A TRANSPARENT FILE FOR A ONE NUMBER CENSUS

C.T. Isaki, M.M. Ikeda, J.H. Tsay , Bureau of the Census; W.A. Fuller, Iowa State University
Cary T. Isaki, Bureau of the Census, Washington, D.C. 20233

Key Words: Post enumeration survey; regression; raking; controlled rounding

I. INTRODUCTION

Current plans for the 2000 U.S. Census of Population and Housing include the use of sampling and estimation methods. In the past, two forms, a long form and a short form have been used. The long form has always been sent to a sample of housing units. In 2000, it is planned to use sampling and estimation with regard to the short form as well. The short form data consist of the tenure class of the housing unit; the age, race and sex of persons in the housing unit; and the relationships among members of the household.

In the 1990 Census, short form data were collected from all persons that could be contacted. Census questionnaires were first mailed to all units. Then enumerators were sent to all nonresponding units to obtain short form information in an operation called nonresponse follow-up (NRFU). In addition, enumerators were sent to all units identified as vacant by the post office. This was the final operation of the enumeration phase of the Census. In 1990, a post enumeration survey or PES was conducted and the enumeration phase person data were adjusted using data from the PES. Upon completion of the adjustment, there were two sets of population counts - one from the enumeration phase and one produced using the PES. The Secretary of Commerce (the Bureau's parent agency) chose the enumeration data as the official Census results.

The 1990 PES adjusted Census counts suffered from several deficiencies. One of the deficiencies was the way in which the PES estimated total person count was assigned to blocks. The PES was used to estimate persons at the block level, but no estimates of housing units at the block level were attempted. The difference between the PES estimate of persons and the enumeration count at the block level was assigned to a tabulation category called "group quarters of unrelated persons." Data users were uncomfortable with this category because the person estimates were really for persons in housing units.

The 2000 Census is being planned as a 'one number census'. That is, the official counts are to be based on a PES procedure called integrated coverage measurement (ICM). Improvements in Census processing and operations are planned to enable the Census Bureau to provide total population counts for each state and the District of Columbia by December 31, 2000. Improvements in timing are necessary because in 1990, the enumeration phase of the Census provided the December 31, 1990 population figures while the PES results became available in late spring of 1991.

The 2000 U.S. Census plans include the use of sampling and estimation of NRFU cases as well as the 'post office vacants'. We term the combined operation as NRFU. In the following, we describe a methodology that could be used to produce a short form data file of households for a census such as the 2000 U.S. Census. We first present a scenario of sampling and estimation and then introduce the proposed methodology for transparent file construction. We also present the results of the construction of a transparent file for two of the 1995 Test Census sites, Paterson, N.J. and Oakland, CA. Finally, we discuss future work.

II. TRANSPARENT FILE

Given a census operation that includes sampling and estimation, we define a *transparent* file as a census data file that is devoid of any evidence of sampling and estimation.

The 2000 U.S. Census plans include sampling and estimation procedures that can easily produce non-integer estimates. A transparent Decennial Census data file would -

i) have the appearance of an enumeration with unit weights to avoid non-integer estimates
ii) be constructed by duplicating or eliminating housing units on the enumeration phase data file at the block level
iii) contain a listing of housing units and persons with their short form data and block identification and
iv) not assign housing units a street address in the block

In the context of a 'one number Census', a transparent file would provide person and housing unit counts that are both arithmetically and definitionally consistent. Tabulations from such a file would be simple and there would be no need to qualify person versus housing unit counts.

III. METHODOLOGY

Transparent file construction assumes the existence of a data file at the completion of NRFU, called the Census file, Census Enumeration file or initial phase file. It is assumed that estimates of total persons by age-race-sex-tenure categories are available from the ICM by geographic area, e.g. a state. The Census file and the ICM estimates form the basis for the construction of the transparent file.

