DOCUMENT RESUME

05424 - [B0945852] RESTRICTED

Review of the Bureau of the Census Methodology To Develop Per Capita Income Estimates. GGD-78-55; B-146285. March 30, 1978. Released April 7, 1978. 6 pp. + 3 enclosures (9 pp.).

Report to Rep. Virginia Smith; Rep. Charles E. Grassley; by Victor L. Lowe, Director, General Government Div.

Issue Area: Data Collected from Non-Federal Sources: Reliability and Accuracy of Information (3101); Intergovernmental Relations and Revenue Sharing: Administration of the Assistance System (401).
Contact: General Government Div.
Budget Function: Revenue Sharing and General Purpose Fiscal Assistance: General Revenue Sharing (851).
Organization Concerned: Department of the Treasury; Department of Commerce; Bureau of the Census.
Congressional Relevance: Rep. Virginia Smith; Rep. Charles E. Grassley.
Authority: State and Local Fiscal Assistance Act of 1972 (P.L. 92-512). P.L. 94-488.

The methods used by the Bureau of the Census to calculate the per capita income estimates for small communities were examined. These estimates are a major factor in determining the amount of Federal general revenue sharing funds a community receives. The reliability of the per capita income estimates for small communities is questionable, and such communities may not be receiving their proper share of revenue sharing funds. The data used to develop the estimates for small communities are inadequate which could result in misallocations of revenue sharing funds. Misallocations are generally confined to small communities within a county because total funding for counties and States is controlled by other data which economists and statisticians consider more reliable. Although the Congress has provided alternatives for distributing the funds, the alternatives have not been used. In addition, there are no adequate checks or controls on the Bureau's methods for developing the estimates; some quality controls should be built into the process to provide greater assurance of the validity of the estimates. An explanation of the Bureau's calculations of estimates for Aredale, Iowa, and Minden, Nebraska, is included. (RRS)

5851

UNITED STATES GENERAL ACCOUNTING OFFICE
WASHINGTON, D.C. 20548

GENERAL GOVERNMENT DIVISION

MAR 30 1978

RESTRICTED — Not to be released outside the General Accounting Office except on the basis of specific approval by the Office of Congressional Relations.

B-146285

The Honorable Virginia Smith
The Honorable Charles E. Grassley
House of Representatives

Dear Ms. Smith and Mr. Grassley:

In response to your letter of July 26, 1977, we examined the methods used by the Bureau of the Census to calculate the per capita income estimates for small communities. The estimates are a major factor in determining the amount of Federal general revenue sharing funds a community receives. As you requested, we reviewed the estimates for Aredale, Iowa, and Minden, Nebraska, for entitlement period 9 (October 1977 to September 1978).

We found that the reliability of the per capita income estimates for small communities is questionable, and therefore such communities may not be receiving their proper share of revenue sharing funds. Although the estimates may be proper for some communities within a county, the data used to develop the estimates for small communities is inadequate and misallocations of revenue sharing funds to these communities could result.

The total revenue sharing funds available to a county would not change as a result of adjusting the income estimates for the communities within it. Misallocations are generally confined to small communities within a county because total funding for counties and States is controlled by other data, which economists and statisticians consider more reliable.

The Congress has recognized that the data on small communities may be inadequate for use in the general revenue sharing formula. Therefore it has provided alternatives for distributing the funds, but these alternatives have not been used.

GGD-78-55

In addition to the reliability problem, there are no adequate checks or controls on the Bureau's methods for developing the estimates. We believe that some quality controls should be built into the process to provide greater assurance of the validity of the estimates.

RELIABILITY OF ESTIMATES IS QUESTIONABLE

The reliability of the per capita income estimates for small communities is questionable because of possible errors in the base data and limitations on the information used to update the base data. Base data has a greater potential for error for small communities because the samples used are not large enough to be representative. For example, Bureau tests on data accuracy showed that the error rate for communities of less than 500 persons could be as high as 79 percent. A Stanford Research Institute study also concluded that data on communities with populations of less than 2,500 was poor. The Bureau's 1970 census, from which the per capita income base data is derived, includes samples for many communities with populations under 2,500.

In addition, some information used to update the 1970 base data was inaccurate and nonrepresentative. The base data was updated to 1974 to determine the per capita income estimates for entitlement period 9. To accomplish this update, the base data on taxable income was multiplied by the rate of change as measured by tax returns filed by community residents. In many cases, income was coded to the wrong community because taxpayers did not provide sufficient address information to identify the community where they resided. For updating the nontaxable portion of income, such as social security payments, Bureau of Economic Analysis (BEA) income data was used. However, the BEA data used was for the county rather than for the community. This presents a problem in developing community data because the characteristics of the county may not be the same as that of the community.

