200-74

INSTITUTE FOR RESEARCH ON POVERTY DISCUSSION PAPERS

INDEXES OF RACIAL RESIDENTIAL SEGREGATION FOR 109 CITIES
IN THE UNITED STATES, 1940 to 1970
(with METHODOLOGICAL APPENDIX)

Annemette Sørensen, Karl E. Taeuber, and
Leslie J. Hollingsworth, Jr.

Studies in Racial Segregation, No. 1

UNIVERSITY OF WISCONSIN~MADISON

INDEXES OF RACIAL RESIDENTIAL SEGREGATION FOR 109 CITIES

IN THE UNITED STATES, 1940 to 1970

(with METHODOLOGICAL APPENDIX)

by

Annemette Sørensen, Karl E. Taeuber, and Leslie J. Hollingsworth, Jr.

February 1974

INDEXES OF RACIAL RESIDENTIAL SEGREGATION FOR 109 CITIES

IN THE UNITED STATES, 1940 TO 1970

(with METHODOLOGICAL APPENDIX)

ABSTRACT

An index of racial residential segregation for a city may be cal-
culated from census data reporting the number of white and nonwhite
households resident in each city block. Such an index summarizes the
dissimilarity in residential location of the two groups. Segregation
indexes for 109 cities for the years 1940, 1950, and 1960 were published
in 1965 in Taeuber and Taeuber, Negroes in Cities.

Segregation indexes for 109 cities for 1970, together with the
previously published indexes for the three earlier census dates, are
presented in Table 1 of this paper. The trend data for 1940 through
1970 refer to residential segregation of whites and nonwhites. For
1970 only, segregation indexes are also presented comparing whites and
Negroes.

The indexes for 1970 were calculated from Public Use Summary Tapes
issued by the Bureau of the Census. Certain difficulties were en-
countered in use of these tapes for this purpose. A methodological
appendix describes the data source and retrieval procedure for calcu-
lation of city segregation indexes for 1970.

To permit quick dissemination of the indexes to other researchers,
this paper is issued without substantive analysis or interpretation.

INDEXES OF RACIAL RESIDENTIAL SEGREGATION FOR 109 CITIES

IN THE UNITED STATES, 1940 TO 1970

Racial residential segregation is a pervasive aspect of American

urban life. Residential segregation of Negroes and other ethnic minori-

ties has been shown to be both cause and consequence in the intricate

web of relationships among poverty, racial discrimination, education,

access to governmental services and benefits, and the life chances of

children. Systematic study of residential segregation in cities through-

out the country requires data on the racial composition of small areas

within each city. The 1940 Census was the first decennial census to

provide information on the number of white and nonwhite households in

each city block. Various measures were subsequently proposed as indexes

of residential segregation, to provide for each city a summary measure

of the degree to which whites and nonwhites differ in their residential

locations.

In 1965, Taeuber and Taeuber published trend data on racial resi-

dential segregation in 109 cities for the years 1940, 1950, and 1960.

Their monograph, Negroes in Cities, reported a number of analyses of

residential segregation and neighborhood change, utilizing segregation

indexes and a variety of other measures derived from census data. In

1971, The Institute for Research on Poverty at the University of Wis-

consin embarked on the task of updating to 1970 and in other ways

extending these earlier studies of racial residential segregation.

The newly calculated indexes for 1970 together with those for the three earlier census dates are given in Table 1 of this report. The accompanying text provides a description of the index and of some limiations of the 1970 data. A methodological appendix describes in more detail the data source and retrieval procedure used for calculation of the 1970 indexes. To avoid delay in release of this data series, no substantive analysis is reported here. Later reports will present our analyses and interpretations, and we hope that other scholars and planners will also make use of these trend data.

Levels of segregation are measured by the index of dissimilarity (D) between the distributions of nonwhite and white households among city blocks. The formula for the index of dissimilarity is

$$D = (100) \frac{1}{2} \sum \left| \frac{N_i}{N} - \frac{W_i}{W} \right| ,$$

where i = city block, numbered in any serial order;

N_i = number of nonwhite households in block i;

$N = \sum N_i$ = total number of nonwhite households in city;

W_i = number of white households in block i;

$W = \sum W_i$ = total number of white households in city.

The index of dissimilarity is a summary measure of the divergence between two population distributions. If the two populations are distributed identically among city blocks, the index assumes a value of 0, indicating no segregation. If the two populations are distributed with complete disjunction, such that no block has both white and non-white households, the index assumes a value of 100 indicating maximum

segregation. Index values between 0 and 100 indicate the degree to which the residential distribution in the city approaches one or the other extreme.

For convenience, the index is expressed as a percentage. One interpretation of the specific value for a city is that it represents the minimum percentage of either population that would have to change residence to bring about a zero degree of segregation. Thus, an index of 100 indicates that 100 percent of either the white households or the nonwhite households would have to move to blocks containing the other group in order to attain an unsegregated distribution. Alternatively, the index may be interpreted as the percentage of non-overlap or dis-similarity in the two residential distributions.

The index is a methodological tool for characterizing the racial residential pattern of a city. It is not intended as a means of characterizing individual city blocks as segregated or unsegregated. Rather, it provides a single summary measure of the entire residential pattern of a city.

In this report, the areal unit for assessing residential patterns is the city block. The index of dissimilarity may be defined anal-ogously using census tracts, wards, or any other system of areal units. The magnitude of the index value is sensitive to the system of areal units used for its calculation, and it is generally inadvisable to compare indexes calculated from different areal units.

An extended discussion of these and other aspects of "The Measure-ment of Residential Segregation," with citations of the relevant scholarly literature, constitutes Appendix A of Negroes in Cities.