Given the estimates of number of persons by categories, the task is to construct housing unit estimates for all Census blocks. This is done by first estimating factors to be applied to the Census file housing unit counts to create estimates of total housing units and then rounding these estimates to integers. Vacant and occupied housing units are treated differently. Finally, households on the Census file are duplicated or eliminated at the block level to obtain the desired transparent file.

To estimate the number of vacant housing units, dual system estimators of total vacant housing units for a state are divided by the total vacant housing units on the Census file for the state. See Isaki and Ikeda (1996). The ratio is applied to the Census file of vacant counts by blocks. The results are control rounded so that each number is rounded to the integer directly below or above in such a way that the sum of the

rounded numbers equals the rounded dual system estimated total. See Cox and Ernst (1982).

Estimators for occupied housing units are defined by creating weights for a set of housing unit categories, rather than for individual households, to reduce the computational effort required in weight construction. The categories of occupied housing units were defined using factors associated with coverage. Such factors include race, age, and sex of the householder, tenure, presence of spouse, and number of persons in the household. In addition, distinctions between certain minority households according to the presence or absence of certain types of persons were also made. For example, a three-person 30 to 49 aged Black female headed household with no spouse present but with young children less than ten years of age is in a different category from households with no children less than ten years old. Households that in the past were not subject to large coverage error, such as nonminority households, were placed in less detailed categories.

To describe the estimation procedure, let x_{ij}, i = 1, 2, ..., n, j = 1, 2, ..., m, denote the number of persons on the Census file in housing unit category i and ICM person category j. For example, x_{ij} may denote the number of Hispanic male or female renters aged 0 to 17 in Paterson, New Jersey, that reside in households containing five persons headed by a Hispanic female aged 30 to 49 and containing two or more young adult Hispanic males. Let $\hat{X}_j$ denote the dual system estimator of the total number of persons in ICM person category j and let h_i, i = 1, 2 ..., n, denote the number of occupied units in household category i in the Census file for the tabulation area. The tabulation area might be a state or a portion of a state. In addition to the estimates for the tabulation area for ICM person category j, an estimate of total persons is constructed for every block, where the estimate for block k is the number of persons in each category in the block multiplied by the estimated undercount rate for that category and summed. Thus, the estimated total number of persons for block k is

$$\hat{X}_{Tk} = \sum_{i=1}^{n} \sum_{j=1}^{m} x_{ijk} R_{uj} , \qquad (1)$$

where x_{ijk} is the number of individuals in housing unit category i, in ICM person category j in block k on the Census file, R_{uj} is the ICM estimator of the ratio of total persons to census counted persons in person category j for the tabulation area. The estimator in (1) is called the synthetic estimator of the total number of persons in the block.

Weights for housing unit categories within blocks were constructed with an iterative procedure. Let

$$a_{rk} = \hat{X}_{Tk} \left(\sum_{i=1}^{n} \sum_{j=1}^{m} c_{r-1,i} x_{ijk} \right)^{-1} , \qquad (2)$$

where $\mathbf{c_o} = \mathbf{1}$ and r is the iteration index. Given a $\mathbf{c_{r-1}}$, a new vector of weights
$\mathbf{c_r} = \{c_{r,1}, c_{r,2}, ..., c_{r,n}\}$ is chosen to minimize

$$f(c) = \sum_{i=1}^{n} (c_{r,i} - 1)^2 h_i , \qquad (3)$$

with respect to $\mathbf{c_r}$, subject to

$$\sum_{i=1}^{n} c_{r,i} \left(\sum_{k=1}^{B} a_{rk} x_{ijk} \right) = \hat{X}_j , \quad j = 1,2,...,m , \quad (4)$$

$$c_{r,i} > K > 0. \qquad i = 1, 2,...,n ,$$

where B is the number of blocks and K is a chosen lower bound for the weights. A total of four iterations were conducted. The weight for housing unit type i in block k is

$$w_{ik} = c_{4,i} a_{4k} .$$

The weight construction combines elements of raking and of least squares regression estimation. Related regression estimators have been considered by Huang and Fuller (1978), Bethlehem and Keller (1987), and Särndal and Deville (1992). Also, see Zaslavsky (1988).

