## STATISTICS AND THE URBAN REVOLUTION

## Albert Mindlin, Management Office, Government of the District of Columbia

Ι

A convenient way to grasp the present state of metropolitan affairs from the viewpoint of the government statistician is to perceive our urban centers and their governments as being in the throes of two major revolutions. The first revolution is the sheer growth in size of urban centers, together with the tremendous increase in a) the Negro population in the central cities and b) the suburban commuting population. The second revolution is that of computer technology, which potentially can fundamentally transform some basic municipal administrative practices. I will discuss each of these revolutions in turn.

Most people by now are fairly familiar in a general way with the vast population shifts that have been going on in our country in recent years -- the movement of most of our people to metropolitan centers, the increasing concentration of Negroes in the central city, and the vast white migration to the suburbs. The actual statistics of these movements may be less known. They are worth calling to your attention. Let us compare the 1950 and 1960 censuses for the 25 largest cities in 1960. greater family income of the suburbs reflects the fact that these migrants were mainly middle class white families.

As Table 3, bank 1 shows, short of largescale migration of non-whites to the suburbs, or massive return of whites to the central cities, the central cities are bound to be increasingly non-white by natural increase alone, as the female fertility ages become increasingly non-white. This shift in childbearing females is also shown in bank 2. The proportion of childless families remained about the same during the decade for white families in the central city, but decreased substantially for non-whites. In the central cities there has been a substantial increase in the number of children in proportion to the total population (bank 3), and an increasing proportion of these children are non-white. There has also been a relative increase, less dramatic, in the number of aged (bank 4). These are the ages that are the least productive and the most dependent on community services. There has been just as great a proportional increase of children in the suburbs. But the higher incomes of the suburbs make

Table 1 -- Median Family Income

central city (cc)	1950	5MSA	urban fringe
1950 1960		1960	1960
\$3,526 \$5,935	\$3,476	\$6,545	\$7,082

SMSA = Census Bureau's Standard Metropolitan Statistical Area

Median family income between 1950 and 1960 increased substantially in both central city and the entire SMSA. But it increased more in the SMSA than in the central city. Since the SMSA includes the central city, it is clear that the increase in the suburbs was substantially greater than that shown for the SMSA as a whole. This is brought out explicitly in the 1960 comparison between the urban fringe, which is that part of the suburbs to which most people moved, and the central city. The median family income here was nearly 25% higher than in the central city.

Ta	bl	е	2
ıa	DТ	e	<u> </u>

Total Number	960 minus 1950 suburbs	
white	-346,000	2,733,000
non-white	450,000	112,000

Table 2 dramatically shows the racial change in the central cities and the flight to the suburbs of white families. The relatively them much better able to pay for the necessary community services than the central cities.

Drastic changes are also occurring in the character of the non-residential sector of many metropolitan areas. In the core city new construction is increasingly in office buildings. There has been a relative decline in retail merchandising in the core city, corresponding to the phenomenal growth of the suburban shopping centers. This has produced a mounting crisis in the classical "downtown" section of the central city.

## II

The changes outlined above have created vast new metropolitan problems. Some of these are:

a. The critical commuting problem.

b. The massive problem of urban renewal in slums and deteriorating areas.

Local and Federal government together have

		cc	suburb <b>s</b>
1. Females, ages 15-44, 1960/1950	white	0.79	1.51
	non-white	1.34	1.27
Non-white proportion of total	in this		
age group	1950	0.14	0.05
	1960	0.21	0.04
2. Families without children unde	r 18:		
Proportion of all families			
white	1950	0.53	0.45
	1960	0.51	0.38
non-white	1950	0.56	0.49
	1960	0.44	0.43
Non-white families proportio	n of all		
families in this group	1950	0.13	0.04
	1960	0.17	0.03
3. Persons under 18 years of age:			
Non-white proportion of all	persons		
in this age range	1950	0.16	0.05
	1960	0.23	0.04
This age range as proportion	of		
all persons	1950	0.26	0.30
	1960	0.33	0.38
4. Persons 65 years and over, pr	oportion		
of all persons	1950	0.08	0.07
	1960	0.10	0.07

Table 3 -- Median Values, 25 Largest Cities in 1960, and their metro areas

NOTE: In 16 of the 25 cities there were boundary changes in central city or SMSA or both between 1950 and 1960. Changes in the 9 cities with unchanged boundaries were almost the same as for all 25 cities. Thus geographic changes were not a significant factor in these phenomena. Furthermore the changes for the cities taken altogether, where the bigger cities count more than the smaller cities, are roughly the same as when the cities are treated equally by the use of medians. Thus experiences of the larger cities are similar regardless of size differences.

been wrestling with these problems. There are vast highway and urban renewal programs, joint ly financed. Both local and Federal interest in interurban mass rapid transit systems have greatly increased.