Census data for city blocks were first published as part of the 1940 Census of Population and Housing. In the 1940, 1950, and 1960 Censuses, city block data were tabulated and published primarily for cities of 50,000 or more population. The only data released for public use were in published block bulletins for each city. For each block, one line of data was published, covering 18 columns of information in 1960 and fewer in earlier years. The only item of racial information for blocks in each of the three censuses was the number of occupied housing units occupied by nonwhites. In census usage, each occupied housing unit is occupied by a household, and each household has a head whose racial identification is used to classify the household as white or nonwhite. In the earlier study, segregation indexes were computed for all 109 cities for which block bulletins were published for 1940, 1950, and 1960, and which in 1940 contained more than 1,000 nonwhite households.

In the 1970 Census, block data were tabulated for all blocks within urbanized areas. The data were released in published "Block Statistics" bulletins similar to those for prior censuses, and also in "Public Use Summary Tapes." The tapes include approximately 250 items of information for each block, and permit division of nonwhites into "Negroes" and "Other Races." For comparability with the segregation index series for 1940 to 1960, indexes were calculated for 1970 using data for white and nonwhite households. In addition a separate index was calculated for 1970 using data for white and Negro (black) households, ignoring "other races." Both indexes are presented in Table 1.

Segregation indexes for 1970 were calculated from the summary tapes. Certain aspects of the processing of the tapes are discussed in the Appendix to this report. Some of the problems, such as suppression of occupancy data for blocks containing fewer than five occupied housing units and occasional discrepancies between city totals from block tallies and later "corrected" totals, also occur with the earlier censuses. A new and very troublesome problem unique to 1970 was the failure to include either on the tapes or in the published bulletins specification of the location of each block within an urbanized area—in particular, whether the block lies within the limits of the central city or of another municipality. The procedure for overcoming this problem is described in the Appendix. We believe this procedure sufficed to eliminate significant biases in the reported index values. But it remains the case that the indexes for 1970 were calculated from data for a universe of households and blocks that approximately but imperfectly corresponds to the universe of households and blocks actually within the city limits.

Persons who wish copies of future reports in the series "Studies in Racial Segregation" should request the Institute for Research on Poverty to place them on the mailing list for the series. Persons who use these indexes for scholarly research or other purposes are invited to send us copies or notice of their work. Notification of errors or ambiguities and comments on any aspects of this work are eagerly solicited.

FOOTNOTES

1   "The urbanized areas ultimately used in the tabulations of the
1970 census could not be completely recognized or defined until the
final population counts were available. Since the precise delineation
and numbering of blocks is quite costly and is therefore done only
for areas where block statistics will be prepared, a preliminary
determination of each urbanized area was made prior to the actual
enumeration. The preliminary boundaries were purposely drawn some-
what beyond the anticipated final boundaries because the information
on density of settlement was not yet available. Thus, for virtually
all urbanized areas, there is some immediately adjacent territory
for which the information was collected on a block basis. Because
the extra cost of tabulating and publishing the block data for this
type of territory is comparatively minor, data for blocks outside
urbanized areas appear in the reports for those areas."("Introduction"
to each 1970 "Block Statistics" bulletin, p. iv).

2.   "[The block group] designation is new in 1970, and is used in
census-by-mail areas where Address Coding Guides have been prepared. A
block group is a combination of contiguous . blocks having a combined
average population of about 1,000. Block groups are approximately equal
in area (discounting parks, cemeteries, railroads yards, industrial
plants, rural areas, etc.); they are subdivisions of census tracts
which simplify numbering and data control. Each block group is identi-
fied by the first digit of the three-digit block number. Block group
"1" will contain any block in range 101-199, block group "2" in range
201-299, etc. However, normally only the first few numbers in a range
are used. For purposes of providing small-area population and housing
census data, they are the equivalent of enumeration districts within the
mail-out/mail-back areas where Address Coding Guides have been prepared.

      "Block groups (and blocks) are typically defined without regard
to the boundaries of political or administrative areas, such as cities,
minor civil divisions, and congressional districts. When a block group
straddles one or more of these boundaries, data for those parts in
different areas will be tabulated separately. Where such a split occurs,
the tapes contain two (or more) data records having the same block
group number within the census tract but a different place, annexation,
minor civil division, or congressional district code depending on the
situation." (1970 Census Users' Guide, Part I, p. 87).

      The MEDList can be used to identify split block groups.

3.   Each of the cities of interest in this study, with the exception of
East Chicago, had a population in 1970 greater than 50,000. Urbanized
areas are composed of central cities with at least 50,000 population and
the place description code tells whether a given MEDList record is
located in a central city. Hence, it is easy to determine whether the
block group or tract for a given block is at least partially in the city.

TABLE 1

INDEXES OF RESIDENTIAL SEGREGATION BETWEEN WHITES
AND NONWHITES, 1940 TO 1970, AND BETWEEN WHITES AND NEGROES,
1970, FOR 109 CITIES

