National Jewish Population Study

METHODOLOGY



Council of Jewish Federations and Welfare Funds
315 Park Ave. South, New York, N. Y. 10010

THE SAMPLE DESIGN OF THE NATIONAL JEWISH POPULATION SURVEY by DR. BERNARD LAZERWITZ

The sample design of the National Jewish Population Survey (NJPS) had as its necessary goal the selection of a probability sample from the entire Jewish population of the United States. Such a statistical task was made quite complicated by the fact that this meant a sample of an uncommon element of the United States population, Jews who compose around 3 percent of the nation's population. Furthermore, apart from New York City and its surrounding counties, the Jewish proportion of local metropolitan populations seldom reaches as high as 8 percent.

To have relied upon straight area probability sampling would have meant screening many thousands of housing units for Jewish occupants at a prohibitive cost in interviewer time and dollars. Fortunately, a sizable proportion of the addresses of the nation's Jews are known to their local Jewish Federations and are available on lists from these organizations. Since this survey had been commissioned by the Council of Jewish Federations and Welfare Funds, there was official access to, and a fair degree of cooperation from, the various local Jewish Federations. The local Federations organized their lists of local Jewish addresses in some coherent manner and gave them to us, publicly supported the survey, and aided in the reduction of refusals to be interviewed.

There does not exist any one list with a large proportion of the Jews in the New York metropolitan area. Fortunately, there Jews do form a considerable percentage of the total population which enabled sampling them by standard area probability methods.

In summary, then, the basic sampling strategy consisted of using lists of Jews living outside the New York City metropolitan area which were supplemented by an area probability sample of Jews not on such local lists together with an area probability sample of Jews in the New York metropolitan area.

Determining the Overall Sampling Fraction

A basic decision in designing a probability sample is that of the size of the overall sampling fraction. After taking into very careful consideration the estimated cost factors of the survey and the available budget, client and survey researcher decide upon how many interviews can be afforded and their lengths. Given these parameters, it usually is a simple matter to divide the estimated size of the population to be surveyed by the desired number of interviews, with allowances for nonresponse, and obtain all overall survey sampling fraction.

However, the national Jewish population survey had as one of its major goals the determination of the size of the nation's Jewish population. We sought from this survey, in other words, that very piece of information required to design the survey creating a sort of circular situation with the connecting link missing. There were,

however, estimates of the nation's Jewish population. First of all, there has been one survey of United States religious groups conducted by the Bureau of the Census in March, 1957. Data from that survey gave an estimated number of 1,650,030 Jewish households, but this survey was clearly outdated when NJPS was being designed. Nor was nearly enough information available on the growth rate of Jewish housing units to update this 1957 figure. After studying a number of estimates of the United States Jewish population obtained by a variety of demographic methods and applying the collective judgment of the survey design committee, it was agreed to work with an estimate of 5,700,000 U.S. Jews living in private housing units with an average of 3 Jews per housing unit, and 1,900,000 Jewish housing units.

The survey design committee also collectively agreed to work with an estimation that 83 percent of the survey's Jewish housing units would agree to participate in the survey ~~ a figure acceptable as good interviewing performance in standard survey work. Finally, we set a target of an effective 12,000 Jewish interviews. This resulted in:

 $\frac{(1,900,000)}{12,000}$ = 131.4 for an overall sampling fraction which was rounded to 131. This means that every Jewish housing unit in the country had one chance in 131 of being sampled for the survey.

The phrase, "effective jewish interviews", is used above because we did not actually seek that many interviews. Instead, as detailed below, we decided to apply the rate of 1 in 131 only to lists of

Jewish housing units obtained from local Jewish federations and area sampling done in districts with a high proportion of Jews.

In other areas, much sampling was done with considerably smaller sampling fractions. This design complication reduced the number of expected Jewish interviews to 8,000.

Primary Sampling Units

We decided that the costs of creating local field forces were such as to generally prohibit working with a local sample size of less than an expected 100 Jewish housing units. With an overall sampling fraction of 131, this required that each sampled Jewish community contain about 13,100 Jewish housing units (40,000 Jews) or represent that number as part of a stratification system.

Mr. Alvin Chenkin, of the staff of the Council of Jewish Federations and Welfare Funds, for years has been collecting Jewish population estimates for all the nation's Jewish communities. Hence, it was possible to use his information to determine quickly those Jewish communities large enough to definitely be part of this survey.

Then, geographic boundaries had to be given to Jewish communities. For this purpose it was clear that a county or group of contiguous counties established the best community limits and should form our initial primary sampling units (psu).

Stratification

Some system of stratification was required to enable a sub-selection among the smaller Jewish communities so that the sub-selected communities could represent both themselves and other such Jewish communities. Accordingly, the following stratification criteria were adopted:

- The entire United States was divided into primary sampling units each of which consisted of a single county or of a group of contiguous counties.
- As far as possible, each such primary sampling units contained a known Jewish community.
- 3) All primary sampling units without any known Jewish communities were grouped into one stratum called the "no known Jewish population" stratum.
- 4) Primary sampling units were assigned Jewish population estimates based upon the information furnished by Mr. Chenkin.
- 5) Those primary sampling units with 40,000 or more Jews automatically came into the survey. A few primary sampling units with somewhat less than this number of Jews were also brought into this survey because of the importance of their local Jewish communities.
- 6) The other primary sampling units with less than 40,000 Jews were grouped together until strata of approximately 40,000 Jews were obtained.

- 7) As far as possible, such strata were formed of primary sampling units lying within the same state or a contiguous group of states.
- 8) Also, as far as possible, the strata contained local

 Jewish communities with approximately the same estimated

 Jewish populations.

From each strata containing two or more primary sampling units, one such unit was selected with probability proportionate to the size of its Jewish population.

As the Jewish population of primary sampling units decreased, it was sometimes necessary to form primary sampling units of groups of counties that were scattered about a state, or adjacent group of states. This was done to avoid going into any local Jewish community and interviewing too large a proportion of its population. It was thought that this would be undesirable from the viewpoints of response rates, anonymity of respondents, and statistical heterogeneity.

The counties forming the so-called "no known Jewish population" stratum were sampled with equal probabilities. This was feasible because of their supposed lack of Jews.

When this complex and time-consuming stratification work was done and sample primary sampling units selected, we obtained 18 primary sampling units that represented just themselves; 20 primary sampling units representing themselves and other units, and 14 counties representing the "no known Jewish population" stratum.

