MEASURING CONSISTENCY OF POST-STRATIFICATION VARIABLES FOR HOUSING UNITS IN THE CENSUS 2000 ACCURACY AND COVERAGE EVALUATION

Joseph Burcham, U.S. Bureau of the Census
4301 Suitland Road, Rm 1107/2, Suitland, MD 20746

Key Words: Dual System Estimation, housing unit coverage

1 Purpose

To examine the level of consistency of different post-stratification variables in classifying housing units to categories of interest on two different address listings.

While this study focuses on inconsistency of housing units in the census, a similar study to this one was conducted on persons in the census (Farber and Davis, 2001).

2 Dual System Estimation

The Census Bureau used dual system estimation to measure the coverage of housing units in Census 2000. In dual system estimation, a sample of housing units from the census is matched to a sample of housing units from an independent listing.

When an address appears on the independent list but not the census list, this might mean the census missed a housing unit. When an address appears on the census list but not the independent list, this could mean that the census erroneously enumerated a unit. Matching results are used to estimate the proportion of housing units missed and the proportion of housing units erroneously enumerated by the census. The dual system estimation process involves using both of these components to produce coverage estimates. See dual system estimation formula in Barrett (2001).

Estimates are produced for several different post-strata. Five post-stratification variables were used in Housing Unit Dual System Estimation for the Census 2000 Accuracy and Coverage Evaluation (A.C.E.):

- Occupancy
- Race Domain of Householder
- Size of Structure (# of units in structure)
- Metropolitan Statistical Area (MSA) / Type of Enumeration Area (TEA)
- Census Region

A certain combination of the levels of these variables defined the post-strata. Housing units in the census sample and housing units in the independent sample are classified into the post-strata. This classification is done independently for each sample. For matches, a unit is consistent if it was assigned to the same category of interest in the census listing as it was in the independent listing. Otherwise, the unit is inconsistent. A category of interest may be a post-stratum or it may be a particular level of a post-stratification variable, depending on the level of consistency that we are examining.

Of the five original variables, we included occupancy, race domain, and size of structure in our consistency analysis. We did not include MSA/TEA or Census Region because these variables came from the same source for both listings. So, it was almost impossible for these variables to be inconsistent.

3 Inconsistency and the Effects of Inconsistency

Inconsistency is associated with several different factors:

- *time lag* - this is the amount of time between the point that a piece of information was collected in the census and the point that it was collected in the independent listing. Because characteristics of a unit can change over time, we can see inconsistent information for a unit between the two points.

  One example of time lag deals with the occupancy variable. For one listing, a lister may have visited a unit when it is occupied. Then, a few months later when conducting the second listing, a lister discovered that the same unit was then vacant.

- *different respondents in the same household* - we may contact one respondent in one listing and a different respondent in a different listing and there could be some confusion, such as who the householder is, so the two respondents might give us two different races for the householder.

- *differences in the size of structure assignment* - The size of structure variable was created a little differently in the census listing than it was in the independent listing, which

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

could cause units to be classified to different size of structure categories.

One example of this difference has to do with mobile home parks. In the census we counted all units within a common house number/street name. For mobile home parks in the independent listing, we counted units within mobile home park name. It is possible that there could be multiple house number/street name combinations that correspond to a single park name. Because of this, units can fall into different size of structure categories on the two listings.

- *differences in imputation methodology for the census and independent samples* - imputation methods may have differed on the two listings, causing a unit to be assigned to a different category on the census listing as the category it was assigned to on the independent listing.

The post-strata are created with some expectation that:

- there is a homogeneous capture probability in the census among the housing units within a post-stratum, and

- heterogeneous capture probabilities for housing units in different post-strata

When a housing unit shows inconsistency for a given variable in the two listings, this inconsistency can generate heterogeneity bias in the dual system estimates if the capture probabilities between the two post-strata differ significantly See Griffin (2000) for more information on heterogeneity bias. For example, suppose that small multi-unit structures have a much lower capture probability than single unit structures. Suppose that in the census sample, many small multi-units were incorrectly classified as single units. This classification error of multi-units to single units would result in the single unit cell appearing to have a lower capture probability, closer to that of the small multi-unit cell, but in reality this single unit capture probability would be biased downward.

4 Measuring Consistency

Inconsistency applies only to matched addresses. Ideally, we would like to examine whether or not an address (matched or unmatched) is classified in error to a particular category of interest. However, we had to limit our consistency analysis to matched addresses because these are the only addresses where we have evidence of classification error. The fact that the two lists disagree on the category in which to place the address provides evidence that at least one list classified the address in error. The nonmatches could be classified in error also but we do not have information telling us about this.

