

2b

2b

|  |  |  |  |
| --- | --- | --- | --- |
| TO: | File | DATE: | July 16, 2020 |
| FR: | Shimon Israel and Elizabeth Theocharides |  | W.I. 1122 |
| RE: | 2015 TM 1.5 TAZ 1454 Land Use Documentation | | |
|  | This memorandum documents the process for creating the Travel Model 1.5 TAZ Land Use data for the 2015 base year. Data are collected, aggregated, and manipulated via the R script “ACS 2013-2017 create TAZ data for 2015.R.” Where data are available from the Decennial Census and the American Community Survey (ACS), they are compiled from Census 2010, ACS 2013-2017 standard tabulations, and the ACS 2013-2017 Public Use Microdata Sample (PUMS). There is a mix of cached and API-accessed Census/ACS data. In other cases, other (non-census-derived) variables were carried forward from previous Plan Bay Area 2040 land use datasets, and other variables were derived from yet other datasets, as is documented in the “TAZ1454 2015 Land Use.xlsx” document in the main Petrale GitHub folder. Comprehensive documentation for this dataset comes from this memo, the “Data Sources” tab in “[TAZ1454 2015 Land Use.xlsx](https://github.com/BayAreaMetro/petrale/blob/master/applications/travel_model_lu_inputs/2015/TAZ1454%202015%20Land%20Use.xlsx)”, and the R script “[ACS 2013-2017 create TAZ data for 2015.R](https://github.com/BayAreaMetro/petrale/blob/master/applications/travel_model_lu_inputs/2015/ACS%202013-2017%20create%20TAZ%20data%20for%202015.R).” Additionally, links to other scripts that produced data for this work, most notably PUMS scripts, are provided within this document. General Approach for Census- and ACS-Derived Variables For census/ACS-derived variables, block- and block group-level data were used when available (i.e., not suppressed by the Census Bureau) and tract-level data otherwise. A Census block/MTC TAZ 1454 equivalency file was created via GIS for use in translating Census/ACS block, block group, and tract data to MTC’s 1454 TAZ system. The block share of a block group or tract was determined by the block’s population share from 2010. This share was then applied to decennial and ACS variables to build up TAZ variables. The R script referenced above demonstrates how this is done.  After the above step, resulting Census- and ACS-derived data are rounded, and then small adjustments are made such that categorical subtotals sum to match marginal totals. For example, households by income should precisely sum across income categories to match the total households variable total. This was the approach taken in the below variables:   * TOTHH, hhlds – total households * HHPOP – population living in households * EMPRES – employed residents * SFDU – single-family dwelling units * MFDU – multi-family dwelling units * HHINCQ1-HHINCQ4 – households by household income * SHPOP620, AGE0004, AGE0519, AGE2044, AGE4564, AGE65P – persons by age  School Enrollment School enrollment data includes high school, part-time college, and full-time college enrollment. Separate processes were used to gather high school and college enrollment data:   1. Public and private high school location and enrollment data are provided by the California Department of Education. The school address information was used to geocode the school locations, and school enrollment information was then summarized by TAZ 1454. 2. College enrollment data were obtained from various university websites. Part-time and full-time enrollment numbers were largely carried over from 2010 data. Where updated part-time and full-time enrollment data were found, they replaced the 2010 data. Where updated aggregate enrollment data were found for a school, but the part-/full-time distribution was lacking, the 2010 part-time and full-time proportions were applied to 2015 totals to derive updated cell values.  Parking Cost Parking cost data are taken from Plan Bay Area 2040, the 2015 land use data. Group Quarters Population Group quarters population disaggregated by institutional/non-institutional type is not provided for small-area ACS data. In order to overcome this data limitation, the following steps were undertaken to develop TAZ-level non-institutional group quarters population:   1. Non-institutional group quarters (university, military, and other) were summed from the 2010 decennial data and partitioned to TAZs, as is described above for other variables. 2. MTC staff [researched college dorm (non-institutional) growth in three TAZs](https://github.com/BayAreaMetro/petrale/blob/master/applications/travel_model_lu_inputs/2015/Group%20Quarters/gq_add_00051015.csv) – 353, 354, and 1008. The GQ growth was added to the 2010 GQ values for these TAZs. 3. Non-institutional GQ population for 2015 was determined at the county level using PUMS 2013-2017 data. County-level correction factors were created by comparing the adjusted 2010 numbers (determined in steps 1 and 2) with the 2015 totals. 4. County-level scaling correction factors were then applied at the TAZ level to produce the final TAZ GQ values, such that their sum equals the “2015” (ACS PUMS 2013-2017) regional total.  Employment The total employment (jobs, not employed residents) is taken from the REMI regional model, and corrected for net incommuters. The travel model assumes that all jobs are held by workers within the region, and that all Bay Area workers work within the region. As such, workers living outside the region and commuting in (incommuters) were subtracted from the jobs total, while workers inside the region commuting out (outcommuters) were added to the jobs total. The steps for calculating net incommute is described below (and is fully detailed in [ACS 2013-2017 Incommute by Industry.xlsx](https://github.com/BayAreaMetro/petrale/blob/master/applications/travel_model_lu_inputs/2015/Employment/Incommute/ACS%202013-2017%20Incommute%20by%20Industry.xlsx))::   1. The Bay Area TAZ 1454 employment total (jobs) for 2015 is a controlled value of 4,005,318, and comes from the REMI regional model.