The specific locations and compositions of these NJPS sample points are:

Principal Jewish Community		Counties of Sample Units	
1)	Allentown, Pa.	Lehigh and Northampton (non-self-representing)	
2)	Atlanta, Ga.	Fulton, DeKalb, Cobb, Rockdale, Walton, Gwinnett and Clayton (non-self-representing)	
3)	Atlantic City, N.J.	Atlantic (non-self-representing)	
4)	Baltimore, Md.	City of Baltimore, Baltimore County, Anne Arundel, Carroll, Hartford, Howard, Cecil, Kent and Queen Annes (represents itself)	
5)	Boston, Mass.	Suffolk, Middlesex, Norfolk and Essex (represents itself)	
6)	Buffalo, N.Y.	Erie, Nîagara and Wyoming (non-self-representing)	
7)	Bakersfield and Central California	San Benito, Kings, Napa, San Lois Obispo, Santa Cruz, Merced, Solano, Sonoma, Stanislaus, Tulare, Santa Barbara, Monterey and Kern (non-self-representing)	

Principal Jewish Community

Counties of Sample Units

8) Charleston, W. Va. and Williamsburg, Va.

Kanawha, Boone, Clay, Fayette and Nicholas, W. Va. Goucester, Mathews, James City and the independent city of Williamsburg in Va. (non-self-representing)

9) Chicago, Ili.

Cook, DuPage, Lake, Kane, Will, McHenry, DeKalb, Boone and Kendall in Ill.
Porter and Lake in Ind.
(represents itself)

10) Cincinnati, O.

Hamilton, Butler, Ross, Clermont, Warren, Clinton, Brown and Highland in O. Campbell, Kenton and Boone in Ky. Dearborn in Ind. (represents itself)

11) Cleveland, 0.

Cuyahoga, Ashtabula, Geauge, Lake, Lorain and Medina (represents itself)

12) Denver, Colo.

Denver, Adams, Arapahoe, Bolder and Jefferson (non-self-representing)

13) Detroit, Mich.

Wayne, Oakland and Macomb (represents itself)

14) Erie, Pa. and Elmira, N.Y.

Crawford, Erie and Bradford in Pa.
Chemung in N.Y.
(non-self-representing)

15) Hartford, Ct.

Hartford
(non-self-representing)

16) Houston, Tex.

Harris, Brazoria, Fort Bend, Liberty, Montgomery, Galveston, Chambers, Hardin, Orange, Waller and Jefferson (non-self-representing)

17) Knoxville, Tenn.

2 counties in 111.
6 counties in Ky.
14 counties in Tenn.
21 counties in Miss.
5 counties in Ark.
1 county in La.
(non-self-representing)

Principal Jewish Community

Counties of Sample Units

18) Los Angeles, Cal.

Los Angeles, Orange, San Bernardino, Ventura and Riverside (represents itself)

19) Louisville, Ky.

Jefferson, Hardin, Nelson, Shelby, Bullitt, Oldham and Spencer in Ky. Floyd and Clark in Ind. (non-self-representing)

20) Miami, Fla.

Dade and Broward (represents itself)

21) Milwaukee, Wisc.

Milwaukee, Waukesha, Fond du Lac, Dodge, Washington, Ozaukee and Green Lake (represents itself)

22) Minneapolis - St. Paul, Minn.

Hennepin, Ramsey, Anoka, Dakota, Washington, Wright, McLeod, Scott, Carver and Meeker in Minn.
St. Croix, Dunne, Pierce, Buffalo and Pepin in Wisc. (represents itself)

23) Asbury Park, N.J.

Monmouth (non-self-representing)

24) Nashville, Tenn.

22 counties
(non-self-representing)

25) Newark, N.J.

Essex, Bergen, Union and Hudson (represents itself)

26) New York City and surroundings

Boroughs of Bronx, Manhattan, Brooklyn, Queens, Staten Island, counties of Nassau, Suffolk, Westchester and Rockland (represents itself)

27) Philadelphia, Pa.

Philadelphia, Montgomery, Delaware, Chester and Bucks in Pa. Camden, Burlington and Gloucester in N.J. (represents itself)

Princ	ipal	Jewish	Commur	iity

Counties of Sample Units

28) Pittsburgh, Pa.

Allegheny, Westmoreland, Beaver, Armstrong and Indiana (represents itself)

29) Poughkeepsie, N.Y.

Dutchess
(non-self-representing)

30) Providence, R.I.

All of R.I., New London in Conn., Barnstable, Dukes, Nantucket, Plymouth and Bristol in Mass. (represents itself)

31) St. Louis, Mo.

St. Louis City, St. Louis County, Jefferson, St. Charles, Franklin and Warren in Mo. St. Clair and Madison in Ill. (represents itself)

32) San Francisco, Cal.

San Francisco, Alameda, Contra Costa, Merin and San Mateo (represents itself)

33) Seattle, Wash.

- King, Pierce, Thruston, Kipsap, Snohomish and Mason (non-self-representing)
- 34) South Bend Ft. Wayne -Evansville, Ind.
- 14 counties in Ind.,
 1 county in Ky.
 (non-self-representing)

35) Stamford, Conn.

Towns of Stamford, Greenwich and Darien of Fairfield County (non-self-representing)

36) Tucson, Ariz.
and Las Vegas, Nev.

Pima in Ariz. Clark in Nev. (non-self-representing)

37) Utica and Rome, N.Y.

Cayuga, Hamilton, Herkimer, Oneida and Otsego (non-self-representing)

38) Washington, D.C.

Washington, D.C.
Montgomery, Prince Georges
and Calvert in Md.
Arlington, Fairfax, Loudon and
Prince Williams in Va.
The independent cities of
Alexandria, Fairfax and Falls
Church in Va.
(represents itself)

Principal Jewish Community

39) The No-Known-Jewish Population Stratum

Counties of Sample Units

County	State		
Bent	Colo.		
Furnas	Neb.		
Gillespie	Tex.		
Gr a dy	Ga.		
Hodgeman	K a n.		
La Crosse	Wisc.		
Laclede	Mo.		
Lawrence	So. Dak.		
Lyon	Minn.		
McCurtain	Okla.		
Moore	N.C.		
Rich	Ut a h		
Schuyler	111.		
Scott	Κy.		
(non-self-representing)			

Internal Stratification

Having selected sample primary sampling units, the actual sample housing units were selected in a two-step process. First, for each Jewish community in the sampling units, the various Jewish federations were contacted and their lists of Jewish housing units were obtained. Then, an area probability sample was selected of those housing units not on federation lists. The housing units of the area sample were screened for Jewish occupants who were subsequently approached for interviewing.

