MEASUREMENT ERRORS IN THE 1960 CENSUS OF POPULATION

Leon Pritzker and Robert Hanson
Bureau of the Census

(This paper presents the personal views of the authors, and not necessarily the official position of the Bureau of the Census.)

1. Mathematical model.

The following mathematical model, described in detail by Hansen, Hurwitz, and Bershad (2), has aided us in the design of our program of evaluation:

a. The basic postulates of the model — not attainable in the real world — are that the process of recording a response for any individual is 1) repeatable and 2) gives rise to a random variable whose value at trial t is not correlated with its value on any other trial (t + k) and whose expected value is constant over repeated trials.

b. Thus, we shall be dealing with a random variable, x_{jtG}, whose value is:

x_{jtG} = 1, if the recorded response classifies individual j in class x of some characteristic on trial t of a census conducted under a set of general conditions G. (x may denote the age-class "0 - 4," the income class "$10,000 and over," etc.)

= 0, if otherwise.

c. Then we can define the proportion classified in x in trial t of a census of n_t persons:

$$p_{tG} = \frac{1}{n_t} \sum_{j}^{n_t} x_{jtG} \qquad (1)$$

d. The task of evaluation, when viewed from this standpoint, is to obtain estimates of the mean square error of p_{tG} for the classes of the population for which census data are tabulated, for example, 5-year age classes, the categories into which income reports are combined, and the number of years of regular schooling completed. Thus:

$$MSE_{p_{tG}} = E(p_{tG} - \bar{U})^2 \qquad (2)$$

where the expectation is taken over trials and where $\bar{U}$ is the true proportion.

e. The mean square error can be divided into its two main components by subtracting and adding $E(p_{tG})$ inside the parentheses of equation (2). We define $E(p_{tG}) = P_G$ and:

$$MSE_{p_{tG}} = E(p_{tG} - P_G)^2 + (P_G - \bar{U})^2 \qquad (3)$$

$$= \sigma^2_{p_{tG}} + B^2_{p_{tG}} \qquad (4)$$

where the first term is the total variance of p_{tG} and the second term is the square of the bias of p_{tG}.

f. To express the variance of p_{tG} in a particularly fruitful way, let:

$$E(x_{jtG}) = P_{jG}, \quad 0 \le P_{jG} \le 1,$$

where the expectation is over trials for a fixed person.

Then, as a device to be used for separating the response variance from the total variance, let:

$$\hat{P}_{tG} = \frac{1}{n_t} \sum_{j}^{n_t} P_{jG} \qquad (5)$$

This is the average of the expected values for the sample drawn on trial t.

Then:

$$\sigma^2_{p_{tG}} = E(p_{tG} - \hat{P}_{tG} + \hat{P}_{tG} - P_G)^2$$

$$= E(p_{tG} - \hat{P}_{tG})^2 + E(\hat{P}_{tG} - P_G)^2 + 2E(p_{tG} - \hat{P}_{tG})(\hat{P}_{tG} - P_G) \qquad (6)$$

In equation (6), the first term is defined as the response variance, the second term as the sampling variance, and the third as the interaction.

g. Our interest in this paper lies, in addition to the bias, in the response variance. Let:

$$\sigma^2_{\bar{d}} = E(p_{tG} - \hat{P}_{tG})^2 \qquad (7)$$

$$d_{jtG} = (x_{jtG} - P_{jG}),$$

the deviation of the response recorded for individual j on trial t from the expected value of the responses over all trials.

Then the evaluation of the expected values indicated in equation (7) for a fixed sample size $n_t = n$ gives:

$$\sigma^2_{\bar{d}} = \frac{1}{n} E(d^2_{jtG}) + \frac{(n-1)}{n} E(d_{jtG}\, d_{ktG}) \quad (8)$$

h. The partition of the response variance provided by equation (8) gives a useful tool for the analysis of sources of unreliability.

Since $E(d_{jtG}) = 0$, $E(d^2_{jtG})$ is a variance, σ^2_{dG}. It is the basic trial-to-trial variability, averaged over individuals. By analogy with population sampling, it is defined as the "simple response variance." It can be shown that:

$$\sigma^2_{dG} = \frac{1}{N} \sum_j^N P_{jG}(1 - P_{jG}) \quad (9)$$

where N = the total number of persons in the population.

