Volume Title: Regional Income

Volume Author/Editor: Conference in Research in Income and Wealth

Volume Publisher: NBER

Volume ISBN: 0-870-14177-5

Volume URL: http://www.nber.org/books/unkn57-3

Publication Date: 1957

Chapter Title: Development of Postcensal Population Estimates for Local Areas

Chapter Author: Henry S. Shryock, Jr.

Chapter URL: http://www.nber.org/chapters/c7608

Chapter pages in book: (p. 377 - 400)

Development of Postcensal Population Estimates for Local Areas

HENRY S. SHRYOCK, JR., BUREAU OF THE CENSUS

Over the half century that the Bureau of the Census has been concerned with the problem of making population estimates, we have noted a growing demand for current population estimates for counties, cities, and other local areas. More and more public and private agencies have expressed their needs, and the kinds of detail for which such estimates are needed have also been expanding. Unfortunately, the supply of good estimates has not kept up with the demand. Few new methods have been developed. The adoption of the relatively better methods has proceeded slowly. Many agencies are trying to produce estimates, but there is little coordination of their efforts.[1] The Bureau of the Census has not had the resources to go beyond the publication of population estimates for states, but it has carried on some experimental work for large cities and counties.

Experimental Work at the Bureau of the Census

Part of this experimental work has consisted of applying to these areas the methods that it uses in making its official postcensal estimates of state population. The resulting estimates for 1950 were then compared with the decennial census figures. For comparison, estimates were also computed by several other commonly used methods.

DESCRIPTION OF METHODS

The methods developed or used by the Bureau of the Census have been fully described elsewhere, and detailed illustrations are available in print.[2] The method that has usually given best results involves the separate estimation of migration and natural increase and is called, for convenience, "migration and natural increase,

NOTE: The author wishes to acknowledge the assistance of Benjamin Greenberg, Statistician, Population Estimates and Projections Branch, Population Division, Bureau of the Census.

[1] Henry S. Shryock, Jr., *Coordination of Population Estimates Used by Federal, State, and Local Agencies*, Bureau of the Census, Current Population Reports, Population Estimates, Series P-25, No. 81, October 12, 1953.

[2] Benjamin Greenberg, *Illustrative Example of a Method of Estimating the Current Population of Subdivisions of the United States*, Bureau of the Census, Current Population Reports, Population Estimates, Series P-25, No. 133, March 16, 1956.

method II" (to distinguish it from a simpler variation called "migration and natural increase, method I").

Migration and Natural Increase, Method II. In method II, the numbers of registered births and deaths are corrected for underregistration. The net migration rate for each state is estimated from the comparison of the observed and expected elementary school enrollment. On the basis of studies of gross interstate migration by age, it is assumed that the net migration rate for the population of all ages is in the same direction, but 1.2 times as large as the corresponding rate for the population of elementary school age. The sum of the estimated net migration over all states is then adjusted to equal an independent estimate of net migration from abroad based on records of the Immigration and Naturalization Service. Allowance is also made for net loss of civilian population to the armed forces.

The shortcomings in this method have been discussed in an article by Lawrence and the writer.[3] The sources of error may be classified into those arising from errors or inconsistencies in the basic data and those arising from the assumption about relative rates of net migration of the child population and the total population.

A basic difficulty is presented by the variability in time and space of the ratios between the net migration of different age groups. Both the 1 to 1 ratio used before 1950 (and in the tests reported below) and the 1.2 to 1 ratio now being used represent national average experiences. Despite the high variability of the ratios about these averages, they have been applied to all areas (states, cities, counties, etc.) alike. Even if we used the ratio last observed for the particular area in question, we should still need to make an assumption about how the ratio had changed in the meantime. The Bureau of the Census is now investigating several different bodies of data on interstate migration by age to see if there are patterns that would permit more flexibility in the basic assumption of method II. A more promising approach, however, seems to be the compilation of series of data that reflect the migration of adult age groups. This possibility will be discussed later in this paper.

Migration and Natural Increase, Method I. Method I differs from method II in that net migration is estimated from school enrollment in a simpler way. In this method, it is estimated on the basis of the difference between the percentage change in the population of school age in a given area and the change in the United States. The

[3] Henry S. Shryock, Jr. and Norman Lawrence, "The Current Status of State and Local Population Estimates in the Census Bureau," *Journal of the American Statistical Association*, June 1949, pp. 157–173.

Bureau has not actually published estimates based on this method; but, since it is much less time-consuming than method II, there has been some interest in it.

Vital Rates. The "vital rates" method was developed by Donald J. Bogue and, in combination with method II, has been used since 1950 by the Bureau of the Census for its state estimates.[4] It involves (1) computing two provisional estimates of population by applying (a) estimated birth and death rates to (b) the number of births and deaths, respectively, and (2) averaging the two provisional results. The estimated birth and death rates are obtained from the known rates for the United States in the postcensal year and the relationship between the rates for the United States and for the area in question in the preceding census year.

Other Methods. The "natural increase" method makes use of birth and death statistics, but assumes that postcensal net migration was equal to zero. Arithmetic and geometric progression probably do not require definition here.

