COVERAGE MEASUREMENT'S IMPACT ON THE CENSUS.

Glenn Wolfgang, Bureau of the Census[1]
Statistical Support Division, Washington, D.C. 20233

KEY WORD: Post-Enumeration Survey.

I. INTRODUCTION

Does the Post-Enumeration Survey have any impact on the census? That question is important to the credibility of a census and its evaluation. However, it is difficult to answer. This investigation of that question uses data from the 1988 Dress Rehearsal Census and its Post-Enumeration Survey (PES).

A census is meant to count all the persons residing in a specified area. When people are missed, not only are the census counts wrong but the census data alone do not provide useful clues about the number who don't respond. A PES provides an independent source of information to be used, along with census data, in dual system estimation to measure the rate of coverage, or completeness, in the census.

If the census data were influenced by the PES, the survey would be viewed as introducing errors in the census itself. If that alone is not serious enough, any errors would very likely be compounded in the coverage measurement estimates for which the survey was conducted. For the most part, it seems reasonable to presume that the PES could not influence census results, because census enumerations are captured before most respondents are aware of the PES. However, occasionally a sample block resident is asked a question before Census Day to confirm a PES address listing. A small fraction of census followup contacts are made after the beginning of PES interviewing. It is also possible that somehow in processing census data the PES sample blocks are handled differently. These are sufficient causes to pursue the issue.

The difficulty is that showing the PES has no impact on the census essentially requires proving an infinite supply of null hypotheses. We may test whether PES areas differ from areas where no PES was done, using specific census variables. The research question places no limits on the kinds of census outcomes open to testing. Furthermore, even an insignificant hypothesis test cannot rule out the existence of an actual difference. To confuse matters more, it is quite possible that any observed differences are due to other causes which happen to coincide with the PES.

To avoid such problems, academic programs train the researcher to redefine the research question or design--usually more extremely than is feasible in this situation. The census is not a closed experiment where all conditions, especially those not yet identified, are under control. Narrowing the question only leaves the study less adequate to its purpose. Still, the more comprehensive the set of census outcomes tested, the more chance meaningless (irrelevant, uninformative, or redundant) variables are included and the greater risk of falsely rejecting a null hypothesis or, alternatively, reducing the statistical power of individual tests to the point where true rejections are missed. The methods for balancing and optimizing these concerns are not fixed. They deserve study as well.

II. EXECUTIVE SUMMARY

The motivation for the study is to seek any evidence that census data was affected by the Post-Enumeration Survey. Testing a variety of census outcomes, no differences attributable to the PES have been found between blocks in which the survey was conducted and paired blocks not surveyed.

The second focus of the report is evaluation of the methodology used to address that primary purpose. For the current data, the methodology provides reasonable support for the conclusions offered, but involves some steps at which different choices could, given more ambiguous data, lead to different conclusions. Suggestions are offered for balancing and improving the analysis' overall comprehensiveness, meaningfulness, statistical power, and control of false null hypothesis rejections.

III. PROCEDURES

There are several steps in the development of the data which determined the breadth, relevance, and error control of the study. First, a sample of blocks outside but paired with those in the PES sample was drawn. For each block (both PES or not), a relevant selection of census data was extracted, reshaped (aggregated and redefined), culled again for relevance and nonredundancy, and submitted to test comparisons.

A. Design of Paired Samples.

Design of the paired samples began with adoption of the sample previously drawn for the PES. PES sample area units are blocks demarcated by roadways or other natural boundaries. All census blocks were assigned to strata designed to be as homogeneous as possible on characteristics such as region of the country, rural-urban location, race/ethnicity of residents, and whether residents own or rent their home. PES blocks were selected with predetermined probability from each stratum. In 1988 each stratum was entirely contained within that area served by one of three district offices (Washington, East Central Missouri, and Saint Louis, MO), field offices from which interviewers were dispatched and their work collected, checked, and shipped to a processing center. For more detail on sampling and other methodology of a PES, read Diffendal (1988).

One rural stratum with a "list/enumerate" type of enumeration was eliminated from the paired comparisons. Since no census or PES address listings were done in those areas before the census, there was less concern about respondent reaction in those blocks. Variables like mailback and delete data were not relevant to these blocks and others were processed on a different schedule. The major results were expected in the areas covered by this study.