[[1]](#footnote-1) TAZ-level distribution comes from [2015 ESRI Business Locations Data](https://github.com/BayAreaMetro/petrale/blob/master/applications/travel_model_lu_inputs/2015/Employment/Summarize%20ESRI%20with%20NAICS2%20and%20ABAG6%20Equivalency%202015.R), and the employment total for each TAZ is scaled up such that the regional total matches the REMI value. 2. The total net incommute is determined by using ACS data. [Workers who both live and work within the Bay Area](https://github.com/BayAreaMetro/PUMS-Data/blob/master/Analysis/ACS%20PUMS%202013-2017/Worker%20Research/ACS%202013-2017%20intra_regional%20commute%20by%20sector.R) (from PUMS 2013-2017) are subtracted from total workers at work (ACS 2013-2017 Table B08526). 3. [Outcommute workers](https://github.com/BayAreaMetro/PUMS-Data/blob/master/Analysis/ACS%20PUMS%202013-2017/Worker%20Research/ACS%202013-2017%20outcommute.R) are added to the total from Step 2, above, to calculate net incommuters (total incommuters - outcommuters). Initially the resulting data were segmented by industry, but in the end only aggregate values were used. 4. Net incommuters were subtracted from TAZs around the region using a [distribution based on CTPP 2012-2016 place-level data](https://github.com/BayAreaMetro/petrale/blob/master/applications/travel_model_lu_inputs/2015/Employment/Incommute/2012-2016%20CTPP%20Places%20to%20Superdistrict%20Equivalency.xlsx) (aggregated to the 34 MTC superdistricts).  Residential Workers The number of total employed residents is derived from ACS 2013-2017, Table B23025. The initial distribution of households by household workers comes from Table B08202 for the same period. This household distribution, however, is skewed (relative to PUMS data) toward zero-worker households because its universe only includes workers at work during the ACS survey reference week. That is, workers with a job but not at work (e.g., employees who are ill, on vacation, at personal appointments, etc.) are not included in Table B08202. In addition, the PUMS household weights appear to undercount workers in larger households (with more 3-plus workers), and the PUMS person weights appear more accurate. The approach used for reconciling households by number of workers relies on PUMS person weights to correct for worker undercounts.  [PUMS data from 2013-2017 PUMS](https://github.com/BayAreaMetro/PUMS-Data/blob/master/Analysis/ACS%20PUMS%202013-2017/Worker%20Research/ACS%202013-2017%20PUMS%20HH%20and%20Person%20Worker%20Research.R) were used to develop correction factors for TAZ-level households by number of workers. The process is described in the below steps (with a supporting data summary in Table 1, below, and in [ACSPUMS\_WorkerTotals\_2013-2017\_Comparisons.xlsx](https://github.com/BayAreaMetro/petrale/blob/master/applications/travel_model_lu_inputs/2015/Workers/ACSPUMS_WorkerTotals_2013-2017_Comparisons.xlsx)):   1. The number of households by workers and implied workers in each household category are derived from ACS 2013-2017 Table B08202 (Column 1). The average number of workers for the 3-plus worker category was calculated from the 2013-2017 PUMS as approximately 3.43 (3.430503 was applied). 2. Total workers within each household category (from the person weights in the 2013-2017 PUMS), along with implied households (1-worker HH/ 1, 2-worker HH/ 2, and 3-plus worker / 3.43) are listed in Column 2. 0-worker households are calculated as the difference of total households minus the sum of the 1-, 2-, and 3-plus-worker households. 3. In Column 3, the ACS 2013-2017 aggregate household county totals come from ACS 2013-2017 Table B08202, and the aggregate worker county totals for the same period come from ACS 2013-2017 Table B23025. Next the relative distributions for households by number of workers within each county need to be reconciled. In order to do this, worker distribution shares from the PUMS 2013-2017 within each county are used. Category values are then inflated so that the inflated category values will sum to ACS table county totals; the inflation factor is Column 3 Total Workers / Column 2 Total Workers, respective to each county (e.g., 827,795 / 826,597 for Alameda County). The implied household distribution is derived from the worker counts as above, in Step 2. The formulas for this can be better understood by looking at [ACSPUMS\_WorkerTotals\_2013-2017\_Comparisons.xlsx](https://github.com/BayAreaMetro/petrale/blob/master/applications/travel_model_lu_inputs/2015/Workers/ACSPUMS_WorkerTotals_2013-2017_Comparisons.xlsx). 4. The final step creates factors for correcting ACS 2013-2017 Table B08202 values at the TAZ level (Column 4). These factors are the quotient of Column 3 Implied Households divided by Column 1 Total Households. The resulting factors were applied in the TAZ-building script ([ACS 2013-2017 create TAZ data for 2015.R](https://github.com/BayAreaMetro/petrale/blob/master/applications/travel_model_lu_inputs/2015/ACS%202013-2017%20create%20TAZ%20data%20for%202015.R)), respective to each county and workers per household category. 5. It is noteworthy that group quarters workers (e.g., working college students in dorms) are included in the employed residents data, ACS Table B23025, and in the PUMS person weights. They are, however, excluded from households by number of workers, Table B08202, and from the household weights in the PUMS data (group quarters records are actually given a 0-weight placeholder value in the household file). The approach outlined here uses the full worker universe, those in both households and group quarters, and includes group quarters workers in the final outputted dataset – both in total employed residents and in households by number of workers. The distribution within the households by number of workers accounts for both household and group quarters workers, such that the total implied workers sums to the regional employment total. | | |
|  |  | | |
|  |  | | |