The address lists obtained from federations were usually arranged alphabetically within postal zones. Where feasible, lists were reorganized alphabetically by street names and -- within streets -- by address number. This latter type was more readily handled clerically for purposes of integrating list and area samples.

After the lists were organized, each address entry was examined in order to remove multiple listings for the same family at the same address, to remove non-residential entries, and any addresses outside the primary sampling unit boundaries. For sampling purposes, these lists were conceptualized as composed of housing units with high probabilities of containing Jewish families. It was not assumed that they still contained Jewish occupants, nor were they regarded as lists of Jewish persons. In this way, we could select a sample of housing units from the lists and, subsequently, integrate area samples with lists. Such integration was accomplished by eliminating from the area samples those housing units whose addresses appeared on the lists from the federations. Hence, housing units had one, and only one, way of entering our sample either through the list, or if not on the list, through the area sample, but never from both list and area samples.

After the lists were made ready, the proper within-primary sampling unit rates were applied and the "high probability of being Jewish" housing unit list samples were selected. Given below is the sampling equation used for this phase.

The territory of each sample primary sampling unit was divided into areas thought to have high numbers, moderate numbers, and low numbers of Jews. Such divisions were made by classifying postal zones, and sometimes census tracts, according to the proportion of their telephone addresses that contained occupants with "distinctive Jewish names".

Such names, and the proportions they are thought to bear to total

Jewish population, have been developed from research done in Los Angeles by Professor Massarik. After such three strata systems had been prepared by a Los Angeles research staff (working under the general guidance of Professor Massarik), they were critically examined by the sampling staff (under the direction of Professor Lazarwitz) and any needed adjustments in unit classification or the boundaries of these three "Jewish" strata were then made. Sometimes, depending on Jewish population numbers and concentrations, a sampling unit received just one or two such internal strata.

This process, then, produced both an internal area sampling stratification system and the sub-units, usually postal zones, forming said strata. These, in their turn, were sub-selected. With the internal stratification system for the area sample now presented, it's time to study the actual design sampling equations.

Sampling Equations

Ten sampling equations govern the sampling plan for this survey. They are:

A. For list samples

$$\frac{1}{131} = \frac{\text{psu J. pop.}}{\text{stratum J. pop.}} \chi \frac{\text{stratum J. pop.}}{\text{(psu J. pop.)}} (131)$$

Here 131 is the basic survey sampling rate,

psu J. pop.
stratum J. pop.

is the probability of selection for a sample primary sampling unit which consists of its estimated Jewish population divided by the estimated population of its sampling stratum. For self-representing psu's this factor is 1.

stretum J. pop. is the inverse of the above term together with (psu J. pop.) (131)

131 In its denominator. This is the value used to actually select list housing unit addresses.

Together, the two right-hand terms of this equation balance out to 131. For self-representing psu's as the equation reveals, the basic list selection rate was 131.

- B. The high numbers area sample stratum:

 Here, there is one basic sampling equation and two additional ones.
 - 1. The basic equation applied to housing units whose occupants were judged to have a substantial probability of being Jewish is:

- here. a) the psu term is the same as for the list sample;
 - sampling units within this area stratum. Its measure of size is the average estimated Jewish population assigned to the postal zone, or census tracts forming this high numbers stratum by the Los Angeles operation which established both secondary units and strata. In those instances in which any particular high numbers stratum was relatively condensed in area, there was no secondary unit sub-selection. Instead, all the territory of the high numbers stratum was selected with certainty. For those cases, this term has a value of 1.
 - c) The <u>block</u> is the probability of selection of blocks within secondary units.
 - d) The within-block housing unit selection rate for this stratum 1/1 is 2. Taking every other housing unit into the sample was a good way of guarding against the disadvantages of getting too large a cluster of sample housing units in blocks with large numbers of Jews.

Still another device was employed to prevent the appearance of too many sample housing units from highly Jewish populated blocks. Use was made of local census block statistical data, and knowledge of densely build-up Jewish areas to assign housing unit counts to blocks. Then, blocks were given one "measure of size" for each group of 50 housing units.

Then, these "measures of size" factors were entered into the numerator

within-block term. In this manner, such blocks were sampled with probability proportional to their measure, of size, but within-block yields were kept to an expected 25 sample wusing units for each such selected measure of size.

chaily, a two-phase sampling device was employed on all area sample blocks to reduce as much as possible the sizable field work costs of contacting very many housing units merely to find that a considerable number of them had no Jewish occupants. Instead, all sample blocks (or "measure of size" groups) were visited once and as much information as possible on sample occupants were obtained without actually talking to them. These were items such as last names, ethnic characteristics of the block, presence or absence of a mezuzah on housing unit doors, etc.

Then, such block fisting information was sent to survey field head-quarters where qualified members of the field staff classified all such housing units into three groups: a) sizable propability of having Jewish occupants; b) don't know whether occupants are Jewish or not; c) sizable probability that occupants are not Jewish. This done, we brought into the sample all housing units with a sizable probability of having Jewish occupants. Hence, the final term of in the above sample equation for Jewish hu rating.

When housing units were classified as "don't know if Jewish", just 1 in 4 of them were brought into the sample. When housing units were classified as "probably not Jewish", just 1 in 10 of those were brought into the sample. Then, all housing units that came into the sample were contacted and, if their occupants turned out to be Jewish, they were interviewed.