It can further be shown that:

$$\sigma^2_{dG} \leq P_G (1 - P_G) \quad (10)$$

And thus, we can define the "index of inconsistency":

$$I_{dG} = \frac{\sigma^2_{dG}}{P_G (1 - P_G)} \leq 1 \quad (11)$$

$P_G(1 - P_G)$ will be recognized as the "sampling variance" of an estimated proportion for a sample of one element in simple random sampling. In this model, however, $P_G(1 - P_G)$ also includes the "simple response variance."

We have found the index, I_{dG}, useful in determining the inherent reliability or "measurability" of any 0, 1 variate included in a census. Estimates of this index enable us to compare the inherent reliability of measurement of, for example, one five-year age class with any other five-year age class, or a broader age grouping with a narrower age grouping, or an age class with an income class.

Under certain circumstances, the more important component of the response variance is the second term of equation (8), reflecting the covariance between response deviations. Results bearing on this component are not yet available. However, a major effort is now in progress to provide estimates.

2. Discussion of two components of the mathematical model.

The two components of interest are:

$$B_{p_{tG}} = P_G - \bar{U} \quad (3a)$$

$$I_{dG} = \frac{\sigma^2_{dG}}{P_G (1 - P_G)} \quad (11)$$

In evaluating the national summary statistics of a census of population, of these two components the one of significance — except for very rare characteristics — is $B_{p_{tG}}$. The term, I_{dG}, is of interest in evaluating the precision of the census-taking process but not the accuracy of the statistics.

3. Estimators.

a. Bias.

Recall that

$$B_{p_{tG}} = P_G - \bar{U} \quad (3a)$$

For P_G, we write the census statistic itself, p_{tG}.

The problem, then, is to find a valid estimator for $\bar{U}$. This comes down to devising a method of measurement that can be agreed upon as "standard," or failing that, devising a "preferred" method that, by judgment or by test, is found to produce the required statistics with significantly smaller mean square errors than the statistics of a census. Then the estimator of $B_{p_{tG}}$ is:

$$b_{p_{tG}} = p_{tG} - p_{tG'} \quad (12)$$

where $p_{tG'}$ is an estimate produced by a standard or a preferred method.

The tables we have provided below give estimates of the bias, $b_{p_{tG}}$. These estimates come from the Evaluation and Research Program of the 1960 Census in which the following studies have been conducted to obtain estimates of the bias in national summary statistics:

(1) CPS - Census Match. (Study EP-23)

The Current Population Survey (CPS), conducted monthly by the Bureau of the Census, is regarded as a "preferred method," in relation to a census of population, for collecting data on the size and composition of the labor force. The data for a sample of about 8,000 house-

holds enumerated in both the April 1960 CPS and in the 1960 Census (conducted primarily in April 1960) have been brought together and analyzed. Some results of this match, as well as of a match conducted between the April 1950 CPS and the 1950 Census, are presented below.

(2) Population Content Evaluation Study. (Study EP-10)

This study was based on what we term an "intensive-interview" approach to obtaining census data. The "intensity" of an interview is a relative matter. We provide some notion of the "intensity" in Attachment A, in which are reproduced the questions employed in the inquiry to determine the age of each person in the sample for Study EP-10. We would assert that the interview designed for Study EP-10 was considerably more "intensive" than the interview designed for obtaining data for characteristics of the population in the Post-Enumeration Survey of the 1950 Census.

Data were obtained on the following characteristics by intensive interview: usual place of residence, age, place of residence on April 1, 1955 (providing a measure of mobility of the population), school enrollment and attainment, number of children ever borne by women who have ever been married, income from self-employment, and income from sources other than earnings. At the outset of each intensive interview, the interviewer also recorded the sex and color of the person for whom data was being obtained. In most cases the interviewer talked with the person himself when the person was an adult, and to the parent or guardian when the person was a child.

The results presented below are based on a sample of about 9,500 persons and reflect differences between the "best" answer obtained by the interviewers in Study EP-10, after reconciliation of differences between Census and EP-10 responses, and the Census responses recorded on the FOSDIC schedules of the 25 percent Census sample.