STATE AND COUNTY PER CAPITA INCOME ESTIMATES DERIVED FROM DIFFERENT DATA SOURCES

Sources for data on income estimates in State and county jurisdictions differ from those for the local governments. BEA data,the primary source for the State and county distributions, is derived from government administrative records and because of the greater size of the sample is considered more reliable than the sources for community per capita income estimates. Changes to estimates for small local governments will affect their revenue shares but it will not alter the overall amount of funds available to the county and State. A change in the amount allocated for the county or State can only be accomplished by changes in the BEA data.

BEA is currently revising its farm income data in accordance with the most recent agricultural census. It expects to have the final figures around June 1978. When this data is incorporated into the revenue sharing formulas, it should affect the revenue shares to all levels of governments.

CHECKS AND CONTROLS ON THE PER CAPITA INCOME ESTIMATES ARE NOT ADEQUATE

In addition to the reliability problems, there are no adequate checks or controls on the Census Bureau's methods for developing the estimates. The Bureau's methods are not formally reviewed, and communities cannot effectively challenge the estimates.

The methods used to develop the estimates are not subject to any formal reviews within or outside the Bureau of the Census. However, the methods are presented in papers distributed to statistical societies and are informally reviewed by State advisory groups.

Although regulations permit a community to challenge the Bureau's data, communities do not have access to the data necessary to support such challenges. Entitlement period 9 (fiscal year 1978), for example, is based on 1974 income estimates and there is no practical way for a community to reconstruct the income of its residents after the year has passed. Communities do not have access to Federal tax records which are an important data source in developing the estimates. Taking a current census to obtain 1974 income data would be impractical because some of the community's former residents could have moved away or died. Furthermore, if a community did challenge the estimates, the challenge would be adjudicated by the same group that developed the methods for and supervised the calculation of the estimates.

The Bureau usually responds to such challenges with form letters advising the communities that the Bureau's estimates appear reasonable and are the best available at the time.

There have been some successful challenges. In those cases, the Bureau reviewed the estimate and found that there was a significant error or that the estimate was obviously out of line. If the Bureau believes that the challenge is valid, it substitutes data, such as county data, in the methodology. This procedure is illustrated in the Bureau's adjustment of the Aredale per capita income estimate. (See enc. II) However, less than 10 percent of the challenges to the estimates have been successful. For entitlement period 9 there have been about 100 challenges from local governments.

The prospect of a community successfully appealing the Bureau's estimates in the courts is not bright. In a recent revenue sharing case, a district court ruled that the Secretary's discretion in using estimates was not reviewable. 1/

Estimates which determine the distribution of revenue sharing funds should be carefully reviewed. About $111 million of the $6.8 billion fiscal year 1978 revenue sharing funds are distributed to 13,764 communities with populations under 2,500. Considering the importance of these funds to small communities and the lack of reliability of the per capita income estimates, we believe some quality controls should be built into the process to provide greater assurance of the validity of the estimates. A group, such as the Office of Federal Statistical Policy and Standards, Department of Commerce, could independently evaluate the Bureau's methods and a party not responsible for the methods and calculations could review the challenges.

Bureau officials agreed with our position and told us of two recent actions taken to improve its controls over over the per capita income estimates and methodology. The Bureau is negotiating with the National Research Council, National Academy of Sciences, for a review of the Bureau's procedures for making intercensal estimates of population and per capita income for small areas, particularly the estimates developed for allocation of general

1/ City of Newark, New Jersey et al. v. W. Michael Blumenthal, Secretary of the Treasury, et al. no 74-548 (D.D.C. Jan. 17, 1978) revenue sharing funds. The initiative was taken by the Bureau in October 1977. In its draft proposal to the Bureau, the Academy stated that it "would examine the data sources now used, the availability and suitability of alternate data sources, the methods of estimation now used, and the possibility of using additional or alternate methods." The Bureau is also considering the establishment of a more formal system of appeals, using an independent group within the Bureau to review challenges to its data.

ALTERNATIVES TO USING PER CAPITA INCOME ESTIMATES ARE AVAILABLE

Alternatives to the present method of distributing funds to small communities are provided in the State and Local Fiscal Assistance Act of 1972. The Congress recognized that data on small communities might be inadequate for use in the general revenue sharing formula. The act therefore provides that the Secretary of the Treasury may choose to distribute the funds to communities with populations of 500 or less on the basis of population alone, if he believes that the data available is inadequate for application of the general formula.