But efforts to solve these problems have created new problems. a. The highway program and the rapid transit program compete with each other for attention, priority, and for rights-of-way. This competition has created one of the liveliest, at times bitter, controversies in contemporary urban politics. At stake are competing philosophies about how people should live, competing economic interests, competing views on the tax base of the community, and competing arguments on how to keep the city alive. b. Large urban renewal and highway construction programs have created a sizeable problem of displacees. So the city is faced with a big new headache -- relocation.

c. With an increasing proportion of the core city's population in the young and old dependent age groups, community services in health, education, welfare and recreation are inevitably increasing, with an attendant spiraling of budget and taxes. These increasing taxes rest on a proportionally diminishing base of income-earning producers.

d. This proportionally diminishing base of income earning producers in the central cities is proportionally increasing non-white and poor. Clearly an improvement in the quality of community life depends on an accelerated improvement in the productivity and income of these people. There is now going on a dramatic shift in emphasis in the civil rights movement, a joining hands of that movement with the war on poverty. Only about one fourth of the nation's families with income under \$3,000 in 1960 were non-white, but in the large cities poverty has become increasingly a non-white problem. Here the war on poverty and the civil rights movement are merging into a single new frontier of social reform.

e. Many urban problems are metropolitan-wide in character. They cannot be solved without close cooperation of the many independent jurisdictions that compose the modern metropolis. Transportation is only the most visible problem. Another is water -- supply, siltation, pollution, and sewage. A third is the preservation of park land and open spaces. A fourth is the racial distribution of the metropolis. The suburbs so far have largely resisted the pressure of out-migrating middle class Negroes. This resistance cannot be maintained much longer without developing serious friction between core city and suburbs.

III

What is the role of the government statistician in this urban revolution? This paper will not discuss the statistician's role in internal management improvement. The panoply of statistical technics such as quality control, work sampling, operations research, and other statistical tools are every bit as applicable to local and regional management as they are to Federal Government and industry management; and indeed it is by such internal management applications that a statistician can often play his most prominent role. But the focus of this paper is on research design and data needs of community programs -- local, regional and Federal.

There is great overlap in the data needs of urban social and economic programs.

a. A great deal of demographic data are needed -- population by age, sex and color; family composition; education; housing data such as condition and overcrowding; place of work related to place of residence.

b. Economic data are also needed -- personal and family income; housing costs, trans portation costs, and medical costs.

c. The geographic distribution of land use is needed -- where different kinds of households live, where different kinds of non-residential uses are located; and how these are moving.

d. Most of these data are needed for small areas -- blocks, census tracts, neighborhoods.

e. Most programs require projections into the future of some sort, so that current and continuing data are needed.

f. For effective regional cooperation, the data are needed with comparable definitions and reconcilable, if not standardized, classification systems.

It is the responsibility of the urban statistician 1)to locate as much of this information as possible; 2)to generate as much of it as possible; and c)to design and carry out statistical research that will provide valid underpinning of community programs -and equally important altho less recognized, to test proposals and programs for usefulness and validity, to the extent that they are statistically testable.

IV

It is at this point that we turn to the second major revolution that is overtaking the local governments -- computer technology. There is no question that the new technology can revolutionize municipal administration -and is actually on the threshold of doing so in many communities. To the statistician this revolution is partly real, partly mythical -- and it is important that the myths be separated from the reality, for there is much wishful thinking on this subject.

When the statistician seeks data to solve community problems, if he does not generate original data where can he turn? Aside from the periodic national censuses, the most obvious information sources are records of operating agencies, both local and national. It is when we examine closely these records in relation to our needs that we perceive the real and the mythical aspects of the computer revolution.