**Table 1: Households by Number of Workers**

|  |  | 1 | | 2 | | 3 | | 4 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | ACS 2013-2017 Table B08202 - HH Size by Number of Workers in HH | | PUMS 2013-2017 Workers (ESR==1,2,4,5), Using Person Weights | | ACS 2013-2017 Refactored with County Total Households (Table B08202) and Workers (Table B23025) | | County-Level Correction Factors for TAZ Application |
| County | HH Workers | Total Households | Implied Workers | Implied Households | Total Workers | Implied Households | Total Workers | Households |
| Alameda | 0\_workers | 140,827 | 0 | 86,734 | 0 | 86,036 | 0 | 0.611 |
|  | 1\_worker | 213,190 | 213,190 | 226,429 | 226,429 | 226,757 | 226,757 | 1.064 |
|  | 2\_workers | 168,962 | 337,924 | 191,141 | 388,282 | 194,422 | 388,845 | 1.151 |
|  | 3p\_workers | 46,091 | 157,030 | 61,765 | 211,886 | 61,855 | 212,193 | 1.342 |
| Alameda Total | | 569,070 | 708,144 | 569,069 | 826,597 | 569,070 | 827,795 | 1.000 |
| Contra Costa | 0\_workers | 92,670 | 0 | 75,568 | 0 | 75,463 | 0 | 0.814 |
|  | 1\_worker | 145,079 | 145,079 | 148,046 | 148,046 | 148,095 | 148,095 | 1.021 |
|  | 2\_workers | 117,893 | 235,786 | 126,824 | 253,648 | 126,866 | 253,732 | 1.076 |
|  | 3p\_workers | 33,955 | 115,683 | 39,160 | 134,338 | 39,173 | 134,382 | 1.154 |
| Contra Costa Total | | 389,597 | 496,548 | 389,598 | 536,032 | 389,597 | 536,209 | 1.000 |
| Marin | 0\_workers | 27,347 | 0 | 23,050 | 0 | 22,412 | 0 | 0.820 |
|  | 1\_worker | 41,727 | 41,727 | 43,335 | 43,335 | 43,673 | 43,673 | 1.047 |
|  | 2\_workers | 30,126 | 60,252 | 31,836 | 63,671 | 32,084 | 64,168 | 1.065 |
|  | 3p\_workers | 5,646 | 19,236 | 6,625 | 22,728 | 6,677 | 22,905 | 1.183 |
| Marin Total |  | 104,846 | 121,215 | 104,846 | 129,734 | 104,846 | 130,747 | 1.000 |
| Napa | 0\_workers | 12,572 | 0 | 8,829 | 0 | 8,929 | 0 | 0.710 |
|  | 1\_worker | 16,470 | 16,470 | 17,614 | 17,614 | 17,570 | 17,570 | 1.067 |
|  | 2\_workers | 15,313 | 30,626 | 16,654 | 33,307 | 16,612 | 33,224 | 1.085 |
|  | 3p\_workers | 4,689 | 15,975 | 5,948 | 20,403 | 5,933 | 20,352 | 1.265 |
| Napa Total |  | 49,044 | 63,071 | 49,044 | 71,324 | 49,044 | 71,146 | 1.000 |
| San Francisco | 0\_workers | 82,119 | 0 | 59,260 | 0 | 59,486 | 0 | 0.724 |
|  | 1\_worker | 139,605 | 139,605 | 147,392 | 147,392 | 147,281 | 147,281 | 1.055 |
|  | 2\_workers | 106,022 | 212,044 | 114,314 | 228,628 | 114,228 | 228,456 | 1.077 |