We did not take additional steps to measure classification error for the nonmatching addresses because:

- these additional steps would require very expensive field work,

- the number of nonmatches is small compared to the matches (from the housing unit coverage study, only four percent of addresses in the independent sample are nonmatches), and

- the distribution of classification error in the nonmatches should be the same as the distribution in the matches. This is true because variables involved in the matching process are not related to post-stratification variables.

To measure the magnitude of consistency, I produced several consistency tables which are shown in this paper.

5 Results

I produced consistency tables for each variable individually. Each consistency table compares the responses from the census sample to those of the independent sample. I also tested for significance of consistency and compared the consistency levels among the different variables.

5.1 Occupancy

Table 1 at the end of this paper is the consistency table for occupancy. The shaded cells in the diagonal are the consistent cases while the off-diagonal cells are inconsistent. The interesting thing about this table is that the units classified as occupied in the census and vacant in the independent list are more than twice as many than those classified as vacant in the census and occupied in the independent list.

The overall proportion consistent for occupancy is 0.957. Non-imputed cases show a much higher consistency (0.958) than imputed cases (0.634), although we did not test the significance. The imputed cases are so few in number that they do not have a big impact overall on proportion consistent.

5.2 Size of Structure

Table 2 is the consistency table for size of structure. As shown in the table, small multi-units are the most inconsistent when looking at either sample. However, for units classified as small multis in the independent sample, most of the switching occurs with large multis, but for units classified as small multis in the census, most of the switching occurs with single units. It is not clear to us yet why the switching would go in two different directions depending on the sample.

5.3 Race Domain

Table 3 is the consistency table for race domain. The most consistent category is the White/American Indian off reservation category. The least consistent category is the Hawaiian/pacific Islander category. For most of the categories, most of the switching happens with the White/American Indian off reservation category.

One thing worth noting about the table is the possible effect of inconsistency on heterogeneity bias. A fair amount of switching occurs between the Asian category and the Hawaiian/Pacific Islander category. We know from the housing unit coverage study (Barrett 2001) that the capture probabilities for these two groups are significantly different. The inconsistency causes the categories to become more similar. The effect on the Asian category would be smaller, due to the much larger number of units in this category than in the Hawaiian/Pacific Islander category. However, the Asian category is relatively small, so classification error here could cause the Asian capture probability to become more similar to the White/American Indian off res category. If there was enough inconsistency between the two groups, the capture probabilities of the groups could become similar enough to be not statistically different.

The overall proportion consistent for race domain is 0.951. Non-imputed cases show a higher consistency (0.961) than imputed cases (0.844), although we did not test the significance. The imputed cases are so few in number that they do not have a big impact overall on proportion consistent.

Something to note about all of these tables is that we are looking at inconsistency only for the levels of an individual post-stratification variable. If we were looking at a cross-classification of the levels of two or more variables, there would be more cells and therefore more inconsistency.

5.4 Significance Testing for Consistency

I computed the kappa statistic (Agresti, 1990) for each of the variables in this study to measure the strength of agreement between the census sample and the independent sample. We use this formula on a consistency table. The formula gives us values that range from 0 to 1 that indicate the strength of agreement for a variable of interest between any two listings. If the agreement that we see between the two listings is due to random chance, the formula gives us a value of zero. If we have perfect agreement between the two listings based on variable of interest, the formula gives us a value of one.

For our consistency tables, we would expect the values to be between 0 and 1 but closer to 1 because both listings are supposed to get the same information for a housing unit.

Table 4 shows 90% confidence intervals for kappa for each of the variables.

Table 4. Confidence Intervals for Kappa Statistic

Variable	Confidence Interval
Occupancy	(0.702,0.709)
Size of Structure	(0.853,0.856)
Race Domain	(0.891,0.893)

Consistency appears to be pretty high for all of the variables. We see that occupancy appears to have a lot lower consistency than the other two, and this difference is statistically significant. This low consistency may be due to the time lag associated with the occupancy variable.

6. Conclusions

Major conclusions of this paper are as follows:

- There are several contributing factors to inconsistency, such as time lag, different respondents, and differences in the way variables were created.
- Occupancy shows a significantly lower consistency level than size of structure or race domain.
- Inconsistency can affect heterogeneity bias in the coverage estimates.
- In the future, the Census Bureau should conduct more research into ways to reduce inconsistency. Some inconsistencies would be harder to reduce than others, such as different respondents providing different answers. We feel that the size of structure algorithm could be made more similar between the two listings.

