the collected data are to be used. The survey should provide a sound basis for decision making. The decision could be: the selection of a course of action from alternative proposals; defining a course of action or, allocation of resources. It cannot be overemphasized that the establishment of objectives is a policy function and one that should be performed at a high policy making level. Although the services of a survey statistician are required in the process, his role is to quantify and interpret, and to estimate costs. His job is not to establish objectives. The interaction between the statistician and the policy maker must be collaborative.

REQUIRED RESOURCES

General

The survey can be carried out completely with "in-house" staff. However there are some tasks, within the survey, which could be contracted to individuals, firms, or other parts of your agency. These opportunities are identified below.

Supervisors

Choosing supervisors is extremely important for their performance is crucial to the success of the project. They are responsible for hiring and training the inspectors, maintaining the work flow, scheduling inspection team assignments, checking the equipment, reviewing forms, and following departmental policies and procedures.

Ideally one would like the supervisors to have experience in survey supervision and to be knowledgeable in housing and public health matters. The former, however, is the more critical. These people may have to be recruited from outside the department, or the most nearly qualified candidates may have to be trained in one or more of the critical areas.

Some projects have hired university graduate students in the appropriate fields for these supervisory positions. However, in our experience this has not been a satisfactory approach. The students often have little work background, even less supervisory experience and have difficulty maintaining authority over people of their own age and educational background.

The capability of people that you hire for these positions will, of course, depend on the candidates available. Since these are very important positions for the survey, every effort should be made to obtain the most qualified candidates.

One supervisor has been found to be appropriate for each group of six inspection teams, that is, one supervisor for twelve inspectors.

Data Base Manager

The data base manager (computer analyst) is responsible for programing, designing and operating the system which stores the data for future analysis. This person might be detailed from a budgeting department or some other department of your agency for the duration of the project rather than hired from the outside.

This is one of the functions which could be performed by an outside contractor, such as a reliable data analysis group either inside or outside of government. The same contract could also include the services of a survey statistician to design and perform the data analysis.

Inspection Staff

The number of inspectors needed can be determined from the number of houses to be inspected. For instance, if 5-6 completed inspections per day per two-man team are assumed (in our experience this is a reasonable estimate), then 165-200 team days are needed to complete 1000 inspections. If the inspectors are hired for a three month period, they will be able to work 60-65 days. Thus, 3 teams would be needed to perform 1000 inspections during those three months. These estimates are based on a five day-forty hour work week, but the inspection teams will in fact be required to perform some inspections during evenings and weekends at the convenience of the occupant.

Statistical Consultant

A survey statistician will be needed at various times during the planning and execution of the survey. During the planning stage, he should be involved in the establishment of objectives; in discussing all of the questions concerning the sample and in the design and testing of the questionnaire. The statistician should, with the data base manager or the data analysis group, design the analysis procedures to be used.

If possible the statistician should be continuously available on a consultant basis, both during the survey, to help in solving whatever problems arise, and during the analysis phase to help interpret the results. Possible sources for this consultant include your agency, a nearby university or a commercial firm.

Clerk/Telephonist

Usually a department will have a clerical worker who can become part of the lead survey group. In addition to maintaining record files, that person also can be in charge of mailing, maintaining mailing records, and arranging inspection appointments. Just as the inspectors may have to work Saturdays and evenings, so might the telephone scheduler. If all the adults in a household work, the only time they can be reached for an appointment is at night or on the weekend.

Equipment

Portable X-ray fluorescence (XRF) lead detectors are the most commonly used and acceptable means for making large numbers of measurements of lead paint in housing. They are non-destructive, and an inspector can perform a measurement every 10-20 seconds. A digital readout displays the measurement in terms of milligrams of lead per square centimeter of surface area (mg/cm^2). These devices are limited in accuracy and durability.

They will measure the lead content of surfaces with good reliability (with an error of less than 10%) above a level of 2 mg/cm^2. Below that level, readings are suspect but can be confirmed with multiple readings at the same location.

New XRF devices of increased accuracy and improved general performance are under commercial development. If present generation instruments are used for the survey, it would be advisable to have a repair service agreement with the manufacturer. Each survey team will need a detector. In addition, one or two spare detectors should be kept on hand to substitute for those returned to the manufacturer for repair.

There are other methods that can be used for the detection of lead paint in houses. Chemicals such as sodium sulfide will react with lead compounds to form dark colors which are indicative of the presence of lead. Chemical spot tests, however, are not quantitative and can only give a rough indication of lead's presence. Under some circumstances they will give false negative reactions to lead, that is, indicate that no lead is present when in fact it is present.

Lead in paint can be accurately analyzed in the laboratory using sophisticated chemical techniques and instruments. Such procedures, however, require that paint chips be scraped or peeled from a number of surfaces, catalogued and transmitted to a laboratory for analysis. These alternatives to the use of XRF detectors, in addition to being very time consuming and expensive, are likely to be unacceptable to most dwelling occupants.

PART II

PREPARING FOR AND PERFORMING THE SURVEY

SELECTING THE SAMPLE

Simple random sampling from a city housing director 1/ has been found to be an efficient means of choosing residences to survey. It is the method least likely to be affected by bias. Every unit has an equal likelihood of being selected, with prior selections having no bearing on subsequent choices.

Housing directories are available for many United States cities in the population range of 10,000 to 1,000,000 and can be found at libraries and real estate offices. They are updated every two or three years. In a lead based paint survey, where the emphasis is on the older housing stock, the omission of new housing units built after the most recent directory publication can be overcome by including a correction for population size in the analysis or by using a better population size estimate from a more up-to-date source.

Most directories are divided into two sections: the first is an alphabetical listing of all city residents and their addresses; the second, an alphabetical listing of streets with the street numbers listed numerically. In conducting the survey, it is easier to go from street to street than from name to name, and so the second directory section is more useful.

Errors do exist in these lists, but our experience has been that there are not enough of them to invalidate the analysis. Someone with a fair knowledge of the city should make a cursory check of the directory, however, to assure that no major publication errors exist, such as the omission of an entire section of the city. In any case it is ultimately the responsibility of the survey statistician to reconcile the population data source to the survey requirements.

1/Several procedures from drawing samples are described in Appendix A.

Random Sampling and the Binomial Distribution

A door to door, unit by unit census of all dwelling units in a city, to determine the extent and distribution of the community's lead paint hazard is an extremely costly and time consuming process which most jurisdictions will want to avoid. Random sampling is faster and more economical and can yield very accurate results.

Random sampling involves choosing a fraction of all the dwelling units from some population in such a manner that each unit has an equal chance of being selected. By applying the mathematics of probability to the sample findings, the relevant characteristics (lead paint content in this case) for the totality of dwellings can be estimated. In a lead paint survey, it may be desirable to deal with a number of subpopulations. The incidence of lead paint may differ according to some physical characteristic of the unit (age, for example) or the objectives of the survey may require that incidence of lead paint be estimated separately for different tenancy groups (public, subsidized, owner-occupied, etc.).

You, the planners and administrators of the lead paint survey, should have an understanding of the sampling process even if you plan to have a statistician handle this portion.

A basic statistical concept which underlies the remainder of this section (and Appendix A) is that of a binomial distribution. If a characteristic of interest is defined in terms of yes-no, true-false, or hazardous-non-hazardous, the characteristic is binomially distributed. The binomial distribution accurately reflects the ways in which laws or codes are usually written and enforced. A typical code might require that a surface with more than 2.0 mg/cm^2 of lead be deleaded according to some acceptable procedures but would not require distinct action for each lead level, such as 2.5, 5.0, or 10.00 mg/cm^2. The binomial distribution does not accommodate the notion of degrees or gradation of hazard, but can only distinguish between hazardous and non-hazardous levels.

Sample Accuracy

There are many ways in which inaccuracies may be introduced into the estimates resulting from a survey. Errors are of two general types: those which arise because of the use of a sample to present an entire population and those which arise from the data collection and/or analysis procedures and should be expected to occur even if a complete census were taken. Errors of both types will occur; they cannot be completely eliminated. They must be controlled to the extent that the error is either negligible or predictable. Errors introduced by the use of a sample can be estimated and can be controlled to produce any accuracy required by use of the methods of statistics. The effect of errors introduced from sources other than the sampling procedure is not as easy to predict; if however, the number of such errors is kept small, their effect will be negligible. The term "accuracy" refers to all errors; the term "precision" refers only to sampling errors (the more precise, the smaller the error).

Sample Size

If the size of the sample (hereafter denoted by n) is close to the number of dwelling units (called the total population and denoted by N), precision, as one would expect, is very good. However, in almost all practical situations the sample may be quite small relative to the total population. The sample size is the dominant factor in the statistical precision of the sample; the size of the total population does not significantly affect the precision 1/. The Gallup and Harris polls, for example, interview only about 1500 persons to represent nationwide opinion.

Maximum efficiency, in terms of precision per sample unit, can be attained by the use of a simple random sample 2/. The basic principle that produces accurate estimates from small samples is that each unit of the total population must be equally likely to appear in the sample.

There is some element of risk involved in accepting the findings from any sample as representative of the total population. Tables 1 and 2 illustrate the magnitude of this risk. If a very large number of samples of the indicated size were drawn from an infinite population, 95 percent (Table 1) of these samples would contain the true percent of hazardous units of the total population within the given confidence intervals. For any given sample, there is a five percent risk that the true value is not contained within the given confidence intervals. For example, if the percentage of hazardous dwellings in the sample is 10 percent, a sample size of 100 insures that there are 95 chances out of 100 that the percentage of hazardous dwellings for the total population lies between 4.1 and 15.9 (Table 1), whereas there are 90 chances out of 100 that the true percentage of hazardous dwellings lies between 5.0 and 15.0 (Table 2).

Note that an increase in sample size (for a given percentage of hazardous dwellings) always improves the precision, but the improvement

---

1/ This assumes the total population is "large" relative to the sample. See Appendix A for procedures appropriate for small populations.

2/ Other sampling schemes beyond the scope of this manual may do as well in terms of precision per sample unit and better in terms of precision per dollar spent. Cluster sampling for example, will lower the unit cost of inspection, but will require a larger sample to achieve the same precision. Stratified samples may be considered as combinations of simple random samples. These more sophisticated techniques may be appropriate for particular objectives.

Table 1

Relation Between Sample Size and Precision

95 Percent Confidence Intervals

True % Hazardous Units	Sample Sizes					
	50	100	200	400	500	1000
2	0-5.9	0-4.7	0.1-3.9	0.6-3.4	0.8-3.2	1.1-2.9
5	0-11.1	0.7-9.3	1.9-8.1	2.8-7.2	3.1-6.9	3.6-6.4
10	1.8-18.2	4.1-15.9	5.9-14.1	7.1-12.9	7.5-12.5	8.1-11.9
20	8.8-31.2	12.2-27.8	14.5-25.5	16.1-23.9	16.5-23.5	17.5-22.5
50	36.1-63.9	40.2-59.8	43.1-56.9	45.1-54.9	45.7-54.3	46.9-53.1
80	68.8-91.2	72.2-87.8	74.5-85.5	76.1-83.9	76.5-84.5	77.5-82.5
90	81.8-98.2	84.1-95.9	85.9-94.1	87.1-92.9	87.5-92.5	88.1-91.9
95	88.9-99.9	90.7-99.3	91.9-98.1	92.8-97.2	93.1-96.9	93.6-96.4
98	94.1-99.9	95.3-99.9	96.1-99.9	96.6-99.4	96.8-99.2	97.1-98.9

Table 2

Relation Between Sample Size and Precision

90 Percent Confidence Intervals

True % Hazardous Units	Sample Sizes 50	100	200	400	500	1000
2	0-5.3	0-4.3	0.4-3.6	0.8-3.2	1.0-3.0	1.3-2.7
5	0-10.1	1.9-8.1	2.4-7.6	3.2-6.8	3.4-6.6	3.8-6.2
10	3.1-16.9	5.0-15.0	6.5-13.5	7.5-12.5	7.9-12.1	8.4-11.6
20	10.6-29.4	13.4-26.6	15.9-24.1	16.7-23.3	17.0-23.0	17.9-22.1
50	38.3-61.7	41.8-58.2	44.2-55.8	45.9-54.1	46.4-53.6	47.4-52.6
80	70.6-89.4	73.4-86.6	75.9-84.1	76.7-83.3	77.0-83.0	77.9-82.1
90	83.1-96.9	85.0-95.0	86.5-93.5	87.5-92.5	87.9-92.1	88.4-91.6
95	89.9-99.9	91.9-98.1	92.4-97.6	93.2-96.8	93.4-96.6	93.8-97.2
98	94.7-99.9	95.7-99.9	96.4-99.6	96.8-99.2	97.0-99.0	96.3-98.7

is only modest in going from a reasonably large sample to a still larger one. For example, in Table 1, for 50 percent hazardous units, doubling the sample size from 50 to 100 narrows the confidence interval by 8.2 percentage points wheras doubling the sample size from 500 to 1000 narrows the confidence interval by only 2.4 percentage points. Note that the case in which 50 percent of the units are hazardous yields larger confidence intervals than any other case for a given sample size. Thus it represents the worst case for a given sample size. Note also that the magnitude of the confidence interval when m percent of the units are hazardous is equal to that for (100-m) percent, and becomes smaller as m approaches zero or 100.

There is no "right" confidence level. The particular level to aim for - 95%, 90%, or some other level - is a policy decision which must be based on the following factors:

1. The cost of performing the survey at various sample sizes.
2. The confidence interval which is required. This is a threshold determination problem; presumably a different course of action will be undertaken depending on whether the percentage of hazardous units is greater or less than m%. How precisely must m be known in order to make a good decision? What are the consequences if m is off by 5% or by 10%, etc.?
3. The risk involved for example, in adopting a course of action based on a 90% confidence level rather than on a 95% or 99% one. If alternative courses of action differ greatly in cost (whether in dollars, time, difficulty, or whatever), the acceptable risk should be small.

Non-Sampling Errors

Non-sampling errors may be either conceptual or mechanical. There are no statistical formulae which can predict how seriously these errors will affect the survey results. Such errors cannot be totally avoided or controlled, and great care at every stage of the survey is essential to minimize them. This care includes, of course, conscientious performance of all tasks; it also requires the development and implementation of procedures at extra cost, for assuring that the number of errors is small. Your consulting statistician can furnish valuable advice based on your specific objectives and circumstances. Here are some of the non-sampling errors which are likely to arise in a lead paint hazard survey:

1) Non-random sampling

Any factor which introduces an unknown bias into the sample can introduce errors into the findings. For example, use of a telephone directory for the selection of households would preclude selection of households with no telephone and those with a unlisted telephone number, and would offer a multiple chance of selection of households with more than one telephone.

2) Inaccuracy of listing

If the list from which the sample is drawn is incomplete or seriously out of date, errors will be introduced into the findings. For example, a city directory cannot be used to select a sample to represent a metropolitan (city and suburbs) area.

3) Survey questionnaire ambiguities

The data collection procedure must be completely reproducible: that is to say, any team of inspectors should produce equivalent data for the same housing unit. This requires a good form design and well-trained inspectors.

A common procedure in many household surveys is to re-interview a sample of the respondents. Re-inspection is not recommended for a lead paint survey for the following reasons. First, the burden on the occupant (time) and the intrusion into his privacy (each room of the dwelling must be entered and inspected) is greater than in the usual survey. Second, to include enough information on the data collection form to uniquely identify each surface and room of the dwelling (for re-measurement) requires a significant increase in the volume of data collected, and these extra data have no intrinsic value. A post-inspection interview, performed by the supervisors, can and should be done for some fraction of the respondents. This can be either a doorway or telephone interview. This should include verifying all of the questionnaire information supplied by the respondent. Also it should include questions such as:

When did the inspectors arrive?

How long did the inspection take?

Were all rooms of the unit inspected?

The inspectors will be aware of the post-inspection interview; thus it has value as a supervisory device as well as checking on the consistency of the respondents.

4) Detection instrument errors

This is one of the errors peculiar to a lead paint survey. The detection devices are simply not as accurate and reliable as one would like. This source of error <u>cannot</u> be eliminated, but the error can be minimized by use of proper calibration procedures and good inspector training.

5) Inspector performance

Most of the data gathered in the survey are instrument readings; the remainder are simple "yes-no" or answers to multiple choice questions, so there is somewhat less chance for the inspector to "lead" the respondent into the "right" answer. The consistency and reliability of the inspectors in making all of the appropriate and necessary measurements of painted surfaces, in the inspected dwellings, is of major importance in minimizing survey errors. In addition to diligence, they must be accurate in recording lead measurements and other pertinent information on the survey questionnaire.

6) Respondent accuracy

This is related to inspector performance and questionnaire ambiguities. This factor can be a serious problem in a general household survey but is less likely to be serious in a lead paint survey.

7) Mechanical and clerical errors

These errors will occur; they include transcription and coding errors (including omissions), key punch errors, illegibility, loss,destruction or mutilation of data collection forms and arithmetic errors in manual tabulation. Although they will presumably be random and therefore will tend to cancel each other, all reasonable precautions to avoid these errors should be taken. Such precautions should include keypunching directly from the data collection form (no transcription errors can occur if transcription is avoided), verifying all card punching, close monitoring of all forms by supervisors, etc.

8) Non-existent unit

A dwelling unit selected into the sample may not exist. This could be due to an address error, demolition of a unit, conversion to non-residential use, etc.

9) Non-responses

Inevitably, entry to some dwellings will be impossible. The unit may be vacant, the occupant may not be at home or the occupant may refuse entry to the inspectors. Every effort should be made to keep non-responses to a minimum, but for each non-response, those data which can be obtained by observation should be collected. These data are useful in detecting any systematic bias induced by non-responses. The actions required differ according to the cause of the non-response. For non-existent, vacant, or non-residence units, nothing can be done; for not-at-homes, efforts should be made to reschedule; for refusals, rescheduling should be attempted and the occupant's reason for refusal should be determined -- this information may be useful in retraining, pairing of inspection teams, etc.

PROPER TIMING FOR THE SURVEY

After you have established the objectives, the number of units to be inspected, and the data to be collected for each unit, planning for the operational part of the survey can begin.

Our experience in surveys for lead paint has shown that the project is most efficiently performed during the summer. There are no slushy or icy streets and no heavy coats and boots to encumber the inspectors, and most importantly, householders are more likely to permit inspection in good weather. In some temperate climates, the season may be of little importance in choosing the period for the survey.

College students who plan to return to school in the fall tend to work faster than other inspectors who know that the sooner the project ends, the sooner they will be unemployed. During one survey, for example, student teams completed an average of 5.5 units inspections a day while other teams averaged only 4 units.