	INDEXES OF RESIDENTIAL SEGREGATION				
	White vs. Negro	White vs. Nonwhite			
CITY	1970	1970	1960	1950	1940
Akron, OH	82.5	81.2	88.1	87.6	82.2
Asheville, NC	89.3	88.5	92.3	89.2	88.6
Atlanta, GA	91.9	91.5	93.6	91.5	87.4
Atlantic City, NJ	87.9	86.9	89.2	94.0	94.6
Augusta, GA	93.8	93.3	93.0	88.9	86.9
Austin, TX	90.2	84.6	93.1	92.0	84.8
Baltimore, MD	89.4	88.3	89.6	91.3	90.1
Beaumont, TX	90.4	89.7	92.3	89.6	81.0
Berkeley, CA	75.4	62.9	69.4	80.3	81.2
Birmingham, AL	91.8	91.5	92.8	88.7	86.4
Boston, MA	84.3	79.9	83.9	86.5	86.3
Bridgeport, CT	73.5	71.7	69.7	74.4	78.8
Buffalo, NY	87.3	84.2	86.5	89.5	87.9
Cambridge, MA	63.4	52.6	65.5	75.6	74.3
Camden, NJ	68.3	67.4	76.5	89.6	87.6
Canton, OH	83.8	82.4	81.5	89.3	89.9
Charleston, SC	86.8	86.5	79.5	68.4	60.1
Charleston, WV	77.0	74.3	79.0	79.6	80.3
Charlotte, NC	93.7	92.7	94.3	92.8	90.1
Chattanooga, TN	90.3	89.9	91.5	88.5	86.5
Chester, PA	82.7	82.2	87.4	88.1	85.1
Chicago, IL	93.0	88.8	92.6	92.1	95.0
Cincinnati, OH	84.2	83.1	89.0	91.2	90.6
Cleveland, OH	90.1	89.0	91.3	91.5	92.0
Columbia, SC	87.6	86.7	94.1	88.1	83.0
Columbus, OH	86.2	84.1	85.3	88.9	87.1
Covington, KY	89.1	86.9	87.8	85.0	80.6
Dallas, TX	95.9	92.7	94.6	88.4	80.2
Dayton, OH	91.1	90.1	91.3	93.3	91.5
Denver, CO	88.9	77.6	85.5	88.9	87.9

TABLE 1 (cont.)

CITY	1970	1970	1960	1950	1940
Des Moines, IA	83.7	79.2	87.9	89.3	87.8
Detroit, MI	82.1	80.9	84.5	88.8	89.9
Durham, NC	88.3	87.5	92.7	88.8	88.2
East Chicago, IN	80.9	79.0	82.8	79.6	74.5
East Orange, NJ	61.4	60.8	71.2	83.7	85.3
East St. Louis, IL	76.9	76.8	92.0	94.2	93.8
Elizabeth, NJ	79.1	75.5	75.2	76.1	75.9
Evanston, IL	85.3	78.3	87.2	92.1	91.5
Evansville, IN	90.3	88.6	91.2	92.4	86.2
Flint, MI	82.9	81.7	94.4	95.3	92.5
Fort Worth, TX	95.4	92.6	94.3	90.4	81.3
Galveston, TX	79.8	77.4	82.9	78.3	72.2
Gary, IN	83.5	82.9	92.8	93.8	88.3
Greensboro, NC	93.0	91.4	93.3	93.5	93.1
Harrisburg, PA	77.3	76.2	85.7	89.8	87.2
Hartford, CT	79.4	77.4	82.1	84.4	84.8
Houston, TX	92.7	90.0	93.7	91.5	84.5
Huntington, WV	89.3	85.9	88.8	85.8	81.6
Indianapolis, IN	89.6	88.3	91.6	91.4	90.4
Jacksonville, Fl	94.3	92.5	96.9	94.9	94.3
Jersey City, NJ	79.0	75.6	77.9	80.5	79.5
Kansas City, KS	87.0	84.7	91.5	92.0	90.5
Kansas City, MO	90.5	88.0	90.8	91.3	88.0
Knoxville, TN	92.2	89.6	90.7	89.6	88.6
Little Rock, AR	90.6	89.7	89.4	84.5	78.2
Los Angeles, CA	90.5	78.4	81.8	84.6	84.2
Louisville, KY	89.7	88.9	89.2	86.0	81.7
Macon, GA	90.3	90.2	83.7	77.1	74.9
Memphis, TN	92.4	91.8	92.0	86.4	79.9
Miami, FL	92.0	89.4	97.9	97.8	97.9
Milwaukee, WI	88.0	83.7	88.1	91.6	92.9
Minneapolis, MN	80.4	67.9	79.3	86.0	88.0
Mobile, AL	91.5	91.0	91.9	89.4	86.6
Montgomery, AL	93.6	93.2	94.7	90.5	86.8
Mt. Vernon, NY	79.5	78.4	73.2	78.0	78.9
Nashville, TN	90.3	89.0	91.7	88.7	86.5
Newark, NJ	76.4	74.9	71.6	76.9	77.4
New Bedford, MA	78.7	72.7	81.6	86.8	83.4
New Haven, CT	71.5	69.1	70.9	79.9	80.1
New Orleans, LA	83.9	83.1	86.3	84.9	81.0

TABLE 1 (cont.)