These two additional <u>Jewish hu rating</u> factors give two additional sampling equations that are the same as the one given above for "probably Jewish" housing units apart from the final <u>Jewish</u> hu rating term. Then, they become:

2. High Numbers Stratum ' "Don't know if Jewish" Housing Units

$$\frac{1}{524} = (same equation as above) \frac{\text{Jewish hu rating}}{\binom{1}{4}}$$

3. High Numbers Stratum - "Probably Not Jewish" Housing Units

$$\frac{1}{1310} = (\text{same equation as above}) \left(\frac{1}{10}\right)$$

C. Moderate Numbers Area Sample Stratum

The moderate numbers area sample stratum was created in much the same way as was the high numbers stratum. Of course, since these stratification techniques were applied within sample psu's, the probability of psu selection doesn't change here. However, the remaining terms were adjusted as follows:

$$\frac{PSU}{t} = \frac{psu J. pop.}{stratum J. pop.}$$

$$\chi \qquad \frac{\text{Secondary}}{\text{(sec. unit J. pop.)(4)}}$$

$$\chi \qquad \frac{\text{(sec. unit J. pop.)(4)}}{\text{(sec. stratum J. pop.)(3)}}$$

Block

As we well know, most Jews live in a relatively small proportion of the housing units of sample primary sampling units. Hence, the number of sample housing units to be screened for relatively few Jews became quite large in this stratum. Therefore, to reduce considerable field work time and costs, this new design selected sample housing units at one-third of their rate in the high numbers stratum. The overall survey rate for this stratum is 1 in 393.

The secondary selection and block selection terms are adjusted so as to give a somewhat greater spread and coverage of secondary selection units at the price of fewer sample blocks per sample secondary unit. Typically, one secondary unit was selected for an estimated 5,000 Jewish housing units. Then all housing units on sample blocks were screened for Jewish respondents. Here, too, the screening was the first phase of the identical two-phase selection process just described for the high numbers stratum. This resulted in again assigning housing units to a "sizable probability of being Jewish" group; a "don't know

Jewish' group; and a "probably not Jewish" group with the same 1, 4, or 10 sub-selection rates of sample housing units. This gives two additional "moderate numbers" sampling equations with overall sampling fractions of 1572 for the "don't knows" and 3930 for the "not Jewish" hu groups.

The Low Numbers Area Sampling Stratum

This low numbers stratum covered still more people and territory with a lot fewer Jews than did the moderate stratum. Hence, to reduce field work costs and time, the sample housing units of this stratum were selected at five times the basic survey rate or at 1 in 655. Then the secondary selection units were formed into strata of around 80,000 total housing units and one sample secondary unit was selected per stratum. Unlike the other strata, here total housing units not Jewish population were assigned as secondary unit measures of size. As in the other two strata, systematic selection was used to obtain sample secondary units. Also, the two-phase name rating technique was utilized. All this resulted in these three sampling equations:

a)
$$\frac{1}{655} = \frac{\text{psu J. pop.}}{\text{stratum J. pop.}} \chi \frac{\text{Secondary}}{\text{sec. unit T. hu.}} \frac{\text{sec. unit T. hu.}}{\text{sec. stratum T. hu.}}$$

b)
$$\frac{1}{2620}$$
 = (same as above) $\frac{\text{Jewish hu. rating}}{(1/4)}$

c)
$$\frac{1}{6550}$$
 = (same as above) $\frac{\text{Jewish hu. rating}}{(1/10)}$

Mere, Tohu. stands for total housing units in secondary units and secondary strata.

The Survey Respondent

Sample housing units were screened for the presence of one or more

Jewish respondents by going through a five to ten minute screening

interview which determined the residents of the housing unit and asked
which of them had been born Jewish, had either parent born Jewish, or
regarded themselves as being Jewish. If any one of these points

applied to any resident, that housing unit was considered eligible
for interviewing. Otherwise, the members of the field force were
instructed to terminate the interviewing process.

f

In those housing units eligible for the survey, any adult respondent who knew the required information was interviewed as to family characteristics and individual behavioral items. Then, the Kish adult selection table technique (see Kish: 1949) was applied to obtain one adult at random from among all the Jewish adults of the sample housing unit. Only that sampled adult was then asked the series of attitude and behavior questions of interview schedule Section 1.

An adult was defined as anyone 21 years of age or older. In those few cases where the Jewish occupants of a sample housing unit were married and under 21 years of age, they were also interviewed.

Members of the families temporarily away -- say at college -- were assigned to the involved sample housing unit if they were then living in some form of group quarters or an institutional setting. Otherwise they were dropped from that housing unit's list of occupants and had their chance of entering the survey via their own housing unit. Family members living in group quarters or institutional settings for a lengthy or indefinite period were not regarded as members of the involved sample housing unit. With these respondent definitions, we excluded from survey coverage those Jews in homes for the aged, prisons, or custodial care in mental hospitals and other equivalent settings.

Some Design Exceptions

A. The New York City PSU:

While the above sampling equations hold, for the most part in New York City and surroundings, sample psu and cost problems necessitated a slight reduction in that psu's sample size. There, the overall sample fraction was reducted to 1 in 197 and the 131 number in the various terms of the area sample design equations just presented has to be changed to 197 for New York.

B. Subsequent Cut-Back:

After the survey had been in the field for around nine months, we encountered spiraling costs of interviewing which forced adoption of a decision to make a considerable reduction in the size of the total sample. Accordingly, all psu samples that had not yet been fielded were reduced by one-half. Those samples then in the field with substantial amounts of work still to do were cut-back by one-half. Those samples in the field that were almost done were left as is.

The reduction of one-half in those samples not yet fielded was done by raising their sampling fractions to 1 in 262 from the earlier 131. Then in the area of sampling equations, the interval for the selection of sample blocks was doubled so that half the previously expected number of sample blocks were obtained.

Only the uncontacted list sample housing units assigned odd-numbered sample book numbers were retained. Of course, those list sample housing units which were uncontacted by the cut-back date of April 1, 1971, were so as a result of all the vagaries of the field process. Nevertheless, it is thought that this reduction mechanism is not a biased procedure. Instead, one can regard it as dividing the list sample housing units into two additional strata. Stratum 1 is composed of those sample housing units that were contacted before the cut-back date; stratum 2 is composed of those sample housing units contacted after the cut-back date. These two strata were formed by a subjective allocation procedure evolving out of the field process, but subjective allocation is permissible in stratification procedures.

The price paid for such a belated cut-back was a decrease in sample and operational efficiencies. The best way to cut-back a sample is across the board on the entire sample design. Here, we were forced to resort to an approach that is akin to disproportionate allocation without any recanningful between-strata cost differentials.

