No Block Group Data For You

Unlike Seinfeld's "Soup Nazi" the loss of transportation Block Group data in the 5-year ACS had nothing to do with how we asked for it. Instead it has to do with how the Census Bureau's Disclosure Review Board is interpreting and applying Title 13 to 5-year ACS standard data products. This story is not about how disclosure works, the problems with it or why the interpretation might be flawed. Instead it is about the expected impact of one small rule made by the review board and released to the general public in the form of supplementary material to a Federal Register notice issued March 6, 2009 that was looking for comments on 5-year ACS standard products.

Register Notice
http://www.census.gov/acs/www/Downloads/FedRegNotice_E9-4803.pdf

Census Proposes not to release any Block Group data for "Means of Transportation to Work" when it is crossed with any other variable

Based on the best analysis possible, NO Block Group data will be released by the Census Bureau for "Means of Transportation to Work" when it is used in a crosstab. In fact, little to no Tract level data for "Means of Transportation to Work" data will ever make it past the Rule 7 filter. Not even when it is collapsed or crossed with "Sex" a small 2-category variable. Analysis done by the Census Bureau for FHWA confirms this.

What is Rule 7?

Rule 7 states:
> "For the residence and workplace tables where means of transportation (mode) is crossed with one or more other variables, there must be at least three unweighted workers in sample for each transportation mode in a given place for the table to be released. Otherwise the data must be collapsed or suppressed and complementary suppression must be applied. There is no threshold on univariate tables."
> http://www.census.gov/acs/www/Downloads/proposal_acs5yearproducts.pdf

What the Analysis Shows

In some recent work done for FHWA by the Census Bureau, it was shown that upwards of 73 percent of the Tract-level "Means of Transportation to Work" tables would not make it past the Rule 7 filter. When one considers that Tracts have a minimum size of 1,200 people and Block Groups average 600 people, it becomes easy to understand why no Block Group data will likely ever pass the filter. Of course this assumes that the user is interested in seven transportation modes which would include drive alone, carpool, public transit, bike or walk, other, work at home and a total. Yes, the total counts as mode.

The percent of Tracts suppressed does get worse. For example if there were eleven modes, approximately 95 percent of the Tracts will likely be suppressed. However, if the user can live with only five modes including the total, only 25 percent of the Tract-level tables will be suppressed. Unlucky for the data user, the Census Bureau is not proposing a five mode breakdown. Instead the smallest group of modes that can be found in the proposed "collapsed" tables is six which will have greater suppression than five or four or even three. When contemplating what the level of suppression might be, keep in mind that the analysis applies to Tracts which have roughly twice as many samples available as Block Groups.

The Data Used for the Analysis

The analysis was possible because of the 5-year ACS test data. It used data from five of the ACS test counties, Multnomah (OR), Lake (IL), Broward (FL), San Francisco (CA) and Bronx (NY). The counties all tend to be urban and have been used for a lot of the Census Bureau's ACS tests and analysis.

In doing the suppression analysis, the data was examined using the same group of modes that were being considered for a special tabulation of 3-year ACS data under the Census Transportation Planning Products (CTPP) program. This distribution of modes was used because the CTPP designers at the time were investigating what the effect of the rule would be if applied to 5-year data under the CTPP program. Prior to the March 6th Federal Register there was no mention that disclosure rules would ever be applied to standard products so there was no reason to analyze a different group of modes that would be specific to the standard products. Besides, the differences are so subtle that in all likelihood they would have a minimal affect.

Presented below are three tables that lie at the nucleus of the analysis. The first two tables show the modal groupings for the CTPP and 5-year ACS followed by the analysis results.

Table 1 shows the list of modes used for the 3-year CTPP followed by Table 2 that has the list and grouping of modes that are proposed for the 5-year ACS standard products. As can be seen, the differences between the two are ever so slight. For example, in one instance "Walk and Bike" are grouped together while in another "Walk" is alone and Bike is grouped into "Other".