The w_{ik} are adjustment factors for occupied units on the Census file in housing unit category i and block k. If h_{ik} denotes the Census file count of occupied housing units in category i and block k then $w_{ik} h_{ik}$ would be the estimator of the number of category i housing units in block k. Using the estimators, one obtains a two-dimensional set of housing unit counts, housing unit category by block, but such estimates are noninteger. The category by block housing unit estimates are converted to integers in three steps of a controlled rounding algorithm. Let t_{ik} denote the integer valued estimate of housing unit counts in category i in block k obtained by controlled rounding. Then $U_{ik} = t_{ik} - h_{ik}$ represents an undercount for category i in block k if U_{ik} is positive and represents an overcount for category i in block k if U_{ik} is negative. The data set is created by either adding U_{ik} category i housing units to block k of the data file if U_{ik} is positive, or removing U_{ik} category i housing units from block k of the data file if U_{ik} is negative.

The required housing units are selected at random from the housing units in the block. In the first step the required number of housing units are selected for housing units with more than three persons. As households with four or more persons are selected, a sequential count of total persons (the original Census plus those in selected households) is maintained and compared to the synthetic estimate of persons. Household selection is terminated when the required number of housing units have been selected or if the selection of an additional household produces a sum of individuals (the original census plus individuals in the selected households) that exceeds the synthetic estimate of total persons for the block. For the created file of households with four or more persons, the number of persons in each category is computed. Let this vector be $\hat{\mathbf{X}}_{(1)}$. Then a new control vector for one-, two-, and three-person households is

$$\hat{\mathbf{X}}_{(2)} = \hat{\mathbf{X}} - \hat{\mathbf{X}}_{(1)},$$

where $\hat{\mathbf{X}}$ is the original vector of controls.

A set of weights for the one- two-, and three-person households is computed beginning with the block factor

$$a_{4k} = (\hat{X}_{Tk} - \hat{X}_{(1)k}) \left(\sum_{i(1,2,3)} \sum_{j=1}^{m} c_{4i} x_{ijk} \right)^{-1} .$$

Using this block factor, a new set of c-factors is computed for the categories of one-, two-, and three- person households and iterated as described previously. A controlled rounding procedure is applied to obtain integer estimates for the one-, two-, and three-person households. Then two- and three-person households are selected for duplication or deletion.

As at the first step, household selection proceeds so that the sum of individuals, in the original and duplicated households, is less than or equal to the synthetic estimate of total persons for the block.

Finally, weights are created for the one-person households using the control total

$$\hat{X}_{(3)} = \hat{X} - \hat{X}_{(1)} - \hat{X}_{(2)},$$

where $\hat{X}_{(2)}$ is the vector of totals from the created file of two- and three-person households. Controls were also imposed on the synthetic estimator of number of persons by race by tract. In Paterson the race groups used were Black, Hispanic and Other. In Oakland, an Asian group was also used. The one-person household counts summarized on a tract by race basis were allocated to one-person HU categories in proportion to their count in the one, two - and three person household controlled rounding step. The results served as a set of tract HU category controls in a two-way raking procedure of one-person HU counts. The other controls were synthetic estimates of persons in one-person households by block, where the one-person numbers are totals for the blocks less persons in households of two or more persons as reflected in the created data file. The raked one-person households were then control rounded and households duplicated or deleted to complete the transparent file.

For some tracts in Oakland, the difference between the synthetic race total and the combined transparent file 2^+ person household and Census file single person count was negative. Our procedure for transparent file construction focused primarily on maintaining the synthetic estimated total population for the block. (In future applications, we will need to focus on maintaining the synthetic estimated total race population for the block). To correct the current deficiency, we re-visited each block in a problem tract, identified the race involved and eliminated households of two or more persons, containing persons entirely of the race, from the transparent file, until the difference at the block level was positive.