(3) Record checks and related studies.

The results are not yet available, but the Bureau of the Census expects to have, as part of the Evaluation and Research Program, estimates of bias in summary statistics from the following sources:

(a) Age and sex for children under 10 years of age from birth certificates.

(b) Occupation and industry of employed persons from employer records.

(c) Occupation, industry, and class of worker (the component of bias due to errors in coding) from a study being conducted by Fasteau, Ingram, and Mills (1).

b. Index of inconsistency.

Recall that:

$$I_{dG} = \frac{\sigma^2_{dG}}{P_G(1 - P_G)} \qquad (11)$$

We can write p_{tG} for P_G. Thus, the problem is to find an estimator for σ^2_{dG}.

Recall that:

$$\sigma^2_{dG} = E(d^2_{jtG}) \qquad (9a)$$

$$= E(x_{jtG} - P_{jG})^2$$

Now consider a theoretical repetition (t') of the Census (t) under the identical conditions, G. The repetition is theoretically independent. For any 0, 1 variate, for example age 15 - 19, the Census and its repetition would generate the following type of table:

Repetition \ Census	$x_{jtG} = 1$	$x_{jtG} = 0$	
$x_{jt'G} = 1$	a	b	a + b
$x_{jt'G} = 0$	c	d	c + d
	a + c	b + d	n_t

$$n_t = a + b + c + d = n_{t'} = n$$

Define the gross difference rate, g:

$$g = \frac{b + c}{n} \qquad (13)$$

It can be shown that under the specified conditions:

$$\sigma^2_{dG} = \tfrac{1}{2} E(g) \qquad (14)$$

Then, writing $\hat{I}_{dG}$ as an estimator of I_{dG},

$$\hat{I}_{dG} = \frac{g}{2p_{tG}(1 - p_{tG})} \qquad (15)$$

As pointed out by our colleague, Max Bershad, this term is equivalent to χ^2 for a two-by-two contingency table, as defined by Karl Pearson (3).

In practice the conditions do not hold and thus $\hat{I}_{dG}$ is a biased estimator of I_{dG}. In particular, there is a growing body of evidence that the division by 2 cannot be justified where the "repetition" is followed by reconciliation with the Census. In our judgment, the division by 2 has provided estimates that understate the response variability of the 1960 Census statistics. We believe, however, that $\hat{I}_{dG}$ is a useful estimator for helping us assess the relative consistency of recorded responses, as between characteristics and as between censuses. Biased estimates of I_{dG} will be available for selected characteristics, for the United States as a whole, from the following studies conducted as part of the Evaluation Program of the 1960 Census:

(1) CPS - Census Match. Study EP-23 — described above.

(2) Population Content Evaluation Study, Study EP-10 — described above.

(3) Replication Study. Study EP-18 This is a study based on the re-enumeration of 6,000 households included in the 25-percent sample of the 1960 Census. Results are not yet available.

4. Results and interpretations.

The results presented below are preliminary in three respects. First, there will probably be some corrections in the estimates themselves. Second, the estimates do not take account of the effect of errors in the coverage of the population and in the coverage of housing units on accuracy and reliability. Third, the estimates do not take account of the effect of nonresponse on accuracy and reliability.

The results are based on estimated "identical populations." For each characteristic studied, the "identical population" is projected from those persons (or housing units) for whom responses were recorded both in the Census and in the survey used as the standard to evaluate the Census.

a. Labor-force status. As indicated above, the standard for evaluating the Census is the CPS. Tables 1 and 2 present the results of two CPS-Census Matches. Table 1 permits us to compare the biases in the statistics for labor-force status of the 1960 Census with those of the 1950 Census. The statistics for females appear to have been uniformly improved in 1960 over 1950. The statistics for males also appear to have improved in 1960 over 1950 for unemployment. For employment in agriculture, the 1960 figures appear to be worse, and, for the other components, about the same. However, there are still estimates of bias in the 1960 Census statistics that some analysts would probably regard as important. For example, the male, civilian labor-force total in the 1960 Census is understated by more than 2 percentage points and the female, civilian labor force is understated by about 1 percentage point.