Besides the minor differences in which modes are collapsed into each other, there are small differences in the number of elements in each group between the two tables. For the CTPP the grouping represent 5, 7 and 11 modes while the 5-year ACS proposes groupings of 6, 7, and 14 modes. (For a listing of all the proposed tables for the 5-year ACS before any suppression tests are applied go to http://www.census.gov/acs/www/Products/users_guide)

Following Tables 1 and 2 is Table 3 which shows the results of processing the actual ACS test data, tabulating it according to the modal groupings in Table 1 and checking

what Tracts would pass the Rule 7 filter. The work was done by Census Bureau staff and what is in Table 3 is what was approved for release by the DRB.

Table 1: CTPP Modal Groupings

Five	Seven	Eleven
Total	Total	Total
Drove Alone	Drove alone	Drove alone
Carpool	Carpool	2-person Carpool
Public Transit	Public Transit	3-person Carpool
All Other	Bike or Walk	4^+ person Carpool
	Other	Bus
	Work at Home	Streetcar/subway
		Railroad
		Bike or Walk
		Other
		Work at Home

Table 2: Modal Groupings for 5-year ACS Standard Products

Six (Collapsed)	Seven (Basic)	Fourteen
Total	Total	Total
Drove Alone	Drove Alone	Drove Alone
Carpooled	Carpooled	2-person Carpool
Public Transit	Public Transit	3-person Carpool
Other	Walk	4^+ person Carpool
Work at Home	Other	Bus
	Work at Home	Streetcar/Trolley
		Subway/Elevated
		Railroad
		Ferryboat
		Bike
		Walk
		Other
		Work at Home

Table 3: Tract Level Tables Suppressed Due to Rule 7 for ACS Test Counties

Modal Categories	Multnomah[1]	Lake[2]	Broward[1]	San Francisco[1]	Bronx[3]	Weighted Average
Five	5.3%	19.3%	37.3%	9.7%	33.0%	24.3%
Seven	65.3%	80.7%	84.6%	47.7%	80.3%	73.9%
Eleven	94.1%	96.7%	97.1%	92.1%	95.8%	95.4%
Total Number of Tracts	170	150	279	176	330	1105

notes:

1-All tracts had 3 or more observations.

2-One tract had fewer than 3 observations and is not included.

3-25 tracts had fewer than 3 observations and are not included.

Findings

After reviewing Table 3 it should become ever so clear that the probability of ever seeing "Means of Transportation to Work" in a crosstab from ACS at the Block Group data is virtually zero. Simply the size and number of records in the ACS sample at 5-year intervals makes it impossible to pass Rule 7. One can speculate but in order to make sure there are some Block Group cross tabs for "Means of Transportation to Work" there would likely have to be only 2 to 3 groupings--Traveled in "Car, Truck or Van", Other and total.

Putting Block Groups aside, the same findings hold true for Tracts. While a small number of Tracts would pass the filtering for the smallest grouping, six modes, it is very probable that the percent of Tracts suppressed would be relatively close to the number suppressed with seven categories. The reason for this is the inclusion of "Work at Home" in the 5-year six-category tables. Although the data was not released there was some verbal confirmation in discussion that were held regarding the analysis that "Work at Home" is very sensitive to the Rule 7 threshold of 3. For some reason "Work at Home" has a hard time passing the test.

Although this all sounds very dismal and odd, after all why would anyone pay to collect all this data and then just not release it? There is one table, Table B08301 that will be released for all geographies without any restrictions known to date (March 20, 2009). It is a Univariate table that contains all 17 modes, 18 with the total. However, this table will only be available for residents since the "new" place of work tables that the Census Bureau has been promoting will be subjected to even more restrictive rules.

Regarding the new "Means of Transportation to Work for Workers at Work" tables there will only be Place level and above (County, State…) data available. No Tract or Block Group level data will ever be released. In addition to the geographical restriction, the workplace tables will not only be subjected to Rule 7 but also to Rule 6. Rule 6 states that in addition to having 3 unweighted records for each modal category the total for the table must contain at least 50 unweighted or 300 weighted records. Exactly how many tables might pass Rule 7 but not Rule 6 is anyone's guess. While some may be concerned about the level of geography available for the work place tables, in all fairness it must be noted that the workplace tables can not really be tabulated below the "Place" level because of inferior geocoding and imputation techniques used by the Bureau. But that is a whole different issue.