The Congress also recognized that State governments may believe that the formula does not allocate funds among the counties and municipalities in accordance with the basic purposes of the act. It therefore provided that a State may enact legislation to employ alternative formulas within certain guidelines to distribute the State's share.

Although neither of these two options has been exercised, we believe that they should be carefully considered and evaluated in light of the difficulty of developing accurate per capita income estimates for small communities.

- - - -

During the latter part of our review the Bureau of the Census made changes in its methodology. These changes should affect per capita income estimates for many small communities. The Bureau is currently recomputing these estimates. In addition, the Bureau of Economic Analysis is revising some of its income data. When this data is incorporated into the revenue sharing formulas, it should affect the revenue shares at all levels.

Enclosed is more detailed information describing revenue sharing with particular emphasis on the methods used for developing per capita income estimates for small communities. We have also included a detailed explanation of the Bureau's calculations of the estimates for Aredale, Iowa, and Minden, Nebraska.

In this review we interviewed officials of the Bureau of the Census and Bureau of Economic Analysis, Department of Commerce, and the Office of Revenue Sharing, Department of the Treasury. At your request, we did not take the additional time to obtain formal written agency comments on matters discussed in this report. We did, however, obtain the views of agency representatives in informal conferences and have considered their views in preparing the report.

We plan no further distribution until 30 days from the date of this report unless you publicly announce its contents earlier. At that time we will send copies to interested parties and make copies available to others upon request.

Sincerely yours,

Victor L. Lowe

Victor L. Lowe
Director

Enclosures - 3

Enclosure I

BUREAU OF THE CENSUS METHODS FOR ESTIMATING PER CAPITA INCOME FOR SMALL LOCAL GOVERNMENTS AS USED IN GENERAL REVENUE SHARING

DISTRIBUTION OF REVENUE SHARING FUNDS

The program popularly called "general revenue sharing" was established under authority of the State and Local Fiscal Assistance Act of 1972 (Public Law 92-512), which was amended by Public Law 94-488 in 1976. Distribution of about $30 billion to State and local governments was authorized under the original legislation which covered the 5-year period ended December 31, 1976. The 1976 amendment covered the period January 1977 through September 1980 and authorized about $25.6 billion for distribution.

Revenue sharing funds are disbursed to over 39,000 State and local governmental units. Complex formulas govern how much the State and local governments receive during entitlement periods. The formulas take into consideration such factors as population, the tax effort of governments, and the relative poverty of a community as measured by per capita income (PCI).

One third of the funds are disbursed to the States and the remainder to the local governments. The funds distributed to the local governments are allocated to the various counties and then to local governments within the counties. Funds allocated to a county are divided between the county government and local governments within it on the basis of the relative adjusted taxes raised within the county. A local government's share depends on its population compared to the population of all other eligible local government units within the county, its relative tax raising effort, and its relative income. The final formula for distributing funds to the local governments simplifies to:

$$\frac{(\text{Adjusted taxes}) \times (\text{Per capita income for the county})}{(\text{Per capita income for the local government})^2}$$

Thus the local governments' PCI is an important factor in distributing funds. As shown by the following chart, about $111 million of the $6.8 billion in revenue sharing funds for fiscal year 1978 will be distributed to small communities.

Population size	Number of local governments	Funds received (millions)	Percentage of all funds disbursed
0 to 499	6,324	$ 13.7	0.2
500 to 999	3,336	24.2	.4
1,000 to 2,499	4,104	73.4	1.1
	13,764	$111.3	1.7

RELIABILITY OF PER CAPITA INCOME ESTIMATES FOR SMALL LOCAL GOVERNMENTS IS QUESTIONABLE

The reliability of the estimates for all local governments is questionable because of possible errors in the base data and limitations on the information used to update the base data.

Base data is not reliable

The base data for entitlement period 9 is the Census Bureau's 1970 decennial census. In that census 20 percent of the households were asked about their income for 1969. The 1970 census determined that Aredale, Iowa, had a population of 126 persons, residing in 47 households. The 20 percent sample consisted of 16 persons, 9 having incomes, residing in 8 households. From the data collected, the Bureau calculated a base PCI of $1,889. However, statistically, considering the size of the sample and the range of the incomes, the true PCI could be anywhere from $561 to $3,217.

Based on the sample size and the range of the incomes in the sample, the true PCI for Minden, Nebraska could be anywhere from $2,832 to about $3,744. The base PCI of $3,307 for this town of 2,669 persons was developed from information obtained from 531 persons with 340 having incomes. About 10 percent or 35 of the incomes were derived by statistical inference because the persons did not fully respond to the census questionnaires.