Table 2 provides a basis for comparing the data-collection procedures in the 1950 and 1960 Censuses by showing estimates of the indexes of inconsistency, $\hat{I}_{dG}$, for the labor-force status categories. The estimates are biased in that the CPS, at the time of each Census, has been taken to be the "repeated trial." We recognize that the CPS is not taken under the same conditions as the Census. In fact, we regard the CPS to be better for determining labor-force status. However, if the following two reasonable assumptions hold, we can conclude that the estimates of the index are underestimates:

(1) The covariances between the CPS and the Census response deviations are zero or positive.

(2) The simple response variances, σ^2_{dG}, of the CPS are less than the simple response variances of the Census for labor-force classifications.

There is some reason to believe that the 1960 CPS was of higher quality than the 1950 CPS. However, if we further assume that simple response variances of the CPS in 1950 are about equal to those of the CPS in 1960, we can make some rough inferences about the quality of the 1960 Census procedures as compared to the 1950 procedures. With possibly one or two relatively minor exceptions, the 1960 Census procedures appear to be at least the equal in reliability of the 1950 Census procedures. However, we can see by comparing Table 2 and Table 3 that the estimated indexes for labor-force characteristics are several times greater than for sex or color. Both the 1950 and the 1960 Census procedures for determining labor-force status generated a considerable amount of noise, i.e., response variance, particularly in the unemployment classification.

b. Other characteristics of the population. Table 3 presents estimates of the bias, $b_{P_{tG}}$, and of the index of inconsistency, $\hat{I}_{dG}$, for sex and for color. Table 4

presents the same types of estimates for age; Table 5, for mobility status; and Table 6, for educational attainment of the population 25 years old and over. Study EP-10 described above has been used as the source of both types of estimates for the 1960 Census. No corresponding estimates are available for sex and race in the 1950 Census. For age, mobility, and educational attainment, the Post-Enumeration Survey, or "PES" (4), has been used as the source of both types of estimates for the 1950 Census. All of the estimates — for 1960 and for 1950 — are undoubtedly biased. There are some a priori reasons for believing that Study EP-10 in 1960 was more "intensive" and had smaller mean square errors than the PES in 1950.

We offer the following comments about Tables 3 - 6:

(1) The reliability of the classification of the population by sex and by color was high in comparison to the reliability of classification of most other characteristics studied in censuses and surveys. However, there is some indication — to be investigated further — that there was a bias in the direction of understating the number of females and of nonwhites. (Recall that these estimates do not reflect coverage error or nonresponse.) The females are understated by 0.4 percent and the nonwhites by 1.7 percent, relative to the estimates of Study EP-10.

(2) The estimates of the index of inconsistency, $\hat{I}_{dG}$, indicate almost uniformly more reliable classifications in the 1960 Census than in the 1950 Census. As we have indicated earlier, the indexes are biased. It is within the realm of possibility that these results can be accounted for entirely by more reliable processes of classification in Study EP-10 than in the PES of the 1950 Census. It would be difficult for us, however, to believe that the improvement in our method of evaluation was so great as to mask a decline in the reliability of the classification processes in the 1960 Census itself. We would conclude that the quality of classification in 1960 was at least equal to or perhaps better than the quality in 1950 for age, mobility status, and educational attainment.

(3) Pending tabulation and analysis of the final results, we would like to withhold judgment concerning the estimates of bias, except for the following observations: First, the bias in the age class, 65 - 69, is consistent with the direction indicated by demographic analysis. In 1950, the PES result was not consistent. Second, the fact that the estimates of bias for mobility status in 1960 are generally higher than those in 1950 has to be viewed against the fact that the 1960 classifications cover a five-year period and the 1950, a one-year period. Third, the estimates of bias in educational attainment of the population 25 years old and over both for 1950 and for 1960 are generally consistent with the hypothesis of a net tendency to overstate educational attainment.

5. Concluding remarks.

The general method of measurement that we have illustrated has provided useful information to the Bureau of the Census. However, it has also been a source of great frustration. The index, I_{dG}, is a measure of the noise in the census-taking process. The bias, $B_{P_{tG}}$, is a measure of the misinformation in the process. The specific techniques developed thus far to estimate these quantities also have high noise levels and also give erroneous signals. The search for standards of measurement has a long way to go.