The area sample cut-back was done by a variety of techniques. First, we doubled the block selection rate in samples not yet fielded. The same doubling of the block selection intervals, within secondary selection units, was applied to samples already in the field. If many of the blocks that were thus eliminated had been visited by the field force, the cut-back was abandoned. Otherwise, they, and their sample housing units, were entirely dropped from the survey. Finally, up to the cut-back date we had not actually been using the Jewish hu rating method in the high numbers stratum. Now, we started to apply it.

These emergency methods considerably reduced field costs while still retaining a probability sample. Of course, all cut-back sample housing units require, and receive, an additional weight factor of 2.

C. Integration of List and Area Samples:

The proper, and full, integration of the list and area samples was insured by a number of techniques. First of all, all lists were classed as either types A or B. Type A lists had all their housing unit addresses numerically ordered by streets and streets listed in alphabetical order. Type B lists had the names of housing unit occupants alphabetically ordered within postal zones and these zones listed in numerical order.

The addresses of the type A lists were considered those of one or more housing units. When a list sample address was selected that had but one entry on a federation list, interviewers were instructed to screen for eligible respondents at all housing units at the given address. If any such corress turned out to have 5 or more housing units, the interviewers have instructed to tell sampling and field headquarters how many housing units were in the structure, but not to do any screening at that time. Sampling then reviewed the situation and its possible screening and interviewing work loads and decided which one to employ out of a variety of standard area sampling techniques used to deal with unexpected cluster sizes. Customarily, we decided to screen all involved housing units for Jews, found few, and included them into the survey.

The situation was somewhat simpler if a list sample address had two or more entries on the federation list. Then, we applied an interval to all the housing units found at that address which was equal to the proportion of federation entries for that address selected into the sample. For example, suppose an address selected into the sample from a type A list actually appeared on the list ten times because ten Jewish families lived at that address. If we selected one out of ten of the entries for that address, we went to the structure at that address, listed all its housing units (customarily appartments), and screened one out of ten of such units. This gave all the housing units at a multiple listed federation address their proper chance of coming into the list sample and

eliminated them from the area sample. The technical details of such a type A list sample-area sample combination is the same as that for a multiple entry city directory address.

Area samples from type A places were integrated with these kinds of the samples by furnishing field workers with a list of those addresses on sample blocks that appeared on the full federation list. They were told to exclude all housing units at such addresses from the block area samples (unless otherwise instructed). All housing units at addresses not on federation lists were accepted into the first phase of the area sample, processed by the Jewish hu rating technique, and screened for survey eligibility. When the Jewish hu rating method was applied to area sample selections, they were again checked with the proper federation lists to make certain that no list sample housing units were mistakenly permitted into the area sample.

The addresses on type B lists were regarded as those of just single housing units. When a sample address from such a list turned out to have two or more housing units, the interviewer was instructed to list all housing units at the sample address, together with occupants' names, and to return the list to field headquarters, but not to do any screening at the involved address. Then, the local Jewish federation was contacted for the specific apartment number associated with the family they had listed at that address. Only this specific apartment number was then screened for survey eligibility and interviewing.

When an area sample block was selected from a type B list place, all its housing units fell into the survey if their addresses and not appear on the federation list entries. When an area sample address contained two or more housing units, an interviewer was instructed to list all housing units at that address and return the listing to field headquarters without doing any screening at the involved address. We then determined the specific apertment numbers at that address which appeared on the local type B federation listing. These apartments were excluded from the area sample while all other apartments at the involved address were eligible for the area sample portion of the survey's sample. Again, all area sample housing units eligible for inclusion in the survey were checked against federation listings to make sure they had not been mistakenly included in the area sample.

Most survey lists were type B. They were much more difficult to work with for entire postal zone list entries had to be carefully gone through to find addresses on area sample blocks. The relatively few type A lists, being by address number within an alphabetical listing of streets, were far easier to process.

The Crucial Design Decision

The most crucial design decision for this survey hinged on the sample size to be sought. Whatever the sample size, it would be

a compromise between anticipated analysis benefits and cost assumptions. There was much debate among members of the design team on what would constitute the "optimum" compromise in sample size.

The final decision saw a split in the design team. The majority felt that a sample size was needed which would serve as a base for detailed cross-tabulations, and which would enable the study to produce regional and size strata statistics. The minority felt that a 3,000 to 4,000 sample size was the best choice since a larger sample would produce costs higher than the estimated budget and further, that too large a sample would run the additional disadvantage of exceeding the field force capacity of available quality survey research organizations.

The survey sponsor, the Council of Jewish Federations and Welfare
Funds, chose the majority report and with that decision, the basic
design characteristics of the survey were set and had to be lived with.

While the study has been able to develop a larger data base for detailed cross-tabulations, experience in the field demonstrates that the minority report was correct in some of their views. Certainly costs were larger than initially expected and survey time in the field exceeded expectations with concommitant survey difficulties.

As the belated cut-back in sample size just discussed indicates, in many ways the minority report was correct. Certainly, costs were larger than initially expected and survey time in the field exceeded expectations.

Weighting

A. For Housing Units and Families

In such a complex survey as this, with an involved sample design employing disproportionate sampling as a major cost reduction approach, careful attention has to be paid to an effective weighting scheme. The sample design strata detailed above require these weights to get all sample housing units to an effective 1 in 131 sample.

Sample Stratum		Sample Fraction
1. Original list sample	1	131
2. List sample after cut-back	2	262
3. High numbers stratum before cut-back		
a. "Probably Jewish" housing units	1	131
b. "Don't know if Jewish" housing units	4	524
c. "Probably not Jewish" housing units	10	1310
4. Moderate high numbers stratum before cut-back	(
a. "Probably Jewish" housing units	3	393
b. "Don't know if Jewish" housing units	12	1572
c. "Probably not Jewish" housing units	30	3930

Sample Stratum	Weight	Sample F raction	
5. Light numbers stratum before cut-back			
a. "Probably Jewish" housing units	5	655	
b. "Don't know if Jewish" housing units	20	2,620	
c. "Probably not Jewish" housing units	50	6,500	
6. High numbers stratum after cut-back			
a. "Probably Jewish" housing units	2	262	
b. "Don't know if Jewish" housing units	8	1,048	
c. "Probably not Jewish" housing units	20	2,620	
7. Moderate numbers stratum after cut-back			
a. "Probably Jewish" housing units	6	786	
b. "Don't know if Jewish" housing units	24	3,144	
c. "Probably not Jewish" housing units	60	7,860	
8. Light numbers stratum after cut-back			
a. "Probably Jewish" housing units	10	1,310	
b. "Don't know if Jewish" housing units	40	5,240	
c. "Probably not Jewish" housing units	100	13,100	
9. New York Area Sample (no cut-back)			
A. High Numbers Stratum			
a. "Probably Jewish" housing units	496 out 1,000 weig 1; remaind	hted	
b. "Don't know if Jewish" housing units	935 out 1,000 weig by 6; rema by 7	hted	
c. "Probably not Jewish" housing units	952 out of 1,000 weighted b 15; remain by 16	У	