CITY	1970	1970	1960	1950	1940
New Rochelle, NY	75.1	70.7	79.5	78.9	80.6
New York, NY	77.3	73.0	79.3	87.3	86.8
Norfolk, VA	93.5	90.8	94.6	95.0	96.0
Oakland, CA	70.4	63.4	73.1	81.2	78.4
Oklahoma City, OK	95.6	81.8	87.1	88.6	84.3
Omaha, NB	89.6	85.6	92.0	92.4	89.5
Pasadena, CA	85.5	75.0	83.4	85.9	84.2
Paterson, NJ	72.0	70.3	75.9	80.0	79.8
Philadelphia, PA	84.4	83.2	87.1	89.0	88.0
Pittsburgh, PA	85.9	83.9	84.6	84.0	82.0
Port Arthur, TX	87.6	87.0	90.4	91.3	81.7
Portland, OR	86.2	69.0	76.7	84.3	83.8
Providence, RI	76.8	72.0	77.0	85.5	85.8
Richmond, VA	91.4	90.8	94.8	92.2	92.7
Roanoke, VA	92.7	91.8	93.9	96.0	94.8
Rochester, NY	76.5	73.8	82.4	86.9	85.5
Sacramento, CA	71.1	56.3	63.9	77.6	77.8
St. Louis, MO	90.1	89.3	90.5	92.9	92.6
St. Paul, MN	87.1	76.8	87.3	90.0	88.6
San Antonio, TX	89.7	81.8	90.1	88.3	79.6
San Diego, CA	85.6	71.6	81.3	83.6	84.4
San Francisco, CA	75.0	55.5	69.3	79.8	82.9
Savannah, GA	91.8	91.2	92.3	88.8	84.2
Seattle, WA	82.2	69.2	79.7	83.3	82.2
Shreveport, LA	97.8	97.4	95.9	93.2	90.3
Springfield, OH	81.9	81.1	84.7	81.6	80.9
Tampa, FL	92.0	90.7	94.5	92.5	90.2
Terre Haute, IN	87.4	82.5	90.1	89.8	86.6
Toledo, OH	89.1	86.7	91.8	91.5	91.0
Topeka, KS	79.4	74.1	83.5	80.7	80.8
Trenton, NJ	78.1	77.2	79.6	83.0	81.9
Tulsa, OK	94.5	76.4	86.3	91.2	84.6
Waco, TX	88.9	86.8	90.7	87.0	80.1
Washington, DC	78.8	77.7	79.7	80.1	81.0
Wichita, KS	93.0	85.0	91.9	93.3	92.0
Wilmington, DE	70.5	69.8	79.8	86.2	83.0
Winston-Salem, NC	94.6	94.0	95.0	93.8	92.9
Yonkers, NY	73.1	68.0	78.1	81.7	82.0
Youngstown, OH	75.9	74.9	78.5	83.5	80.0

REFERENCES

Gates, William.
    1972   THE USE OF MEDLIST TO DEFINE CITIES ON THE CENSUS THIRD COUNT
            or WHY WE HASHED THE MEDLIST FILE. Computing Report: Working Paper
            CR-1. Center for Demography and Ecology. Madison: University
            of Wisconsin.

Sørensen, Annemette.
    1974a   RETRIEVAL OF CITY-BLOCK RECORDS FROM THE 1970 CENSUS 3RD COUNT
            SUMMARY TAPES. Technical Documentation No. 734-016. Center for
            Demography and Ecology. Madison: University of Wisconsin.

    1974b   INDEXES OF DISSIMILARITY FOR RACE AND RACE BY TENURE GROUPS.
            Technical Documentation No. 734-017. Center for Demography
            and Ecology. Madison:University of Wisconsin.

Taeuber, Karl E., and Alma F. Taeuber.
    1965   NEGROES IN CITIES: RESIDENTIAL SEGREGATION AND NEIGHBORHOOD
            CHANGE. Chicago: Aldine Publishing Company.

U. S. Bureau of the Census.
    1970   1970 CENSUS USERS' GUIDE. Washington: U. S. Government Printing
            Office.

APPENDIX

DATA SOURCE AND RETRIEVAL PROCEDURE FOR CALCULATION

OF CITY SEGREGATION INDEXES FOR 1970

The primary data source for the present study is the computer

tape file from the 1970 Census, "Third Count Population and Housing

Summaries for City Blocks and Census Tracts in Blocked Areas." All

Third Count summary data for a given state are together on one logical

file on one or more reels of tape. The Third Count summary tapes have

two separate sequences of data, one for Block Publishing Areas and

another for Contract Block Publishing Areas. The first portion of the

file, for Block Publishing Areas, contains the data used in this study.

This portion of the file has data on blocks and tracts in urbanized

areas and some adjacent areas.[1] For each block and tract in the urban-

ized area, 38 tables are published. Five of these contain population

data and 33 contain housing data. One of these, Table 12, "Tenure and

Race of Head," was used for the computation of indexes of dissimilarity

for nonwhites/whites and other race and race-by-tenure groups. Table

A-1 in the report gives a listing of the data available in Table 12

of the Third Count summary tapes.

To make the 1970 segregation indexes comparable to the indexes

previously calculated for 1940, 1950, and 1960, data were needed for

blocks within each of 109 cities. Each block that was in a city of

interest to this study had to be identified. Numerical identification

codes are listed in the geographical field of each record of the Third

Count tapes. A list of the geographic information is given in Table
A-2 of this report. These geographic codes, singly or in combination,
are insufficient to determine for every block whether it is located in
a particular city within an urbanized area. Lack of a city code on
the Third Count summary tapes greatly complicates the task of obtaining
data for blocks within the central city. (The same problem confronts
the user of the published block bulletins for 1970.)

A computerized procedure was devised to retrieve data for city
blocks from the summary tapes. An account of the computer programming
involved in this retrieval is given in Gates (1972) and Sørensen (1974a).
Here a description of the logic of the procedure will suffice. A
diagram of the procedures is given in Figure A-1 of this report.

To retrieve data for city blocks from the Third Count summary
tapes, additional geographic information has to be obtained for each
block in the urbanized area. This information is provided by the
Master Enumeration District List (MEDList). The MEDList provides
a listing of geographic codes and names of different geographic areas,
ranging from the county to the block group level (see Table A-3).[2]

The MEDList does not provide a listing of codes and names for
blocks. Consequently, the unit for which additional geographic
information was obtained was not the block but instead the smallest
unit identifiable on both the MEDList file and the Third Count summary
tapes. This was either the block group or the census tract. Block
group records were used where available. For cities where no block
group codes had been assigned, tract records were used as the source
of additional geographic information.