6. Acknowledgments.

The Evaluation and Research Program of the 1960 Census required substantial contributions by a large number of people. The program is under the general direction of Conrad Taeuber, Morris H. Hansen, and William N. Hurwitz. Technical planning was the primary responsibility of Leon Pritzker and Joseph Steinberg, supported by other staff members. The operation of the program was the primary responsibility of Joseph Steinberg.

The authors wish to acknowledge especially the contributions to the studies described in this paper by Donald E. Ball, Stanley Greene, Eli S. Marks, Ruth H. Mills, and Martha Thomas.

REFERENCES

1. Fasteau, H. H., Ingram, J. J., and Mills, R. H. Study of the reliability of coding of census returns. Paper prepared for September 1962 meetings of the American Statistical Association.

2. Hansen, M. H., Hurwitz, W. N., and Bershad, M. A. Measurement errors in censuses and surveys. Bulletin of the International Statistical Institute, v. XXXVIII, pt. 2, pp. 359-74.

3. Pearson, Karl. On the probability that two independent distributions of frequency are really samples from the same population. Biometrika, v. 8, 1911-12, pp. 250-4.

4. U. S. Bureau of the Census. The Post-Enumeration Survey: 1950. Bureau of the Census Technical Paper No. 4, Washington, D. C., 1960.