Averages of Methods. The methods that involve an average of the results of two independent methods remain to be discussed and, perhaps, justified. Other tests made prior to those presented here had shown that, when the results of two independent and relatively accurate methods were averaged, the average error of the averages tended to be lower than that of either separate set of estimates. One benefits from the fact that, when a positive error from one method is paired with a negative error from the other method, their mean error may often be lower than either original error. It may be argued that sometimes the error of the averaged results may be higher for a particular area than that from the preferred method used alone. On the other hand, the general tendency is clearly toward improved accuracy of estimation. Furthermore, one tends to get rid of the extreme errors—a very important consideration in the actual use of the population estimates.

RESULTS OF TESTS

Table 1 presents a comparison of the errors in the results of the various methods as obtained by checks against the 1950 census counts. This is a longer time span since the last census—ten years—than we normally have to deal with in making postcensal estimates. Method II (as used here) represents the procedure used before 1950. Since then a few changes have been introduced.

[4] Donald J. Bogue, "A Technique for Making Extensive Population Estimates," *Journal of the American Statistical Association*, June 1950, pp. 149–163.

TABLE 1

Summary of Percentage Deviations from the 1950 Census Count of Population Estimates by Various Methods, for Cities of 100,000 or more. Metropolitan Counties, and Standard Metropolitan Areas

Area and Measure	*Migration and Natural Increase* I	II	*Vital Rates*	*Average of II and Vital Rates*	*Arithmetic Progression*	*Geometric Progression*	*Natural Increase*
Cities (92):							
Average deviation	8.34	6.53	9.33	4.93	9.60	9.33	7.73
Quadratic mean deviation	9.86	8.52	12.56	7.25	12.15	11.75	10.03
Deviations of 10 per cent or more	29	18	31	7	33	31	26
Deviations of 5 per cent or more	65	50	59	34	65	64	56
Positive deviations (46 expected)	30	22	79	63	11	13	56
Counties (102):							
Average deviation	9.21	6.57	6.29	4.77	18.25	16.42	15.66
Quadratic mean deviation	12.41	8.87	8.08	6.17	21.58	19.46	20.01
Deviations of 10 per cent or more	41	21	19	9	73	70	56
Deviations of 5 per cent or more	69	52	52	37	95	92	74
Positive deviations (51 expected)	48	49	69	68	2	4	12
Metropolitan areas (32):							
Average deviation	5.99	3.70	5.45	3.39	15.05	13.80	10.90
Quadratic mean deviation	7.91	4.97	6.52	4.34	16.83	15.53	14.07
Deviations of 10 per cent or more	7	2	3	0	23	21	13
Deviations of 5 per cent or more	16	8	16	8	30	29	19
Positive deviations (16 expected)	17	14	28	25	0	0	1

Source: Bureau of the Census (see the text for a brief description of methods).

An analysis of the errors of the city estimates, including the statistical significance of the differences, has already been published.[5] The tests for metropolitan counties and standard metropolitan areas, both with central cities of 250,000 or more, are presented here for the first time. Omissions from the pertinent list of areas represent cases for which the necessary school data were not readily available.

In terms of the average deviation (disregarding sign), it may be seen that method II compares favorably with the other independent

[5] Jacob S. Siegel, Henry S. Shryock, Jr., and Benjamin Greenberg, "Accuracy of Postcensal Estimates of Population for States and Cities," *American Sociological Review*, August 1954, pp. 440–446.

methods in all three types of local areas, although not all the differences are statistically significant. The vital rates method performed relatively well for large standard metropolitan areas and their individual counties, but not very well for cities. The improvement introduced by averaging the results of these two leading methods is also apparent. The last three methods in the table are obviously inferior; but two of them are nonetheless widely used.

The next three measures in the stub (quadratic mean deviation, deviations of 10 per cent or more, and deviations of 5 per cent or more) were all employed to get at the relative frequency of extreme error. A method having a slightly larger average error may still be preferred if it yields fewer extreme errors, since it will be a "safer" method to use for a particular area. The various methods rank roughly the same way by these measures as they did by the first, however.

The number of positive deviations is shown as a measure of directional bias in the estimates. For example, it can be seen that projection of the absolute changes of the 1930–1940 decade (arithmetic progression) always resulted in 1950 estimates for standard metropolitan areas that were too low. In the case of our state estimates, the process of adjusting to an accurate national total eliminates any directional bias for the set. When estimates are made for all counties in a state, they can be adjusted to add to the independently estimated state total. This process should usually reduce the average county error.

The 1950 estimates for large standard metropolitan areas by the best methods are about as accurate as those for states. It can be seen that average errors for the parts of metropolitan areas tend to be higher than for the totals. How good postcensal population estimates need to be is an open question; and the answer depends, of course, on the use to which they are to be put. The consumers, as well as producers, of such estimates are in a position to exercise judgment on this point and to act accordingly in their choice of methods (when a less accurate but adequate method is cheaper than a more accurate one). Where population estimates are used only in the computation of per capita income figures, the accuracy of the income estimates is clearly relevant to the problem. If, for example, the income estimates are subject to an average error of 10 per cent, one need not put in a lot of extra effort to reduce the average error of the population estimates from 5 per cent to 3 per cent.

In making an estimate for a particular local area, is it wise to use a method that has worked well in the past for that area even though its over-all performance is relatively inferior? Here one needs to

know whether the given method is likely to continue to give good results for the area or whether its past good performance was simply a random fluctuation. In other words, some knowledge of the relationship between the peculiar nature of the area and the peculiar nature of the method is required. In any case, one "success" should not be given too much weight. On the other hand, where good data symptomatic of population change are available for some areas but not for others, there is no virtue in consistency of method.