There are, of course, administrative and managerial problems involved in such a concentrated effort. The administrator may be running several other programs and have only a limited amount of time to spend on the lead paint survey. Money may be tightly budgeted. But it is most efficient, in the long run, to hire as many inspectors as is necessary to complete the survey during the summer.

CHARACTERISTICS TO LOOK FOR IN HIRING INSPECTORS

Experience in previous surveys has shown that the teams most successful in being accepted by housing occupants, were made up of a male and a female; or a black and a white. You should, therefore, consider these public preferences in hiring.

While enthusiasm and courtesy are extremely important, a certain adaptability and resilience are indispensable in dealing with the public. Inspectors should not overreact for example, if an occupant answers the door clad only in a wristwatch, as has happened.

One requirement for employment may have to be ownership of a car. An inspection team must have a car to get to its daily assignments. If the sponsoring agency cannot provide each team with a vehicle, then the inspectors will have to use their personal cars and be reimbursed for expenses as agreed. The inspectors should have a general knowledge of the city and be able to get around in it.

The prospective inspectors should also understand, when they are hired, that they may be required to work irregular shifts and/or Saturdays. Since a great many occupants have full-time jobs, they are not available between 9 AM and 5 PM, to open their homes for inspection.

Also remember, in hiring, that the lead detection equipment weighs approximately twenty pounds. Your inspectors will have to be able to carry this weight a good part of the day.

SUPERVISOR TRAINING

While the publicity campaign (page 34) is still underway and approximately ten to fifteen days before the inspectors come on board, the supervisors should begin training for the survey. If they are not part of the department already, and are not familiar with the lead paint problem, they must be completely familiarized with the project, besides learning all the particulars of running the survey.

Since the supervisors will be leading the training sessions for the inspectors they will have to learn how to operate the XRF lead detector, how to complete the survey forms, and generally how to conduct the inspection. This includes learning all about the various housing materials and conditions, so that they can instruct the inspectors in filling out the survey form (Figure 1) and check the completed forms for accuracy. (This is why people with experience in housing were suggested for this job.)

The pretest (a mini-survey held in advance to check for problems in the survey plan) would most conveniently be held at this time. It's an important testing ground for all the tasks that the supervisors must learn and later teach the inspectors. During this period the supervisors gain first-hand experience of the actual survey tasks as well as insight into the human problems that are likely to be encountered during the survey.

NBS-744
(REV 7-74)

U.S. DEPARTMENT OF COMMERCE
NATIONAL BUREAU OF STANDARDS

Figure 1
Inspection Questionnaire Form

SURVEY OF LEAD-BASED PAINT HAZARD IN DWELLING UNITS

SECTION I - IDENTIFICATION

ADDRESS

SERIAL NUMBER	CODE	TRACT	BLOCK	ZIP CODE	V	STREET NAME AND NUMBER
1 - 6	7 - 8	9 - 14	15 - 17	18 - 22	23	24 - 59
	0 1					

CODE	XRF SERIAL	TEST BLOCK	ZERO READING	ZERO VERNIER	CALIBRATE READING	CALIBRATE VERNIER	PEAK READING	PEAK VERNIER	INSPECTORS	INSPECTORS	DATE MONTH	DATE DAY	
7 - 8	9 - 11	12 - 14	15 - 17	18 - 20	21 - 23	24 - 26	27 - 29	30 - 32	33 - 35	36 - 38	39 - 41	42 - 44	45 - 59
0 2													

CODE	SEQ	TEST BLOCK	TYPE	XS	OC	YEAR	OWNER	MORT	PUBLIC	SUB	CHILD		
7 - 8	9 - 11	12 - 14	15 - 17	18 - 20	21 - 23	24 - 26	27 - 29	30 - 32	33 - 35	36 - 38	39 - 41	42 - 44	45 - 59
0 3													

SECTION II - INTERIOR

	CODE	WALLS AND CEILING COND	READING WALL 1	READING WALL 2	READING WALL 3	READING WALL 4	READING CEILING	TRIM COND	WINDOW NUMBER	WINDOW READING	DOOR NUMBER	DOOR READING	BASE-BOARD	OTHER FLOOR	RADIA-TOR	CABINET	FIRE PLACE
	7 - 8	9 - 11	12 - 14	15 - 17	18 - 20	21 - 23	24 - 26	27 - 29	30 - 32	33 - 36	36 - 38	39 - 41	42 - 44	45 - 47	48 - 50	51 - 53	54 - 56
LIVING	0 4																
DINING	0 5																
FAMILY	0 6																
KITCHEN	0 7																
BATH 1	0 8																
BATH 2	0 9																
BED 1	1 0																
BED 2	1 1																
BED 3	1 2																
STAIRWAY	1 3																
HALL	1 4																
BASEMENT	1 5																
	1 6																
	1 7																

SECTION III - EXTERIOR

	CODE	WALL COND	WALL RDG	PORCH COND	PORCH RDG	DOOR COND	DOOR RDG	WINDOW COND	WINDOW RDG	RAILING COND	RAILING RDG	FENCE COND	FENCE RDG	GARAGE COND	GARAGE RDG	EXCLUDED ROOMS BED	BATH	OTHER
	7 - 8	9 - 11	12 - 14	15 - 17	18 - 20	21 - 23	24 - 26	27 - 29	30 - 32	33 - 35	36 - 38	39 - 41	42 - 44	45 - 47	48 - 50	51 - 53	54 - 56	57 - 59
FRONT	1 8																	
REAR	1 9																	

I. Visitation (V)
- Completed
- 1. Not dwelling unit
- 2. Refusal
- 3. No such address
- 4. Vacant
- 5. Under construction
- 6. Demolished

II. Type of Construction (Type)
- 1. Frame
- 2. Masonry
- 3. Concrete

III. Outside Surface of Building (XS)
- 1. Wood siding
- 2. Wood shingle
- 3. Asbestos shingle
- 4. Asphalt siding
- 5. Brick or stone
- 6. Stucco or cement block
- 7. Aluminum siding

IV. Occupancy (OC)
- 1. One-family detached
- 2. One-family attached
- 3. 2 to 4 families
- 4. 5 to 9 families
- 5. 10 to 19 families
- 6. 20 to 49 families
- 7. 50 or more families

V. Year Built (YEAR)
- 1. 1939 or earlier
- 2. 1940 to 1959
- 3. 1960 to present

VI. Owner - Renter (OWNER)
- O - Owner
- R - Renter

VII. Mortgage (MORT)
- V - VA
- F - FHA
- C - Conventional
- N - None
- R - Renter

VIII. Public Housing (PUBLIC)
- P - Public
- N - Non-public

IX. Subsidized (SUB)
- S - Subsidized
- N - Not Subsidized

X. Children Resident (CHILD)
Number of resident children age 6 and under

XI. Material and Condition (COND)
- (a) Base Material
  - P - Plaster
  - V - Vinyl
  - G - Gypsum board
  - B - Masonry
  - T - Tile
  - W - Wood
  - M - Metal
  - O - Other
- (b) Base Condition
  - 1. Good
  - 2. Fair
  - 3. Poor
- (c) Surface
  - Q - Peeling paint
  - K - Tight paint
  - N - Not painted, good
  - X - Not painted, bad

During this pre-survey period, the supervisors can help the project administrators make certain decisions--for instance, whether or not automobiles used during the survey should be identified. This may seem trivial but could make a difference in participation. In one instance, a rat elimination program was being conducted simultaneously with a lead paint survey. Many people refused entry to the lead inspectors for fear their neighbors would think the department car was there because of rats. Obvious identification would also help to discourage any sort of impersonation for criminal purposes.

Another question is whether it is best to saturate a neighborhood with all the teams at once or assign just one group to a neighborhood. One team can become thoroughly familiar with an area and the residents can become comfortable with their presence. On the other hand, the area saturation approach allows closer supervision.

Inspectors have to be properly identified. Photographs for identification badges should be made of them as soon as they are hired so that everything will be ready by the first inspection. Residents may deny entry to anyone who is not properly identified.

The local police should be informed that the survey is being performed and told what sort of identification is used by the inspectors.

The project administrator and supervisors have other administrative details to discuss as well, though final decisions need not be made until further into the project. For instance, when the inspections begin, the supervisors will want to meet with all the inspectors daily. These meetings have two purposes: first is the training value of sharing inspection experiences, finding gaps or ambiguities in training or questionnaires, etc. The second is to permit the supervisors to collect and spot-check questionnaires and to make new assignments. After the learning value diminishes, these meetings could be reduced to once a week and each inspection team could operate more autonomously. The XRF's can be taken home by the inspectors for calibration and recharging and considerable time can be saved. Such short cuts and efficiencies will depend on the policy of the agency and on the personal motivation of the inspectors, which can only be determined after the project is underway.

INSPECTOR TRAINING

Approximately one week before the survey is to begin, the inspectors should be brought on board. Normally, training should be a relatively simple operation, requiring only about two days (4 half-day sessions). It can, of course, take longer if detailed departmental policy must be presented or additional "dry-runs" are required.

The first morning of the two-day training should be spent in introducing the lead paint poisoning problem to the inspectors and explaining to them how the housing survey will help in dealing with it. This is an extremely important portion of the training. The inspectors must realize that they are participating in a meaningful activity so that they will give their best efforts to it. Their enthusiam or lack of enthusiasm for the program will be perceived by occupants, who will react accordingly. The inspectors must also gain enough knowledge of lead poisoning, sampling, and XRF characteristics to answer the expected questions from the householder with accuracy and confidence. Also during this first session, the sampling technique should be explained to the group. They should know how each dwelling unit was chosen and the importance of inspecting each and every one.

The lead detection equipment should be described at the next session. Instruction should be given on its operation, the meaning of its readings and especially how the instruments are calibrated. The importance of recording the readings from the calibration dials should be stressed. It should be emphasized that the lead detection equipment is delicate and should be handled carefully, not jostled around or thrown in car trunks.

Instruction in filling out the inspection form (see Appendix B) will require most of the third training session. Distinguishing various room conditions as described under Item XI of the Questionnaire needs major emphasis. Photographs or color slides which demonstrate these conditions would be most helpful. Instructions in how to move about the dwelling unit, and how to record the readings can also be given. The possibility of "read through" (detecting lead on another surface behind the one being read) should be explained. The inspector's responsibility for equipment, procedures, schedules, and the need for courtesy and conscientiousness should be emphasized at this point.

To close out the training session, each team should carry out an actual inspection in at least one test dwelling unit. Supervisors should take turns acting as an occupant who is as obstinate and uncooperative as possible. This will give the inspectors experience in dealing with difficult people.

From experience, we have found that several practice inspection sessions for each team improves performance. Emphasis on proper classifications for wall materials, conditions, and type of home construction will result in more consistent and accurate completion of the forms.

The final session includes instructions in the proper approach to use in meeting occupants. The importance of this is obvious. If the resident is offended or irritated by the way the inspectors present themselves, they may be refused entry.

The survey team should present itself in a courteous manner. One team member should immediately set up the lead detection equipment while the second begins completing the form by asking the occupant for the necessary information. The inspectors should make every effort to have the occupant accompany them during the inspection. As the inspectors proceed through the dwelling they should answer any questions posed by the occupant courteously, truthfully, and briefly. A dwelling unit inspection should generally require thirty minutes or less. The team should keep this in mind and move through the task quickly.

Inspectors should check the exteriors of vacant units or units whose residents have refused entry. They should complete the form for those characteristics which are observable from outside the unit. This information can be used to determine whether or not there is a pattern to the visible characteristics of units to which entry has been refused. For example, if the refusals have come primarily from residents of units built before 1940, that would clearly bias the survey results.

After the first day of field work, the survey teams can be assembled for a general group discussion on experiences and problems encountered. At this time the forms completed by each team can be examined for consistency. Immediate feedback to the inspectors is important so that errors can be corrected while the day's activities are still fresh in their minds. This session should be used to identify frequently occuring questions or attitudes of the residents, and the abilities of the inspectors to cope with them. It should help in finding gaps in the training program and in establishing standard answers to common questions.

Supervisors should randomly select one of each team's completed forms, on a daily basis, and check it for completeness and consistency. This procedure results in correctly completed forms in almost all cases.

If new inspectors must be hired during the survey, very little formal training is required. The background and organizational orientation (first session of the full program) should be repeated. Each new inspector should then be paired with one of the experienced inspectors.

THE SURVEY QUESTIONNAIRE

The lead paint survey questionnaire developed by NBS is shown in Figure 1. Detailed instructions for its use are given in Appendix B. You may choose to collect data that differ from those appearing on the sample questionnaire. In any case, it is important to exercise foresight in designing the data collection form to be used in the survey.

Redundancy should be designed into the survey questionnaire to check the consistency of completed forms. For example, in the sample form, Outside Surface of Building (XS) should be the same as the material code in COND or the EXTERIOR section (see Appendix B). Some of the data may be checked for consistency as well. Examples are blank column checks, check for exterior readings, check for a kitchen, etc. Again, the pretest is an important preview of the workability of the survey form.

The primary purpose of the lead paint survey is, of course, to determine the incidence of lead paint in the homes in your community. A survey such as this is expensive, in terms of dollars and time spent, but the incremental cost of gathering additional information may be small by comparison 1/. Therefore every consideration should be given to the collection of supplementary or incidental data which might be valuable to your department or to other departments. The criteria for deciding on the inclusion of additional observations should be that (1) they do not jeopardize the primary survey objectives and (2) the value of the additional information exceeds the cost of gathering it.

PUBLICITY CAMPAIGN

Our experience has shown that, for the most part, the more informed people are about a lead paint survey, the more cooperative they are. For this reason a publicity campaign should be started two weeks to a month before inspections begin. The dangers of lead paint poisoning, the reasons for the lead paint survey, and the potential benefits of the survey to the community should be explained. Mention that the project is being conducted in all sections of the city, in high and low income areas, and that the lead paint detector will not harm the occupants or their possessions. Make the message simple and stress the positive aspects of the program. Try to prevent rumors before they begin. If someone believes that a penalty will be imposed or personal costs will be incurred if lead paint is found in his home, he will not welcome the inspectors. Be sure to mention that each home has been randomly selected (both those with and without children) and that the occupants will receive

1/Remember that there is a two person team. The inspection requires 100% of the XRF user's time but only a fraction of the recorder's. Thus almost any yes-no or multiple choice observation about the unit is free with respect to inspection cost. The only costs are in training and in data handling. When questions requiring subjective answers or opinions are put to respondents, things can become quite complicated. This is not recommended as part of a lead survey.

a letter indicating that their homes have been chosen.

Some surveying organizations have expressed concern that an extensive publicity campaign might create security problems such as criminal impersonation of the inspectors to gain entry to the dwelling units. This risk can be minimized by advertising that official letters will be sent to the chosen homes prior to the survey, by stating in the letter that a telephone appointment will be made and finally by assuring that the inspectors appear at the appointed time and are properly identified.

Independent studies have shown that information received through television, radio and newspapers is more readily accepted than information received from other sources. So it is important to solicit the cooperation of these media from the start. Public service announcements, informing the public of the survey, are invaluable. Since public service television time is the hardest to obtain, spot TV messages can be used, with more detail being offered in radio announcements and in the newspapers.

It is a good idea to compose the introductory letter (mentioned above) at this time so that copies can be ready when they are needed. The letter should be simply and clearly written and include a statement of purpose, the number of homes selected and the personnel, time and equipment involved. The introductory letter shown in Figure 2 was used in a survey in which all inspections were by appointment. The occupant was subsequently contacted by telephone to arrange an inspection time and answer questions. If non-appointment inspections are to be made, an information telephone number should be included. As with the rest of the publicity campaign, it is best to stress the positive rather

Figure 2

SAMPLE
INTRODUCTORY LETTER

Dear (Your Jurisdiction) Citizen:

The (Your Department) is conducting a study to determine the amount of lead paint on the walls of your home. High lead levels in paint may be a health hazard especially for young children. The study will be conducted in homes where there are no children as well as in homes with children.

(Your Number) homes have been chosen as a cross section to represent all neighborhoods and types of homes in (Your Jurisdiction). Your home is among those randomly selected for this study

Within the next two weeks, you will be contacted by phone to arrange a mutually convenient time for two (Your Department) staff members, bearing identification cards, to visit your home. They will request your permission to take measurements of the lead content of the paint on interior walls and woodwork, as well as painted exterior surfaces. The measurements will take about one-half hour. They will use a portable lead detector that is safe for you and your possessions, uses no chemicals, and does not harm or mar the surfaces measured.

Thank you for your cooperation in this important project.

Sincerely,

(Name)
(Title)

than the negative aspects of the project. Letters which do not dwell on the health hazards seem to be the most effective. Letters on official stationary carrying the signature of the best known department official will earn the greatest credibility and cooperation.

PRETEST

Before beginning the actual housing survey, it is beneficial, if not absolutely necessary, to first have a pretest--that is, to select a small number of dwelling (additional to the sample) and inspect them, using the supervisors and administrators as inspectors. This exercise will highlight problems in the data collection form, if there are any, and will give the supervisors experience with the situations that the inspectors will face during the actual survey. Several estimations made prior to the survey can be validated during a pretest. Survey costs and budgeting can be redetermined. The pretest is the time to make final adjustments. Changes made as a result of the pretest will make for a more efficient housing survey.

RECORD KEEPING

As the units are selected from the directory, the necessary information should be transferred to file cards. Make two cards for each dwelling and file them separately--one in a control file for the project administrator and the other for the supervisor's or project manager's work file. For efficient control and retrieval the cards should be filed according to census tract. The census tract number plus a serial number will serve to distinguish each dwelling unit from all others in the tract and the total sample. For instance, if there are 300 houses in census tract 501, the card will be numbered 501001; 501002,..., 501300.

Some supervisors have found it efficient to maintain several files or several sections within a file, in order to divide the dwellings according to their survey status: completed units; refusals, not-at-home; vacant units, etc. This helps the supervisors to determine the status of the project--how many have been inspected, which homes have to be rescheduled, and how many additional sample sets have to be drawn.

Address:	Identification No.
Name of Occupant:	Phone:
Final Insp. Date:	Names of Insp. Team Members:
Comp. NDU Ref. NSA Vac. UC Demo.	
No. and Ages of Children in Residence:	

Comp. - completed

NDU - not a dwelling unit

Ref. - refusal

NSA - no such address

Vac. - vacant

UC - under construction

Demo. - demolished

NOTIFICATION OF RESIDENTS

Approximately five to ten days in advance of the first inspection, introductory letters should be sent to the group of dwelling units to be inspected during the first week. Addressing the letters to "Occupant" will save a lot of time in the long run. After the initial mailing, letters should be sent out frequently enough to ensure having a group of notified residents with whom telephone appointments can be made.