Some block differences were over 50 persons. We arbitrarily chose to limit elimination to two and three person housing units when the difference was less than 11 persons and to use a proportional allocation among all sizes of housing units when the difference was 11 or more.

For U_{ik} additions, we randomly selected as donors housing units from among the h_{ik} housing units in the Census file beginning with Census housing units in block k. In a few cases, some of the h_{ik} housing units were used as donors three times.

For the U_{ik} deletions, the source of the h_{ik} was used to determined which units were deleted. Donors for the U_{ik} deletions were first selected from imputed housing units. Imputed units are units on the Census file that were either identified as not containing sufficient information or were nonrespondents to the mailing and not selected for nonresponse follow-up. If more housing units than imputed units are required for deletion, then other units are deleted. Real respondents that are deleted from a block are used to replace housing units in the same housing unit category in another nearby block. This minimizes the number of actual responding housing units deleted in the file creation operation.

IV. APPLICATION-95 TEST

A Census pre-test, called the 95 Census Test, was conducted in Paterson, New Jersey and Oakland, California. Included were a test of methodology for sampling nonrespondents and a test of post enumeration measurement methods, called ICM in the test. The original objectives did not include transparent file construction. Consequently, as in 1990, survey data processing was oriented toward editing and imputation for persons, as opposed to a 'persons in housing unit' treatment. Such data processing produced some strange ICM microdata observations, such as a household composed only of babies. Such occurrences were edited on a case-by-case basis.

Transparent files were later constructed for both the Paterson and Oakland sites. We present estimates for both sites. The 1995 Test Census reported 127,954 persons in 42,516 occupied housing units and 3,239 vacant units in Paterson. The direct ICM person estimate of number of persons was 145,508. There are 984 blocks in Paterson ranging in size from one to 792 housing units. Twelve percent of the blocks contained fewer than eleven housing units.

In Oakland, the Census reported 333,902 persons in 137,684 occupied housing units and 18,576 vacant units. The direct ICM person estimate of number of persons was 361,546. There are 3315 blocks in Oakland ranging in size from one to 723 housing units. Thirteen percent of the blocks contained fewer than eleven housing units.

Table 1 contains estimates for some types of households, where the estimator is of the form

$$\hat{Y} = \sum_{k=1}^{B} \sum_{i=1}^{n} w_{ik} y_{ik}, \qquad (5)$$

where y_{ik} is the census total for household category i in block k. The weights were chosen to minimize (3), with three iterations of the step (2). The constant K in (4) was set equal to 0.5 in the minimization. In Paterson the number of housing unit categories was about 350 and the number of person categories was 42. In Oakland the number of housing unit categories was about 460 and the number of person categories was 56.

We summarize some properties of the ICM estimates and the constructed transparent files for both sites. Table 2 exhibits persons as obtained in the Census operation, as estimated in the ICM, and as they appear in the transparent file for some ICM publication categories. The last column gives the differences between the ICM estimates and the transparent file. The differences are due entirely to rounding. Rounding includes rounding to integers and the effect of donor selection of households of different sizes in constructing the transparent file. There would be zero difference if real valued household weights were used. The differences are negligible, relative to the standard errors of the ICM estimators.

Our research began under the assumption that the transparent file would provide the official estimates of both

persons and households. Under this assumption, rounding error that is small relative to estimation error is not important. In our original formulation, we constructed the regression estimator without controlling to the block synthetic estimator. Results of this estimation were presented to the National Academy of Sciences Panel on Census 2000 Methodology. At that time it was clear that, at most, the Census Bureau would publish only synthetic estimates of persons in 2000. It was the panel's opinion that household estimates should give block estimates of persons "close" to the block synthetic estimates. Furthermore, it was the opinion of the panel that the block estimates of number of persons constructed by the regression estimator using total controls were "not close enough" to the synthetic estimates. In response to the panel's position, we developed the procedure described in the previous section.