Let us return to the tests previously described to see what factors are associated with accuracy in the population estimates for local areas. Table 2 deals with the factor of population size at the begin-

TABLE 2

Average Percentage Deviation from the 1950 Census Count of Population Estimates by Various Methods, for Cities of 100,000 or more and Metropolitan Counties, by Population Size

Area and Population Size	*Migration and Natural Increase* I	II	*Vital Rates*	*Average of* II *and Vital Rates*	*Arithmetic Progression*	*Geometric Progression*	*Natural Increase*
Cities							
Larger (46)	8.05	6.32	9.06	4.22	10.27	9.97	7.86
Smaller (46)	8.63	6.73	9.60	5.63	8.92	8.69	7.61
Counties							
Larger (51)	7.81	4.21	5.53	3.23	14.06	12.99	10.78
Smaller (51)	10.62	8.94	7.05	6.31	22.45	19.85	20.54

Source: Bureau of the Census.

ning of the period (1940). When the cities and counties are each divided into two groups on this basis, we find that better results were obtained, on the average, for the larger cities and counties than for the smaller ones. This directional difference is consistent for the better methods but not for the poorer ones. Furthermore, the factor of size seems to be more important for counties than for cities. On the basis of this evidence, it seems safe to say that average errors by these methods would be even larger for cities under 100,000 and probably for nonmetropolitan counties.

From Table 3 it can be seen that, with only one exception, the average error by each method was greater for those areas with the more rapid rate of population change from 1940 to 1950. The superiority of the better methods was more pronounced for these more rapidly changing areas than for those whose populations were more static.

Several persons have pointed out to us that method II tends to give estimates that are much too low for suburban areas undergoing

TABLE 3

Average Percentage Deviation from the 1950 Census Count of Population Estimates by Various Methods, for Cities of 100,000 or more and Metropolitan Counties, by Rate of Change

Area and Rate of Change [a]	*Migration and Natural Increase* I	II	*Vital Rates*	*Average of II and Vital Rates*	*Arithmetic Progression*	*Geometric Progression*	*Natural Increase*
Cities							
Larger (46)	8.34	7.45	10.51	5.79	14.91	14.37	9.69
Smaller (46)	8.34	5.61	8.15	4.06	4.28	4.29	5.78
Counties							
Larger (51)	11.91	8.92	6.60	6.10	27.11	23.66	26.11
Smaller (51)	6.25	4.22	5.97	3.43	9.39	9.17	5.21

[a] Regardless of direction.
Source: Bureau of the Census.

their major settlement phase. Some of these critics have given this situation as a reason for not adopting the method at all. The downward bias in the estimates undoubtedly results from the fact that movers to new suburban developments tend to be young couples with children of preschool age. Hence, there is a lag of several years before these migrant families are reflected in the school statistics. This problem was examined for all methods by dividing the metropolitan counties into those containing and those not containing the central city. The latter may be regarded as suburban counties, although some of them are of the type that has already passed through its major settlement phase. The results of this comparison are presented in Table 4.

TABLE 4

Average Percentage Deviation from the 1950 Census Count of Population Estimates by Various Methods, for Central and Suburban Metropolitan Counties

Type of County	*Migration and Natural Increase* I	II	*Vital Rates*	*Average of II and Vital Rates*	*Arithmetic Progression*	*Geometric Progression*	*Natural Increase*
Central (39)	5.84	4.11	5.69	3.25	13.21	12.23	9.28
Suburban (63)	11.30	8.10	6.66	5.71	21.38	19.01	19.61

Source: Bureau of the Census.

The average errors show that not only method II but also all the other methods gave better results for metropolitan counties containing a central city than for suburban counties. In the latter group, the vital rates method may be a better choice than method II, but it also seems desirable to average the results of the two methods.

Since suburban counties tend to have smaller populations than central counties and also grow more rapidly in the last intercensal decade, Tables 2 to 4 deal with overlapping factors. If we subdivide the 102 metropolitan counties simultaneously by all three factors, we find that the average error of the population estimates by method II was 3.36 per cent for the twenty-six large, slowly growing, central counties and as much as 11.04 per cent for the thirty-two small, rapidly growing, suburban counties. In fact, with all of the other methods, smaller average errors were obtained for the same central counties than for these suburban counties. In the latter group of counties, the vital rates method was more accurate, on the average, than method II (7.24 versus 11.04 per cent).

Tests on West Virginia Counties

Recent estimates have not been made and tested for nonmetropolitan counties by the Bureau of the Census. Experimental estimates for 1950 were prepared at the National Office of Vital Statistics, however, for the fifty-five counties of West Virginia.[6] Two variations of method II were tested, one using enrollment in grades four through eight to estimate net migration, and the other, enrollment in grades four through ten.[7] The results are summarized in Table 5.

TABLE 5

Summary of Percentage Deviations from the 1950 Census Count of Population Estimates by Two Variations of the Migration and Natural Increase Method, for West Virginia Counties, by Size in 1940

Measure	USING GRADES 4 TO 8			USING GRADES 4 TO 10		
	All Counties (55)	*25,000 and over* (23)	*Under 25,000* (32)	*All Counties* (55)	*25,000 and over* (23)	*Under 25,000* (32)
Average deviation	5.53	4.20	6.48	4.75	2.68	6.25
Deviations of 10 per cent or more	6	0	6	8	1	7
Deviations of 5 per cent or more	28	8	20	22	4	18
Positive deviations	21	5	16	35	13	22

Source: Adapted from a study at National Office of Vital Statistics.