FIGURE A-1

DIAGRAM OF RETRIEVAL PROCEDURE
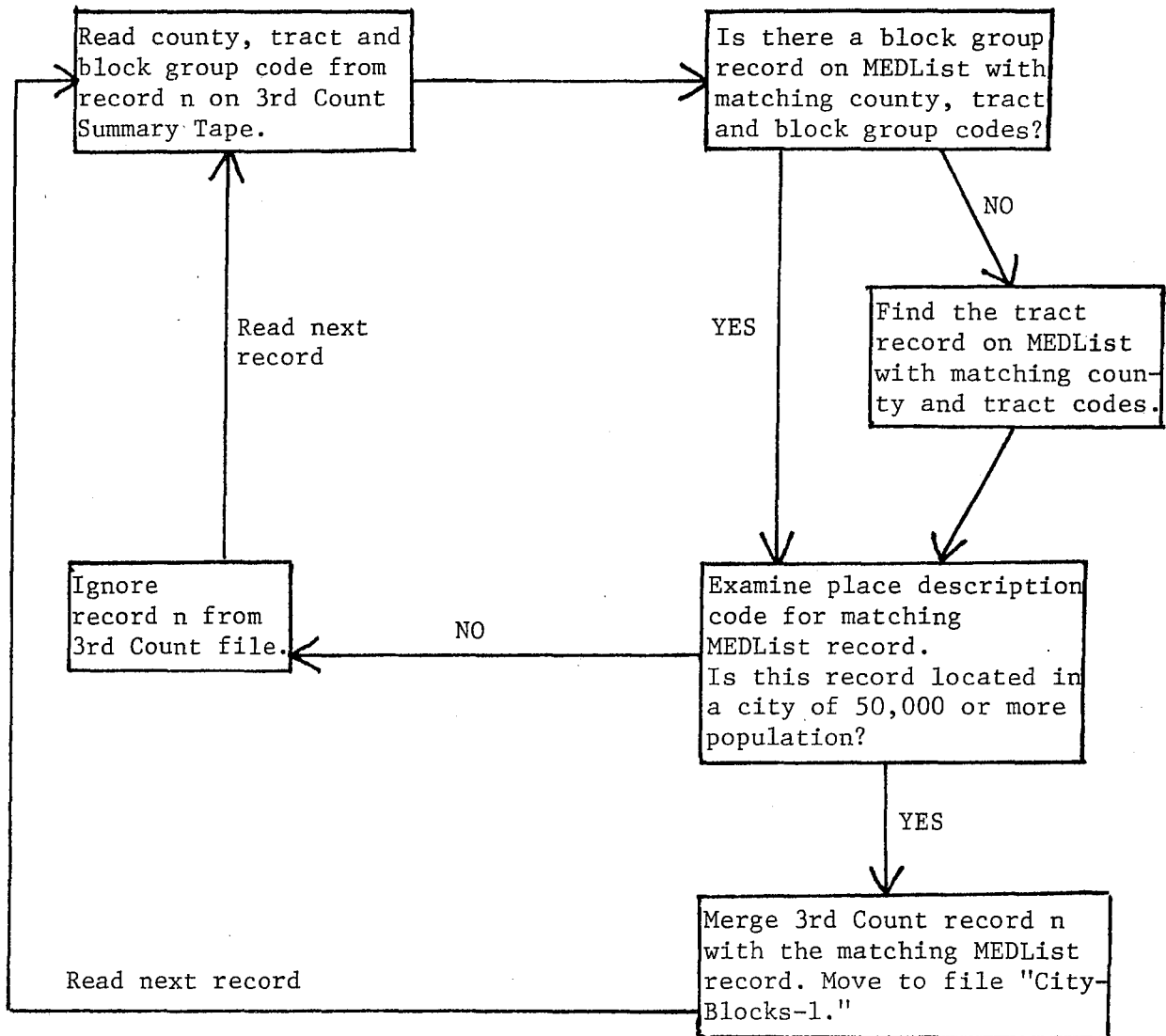
For each block record on the Third Count tapes, the MEDList record for the corresponding block group or tract was examined. The place description code (item 8 on the MEDList record) specifies whether the block group (or tract) is located within the central city of an urbanized area.[3] If so, both the Third Count and MEDList records for the block were merged into a file titled "City-Blocks-1."

Before proceeding with calculation of segregation indexes from the block data on City-Blocks-1, we conducted a check on whether the data on this newly created file were what we thought they were. The checking procedure is described more fully below, but its basic feature was a comparison between the number of households tallied for a city from City-Blocks-1 and the number of households reported for that city in the official Census count. There are several reasons not to expect an exact correspondence, but we were surprised to find for some cities that discrepancies between the two counts considerably exceeded 3 percent. We thought these large discrepancies probably arose from one particular source, the inclusion on City-Blocks-1 of an unknown number of households located outside of the city but in block groups or tracts that lie partially within the city.

For many cities, particularly those that are still growing by annexation, the boundaries of block groups and census tracts sometimes cross the boundaries of the central city. Indeed, the boundaries of the city may dissect individual blocks. The Bureau of the Census has--or could create--a file for each city containing only those blocks and portions of split blocks that are located within the city boundaries. The geographic identification scheme made available for

public use did not permit such precise delineation. To improve substantially our identification of the location of blocks it was necessary to follow a tedious clerical procedure.

Most of the cities with large discrepancies were cities for which blocks are grouped into census tracts but not into block groups. Because block groups are relatively small units (see footnote 2), inclusion of noncity blocks from the block groups that are split by city boundaries seldom caused a great increase in the number of households carried onto City-Blocks-1 as being in the city. Census tracts are larger and seem to be cut by city boundaries more often. Inclusion in our working file of all blocks in split tracts led in many cities to a serious inflation of the household count. The correction procedure is described by reference to census tract data, but again block groups were used if available.