For each PES sample block, another block was selected randomly from remaining blocks of the same sample design stratum which also shared housing unit distribution, as defined by five categories: mostly apartments or other units in large multi-unit structures, mostly units in small multi-unit structures, mostly single units, half single units with half small multiunits, and some of each type structure. Notably, pairings were not explicitly controlled on estimated housing unit or group quarters counts.

B. Extraction of Data from Census Records.

The census data used in these analyses were extracted from census data control and results files in October and November of 1988, after late census records had been added. To properly serve the purpose of the study, the data extracted should be comprehensive, yet focussed on meaningful measures of potential PES impact. A meaningful variable is relevant to the research question (e.g. identification numbers say nothing about PES impact). Variables were judged relevant if they appeared to be related to respondent reaction or differences in office handling. Also, meaningful variables are not constant or largely missing. An edit flag that was never checked would be useless in this analysis. Long-form sample data were too incomplete. Meaningful variables add information to that represented in other variables. This criterion leads to reducing redundancy among variables, as described in a later step.

To guard against attributing to the PES some extraneous influence on the census, variables which appeared to be out of the range of PES influence but could influence the other variables selected were sought. For example, the percentage of short census forms administered in the block is determined long before PES listings and should not be altered by any subsequent census or survey operation. But if that variable caused differences in response or edit rates while also covarying with PES sample membership, it could mediate a false conclusion about the PES.

C. Computing Block-Level Variables.

The information extracted from the person records of the census files were collapsed to the block level. Some variables were descriptive of the household, or even the block as a whole, rather than a person. When aggregating over person records to block level, only one record per person, address, or occupied housing unit, as appropriate, was counted. Since the size of the block could be a common factor in all block data and the effect of block size was not controlled in block pairing, block size could mask or confound other effects of interest. The count of persons, of housing units or of occupied units in the block were kept for background analysis. Other variables were computed as proportions of the appropriate block size count. When a unit had missing values for a variable, it was not figured into numerator or denominator of the block level result.

Also, the census data were recoded in an attempt to better isolate different dimensions of census impact. When categorical variables had more than two values, such as unit status (occupied, vacant, or deleted), the block level record contained a sum or proportion of households in each category, rather than a harder to interpret statistic confounding all those categories.

D. Identifying Redundant or Irrelevant Information.

Being comprehensive in the early stages of the study meant retaining variables that were not necessarily meaningful. Review of their descriptive statistics or simple correlations helped weed out useless variables. A factor analysis reported below continued the process. Variables were not dropped just because the name implied redundancy or constancy; empirical evidence was stressed.

E. Multiple Hypothesis Testing.

The variables of interest were submitted to t-tests and Signed Rank, i.e. Wilcoxon Matched Pairs, tests (Marascuilo and McSweeney, 1977, pp. 337-339). The differences from sample to out-of-sample blocks are the focus of either test. The t-test naturally produces the common point and variance estimates that meaningfully describe the variables being tested, but its assumption of normality in the variable being tested is important when there are fewer than thirty units. Since the variables here are proportions, frequently with values tending toward the floor or ceiling values, 0.0 or 1.0, the distributions are not generally expected to approximate normality. The consistency of the observed test probabilities for each variable should show that the tests were properly applied.

A concern for the likelihood of erroneous conclusions of a difference (Type I errors) arises from doing many hypothesis tests for one basic research question. The more independent tests one conducts at a given level of significance, the more likely one will encounter by chance a significantly rare observed probability and falsely conclude there is a difference. One may adjust the level of the criterion test statistic or of the criterion probability for the individual tests so that the probability of observing a difference (when there really is none) for the whole set of tests remains at the desired nominal level of significance, traditionally 10% in Census Bureau research. A modified Bonferroni procedure (Games, Paul, 1971, "Multiple Comparisons of Means", American Educational Research Journal, 8, 531-565.) computes such a reduced criterion probability value. The per-comparison reduced significance level, α_R, is a function of the nominal level of significance, α_N, and the number of comparisons, c, specifically involving the c^{th} root:

$$\alpha_R = 1 - (1 - \alpha_N)^{**(1/c)}.$$

Independence of the hypothesis tests is central to this application. Testing the same variable under different names will simply duplicate results while confusing Type I error interpretations. That is the importance of the redundancy analyses. About twenty-one variables were judged to be unrelated enough to be tested independently. If one wishes to control the overall rate to 0.10, the reduced, per-comparison significance level computed by the modified Bonferroni procedure is 0.005. Tests with an observed probability less than 0.005 are considered significant.