[6] "Study of Population Estimates Made for Each County in West Virginia, as of April 1, 1950," processed, National Office of Vital Statistics, presented by Robert D. Grove at the Third Annual Meeting of the Public Health Conference on Records and Statistics, Washington, D.C., April 23, 1951. The errors were recomputed using final instead of preliminary census figures.

[7] "According to the original plan of the study, independent estimates were to be made by use of data on telephone installations, elcctric, gas, and water meters, building permits, and other related data. Considerable time was spent in trying to collect such information. The results were disappointing. Only fragmentary data could be obtained which seemed unusable" (*ibid.*, p. 1).

Better estimates, on the average, were obtained for these West Virginia counties than were obtained by a similar method for the counties in large metropolitan areas—despite the smaller size of the former. Part of the explanation may lie in the fact that the West Virginia counties tended to have moderate rates of population change during the 1940's compared with the metropolitan counties, the respective medians (disregarding sign) having been 7 and 29 per cent. The maximum error by either method for a West Virginia county was 18 per cent for one with 5,000 inhabitants.

Estimates from Symptomatic Data and from Sample Surveys

In the fall of 1946 and in April 1947, the Bureau of the Census conducted sample surveys of a number of metropolitan districts. (These samples contained between 3,100 and 4,200 dwelling units.) Estimates of the total population of the central cities were prepared from these survey results. As a check, independent population estimates were also prepared by migration and natural increase method II. After the 1950 census figures became available, intercensal estimates were made for these dates and the original postcensal estimates were compared with them. Of course, these intercensal estimates do not constitute as valid a standard of comparison as the census figures themselves, but deviations from them do represent rough measures of accuracy. In general, it can be said that the estimates by method II were at least as accurate as those from the sample surveys. (Of course, these surveys provided many other data in addition to total population.)

Work of State Agencies

In response to a resolution passed at the Public Health Conference on Records and Statistics in Washington in March 1954, the Bureau of the Census made a mail canvass of state agencies to find out about the work they were doing on population estimates.[8] A questionnaire was sent to seventy-eight state agencies and to five independent city health departments. Of the seventy-five that replied, fifty-eight indicated that they are engaged in making population estimates at the present time. All but seven states have at least one state agency preparing estimates for local areas. In another nine, there are two or more agencies making such estimates. The following types of agencies are active:

[8] *The Current Status of Population Estimates Prepared by State Agencies,* Current Population Reports, Population Estimates, Series P-25, No. 116, June 6, 1955.

Department of health	37
University bureau of business research	8
Employment service	2
Other	11

Bureaus of business research have entered this field relatively recently, but appear to have some of the most extensive programs.

The report shows (1) whether population estimates are made, (2) the agency making the estimates, (3) the method used, (4) the areas and population characteristics shown, (5) the date of latest estimates, (6) whether estimates are published, and (7) whether unpublished estimates are available upon request. The following methods were used to make population estimates for counties:

Annual census		1
Migration and natural increase		16
Census Bureau's method I	2	
Census Bureau's method II	8	
Other	6	
Combination or selection		5
Involving migration and natural increase	2	
Other	3	
Natural increase alone		6
Adjusted to state estimate from Census Bureau	1	
Not adjusted	5	
Censal ratio		5
One series	4	
More than one series	1	
Proration of Census Bureau's state estimate by a current series		3
Proration of state estimate by 1950 distribution		2
State estimate from Census Bureau	1	
Other state estimate used	1	
Arithmetic projection		9
Other		3
Not reported		1

Perhaps the only additional method that requires definition is the "censal ratio." This method involves (1) a computation of ratios

of census counts of population to symptomatic data (school statistics, births, utility consumers, etc.) for each county, and (2) the application of these ratios to postcensal symptomatic data to obtain the county population estimates.

Not much work on validation of these methods was reported by these state agencies, and only a few results of such accuracy tests were transmitted.

Utility Data and Building Permits

Greater use of economic series symptomatic of population change may have been made for cities than for counties. Such data, which include residential electric, gas, and water meters, residential telephones, and building permits for residential units, are rarely available for rural areas. In past decades, the utility series were too much affected by business conditions, technological changes, and wartime materials shortages to be anything but notoriously poor indicators of population growth. In areas where use of a given utility has become fairly universal and where business conditions have remained relatively constant (no pronounced cyclical changes), utility data may be more useful. Building permits can be used to estimate urban population growth with some prospect of success if allowances are made for such factors as the lag between permit and completion, conversions, demolitions, vacancies, and quasi-household population. Some assumption must be made about the average size of both old and new households, and a good deal of variability exists here in both time and space. Unfortunately, very little validation has been reported for population estimates based on utility and housing data.

Robert C. Schmitt has reported, on tests against 1950 census counts, on various symptomatic series by both the Proration and Censal Ratio methods for Washington counties.[9] The series used were births, deaths, school enrollment, auto registration, registered voters, and welfare recipients. Average errors were fairly high, but were reduced appreciably by several methods of averaging the results of the independent methods. He also checked, against the 1950 census counts, population estimates based on building, conversion, and demolition permits for the seventy-nine census tracts in Seattle and found an average error of 8.5 per cent.[10]

[9] Robert C. Schmitt, "Short-Cut Methods of Estimating County Population," *Journal of the American Statistical Association*, June 1952, pp. 232–238.