A listing of all split tracts in a city subject to review was made from MEDList. The maps published in the "Block Statistics" bulletin for the city were examined visually. This is quite tedious because the maps for a large city are spread over 5 or more sheets, and all boundaries relevant to census use are portrayed by means of a complex symbol system. From this examination a list was made for each split tract of all blocks located outside the city boundary. (No correction was possible for blocks split by city boundaries; they are retained in their entirety on the City-Blocks files.) The list was punched for computer input and compared with the City-Blocks-1 list of blocks. All blocks identified in this way as being located outside the city were deleted from the file. The new file with revised extracts for 34

cities and the original extracts for 75 cities is titled City-Blocks-2.

The City-Blocks-2 file is clearly better for our purposes than the City-Blocks-1 file. But given the difficulties in its compilation and the unavailability of precise checking procedures, is City-Blocks-2 a satisfactory file for calculation of segregation indexes? Again, the basic checking procedure is a comparison between the tally of households from our created file and the official 1970 Census count for each city. Discrepancies between the two counts can arise from several sources; we were able to make approximate adjustments for two of these sources. The results of this checking procedure are presented in Table A-4.

One source of discrepancy is suppression of data for small areas. In its publications and on its public use tapes, the Bureau of the Census deletes fine detail that might permit a user to violate the Bureau's promise that individual information reported to it are confidential. On the Third Count tape and in the "Block Statistics" bulletins, racial identity and other occupancy traits are omitted for all blocks having fewer than five occupied housing units. The number of households omitted from City-Blocks-2 because of such suppression cannot be determined precisely. Instead we tallied from Table 10 of the Third Count tapes the total number of housing units (vacant and occupied) in suppressed blocks. This tally is given in our Table A-4, col. 2. This tally yields a slight overcount of the occupied housing units omitted from City-Blocks-2 because of suppression. There are three reasons for an overcount. First, not

all suppressed housing units are occupied. An adjustment will be
made for this, as described below. Second, some suppressed units
may be in split blocks, split block groups, or split tracts and may
not lie within the city boundaries. Third, some suppressed units may
be in blocks for which there is census error (see below), and we may
double-count them in our file evaluation procedure.

The Bureau of the Census has a massive data processing job in
handling the decennial censuses, and from time to time it detects
errors in previously published or released files. On page XII of
the "Introduction" to many of the "Block Statistics" bulletins
for 1970 there is a list of "Correction Notes." For our purposes
we distinguish two kinds of known census error. One is error
detected in the census file of block data, but of the sort that does
not alter totals for block groups, tracts, and cities. Blocks for
which this kind of error was detected are identified by a dagger
symbol in the bulletins. The "Correction Notes" indicate the correct
number of persons and of year-round housing units for each such
block. By a clerical procedure, making use of the block maps,
"Census Tract" bulletins, and information on how we had already handled
split tracts or block groups in each city, we tallied the approximate
number of year-round housing units erroneously omitted from or included
in the Third Count file and located within the city boundaries. This
estimate is given in Table A-4, col. 3. In the "Correction Notes"
in the "Block Statistics" bulletins and other census publications,
additional census errors are listed. These errors were not corrected
in the Third Count tapes or in the official household count for

each city that is the standard against which we compare the household

tally from City-Blocks-2. Hence, we ignore these errors, which

fortunately are quite small.

These estimates of housing units suppressed and of year-round

housing units omitted from the Third Count tapes (and hence also from

City-Blocks-2) are presented in columns 2 and 3 of Table A-4. These

estimates were converted to estimated numbers of occupied housing

units by multiplying by the ratio of occupied housing units to total

housing units or year-round housing units, respectively, using the

official 1970 Census counts for each city. The estimated numbers of

occupied housing units (households) from these two sources were

summed and the total then added to the tally of households given

in Table A-4, col. 1. This procedure yields an Adjusted City-Blocks-

2 Tally of Occupied Housing Units (Table A-4, col. 5). This Adjusted

Tally incorporates approximate corrections for census suppression

and census error.

The Adjusted City-Blocks-2 tally of households (Table A-4, col.

5) should resemble the census count of household (Table A-4, col.

4). The discrepancy between these two counts, expressed as a

percentage of the census count, is given in Table A-4, col. 6. For

26 cities the discrepancy is less than 0.05 percent, for 14 other cities

it is negative, and for 69 other cities it is positive. For 13

of the 109 cities the discrepancy is greater than 3 percent.

The discrepancies could not all be reduced to zero by any standard

adjustmant procedure. There are 6 different sources of discrepancy,

aside from any data processing errors that we may have made.

(1) The Third Count tapes and "Block Statistics" bulletins do not contain the information necessary to designate location inside or outside the city for those housing units in blocks split by the city boundary. Where the city boundary is not a street or major feature easily noted in the field, as when it is the line 100 yards south of highway N, there may be many split blocks with very long and rambling perimeters and a sizable number of households.

(2) We did not carry out for all cities the tedious process of determining which blocks in split block groups or split tracts lie within the city boundaries. The 35 cities for which this procedure was followed are identified in Table A-4 by an asterisk.

(3) For those cities for which we did check the maps and try to allocate individual blocks, there are undoubtedly some clerical errors and ambiguities caused by the welter of detail. Errors in the maps may have affected this task.

(4) The effects of suppression are estimated, not tallied precisely.

(5) The effects of known census error are estimated, not tallied precisely.

(6) There may be unknown census error.