An additional perspective on controlling Type I errors is provided by Saville (1990). He argued that an applied statistician should use the same test criterion for multiple comparisons as for a single test (e.g. 0.05) and depend on replication to justify conclusions of significant alternative hypotheses. Most of the argument rests on the importance of consistency, which other scholars may define differently. He also asserts that concern over Type I errors is overblown because true differences are likely where hypothesis tests are employed and "Thus in general there is more opportunity for Type II errors to occur than Type I errors." This may be true in academic settings where experiments can be selected and designed to ensure more striking conclusions based on significant differences. If anything, in this study there should be few true differences, but to presume so without prior empirical results begs the question of the present analysis. So, replication becomes an important part of confirming any conclusions about differences.

Each district office's data is considered a replicate, i.e. an independent data set for investigating PES impact on the census. That view, if acceptable, provides additional control of Type I error. Significant differences not observed in all or at least two of the replicates will be presumed significant by chance. If a specific hypothesis based on differences in the replicates can explain the pattern of results, it may be recommended for future testing.

IV. DESCRIPTION OF VARIABLES

The variables measuring respondent reaction and its covariates may be organized into groups: block size, miscellaneous potential covariates, population coverage, housing unit status, mailback, field response, and edit/quality. The specific variable names and descriptions are listed in Table 1. The variable names used in that table will be used in subsequent tables and text since the descriptions are often unwieldy and distracting in a sentence. The rationales for attributing differences in groups of variables to PES intervention are presented here.

After eliminating the most redundant and noninformative variables, only one potential mediating covariate remained. Since long forms are notorious for discouraging response, it is possible that response patterns are influenced more by form length than by PES contacts. If short form rate consistently varies with any significant response variable, further analyses may be necessary to remove its effect on that variable.

Population coverage variables are proportions of persons in the block who fall into demographic subgroupings used in final PES estimation. Differences that may be difficult to detect in the simple population counts may stand out when the focus is on demographic proportions. Furthermore, if respondents react to the PES in any way, it may be in the same way as to other causes of undercount and may show up in demographic breakdowns shown to relate to census undercount (Fay, Passel, and Robinson, 1988). Such findings would be especially damaging to coverage measurement, since undercount is its primary concern. Most of these variables are defined on only one demographic characteristic used in poststrata breakdowns to catch overall effects, but one, BM2044, used age and race and sex to specify a group that has been identified as among the most undercounted.

Housing unit status variables, proportions of the block's units that were added to the address lists after Census Day or were classified vacant, deleted, or occupied, could be sensitive to community reaction against extra survey involvement.

Mailback rates are viewed as prime gauges of respondent reaction. This group of variables is represented by mail return rates (based on occupied housing units) and mail response rates (computed over all housing units) as well as mean mail check-in date (measuring promptness of respondents and of office handling).

Field response variables include the mean date of check-in for both non-response and edit followups along with the rate of the latter. Reactions related to PES activities could show up in mailback nonresponse or a questionable housing unit status, leading to field followup.

Edit and quality variables are derived mostly from office-use-only flags marked on the census forms in processing. Some of them with questionable importance and reliability were finally designated as meaningless due to near constancy, less than five percent of households' forms were marked.

V. REDUNDANCY ANALYSIS

The first step in eliminating redundant variables used simple correlations as a basis. Whenever the average for two variables of the absolute values of the correlations from the six data sets (in-sample and paired block data crossed by three district offices) was more than 0.90, only one of the variables was kept. This and elimination of variables which were nearly constant across all replications left 39 variables to consider.

Factor analyses were done to identify other combinations of variables acting in concert. Factor analysis was not designed to settle what we really wish to decide, "How many independent dimensions worth testing are in the data?" At best, it empirically describes dimensions in the data, guiding the essentially subjective choices.

There were nine replications for factor analyses--in-sample block data, paired-block data, and the

Table 1. Census Variables Describing Blocks.