[10] Robert C. Schmitt, "Estimating Current Populations of Census Tracts," *Sociology and Social Research*, September–October, 1952, pp. 12–15.

Prospective Developments

Perhaps enough has been presented about the absolute and comparative accuracy of estimation methods now being used at the Census Bureau for the reader to decide whether these methods are adequate for his own particular purpose. The Bureau has collected the necessary source materials to prepare estimates for standard metropolitan counties, their component counties, and large cities for years since 1950. The necessary manpower to complete the computations and checking has been lacking, however. Meanwhile, we should be pleased to see more local use made of these methods in preference to demonstrably inferior methods. At the same time, we are actively exploring additional methods for both states and local areas.

SPECIAL CENSUSES

Any discussion of population estimates would be incomplete without mention of more frequent censuses, for which the estimates are merely inadequate substitutes. Unfortunately, Massachusetts and Kansas are the only survivors of a large number of states that once conducted their own censuses. The Bureau of the Census will take a special census at the request and expense of a local government. Since April 1950, the Bureau has conducted over five hundred such censuses for municipalities ranging in size to that of Los Angeles and for a few entire counties.[11] Almost all of these censuses were purchased because the state concerned accepted the certified result as official and the local government stood to gain financially by this recognition. Probably many more such censuses would be requested if it were felt that the usefulness of the results in planning and research were worth the cost.

SOCIAL SECURITY DATA

It was previously mentioned that the Bureau of the Census is now investigating some bodies of data that may reflect the migration of adult age groups. Estimates of net migration for adults could be combined with estimates of net migration for children (based on school data) in migration and natural increase, method II. Postcensal population estimates by broad age groups may also be possible.

Covered Workers. To this end, a cooperative project is now under way with the Bureau of Old-Age and Survivors Insurance.

[11] *Special Census of . . .*, Current Population Reports, Special Censuses, Series P-28, various numbers and dates.

As a pilot study, changes in state of employment between the beginning and the end of 1951 have been tabulated for covered wage and salary workers from the 1 per cent continuous work history sample. Possibly a similar tabulation can be made for a sample of self-employed workers. Some of the shortcomings of estimates of net migration for all persons eighteen to sixty-four years old are:

1. Sampling error
2. Differences between state of employment and state of residence (In the record linkage of forms from the Internal Revenue Service and the Bureau of Old-Age and Survivors Insurance, planned to begin shortly, information on post-office address will be available.)
3. Differences in net migration rates of covered workers and of all adults (It will be possible to compare the difference at the national level in terms of *gross* interstate migration, however; and perhaps an adjustment factor can be drawn from this source.)
4. Migration estimates applicable to calendar years (For July 1 estimates, interpolations would have to be made.)
5. Delay in the tabulations of the O.A.S.I. records (This arises partly from 4. According to present arrangements, the tabulations for a given calendar year would not be available until near the end of the next calendar year.)

Beneficiaries. Finally, to round out the coverage of broad age groups, the Bureau of Old-Age and Survivors Insurance is experimenting with tabulations of annual changes of state of residence for O.A.S.I. beneficiaries sixty-five years old and over. These tabulations would be on a 100 per cent basis, and the lag would be minimal. State of residence, however, has to be obtained from mailing addresses grouped by field offices, some of which cross state lines.

Composite Estimates. Thus, from several different bodies of data, net migration could be estimated by a composite, additive method. It would not be necessary to generalize from the experience of only one broad age group (children of elementary school age). If this approach is found to yield much better postcensal estimates of total population (and of broad age groups) for states, the next step would be to apply it to local areas. Here, except for a few large metropolitan areas, the 1 per cent sample of workers aged eighteen to sixty-four would be much too small. The problem of a larger sample is mostly a problem of added cost and should not be insoluble. Differences between place of work and place of residence also become more important as we proceed from states and metropolitan areas

to counties and cities. The proposed change in the O.A.S.I. record system may remove this particular difficulty, however.

Summary and Conclusions

This paper describes the most widely used methods of making current estimates of total population for local areas. Particular attention is devoted to those methods that can be used for counties as well as for cities. Comparisons of test estimates with census results indicate that methods involving the separate estimation of natural increase and net migration, with the latter estimated from *current, symptomatic data,* tend to be more accurate than other methods. Averaging the results of two or more independent methods, all of the same rough order of accuracy, also tends to improve the average estimate of a set. Whether or not a particular method is adequate for a particular kind of area depends on the use to which the estimates are to be put; but there is obviously room for improvement, for some purposes, in any of the methods now available.

Included in the paper is a summary report on an inventory, conducted by the Bureau of the Census, of the activities of state agencies in this field. Fifty-eight agencies reported programs of current population estimation, but seven states had no program. Most of the work is done in the department of health, but eight university bureaus of business research make population estimates for postcensal dates. Some kind of migration and natural increase method is used in sixteen agencies for making county estimates. On the other hand, nine agencies use arithmetic progression, and six agencies merely add the natural increase to the 1950 base.

If adequate funds can be found, the best way of obtaining up-to-date population figures is from a special census. Special censuses are especially appropriate for rapidly growing local areas where estimation methods tend to give the poorest results, but where public interest tends to be greatest. As far as estimates themselves are concerned, the most promising approach seems to the writer to be the improvement of the Census Bureau's migration and natural increase, method II. The greatest present weakness of this method is that the net migration of other age groups has to be deduced from the estimated net migration of children of school age. Hence, if the net migration of the other age groups could be estimated from independent, current, symptomatic data, substantial improvement should be possible. The Bureau of the Census is now working with the Bureau of Old-Age and Survivors Insurance on several cooperative projects that would measure, respectively, the net interstate migra-

tion of covered workers and of aged beneficiaries. Given a larger sample of O.A.S.I. records for covered workers, the procedure could perhaps be extended to some types of local areas.