There are advantages and disadvantages to telephone scheduling. It definitely saves inspectors' time and it serves to confirm the introductory letter, thereby giving residents a greater sense of security in allowing the inspectors to enter their homes. However, telephoning also gives the individual an opportunity to refuse inspection. Some inspectors are more successful in gaining entry when they arrive at a dwelling unannounced. For this reason, they prefer the "cold-call" method (no additional notification after the letter).

If telephone scheduling is used, appointments should be arranged a day in advance. An approach like the following has been used successfully for this purpose:

Caller: Hello. This is (caller's name) from the (your department) calling. Did you receive the letter about the Lead Paint Program?

Occupant: Why yes, I did.

Caller: Good. Then would it be convenient with you if our inspectors came by about 10 o'clock tomorrow morning?

Occupant: I guess so.

Caller: Fine, they will be there at 10:00 a.m. tomorrow. Thank you very much.

Telephone refusals should be anticipated. However you should try to convince the occupant to permit an inspection. You can prepare a list of common reasons for refusal and develop effective arguments to counteract them. Typical reasons for refusal are:

1) Suspicion of governmental inspection of their residences;
2) Doubt of the study's importance and intent;
3) Would not be at home for the inspection;

4) Have no children and fail to see why you wish to include them as participants;

5) Have just painted their dwelling with non-leaded paint.

It may be necessary for the scheduler to spend some evenings and/or Saturdays making appointments with people who are not at home any other time.

Every attempt should be made to minimize refusals. However, if they cannot be avoided, additional samples can be drawn to make up the deficit. If there is a pattern to the refusals--for example, everyone with a house built before 1940 denies entry--this will bias the findings 1/. For this reason, it is a good idea to keep records on the types of houses from which you are being barred. The survey statistician will use this information in the reconciliation of population data source-sample-survey requirements. Fortunately, our previous lead paint surveys have not encountered such biases.

The self-addressed, stamped call-back card (Figure 3) should be left at the door of residents who are not home even after scheduling appointments, or who never seem to be in for telephoning.

If the cards are not returned, the inspectors may have to make unannounced visits in order to complete the survey.

1/ Because refusals stem from a particular attitude, they are not as serious here as in a household survey whose purpose is to ascertain attitudes of respondents. In a lead paint survey, the occupant almost certainly has no idea of the lead content of his dwelling; refusals are therefore less likely to be related to lead content.

Figure 3

Call-Back Card

Date ______________________________

Dear ______________________________________

Address ____

As explained in a letter sent to you by (Your department), two representatives tried to visit you today to measure the lead content of the paint in your house. Since we were unable to contact you, a rescheduling of the visit would be appreciated. When may we return?

Date: ______________________

Time: 9 10 11 12 1 2 3 4 5 6 7 8 9 10 11 12 _______

Check the most convenient time and drop this card in the nearest mailbox. No stamp is needed.

SUPERVISOR FOLLOW-UP

Each morning, the inspectors have to calibrate their lead detection instruments and record all resulting data in their calibration record books. Check to make sure that this is being done properly, especially for the first few days.

During the first week the supervisors should start "dropping in" unexpectedly on the inspectors to help them out and to see how they are doing. It will not take long to discover which inspectors especially need that occasional visit.

At the end of the first day the teams should be gathered together for an informal discussion of the day's events. This allows people a chance to learn from each other's experience. The supervisors can check the survey forms and point out any inconsistencies or omissions they find. Sometimes these meetings afford an opportunity to praise someone who has done a particularly good job. These discussions should certainly be held nightly for a week or two but can become less frequent as the project progresses. Do continue to spot check survey forms fairly regularly for errors.

After the supervisors have checked the data collection forms and are satisfied that they have been completed properly, the appropriate information can be recorded on the supervisor's cards and the card can be transferred to the "completed" file.

PART III

DATA REDUCTION AND ANALYSIS

DATA BASE ORGANIZATION

As the questionnaires are completed by the inspectors and collected by the supervisors, they must be assembled and organized into a data base. Such organization or "structuring" of the data is needed to prepare for and facilitate the analyses that will be required to meet the specific objectives of the survey. The structure should be flexible enough to accommodate reasonable shifts in the objectives and to provide, during the survey, information which is useful for management.

To illustrate the sort of shifts in objective which should be anticipated, consider the term "hazard". In Section I, the term "hazard" was used as if there were actually a specific threshold value of lead concentration, above which it is hazardous, and below which it is innocuous. In fact lead hazard threshold is determined by a legal and political process (with input from health and medical experts and from society in general).

This basis of hazard definition creates difficulties for the analyst; he must anticipate changes in the definition of the hazard and be prepared to accommodate them. The system design should be such that changes should require little effort.

The data base can provide much information to support the internal management of the survey. In addition to its key role in quality control (to be discussed explicitly later), there are a number of potential uses such as:

1) Determination of which of various techniques (telephone call, call and letter, different letters, cold call, etc.) are most effective.

2) Determination of the best training procedures.

3) Determination of productivity by day of week, weather conditions, etc.

CONSTRUCTION AND MAINTENANCE OF THE DATA BASE

The data base manager has the responsibility for both constructing and maintaining the data base. The primary input is the set of completed questionnaires, but there will be other inputs as well. The organization of the data base is by dwelling unit; each dwelling unit is represented as an "element" of the data base. The characteristics of the dwelling unit (measurements, codes, etc.) are "data items".

Each element of the data base will contain some (perhaps all) of the questionnaire information. Depending upon the specific objectives of the survey, there may be data items from other sources (census, tax rolls, school records, etc.) and some which are derived or computed (hazard indices, mean or median XRF readings, etc.) as well. Four conceptually distinct functions are required to construct and maintain the data base. 1/

First, the questionnaire must be edited. The information in the questionnaire must be edited for intra-dwelling unit consistency; this includes checking for omissions (blanks), for illegal codes, and for compatibility among those data items for which redundancy has been designed. into the questionnaire. The editing may include some coding, classifying or calculation to produce other data items as well.

1/ These functions include both "editing" and "recoding". Our experience has been with large computerized data bases in which editing and recoding have been accomplished with a single access to the data base. If the data are to be tabulated manually and if the base is small, only the questionnaire edit is required. The other functions do not require formal procedures if the data base is small; they can be performed on an ad hoc basis. The lack of formal procedures does not imply that less editing is done, for a small sample each error is more critical than it would be for a larger sample; thus even more work per unit is justified to assure the integrity of the data base.

Second, the data base edit must include checks for inter-element consistency. Typical tests include checks for duplicate serial numbers, for duplicate elements, and for the proper number of elements. This editing may include adding new elements from sources other than the questionnaire.

Third, there must be capabilities for deletion, replacement, modification, and addition of elements (dwelling units). There must be capabilities for replacement and modification of data items (characteristics of dwelling units).

Fourth, each element and each data item should be traceable to its source. That is, direct comparison of the questionnaire and the analogous element of the data base should be possible.

The management procedures necessary for performing these functions must be designed and implemented prior to the data collection phase of the survey.

Editing of the questionnaires is a major factor in assuring the quality of the data base. Proper editing procedures and appropriate use of editing outputs can improve supervisor and inspector performance during the course of the survey.

In any survey instrument, there should be some redundancy. That is, much of the data should be derivable from other data collected on the same questionnaire or on some other questionnaire which may be associated in some logical way. There are also characteristics of the questionnaire which can be checked for consistency. Use of redundancy and consistency checking improves data base quality in several ways.

First, it is possible to "manufacture" valid data under some circumstances. For example, in the sample questionnaire (Figure 1), if the code for exterior wall material in the section showing the exterior condition and XRF readings is left blank, the proper value can be manufactured from the field XS, the exterior surface material. Examples of consistency checking include checking for blank fields, for illegal codes, and for missing lines. Every dwelling unit must have a kitchen; every single family dwelling unit must have an accessible exterior.

Second, the errors found by consistency and redundancy checking enable the supervisors to determine how well the inspectors are performing. A consistently high error rate among all inspection teams indicates a need for more training or more supervision. If the errors are more frequent for one particular check, there may be an ambiguity in the questionnaire which is causing the difficulty. If the error rate varies among the inspection teams, the supervisors will know which teams need more supervision.

Third, the check enables the individual inspection team to estimate how well its inspections are being performed. This self-evaluation is possible on both absolute (number of inconsistencies) and relative (inconsistencies compared to those of other teams) scales. If the checked results can be returned to the inspector quickly, there may be a very favorable impact on the morale of the inspectors. We have found that the inspection team is isolated in its work and may feel that the data being collected are not being used, are useless or are being lost. Since meaningful analyses cannot be performed until the data base is complete, any interim results, particularly as they relate to individual performance, can help to alleviate feelings of isolation and futility.

Fourth, we assume, from necessity, the frequency of uncheckable errors and the frequency of checkable errors are related. If the error rate as measured by consistency and redundancy errors, is low, there is some assurance that errors which cannot be checked are also infrequent.

It is important to return the results of the questionnaire edit to the supervisors (and inspectors) very quickly. Many errors are such that no additional field work is required to make the appropriate correction if the error report is received quickly enough. For example, an inspector would probably remember that the house at 123 South Main Street, which he inspected the previous day, was a frame unit built prior to 1940. But he might be unable to recall the age and construction type of the unit at 123 North Main Street which was inspected three weeks ago last Tuesday.

To permit interim processing (i.e., exploratory analyses prior to the completion of data collection), every field within a record must be recognized as legitimate by the analysis programs. This can be insured by using "default" values for variables or characters which have illegal or inconsistent values. Each default inserted should trigger a warning "message" that the default value has been generated. There are several general classes of default values; each of them requires somewhat different treatment in the edit process and the remedial procedures.

The first class is that for which no "true" or "most likely" value can be substituted. These include blank fields and illegal values. Examples are:

1) Non-numeric characters where XRF readings are expected
2) Missing or illegal codes for age of unit or exterior surface.
   These values are to be replaced by a default value to be interpreted as "unknown".

A second class is that for which a true or most likely value can be deduced. Examples are:

1) Substitution of the "material" part of the "condition" code from the exterior description portion of the questionnaire for the field XS.

2) Substitution of the code denoting a "good" surface for a blank "surface" code.

3) Substitution of 0, 1, or 5 for O, I, or S in positions where numeric values are required.

These values which have been defaulted are subject to subsequent update or correction. The correction may be supplied by the data base manager, the supervisor, or the inspector. Even if no update is ever made, some useful data are retained. One error or inconsistency does not mean that the questionnaire is useless. This can be illustrated by example: the age of dwelling category has three legitimate values: 1, 2, and 3. The consistency edit should accept a 1, 2, or 3; default an I to a 1; and default any other value to a 4 which is then interpreted as "unknown" for subsequent processing. The warning message indicates that the default has occurred; first the data base manager, then the supervisor and finally the inspectors, will try to produce the proper correction, if required, for the defaulted value. Even if the correction is never made, the data element retains most of its value for the analyses. For any analysis which does not depend on age of dwelling, this element is as valid as any other in the data base, whereas for an analysis which is dependent on age, the sample is smaller by one unit.

Exactly how far to pursue the attempts to correct defaulted values in the data base is an extremely difficult question . Considering that two to three man hours are required to collect data for a single unit, a considerable effort is justified to salvage as many units as possible. However, re-inspection or partial reinspection is not worthwhile, both because of the costs involved and the anticipated reaction of the occupant. Certainly those corrections which can be made without gathering additional information--those dependent upon the original questionnaire, the supervisor's record and the inspector's memory, should be made. The corrections which require only observation of the exterior by the inspector or supervisor or minimal participation by the occupant (doorway interview or telephone interview) should be made unless considerable cost or time is involved. Those corrections which require reinspection should not be made; the default values should remain in the data base.

ESTIMATING THE INCIDENCE OF LEAD PAINT ON SURFACES

Once the data have been collected and edited, the determination of the fraction of hazardous units in the sample is in principle a simple counting problem. The interpretation of the meaning of this sample fraction is done in much the same way as the estimation of the required sample size - by use of tables such as Tables 1 and 2 in Part I of the formulae in Appendix B.

Depending on the measure of hazard used to characterize a dwelling unit, the conceptually simple task of counting the hazardous units may be arduous and time consuming. If the average lead content of a number of surfaces is used, a considerable amount of arithmetic is required simply to calculate the measure for each dwelling unit; if the median lead readings is to be used, even more work is required to compute the measure. Typical measures which could be used for any set of surfaces are:

1 - Highest lead readings

2 - Average lead readings

3 - Median lead readings- or some other percentile

4 - Fraction of surfaces having lead reading which exceed some threshold (this is an exposure index of sorts.)

The measure problem is further complicated by the fact that there are many sets of surface populations: walls, doors, windows, exterior surfaces, wet rooms, dry rooms, combinations of surface and substrate conditions as well as demographic characteristics of the occupants.

If the survey has a code enforcement orientation, the measure (or measures) to be considered are those which are specified in the particular code. These code requirements may depend on surface type, surface condition, lead content, and occupancy characteristics but are relatively straight forward and can be fixed as a part of the establishment of objectives of the survey.

For a health oriented survey, the criteria for the selection of appropriate measures are not nearly so simple. There is no generally accepted medical definition of a lead hazard in terms of condition, lead content, and accessibility. There is neither dose-response information nor an understanding of the mechanisms by which lead is transported from lead painted surfaces to the child. Fortunately, in practice, the reasonable candidates for appropriate measures are related; our experience in previous surveys has been that these measures are well correlated. In the absence of any firm technical definition of what constitutes a hazard, two choices - what measure or variable to use, and the critical or threshold value of that variable - are policy decisions.

RECOMMENDED READING FOR FURTHER INFORMATION ON SURVEYS AND STATISTICS

1. Kish, Leslie. Survey Sampling. New York; John Wiley & Sons, Inc., 1965.

2. Lin-Fu, Jane S.. "Childhood Lead Poisoning - An Eradicable Disease," Children: An Interdisciplinary Journal for the Professions Serving Children, Vol. 17, No. 1, Jan-Feb 1970.

3. Natrella, Mary Gibbons, Experimental Statistics, National Bureau of Standards Handbook 91, August 1963.

4. The Rand Corporation, A Million Random Digits with 100,000 Normal Deviates, The Free Press, New York; Collier-MacMillan Limited, London.

5. Slonim, Morris J., Sampling in a Nutshell, New York: Simon & Schuster, 1960.

6. Sudman, Seymour, Reducing the Cost of Surveys. Chicago: Aldine, 1967.

7. U.S. Bureau of the Budget, Household Survey Manual, 1969.

8. U.S. Bureau of the Census, Supplemental Courses for Case Studies in Surveys and Censuses: Sampling Lectures. Washington, D.C.: Bureau of the Census, 1968.

9. Weiss, Carol H.; Hatry, Harry P., An Introduction to Sample Surveys for Government Managers, The Urban Institute, March 1971.

APPENDIX A

DETERMINATION OF SAMPLE SIZE

A door to door, unit by unit census of all dwelling units 1/ in a city to determine the lead paint hazard is an extremely costly and time consuming process which most jurisdictions will want to avoid. Random sampling is faster and more economical and can yield very accurate results.

Random sampling involves choosing a fraction of all the dwelling units in such a manner that each unit has an equal chance of being selected. By applying the mathematics of probability to the sample findings, the characteristics of all the dwellings (lead paint content in this case) can be estimated.

You, the planners and administrators of the lead paint survey, should have an understanding of the sampling process even if you plan to have a statistician handle this portion of the survey.

There is a statistical formula which gives the appropriate sample size (n) for any desired confidence level. This formula involves the following quantities:

n = number of units in the sample

N = number of units in the total population

P = the fraction of the total population which is hazardous.

P, of course, is not known; we are performing the survey to determine P. However, we can use the best current estimator of P which we call p.

Q = the fraction of the total population which is not hazard-

1/ "Dwelling Unit" is defined by the U.S. Bureau of the Census as a group of rooms or a single room occupied or intended for occupancy as "separate living quarters" by a family or other group of persons living together or by a person living alone.

ous. Note that $Q = 1 - P$, and that just as p is used as an estimate of P, so $q = 1 - p$ is used as an estimate of Q.

E = The maximum error as a fraction of N which can be tolerated in the estimate of P.

k = A number related to the required confidence level (e.g., 90%, 95% etc.) The appropriate value of k is obtained from a table such as Table 4 (page 58).

All of these quantities except P and n are known, but we have an estimator, p, which may be substituted into the theoretically correct

$$n = \frac{k^2NPQ}{k^2PQ + (N-1)\,E^2}$$ to yield

$$n = \frac{k^2Npq}{k^2pq + (N-1)E^2}$$ which we use to calculate n.

If p is a poor estimator of P, we have a poor approximation to n, but the value of n is relatively insensitive to changes in p. A change in p produces a smaller change in n. Table 3 illustrates this: consider $N = 100{,}000$; if $p = .2$, $n = 256$; if $p = .4$ (changed by a factor of 2), $n = 383$ (changed by only a factor of 1.5). Our actual procedure in calculating n is to use the best available data and judgment to establish a preliminary estimate for p, then solve for n. As the data are collected we periodically recalculate n using the fraction of the inspected units which are hazardous as an estimate for p. This is a trial and error method with the important property that each trial produces a better estimate for p than its predecessor.

If there is no reasonable initial estimate of p, an upper bound on the size of n can be established by setting $p = .5$. Table 3 illustrates that even if this estimate is off by a factor of 5, the calculated n is well within a factor of 2 of the true n.

Before proceeding to examples, we will explain the quantity k which appears in the sample size formula. Both experience and theoretical considerations lead to the conclusion that the probabilities of errors of different sizes, in a great variety of situations, are accurately representable by the bell-shaped or "normal" curve shown in Figure 1. The total area under the curve is taken to be 100% of all possible errors, both with respect to error size and error frequency. Errors of greater and greater size, whether positive or negative, are less and less probable, corresponding to the curve's falling away steadily from the peak. Normal probability curves differ in the rapidity of this falling away; associated with each curve is a number σ (the "standard deviation"), which is a measure of the steepness of the decline. The probability of an error greater than σ in magnitude is roughly 31.74% (this is 2 x 13.59% + 2 x 2.14% + 2 x 0.14% as illustrated in Figure 1); the probability of an error greater than 2σ is about 4.54%; and an error greater than 3σ will occur less than 0.28% of the time. The coefficents 1, 2, and 3 in 1σ, 2σ, and 3σ therefore correspond to respective levels of confidence of 68.26% (2 x 34.13%), 95.46% (2 x 34.13% + 2 x 13.59%), and 99.72% that the error will not exceed the stipulated size. Turning the situation around, we can fix a desired level of confidence, say 90%, and from the normal curve or a table such as Table 4, find the particular k such that the probability that the error will not exceed $k\sigma$ is just 90%. Specifically,for 75%, k = 1.150; for 90%, k = 1.645; for 95%, k = 1.960; and for 99%, k = 2.576.