An evaluation of the general revenue sharing program by the Stanford Research Institute for the Office of Revenue Sharing concluded that data on communities with populations of less than 2,500 is poor.

The Bureau recognizes the lack of reliability in the data for small communities. Before entitlement period 9, the Bureau substituted county PCI data for places with populations of less than 500. However, the Bureau later made tests and found that the method of substituting county data was not as accurate as using adjusted data for places with populations of less than 1,000.

One of the Bureau's tests was a special census of 24 local governments, 17 with populations under 500 and 7 with populations between 500 and 1,000. The Bureau compared the PCI obtained from the special census to the PCI derived from using the Bureau's methodology. For the 17 smaller governments, the average PCI error using the substituted county data was 31.6 percent compared to 22 percent using the adjusted local government information. However, even with the use of the adjusted data, the error for the 17 ranged from 1.4 to 78.7 percent. For the seven larger governments, the average error using the local government data was also lower than that using the county data, 15.6 compared to 19.3 percent. But even for these larger governments, the error using the adjusted local government data ranged from 2.7 to 34.1 percent. As a result of these tests, the Bureau changed its methodology and now uses adjusted local government data for places with populations of less than 1,000.

For the 1980 decennial census the Bureau, recognizing its problems with using small samples from the 1970 census, is planning to increase to 50 percent the sample number of households to be asked income questions in communities with populations of less than 5,000. In fact the Bureau official responsible for the PCI methodology believes that a sample greater than 50 percent would be needed for local governments with populations of 100 to 200.

Factors used to update the base data include flaws

The 1970 base data was updated to 1974 for entitlement period 9 by multiplying it by the rate of income change calculated from Federal tax returns and Bureau of Economic Analysis (BEA) estimates. BEA uses information from such agencies as the Social Security Administration and the Department of Health, Education, and Welfare for making its estimates. The updating was done in two steps, 1969 to 1972 and 1972 to 1974. The taxable income portion is derived from the adjusted gross income (AGI) divided by the number of exemptions as shown by the tax returns filed by persons residing in the community.1/ The other types of income (transfer payments) were obtained from BEA.

The updating factors are flawed. The AGI, for tax purposes, is not the same as the definition used in the census. For example, unlike the census data, AGI includes capital gains and losses. Another example is farm income which on the tax returns is governed by complex tax provisions that do not apply to the simpler definition used in the census. Moreover, many taxpayers did not provide sufficient information on their returns to positively identify their place of residence. For example, about 30 percent of the 1972 tax returns did not provide complete geographic information. To code these returns the Bureau of the Census developed a special methodology. The Bureau recognizes that some income may be associated with the wrong place.

BEA does not maintain data at the local government level. Therefore county data is used to update the transfer payments portion of the base data. This presents a problem in developing local government data because the characteristics of the county may not be the same as that of the local government. However, transfer payments comprised only about 10 percent of total income.

Adjustments to reduce error

The updated local government estimates are adjusted to bring them in line with county and State income totals used in the county and State revenue sharing formulas. This adjustment, called a "rake," is usually small. However, in some farm counties it can be large.

1/At the time our report was being drafted, the Bureau of the Census was changing its methodology. One of these changes was to use substitute rate of income increases for local governments where there were less than 100 tax returns. The substitute increases are derived from tax returns from larger geographic areas, such as a county. This change was made because of serious limitations on Internal Revenue Service data for places with less than 100 tax returns.

To limit radical changes in the estimates from year to year for local governments, the Bureau employs several constraints.1/ One of the constraints to the small local government PCI is to limit the base figure to one standard error from the raw data. Another constraint limits the factor representing the increase in taxable income applied to the base data to 150 percent of the county rate of change. We were unable to obtain the rationale for the constraints other than that they are based on professional judgment.

Census uses a mathematical test to aid in detecting tax return geocoding errors. It compares a ratio of exemptions as shown on the tax returns to the Bureau's population estimates. If the ratio shows a change of more than 15 percent from one period to the next, the county rate of change is used instead of the local government's rate of change.

The effects of the adjusting factors and the constraints on the PCI calculations for Aredale, Iowa and Minden, Nebraska, are shown in enclosures II and III.

1/Subsequent to our fieldwork the Bureau of the Census changed several constraints.