Sample Weight Fraction

B. Moderate Numbers Stratum

- a. "Probably Jewish" housing units 490 out of 591 1,000 weighted by 4; remainder by 5
- b. "Don't know if Jewish" housing units 949 out 2,364 of 1,000 weighted by 18; remainder by 19
- c. "Probably not Jewish" housing units 885 out 5,910 of 1,000 weighted by 45; remainder by 46

C. Light Numbers Stratum

- a. "Probably Jewish" housing units 480 out 985 of 1,000 weighted by 7; remainder by 8
- b. "Don't know if Jewish" housing units 924 out 3,940 of 1,000 weighted by 30 remainder by 31
- c. "Probably not Jewish" housing units 810 out 9,850 of 1,000 weighted by 75; remainder by 76

Not only do these weights apply to housing units, but they apply equally as well to measurements based on the entire elements of housing units, such as families, heads of families, spouses of heads, etc.

B. For Persons, Adults, and Attitudes

However, when we want to work with the data derived from specific persons or adults, or attitudes by individuals, we must add to the agove weighting scheme. This additional weighting state has to be handled with considerable

caution because it is a complex operation. If, for example, we want to deal with some item pertaining to individuals, such as age, and we work with the age data on all individuals in sample housing units, then just the housing unit weights are needed. However, if we work with a randomly sub-selected person, then the data for that person must be weighted by the product of the housing unit weight and the sub-selection factor used to obtain the sub-sample of people. this survey, we selected a cross-section sample of Jewish adults within all sample housing units. Only to such adults did we administer an interview Section 1. The items which compose Section 1 have become decks 71, 72, and 73 of the survey codes. Whenever one works with the data from these code decks, he must use the adult cross-section sample weight which is the product of the weight of a sample housing unit times the number of Jewish adults in the sample housing unit -- up to six in number. (Even if there are more than six such adults in a sample housing unit, the Kish technique used to select the adult subsample limits this weight to a maximum of 6.)

Whenever one wants to work with information on the adult cross-section sample respondent, he must also use the adult weights. Of course, to make analysis feasible for anyone, we have incorporated both the required housing unit weights and adult cross-section sample weights into the data codes and in the computer variable analysis system. Hence, for statistical work with programs that permit weight factors, it is just a matter of naming the right weight variable within the computer program requirements.

Yet another word of caution is required with regard to decks 71, 72, and 73. In the New York psu, we actually administered the section 1, which gathers the data for these decks, to a random half-sample of the survey's Jewish housing units. Hence, the New York psu's adult cross-section weights are twice the product of the housing unit weights and the Jewish adults in the housing unit. In itself, this is no problem for the correct weights are present in the survey's codes as just explained. Of course, whenever one works with any data from code decks 71, 72, or 73 which includes the New York psu, he must work with the adult cross-section sample weights and with just those interviews in the New York psu which received the section 1 portion of the survey interview. Such sample housing units in New York (and in the rest of the sample apart from survey non-respondents) are coded 1 in column 75 of code deck 11.

Whenever one is working with data other than those from New York's decks 71, 72, or 73, he can choose to work with the sample of New York City adults, and can change the New York adult weights to just the product of their housing unit weights and the number of Jewish adults in the housing unit. This is one-half the value of the adult cross-section weights presently given to the New York psu based on the half-sample of that psu.

The Cut-Off Bias

In areas with little or no known Jewish population this survey faced two sampling problems. First of all, seldom were local lists of Jews available. Secondly, any area sample approach would be extremely expensive and would obrain very few, maybe no, Jewish interviews.

In some extent, such areas were covered by nearby Jewish federations. Those federations located in relatively small Jewish communities do extend their functional territory throughout neighboring counties. Hence, forming primary sampling units out of the counties containing known Jewish federations and their surrounding tier of counties resulted in some degree of additional list sample coverage by local federations.

However, not all the territory of sample primary sampling units was covered by interior Jewish federations. Then, a few sample primary sampling units had no local federations. When these situations occurred, lists were created by obtaining the local telephone directories and searching out the addresses of those names thought to be distinctively Jewish, according to research done on this tactic by Dr. Massarik. Then, such lists were sub-sampled at the proper within-psu sampling intervals to obtain their contribution to the national sample. Next, such sampled addresses were

contacted by long distance telephone by our field headquarters staff and screened for Jewish housing units. These, when found, were interviewed by phone. Also, such Jewish respondents were used to add more addresses to the local lists of Jewish housing units. Whenever such localities contained, or were near synagogues, these were contacted for a list of the addresses of their membership residing in the sample counties. Unfortunately, the majority of such synagogues did not cooperate. Hence, all in all, only fair to good coverage of such local Jews could be obtained by this list development technique, but that is considerably better than no coverage at all of an especially rare local population type.

Then, we did not screen any area sample of housing units in several of the outlying light stratum counties of some of our sample primary sampling units. The same dropping of the area sample was adopted in ten sample primary sampling units which were entirely light stratum and whose lists were thought to cover a very high percentage of the local Jewish population, or where the number of local Jews was so small as to make area sampling extremely expensive. Hence, for ten sample primary sampling units and 34 outlying light stratum counties of other sample primary sampling units, a cut-off bias was accepted in the area sampling process and somewhat in the thoroughness of list sample coverage. Such a bias actually missed

very few Jews and, in turn, removed the need to expend a considerable amount of effort and money to try and seek out a mere handful of Jewish interviews. We estimate the extent of this bias in the report on the total Jewish population of the United States.

Field Forms

Throughout this presentation of the sample design created for the national Jewish population survey, mention has been made of the various techniques used to control and guide field operations. The forms for such techniques are shown in Appendix I following the text. Form 7, the first field form shown, was used to define sample primary sampling units. Here, by psu, was given the name of any specific sample primary sampling unit and, then, its rate of selection from within its original stratum. Then came a listing of all the counties forming a specific sample primary sampling unit.