ATTACHMENT A

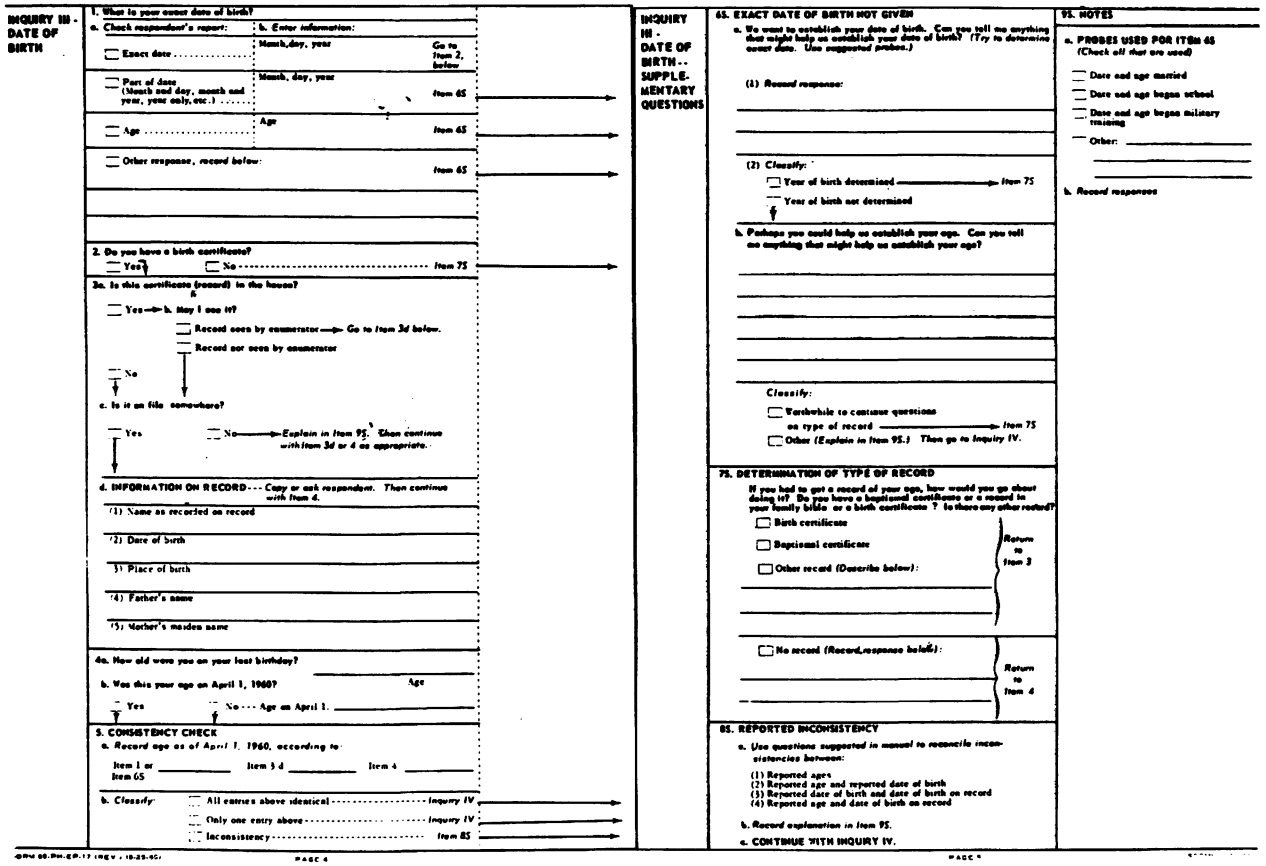

Table 1.--PRELIMINARY ESTIMATES OF THE BIAS IN THE STATISTICS ON LABOR-FORCE STATUS IN THE 1960 AND 1950 CENSUSES OF POPULATION, FOR THE "IDENTICAL POPULATION" FOURTEEN YEARS OLD AND OVER, BY SEX

Sex and labor-force status	1960 Census				1950 Census				Difference between relative biases 2/
	Estimated percentage in class according to		Estimated bias 1/		Estimated percentage in class according to		Estimated bias 1/		
	1960 Census (P_{tG})x100	CPS ($P_{tG'}$)x100 April 1960	Absolute 3/ ($b_{P_{tG}}$)x100	Relative 3/ ($b_{P_{tG}}$) as percent of $P_{tG'}$	1950 Census (P_{tG})x100	CPS ($P_{tG'}$)x100 April 1950	Absolute 3/ ($b_{P_{tG}}$)x100	Relative 3/ ($b_{P_{tG}}$) as percent of $P_{tG'}$	|(8)| - |(4)|
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Males	100.0	100.0	--	--	100.0	100.0	--	--	--
1. In the civilian labor force....	79.1	81.3	-2.2	-2.7	82.1	84.0	-1.8	-2.2	-0.5
a. Employed.....	75.5	77.5	-2.0	-2.6	78.2	79.2	-1.0	-1.2	-1.4
(1) In agriculture.....	7.2	8.3	-1.1	-13.6	11.9	12.5	-0.7	-5.5	-8.1
(2) In nonagricultural industries.....	68.3	69.3	-0.9	-1.3	66.4	66.7	-0.3	-0.4	-0.9
b. Unemployed.....	3.6	3.8	-0.2	-4.4	3.9	4.8	-0.9	-18.1	+13.7
2. Not in the civilian labor force	20.9	18.7	+2.2	+11.8	17.9	16.0	+1.8	+11.5	-0.3
Females	100.0	100.0	--	--	100.0	100.0	--	--	--
1. In the civilian labor force....	35.1	36.1	-1.0	-2.8	29.6	31.8	-2.2	-6.8	+4.0
a. Employed.....	33.3	34.2	-0.8	-2.4	28.3	30.2	-1.9	-6.2	+3.8
(1) In agriculture.....	1.1	1.2	-0.1	-9.1	0.9	1.7	-0.8	-46.4	+37.3
(2) In nonagricultural industries.....	32.2	32.9	-0.7	-2.2	27.4	28.5	-1.1	-3.8	+1.6
b. Unemployed.....	1.8	2.0	-0.2	-10.3	1.3	1.6	-0.3	-18.9	+8.6
2. Not in the civilian labor force	64.9	63.9	+1.0	+1.6	70.4	68.2	+2.2	+3.2	+1.6

1/ Minus sign indicates understatement in Census; plus sign indicates overstatement.
2/ Minus sign indicates larger bias in 1960 Census than in 1950 Census; plus sign indicates larger bias in 1950 Census.
3/ Computed from unrounded figures.