Variable		Description(#=count; %=rate)	Base
		----- BLOCK SIZE -----	
1	P	# Persons not in Group Quarters	
2	HU	# Housing Units	
3	OCC	# Occupied Housing Units	
		----- OTHER POSSIBLE COVARIATE -----	
4	SHORTF	% Short Forms	~hu
		----- POPULATION COVERAGE -----	
5	HHSIZE	Mean Block Household Size	occ
6	HHHEAD	% Household heads/mates	~p
7	NOTKIN	% Unrelated roomers	~p
8	MALE	% Males	~p
9	WHITE	% Whites	~p
10	BLACK	% Blacks	~p
11	HISP	% Hispanic	~p
12	MARST	% Married (not separated)	~p
13	AGE0	% Aged 0-9	~p
14	AGE10	% Aged 10-19	p
15	AGE20	% Aged 20-29	p
16	AGE30	% Aged 30-44	p
17	AGE45	% Aged 45-64	p
18	AGE65	% Aged Over 64	p
19	BM2044	% Black Male Aged 20-44	~p
20	OWN	% Homes Owned (Tenure)	~hu
21	MU	% Multiunit/apartment Structures	~hu
		----- HOUSING UNIT STATUS -----	
22	OCCRATE	% Occupied Units	hu
23	ADD	% Units Added after First Listing	~hu
24	VAC	% Unit Status: Vacant	~hu
25	VACFIN	% Final Unit Status: Vacant	hu
26	DEL	% Unit Status: Deleted	~hu
27	DELFIN	% Final Unit Status: Deleted	hu
		----- MAILBACK -----	
28	MAILFOR	% Mail Response (forms)	~hu
29	MAILRET	% Mail Return	occ
30	MAILDAY	Mean Mail Return Check-in Date	~hu
		----- FIELD RESPONSE -----	
31	NRFUDAY	Mean Julian Date of Nonresponse Followup Check-in	~hu
32	FFUDAY	Mean Field Followup Check-in Date	~hu
33	FFURATE	% Field Followup Check-in	hu
		----- EDIT & QUALITY -----	
34	LASTR1	% Last Resort Data	hu
35	LASTR2	% More Last Resort	hu
36	EDITOK	% Passed Edit without Followup	hu
37	TELEFU	% Telephone Followup	hu
38	INMOVE	% Inmovers Respond	hu
39	TELEINT	% Telephone Interview	hu
40	CLOSE	% Close Out	hu
41	RE	% Replaces Data from Prior Form	hu
42	QASSIST	% Questionnaire Assistance	hu
43	GOLDP	% Goldplate to Pass Edit	hu
44	NODAY	% No Check-in Dates	hu

NOTE: The tilde (~) in the "Base" column denotes when missing or irrelevant data in the original ACF/DCF/ID variable were excluded from the computation's numerator and denominator. The count used as the denominator of a rate also specifies whether the numerator was tallied over households or persons.

difference score data for each of three district offices. The difference score was the statistic analyzed by both parametric and non-parametric test procedures, but research conclusions are based on the undifferenced variables, so all three seem legitimate for analysis of variable interrelationships.

Summary statistics from an unrotated principal components analysis are presented in Table 2. Replications from the Washington state district office are left out of this table because there are fewer blocks than variables. All block size correlates, P, HU, and OCC, were carried along in analyses and tables because of their importance to the other variables.

Table 2. Number of Principal Components for Percents of Total Census Variance in Six Replications.

District Office Score	Central MO PES	Not	Diff	St. Louis PES	Not	Diff
% of Total Variance per # Components						
80%	10	10	17	8	7	16
90%	16	15	23	13	12	21
95%	20	18	26	17	16	25
98%	24	23	30	22	20	29
# Components Exceeding 1% of Total Variance						
	26	26	27	24	23	25
% of Total Variance in the First Component:						
	45	47	10	56	57	18

The overall picture emerging from the principal components results is of many dimensions gradually (rather than suddenly) dropping in size. That pattern may be a sign that the aim of identifying a broad range of independent variables had been accomplished. About twenty-six difference score components cover at least 95% of the total data variance and are at least as large as 1% of the total data variance (about equal to 40% of the average component variance). By choosing more or less stringent percentages from the table, one might suggest there are from fifteen to thirty important independent dimensions in this data.

Next, a varimax rotation of principal components in each of the nine replication was conducted. Washington State replications could be included here because the number of factors of interest are limited to less than the number of blocks. The rotation expresses orthogonal dimensions of the data as strongly as possible in terms of original variables. Table 3 summarizes twenty-six variables that were organized into nine factors by the following criteria: (1) Consistency--The same pattern, in terms of size and sign, of factor loadings were evident in all replications. (2) Magnitude--Over half, across replications, of an included variable's factor loadings had absolute values over 0.5. (3) Exclusiveness--Other variable's factor loadings were between -0.3 and 0.3.