7. Further Research

We have plans to look into the effects of inconsistency on the coverage estimates. To do this type of research, ideally we would like to know whether the census or independent list has the correct value for each record. But we do not have this information. So, to measure the effect of inconsistency on the coverage estimates, we will compute coverage estimates under two different assumptions:

- the assumption that the census classified all addresses correctly and any misclassification is due to independent list misclassification
- the assumption that the independent list classified everything correctly and any differences are due to census misclassification

This will give us a range of values showing the worst-case scenario of how the coverage factors might be affected.

Another study I plan is to recompute consistency estimates and coverage estimates using a revised size of structure algorithm on the independent listing. As mentioned earlier in the paper, the size of structure variable was created differently in the census listing than it was in the independent listing. There were different reasons that make the two assignments different, depending on the type of unit. For mobile home parks inside areas that contain predominantly house number/street name addresses, size of structure was created by counting units within house number/street name on the census listing and by counting units within mobile home park name on the independent listing.

We have recently been working on splitting addresses in the independent sample into house number and street name, and then re-computing size of structure. Using the new size of structure assignment, I plan to compute consistency estimates to determine if there was an improvement in the new size of structure assignment, and if so, how much of an improvement there was.

Something else that would be interesting would be doing more research into reasons for inconsistency. For example, in this research we could determine whether most of the addresses are inconsistent due to time lag, different respondents, or something else. But in order to do this we would probably need to do extra field work.

Finally, an interesting project would be examining gross misclassification error for both of the listings. When simulating the coverage estimates in the way that I discussed earlier in this section, if there is similar misclassification on the two listings, the effects of misclassification will cancel out.

8. References

Agresti, A. (1990), *Categorical Data Analysis*, New York: John Wiley and Sons.

Barrett, Diane F. (2001), *Executive Steering Committee on Accuracy and Coverage Evaluation Policy II Report Number 17: Census 2000 Housing Unit Coverage Study*, Decennial Statistical Studies Division Census 2000 Procedures and Operations Memorandum Series #U-10R2, Bureau of the Census.

Farber, Jim, and Pete Davis (2001), "Consistency of Census 2000 Post-stratification Variables," *Proceedings of the Section on Survey Research Methods*, American Statistical Association.

Griffin, Richard (2000), *Accuracy and Coverage Evaluation Survey: Definition and Explanation of Correlation and Related Heterogeneity Bias*, Decennial Statistical Studies Division Census 2000 Procedures and Operations Memorandum Series #Q-35, Bureau of the Census.

Table 1. Consistency Table for Occupancy

Total Matched Cases		Independent sample: Occupied	Independent sample: Vacant	Total	Prop. Consistent
Census sample	Occupied	226,557	7,533	234,090	0.968
	Vacant	3,351	14,627	17,978	0.814
	Total	229,908	22,160	252,068	
	Prop. Consistent	0.985	0.66		0.957

Table 2. Consistency Table for Size of Structure

Total Matched Cases		Independent sample: Single	Small Multi	Large Multi	Total	Prop. Consistent
Census sample	Single	178,015	1,824	2,329	182,168	0.977
	Small Multi	5,950	24,245	888	31,083	0.78
	Large Multi	1,468	3,332	34,017	38,817	0.876
	Total	185,433	29,401	37,234	252,068	
	Prop. Consistent	0.96	0.825	0.914		0.937

Table 3. Consistency Table for Race Domain

Total Matched Cases		Independent Sample: AI on res	White/AI off res	Hisp.	Black	H/PI	Asian	Total	Prop. Consist.
Census Sample	AI on res	3,126	96	19	3	0	1	3,245	0.963
	White/ AI off res	91	158,152	2,312	1,182	93	744	162,574	0.973
	Hispanic	22	2,425	19,610	365	23	112	22,557	0.869
	Black	4	1,566	416	27,515	15	96	29,612	0.929
	Hawaiian / Pac Isl.	0	106	30	7	441	70	654	0.674
	Asian	0	944	186	89	93	6,603	7,915	0.834
	Total	3,243	163,289	22,573	29,161	665	7,626	226,557	
	Prop. Consist.	0.964	0.969	0.869	0.944	0.663	0.866		0.951