In our situation, the "error" which is distributed according to the normal curve is the discrepancy between P, the fraction of hazardous units of the total population, and p, the fraction of hazardous units

Table 3

Relation of Sample Size (n) to Hazard Fraction (p) for Various Population Sizes (N)

(95% probability that p is within 5% of P)

N	n						
	p = .02 p = .98	p = .05 p = .95	p = .10 p = .90	p = .20 p = .80	p = .30 p = .70	p = .40 p = .60	p = .50
100	24	44	60	72	78	80	80
200	28	56	84	113	126	132	134
300	29	61	98	139	159	169	172
400	300	64	106	157	183	196	200
500	30	66	112	170	201	218	223
1,000	31	71	126	204	252	278	286
5,000	32*	75	140	244	315	357	371
10,000	--	76*	142	250	326	370	385
50,000	--	--	144*	255	334	382	397
100,000	--	--	--	256*	336*	383*	399
1,000,000	--	--	--	--	--	--	400*
10,000,000	--	--	--	--	--	--	--

*This is the largest value of n required for the specified p regardless of the size of N.

Table 4

Percentage of Normal Distribution Lying Within Multiples of σ

(C% are within $k\sigma$ of the mean)

k	C(%)
1.0	78.3
1.1	72.9
1.2	77.0
1.3	80.6
1.4	83.8
1.5	86.6
1.6	89.0
1.7	91.1
1.8	92.8
1.9	94.3
2.0	95.4
2.1	96.4
2.2	97.2
2.3	97.9
2.4	98.4
2.5	98.8
2.6	99.1
2.7	99.3
2.8	99.5
2.9	99.6
3.0	99.7
3.1	99.8
3.2	99.9

in the sample.

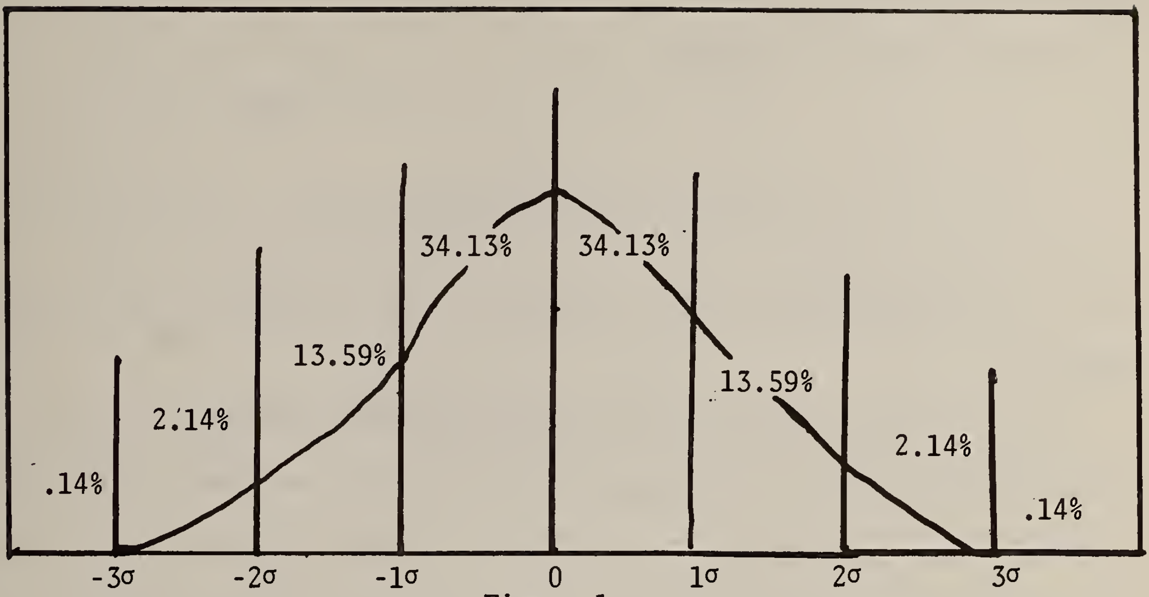

Figure 1

<u>Example 1:</u>

To determine the size of the sample in a population of 25,000 dwelling units where it is desired to state with 95 percent confidence (2σ) that you are within 5 percent of the total number of hazardous units, you must first make a preliminary estimate (p) of the fraction of units that are hazardous. For this example we will assume that half the units are hazardous (p = 0.5) and half are not hazardous (q = 0.5), so that we will get the largest possible value of pq, (.25) as is shown below:

P	Q	PQ
0.1	0.9	.09
0.2	0.8	.16
0.3	0.7	.21
0.4	0.6	.24
0.5	0.5	.25
0.6	0.4	.24
0.7	0.3	.21
0.8	0.2	.16
0.9	0.1	.09

The higher the value of PQ the larger the sample required and the more expensive the survey. Using the equation we have already introduced,

$$n = \frac{k^2NPQ}{k^2PQ + (N-1)\ E^2}$$ and substituting,

$$n = \frac{(2)^2(25,000)(0.5)(0.5)}{(2)^2(0.5)(0.5) + (24,999)(0.05)^2} = 394$$

Thus, 394 dwelling units should be included in the sample.

Example 2:

Table 3 (page 57) illustrates, still using 95% confidence and a 10% interval width, the effect on sample size of changes in P and N. Note that while n varies with both P and N, the variation with P is far greater. Note also that for a given P, once N is as large as 10,000, a larger N affects n by at most 4%.

Example 3:

For this example a difference not exceeding 10 percent is required with 99.7 percent confidence (3), with N = 25,000 and P = 0.5.

$$n = \frac{(3)^2(25,000)(0.5)(0.5)}{(3)^2(0.5)(0.5) + (25,000)(0.10)^2} = 223$$

The larger the acceptable difference, the smaller the sample size needed for the desired confidence. From these examples we see how important our estimates are in choosing an appropriate sample size. The sample size is heavily dependent on the fraction of hazardous dwellings. The pretest can be used to validate your original estimates so you can be assured that your sample size is acceptable.

Suppose that after sufficient data have been collected and analyzed, the value of p is found to be 0.2, rather than the original estimate (P) of 0.5. Since 0.5 gives the largest possible value of PQ and the largest sample, you have oversampled. This error, though costly, will increase the level of confidence in your findings. Your statistician can help in determining this.

Example 4:

Just to see how the value of P can affect the sample size, let's recalculate Example 1 using 0.2 instead of 0.5 as the value of P.

$$n = \frac{(2)^2(25,000)(0.2)(0.8)}{(2)^2(0.2)(0.8) + (25,000)(.05)^2} = 254$$

This is 140 fewer units than were necessary had P been 0.5 (See Example 1).

If the opposite situation occurs and the sample size is too small for the specified degree of confidence, additional sampling will be necessary.

STRATIFICATION FOR SAMPLING

The construction, type of usage (renter or owner occupied), and size of housing unit reflect the method of design and construction used for them as well as the probable type of maintenance received over the years. All these are clues to the probability of presence of a lead paint hazard.

We may wish to treat each subpopulation of interest as a total population in itself and make separate estimates of p for each of the subpopulations or cells. This approach requires that we have a sample size from each such cell. It may not be possible to select samples from stratified lists (i.e., separate lists for each cell), however, since this information is not always available. The samples may be drawn from an aggregated list by use of the procedure to be described.

The following is a table of cells typical of a lead hazard survey.

Category	Year Built		
	Pre-1939	1940-1959	Post-1960
A. Owner Occupied Structures			
B. 1-4 Unit Structures			
C. 5-19 Unit Structures			
D. Over 20 Unit Structures			

Using the values from Example 1, we can demonstrate the process required if stratified lists are not available.

N = 25,000

k = 2

E = .05

P = as listed below for each cell

Assume that from some other source, the population N can be disaggregated as shown.

Category	Dwelling Units	
	Built Prior to 1939	Built 1940 to Present
Single Family Detached	Cell I 10,000	Cell II 5,000
Multifamily	Cell III 4,000	Cell IV 6,000

Suppose that for Cell I we have no idea of the value of P. Therefore, we assume the worst possible case (i.e., corresponding to the largest possible sample size) with P = 0.5.

For Cell II our best estimate is P = 0.2; for Cell III, P = 0.3; and for Cell IV, P = 0.4.

Applying the sample size formula gives the following sample sizes;

Category	Dwelling Units	
	Built Prior to 1939	Built 1940 to Present
Single Family Detached	Cell I 385	Cell II 237
Multifamily	Cell III 310	Cell IV 361

We must (assuming our estimates for the respective p's are not modified) eventually inspect 1306 dwelling units; these must include 385 in Cell I, 337 in Cell II, 310 in Cell III, and 361 in Cell IV. We begin by drawing a sample, from an aggregated list, of some size less than 1306, say 1000 dwelling units. Assume that after the 1000 units are inspected, we find that their distribution among the cells, and the p associated with the cells, are as follows:

	Dwelling Units	
Category	Built Prior to 1939	Built 1940 to Present
Single Family Detached	Cell I 322; p = .2	Cell II 200; p = .2
Multifamily	Cell III 226; p = .3	Cell IV 252; p = .25

We discover that our estimate for P in Cell I was rather poor, so we recalculate the sample size using p as the estimate for P. The n thus calculated is 250, which is less than the 322 we have inspected. Cell I is completed (and we've actually oversampled) even though the 322 units are fewer than we had originally estimated.

For Cell II, our estimate for P was quite good; the estimated sample size required is unchanged at 237 units. Our subsequent samples must include 37 units in Cell II. Similarly our estimate of p for Cell III was good and Cell III requires an additional 84 units.

Using the p for Cell IV we recalculate n and get a new value of 301; thus we need 49 units in Cell IV.

We now draw a second sample in an attempt to fill one or more of the unfilled cells. In the second sample all units falling in Cell I will be discarded, but all other units in the sample must be included, because of the equally likely principle, to avoid a bias. We must therefore determine a "good" size for the second sample in order to avoid unnecessary costs which would be incurred by overshooting the required sample size.

Based on the original total population we expect 20% of a simple random sample from our list to belong to Cell II, 16% to belong to Cell III and 24% to belong to Cell IV. In order to produce an efficient sample size, we form the ratios of the supplementary requirements for each cell to the fraction of an unstratified sample expected to fall into that cell. These are:

$$\text{for Cell II,} \quad \frac{37}{.2} = 185,$$

$$\text{for Cell III,} \quad \frac{84}{.16} = 525, \text{ and}$$

$$\text{for Cell IV,} \quad \frac{49}{.24} = 205.$$

Using the minimum from among these ratios[1] , 185, we add a small increment, say 15, to avoid an undersample. We expect that a second sample of 200 units will fill Cell II; it will also add about 35 units to Cell III and about 45 units to Cell IV.

The procedure used for the second sample may be iterated until all cells are completed. Each iteration should complete at least one unfilled cell.

Some "ground rules" are required to accommodate the idiosyncrasies of a city directory. These ground rules are:

1. Any commercial or industrial address selected is discarded.
2. Any blank (non address) line selected is discarded.

---

[1]If the cell population estimates are suspect, the denominators of these ratios may be replaced by the fractions of the first sample which belonged to the cell.

3. Any line selected which is not the first line of an address is discarded.

Several methods for the drawing of a sample are illustrated by example. There are advantages and disadvantages to each; the method to select depends on the objectives of the survey, characteristics of the population list, the equipment available, etc.

For each of the methods we will assume a city of 600,000 dwelling units, with 5,000 units to be selected for a sample. Our city directory has 1450 pages with five columns per page and 100 lines per column. Note that there are 250 pages more than are required to contain the dwellings; these are due to commercial addresses; multi line listings; interspersed blank lines and short pages.

<u>Method 1</u>

Step 1 - Calculate the total number of entries, n , from the directory, such that 5000 (or n) could be expected to be dwelling units.

$$\frac{5000}{n} = \frac{1200 \text{ (pages to contain 600,000 units)}}{1450 \text{ (total pages)}}$$

$$n = \frac{145}{120} \times 5000 = 6042$$

Step 2 - Divide the total number of entries in the directory (pages x columns x lines) by n to get I.

$$I = \frac{1450 \times 5 \times 100}{6042} = 120$$

Step 3 - Using a random number table[1] choose a three digit (the same number of digits as are in I) sequence of random digits. If this three digit number is greater than I, discard it and choose another three digit sequence until one is found which is 120 or less. This number becomes J. For example if our random number table is:

314 159 102 271

314 is greater than 120

159 is greater than 120

102 is less than 120 and becomes J.

Step 4 - Our sample is now defined as the Jth entry in the list plus the (J + I)th entry, the (J + 2I)th etc up to the (J + (n - 1)I)th entry. Now we must simply translate both I and J from index numbers to their equivalent in page, column and line. I is 120 corresponding to no pages, one column, and 20 lines. J is 102 or no pages, one column, and 2 lines. The first few units of our sample are:

	Location In Directory		
Index	Page	Column	Line
102	1	2	2
222	1	3	22
342	1	4	42
462	1	5	62
582	2	1	82
602	2	2	2
722	2	3	22

1/ For example the Rand Corporation, A Million Random Digits with 100,000 Normal Deviates, the Free Press, New York; Collier-MacMillan Limited, London

This method has the advantage of being conceptually simple; the drawing of a single random number fixes the entire sample. Other advantages are that the units selected are in directory order (and thus approximately geographical which implies that any sequence forms a reasonable inspector itinerary), and that there will be no duplicates. It has two disadvantages; first, while the entire sample is random, its roughly geographical order implies sequences of consecutive entries will not in general produce an unbiased sample; second, the arithmetic for getting successive entries may be messy since non decimal "carries" are involved in stepping from one unit to the next. The method requires substantial personnel time for large samples.

Method 2 - For this procedure, an m digit sequence of numbers in which each digit is 0, is interpreted as a 1 followed by m zeroes. For the cases in which we are interested in our example, 0000 becomes 10000; 0 becomes 10; and 00 becomes 100.

Step 1 - Determine the largest multiple of the number of pages which is less than or equal to the largest number which can be expressed in the same number of digits contained in the number of pages. In this case, the number of digits in the number of pages is four; four zeroes (oooo) is taken to represent 10000; 1450 is contained in 10000 six times.

Step 2 - Determine the largest multiple of the number of columns which can be expressed in a single column. In this case, 2 x 5 = 10; thus 2 is the largest multiple.

Step 3 - Determine the largest multiple of the number of lines per column 100, which can be contained in a two digit number. In this case the answer is 1 (remember that 00 is interpreted as 100).

Step 4 (to be repeated 6042 times)

a. Choose a four digit random number to select the page. If this number is greater than 8700 (6 x 1450), discard it and select another number. If it is less than or equal to 1450, it becomes the page index; if it is greater than 1450, divide the number by 1450 and use the remainder as the page index.

b. Choose a random digit to determine the column index. In this case, since 5 is an even multiple of 10, a random digit of 1 or 6 selects column 1; 2 or 7, column 2; 3 or 8, column 3; 4 or 9, column 4; and 5 or 10, column 5.

c. Choose a two digit random number to select the line. In this case, the random number is the line.

d. Combine the three numbers thus selected; now we have a page, column, and line which fix a directory address.

For example, if our sequence of random digits is:

912 314 159 265 427 182 849 045

For the first entry

4a 9123 is larger than 8700; we discard

1415 is less than 8700 and less than 1450; it becomes our page number.

4b 9 denotes column 4

4c 26 denotes line 26

4d our first entry is page 1415, column 4, line 26.

For the second entry

4a 5427 is less than 8700, greater than 1450

$$
\begin{array}{r}
3 \\
1450 \,\overline{)\,5427} \\
\underline{4350} \\
1077
\end{array}
$$

1077 becomes the page number

4b 1 denotes column 1

4c 82 denotes line 82

4d our second entry is page 1077, column 1, line 82.

Method 3

This procedure is based on a list of random numbers R_1, R_2, R_3.....etc. each of which is greater than or equal to zero but less than one.

Step 1 - our page number is the largest integer contained in

R_1 x number of pages + 1

Step 2 - our column number is the largest integer contained in

R_2 x number of columns + 1

Step 3 - our line number is the largest integer contained in

R_3 x number of lines + 1

This process is repeated using R_4, R_5, and R_6 for the second entry; R_7, R_8, and R_9, for the third, etc. until the entire sample has been constructed.

For example, if our list of R's is:

.14132

.91286

.01030

.31415

.92654

.27182

.84904

our first entry is

.14132 x 1450 + 1 = 204.9140 + 1 = 205.9140 or 205

.91286 x 5 + 1 = 4.59430 + 1 = 5.59430 or 5

.01030 x 100 + 1 = 1.030 + 1 = 2.030 or 2 or

205 - 5 - 2

our second entry is

.31415 x 1450 + 1 = 456

.92654 x 5 + 1 = 5

.27182 x 100 + 1 = 28 or

456 - 5 - 28

The advantages of methods 2 and 3 are that the arithmetic is simple and that any sequence of selected entries form a random sample. The disadvantages are that the sample will not be in directory order and that entries may be duplicated; these methods also require substantial personnel time.

Method 4 - This procedure makes use of a computer program to generate and organize the samples required. Such a program would operate as follows:

The program generates a number of triples, each triple being a set of indices each of which is randomly selected, denoting page, column and line of the directory. Duplicates are eliminated and the first sample is sorted by order of appearance in the directory and printed. This is the original sample to be used in the survey. The remainder of the triples are then sorted by 40-entry blocks into directory order and these samples are printed.

Each of the samples thus generated is a random sample, as is any combination of these samples. If any entry from a sample is used, then the entire sample must be used in order to preserve the validity (absence of bias) of the procedure.

In using the samples, the entire first sample is used to start a survey. Some entries, because of directory anomalies such as short pages, commercial addresses, blank lines, etc., can be identified in advance as not denoting dwelling units; others will be found, in the field, to denote non-existent or nonresidential buildings. If the remaining sample is smaller than that required, one or more of the 40-entry samples must be added to the survey sample. This process of adding 40-entry samples to the survey sample may be iterated until the survey sample is of sufficient size for the desired confidence level.

A program is provided in Appendix D along with specimen inputs and outputs.

Advantages of this method are that the creation of the sample is economical, quick, and easily reproducible. The disadvantages are that there may be lag time until the program is created (or adapted) for the available computer configuration, and, of course, access to the computer is required.

APPENDIX B

ESTIMATING INCIDENCE OF LEAD PAINT

The procedures used to estimate the hazard fractions after the data collection has been completed are similar to those described in the preceding section for estimation of sample size. For the present analyses, the sample size and hazard fraction for the sample are given; the precision with which the sample fraction P represents the total population fraction P must be determined. To illustrate we will consider several examples.