As one would expect with factor analysis, patterns did not emerge without exceptions. The consistency criterion was not clearly met in five of 81 replicated factors represented in the table because the group of variables appeared to split in partial alignment with other variables on different factors. In the three replications for one district office's edit followup factor, erroneous, near-zero EDITOK data caused a break in the pattern. MARST, AGE0, and FFURATE (under the deletes factor) were close but did not satisfy the magnitude criterion. The exclusiveness criterion was not met in seven instances where two factors aligned together as one; only one of those was outside Washington, where limited sample size packed variables onto fewer dimensions. Frequently, single variables otherwise excluded from these factors would also join a factor's dimension. Those alignments appeared to be random across replications, however, and are considered effectively independent in the universe which these replications represent.

Table 3. Multivariable Factors from Varimax Rotations.

Factor Name	Rank	Size	Variable Name	Loading
Vacants	1-1-5	3.3	VAC	89.3
			VACFIN	81.8
			FFURATE	67.7
			OCCRATE	-61.8
Block Size	1-3-4	3.1	P	96.1
			HU	95.6
			OCC	96.1
Race	1-3-4	2.9	WHITE	-76.7
			BLACK	86.1
			BM2044	68.9
Single Head with Kids	2-4-10	2.5	HHSIZE	60.9
			HHHEAD	-75.2
			AGE0	35.3
			AGE10	63.3
			AGE65	-50.0
Deletes	1-5-6	2.4	DEL	72.7
			DELFIN	87.4
			OCCRATE	-53.7
			FFURATE	34.2
Homeownership	2-7-13	2.0	OWN	-65.7
			MU	80.8
			MARST	-36.4
Mail Return	1-6-11	1.8	MAILFOR	70.4
			MAILRET	74.6
Middle Age	3-10-12	1.5	AGE30	80.2
			AGE45	-54.0
Edit Followup	6-8-9	1.9	EDITOK	-72.4
			TELEFU	55.77

For the most part variables did seem to align on factors that on their face measure much the same thing (e.g. mail response and mail return rates). The high loading of BM2044 with BLACK gives the first clue that, in this data, highly undercounted persons react much like others of their own race. The most surprising factor is "single head of large household with youths". It comes from a correlation among block-level age and household composition variables; larger household sizes go along with lower proportions of mates for the head of household, fewer persons aged over 64, and more aged under twenty. Some other results are understandable but might not have been anticipated. For example, homeownership is likely to be low in blocks comprised largely of multiunits or where most persons are single. Similarly, occupancy rate is a complement to vacancy and delete rates and field followup consists mostly of checking vacants and deletes, so it is appropriate that OCCRATE and FFURATE joined those factors.

The factor analyses thus provides some different clues about which variables are essentially different. If the eight multivariable factors relevant to the research question came from twenty-three variables, fifteen variables are redundant, leaving twenty-five for testing. Among the office edit variables, four were found to have so little information that they provided no variance to the factor analyses in at least one district office and little to the others. They were among those viewed as having uncertain reliability and are dropped as unimportant. The conclusion of the redundancy analysis is that about twenty-one meaningful, independent tests can be run on the census data.

One weak step in this reasoning deserves clarification. Just because variables covary highly in all current replications, does not mean that they bear no independent information. The rationale for dropping redundant variables is that, after accounting for the common variance, the remaining variance is unimportant. The author admits to some unease in some cases, specifically HHSIZE, MARST, and OWN, but accepts the compromise in the face of the need to control Type I and II errors.

VI. MULTIPLE COMPARISON RESULTS

To assess the possibility that PES impact on the census distorted either census or coverage measurement results, all worthy variables were compared. A reduced significance level computed with a modified Bonferroni criterion and a replication approach were adopted to control Type I hypothesis testing decision errors while still preserving power to observe truly significant differences. As described in the preceding section, for the main research question an overall significance criterion of 0.10 and an estimated number of independent comparisons set at 21 were used. The reduced probability level, α_R, was then 0.005.

To detect extraneous causes of differences in the paired data, block size, three variables treated as one because of redundancy, and short form rate were tested. Only in the St. Louis office data was block size found to differ with PES contact (using $\alpha_N = 0.10$ and $\alpha_R = 0.051$); short form rate was not found to differ. Since neither related consistently with any other variable, it is not likely that either could mediate a false conclusion that PES affected the census.