The percentage discrepancy is the best guide we could devise to the question of whether our data processing procedures performed as intended. The choice of what level of discrepancy is tolerable is an arbitrary one. We wished to have a check that would assure us that (1) the right tapes were mounted, (2) the correct records from each tape were read and processed by the computer, and (3) the computer program had no major errors. We are satisfied that City-Blocks-2,

prepared from publicly released data, is reasonably accurate. It is

a good approximation to the near-perfect file that could have been develop-

ed by the Bureau of the Census were it to undertake computation of city

segregation indexes. On the basis of considerable experience in cal-

culating segregation indexes from computer tape files and by pencil and

paper methods from printed block bulletins, we believe that segregation

indexes from City-Blocks-2 would rarely differ by as much as one per-

centage point from the indexes that would be obtained from a more accurate

file.

The formula for the index of residential segregation is given

in the main text of this paper. Further information on computational

aspects is given in Sørensen (1974b).

TABLE A-1

CONTENTS OF DATA TABLE 12 FROM THE THIRD COUNT SUMMARY TAPES

TENURE AND RACE OF HEAD

Count of Occupied Units

By: Tenure (2)   By: Race of Head (3)

    Owner occupied:
           Total   (Includes white, Negro, and other races)

           White

           Negro

    Renter occupied:

           Repeat   Race of Head (3)

Source: 1970 Census Users' Guide, Part II, p. 3rd Count-11.

TABLE A-2

GEOGRAPHIC FILE OF THIRD COUNT SUMMARY TAPE

Item	Character(s)
Record Type (Urbanized Area)	1
1970 State Code	2-3
Padding	4-6
1960 State Code	7-8
Padding	9-24
Tract (Basic) Code	25-28
Tract (Suffix) Code	29-30
Padding	31-36
Central Business District Code	37
Padding	38-69
1970 County of Tabulation Code	70-72
Padding	73-84
Block Publishing Area (UA)	85-90
Summary for a tract	91-96
Padding	97-102
County No.	103-108
Tract No.	109-114
Block No.	115-117
Padding	118-119
$ Symbol	120

Source: 1970 Census Users' Guide, Part II, p. 3rd Count -6.

TABLE A-3

LISTING OF MEDLIST ITEMS

1. 1970 State Code
2. 1960 State Code
3. Federal Standard County Code
4. County of Tabulation Code
5. Central County Code
6. Minor Civil Division or Census County Division Code
7. Place Code
8. Place Description Code
9. Size of Place Code
10. Standard Consolidated Area Code
11. Standard Metropolitan Statistical Area Code
12. Urbanized Area Code
13. Tracted Area Code
14. Universal Area Code Prefix
15. Universal Area Code
16. State Economic Area Code
17. Economic Sub-Region Code
18. Central Business District Code (Blank in GACI)
19. Area Name
20. Basic Tract Code
21. Tract Suffix Code
22. Block Group Code
23. Enumeration District Code
24. Enumeration District Suffix Code
25. Urban/Rural Classification Code
26. Ward Code (ED records only)
27. Congressional District Code
28. Housing Count
29. Population Count

Source: Technical Document No. GT-1, National Data Use and Access
        Laboratories, July, 1971.

TABLE A-4

COMPARISON OF NUMBER OF HOUSEHOLDS IN TALLY FROM CITY-BLOCKS-2
WITH NUMBER OF HOUSEHOLDS IN OFFICIAL CENSUS COUNT, FOR 109 CITIES

City	From City-Blocks-2 Tape File		From Census Publications		Derived Figures	
	Tally of House- holds	Tally of Total Housing Units Suppressed	Estimated Error in Year-round Housing Units	Official Count of House- holds	Adjusted Household Tally From City-Blocks-2	Percent- age Discre- pancy
	(1)	(2)	(3)	(4)	(5)	(6)
Akron, OH	91131	507	0	91593	91616	0.0
*Asheville, NC	20704	291	354	20061	21310	6.2
Atlanta, GA	166253	1124	757	162291	168039	3.5
Atlantic City, NJ	19212	129	333	19561	19603	0.2
Augusta, GA	19744	190	0	19877	19922	0.2
*Austin, TX	79303	582	370	78570	80176	2.0
Baltimore, MD	287034	1732	1617	289349	290206	0.3
Beaumont, TX	36553	523	0	37027	37036	0.0
Berkeley, CA	44331	104	1298	45655	45682	0.1
Birmingham, AL	101509	918	161	99956	102533	2.6
Boston, MA	215002	815	2484	217622	218091	0.2
Bridgeport, CT	52761	218	0	52924	52972	0.1
Buffalo, NY	157679	163	98	157951	157927	0.0
Cambridge, MA	36261	130	23	36411	36409	0.0
Camden, NJ	32385	415	0	32565	32776	0.6
*Canton, OH	36916	1180	0	37146	38038	2.4
*Charleston, SC	21219	146	0	21213	21351	0.7
Charleston, WV	25242	523	110	25595	25839	1.0
*Charlotte, NC	77848	565	0	76992	78387	1.8
*Chattanooga, TN	40512	599	1198	40856	42186	3.3
Chester, PA	17754	126	0	17869	17874	0.0
Chicago, IL	1132015	1235	532	1137854	1133679	-0.4
Cincinnati, OH	162161	429	0	159838	162558	1.7
Cleveland, OH	245850	575	3019	248280	248289	0.0
*Columbia, SC	28256	277	31	27952	28542	2.1
*Columbus, OH	175846	568	835	173056	177177	2.4
*Covington, KY	17947	70	0	18134	18012	-0.7
*Dallas, TX	279880	1581	743	280993	282033	0.4
*Dayton, OH	81290	363	70	81597	81704	0.1
Denver, CO	184353	851	167	185331	185327	0.0

TABLE A-4 (cont.)