Table 2.—PRELIMINARY ESTIMATES OF THE "INDEX OF INCONSISTENCY" FOR LABOR-FORCE CLASSIFICATIONS IN THE 1960 AND 1950 CENSUSES OF POPULATION, FOR THE "IDENTICAL POPULATION" FOURTEEN YEARS OLD AND OVER, BY SEX

Sex and labor-force status	Index of inconsistency, $\hat{I}_{dG}$		Difference [1]
	1960 Census	1950 Census	(2) - (1)
	(1)	(2)	(3)
Males			
1. In the civilian labor force	.177	.205	+.028
a. Employed	.170	.196	+.026
(1) In agriculture	.224	.144	-.080
(2) In nonagricultural industries	.132	.140	+.008
b. Unemployed	.500	.513	+.013
2. Not in the civilian labor force	.177	.205	+.028
Females			
1. In the civilian labor force	.192	.195	+.003
a. Employed	.175	.180	+.005
(1) In agriculture	.593	.957	+.364
(2) In nonagricultural industries	.156	.145	-.011
b. Unemployed	.720	.751	+.031
2. Not in the civilian labor force	.192	.195	+.003

1/ Minus sign indicates greater unreliability in 1960 Census than in 1950 Census; plus sign indicates greater unreliability in 1950 Census.

Table 3.—PRELIMINARY ESTIMATES OF THE BIAS IN THE STATISTICS AND OF THE "INDEX OF INCONSISTENCY" FOR SEX AND COLOR IN THE 1960 CENSUS OF POPULATION, FOR THE "IDENTICAL POPULATION"

Characteristic and category	Bias, $b_{p_{tG}} \times 100$ [1]	Relative bias [2]	$\hat{I}_{dG}$
	(1)	(2)	(3)
Sex			
Male	+0.2	+0.4	.018
Female	-0.2	-0.4	.018
Color			
White	+0.2	+0.2	.045
Nonwhite	-0.2	-1.7	.045

1/ Minus sign indicates understatement in Census; plus sign indicates overstatement.

2/ Computed from: $[\frac{b_{p_{tG}}}{p_{tG'}}] \times 100$, where $p_{tG'}$ is the estimate from Study EP-10.

Table 4.--PRELIMINARY ESTIMATES OF THE BIAS IN THE STATISTICS AND OF THE "INDEX OF INCONSISTENCY" FOR FIVE-YEAR AGE CLASSES IN THE 1960 AND 1950 CENSUSES OF POPULATION, FOR THE "IDENTICAL POPULATION"

Age class	Bias, $b_{p_{tG}}$ x 100 [1]		Relative bias [2]			Index of inconsistency, $\hat{I}_{dG}$		
	1960 Census	1950 Census	1960 Census	1950 Census	Difference [3] |(4)| - |(3)|	1960 Census	1950 Census	Difference [3] (7) - (6)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
0-4	+.01	-.18	+.06	-1.63	+1.57	.020	.025	+.005
5-9	+.02	+.08	+.16	+ .92	+ .76	.029	.028	-.001
10-14	+.05	+.01	+.47	+ .11	- .36	.024	.034	+.010
15-19	-.07	+.11	-1.00	+1.64	+ .64	.029	.040	+.011
20-24	-.04	+.02	-.79	+ .26	- .53	.037	.051	+.014
25-29	+.08	--	+1.53	- .03	-1.50	.036	.062	+.026
30-34	-.03	+.04	-.49	+ .48	- .01	.043	.076	+.033
35-39	+.12	+.06	+1.85	+ .78	-1.07	.058	.075	+.017
40-44	+.03	+.09	-.44	+1.38	+ .94	.078	.088	+.010
45-49	-.12	--	-1.85	- .07	-1.78	.071	.101	+.030
50-54	+.03	+.02	+.59	+ .30	- .29	.078	.112	+.034
55-59	+.10	-.16	+2.11	-3.11	+1.00	.063	.103	+.040
60-64	-.10	-.04	-2.77	-1.04	-1.73	.098	.084	-.014
65-69	+.09	-.02	+2.63	- .52	-2.11	.078	.090	+.012
70-74	-.11	--	-.40	+ .12	- .28	.095	.095	--
75 and over	-.05	-.03	-1.80	-1.07	- .73	.032	.051	+.019

1/ Minus sign indicates understatement in Census; plus sign indicates overstatement.

2/ Computed from: $[\frac{b_{p_{tG}}}{p_{tG'}}]$ x 100, where $p_{tG'}$ is the estimate from Study EP-10.