A great deal of money and manpower is being expended in this country on population estimates by federal, state, and local agencies —although perhaps not so much as their usefulness would justify. There is some duplication of effort and considerable use of obviously inferior methods. If the present level of effort and expenditure could be brought to bear in a coordinated program, I believe that much more accurate data would result.

COMMENT

JOHN N. WEBB, University of Florida

I have selected three statements of Henry S. Shyrock, Jr. that seem to me important in relation to the future course of postcensal population estimates for local areas:

1. During the past half century the Bureau of the Census has noted a growing demand for current population estimates in local areas.
2. All but seven states now have one or more state agencies making population estimates.
3. The accuracy of estimates would be improved by a coordinated program.

Why is it that the decennial population counts, which for so long a time served the need for local information, are no longer adequate? Why is it that, according to Shryock, some five hundred localities have felt it worthwhile to lay out the very considerable sums required for special censuses so soon after the 1950 enumeration?

The answer, I think, lies in the changing relation of the local area —town, city, and county—to the state. The local administrative unit is no longer so self-supporting, so independent, as it once was. While many cities are growing outward physically, counties as well as cities are moving in the opposite direction financially. Centralization is a trend that legislators may deplore, but steadily, and under local pressures, legislatures enact laws that transfer responsibility from localities to the state.

Little wonder that this is true! The local property tax is no longer able to bear the entire financial burden of the schools, the roads, the welfare services demanded today. Special local taxes discrim-

inate against local business and industry, or overlap with state levies. The answer to growing local needs is to tap the general revenue that flows from the broader tax base of the state.

The first step in stating the local case for state aid is often to point to population increase. In fact, size of population is, in many instances, the whole case for state assistance. And it is not the size one, two, or five years ago; it is the size *now*.

Let me illustrate the current trend in terms of my own state, Florida. In recognition of the inadequacy of the local real property tax to support municipal services, and in recognition also of the discriminating effect of special local sales taxes, the state collects a 5 cents per package tax on cigarettes and refunds this money monthly to the incorporated towns and cities. Based on similar reasoning, one-half of the license fee for the sale of alcoholic beverages (a fee graduated by county population size) is refunded to corporate communities. Part of the state-imposed gasoline tax comes back to the counties. Since 1947, the state, out of general revenue, has underwritten a minimum program for public schools, thereby making expenditures for public schools the largest of the biennial appropriations. The administration of the public schools remains in county hands, but the state pays most of the bill. How much each county gets depends in large part on how many students it has, which in turn depends upon population. The number of circuit judges, the salaries of local officials, county planning programs, and fringe area annexations are other local concerns that require population counts currently, and the list could be lengthened.

I suggest, then, that in stating a trend based upon the experience of the Bureau of the Census during recent years, Shryock has also predicted the trend for the future: a demand for, or perhaps better, the necessity for more frequent local population counts and estimates.

Secondly, there is the problem connected with the many local intercensal estimates of population noted by Shryock. These estimates may not agree with estimates for the same areas computed by the Bureau of the Census for one of two reasons: (1) the methods may be different (a fact brought out by Shryock's report on the mail canvass of state agencies making estimates); and (2) the estimates may differ when the same method is used but different local data are used, or the same data are used differently. To take Florida again as an illustration, our estimates for the state, constructed by the method Shryock refers to as migration and natural increase, method II, differ from those published by the Bureau of the Census because we estimated white and nonwhite populations separately.

Further, the school attendance series we used was not the one used by the Bureau of the Census.

Our estimate was considerably *less* than that published by the Bureau of the Census. We believe the difference comes mainly from estimating white and nonwhite populations separately (an idea we got from reading the census bulletin describing method II [1]). The migration rate for white population is high and positive, but the rate for the nonwhite is low, or even negative. Nonwhite population stands at about 22 per cent of the total and so can introduce a considerable divergence in the results.

Also, the State of Florida gathers and publishes three school attendance series: average daily attendance, original entry and transfer, and attendance at the end of term. Using Shryock's method II and the period 1940–1950 (because of census counts at both ends), we tried each of these series for the state and for selected counties. The average daily attendance figure consistently gave the best results in all parts of the state and for both color groups. The original entry series, which on a priori grounds should be most sensitive to migration, proved to be most unreliable in the largest population centers.

We asked the state Board of Education statisticians for their views on this outcome. They were cautious, but they pointed to the fact that the very large sums of state money supplied to the county public school systems are allocated on a basis of average daily attendance in each school.

Whether this series is actually a better one for use in making local population estimates than the one the Census Bureau selected may be debated. But there can be no doubt that it is the one most subject to constant public scrutiny, and this may be good for the accuracy of a series. Recently this series and its compilation made the headlines of most Florida newspapers because in one county a school teacher told the press that she was being forced to pad her reports by allowing absentees to make up their absences. The story was that the school superintendent approved, if he did not actually originate, the idea of make-up, because under the minimum foundation program of state aid, one "instructional unit" (read "teacher") was allowed for each twenty-seven units of average daily attendance in the school involved, and the school needed more teachers. This incident had an immediate effect. Newspapers in most of the counties assigned reporters to see whether or not average daily attendance records were

[1] *Illustrative Examples of Two Methods of Estimating the Current Population of Small Areas,* Current Population Reports, Population Estimates, Series P-25, No. 20, May 6, 1949.

being padded in their areas. No other instance was found. It is my belief that a statistical series subject to this kind of scrutiny is likely to stick pretty close to facts in self defense.