City	(1)	(2)	(3)	(4)	(5)	(6)
Des Moines, IA	69920	517	128	68506	70531	3.0
Detroit, MI	490911	854	7173	497753	498463	0.1
*Durham, NC	30294	357	596	30097	31198	3.7
East Chicago, IN	14620	24	0	14639	14642	0.0
East Orange, NJ	28591	29	0	28618	28619	0.0
East St. Louis, IL	21304	166	424	21516	21842	1.5
Elizabeth, NJ	38486	606	0	38575	39079	1.3
Evanston, IL	26832	67	0	27173	26898	-1.0
Evansville, IN	46100	333	0	46404	46414	0.0
Flint, MI	60672	272	0	60931	60930	0.0
*Fort Worth, TX	131910	1588	382	129960	133750	2.9
Galveston, TX	20935	196	0	21024	21114	0.4
Gary, IN	51326	282	0	51598	51594	0.0
*Greensboro, NC	43829	392	1385	43696	45534	4.2
Harrisburg, PA	25368	391	0	25742	25727	-0.1
Hartford, CT	54818	89	1245	56024	56096	0.1
*Houston, TX	397141	2932	1749	393555	401447	2.0
*Huntington, WV	25835	159	0	26117	25985	-0.5
Indianapolis, IN	238342	1640	692	235772	240520	2.0
Jacksonville, Fl	159019	2422	486	161666	161716	0.0
Jersey City, NJ	86752	80	981	87853	87766	-0.1
Kansas City, KS	56361	341	0	54896	56685	3.3
Kansas City, MO	179999	1209	0	176373	181108	2.7
*Knoxville, TN	58059	676	0	57059	58691	2.9
*Little Rock, AR	45014	527	18	45294	45523	0.5
Los Angeles, CA	1023628	2160	116	1027374	1025799	-0.2
*Louisville, KY	123217	389	886	122683	124423	1.4
*Macon, GA	38882	437	533	38359	39787	3.7
Memphis, TN	189778	568	1085	190006	191364	0.7
*Miami, FL	119685	483	0	120393	120149	-0.2
Milwaukee, WI	235535	546	1985	236981	237973	0.4
Minneapolis, MN	160673	489	0	161141	161144	0.0
*Mobile, AL	57150	512	114	56938	57729	1.4
*Montgomery, AL	42107	225	0	41569	42314	1.8
Mt. Vernon, NY	24975	66	0	25051	25040	0.0
Nashville, TN	139167	737	566	140409	140409	0.0
Newark, NJ	120082	5205	801	121041	125787	3.9
New Bedford, MA	35155	279	0	35423	35425	0.0
New Haven, CT	46067	142	559	46741	46737	0.0
New Orleans, LA	189206	731	0	191363	189877	-0.8

TABLE A-4 (cont.)

City	(1)	(2)	(3)	(4)	(5)	(6)
New Rochelle, NY	24400	63	0	24463	24462	0.0
New York, NY	2783880	3540	-17	2886872	2787298	-1.7
Norfolk, VA	83664	959	2591	86742	87047	0.4
Oakland, CA	136435	452	2638	138831	139361	0.4
Oklahoma City, OK	119545	2358	2773	126954	124251	-2.1
Omaha, NB	110786	1346	2994	111223	114879	3.3
*Pasadena, CA	45070	252	0	44653	45309	1.5
Paterson, NJ	46729	1126	1042	47746	48827	2.3
Philadelphia, PA	631870	1458	3189	642145	636301	-0.9
Pittsburg, PA	176430	953	718	178016	177997	0.0
*Portland, OR	145739	1589	889	145082	148103	2.1
Port Arthur, TX	18262	217	0	18455	18461	0.0
Providence, RI	62800	664	0	63148	63415	0.4
Richmond, VA	81605	671	587	82769	82801	0.0
Roanoke, VA	31464	698	0	31928	32127	0.6
Rochester, NY	100965	283	0	101238	101236	0.0
*Sacramento, CA	91165	395	1046	91697	92533	0.9
St. Louis, MO	210833	499	0	215479	211284	-1.9
St. Paul, MN	103701	437	0	104128	104123	0.0
*San Antonio, TX	191857	1367	1177	190692	194244	1.9
*San Diego, CA	232060	794	996	227006	233746	3.0
San Francisco, CA	293002	289	1977	295174	295157	0.0
*Savannah, GA	37677	567	86	37824	38295	1.2
Seattle, WA	204892	1013	275	206092	206088	0.0
*Shreveport, LA	58108	405	784	58439	59209	1.3
*Springfield, OH	27037	85	0	26731	27118	1.4
Tampa, FL	93221	1754	143	94889	95006	0.1
*Terre Haute, IN	23234	439	0	23446	23649	0.9
Toledo, OH	126557	364	282	125364	127178	1.4
Topeka, KS	43030	434	478	41991	43906	4.6
Trenton, NJ	33377	1374	0	33546	34687	3.4
Tulsa, OK	111957	1506	810	112792	114101	1.2
*Waco, TX	31557	657	642	31504	32747	3.9
Washington, DC	259001	511	3332	262538	262625	0.0
Wichita, KS	94211	544	607	92751	95280	2.7
Wilmington, DE	26713	330	0	27565	27017	-2.0
*Winston-Salem, NC	43605	315	79	42634	43979	3.2
Yonkers, NY	68898	209	0	69093	69103	0.0
Youngstown, OH	43412	301	1055	44655	44713	0.1

TABLE A-4 (cont.)

Notes: Column 5 is derived from Columns 1, 2, and 3 by a procedure described in the text.

Column 6 = 100 [(Col. 5 - Col. 4) ÷ Col. 4].

* Identifies cities for which a special adjustment was made to allocate blocks in split block groups or split tracts.