3/ Minus sign indicates higher level of error in 1960 Census than in 1950 Census; plus sign indicates higher level of error in 1950 Census.

Table 5.--PRELIMINARY ESTIMATES OF THE BIAS IN THE STATISTICS AND OF THE "INDEX OF INCONSISTENCY" FOR MOBILITY-STATUS CLASSES IN THE 1960 AND 1950 CENSUSES OF POPULATION, FOR THE "IDENTICAL POPULATION" FIVE YEARS OLD AND OVER

Mobility-status classes [1]	Bias, $b_{p_{tG}}$ x 100 [2]		Relative bias [3]			Index of inconsistency, $\hat{I}_{dG}$		
	1960 Census	1950 Census	1960 Census	1950 Census	Difference [4] |(4)| - |(3)|	1960 Census	1950 Census	Difference [4] (7) - (6)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Same house	+1.4	+0.3	+2.6	+0.3	-2.3	.072	.223	+.151
Different house, same county	+0.3	-0.7	+1.2	-5.8	+4.6	.125	.260	+.135
Different county, same State	-0.7	+0.1	-7.6	+2.3	-5.3	.108	.274	+.166
Different State	-0.9	+0.2	-11.2	+8.3	-2.9	.107	.336	+.229
Abroad	-0.2	+0.2	-13.3	+278.3	+265.0	.187	.584	+.397

1/ Residence five years prior to the Census date for the 1960 Census; residence one year prior to the Census date for the 1950 Census.
2/ Minus sign indicates understatement in Census; plus sign indicates overstatement.

3/ Computed from: $\left[\frac{b_{p_{tG}}}{p_{tG'}}\right] \times 100$, where $p_{tG'}$ is the estimate from Study EP-10.

4/ Minus sign indicates higher level of error in 1960 Census than in 1950 Census; plus sign indicates higher level of error in 1950 Census.

Table 6. - PRELIMINARY ESTIMATES OF THE BIAS IN THE STATISTICS AND OF THE "INDEX OF INCONSISTENCY" FOR EDUCATIONAL ATTAINMENT CLASSES IN THE 1960 AND 1950 CENSUSES OF POPULATION, FOR THE "IDENTICAL POPULATION" TWENTY-FIVE YEARS OLD AND OVER

Educational attainment class	Bias, $b_{p_{tG}}$ x 100 [1]		Relative bias [2]			Index of inconsistency, $\hat{I}_{dG}$		
	1960 Census	1950 Census	1960 Census	1950 Census	Difference [3] |(4)| - |(3)|	1960 Census	1950 Census	Difference [3] (7) - (6)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
None.....	-0.1	-1.0	- 0.7	-29.0	+28.3	.238	.554	+.316
Elementary, 1-4 years.....	-0.5	+0.3	+ 8.5	+ 3.6	- 4.9	.309	.360	+.051
Elementary, 5-6 years.....	-0.8	-0.6	-11.0	- 5.7	- 5.3	.333	.479	+.146
Elementary, 7 years.....	-0.8	-1.2	-11.1	-14.3	+ 3.2	.399	.604	+.205
Elementary, 8 years.....	+0.7	+1.4	+ 4.6	+ 7.3	+ 2.7	.300	.400	+.100
High school, 1-3 years.....	+0.7	-0.7	+ 3.6	- 4.0	+ 0.4	.240	.375	+.135
High school, 4 years.....	-0.5	+0.3	- 2.0	+ 1.3	- 0.7	.186	.263	+.077
College, 1-3 years.....	+1.0	+1.0	+11.4	+15.0	+ 3.6	.224	.339	+.115
College, 4 or more years.....	+0.2	+0.5	+ 3.1	+ 8.7	+ 5.6	.074	.170	+.096

1/ Minus sign indicates understatement in Census; plus sign indicates overstatement.

2/ Computed from: $[\frac{b_{p_{tG}}}{p_{tG'}}]$ x 100, where $p_{tG'}$ is the estimate from Study EP-10.

3/ Minus sign indicates higher level of error in 1960 Census than in 1950 Census; plus sign indicates higher level of error in 1950 Census.