Often, it was necessary to sub-select from among the many light numbers stratum counties of a sample primary sampling unit. When this was done, part A under secondary sample units gave such counties their stratum, selection rates, and resultant within-rates for internal sampling needs. Those counties selected with correlately were also listed in this section of the form.

Under part B, within counties, was listed all the secondary selection units, postal zones, census tracts, etc., that are the secondary sample elements within the sample counties together with their

strata, selection rates, and within-rates for the selection of sample blocks and housing units.

Form 8 of the next illustration, has been used to list the specific sample blocks selected for the survey within the set of sample secondary units given on form 7. Each such sample secondary selection unit was given under place and then all its sample blocks were listed in the next column. The third column gave any within-block sample rate used to obtain specific portions of a sample block when it contained more than one block measure of size. Then, the next column gave within-measure sub-selection rates, such as one-half, when these were required.

Form 5, shown in the third illustration, gave a block sketch to interviewers so they could determine the boundaries of sample blocks when in the field. Next comes a copy of the field instructions guiding area sampling operations for type B list places. Then comes the form 6, used to list those housing unit addresses on area sample blocks to be excluded from the area sample because they were on the local Jewish federation list of "probably Jewish" housing units.

Finally, form 4 was that form used to guide the application of the Jewish hu rating system to area sample blocks. Interviewers received a supply of these field area sample housing unit listing

forms for each sample block. The necessary heading items as to psu, place, and sample block were filled in prior to field operations. Then followed a set of field headquarters instructions to be ignored by interviewers. The instruction box on the left gave the random starts to be used for the I in 4 sub-sampling of those block housing units judged to be "don't know if Jewish" hu's and the I in 10 judged to be "probably not Jewish" hu's. The sampling instructions to the right were used for the "probably Jewish" hu's and gave the variety of possible sampling instructions for this hu category.

The work procedure with form 4, and an area sample block went like this:

Interviewers received:

- 1. A block sketch, form 5, a filled-in form 6, and its instructions (if the place had a type B list), and a form 4 for the same block with its heading items filled in.
- 2. The interviewer next went to the sample block and checked its boundaries. If boundary problems developed, they were resolved by field headquarters and sampling before proceeding further. Otherwise the interviewer listed all block housing units together with occupants! names on the form 4 (using additional blank form 4's for extra pages). The entire group of field forms were returned to national field headquarters for the next step.

- 3. In national field headquarters, a group of specially trained people checked the listing procedures and assigned each housing unit on the block to one of the three Jewish hu rating categories by entering checks in columns 4, 5, or 6. If a housing unit was thought to be "probably Jewish", it received a check in column 4; if thought to be "don't know if Jewish", it received a check in column 5; if thought to be "probably not Jewish", it received a check in column 5; if thought to be "probably not Jewish", it received a check in column 6.
- 4. For the next step, the Jewish hu rating people applied the same instructions given at the top of form 4 to the listing of housing units. This was done by circling in red those checks of columns 4, 5, and 6 which came into the sample by the application of the sampling instructions on the top of the form 4 page. For example, if the right side box "screen odd lines" was checked, the first, third, fifth, etc., checks for "probably Jewish" hu's entered in column 4 were circled in red. If the left side box had a 2 entered for the 1 in 4 "don't know if Jewish" category, the second, sixth, tenth, etc. checks entered in column 5 were circled in red. If the same box had an 8 entered for the 1 in 10 "probably not Jewish" hu's, then the eighth, the eighteenth, the twenty-eighth, etc., checks in column 6 were circled in red.
- 5. Finally, the entire set of field forms were returned to the original interviewer with instructions to return to the involved

sample block and screen for survey eligibility just those housing units with red circles in columns 4, 5, and 6. Those "checked and circled" housing unit occupants found to be eligible for the survey were then interviewed as part of the national Jewish population survey.

The Sampling Error Model

The final order of business was to establish a sampling error computational model which permits the determination of both specific and generalized sampling errors for the totals, means, and percentages of this survey and, hence, the ability to establish confidence limits around the data of this survey.

Several design features were built into the sample to enable the construction of a correct sampling error model. First of all, the various strata composing the national sample were created in a definite hierarchy from New York City down through the "no known Jewish population" stratum number 39. Then, the non-self-representing psu's form their own sampling error units. The sample secondary selection units of the self-representing psu's are their sampling error components since these secondary units are the first sub-selected elements for this type of psu.

Secondly, the best sampling error computational model for this design with its own computer program package is that developed at the Survey Research Center of the University of Michigan by Professors Kish and Frankel, and Irene Hess. The details of their work are available in Kish (1965), Kish and Hess (1959), Frankel (1971), (Kish and Frankel (1970), and Kish, Frankel, and Van Eck (1972). Hence, our sampling error model has been adjusted to both their mathematical frame and computer program.

To apply this model, each non-self-representing sample primary sampling unit was assigned one sampling error computational number starting with the "no known Jewish population" stratum 39 which has been numbered 1 down through the Hartford psu, which has been numbered 21. These sampling error numbers apply to both the list sample and area sample interviews obtained in these 21 non-self-representing psu's.

The 18 self-representing psu's were arrayed in order from Cincinnati down through Rockland County of the New York City psu. Then, their area sample secondary units were listed within each of these psu's in order of selection starting with the high numbers stratum. They were each assigned single sampling error numbers starting with 22 for the first sample secondary unit of the high stratum in Cincinnati down to number 377 for the last light numbers stratum secondary selection unit of Rockland County. Following the rules for sampling

error numbers permissible with the Kish-Hess model, some area sample secondary units, customarily belonging to the same Jewish numbers stratum, were collapsed together into the same sampling error number when they had no, or very few, sample Jewish housing units.

The sample from the various self-representing psu's is composed of two basic parts. One part is the area sample and the other major part is the list sample. While the area sample is a complex, multistage disproportionate design, the list sample is just a systematically selected one deriving from the series of master lists provided by the local Jewish federations. While simple random sample variance formulas can be applied to list sample data, it is more efficient to link the variance computations for both area and list sample parts. This the Kish-Hess-Frankel model permits, provided individual list sample housing units are grouped into small clusters. This was accomplished by listing each self-representing psu's list sample in systematic selection order, clustering them by groups of four sample housing units within the same psu. and assigning each such cluster a sampling error number. Again, such psu lists were arrayed from Cincinnati through Newark (with New York missing since it has no list sample). These list sampling error numbers run from 378 to 970.