A. The real revolution. An overwhelming barrier to effective statistical use of local records is that the bulk of them are not mechanized. Operating agencies as a general rule mechanize only the items which are critical for the regular operations of their program. Only rarely can such mechanized data fit, even reasonably well, the needs of other agencies and projects for which they were not intended, designed, or gathered.

Much extremely valuable information which is obtained almost incidentally as part of an operating program is left unmechanized. For example, many jurisdictions gather land-use and structure-use information as part of their assessment program -- landuse classification, number of living quarters in the structure, housing condition and other extremely valuable items -- but do not mechanize it. At times, needed information is laboriously copied from these records by hand; at other times, original field surveys are made to obtain information already present in inaccessible assessment records; most often, action programs that would bene-fit greatly by this information simply do the best they can without it.

The new technology has opened up new horizons for the statistical utilization of ad-

ministrative records. To continue the example of the last paragraph, for the first time mechanized assessment files can become widely accessible by being taped. A deck of hundreds of thousands of punch cards is not a viable tool to be used by many different agencies or to be statistically manipulated --but three or four reels of tape are. On the computer we can apply statistical methods with mass data; and on tape it becomes feasible to add to the mechanical record many items not previously mechanized. No operating agency will automate primarily to satisfy statistical research or the needs of community programs. It will automate only when that is shown to be a more effective method of conducting operations. But once automated, it is far easier and cheaper than ever before to add in the extra information so important to other community programs. Furthermore it becomes feasible to bring together in a single record information about the same property or person from different agencies. The vision of a "data bank", so dear to the hearts of planners and some statisticians, has become a real possibility.

B. <u>The mythical revolution</u>. The possibilities of the new technology are so exciting and the pressures of salesmanship and "machine politics" so great, that a number of myths have arisen. These must be dissipated.

One myth is that there is great duplication in local government records that will be eliminated by an integrated electronic data processing system. In general, neither part of this assertion is true. There is not great duplication in records, and what duplication there is will not for the most part be eliminated by automated processing. For the most part municipal agencies do not duplicate substantive information. The Fire Dep't., Health Dep't., Assessment Office, License Office, Housing Code Division, and Highway Dep't., do not maintain on their records the same substantive information about a piece of property. They operate different programs and maintain different information. They do maintain the same non-substantive information for identification -- block, lot, census tract, etc. But this duplicated information would have to be maintained even under an automated system in order to query or enter the central system for a unique record.

Another myth much more important to the statistician is that the bulk of our statistical problems can be solved with records now collected by operating agencies. A third myth is that the main methodological problem is simply how to integrate these records in a mechanical system and make them statistically available. Both of these are indeed serious myths that have created much delusion among administrators and non-statistical program planners and researchers. In fact, even if most presently gathered administrative information were mechanized and accessible, this would not, by itself, ade-

quately satisfy many, including some of the most critical, statistical needs. Almost all information gathered by local agencies, either public or private, is designed and tailored to serve primarily operating programs. In spite of the seemingly large amount of information gathered, some of the largest and most pressing needs of non-operating agencies, such as urban renewal offices, planning commissions, and mass transit planning agencies, cannot be met, or can be met only minimally, very inefficiently, and unsatisfactorily, by data regularly gathered by operating agencies. The same statement holds for problems constantly coming before the operating agencies themselves.

Two examples will suffice. 1)Much useful information about housing condition, overcrowding, etc. is often gathered by a municipal housing agency. But aside from licensed premises, such as apartments, and specific programs in designated urban renewal areas, this information is usually generated only by complaints. Hence it is fitful, irregular and spotty, and therefore largely unusable for probability sampling or for estimates. 2)The small area identification of place-of-work in relation to residence, so important to transportation, city planning, and civil defense, simply does not exist because it is not part of any operating program.

A much more central problem than making accessible current records, is the absence of professional statistical competence. Most municipal and regional governmental agencies do not have the statistical know-how to either generate or analyze statistical data in a sophisticated or efficient way. There is a drastic paucity of well-qualified professional statisticians in most metropolitan areas. For example, many statistical needs that cannot be filled by existing records can be filled by properly designed sample surveys. But the professional talent to design sophisticated sample surveys is often lacking. All of the automating in the world will not produce non-existent input data and non-existent statistical talent.

v

In light of all of the above, what are some genuinely realizable elements of a modern local statistical program? The following proposals are advanced, in no special order of importance.