First, assume an extremely small total population (N=400), with a sample size, n of 200, and that the number of hazardous units in the sample is 100. Thus p = .5 and Tables 1 and 2 (in Part I) indicate that: 1) there is a 95% probability that P lies between 43.1% and 56.9%, i.e, that the actual number of units lies between 172 and 228, and 2) there is a 90% probability that P lies between 44.2% and 55.8%, i.e. that the actual number of hazardous units lies between 176 and 223. However, since the confidence ranges in Tables 1 and 2 are based on "large" populations and our N of 400 is "small", we also consult Table 3 in Appendix A. This shows that there is at least a 95% probability that P is between 45% and 55% (180 to 220 hazardous houses), because a "small" N means that n = 200 is relatively quite "large", the range for P is reduced by 3.8 percentage points.

As a second example, we consider a case in which n does not appear in Tables 1 and 2. Let N = 5000, p = .2, and n = 150. Since we have no confidence ranges in these tables for n = 150, we must use the entry from the table for the largest sample size which is less than n. In this case the value is 100 and we can say: 1) there is at least a 95% probability that P lies between 12.2% and 27.8%, and, 2) there is at least a 90% probability that P lies between 13.4% and 26.8%.

From Table 3 we see that the effect of having a relatively small N is not enough to assure a 95% probability that P lies between 15% and 25%.

We can, however, estimate the confidence interval for any given confidence level (95% in the example following) by a direct calculation from the sample size equation:

$$n = \frac{k^2NPQ}{k^2PQ + (N-1)E^2}$$

Substituting p for P and q for Q and solving for E yields:

$$E^2 = \frac{k^2pq}{n} \frac{(N-n)}{(N-1)}$$

$$E^2 = \frac{(1.96)^2 \times 2 \times 8}{150} \frac{(5000 - 150)}{(5000 - 1)}$$

$$E^2 = .00399 \text{ or}$$

$$E \quad 6.3\%$$

Thus there is a 95% probability that P lies between (20.0 - 6.3%) and 20 + 6.3%), i.e., between 13.7% and 26.3%. This narrows the confidence interval from that which was the best attainable from the table.

Alternatively, we can calculate the confidence level for any prescribed interval by solving the sample size equation for k. We give two examples:

A. For a confidence range of 15 - 25%

$$k^2 = \frac{E^2n}{pq} \frac{(N-1)}{(N-n)}$$

$$k^2 = \frac{.05^2 \times 100}{.8 \times .2} \frac{(5000 - 1)}{(5000 - 100)}$$

$$k^2 = 1.60 \text{ or}$$

$$k = 1.3$$

Using Table 4 of Appendix A, we find that k = 1.3 corresponds to C = 80.6; thus there is a probability of 80.6% that P lies between 15% and 25%.

B. For a confidence range of 10 - 30%

$$k^2 = \frac{.1^2 \times 100}{.8 \times .2} \frac{(5000 - 1)}{(5000 - 100)}$$

$$k^2 = 6.37 \text{ or}$$

$$k = 2.5$$

Using Table 4 we find that there is a 98.8% probability that P lies between 10% and 30%.

APPENDIX C

INSTRUCTIONS FOR COMPLETING THE SAMPLE SURVEY FORM

The instructions for completing the data collection were used in a survey in which some 4000 dwelling units were inspected. The specification as to legitimate codes, conventions, etc. reflect the format of the instrument itself, the editing and analysis software used, and the objectives of the survey. They are not presented as doctrine but rather as an example of instructions which were effective in a particular survey.

General

The Data Collection Form (NBS-744), hereafter abbreviated as DCF, is the sole reporting instrument for the housing portion of a survey. There are several different types of data appearing on the DCF. Some of the data will have been entered prior to the inspector's receipt of the DCF; others will be transcribed from the XRF log book; the remainder will be collected on-site from the individual dwelling unit.

In the description of the data elements, the term "field" is used to denote a sequence of characters or digits which are treated as an entity. DU is used as an abbreviation for dwelling unit.

The DCF consists of 19 lines each pre-numbered in the field CODE. The field SERIAL NUMBER forms the unique identification of the DCF. Generally, if the entity defined by a line of the DCF does not exist (a one story DU, for example, would have no stairway), that line is left completely blank, when the data are collected. The position of the relevant (or non-blank) characters within the field is immaterial e.g. ΔΔ1 is considered identical to Δ1Δ or 1ΔΔ, where Δ is used to denote the character "blank". However, the characters must be successive (a field such as 1Δ2 is meaningless whereas Δ12 and 12Δ are meaningful and equivalent). Decimal points are ignored in entering XRF readings; a decimal point is assumed in subsequent processing. No particular convention must be observed to avoid ambiguity of l's and I's or 0's and o's.

The right hand side (RHS) of the DCF contains a comprehensive list of the codes required in the body of the form.

When the inspection is completed, lines 1, 2, 3 and one or more of the lines from 4 to 9 will be entered. If the DU did not exist (mistake in the directory), only line 1 will be entered. If the DU exists and could not be inspected, lines 1 and 3 will be entered.

<u>Section I - Identification</u>

Line 01 - This line is not the responsibility of the inspector. All the fields here except "V" will be filled before he receives the form. The field "V" will be filled, if necessary, by the supervisor after the form is returned.

<u>SERIAL NUMBER</u> - the arbitrarily assigned unique identifying number for the dwelling unit. It must be numeric.

<u>TRACT</u> - the census tract in which the DU lies.

<u>BLOCK</u> - the census block in which the DU lies.

<u>ZIP CODE</u> - the postal ZIP CODE of the DU mailing address.

<u>V</u> - the visitation code which is entered after the DCF is returned. If an inspection has been made, this field is left blank, otherwise V describes the reason the inspection could not be made. The codes are listed in I - VISITATION on the RHS. The accompanying codes is self-explanatory; note, however, that 3 and 6 are different.

<u>STREET NAME AND NUMBER</u> - self-explanatory. This field is used only by the inspector for locating the DU. It will be dropped in the subsequent construction of the data base.

Line 02 - These data are collected daily and copied onto the DCF from the XRF Log Book. They relate to the day of inspection rather than the specific DU.

XRF SERIAL - the three character numeric serial number of the XRF instrument used.

TEST BLOCK - the XRF reading obtained from the painted panel test block included with each instrument. This is taken after the instrument is calibrated.

ZERO READING - the reading obtained from the non-leaded wood block included with each instrument. This is taken as the instrument is calibrated.

ZERO VERNIER - the zero vernier (dial) setting corresponding to the zero reading.

CALIBRATE READING - the reading obtained from the lead-foil-on-wood block included with the instrument used in this survey. This is taken as the instrument is calibrated.

CALIBRATE VERNIER - the vernier setting corresponding to the calibrate reading.

PEAK READING - the reading obtained from the lead-foil-on-wood block included with each instrument. This is taken as the instrument is calibrated.

PEAK VERNIER - the vernier setting corresponding to the Peak Reading.

INSPECTORS - the initials of each of the members of the inspecting team

DATE-MONTH - conventional numeric codes 1-12 for month,

DATE-DAY - day of month.

Line 03 - This line is completed on-site. The fields YEAR and those following require the active participation of the householder.

SEQ - is the sequence within the day, in which the DU was inspected (first, second, third, etc.). Do not include DU's which were visited but not inspected.

TEST BLOCK - the average of five XRF readings on-site of the medium painted panel test block included with the instrument.

TYPE - coded according to II--Type of Construction--on the RHS. If the type is mixed, enter the code for the predominant type.

XS - the material of the exterior according to III--Outside Surface of Building--on the RHS. If the surface is mixed, enter the predominant type code.

OC - the code according to IV--Occupancy--on the RHS. For multi-unit buildings, a mailbox count is a convenient way of ascertaining the number of units.

YEAR - the code according to V--Year Built--on the RHS. This should be obtained from the householder if possible; otherwise it is to be estimated by the inspector.

OWNER - the code according to VI--Owner-Renter--on the RHS.

MORT - the code according to VII--Mortgage--on the RHS.

PUBLIC - the code according to VIII--Public Housing--on the RHS. Public housing is that which is operated by some agency of government whether federal, state, or local.

SUB - the subsidized housing code according to IX--Subsidized--on the RHS. Subsidized housing is that for which at least some portion of the rent is paid by some agency of federal, state, or local government.

CHILD - the number of children, age 6 and under, resident in the DU.

Section II - Interior

This section of the DCF includes lines 4 through 17. Each of these lines has the same format; the pre-printed CODE field identifies the room type.

No attempt is to be made to move furniture, pictures, etc., or to stand on anything to obtain readings. If the entity is inaccessible (the ceiling, for instance) or does not exist, the field is left blank.

Lines 04 - 17.

Walls and Ceilings

COND - the material and condition of the walls, ceiling, etc. This is a three character field in which one character is a material code according to XI (a) of the RHS; one character is a base or a substrate condition code according to XI (b) of the RHS; and the third character is a surface condition code according to XI (c) of the RHS. The order in which these characters appear is immaterial. The code P2Q, for example is equivalent to PQ2, 2PQ, or QP2.

These codes are for the room in general rather than for each specific surface. If the surfaces are mixed, the code which most nearly characterizes the room at the height of four feet or less should be entered.

Base Condition codes must come from a subjective evaluation of the painted or varnished surfaces. The criterion is an estimation of the magnitude of work involved if redecoration were considered:

1. If no work is required

2. If minor work is required. (small cracks, no deep or large holes.)

3. If major work is required. (Large cracks, deep or large holes, structural repairs).

Since COND reflects the general condition of the room, it is best to complete this field after the room inspection has been completed.

<u>WALL 1, WALL 2, WALL 3, WALL 4, and CEILING</u> - each of these is an XRF reading taken on the indicated surface. The walls are numbered clockwise beginning with the wall to the left of the entry. The wall reading should be taken at any point less than four feet from the floor if such a point is accessible. The ceiling reading is to be taken only if it is possible to do so by standing on the floor or stairs. (Do not stand on furniture or counter tops, etc.) If the surface is inaccessible enter an "x" in the field; if the surface does not exist (three walled room for example) leave the reading field blank. If the surface is not painted or varnished, enter a U rather than an XRF reading.

<u>TRIM COND</u> - as with the COND field for walls and ceilings, this field must be for all trim within the room (i.e., windows, doors, and baseboards).

<u>WINDOW NUMBER</u> - Enter the number of windows in the room ("0" if none)

<u>WINDOW READING</u> - enter the reading obtained from the window frame or sill; if possible this should be from a vertical surface within four feet of the floor.

DOOR NUMBER - enter the number of doors in the room ("0" if none)

DOOR READING - enter the XRF reading from any point on the door's surface which is within four feet of the floor. If the door is of variable thickness, take this reading from the thick part of the door to avoid the read through problem (detecting lead on the back surface of the door).

BASEBOARD - enter the XRF reading from an accessible part of the baseboard from any convenient wall of the room. If the baseboard is inaccessible enter an "x"; if there is no baseboard, leave blank.

OTHER

FLOOR - enter an XRF reading only if the floor is a painted one.

RADIATOR, CABINET, FIREPLACE - enter XRF readings for each of these which exists; leave blank for each non-existent entity.

Section III - Exterior

The exterior section differs from the interior in that there is a condition code for each XRF reading taken. Criteria for each condition are the same as for the interior. Readings are to be made only if the designated areas are painted. As with the interior readings, the XRF readings should be taken at any point at a height of four feet or less.

WALL RDG - the outside wall of the dwelling unit, reading taken from the predominant exterior surface.

PORCH RDG - the floor of the porch.

DOOR RDG - the exterior door. If the door is of variable thickness, this reading is taken from the thick portion.

WINDOW RDG - the window frame or sill.

RAILING RDG - the porch or stair railing.

FENCE RDG - the fence.

GARAGE RDG - the garage wall.

EXCLUDED ROOMS - enter the number of rooms of indicated type which were not inspected.

APPENDIX D

SPECIFICATIONS FOR SAMPLE GENERATION PROGRAM

This appendix contains the specifications for a computer program to generate samples, the listing of the program and an example of the output generated. For the example, 1023 entries were generated with 320 appearing in the first sample. The directory assumed has two columns with sixty lines per page; the names appearing beginning on page 6 and run through page 50; the starting seed (base 8) = 000011111111 012345012345.

Sample Generator

The sample generator creates a set of samples from a hard copy directory. It is parameterized so that any directory may be used. The parameters:

1. Page number of first page containing addresses (must be ≥ 1 and <1024)
2. Page number of last page containing addresses (must be ≥ 1 and <1024)
3. Number of columns per page (must be ≥ 1 and <8)
4. Number of lines per page (must be ≥ 1 and <512)
5. Size of first sample (must be <8190; the sample produced will be the least multiple of 160 which is greater than or equal to the size specified)

The program generates 8191 triples, every triple being a set of randomly generated indices denoting page, column and line of the directory. Duplicate triples are eliminated and the first sample set is put into directory sort and printed. This is the original sample to be used in the survey. The remainder of the triples are then sorted by forty entry blocks into directory sort and these sets are printed as extras.

Each of the sample sets thus generated is a random sample as is any combination of these sets. In order to preserve the validity, if any entry from a group is used that entire sample set must be used:

otherwise a bias occurs (see Section I on sampling).

Because of inconsistencies in the directory (short pages, commercial addresses, blank lines, etc.) and other problems which may be discovered in the field (non-existent or non-residential buildings), extra sets may have to be added to the survey sample until the required number of units are obtained.

This program consists of a main program (DIRECT) and a subroutine (SORT) both written in FORTRAN V, and a function subroutine (RAND) in assembly language. The FORTRAN programs should be readily transportable to any computer system of sufficient size with word length of 36 bits or more. On another system, RAND must be replaced by a function which calculates a random variable, x, $0 \leq x < 1$, with a statement:
X = RANDNO (0, SEED) where SEED is changed within the function subroutine to advance the random variable with successive calls. The method used in RAND is the conventional multiplicative one which requires input of a starting value of SEED and provides for generation of a different sequence of random variables on subsequent runs by printing the last value of SEED at run termination.

Required Input

1. Seed & sentinel (Format 3013)
    a. High order part of SEED
    b. Low order part of SEED
    c. Sentinel of all binary 1's (377777777777)
2. Directory parameters (Format 5I5)
    a. IPS Starting page
    b. IPN Last page

c. IC Number of columns/page

d. IL Number of lines/page

e. IFIRST Size of first sample

Output:

INPUT PRINT

α_1 α_2 α_3

β_1 β_2 β_3 β_4 β_5

NEXT VALUE OF SEED γ_1 γ_2 THERE ARE δ UNIQUE UNITS followed by the samples generated.

α_1 is 12 character octal high order part of initial SEED

α_2 is 12 character octal low order part of initial SEED

α_3 is 377777777777 (α_1 - α_3 are print back of seed & sentinel input)

β_1 is IPS

β_2 is IPN

β_3 is IC

β_4 is IL

β_5 is IFIRST (β_1-β_5 are print-back of directory parameters.)