Table 3 contains measures of closeness between the ICM synthetic person estimates and those from the transparent file for blocks and tracts. Let TF_k, SYN_k and CEN_k denote the transparent file, synthetic estimate, and Census count for a characteristic in the k-th area, where an area can be a block or a tract, and let N denote the number of areas. Define

i) Mean Squared Difference

$$= MSD = N^{-1} \sum_{k=1}^{N} (TF_k - SYN_k)^2$$

ii) Mean Absolute Difference

$$= MAD = N^{-1} \sum_{k=1}^{N} |TF_k - SYN_k|$$

iii) Relative Difference

$$= RD_k = (TF_k - SYN_k)\,[CEN_k]^{-1}$$

Table 3 contains the mean squared difference and mean absolute difference for blocks and tracts in both sites. The transparent file procedure imposed control for race at the tract level. Hence, the race differences at the tract level are due to rounding, where the original tract synthetic estimates were not rounded. Because there was no direct control for tenure at the tract level, the absolute differences are larger for the tenure categories than for the race categories.

The differences at the block level in total persons are due to rounding. The synthetic estimators were not rounded, and the mean squared difference is less than twice the squared difference due to rounding to integers. The differences in other categories are larger than those for total persons, because no direct restrictions were imposed on those categories at the block level.

The five percent and 95 percent points of the empirical distribution of the relative differences are also given in Table 3. The relative differences of large absolute value are associated with blocks with small numbers of persons. For example, of the 50 blocks with relative differences exceeding the 95 percent points, the average number of persons were 16 Black, 19 Hispanic and six Other, respectively. In comparison, the average numbers per block in the site were 51 Black, 56 Hispanic, and 32 Other.

V. COMMENTS AND FUTURE WORK

The U.S. Bureau of the Census has decided not to develop a transparent file for the 2000 Census because of timing and accuracy concerns. Although transparent file methodology as presented will not be used in the 2000 Census, there are other applications within the Census context.

First, a transparent file could be used to produce tabulations based on short form data similar to those provided in previous censuses, especially tabulations for households. Such a file would be constructed after the population counts required for apportionment and redistricting are released. A transparent file, called the research file, will be produced for the 2000 Census Dress Rehearsal.

Second, there are several options for application to long form data, depending on the Census environment. In the 2000 Census Dress Rehearsal, a traditional Census, with short forms for all persons that can be contacted and a long form for a sample, will be conducted in one site. In previous censuses, the Bureau developed two sets of weights for long form data, one for person characteristics and another for household characteristics. The weights were created using a raking procedure that controlled to short form marginals. Under the traditional Census, as to be conducted at one site in the dress rehearsal, two scenarios are possible. One is to use the optimization procedure to obtain a single household weight for the long form sample cases. The second is to construct a transparent file of households for long form data.

In the second site of the 2000 Census Dress Rehearsal, the ICM will provide person control totals which can be utilized to obtain household weights using the procedure such as that discussed in Section III. In addition, a transparent file for long form data could be developed as a second step. Timing is less of a concern in transparent file development for long form data than for short form data. The schedule for 2000 requires state population counts by December 31, 2000 and block person counts by April, 2001, with no tight date for household data.

This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a more limited review than official Census Bureau publications. This paper is released to inform interested parties of research and to encourage discussion.

VI. REFERENCES

Bethlehem, J.G. and Keller, W.A. (1987). "Linear Weighting of Sample Survey Data." *Journal of Official Statistics*, **3**, 141-153.

Cox, L.H. and Ernst, L.R. (1982). "Controlled Rounding." *INFOR* **20**, 423-432.

Deville, J.C. and Särndal C.E. (1992) "Calibration Estimators in Survey Sampling." *Journal of the American Statistical Association*, **87**, 376-382.