|  | 3p\_workers | 31,026 | 105,704 | 37,805 | 129,689 | 37,776 | 129,592 | 1.218 |
| San Francisco Total | | 358,772 | 457,353 | 358,771 | 505,709 | 358,772 | 505,329 | 1.000 |
| San Mateo | 0\_workers | 50,750 | 0 | 35,819 | 0 | 36,056 | 0 | 0.710 |
|  | 1\_worker | 95,393 | 95,393 | 97,044 | 97,044 | 96,944 | 96,944 | 1.016 |
|  | 2\_workers | 86,422 | 172,844 | 93,489 | 186,978 | 93,392 | 186,785 | 1.081 |
|  | 3p\_workers | 29,231 | 99,589 | 35,441 | 121,580 | 35,404 | 121,454 | 1.211 |
| San Mateo Total | | 261,796 | 367,826 | 261,793 | 405,602 | 261,796 | 405,183 | 1.000 |
| Santa Clara | 0\_workers | 118,875 | 0 | 77,162 | 0 | 77,547 | 0 | 0.652 |
|  | 1\_worker | 240,595 | 240,595 | 252,019 | 252,019 | 251,842 | 251,842 | 1.047 |
|  | 2\_workers | 204,894 | 409,788 | 221,596 | 443,191 | 221,440 | 442,880 | 1.081 |
|  | 3p\_workers | 66,087 | 225,155 | 79,677 | 273,332 | 79,621 | 273,140 | 1.205 |
| Santa Clara Total | | 630,451 | 875,538 | 630,453 | 968,542 | 630,451 | 967,863 | 1.000 |
| Solano | 0\_workers | 37,665 | 0 | 30,350 | 0 | 29,839 | 0 | 0.792 |
|  | 1\_worker | 54,266 | 54,266 | 55,914 | 55,914 | 56,158 | 56,158 | 1.035 |
|  | 2\_workers | 41,863 | 83,726 | 45,984 | 91,968 | 46,185 | 92,370 | 1.103 |
|  | 3p\_workers | 13,558 | 46,191 | 15,104 | 51,814 | 15,170 | 52,041 | 1.119 |
| Solano Total |  | 147,352 | 184,183 | 147,352 | 199,696 | 147,352 | 200,569 | 1.000 |
| Sonoma | 0\_workers | 51,489 | 0 | 41,560 | 0 | 42,039 | 0 | 0.816 |
|  | 1\_worker | 69,025 | 69,025 | 72,470 | 72,470 | 72,236 | 72,236 | 1.047 |
|  | 2\_workers | 53,747 | 107,494 | 57,649 | 115,297 | 57,463 | 114,925 | 1.069 |
|  | 3p\_workers | 15,797 | 53,820 | 18,380 | 63,052 | 18,321 | 62,849 | 1.160 |
| Sonoma Total | | 190,058 | 230,339 | 190,058 | 250,819 | 190,058 | 250,010 | 1.000 |
| Bay Area | 0\_workers | 614,314 | 0 | 438,332 | 0 | 437,807 | 0 |  |
|  | 1\_worker | 1,015,350 | 1,015,350 | 1,060,263 | 1,060,263 | 1,060,557 | 1,060,557 |  |
|  | 2\_workers | 825,242 | 1,650,484 | 902,485 | 1,804,970 | 902,693 | 1,805,385 |  |
|  | 3p\_workers | 246,080 | 838,382 | 299,904 | 1,028,822 | 299,929 | 1,028,908 |  |
| Bay Area Total | | 2,700,986 | 3,504,216 | 2,700,984 | 3,894,055 | 2,700,986 | 3,894,851 |  |

SI  
X:\petrale\applications\travel\_model\_lu\_inputs\2015\Documentation\TM 1.5 TAZ 1454 Land Use Documentation Memo.docx

1. https://mtcdrive.app.box.com/file/654134152628 [↑](#footnote-ref-1)