γ_1 is 12 character octal high order part of last SEED

γ_2 is 12 character octal low order part of last SEED

δ is the number of distinct units in the 8191 unit sample after duplicates have been eliminated.

```
@FCR,IS DIRECT
FOR 0E2B-02/23/76-16:34:37 (,0)

   MAIN PROGRAM

   STORAGE USED: CODE(1) 000705; DATA(0) 062027; BLANK COMMON(2) 000000


   EXTERNAL REFERENCES (BLOCK, NAME)

    0003    RANDNO
    0004    SORTP
    0005    SORT
    0006    NINTR$
    0007    NRDC$
    0010    NIO2$
    0011    NPRT$
    0012    NIO1$
    0013    NSTOP$


   STORAGE ASSIGNMENT  (BLOCK, TYPE, RELATIVE LOCATION, NAME)

    0001     000436 110L      0001     000105 147G      0001     000112 154G       0001     000223 175G       0001     000235 204G
    0001     000267 225G      0001     000313 236G      0001     000320 243G       0001     000341 252G       0001     000366 270G
    0001     000376 275G      0001     000254 30L       0001     000447 320G       0001     000504 337G       0001     000570 363G
    0001     000576 370G      0001     000606 375G      0001     000657 420G       0001     000646 430L       0001     000351 90L
    0000     061721 900F      0000     061723 901F      0000     061724 950F       0000     061727 951F       0000     061734 952F
    0000     061743 953F      0000     061754 954F      0000     061762 955F       0000     061770 956F       0000     061777 957F
    0000     062002 958F      0000 I 061667 IC         0000 I 061671 IFIRST      0000 I 061670 IL          0000 I 061705 IPAGE
    0000 I 061666 IPN         0000 I 061665 IPS        0000 I 061677 I1          0000 I 061700 I2          0000 I 061701 I3
    0000 I 061703 I4          0000 I 061704 I5         0000 I 061675 J           0000 I 061713 JL          0000 I 061714 JM
    0000 I 061702 JN          0000 I 061711 J1         0000 I 061712 J2          0000 I 061706 J3          0000 I 061707 J4
    0000 I 061715 J5          0000 I 061716 J6         0000 I 061717 J7          0000 I 061720 J8          0000 I 061710 J9
    0000 I 000000 K           0000 I 020464 K1         0000 I 041150 K2          0000 I 061634 L           0000 I 061662 N0
    0000 I 061663 N1          0000 I 061664 N2         0003 R 000000 RANDNO      0000 R 061673 RC          0000 R 061674 RL
    0000 R 061672 RP          0000 R 061660 SEED       0000 R 061676 X


 00101        1*         DIMENSION K(8500),K1(8500),K2(8500),L(20),SEED(2)                              000000
 00103        2*         N0=8190                                                                        000001
 00104        3*         N1=8191                                                                        000003
 00105        4*         N2=8192                                                                        000005
 00106        5*         N0=1022                                                                        000007
 00107        6*         N1=1023    Remove for Production                                               000011
 00110        7*         N2=1024                                                                        000013
 00111        8*         READ 900 SEED(1),SEED(2),K(1)                                          60      000015
 00116        9*         READ 901 IPS,IPN,IC,IL,IFIRST                                          70      000024
 00125       10*         PRINT 950                                                              80      000035
 00127       11*         PRINT 900 SEED(1),SEED(2),K(1)                                         90      000041
 00134       12*         PRINT 901 IPS,IPN,IC,IL,IFIRST                                        100      000050
 00143       13*         RP=IPN-IPS                                                            110      000061
 00144       14*         RC=IC                                                                 120      000066
 00145       15*         RL=IL                                                                 130      000071
 00146       16*         DO 10  J=2,8500                                                       133      000105
```

```
00151    17*         10 K(J)=K(1)                                                        136   000105
00151    18*      C * * GENERATE 8191 ENTRIES * * * * * * * * * * * * * * * * * * * * *   140   000105
00153    19*           DO 20  J=1,N1                                                     150   000112
00156    20*           X=RANDNO(0,SEED)                                                  160   000112
00157    21*           I1=RP*X+IPS              @ PAGE                                   170   000117
00160    22*           X=RANDNO(0,SEED)                                                  180   000130
00161    23*           I2=RC*X+1                @ COLUMN                                 190   000135
00162    24*           X=RANDNO(0,SEED)                                                  200   000146
00163    25*           I3=RL*X+1                @ LINE                                   210   000153
00163    26*      C SAMPLE SIZE 13 BITS 8191                                             220   000153
00163    27*      C PAGE SIZE   10 BITS  1023                                            230   000153
00163    28*      C COLUMN SIZE  3 BITS   7                                              240   000153
00163    29*      C LINE SIZE    9 BITS   511                                            250   000153
00164    30*           K(J)=(((I1*8)+I2)*512+I3)*8192+J                                  260   000164
00165    31*         20 CONTINUE                                                         270   000176
00167    32*           PRINT 951 SEED(1),SEED(2)                                         280   000176
00167    33*      C * * SORT INTO DIRECTORY ORDER * * * * * * * * * * * * * * * * * *    290   000176
00173    34*           CALL SORTP (K,N1,K1,K2)                                                 000204
00174    35*           DO 25   J=1,N1                                                          000212
00177    36*         25 K(J)=K1(J)                                                             000223
00201    37*           I1=1                                                              310   000225
00202    38*           I2=K(1)/8192                                                      320   000227
00203    39*           DO 30  JN=2,N1                                                    330   000235
00206    40*           I4=K(JN)                                                          340   000235
00207    41*           I5=I4/8192                                                        350   000236
00210    42*           K(JN)=K(8500)                                                     360   000240
00211    43*           IF (I2.EQ.I5)   GO TO 30                                          370   000242
00213    44*           K(JN)=I4                                                          380   000244
00214    45*           I1=I1+1                                                           390   000246
00215    46*           I2=I5                                                             400   000251
00216    47*         30 CONTINUE                                                         410   000255
00220    48*           JN=I1                                                             420   000255
00221    49*           PRINT 952   I1                                                    440   000257
00224    50*           DO 70 J=1,N1                                                      460   000267
00227    51*           I1=K(J)/8192                                                      470   000267
00230    52*           I2=K(J)-8192*I1                                                   480   000271
00231    53*           K(J)=I1+4194304*I2                                                490   000274
00232    54*         70 CONTINUE                                                         495   000300
00232    55*      C * * SORT INTO GENERATED ORDER * * * * * * * * * * * *                500   000300
00234    56*           CALL SORTP (K,N1,K1,K2)                                                 000300
00235    57*           DO 75   J=1,N1                                                          000306
00240    58*         75 K(J)=K1(J)                                                             000313
00242    59*           DO 80 J=1,8499                                                    520   000320
00245    60*           I1=K(J)/4194304                                                   530   000320
00246    61*         80 K(J)=K(J)-4194304*I1                                             540   000322
00246    62*      C * * TOTAL SAMPLE INTO REGISTER SORT * * * * * * * *                  550   000322
00250    63*           CALL SORTP (K,IFIRST,K1,K2)                                             000326
00251    64*           DO 85   J=1,IFIRST                                                      000334
00254    65*         85 K(J)=K1(J)                                                             000341
00256    66*           IPAGE =1                                                          570   000343
00257    67*           J3=1                                                              580   000345
00260    68*           J4=40                                                             590   000346
00261    69*         90 I3=5                                                             600   000351
00262    70*           J9=0                                                              605   000352
00263    71*           PRINT 953 IPAGE,IFIRST                                            610   000353
00267    72*           DO 120  J=J3,J4                                                   620   000366
00272    73*           J9=J9+1                                                           625   000366
00273    74*           I2=1                                                              630   000371
```

```
00274    75*          DO 100  J1=1,121,40                          640    000376
00277    76*          J2=J+J1-1                                    650    000377
00300    77*          L(I2)=K(J2)/4096                             660    000403
00301    78*          I1=K(J2)-L(I2)*4096                          670    000407
00302    79*          L(I2+1)=I1/512                               680    000412
00303    80*          L(I2+2)=I1   -L(I2+1)*512                    690    000414
00304    81*          I2=I2+3                                      693    000417
00305    82*      100 CONTINUE                                     696    000424
00307    83*          IF (I3.LT.5)  GO TO 110                      700    000424
00311    84*          I3=0                                         710    000430
00312    85*          PRINT 954                                    720    000431
00314    86*      110 I3=I3+1                                      730    000436
00315    87*          PRINT 955 J9,(L(J2),J2=1,12),J9              750    000440
00325    88*      120 CONTINUE                                     760    000456
00327    89*          IPAGE =IPAGE+1                               770    000456
00330    90*          J3=J3+160                                    780    000461
00331    91*          J4=J4+160                                    790    000464
00332    92*          IF (J3.LT.IFIRST)  GO TO 90                  800    000467
00334    93*      400 JL=(JN-IFIRST)/160+1                         810    000472
00335    94*          L(16)=0                                      815    000500
00336    95*          DO 450  JM=1,JL                              820    000504
00341    96*          J5=J3+40                                     830    000504
00342    97*          J6=J5+40                                     840    000506
00343    98*          J7=J6+40                                     850    000510
00344    99*          CALL SORT (K,40,J3)                          860    000512
00345   100*          CALL SORT (K,40,J5)                          870    000517
00346   101*          CALL SORT (K,40,J6)                          880    000524
00347   102*          CALL SORT (K,40,J7)                          890    000531
00350   103*          I3=5                                         900    000536
00351   104*          J9=0                                         910    000540
00352   105*          PRINT 956  IPAGE                             920    000541
00355   106*          L(13)=L(16)+1                                930    000546
00356   107*          L(14)=L(13)+1                                940    000551
00357   108*          L(15)=L(14)+1                                950    000553
00360   109*          L(16)=L(15)+1                                960    000555
00361   110*          PRINT 957 (L(J8),J8=13,16)                   970    000561
00367   111*          DO 440  J=J3,J4                              980    000576
00372   112*          J9=J9+1                                      980    000576
00373   113*          I2=1                                         990    000601
00374   114*          DO 420  J1=1,121,40                         1000    000606
00377   115*          J2=J+J1-1                                   1010    000607
00400   116*          L(I2)=K(J2)/4096                            1020    000613
00401   117*          I1=K(J2)-L(I2)*4096                         1030    000617
00402   118*          L(I2+1)=I1/512                              1040    000622
00403   119*          L(I2+2)=I1   -L(I2+1)*512                   1050    000624
00404   120*          I2=I2+3                                     1053    000627
00405   121*      420 CONTINUE                                    1056    000634
00407   122*          IF (I3.LT.5)  GO TO 430                     1060    000634
00411   123*          I3=0                                        1070    000640
00412   124*          PRINT 954                                   1080    000641
00414   125*      430 I3=I3+1                                     1090    000646
00415   126*          PRINT 955 J9,(L(J2),J2=1,12),J9             1100    000650
00425   127*      440 CONTINUE                                    1120    000666
00427   128*          IPAGE=IPAGE+1                               1130    000666
00430   129*          J3=J3+160                                   1140    000671
00431   130*          J4=J4+160                                   1150    000674
00432   131*      450 CONTINUE                                    1160    000700
00434   132*          STOP                                        1170    000700
```

```
00435     133*        900 FORMAT (10013)                                                3000    000704
00436     134*        901 FORMAT (5I5)                                                  3010    000704
00437     135*        950 FORMAT ('0INPUT PRINT')                                       3020    000704
00440     136*        951 FORMAT (' NEXT VALUE OF SEED'2013)                            3030    000704
00441     137*        952 FORMAT (' THERE ARE'I5,' UNIQUE UNITS.')                      3040    000704
00442     138*        953 FORMAT ('1PAGE'I3,' OF ORIGINAL SAMPLE OF'I5,' UNITS')        3050    000704
00443     139*        954 FORMAT ('                                 ')                  3060    000704
00444     140*        955 FORMAT (I5,4(I7,'-',I1,'-',I3,),I6)                           3070    000704
00445     141*        956 FORMAT ('1PAGE'I3,' FORTY UNIT SAMPLES')                      3080    000704
00446     142*        957 FORMAT ('0'I14,3I13)                                          3090    000704
00447     143*        958 FORMAT (' ENTERING SORT                          ')                   000704
00450     144*            END                                                           9000    000704

      END OF COMPILATION:        NO  DIAGNOSTICS.
```

```
@FCR.IS SCRT
FOR 0E2B-02/23/76-16:34:49 (,0)

   SUBROUTINE SORT       ENTRY POINT 000063

   STORAGE USED: CODE(1) 000074; DATA(0) 000023; BLANK COMMON(2) 000000

   EXTERNAL REFERENCES (BLOCK, NAME)

    0003   NERR3$

   STORAGE ASSIGNMENT  (BLOCK, TYPE, RELATIVE LOCATION, NAME)

    0001    000022 106G      0001    000027 112G      0001    000040 20L      0001    000047 40L      0000 1 000001 1
    0000    000004 INJP$     0000 I 000002 11         0000 1 000003 12        0000 1 000000 NL


00101       1*           SUBROUTINE SORT (K,N,L)                                  10     000011
00103       2*           DIMENSION K(1)                                           20     000011
00104       3*           NL=N+L-2                                                 30     000011
00105       4*           DO 30 I=1,N                                              40     000015
00110       5*           I1=0                                                     50     000022
00111       6*           DO 20  I2=L,NL                                           60     000027
00114       7*           IF (K(I2).LT.K(I2+1))  GO TO 20                          70     000027
00116       8*           I1=K(I2)                                                 80     000032
00117       9*           K(I2)=K(I2+1)                                            90     000034
00120      10*           K(I2+1)=I1                                              100     000036
00121      11*        20 CONTINUE                                                110     000042
00123      12*           IF (I1.EQ.0)  GO TO 40                                  120     000042
00125      13*        30 CONTINUE                                                130     000047
00127      14*       40  RETURN                                                  170     000047
00130      15*           END                                                    9000     000073

      END OF COMPILATION:       NO  DIAGNOSTICS.
```

```
@ASM,IS RAND
ASM13B RLIB69 02/23-16:34:50-( ,0)

    0C   000000  71 13 01 13 1 000001   RANDNO*  DL      13,*1,11
         000001  71 12 01 00 0 000000            DS      13,IS
         000002  27 17 14 00 000000              L,017   12,0
         000003  27 00 15 00 0 000001            L       13,IS+1
         000004  32 00 01 00 0 000002            MF      13,M
         000005  73 00 02 00 0 000001            SSC     14,1
         000006  27 00 17 00 0 000000            L       15,IS
         000007  30 00 03 00 0 000002            MI      15,M
         000010  71 10 00 00 0 000017            DA      12,15
         000011  73 12 01 00 0 000001            LSSL    13,1
         000012  73 02 01 00 0 000001            SSL     13,1
         000013  71 12 01 00 0 000000            DS      13,IS
         000014  27 17 14 00 000170              L,017   12,0170
         000015  76 05 00 00 0 000015            LCF     12,13
         000016  27 00 14 00 0 000015            L       12,13
         000017  71 13 01 00 0 000000            DL      13,IS
         000020  71 12 01 13 1 000001            DS      13,*1,11
         000021  51 00 00 13 1 000000            TNZ     *0,11
         000022  74 04 00 13 0 000003            J       3,11
         000023  24 00 14 00 0 000003            A       12,BIT
         000024  76 01 00 00 0 000004            FAN     12,ONE
         000025  74 04 00 13 0 000003            J       3,11
    01   000000  154447730601           $(1),IS  +0154447730601
         000001  255751305264                    +0255751305264
         000002  011060471625           M        +0011060471625
         000003  001000000000           BIT      +0001000000000
         000004  200777777777           ONE      +0200777777777
                                                 END

END ASM. ERRORS : NONE
```

```
@FOR,IS SCRTP
FOR 0E28-02/23/76-16:34:51 (,0)

   SUBROUTINE SORTP      ENTRY POINT 000602

   STORAGE USEO: COOE(1) 000631; OATA(0) 000273; BLANK COMMON(2) 000000

   EXTERNAL REFERENCES (BLOCK, NAME)

    0003    NWCU$
    0004    NIC2$
    0005    NERR3$

   STORAGE ASSIGNMENT  (BLOCK, TYPE, RELATIVE LOCATION, NAME)

    0001     000037 115G      0001     000064 127G      0000     000160 15F       0001     000140 157G      0001     000147 165G
    0001     000163 174G      0000     000201 18F       0001     000176 200L      0001     000204 250L      0001     000211 305L
    0001     000215 310L      0001     000254 320L      0001     000331 330L      0001     000351 340L      0001     000370 350L
    0001     000430 360L      0001     000443 370L      0001     000460 380L      0001     000474 390L      0001     000520 395L
    0000     000224 47F       0001     000076 50L       0001     000114 55L       0000     000131 9F        0001     000130 90L
    0000 I 000121 AMEO        0000 I 000122 BMED        0000 I 000111 HOLD        0000 I 000112 I           0000 I 000044 IL
    0000     000237 INJP$     0000 I 000110 IPR         0000 I 000114 IP1         0000 I 000125 ITT         0000 I 000000 IU
    0000 I 000116 J           0000 I 000130 JMI         0000 I 000127 JMK         0000 I 000117 K           0000 I 000123 L
    0000 I 000126 LMI         0000 I 000115 M           0000 I 000120 MIO         0000 I 000113 NM1         0000 I 000124 TT


00101      1*            SUBROUTINE SORTP(X,N,Y,XPOS)                                                  000013
00101      2*      C     THIS ROUTINE SORTS THE ELEMENTS OF THE INPUT VECTOR X AND PUTS THE SORTEO     000013
00101      3*      C     ELEMENTS INTO THE VECTOR Y.  IT ALSO CARRIES ALONG THE INDEX NUMBER           000013
00101      4*      C     OF EACH OROEREO OBSERVATION--THAT IS, IT CARRIES ALONG THE POSITION OF        000013
00101      5*      C     THE I-TH OROERED OBSERVATION (FOR EACH I) AS IT WAS IN THE ORIGINAL          000013
00101      6*      C     UNCROERED OATA VECTOR X.  THESE POSITIONS ARE PLACEO IN THE VECTOR XPOS.     000013
00101      7*      C     THIS ROUTINE IS USEFUL IN ATTEMPTING TO LOCATE THE MINIMUM, THE MAXIMUM,     000013
00101      8*      C     OR SOME OTHER OROEREO OBSERVATION OF INTEREST IN THE ORIGINAL UNOROERED      000013
00101      9*      C     INPUT VECTOR X.                                                               000013
00101     10*      C     THE INPUT TO THIS ROUTINE IS THE SINGLE PRECISION VECTOR X OF                000013
00101     11*      C     (UNSORTED) OBSERVATIONS, THE INTEGER VALUE N (= SAMPLE SIZE),               000013
00101     12*      C     AN EMPTY SINGLE PRECISION VECTOR Y INTO WHICH THE SORTEO OBSERVATIONS        000013
00101     13*      C     WILL BE PLACEO, ANO AN EMPTY SINGLE PRECISION VECTOR XPOS INTO WHICH THE     000013
00101     14*      C     POSITIONS OF THE SORTEO OBSERVATIONS WILL BE PLACEO.                         000013
00101     15*      C     THE OUTPUT FRCM THIS ROUTINE IS THE SINGLE PRECISION VECTOR Y INTO WHICH     000013
00101     16*      C     THE SORTED OBSERVATIONS HAVE BEEN PLACED, ANO THE SINGLE PRECISION VECTOR    000013
00101     17*      C     XPOS INTO WHICH THE POSITIONS OF THE SORTEO OBSERVATIONS HAVE BEEN PLACED.   000013
00101     18*      C     RESTRICTIONS ON THE MAXIMUM ALLOWABLE VALUE OF N--THE OIMENSIONS             000013
00101     19*      C     OF VECTORS IU AND IL (DEFINEO ANO USEO INTERNALLY WITHIN THIS ROUTINE)       000013
00101     20*      C     OETERMINE THE MAXIMUM ALLOWABLE VALUE OF N FOR THIS                          000013
00101     21*      C     ROUTINE.  IF IU AND IL EACH HAVE DIMENSION K, THEN N MAY NOT EXCEEO          000013
00101     22*      C     2**(K+1) - 1.  FOR THIS ROUTINE AS WRITTEN, THE OIMENSIONS OF IU ANO IL      000013
00101     23*      C     HAVE BEEN SET TO 36, THUS THE MAXIMUM ALLOWABLE VALUE OF N IS                000013
00101     24*      C     APPROXIMATELY 137 BILLION.  SINCE THIS EXCEEDS THE MAXIMUM ALLOWABLE         000013
00101     25*      C     VALUE FOR AN INTEGER VARIABLE IN MANY COMPUTERS, AND SINCE A SORT OF 137     000013
```

```
00101     C      BILLION ELEMENTS IS PRESENTLY IMPRACTICAL AND UNLIKELY. THEREFORE NO        000013
00101     C      TEST FOR WHETHER THE INPUT SAMPLE SIZE N EXCEEDS 137 BILLION HAS BEEN       000013
00101     C      INCORPORATED INTO THIS ROUTINE.  IT IS THUS ASSUMED THAT THERE IS NO        000013
00101     C      (PRACTICAL) RESTRICTION ON THE MAXIMUM VALUE OF N FOR THIS ROUTINE.         000013
00101     C      PRINTING--NONE UNLESS AN ERROR CONDITION EXISTS                             000013
00101     C      THIS ROUTINE IS SINGLE PRECISION IN INTERNAL OPERATION.                     000013
00101     C      SUBROUTINES NEEDED--NONE                                                    000013
00101     C      SORTING METHOD--BINARY SORT                                                 000013
00101     C      REFERENCE--CACM MARCH 1969, PAGE 186 (BINARY SORT ALGORITHM BY RICHARD      000013
00101     C                 C. SINGLETON.                                                    000013
00101     C               --CACM JANUARY 1970, PAGE 54.                                      000013
00101     C               --CACM OCTOBER 1970, PAGE 624.                                     000013
00101     C               --JACM JANUARY 1961, PAGE 41.                                      000013
00101     C      WRITTEN BY JAMES J. FILLIBEN, STATISTICAL ENGINEERING LABORATORY (205.03)   000013
00101     C      NATIONAL BUREAU OF STANDARDS, WASHINGTON, D.C. 20234        JUNE 1972       000013
00101     C                                                                                  000013
00103            DIMENSION X(1),Y(1),XPOS(1)                                                 000013
00104            DIMENSION IU(36),IL(36)                                                     000013
00105            IMPLICIT INTEGER (A-Z)                                                      000013
00105     C                                                                                  000013
00105     C      CHECK THE INPUT ARGUMENTS FOR ERRORS                                        000013
00105     C                                                                                  000013
00106            IPR=6                                                                       000013
00107            IF(N.LT.1)GOTO50                                                            000015
00111            IF(N.EQ.1)GOTO55                                                            000021
00113            HOLD=X(1)                                                                   000024
00114            DO60I=2,N                                                                   000026
00117            IF(X(I).NE.HOLD)GOTO90                                                      000037
00121         60 CONTINUE                                                                    000045
00123            WRITE(IPR, 9)HOLD                                                           000045
00126            DO61I=1,N                                                                   000064
00131            Y(I)=X(I)                                                                   000064
00132            XPOS(I)=I                                                                   000066
00133         61 CONTINUE                                                                    000072
00135            RETURN                                                                      000072
00135         50 WRITE(IPR,15)                                                               000076
00140            WRITE(IPR,47)N                                                              000102
00143            RETURN                                                                      000110
00144         55 WRITE(IPR,18)                                                               000114
00146            Y(1)=X(1)                                                                   000120
00147            XPOS(1)=1.0                                                                 000122
00150            RETURN                                                                      000124
00151         90 CONTINUE                                                                    000130
00152          9 FORMAT(1H ,108H***** NON-FATAL DIAGNOSTIC--THE FIRST  INPUT ARGUME          000130
00152           1NT (A VECTOR) TO THE SORTP  SUBROUTINE HAS ALL ELEMENTS = ,E15.8,6          000130
00152           1H *****)                                                                    000130
00153         15 FORMAT(1H , 91H***** FATAL ERROR--THE SECOND INPUT ARGUMENT TO THE          000130
00153           1 SORTP  SUBROUTINE IS NON-POSITIVE *****)                                   000130
00154         18 FORMAT(1H ,100H***** NON-FATAL DIAGNOSTIC--THE SECOND INPUT ARGUME          000130
00154           1NT TO THE SORTP  SUBROUTINE HAS THE VALUE 1 *****)                          000130
00155         47 FORMAT(1H , 35H***** THE VALUE OF THE ARGUMENT IS ,I8   ,6H *****)          000130
00155     C                                                                                  000130
00155     C      COPY THE VECTOR X INTO THE VECTOR Y                                         000130
00156            DO100I=1,N                                                                  000130
00161            Y(I)=X(I)                                                                   000140
00162        100 CONTINUE                                                                    000147
00162     C                                                                                  000147
00162     C      DEFINE THE XPOS (POSITION) VECTOR.  BEFORE SORTING, THIS WILL               000147
```

```
00162     C      BE A VECTOR WHOSE I-TH ELEMENT IS EQUAL TO I.                000147
00162     C                                                                   000147
00164            DO150I=1,N                                                   000147
00167            XPOS(I)=I                                                    000147
00170        150 CONTINUE                                                     000153
00170     C                                                                   000153
00170     C      CHECK TC SEE IF THE INPUT VECTOR IS ALREADY SORTED           000153
00170     C                                                                   000153
00172            NM1=N-1                                                      000153
00173            DO200I=1,NM1                                                 000156
00176            IP1=I+1                                                      000163
00177            IF(Y(I).LE.Y(IP1))GOTO200                                    000170
00201            GOTO250                                                      000174
00202        200 CONTINUE                                                     000200
00204            RETURN                                                       000200
00205        250 M=1                                                          000204
00206            I=1                                                          000205
00207            J=N                                                          000206
00210        305 IF(I.GE.J)GOTO370                                            000211
00212        310 K=I                                                          000215
00213            MID=(I+J)/2                                                  000216
00214            AMED=Y(MID)                                                  000231
00215            BMED=XPOS(MID)                                               000233
00216            IF(Y(I).LE.AMED)GOTO320                                      000235
00220            Y(MID)=Y(I)                                                  000240
00221            XPOS(MID)=XPOS(I)                                            000242
00222            Y(I)=AMED                                                    000244
00223            XPOS(I)=BMED                                                 000246
00224            AMED=Y(MID)                                                  000247
00225            BMED=XPCS(MID)                                               000251
00226        320 L=J                                                          000254
00227            IF(Y(J).GE.AMED)GOTO340                                      000255
00231            Y(MID)=Y(J)                                                  000275
00232            XPOS(MID)=XPOS(J)                                            000277
00233            Y(J)=AMED                                                    000301
00234            XPOS(J)=BMED                                                 000303
00235            AMED=Y(MID)                                                  000305
00236            BMED=XPOS(MID)                                               000307
00237            IF(Y(I).LE.AMED)GOTO340                                      000311
00241            Y(MID)=Y(I)                                                  000314
00242            XPOS(MID)=XPOS(I)                                            000316
00243            Y(I)=AMED                                                    000320
00244            XPOS(I)=BMED                                                 000322
00245            AMED=Y(MID)                                                  000323
00246            BMED=XPCS(MID)                                               000325
00247            GOTO340                                                      000327
00250        330 Y(L)=Y(K)                                                    000331
00251            XPOS(L)=XPOS(K)                                              000336
00252            Y(K)=TT                                                      000344
00253            XPOS(K)=ITT                                                  000346
00254        340 L=L-1                                                        000351
00255            IF(Y(L).GT.AMED)GOTO340                                      000353
00257            TT=Y(L)                                                      000361
00260            ITT=XPOS(L)                                                  000363
00261        350 K=K+1                                                        000370
00262            IF(Y(K).LT.AMED)GOTO350                                      000372
00264            IF(K.LE.L)GOTO330                                            000400
00266            LMI=L-I                                                      000404
```

```
00267            JMK=J-K                                000407
00270            IF(LMI.LE.JMK)GOTO360                  000412
00272            IL(M)=I                                000415
00273            IU(M)=L                                000417
00274            I=K                                    000421
00275            M=M+1                                  000423
00276            GOTO380                                000426
00277        360 IL(M)=K                                000430
C0300            IU(M)=J                                000432
00301            J=L                                    000434
00302            M=M+1                                  000436
00303            GCTO380                                000441
00304        370 M=M-1                                  000443
00305            IF(M.EQ.0)RETURN                       000445
00307            I=IL(M)                                000452
00310            J=IU(M)                                000455
00311        380 JMI=J-I                                000460
00312            IF(JMI.GE.11)GOTO310                   000462
00314            IF(I.EQ.1)GOTO305                      000465
00316            I=I-1                                  000470
00317        390 I=I+1                                  000474
00320            IF(I.EQ.J)GOTO370                      000476
00322            AMED=Y(I+1)                            000500
00323            BMED=XPOS(I+1)                         000504
00324            IF(Y(I).LE.AMED)GOTO390                000510
00326            K=I                                    000515
00327        395 Y(K+1)=Y(K)                            000520
C0330            XPCS(K+1)=XPOS(K)                      000525
00331            K=K-1                                  000533
00332            IF(AMED.LT.Y(K))GOTO395                000536
00334            Y(K+1)=AMED                            000544
00335            XPOS(K+1)=BMED                         000550
00336            GOTO390                                000554
00337            END                                    000630

     END OF COMPILATION:        NO  DIAGNOSTICS.
```

```
@XQT
MAF27.1 RL71-3 02/23/76 16:34:54

ADDRESS LIMITS     001000 012653      5036 IBANK WORDS DECIMAL
                   040000 126312     27851 DBANK WORDS DECIMAL
STARTING ADDRESS   011747


                SEGMENT $MAIN$          001000 012653     040000 126312

NSWTC$/FOR69            $(1)    001000 001024
NRBLK$/FOR-E2           $(1)    001025 001047
NRWND$/FOR-E2           $(1)    001050 001133    $(2)    040000 040011
NWEF$/FOR-E2            $(1)    001134 001341    $(2)    040012 040031
NCLCS$/FOR-E2           $(1)    001342 001574    $(2)    040032 040057
NWBLK$/FOR68            $(1)    001575 001706
NBSBL$/FOR68            $(1)    001707 001747
NUPCA$/FOR68            $(1)    001750 002003
NBFOO$/FOR-E2                                    $(2)    040060 042261
NBCCV$/FOR64            $(1)    002004 002131    $(2)    042262 042324
NFTV$/FOR-E2            $(1)    002132 002154
NCNVT$/FOR68            $(1)    002155 002376    $(2)    042325 042421
NCTIN$/FOR68            $(1)    002377 002673    $(2)    042422 042425
NFCHK$/FOR-E2C          $(1)    002674 003660    $(2)    042426 042576
                        $(3)    003661 003661    $(4)    042577 042650
NOUT$/FOR-E2C           $(1)    003662 005036    $(2)    042651 042707
NICER$/FOR-E2           $(1)    005037 005230    $(2)    042710 043055
NFMT$/FOR-E2C           $(1)    005231 006113    $(2)    043056 043132
NINFT$/FOR-E2C          $(1)    006114 007265    $(2)    043133 043167
NTAB$/FOR                                        $(2)    043170 043252
NOBUF$/FOR68            $(1)    007266 007326
NERR$/FOR-E2            $(1)    007327 007752    $(2)    043253 043457
NSTOP$/FOR69            $(1)    007753 007777    $(2)    043460 043467
NCSYM$/FOR69            $(1)    010000 010242    $(2)    043470 043471
NIER$/FOR-E2            $(1)    010243 010421    $(2)    043472 043610
NISYM$/FOR69            $(1)    010422 010575    $(2)    043611 043612
NINTR$/NBS              $(1)    010576 011014    $(2)    043613 043717
BLANK$COMMON(COMMONBLOCK)
SORTP                   $(1)    011015 011645    $(0)    043720 044212
                                                 $(2)    BLANK$COMMON
RAND                    $(1)    011646 011652    $(0)    044213 044240
SORT                    $(1)    011653 011746    $(0)    044241 044263
                                                 $(2)    BLANK$COMMON
DIRECT                  $(1)    011747 012653    $(0)    044264 126312
                                                 $(2)    BLANK$COMMON

SYS$*RLIB$. LEVEL 71
END MAP

 INPUT PRINT
 000011111111 012345612345 377777777777
    6    50    2   60   320
 NEXT VALUE OF SEED 141634423274 152315541031
 THERE ARE  929 UNIQUE UNITS.
```