Enclosure II

BUREAU OF THE CENSUS

COMPUTATION OF PER CAPITA INCOME

FOR AREDALE, IOWA, 1969-74

	Income subject to taxation	Transfer payments	Total
Base figure (1970 census-20 percent sample)	$ 1,662	$ 227	$1,889
Adjustment factor (note a)	1.2030	x1.0440	
Adjusted base figure	$ 1,999	$ 237	$2,236
Rate of increase (1969-72)	b/x1.4317	c/x1.3800	
Unadjusted PCI	$ 2,862	$ 327	
Rake factor (note d)	x .9710	x1.0060	
1972 raked PCI	$ 2,779	$ 329	$3,108
Rate of increase (1972-74)	e/x1.4037	x1.2310	
Unadjusted PCI	$ 3,901	$ 405	
Rake factor (note d)	x1.0300	x .9770	
1974 raked PCI	$ 4,018	$ 396	$4,414
Adjustment as a result of the challenge:			
Adjusted base figure	$ 1,999	$ 237	$2,236
Rate of increase (1969-72)	f/x1.2876	c/x1.3800	
Unadjusted PCI	$ 2,574	$ 327	
Rake factor (note d)	x .9584	x1.1280	
1972 raked PCI	$ 2,467	$ 369	$2,836
Rate of increase (1972-74)	x1.4460	x1.0920	
Unadjusted PCI	$ 3,567	$ 403	
Rake factor (note d)	x1.0000	x1.0000	
1974 raked PCI	$ 3,567	$ 403	$3,970

NOTES TO ENCLOSURE II

a/ The factor was developed by using (1) county PCI and the 1969 place PCI estimates in a regression procedure and (2) a multiple univariate rake factor which controls place totals to similar size places throughout the State, as well as to county income totals.

b/ The rate of income increase from Federal tax returns for Aredale was 59.7 percent. One hundred and fifty percent of the county rate of increase (28.8) was 43.2 percent. Since 59.7 exceeds 150 percent of the county rate, the rate of increase used for Aredale was constrained to 43.2 percent. (see enc. II, p. 5)

c/ County rate of increase. Data is not available for local communities.

d/ The factor is used to bring place income data into conformity with county income totals.

e/ The county rate of increase was substituted for the community rate of increase because of apparent mis-coding of tax returns. This miscoding change was disclosed by a mathematical test which compared the ratio of exemptions as shown on the tax returns to the Bureau's estimate of population. If the comparison shows a change of more than 15 percent, the county rate of change is used.

f/ The Census Bureau substituted the county rate of increase for the community rate of increase because of an apparent miscoding of tax returns. This was noted because the number of returns went up (45 to 53) and the number of exemptions decreased (128 to 125) and the population went down (126 to 119).

BUREAU OF THE CENSUS

COMPUTATION OF PER CAPITA INCOME

FOR MINDEN, NEBRASKA, 1969-74 (note a)

Hayes Township	Income subject to taxation	Transfer payments	Total
Base figure (1970 Census 20 percent sample)	$ 2,703	$ 143	$2,846
Adjustment factor (note b)	x1.0180	x .8530	
Adjusted base figure	$ 2,752	$ 122	$2,874
Rate of increase (1969-72)	x1.4240	d/x1.3200	
Unadjusted PCI	$ 3,919	$ 161	
Rake factor (note c)	x .9900	x .9310	
1972 raked PCI	$ 3,880	$ 150	$4,030
Rate of increase (1972-74)	x1.3800	x1.4770	
Unadjusted PCI	$ 5,354	$ 222	
Rake factor (note c)	x1.0075	x1.0450	
1974 raked PCI	$ 5,394	$ 232	$5,626

Lincoln Township	Income subject to taxation	Transfer payments	Total
Base figure (1970 Census 20 percent sample)	$ 3,312	$ 288	$3,600
Adjustment factor (note b)	x1.0070	x .8710	
Adjusted base figure	$ 3,335	$ 251	$3,586
Rate of increase (1969-72)	x1.2150	d/x1.3200	
Unadjusted PCI	$ 4,052	$ 331	
Rake factor (note c)	x .9755	x1.0090	
1972 raked PCI	$ 3,953	$ 334	$4,287
Rate of increase (1972-74)	x1.4570	x1.4770	
Unadjusted PCI	$ 5,760	$ 493	
Rake factor (note c)	x1.0094	x1.0380	
1974 raked PCI	$ 5,814	$ 512	$6,326

NOTES TO ENCLOSURE III

a/ Minden is located in two townships. One part of the city is in Hayes Township, and the other part is in Lincoln Township. Revenue sharing regulations require separate estimates of PCI for each part.

b/ This adjustment was made using the multiple univariate rake factor which controls places to similar size places throughout the State, as well as to county income totals.

c/ The factor is derived to bring place income data into conformity with county income totals.

d/ County rate of increase. Data is not available for local communities.