Deville, J.C. and Särndal C.E., and Sautory, O. (1993). Generalized raking procedures in survey sampling. *Journal of the American Statistical Association*, **88**, 1013-1020.

Fay, R.E. (1990). "VPLX: Variance Estimates for Complex Samples." Proceedings of the Survey Research Section, *American Statistical Association*, 266-271.

Hogan, H. (1992). "The 1990 Post-Enumeration Survey: An Overview." *The American Statistician*, **46**, 261-269.

Hogan, H. (1993). "The 1990 Post-Enumeration Survey: Operations and Results." *Journal of the American Statistical Association*, **88**, 1047-1060.

Huang, E.T. and Fuller, W.A. (1978). "Nonnegative Regression Estimation for Survey Data." Proceedings of the Social Statistics Section, *American Statistical Association* 1978, 300-303.

Isaki, C.T. and Ikeda, M.M. (1996). "Some Estimators of Housing Unit Category Totals in the 95 ICM Via Dual System Estimation and Census Plus." Unpublished manuscript dated 11/13/96, 30 pages.

Isaki, C.T. (1997). "Estimating Occupied Housing Units - More DSE Options." Draft dated 3/21/97, 6 pages.

Isaki, C.T., Ikeda, M.M., Tsay, J.H. and Fuller, W.A. (1998). "A Transparent File for a One-Number Census." Submitted to the *Journal of Official Statistics*, 23 pages.

McLaughlin, G. (1997). "Memo for the Distribution List, Subject: 2000 Post-Census Processing: Transparent File." Dated 3/19/97, 8 pages.

U.S. Bureau of the Census (1996). "The 1995 Census Test: A Compilation of Results and Decisions." DSSD 1995 Census Test Memorandum Series #T-13, 45 pages with attachments.

Zaslavsky, A.M. (1988). "Representing Local Area Adjustments by Reweighting of Households." *Survey Methodology*, vol. **14**, no. 2, 265-288.

Table 1. Household Estimates for Paterson and Oakland

Type of Household	Census	Estimate	s.e.	DSE
1. Paterson				
Black Householder	15,387	18,041	547	18,496
Hispanic Householder	14,764	16,906	308	16,724
Other Householder	12,365	12,852	231	12,813
Owner	13,407	13,812	290	13,686
Renter	29,109	33,987	669	34,347
TOTAL	42,516	47,799	657	48,033
2. Oakland				
Black Householder	57,807	61,415	1,196	61,788
Hispanic Householder	13,693	16,250	520	15,862
API Householder	17,492	18,321	713	17,559
Other Householder	48,692	49,744	1,217	49,369
Owner	56,813	59,551	689	58,798
Renter	80,871	86,179	2,356	85,781
TOTAL	137,684	145,730	2,594	144,579

Table 2: Person Estimates for Paterson and Oakland Pretest

Category	Census File	ICM Estimate	s.e. of ICM	Transparent File	Differences ICM-T File
1. Paterson					
Black Persons	46,673	56,260	2,159	56,266	-6
Black owned	13,767	14,487	767	14,494	-7
Black owned aged 0 to 17	3,528	3,858	248	3,859	-1
Black Rented	32,906	41,773	2,251	41,772	1
Hispanic Persons	52,268	59,476	1,288	59,474	2
Hispanic Owned	14,625	15,929	538	15,915	14
Hispanic Rented	37,643	43,547	1,086	43,559	-12
Other Persons	29,012	29,772	753	29,767	5
Other Owned	14,863	14,741	445	14,713	28
Other Rented	14,149	15,031	590	15,054	-23
TOTAL PERSONS	127,954	145,508	2,746	145,507	1
2. Oakland					
Black Persons	136,997	150,823	3,455	150,815	8
Black Owned	51,199	56,159	1,202	55,978	181
Black Owned aged 0-17	10,775	12,385	480	12,381	4
Black Rented	85,798	94,664	3,017	94,837	-173
Hispanic Persons	48,041	57,205	2,113	57,217	-12
Hispanic Owned	17,562	20,542	704	20,518	24
Hispanic Rented	30,479	36,663	1,854	36,699	-36
Other Persons	93,392	93,949	634	93,966	-17
Other Owned	54,376	54,304	905	54,445	-141
Other Rented	39,016	39,645	2,336	39,521	124
API Persons	55,472	59,569	2,312	59,545	24
API Owned	23,271	24,050	936	24,109	-59
API Rented	32,201	35,519	1,796	35,436	83
TOTAL PERSONS	333,902	361,546	6,938	361,543	3