PAGE 1 OF ORIGINAL SAMPLE OF 320 UNITS

1	6-1- 18	12-2- 50	18-1- 33	22-2- 30	1
2	6-1- 27	12-2- 52	18-1- 54	22-2- 33	2
3	6-1- 29	12-2- 56	18-1- 56	23-1- 11	3
4	6-2- 3	13-1- 5	18-2- 29	23-1- 25	4
5	7-1- 56	13-1- 32	18-2- 31	23-1- 46	5
6	7-2- 1	13-1- 43	18-2- 33	23-2- 11	6
7	7-2- 13	13-1- 52	18-2- 41	23-2- 13	7
8	7-2- 26	13-2- 24	18-2- 49	23-2- 16	8
9	7-2- 32	13-2- 25	19-1- 23	23-2- 40	9
10	7-2- 39	14-1- 19	19-1- 25	23-2- 55	10
11	7-2- 43	14-1- 37	19-2- 7	24-1- 59	11
12	8-1- 33	14-1- 47	19-2- 21	24-2- 9	12
13	8-1- 54	14-2- 1	19-2- 28	24-2- 15	13
14	8-2- 4	14-2- 18	19-2- 29	24-2- 23	14
15	8-2- 5	14-2- 39	20-1- 8	24-2- 38	15
16	8-2- 15	14-2- 45	20-1- 9	24-2- 58	16
17	9-1- 6	14-2- 53	20-1- 41	25-1- 12	17
18	9-1- 7	14-2- 54	20-1- 42	25-1- 19	18
19	9-1- 10	14-2- 55	20-1- 45	25-1- 21	19
20	9-1- 34	15-1- 6	20-2- 27	25-1- 54	20
21	9-1- 42	15-1- 14	20-2- 29	25-2- 33	21
22	9-2- 18	15-1- 49	20-2- 32	25-2- 53	22
23	9-2- 20	15-1- 60	20-2- 40	26-1- 36	23
24	9-2- 35	15-2- 3	20-2- 42	26-1- 38	24
25	9-2- 53	15-2- 31	20-2- 52	26-1- 50	25
26	9-2- 58	15-2- 46	21-2- 3	26-2- 8	26
27	10-1- 2	16-1- 16	21-2- 10	26-2- 12	27
28	10-1- 47	16-1- 18	21-2- 23	26-2- 26	28
29	10-1- 50	16-1- 34	21-2- 27	26-2- 47	29
30	10-2- 24	16-2- 29	21-2- 29	26-2- 49	30
31	11-1- 3	16-2- 39	21-2- 38	27-1- 26	31
32	11-1- 41	17-1- 13	21-2- 41	27-1- 28	32
33	11-2- 2	17-1- 18	21-2- 54	27-2- 26	33
34	11-2- 32	17-2- 17	22-1- 22	27-2- 36	34
35	11-2- 53	17-2- 39	22-1- 48	27-2- 49	35
36	12-1- 57	17-2- 41	22-1- 51	27-2- 54	36
37	12-2- 1	17-2- 44	22-2- 12	28-1- 12	37
38	12-2- 24	17-2- 49	22-2- 19	28-1- 24	38
39	12-2- 27	18-1- 1	22-2- 20	28-1- 25	39
40	12-2- 45	18-1- 11	22-2- 25	28-1- 34	40

PAGE 2 OF ORIGINAL SAMPLE OF 320 UNITS

1	28-1- 36	33-1- 54	39-1- 31	44-2- 18	1
2	28-1- 56	33-2- 20	39-1- 50	44-2- 20	2
3	28-2- 24	33-2- 47	39-2- 37	44-2- 29	3
4	28-2- 37	34-1- 23	39-2- 49	44-2- 49	4
5	28-2- 59	34-2- 5	40-1- 17	44-2- 55	5
6	29-1- 18	34-2- 11	40-2- 13	45-1- 32	6
7	29-1- 35	34-2- 24	40-2- 19	45-1- 33	7
8	29-2- 7	34-2- 32	40-2- 25	45-1- 35	8
9	29-2- 19	34-2- 43	40-2- 28	45-1- 48	9
10	29-2- 21	34-2- 45	41-1- 18	45-1- 55	10
11	29-2- 47	34-2- 54	41-1- 56	45-1- 60	11
12	29-2- 58	34-2- 56	41-2- 20	45-2- 4	12
13	30-1- 7	35-1- 24	41-2- 27	45-2- 59	13
14	30-1- 23	35-1- 57	41-2- 37	46-1- 8	14
15	30-1- 29	35-2- 28	41-2- 46	46-1- 21	15
16	30-2- 2	35-2- 38	41-2- 51	46-1- 27	16
17	30-2- 23	35-2- 48	41-2- 55	46-1- 52	17
18	30-2- 33	36-1- 2	41-2- 56	46-2- 9	18
19	30-2- 41	36-1- 9	42-1- 1	47-1- 1	19
20	31-1- 1	36-1- 30	42-1- 22	47-1- 37	20
21	31-1- 44	36-1- 49	42-1- 29	47-1- 53	21
22	31-1- 48	36-1- 51	42-2- 7	47-2- 3	22
23	31-1- 52	36-2- 13	42-2- 23	47-2- 46	23
24	31-1- 56	36-2- 27	42-2- 29	48-1- 14	24
25	31-2- 13	36-2- 43	42-2- 46	48-1- 31	25
26	31-2- 60	36-2- 59	42-2- 53	48-1- 33	26
27	32-1- 3	37-1- 27	43-1- 5	48-1- 45	27
28	32-1- 19	37-1- 59	43-1- 7	48-1- 52	28
29	32-1- 22	37-2- 12	43-1- 20	48-1- 53	29
30	32-1- 24	37-2- 31	43-1- 30	48-2- 7	30
31	32-1- 47	37-2- 41	43-1- 32	48-2- 26	31
32	32-2- 3	37-2- 47	43-2- 14	48-2- 29	32
33	32-2- 11	37-2- 52	43-2- 32	49-1- 11	33
34	32-2- 24	37-2- 53	43-2- 35	49-1- 31	34
35	32-2- 33	38-1- 34	43-2- 41	49-1- 44	35
36	32-2- 36	38-1- 48	43-2- 44	49-1- 47	36
37	32-2- 41	38-2- 24	43-2- 55	49-1- 52	37
38	33-1- 5	38-2- 54	44-1- 33	49-1- 58	38
39	33-1- 28	38-2- 56	44-1- 44	49-2- 48	39
40	33-1- 51	39-1- 9	44-2- 15	49-2- 55	40

PAGE 3 FORTY UNIT SAMPLES

	1	2	3	4	
1	6-2- 60	7-1- 43	6-2- 15	6-1- 60	1
2	9-2- 32	8-1- 27	6-2- 49	7-1- 28	2
3	10-2- 55	9-1- 35	7-1- 38	7-1- 44	3
4	12-1- 37	9-1- 52	8-2- 60	8-1- 9	4
5	14-1- 43	9-2- 38	13-2- 14	9-1- 11	5
6	14-2- 5	10-1- 60	15-1- 28	9-2- 17	6
7	15-1- 57	13-2- 8	16-2- 25	11-2- 31	7
8	16-2- 23	15-1- 42	17-1- 27	12-2- 5	8
9	17-1- 46	16-2- 4	17-1- 28	13-2- 57	9
10	18-1- 9	17-1- 45	18-2- 2	14-2- 36	10
11	20-1- 59	18-2- 21	18-2- 19	16-2- 60	11
12	24-2- 48	19-1- 21	18-2- 60	17-1- 30	12
13	25-2- 59	19-2- 33	20-2- 47	17-1- 56	13
14	26-1- 1	20-2- 5	21-2- 58	17-2- 7	14
15	27-2- 41	20-2- 37	23-1- 4	18-1- 50	15
16	29-1- 56	20-2- 43	23-2- 51	21-1- 4	16
17	29-2- 37	24-1- 45	25-1- 60	21-1- 48	17
18	30-1- 20	24-2- 12	26-2- 14	21-2- 33	18
19	31-1- 6	24-2- 29	26-2- 15	22-2- 56	19
20	32-1- 54	25-1- 51	28-1- 5	23-2- 57	20
21	34-2- 21	27-1- 20	28-1- 30	28-2- 25	21
22	35-1- 29	28-1- 46	28-2- 35	32-2- 4	22
23	35-1- 45	33-2- 33	29-2- 42	32-2- 14	23
24	36-1- 3	34-1- 8	30-2- 20	33-1- 47	24
25	36-1- 36	34-1- 21	30-2- 26	36-1- 48	25
26	38-2- 5	34-1- 43	31-1- 42	37-1- 5	26
27	39-1- 3	36-1- 28	31-2- 58	37-1- 49	27
28	39-2- 6	36-1- 57	32-2- 58	37-2- 3	28
29	40-1- 31	36-2- 7	35-2- 22	38-1- 35	29
30	40-2- 1	39-1- 55	37-2- 44	38-2- 29	30
31	41-2- 35	41-1- 37	38-2- 58	38-2- 42	31
32	42-1- 44	44-1- 16	39-1- 56	40-2- 8	32
33	43-1- 29	45-1- 21	39-2- 40	41-2- 42	33
34	46-2- 47	45-2- 60	40-1- 50	42-1- 39	34
35	48-1- 38	46-2- 53	42-1- 55	42-2- 48	35
36	48-2- 46	47-1- 4	42-2- 56	43-1- 27	36
37	48-2- 52	47-2- 26	44-1- 28	44-1- 30	37
38	49-1- 39	48-1- 1	47-1- 58	45-1- 25	38
39	49-2- 5	48-1- 24	48-2- 31	47-1- 45	39
40	49-2- 50	49-2- 6	49-1- 48	48-2- 15	40