1. There should be created a professional demographic unit to serve the entire governmental apparatus. This unit should make periodic population estimates by age, sex, color and various other groupings. These estimates should be for areas within the city as small as can be professionally supported. This unit should be responsive to requests for demographic assistance from any governmental agency, e.g. to put together all available census material for a small area and to relate such material meaningfully to local material such as vital statistics and welfare data. This unit must have professional demographers and be centrally located, not dominated by any single Departmental program.

2. There should be developed a regionwide set of comparable subject-matter classification systems -- comparable both for agencies within a jurisdiction and between jurisdictions. For example, three areas where lack of comparable data has hampered not only regional programs but even programs in a single jurisdiction, are land use, juvenile delinquency, and housing condition.

3. There should be developed a master sample survey design of properties that can be used repeatedly and that can form the basis for longitudinal studies. The full design should allow reliable neighborhood estimates.

4. There should be maintained a comprehensive statistical library. This may well be a responsibility of the demographic unit, which would probably use it more than anyone else. This library would maintain the local area output of the Census Bureau, Bureau of Labor Statistics, Internal Revenue Service, and other Federal agencies; would also maintain all other obtainable tabulated data from local public and private sources of more than purely internal workload interest; and would constitute a responsive reference service that would tell an inquirer readily what organization has what data, how complete they are, how current, how available, and how much they cost.

5. There should be developed a comprehensive locational directory. This directory would cross-classify the most important ways that a given location is identified -- for example, address, lot, city block, census block, census tract, census enumeration district, neighborhood, and police precinct. More generally, a feasible master coding scheme might be developed that would identify any property to a very small geographic building block, usable to approximate most geographic classifications. The master code need not be introduced into operating records, but could exist as a tape or deck providing a master switchboard between one classification and another. Thus all available data for an area -- local, state and national -- could be brought together. Indeed research is now going on in several places to locate all properties on a fixed coordinate system that would be independent of any area changes. Property coordinates can be entered into a mechanical data plotter which mechanically constructs data maps.

6. Leadership should be exerted in mechanizing or automating those governmental records that are the most valuable for community research. In carrying forward this task, it is well to perceive clearly two different kinds of records -- property records and social records. It is much easier to automate the first kind than the second, hence most progress has been made in the first area.

a. <u>Property records</u>. Almost every community has one operating file that is geographically universal -- its property assessment file. In most large communities this file is already mechanized to the extent necessary to calculate and distribute property tax bills. However, as stated earlier, the non-mechanized records usually contain a statistical cornucopia of information. The most important of these items should also be mechanized. Such mechanization would serve two fundamental purposes: 1) It would provide critically needed statistical information, and 2) probably of much greater long-range value, it would provide various bases for stratifying and classifying properties in order to draw samples for field surveys. An appropriately mechanized assessment record file would be of inestimable value as a sampling frame for small area samples, or samples of various types of property. Indeed such a mechanized file could virtually eliminate the more cumbersome classical method of area sampling in residential surveys. Furthermore tape reels of this file, which is usually all public information, distributed to a Planning Commission, Urban Renewal Office, Highway Dep't., Mass Transit Agency, and even public utilities and real estate institutes, Board of Trade, would enormously improve community research.

b. Social records. These are records about people and families from the Welfare Dep't., Health Dep't., Police Dep't., Juvenile Court, Unemployment Compensation Board, School Dep't., etc. This is a more difficult problem, but one that is receiving increasing attention. The creation of an integrated file of such records would stimulate effective and intelligent social planning to an extent that cannot be exaggerated. Here the computer technology opens possibilities previously closed. In particular it offers solutions to two of the stickiest problems --1) the necessity to develop identification search rules that will assure the proper match of records about the same person or family; 2) the preservation of confidentiality of information, which is entirely feasible on a computer.

7. Probably the greatest single step forward that could be accomplished for urban statistics is the establishment and adequate staffing of a professional statistical unit available to the entire local and regional governmental complex for consultation, design, analysis, and other statistical research.