The results of the hypothesis tests are in Table 4. The variables marked with an asterisk in the tables are those that should be considered part of the simultaneous testing. The tables include the tests of fifteen variables designated redundant by the factor analysis results and of four marginally unimportant variables and of four potential covariates. They are reported only to help evaluate the methods of this study. Except for the block size variables in the St. Louis district office, none of the tests reached criterion significance. The lack of any significant results leads to a conclusion that no effect on the census is manifest in this data; given the independence of block size from the primary variables and no conflicting results among variables deemed redundant, there is no evidence to refute that conclusion.

VII. SUGGESTIONS FOR IMPROVEMENTS IN THIS STUDY

A. <u>Design of Paired Samples:</u>
1. Include list/enumerate strata, even if separate pairing strategies and different dependent variables must be used.
2. Control for block size, so that predicted housing unit counts are equal from in-sample to not-in-sample pairs.

B. <u>Extraction of Data from Census Records:</u>
1. Expand comprehensiveness--Brainstorm or review census data to find other variables not already extracted that could be responsive to PES impact (e.g,. edit failure rates, both content and coverage).
2. Identify meaningless data--Interview those familiar with field and clerical handling of the data to judge how well data might relate to reactions to the PES. For example, if an edit check box was never marked, it should be dropped from analyses.
3. Empirical methods for selecting variables--Factor analyses do not definitively tell how many important independent tests are in the data, but help in understanding the data and informing the decisions. What other methods would help or suffice?

C. <u>Computing Block-Level Variables.</u>
1. Review variable recodes, considering combinations--Perhaps there are other recodes of census data aggregated to block level which better represent a new or more independent dimension of possible PES impact.
2. Review categorical variable breakdowns. For example, would different recodings of race data serve better for data from other areas.

D. <u>Multiple Hypothesis Testing.</u>
1. Control Type I error--the method used here to approximate a value c for computing a reduced significance level applied to each hypothesis test is exploratory; alternatives with better balance of meaningfulness, comprehensiveness, statistical power, and control of Type I error may be developed.
2. Control statistical power--no attempt has been made to do a formal power analysis. The overall significance level of 0.10 is sometimes considered liberal, but does lend more power to individual simultaneous tests.

VIII. REFERENCES

Diffendal, G. (1988), "The 1986 Test of Adjustment Related Operations in Central Los Angeles County," <u>Survey Methodology</u>, 14, 71-86.

Games, P. (1971), "Multiple Comparisons of Means," <u>American Educational Research Journal</u>, 8, 531-565.

Fay, R.E., Passel, J.S., and Robinson, J.G. with Cowan, C.D. (1988), "The Coverage of the Population in the 1980 Census," 1980 Census of Population and Housing Evaluation and Research Report PHC80-E4, Bureau of the Census, Washington D.C.

Marascuilo, L.A. and McSweeney, M. (1977), <u>Nonparametric and Distribution-Free Methods for the Social Sciences</u>, Monterey, CA: Brooks/Cole.

Saville, D.J. (1990), "Multiple Comparison Procedures: The Practical Solution," <u>The American Statistician</u>, 44, 174-180.

IX. ACKNOWLEDGEMENTS

Mary Mulry and Aref Dajani contributed to this project: They planned and executed and documented the pairing of blocks, and arranged for the first stage of extracting census data.

Table 4. Mean Differences and Probabilities for t-Values and Wilcoxon (Signed Rank) Test Statistics in Three Replicates: the Washington State (WA), East Central Missouri (ECMO), and St. Louis (SL) District Offices.