	5	6	7	8	
1	6-1- 34	6-2- 34	7-1- 24	6-2- 54	1
2	8-1- 53	6-2- 38	7-1- 45	8-1- 56	2
3	11-1- 21	7-1- 8	8-1- 17	9-1- 21	3
4	12-2- 49	8-1- 23	8-2- 45	9-2- 34	4
5	14-1- 53	8-2- 14	10-1- 7	10-1- 39	5
6	14-2- 32	10-1- 28	11-1- 24	12-2- 21	6
7	15-2- 26	13-1- 54	13-1- 13	15-1- 50	7
8	15-2- 60	14-2- 25	13-1- 16	15-2- 57	8
9	16-2- 44	15-1- 25	13-2- 59	17-2- 55	9
10	19-1- 27	16-1- 59	15-1- 10	18-2- 8	10
11	19-2- 27	17-2- 8	15-2- 19	20-1- 57	11
12	20-1- 13	19-1- 32	19-2- 6	20-2- 54	12
13	20-2- 23	19-2- 25	20-1- 25	20-2- 56	13
14	21-1- 10	23-1- 50	22-1- 39	21-1- 21	14
15	21-1- 13	24-1- 50	22-2- 16	22-1- 3	15
16	22-1- 58	25-1- 15	23-2- 18	25-1- 17	16
17	22-2- 24	26-1- 43	23-2- 20	26-1- 35	17
18	26-1- 45	26-2- 6	24-2- 26	27-1- 31	18
19	28-1- 1	27-1- 32	26-2- 57	27-2- 43	19
20	28-2- 7	29-2- 48	27-1- 34	28-2- 12	20
21	31-1- 40	31-1- 26	27-1- 59	29-1- 16	21
22	32-1- 18	33-1- 12	29-2- 45	29-1- 34	22
23	32-2- 8	35-2- 27	30-2- 6	30-2- 44	23
24	35-1- 9	37-2- 16	30-2- 50	31-1- 25	24
25	35-1- 46	37-2- 60	32-1- 26	33-2- 17	25
26	35-2- 42	38-1- 24	33-2- 12	33-2- 25	26
27	36-1- 5	38-1- 55	34-1- 19	34-2- 17	27
28	37-1- 55	38-2- 1	36-2- 30	35-2- 25	28
29	37-2- 19	40-1- 18	36-2- 50	37-1- 6	29
30	38-1- 39	40-2- 48	38-2- 39	37-1- 8	30
31	38-2- 21	41-1- 23	38-2- 51	41-1- 19	31
32	38-2- 37	41-2- 54	39-1- 7	41-1- 54	32
33	40-1- 42	43-1- 60	40-2- 34	41-2- 36	33
34	40-1- 59	44-1- 26	40-2- 49	42-1- 3	34
35	42-2- 51	45-2- 12	44-1- 50	44-1- 11	35
36	43-1- 3	46-2- 40	44-2- 3	45-1- 12	36
37	45-2- 41	47-1- 13	45-1- 10	46-1- 20	37
38	46-2- 25	47-1- 60	47-1- 36	46-1- 22	38
39	46-2- 60	48-1- 32	48-2- 54	47-2- 37	39
40	47-2- 24	48-2- 42	49-1- 18	49-1- 22	40

PAGE 5 FORTY UNIT SAMPLES

	9	10	11	12	
1	6-1- 36	7-2- 10	6-2- 23	9-2- 36	1
2	7-1- 57	8-1- 15	7-1- 49	9-2- 60	2
3	9-1- 57	9-1- 17	7-2- 36	11-2- 15	3
4	11-1- 43	9-1- 55	10-2- 15	12-2- 31	4
5	13-1- 2	9-2- 11	11-1- 1	14-1- 55	5
6	13-1- 25	11-2- 4	11-1- 52	14-2- 23	6
7	13-1- 50	14-2- 40	13-1- 1	15-2- 35	7
8	13-2- 37	15-1- 56	13-1- 20	17-2- 37	8
9	15-2- 8	16-1- 1	15-1- 45	18-2- 4	9
10	17-1- 20	16-1- 37	15-2- 33	18-2- 14	10
11	17-1- 23	16-1- 49	16-1- 48	18-2- 30	11
12	18-2- 20	19-1- 15	19-2- 3	20-1- 54	12
13	20-2- 24	23-2- 34	19-2- 42	20-2- 41	13
14	23-1- 42	24-2- 44	20-1- 43	22-1- 43	14
15	23-2- 21	25-1- 10	21-1- 51	22-2- 11	15
16	24-1- 40	27-2- 1	21-2- 56	22-2- 57	16
17	25-1- 2	28-2- 20	24-2- 16	23-2- 10	17
18	26-1- 37	29-2- 46	25-1- 16	24-2- 13	18
19	28-1- 19	30-2- 52	25-1- 42	24-2- 36	19
20	28-1- 51	31-2- 7	25-2- 22	26-1- 27	20
21	29-1- 21	32-1- 10	25-2- 31	26-2- 29	21
22	29-1- 51	34-2- 38	29-2- 33	27-2- 39	22
23	30-1- 60	35-1- 60	34-2- 13	28-2- 33	23
24	30-2- 11	35-2- 53	34-2- 23	29-1- 14	24
25	31-1- 24	37-1- 32	34-2- 27	30-1- 50	25
26	32-1- 42	39-2- 13	37-1- 53	34-1- 32	26
27	33-2- 4	43-1- 43	38-1- 14	34-1- 57	27
28	33-2- 9	44-1- 15	38-1- 51	35-1- 7	28
29	38-2- 22	44-1- 34	38-2- 46	35-2- 9	29
30	39-1- 13	44-2- 4	41-2- 41	35-2- 14	30
31	42-1- 23	45-1- 3	42-1- 7	38-1- 12	31
32	42-2- 33	45-1- 18	43-2- 5	39-2- 9	32
33	42-2- 49	45-1- 49	43-2- 26	40-2- 31	33
34	43-1- 57	46-2- 20	44-2- 2	42-1- 45	34
35	43-2- 47	46-2- 51	44-2- 47	43-1- 38	35
36	46-1- 51	46-2- 59	44-2- 56	45-1- 17	36
37	47-1- 5	48-1- 23	45-2- 31	45-2- 28	37
38	48-1- 19	48-2- 9	45-2- 57	47-2- 15	38
39	48-1- 30	48-2- 33	46-2- 43	48-1- 13	39
40	49-1- 36	49-1- 30	48-1- 2	49-2- 52	40

PAGE 6 FORTY UNIT SAMPLES

	13	14	15	16	
1	6-1- 31	7-1- 7	6-2- 4	8-2- 16	1
2	8-1- 28	7-1- 55	7-1- 6	23-2- 28	2
3	8-1- 31	9-1- 28	7-1- 14	24-1- 57	3
4	10-1- 16	9-2- 14	8-2- 13	26-1- 23	4
5	10-1- 46	10-1- 21	9-2- 50	29-2- 17	5
6	10-2- 12	12-1- 6	12-1- 18	30-2- 8	6
7	10-2- 28	12-1- 27	14-1- 31	40-1- 58	7
8	14-1- 42	13-1- 48	14-2- 28	41-2- 5	8
9	14-2- 43	14-1- 24	16-1- 58	44-1- 3	9
10	14-2- 47	14-2- 42	17-1- 10	1023-7-511	10
11	16-1- 35	15-1- 32	17-1- 29	1023-7-511	11
12	17-2- 14	16-1- 53	17-2- 40	1023-7-511	12
13	20-1- 24	17-2- 60	18-1- 52	1023-7-511	13
14	20-1- 39	19-1- 1	19-1- 50	1023-7-511	14
15	21-2- 15	21-1- 9	20-1- 40	1023-7-511	15
16	22-2- 50	21-2- 22	22-1- 11	1023-7-511	16
17	22-2- 59	23-1- 1	22-1- 46	1023-7-511	17
18	23-2- 33	23-1- 44	22-2- 39	1023-7-511	18
19	24-2- 32	23-2- 29	26-1- 49	1023-7-511	19
20	24-2- 39	24-1- 32	28-2- 13	1023-7-511	20
21	26-2- 41	24-2- 25	32-1- 7	1023-7-511	21
22	26-2- 51	26-1- 52	34-1- 39	1023-7-511	22
23	27-1- 19	26-2- 4	34-1- 50	1023-7-511	23
24	27-2- 52	26-2- 43	37-1- 31	1023-7-511	24
25	28-1- 29	27-2- 30	37-2- 36	1023-7-511	25
26	29-1- 8	28-2- 34	38-2- 9	1023-7-511	26
27	32-2- 32	30-1- 28	39-2- 19	1023-7-511	27
28	33-1- 44	31-1- 57	40-1- 24	1023-7-511	28
29	33-2- 11	32-2- 16	40-2- 9	1023-7-511	29
30	35-2- 2	33-1- 41	41-2- 9	1023-7-511	30
31	38-2- 7	35-1- 43	42-2- 20	1023-7-511	31
32	40-2- 37	35-2- 36	42-2- 41	1023-7-511	32
33	41-2- 13	36-1- 44	46-1- 19	1023-7-511	33
34	43-1- 58	36-1- 58	46-1- 29	1023-7-511	34
35	44-1- 2	36-2- 42	46-1- 58	1023-7-511	35
36	44-2- 10	38-2- 57	47-1- 21	1023-7-511	36
37	45-2- 52	39-1- 49	47-2- 6	1023-7-511	37
38	46-1- 41	43-1- 1	48-1- 3	1023-7-511	38
39	47-1- 29	46-2- 32	48-1- 29	1023-7-511	39
40	49-1- 25	48-1- 17	49-1- 16	1023-7-511	40

NBS-114A (REV. 7-73)

U.S. DEPT. OF COMM. BIBLIOGRAPHIC DATA SHEET	1. PUBLICATION OR REPORT NO. NBS TN-921	2. Gov't Accession No.	3. Recipient's Accession No.
4. TITLE AND SUBTITLE Survey Manual for Estimating the Incidence of Lead Paint in Housing			5. Publication Date September 1976 6. Performing Organization Code
7. AUTHOR(S) William G. Hall and Lillian T. Slovic			8. Performing Organ. Report No.
9. PERFORMING ORGANIZATION NAME AND ADDRESS NATIONAL BUREAU OF STANDARDS DEPARTMENT OF COMMERCE WASHINGTON, D.C. 20234			10. Project/Task/Work Unit No. 11. Contract/Grant No.
12. Sponsoring Organization Name and Complete Address *(Street, City, State, ZIP)* Division of Energy, Building Technology and Standards Office of Policy Development and Research The Department of Housing and Urban Development Washington, D.C. 20410			13. Type of Report & Period Covered 14. Sponsoring Agency Code
15. SUPPLEMENTARY NOTES			

16. ABSTRACT *(A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here.)*

This manual is intended as a guide for municipal managers in performing a survey to determine the prevalence of lead based paint in their community's dwelling units. There are four parts to the Manual, each is intended for a different audience.

Part I discusses the preliminary considerations for a survey. It is intended for the department head or executive who will initiate plans for the survey. It presents a managerial overview of the processes, the cost determinants, criteria for the establishment of objectives and the resources required.

Part II is intended for the survey manager and the inspector supervisors. It contains more detailed information on the planning, staffing, training, and execution of the data collection phase of the survey.

Part III is for the use of the person responsible for the control and management of the data collected and for the analysis of these data.

The Appendices contain quite detailed information about procedures we have used in previous surveys. These may be used as they are described or may be modified or adapted to meet specific objectives.

17. KEY WORDS *(six to twelve entries; alphabetical order; capitalize only the first letter of the first key word unless a proper name; separated by semicolons)*

Lead paint; lead paint detection; lead paint programs; lead poisoning; portable x-ray fluorescence; random sampling; x-ray fluorescence.

18. AVAILABILITY	19. SECURITY CLASS (THIS REPORT)	21. NO. OF PAGES
[XX] Unlimited [ ] For Official Distribution. Do Not Release to NTIS [X] Order From Sup. of Doc., U.S. Government Printing Office Washington, D.C. 20402, SD Cat. No. C13. 46:921 [ ] Order From National Technical Information Service (NTIS) Springfield, Virginia 22151	UNCLASSIFIED	111
	20. SECURITY CLASS (THIS PAGE) UNCLASSIFIED	22. Price $2.20

USCOMM-DC 29042-P74

☆U.S. GOVERNMENT PRINTING OFFICE: 1976 210-801/375 1-3

NBS TECHNICAL PUBLICATIONS

PERIODICALS

JOURNAL OF RESEARCH reports National Bureau of Standards research and development in physics, mathematics, and chemistry. It is published in two sections, available separately:

• **Physics and Chemistry (Section A)**

Papers of interest primarily to scientists working in these fields. This section covers a broad range of physical and chemical research, with major emphasis on standards of physical measurement, fundamental constants, and properties of matter. Issued six times a year. Annual subscription: Domestic, $17.00; Foreign, $21.25.

• **Mathematical Sciences (Section B)**

Studies and compilations designed mainly for the mathematician and theoretical physicist. Topics in mathematical statistics, theory of experiment design, numerical analysis, theoretical physics and chemistry, logical design and programming of computers and computer systems. Short numerical tables. Issued quarterly. Annual subscription: Domestic, $9.00; Foreign, $11.25.

DIMENSIONS/NBS (formerly Technical News Bulletin)—This monthly magazine is published to inform scientists, engineers, businessmen, industry, teachers, students, and consumers of the latest advances in science and technology, with primary emphasis on the work at NBS. The magazine highlights and reviews such issues as energy research, fire protection, building technology, metric conversion, pollution abatement, health and safety, and consumer product performance. In addition, it reports the results of Bureau programs in measurement standards and techniques, properties of matter and materials, engineering standards and services, instrumentation, and automatic data processing.

Annual subscription: Domestic, $9.45; Foreign, $11.85.

NONPERIODICALS

Monographs—Major contributions to the technical literature on various subjects related to the Bureau's scientific and technical activities.

Handbooks—Recommended codes of engineering and industrial practice (including safety codes) developed in cooperation with interested industries, professional organizations, and regulatory bodies.

Special Publications—Include proceedings of conferences sponsored by NBS, NBS annual reports, and other special publications appropriate to this grouping such as wall charts, pocket cards, and bibliographies.

Applied Mathematics Series—Mathematical tables, manuals, and studies of special interest to physicists, engineers, chemists, biologists, mathematicians, computer programmers, and others engaged in scientific and technical work.

National Standard Reference Data Series—Provides quantitative data on the physical and chemical properties of materials, compiled from the world's literature and critically evaluated. Developed under a world-wide program coordinated by NBS. Program under authority of National Standard Data Act (Public Law 90-396).

NOTE: At present the principal publication outlet for these data is the Journal of Physical and Chemical Reference Data (JPCRD) published quarterly for NBS by the American Chemical Society (ACS) and the American Institute of Physics (AIP). Subscriptions, reprints, and supplements available from ACS, 1155 Sixteenth St. N. W., Wash. D. C. 20056.

Building Science Series—Disseminates technical information developed at the Bureau on building materials, components, systems, and whole structures. The series presents research results, test methods, and performance criteria related to the structural and environmental functions and the durability and safety characteristics of building elements and systems.

Technical Notes—Studies or reports which are complete in themselves but restrictive in their treatment of a subject. Analogous to monographs but not so comprehensive in scope or definitive in treatment of the subject area. Often serve as a vehicle for final reports of work performed at NBS under the sponsorship of other government agencies.

Voluntary Product Standards—Developed under procedures published by the Department of Commerce in Part 10, Title 15, of the Code of Federal Regulations. The purpose of the standards is to establish nationally recognized requirements for products, and to provide all concerned interests with a basis for common understanding of the characteristics of the products. NBS administers this program as a supplement to the activities of the private sector standardizing organizations.

Federal Information Processing Standards Publications (FIPS PUBS)—Publications in this series collectively constitute the Federal Information Processing Standards Register. Register serves as the official source of information in the Federal Government regarding standards issued by NBS pursuant to the Federal Property and Administrative Services Act of 1949 as amended, Public Law 89-306 (79 Stat. 1127), and as implemented by Executive Order 11717 (38 FR 12315, dated May 11, 1973) and Part 6 of Title 15 CFR (Code of Federal Regulations).

Consumer Information Series—Practical information, based on NBS research and experience, covering areas of interest to the consumer. Easily understandable language and illustrations provide useful background knowledge for shopping in today's technological marketplace.

NBS Interagency Reports (NBSIR)—A special series of interim or final reports on work performed by NBS for outside sponsors (both government and non-government). In general, initial distribution is handled by the sponsor; public distribution is by the National Technical Information Service (Springfield, Va. 22161) in paper copy or microfiche form.

Order NBS publications (except NBSIR's and Bibliographic Subscription Services) from: Superintendent of Documents, Government Printing Office, Washington, D.C. 20402.

BIBLIOGRAPHIC SUBSCRIPTION SERVICES

The following current-awareness and literature-survey bibliographies are issued periodically by the Bureau:
Cryogenic Data Center Current Awareness Service

A literature survey issued biweekly. Annual subscription: Domestic, $20.00; foreign, $25.00.

Liquefied Natural Gas. A literature survey issued quarterly. Annual subscription: $20.00.

Superconducting Devices and Materials. A literature survey issued quarterly. Annual subscription: $20.00. Send subscription orders and remittances for the preceding bibliographic services to National Bureau of Standards, Cryogenic Data Center (275.02) Boulder, Colorado 80302.

U.S. DEPARTMENT OF COMMERCE
National Bureau of Standards
Washington, D.C. 20234

OFFICIAL BUSINESS

Penalty for Private Use, $300

POSTAGE AND FEES PAID
U.S. DEPARTMENT OF COMMERCE
COM-215

U.S.MAIL

SPECIAL FOURTH-CLASS RATE
BOOK

AMERICAN REVOLUTION BICENTENNIAL 1776-1976

75 YEARS NBS 1901-1976