VI

The Federal Government could exert great leadership in helping solve the statistical

problems produced by the urban revolution. Federal assistance to improve local statistical programs will directly and greatly benefit the Federal Government itself, because many of the new programs are largely Federally financed. Indeed, in 1960 - 1961 the Federal Government spent the astounding amount of nearly \$7 billion on grants-in-aid programs administered by State and local governments. The current anti-poverty program will add millions more to this. Below is a too-brief, highly incomplete, and superficial discussion of some things the Federal Government could do.

1. A mid-decade population and housing census: It is probably too late to mount a mid-decade census now, but this issue must be kept alive. The national census is by all odds the greatest single source of local information. Our metropolitan growth is simply too dynamic to have to rely on this most fundamental of all statistical tools only once every ten years. There have been hundreds of pages of Congressional testimony on this subject, indeed probably more than on any other Federal statistical subject. There is nothing unprecedented about a shorter census period. Canada, Japan, New Zealand, and Turkey have five-year censuses, and the United Kingdom is about to have its first. The opposition of the Bureau of the Budget must be overcome. The last half of an intercensal decade is the dark ages for urban research. This is the most important single statistical step the Federal Government could take to advance local and regional community planning.

2. Demographic improvements in the census: Experience and new needs have indicated several areas of improvement in the population and housing census data, e.g.:

a. Housing condition: This is one of the most important and least reliable census items. Efforts to improve it must continually be made. In addition to improved condition reporting itself, the information would be more useful if it applied to a structure rather than to a housing unit.

b. Value of structure: Ways should be found to extend this beyond the present highly limited owner-occupied-one-housing-unit detached structure.

c. Block statistics bulletin: This should be greatly expanded.

d. Census blocks: These should be made coterminal with city blocks in order better to integrate census and local data.

If the Census Bureau abandons the enumeration district, as it currently contemplates doing, points c and d become both critical and feasible.

e. There should be a much more refined work-residence item -- if not on a census tract level, at least on some gross neighborhood level.

f. There should be much finer crosstabulations by color. There is now going on a most interesting reversal of thinking on this subject. For many years pressure mounted to remove color from all kinds of records, including census, because color was not only offensive but was often used for racial discrimination. This campaign was successful in many areas. This trend is now being reversed. The civil rights movement has become so pervasive that it now perceives that statistical information by color is one of its strongest weapons for carrying forward its objectives -something that most statisticians have long known. For example, it is impossible to make adequate studies of employment discrimination in local areas without detailed color crosstabulations by occupation, industry, wages, educational attainment, age, and class of worker -- all of which are unobtainable in present census publications.

3. There is a cornucopia of important local information locked in the records of some Federal operating agencies. These agencies must recognize responsibilities beyond their immediate operating programs, and make their data more easily available for community research. Two of the most important agencies to which this applies are the Bureau of Old Age and Survivors Insurance and the Internal Revenue Service. One feasible way to do this is to mechanize one or two more locational codes for ready access to tabulated data on a small area basis. With respect to the two mentioned agencies, another way is a vigorous and imaginative effort to integrate their records. A third is to develop regular series of local statistics from agency data, as the IRS has done in its SMSA "Statistics of Income on Individuals" -- preferably for central city and suburbs separately, since these are often separate political jurisdictions.

4. The Dep't. of Labor, which publishes extremely useful employment and earnings information for SMSA's, should if possible develop separate estimation programs for central cities.

5. The Federal Government could greatly stimulate urban statistical research by an expanded program of planning, demonstration and research grant funds. The variety of possible relevant and important statistical research projects is too numerous to list. Two examples of demonstration projects currently being funded by the Urban Renewal Administration are the study of automation of local records for use by planning commissions centered in Tulsa, Oklahoma, and the effort to develop a reconcilable and operable set of land use and housing classification systems over an entire metropolitan area centered in Washington, D.C.

6. Finally, the Federal Government should explore ways of bringing to local

communities professional statistical assistance which they sorely need. A string of Federally assisted metropolitan statistical centers (not data centers) across the country should go a long way toward channeling our sprawling urban revolution into reasonably ordered and controlled progress.