	WA (n<=32)			ECMO (n<=258)			SL (n<=166)		
Variable	Mean Diff.	t Prob.	Signed Rank Prob.	Mean Diff.	t Prob.	Signed Rank Prob.	Mean Diff.	t Prob.	Signed Rank Prob.
P	-0.62	0.92	0.66	0.81	0.90	0.98	23.19	0.021	0.017
HU	2.43	0.38	0.50	1.45	0.69	0.97	11.72	0.017	0.026
OCC	0.25	0.92	1.00	-0.57	0.80	0.98	8.49	0.022	0.022
SHORTF	-0.07	0.15	0.18	0.01	0.46	0.65	-0.01	0.53	0.97
HHSIZE	-0.26	0.09	0.23	-0.04	0.57	0.41	-0.03	0.70	0.80
*HHHEAD	-0.00	0.91	0.68	-0.01	0.64	0.72	-0.002	0.84	0.71
*NOTKIN	0.02	0.12	0.28	0.01	0.068	0.27	0.001	0.88	0.77
*MALE	0.02	0.49	0.70	-0.01	0.12	0.48	0.002	0.79	0.85
WHITE	0.01	0.88	0.65	0.002	0.88	0.87	-0.01	0.67	0.76
*BLACK	-0.001	0.97	0.46	-0.002	0.83	0.79	0.01	0.71	0.90
*HISP	-0.01	0.83	0.67	0.001	0.58	0.66	0.002	0.28	0.39
MARST	-0.06	0.11	0.05	-0.000	0.99	0.91	0.01	0.60	0.99
AGE0	-0.03	0.29	0.20	0.003	0.73	0.63	-0.004	0.57	0.65
AGE10	0.01	0.90	0.93	-0.003	0.73	0.42	-0.01	0.37	0.57
*AGE20	0.03	0.31	0.48	-0.01	0.67	0.91	0.000	0.96	0.94
*AGE30	-0.02	0.63	0.54	-0.01	0.43	0.54	0.003	0.74	0.66
AGE45	0.06	0.16	0.08	0.02	0.13	0.30	0.01	0.25	0.33
AGE65	-0.04	0.17	0.23	-0.01	0.43	0.27	-0.004	0.78	0.45
BM2044	0.01	0.25	0.20 !	0.001	0.78	0.60	0.01	0.30	0.20
OWN	-0.06	0.49	0.51	-0.01	0.66	0.91	-0.01	0.70	0.80
*MU	0.01	0.76	0.25	-0.01	0.73	0.72	-0.01	0.52	0.54
OCCRATE	-0.05	0.21	0.16	-0.002	0.92	0.63	0.02	0.41	0.45
*ADD	-0.004	0.90	0.89	0.02	0.13	0.26	0.002	0.90	0.62
VAC	0.002	0.96	1.0	0.004	0.72	0.56	-0.01	0.56	0.31
*VACFIN	0.01	0.58	0.83	-0.01	0.42	0.88	-0.01	0.34	0.17
DEL	0.07	0.12	0.15	0.01	0.59	0.95	-0.002	0.90	0.90
*DELFIN	0.03	0.34	0.44	0.01	0.24	0.50	-0.004	0.75	0.97
*MAILFOR	-0.11	0.19	0.21	0.02	0.45	0.62	0.01	0.70	0.86
MAILRET	-0.11	0.15	0.17	-0.01	0.63	0.71	-0.002	0.92	0.85
*MAILDAY	-1.60	0.18 ?	0.24	-0.42	0.50	0.46	0.17	0.63	0.31
*NRFUDAY	-4.52	0.30 ?	0.62	-0.55	0.42	0.64	-0.04	0.96	0.61
*FFUDAY	-1.64	0.29 ?	0.30	-1.44	0.48	0.61	-1.19	0.27	0.98
FFURATE	0.05	0.33	0.50	0.01	0.56	0.70	-0.01	0.43	0.51
*LASTR1	-0.03	0.06	0.13	0.01	0.32	0.41	-0.01	0.10	0.12
LASTR2	#			0.000	0.55	0.61	0.001	0.065	0.11
*EDITOK	0.03	0.41	0.40	-0.01	0.45	0.18	-0.002	0.23	0.35
TELEFU	-0.05	0.13	0.19	-0.01	0.72	0.59	0.01	0.64	0.94
*INMOVE	-0.002	0.44	0.42 !	-0.003	0.29	0.56	0.002	0.54	0.78
*TELEINT	0.01	0.78	0.42	-0.01	0.15	0.14	0.004	0.30	0.23
CLOSE	#			-0.001	0.82	0.66	-0.001	0.83	0.92
RE	#			0.000	0.86	0.81	0.000	0.88	1.0 !
QASSIST	-0.002	0.33	1.0 !	#			0.000	0.14	0.18 !
*GOLDP	-0.01	0.28	0.31 !	-0.002	0.48	0.49	0.002	0.53	0.55
*NODAY	0.01	0.21	0.23 !	0.003	0.04	0.04	-0.000	0.45	0.72

* One of the primary simultaneous hypothesis tests: criterion probability = 0.005.
\# All data constant at zero.
? Due to missing data, there are fewer than thirty differences on which to base test.
! Fewer than ten nonzero differences on which to base test.

[1]This paper reports the general results of research undertaken by Census Bureau staff. The views expressed are attributable to the author and do not necessarily reflect those of the Census Bureau.