Table 3. Summary Statistics of Transparent File and Synthetic Person Estimates at the Block and Tract Levels for Paterson and Oakland

A. Paterson - Blocks	Number of Blocks	Mean Squared Difference	Mean Absolute Difference				
Characteristic	N	MSD	MAD	5% of RD_i	95% of RD_i	Mean of TF_i	Mean of CEN_i
Total Persons	984	0.19	0.31	-0.02	0.01	147.87	130.03
Owners	941	19.00	3.05	-0.18	0.19	47.95	46.00
Renters	970	18.54	2.99	-0.15	0.14	103.49	87.28
Blacks	909	14.30	2.68	-0.29	0.32	61.90	51.09
Hispanics	924	16.94	2.97	-0.17	0.20	64.37	56.72
Others	905	7.78	1.72	-0.23	0.33	32.89	32.15
B. Paterson - Tracts	Number of Tracts	Mean Squared Difference	Mean Absolute Difference				
Characteristic	N	MSD	MAD	5% of RD_i	95% of RD_i	Mean of TF_i	Mean of CEN_i
Total Persons	33	0.09	0.26	-0.0007	0.0002	4409.30	3877.15
Owners	33	450.39	16.39	-0.029	0.044	1368.39	1311.58
Renters	33	451.94	16.43	-0.015	0.015	3041.97	2565.58
Blacks	33	1.55	0.99	-0.002	0.004	1705.03	1407.42
Hispanics	33	1.63	1	-0.002	0.002	1802.24	1588.09
Others	33	2.05	1.15	-0.014	0.012	902.03	881.64

C. Oakland - Blocks	Number of Blocks	Mean Squared Difference	Mean Absolute Difference				
Characteristic	N	MSD	MAD	5% of RD_i	95% of RD_i	Mean of TF_i	Mean of CEN_i
Total Persons	3315	0.41	0.36	-0.02	0.01	109.13	100.72
Owners	3154	11.04	2.20	-0.16	0.18	49.16	46.42
Renters	3162	11.06	2.22	-0.16	0.16	65.35	59.30
Blacks	3010	11.57	2.28	-0.16	0.21	50.10	45.51
Hispanics	2639	10.68	2.16	-0.33	0.66	21.68	18.20
Others	2972	5.10	1.39	-0.24	0.33	31.62	31.42
API	2546	12.60	2.25	-0.41	0.56	23.40	21.79
D. Oakland - Tracts	Number of Tracts	Mean Squared Difference	Mean Absolute Difference				
Characteristic	N	MSD	MAD	5% of RD_i	95% of RD_i	Mean of TF_i	Mean of CEN_i
Total Persons	107	0.20	0.31	-0.0003	0.0003	3378.91	3120.58
Owners	107	309.07	12.76	-0.038	0.044	1449.07	1368.30
Renters	107	306.28	12.68	-0.020	0.028	1929.84	1752.28
Blacks	106	2.16	1.13	-0.007	0.003	1422.78	1292.42
Hispanics	105	0.97	0.80	-0.010	0.017	544.92	457.53
Others	107	2.40	1.20	-0.009	0.014	878.19	872.82
API	106	1.62	1.01	-0.011	0.007	561.75	523.32