I cite this instance as an illustration of why local area estimates may differ because the local agency selects a series different from the one used by an agency outside the state. Of course, many local estimates will not be duplicated by other agencies, and therefore no differences can appear. But since the Bureau of the Census regularly makes and publishes estimates for each of the forty-eight states, there is the possibility of discrepancies whenever a state agency makes an estimate for the state directly, or indirectly, by adding county estimates, or estimates for other subdivisions of the state.

I have no doubt that the general public will accept the Bureau of the Census estimate and question the state agency estimate if they differ. But if you know of Florida's present pride in the rapidity of its population growth, ranking as it does near the top along with California and Arizona, you can recognize the handicap under which we would operate in publishing an estimate for Florida that was lower than the one published by the Bureau of the Census.

Intercensal population estimates for local areas represent a new application of the statistical art. The ingenuity of the workers in this field should produce rapid improvement in methods and in the reliability of some of the basic data from which estimates are derived. Without the leadership of Shryock and his co-workers at the national level, we would not only be poorer in methodology but, more importantly, we would lack a clearing house for the discussions of our problems and for seeking advice.

This brings me to my third and final point: the need for a coordinated program. It is a necessity if we are to maintain and extend public confidence in our local estimates. Coordination also offers a sound basis for a desirable division of labor. The national level has great advantages in estimating population change for the country as a whole, and to a somewhat lesser degree this advantage extends to estimates for each of the forty-eight states. The state level has, I believe, an advantage in estimating the population of local areas.

I would like to explore this point briefly because of a statement Shryock made in discussing the accuracy of the several estimating methods when examined by size of area (his Table 2). "When the cities and counties are each divided into two groups on this basis [size in 1940], we find that better results were obtained, on the average, for the larger cities and counties than for the smaller ones. . . . Furthermore, the factor of size seems to be more important for counties than for cities."

I can confirm Shryock's finding of the relation of accuracy to size in county estimates from some tests we have just completed for Florida counties with the period of estimate April 1950 to July 1954. We made population estimates for each of the sixty-seven counties in Florida by the migration and natural increase, method II, by the vital rates method, and by one other method. These county estimates were for white and nonwhite populations separately, and for the two color groups combined. For controls, we used an estimate of the state as a whole made by each method and the July 1, 1954 estimate of Florida population published by the Bureau of the Census. In general, our results agree with Shryock's statement: the larger the county, the better the agreement in the estimates.

Method II, however, was not clearly superior to the other methods when we added up the separate county estimates and compared these additions with the estimates for the state. We suspect that the elaborate adjustment required by method II to arrive at a migration rate is the reason for the erratic results in the small counties; and we have nearly twice as many counties with less than 10,000 inhabitants as we have counties with 50,000 and over.

So we tried what seemed to be the simplest method of relating population change to school enrollment. We began with the 1950 census data and computed the ratio of school enrollment in grades one through eight to the total population, by color, for each county and for the state as a whole. This ratio was then applied to the school enrollment figures for the estimate year to obtain an estimate of population. In aggregate, these estimates overshot the estimate published by the Bureau of the Census, but by no more than did the results from some of the more complicated methods.

We then examined the rate of increase of population and the rate of increase in school enrollment shown by the Bureau of the Census and by the Florida State Department of Education. In Florida, school enrollment is increasing more rapidly than population for the white group and less rapidly for the nonwhite group. This gave us a correction factor for the ratio of school enrollment to population by color. When the county estimates were recomputed with this correction, our state aggregate of county estimates differed from the census estimate by less than 15,000 where the estimated total was approximately 3.5 million. This was by far the best estimate we have made.

I report this experiment only to illustrate what I think state agencies are likely to be doing in a search for new methods or for adjustments to old methods that will provide consistent estimates under widely varying local conditions. In this area of inquiry, the

state agencies have an advantage in the development of estimates for cities and counties. But state agencies need the controls, the overall view, that Washington agencies can provide. A coordinated program would provide the way and the incentive for an exchange of views and for a discussion of problems. Given this, I feel certain that we can advance rapidly toward the improvement in accuracy that Shryock mentions in the closing lines of his excellent paper.

ORMOND C. CORRY, The University of Tennessee

The background of investigative work upon which Henry S. Shryock, Jr. draws for his paper is substantial. His position is probably best summarized as follows: There is a choice of method when the available source data, desired accuracy, and feasibility, or financial resources of the estimator, are weighed; past performance and logical considerations favor the Census Bureau's method when accurate estimates are needed. His application of census data in tests of the estimates by various methods against enumerated data, his Tables 1 to 5, and the accompanying discussion should prove valuable reference material for many years.

As a person working in a university bureau of business research, I am most interested in the relatively void field of current population estimates for the counties and cities by state. Much work must be done on improving source data, specific applications of the estimating methods, and testing of results before reliable current estimates are available for all local areas in most of the states. The estimates should be made for use as general purpose data. Thus, they should be made by the method promising the most reliable results. While the level of accuracy obtained by the Census Bureau for the state current population estimates is probably too much to expect, no estimating program can hope to survive unless it includes safeguards against extreme errors.