To apply the Kish-Hess-Frankel model to this sample design, the paired differences form of the sampling error model was assigned to the non-self-representing sample primary sample units. To obtain the necessary even number of sampling error units required by this model form, we paired the sampling error number for Hartford (21) with the first area sample number assigned to Cincinnati (22).

The area sample from the self-representing primary sampling units had the successive differences form of the sampling error model applied to secondary selection unit sampling error numbers and the clusters formed from their list sample housing units. Since the model's program also computes the corresponding simple random sampling variances of each complex sample variance in order to obtain cluster sample design effects, it is readily possible to see the extent of variance increase caused by clustering the list sample interviews. This can be done by computing separate sampling errors for just the list sampling error numbers. This gives both the cluster sampling variances for the four housing unit clusters and the simple random sampling variances for the same list sample statistics. The two sets of variances can be contrasted and adjustments made for any substantial variance increases arising from clustering list sample addresses.

APPENDIX 1-a

Selection Rate

(Form 7 11/13/70)

PSU

NATIONAL JEWISH POPULATION SURVEY SAMPLE FRAME

-Primary and Secondary Units-

Composing Counties:				
Secondary Sample Units	Strata	Selection Rates	Within Rates	Comments
A. Subselected Counties				
ago agos, calquing e reconsiderativos e interespendente der administrativos como calculos de administrativos m				-
(
B. Within Counties				

orm	2.
1/1	3/70)

NATIONAL JEWISH POPULATION SURVEY SAMPLE FRAME

1/12/70) PS U		-Block Clusters-	·	Pageof_(
Place	Sample Blocks	Within Block Rates for Sample Meas.	Within Meas. Rate (if not one)	Comments	
				·	
alara. Ay villeya diga payahan yang perkebalan bahaya barr 1900 te				(
·					
				ſ	

APPENDIX 1-c

Form (12/12/69)	NATIONAL JEWISH POPULATION SURVEY SKETCH SHEET		
PSU	PLACE	BLK. CLUSTER	

You are to list all housing units within the designated boundaries of this block cluster on the Form 4 assigned to this block cluster. Then, follow the indicated instructions on the Form 4.

TYPE B LIST AREA

For this block cluster it will be necessary to do an extra step at all the addresses that appear on the attached form 6.

First of all, go to each address within block cluster boundaries that appears on the form 6 and determine how many housing units are there. If there is but one housing unit at every address on form 6, follow the procedures given for this block cluster in your manual for interviewers.

But, if any one of the addresses on this block cluster's form 6 turns out to have two or more housing units, then do the following:

- a) List all housing units at those form 6 addresses with two or more housing units. As often as possible, get the last names of housing unit occupants if this can be done fairly easily.
- b) Finish listing the block cluster as instructed by your manual for interviewers, but do not do any interviewing on the block. Instead, return all the block forms (form 5, form 4¹s, and form 6) along with the listings you have done at the various form 6 addresses with two or more housing units, and a note explaining the situation to:

Bernard Lazerwitz
Public Opinion Survey Unit
University of Missouri
7018 Pershing Avenue
St. Louis, Missouri 63130

c) Also notify your field headquarters of this action on your part.

Shortly, you will receive back the block cluster with the form 6 addresses corrected so that each represents just one housing unit. Then, all you need do is to follow the area sample instructions in your interviewers manual.

One final note of caution. If a form 6 address is that of a specific apartment unit, say 1123 Dayan St., Apt. 3, you are to check for the number of housing units at that apartment itself. In other words, addresses on form 6 are, to the best of our knowledge, those of single housing units. This extra operation is aimed at catching, and correcting, those cases in which our list information fails and a form 6 address turns out to contain two or more housing units.

APPENDIX 1-e

(9/70, NATIC			
PSU	PLACE	BLK.CLUSTER	
Exclude from the above in	ndicated block cluster any housi	ng units listed below.	
Housing Unit Ad	Idresses to be Excluded	Housing Unit Addresses to be Excluded	
,)	δ		
. 2	7		
3	8		
4	9		
5	0		
6	1		
7	2		
8	3		
9	4		
0	5		
	6		
. 2	7		
3	8		
4	Θ		
5	0		

COMMENTS

APPENDIX 1-f

Form 10/3	4 PATIONAL JEW 0/70 BLOCK CLUS	ISH POPULATION SURVEY Sheet_ STER LISTING SHEET Date_	of
PSU_	Flace	Blk.Cluster	
lnte	rviewer	- SAMPLING INSTRUCTIONS	
Co.	dium and Low Density Blocks 1. (5) Starts(1 in 4) 1. (6) Starts(1 in 10)	In Screen Odd Lines	reen Those nes Checked Col. (4) Col. (4) Collow Attached Structions
Line	Address(or Description)of Unit (2)	Occupant's Information (3)	J 1/4 1/10 4) (5) (6)
1			
2			
3			
4			(
5			
6			
7			
8			
9			
0			
1			
2			
3			
4			
5			
6			(
7			
8			
9			

Bibliography

Frankel, Martin.

1971, Inference From Survey Samples: An Empirical Investigation.
Ann Arbor, Michigan: Institute for Social Research,
University of Michigan.

Kish, Leslie.

1949, "A Procedure for Objective Respondent Selection Within the Household". Journal of the American Statistical Association. 44 (September), 380-387.

1965, Survey Sampling. New York: John Wiley.

Kish. Leslie. and Martin Frankel.

1970, "Balanced Repeated Replications for Standard Errors". Journal of the American Statistical Association. 65 (September), 1070-1094.

Kish, Leslie, and Irene Hess.

1959, "On Variances of Ratios and of Their Differences in Multi-Stage Samples". Journal of the American Statistical Association, 54 (June), 416-446.

Kish, Leslie, Martin Frankel, and Neal Van Eck.

1972, <u>Sampling Error Program Package</u>. Ann Arbor, Michigan: Institute for Social Research, University of Michigan.

Lazerwitz, Bernard.

1968, "Sampling Theory and Procedures". In Methodology in Social Research, edited by Blalock and Blalock. New York: McGraw-Hill, 278-328.