The problems of estimating current population for counties parallel those met by the Census Bureau in making the state estimates, and the state estimates are available to provide independently determined totals comparable to the national totals used for the state estimates. If acceptably reliable estimates are sought, the Census Bureau method should be accepted as a flexible estimating procedure adaptable to varied situations.

The current population estimates by county differ from population forecasts, since they are postcensal projections relating to a date already passed, usually recently, for which symptomatic data are available. The estimation process followed in the Census Bureau

method is based on the truism that the current population is equal to the census or base-date population plus changes to the current date. Only changes that are large enough or volatile enough to influence the accuracy of the total population estimates are involved.

The estimates of change required at the state level can be grouped into four components of population net change: (1) natural increase or the net of cumulative births and deaths over the estimate period, (2) the net difference in the armed forces personnel stationed in the state at the enumeration and estimate dates, (3) the net difference in the national total armed force personnel whose prior-service residence was in the state at the enumeration and estimate dates, and (4) civilian net migration of residents of the state over the estimate period. Changes in institutional population and changes in college enrollment at the county level may be volatile enough to require a total of six components of change. The net migration component is the most difficult one to estimate; the others may require only the direct use of source data.

The Census Bureau equation becomes elaborate when the estimate for the fourth component is expanded to account for an averaging of method II and the vital rates method (which includes an average of net migration inferred from separate total population estimates from changes in birth and death rates).[1] To use the equation, thirty or more data values are needed (including some duplications) plus several constants (including those needed for indicating the averaging steps). Each of the data entries is supported (again with some duplications) by working papers which carry through from the original or reported data to the adjusted data used for the entries. The equation yields an estimate of the total current population for one state if, at each step, the aggregates for the forty-eight states are coordinated with corresponding, independently determined national totals.

Many numerical values are required to record the source data series, correct or adjust them, and for the final entries of the work paper estimate sheet. Each is a statistic in an area-oriented, time series. The implications are many. The statistical analyses must be limited to those applicable to nonparametric or distribution-free statistical methods. In the main, one must depend on data correction and adjustment and seek the most probable true values for the

[1] The Census Bureau's experimentation and incorporation of new data sources has now eliminated the use of the vital rates method and the need for this averaging step (see Benjamin Greenberg, *Illustrative Example of a Method of Estimating the Current Population of Subdivisions of the United States*, Bureau of the Census, Current Population Reports, Population Estimates, Series P-25, No. 133, March 16, 1956).

specific purpose for which they are being used in the estimating equation. The accuracy of the total estimates, apart from compensating errors, depends on how efficiently each adjustment problem is solved. However, there are many quite similar problems where the experience at the state level is applicable at the county level, and the experience with county data in one state is applicable to another state. Thus, as Shryock concludes, a coordinated program is desirable.

Coordinated work between the Census Bureau and the state agencies must be based on rather substantial estimating programs or projects in a number of states. What is feasible in each estimating situation? If the county current population estimates are for general purpose use, the program must be broad. But there are other uses for population estimates, as there are for county income payments, where Shryock correctly observes that pursuit of a high degree of accuracy may not be justified when their only use is for calculation of per capita income by county.

In county income work, the use of, or rather the need for using, population estimates extends to many steps in the detailed income estimating procedure. The necessity of relying on indirect allocators and of making adjustments from county-of-employment to county-of-residence sends us back the data from the 1950 census of population all too frequently. The lack of county population estimates (and more extensive employment estimates than at present) for local areas can be given as a basic reason for the questionable reliability of many of our total income estimates. The county income work was undertaken and has been pursued fruitfully by some of the university research bureau members of the Southeastern Economic Research Conference as an approach to or basis for more extensive studies of the state economies. Several of the member bureaus have been understaffed even for the income estimate work. The opportunity for research on local area data exists, however, and a record of shorter-term population changes than are available would provide basic information for many phases of such research.

Many state and local government agencies need current population estimates by local area. Even the municipal planning and economic development people sometimes admit that the current estimates might prove useful checks on their short-range population forecasts. But most of the state agencies have administrative and specialized problems of such scope that dependence on population estimates may be considered of minor importance. Thus, when estimates are required, their production is assigned as an incidental task to personnel who are specialists in other fields and without time,

interest, or background for making accurate population estimates.

To initiate some action, our university research bureau asked a number of the Tennessee state agencies about their needs for better current population estimates by county and city. We also asked whether or not the needs were such that the agency would consider joining in a cooperative research project for the production of generally acceptable estimates. Affirmative answers were received from all except one of the agencies. The project actually to develop, however, was one for making both current estimates and projections, by county and city, as a part of a long-range education study under direction of the Tennessee Legislative Council. The work underway has adequate financial support, cooperation from other agencies, and assurance of being continued beyond the time set for completion of the education study.

The proposal for a cooperative estimating project in Tennessee was based on a decision that the estimates could be undertaken on a budget of not over $15,000, provided some of the work on special tabulations and data improvement was done in the originating offices, especially the departments of education and public health. The principal reason for suggesting that the project be made a cooperative one, however, was to provide for more general acceptance, or what Shryock has referred to as an official status for the estimates. Representatives of several agencies can be brought into the planning and conduct of the project and into the appraisal of tentative estimates in a way to assure confidence in the final estimates.