Decisions

* Don’t eliminate principal from mortgage so that only mortgage interest and charges remain (this code exists in the do file but is commented out at the moment)
* Don’t drop people with missing values for recollected/expected retirement spending
* No longer drop if waves not between 5 and 8
* Keep if age between 50 and 70
* Winsorize by wave (excluding variables with nonmissing values below 1000)
* An individual’s social security income is estimated as a person's highest social security across all waves in order to protect from early claiming at a discount and 0 values
  + To impute missing values, an individual was assigned into tertiles based on the head's years of education (cutoffs at 0-11; 12-15; 16-)

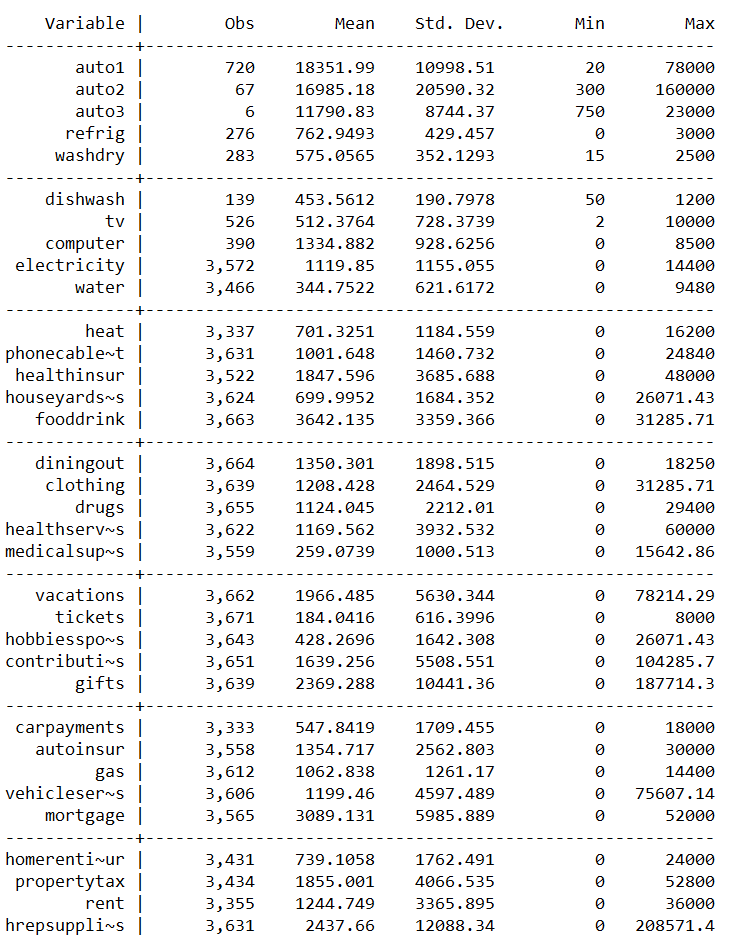
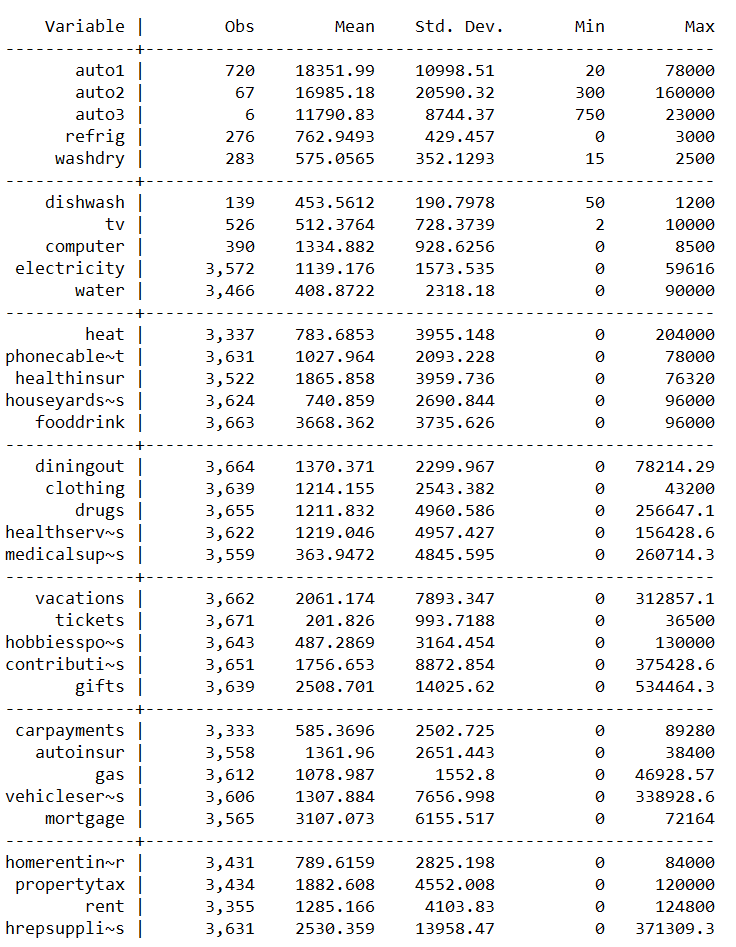
Relevant Links

* Target Paper (Heterogeneity in spending change at retirement; Hurd and Rohwedder): <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3919678>
* HRS and CAMS Codebooks: <https://hrs.isr.umich.edu/documentation/codebooks>
* HRS Question Concordance: <https://hrs.isr.umich.edu/documentation/question-concordance>
  + Text search of all HRS datasets from 1992 to 2014 (select core as filter as well as relevant years)
  + Under cross wave equivalents, click View Xref to pull up a list of the relevant variable’s names across all waves the variable appears in (variable names are often inconsistent across waves)
* HRS, CAMS, and RAND Data Download: <https://ssl.isr.umich.edu/hrs/start.php>

(Important notes on next page)

Notes

* Relevant documentation are described whenever appropriate in the do file as the code goes along
  + The first block of comments in “Prepare Data, Mean Comparison” describes how to download the data sets used in any subsequent code
* The below two drops statements makes food means decrease instead of increase
  + drop if recollect == . & retired == 1
  + drop if expected == . & retired == 1
* Eliminating principal from CAMS by RAND (and thus potentially by Hurd and Rohwedder) is done by age range, and a different ratio (mortgage interest and charges / (mortgage principal paid on owned property + mortgage interest and charges)) is used for 55-64 and >65 which might drive some of the fall in spending at retirement for total/nondurables
* The CAMS spending data is winsorized for each wave (auto price 3 had 6 values in wave 1, so only winsorize if more than 1000 nonmissing values, which basically excludes durables)
  + Before winsorizing, some of the variables have large maxes like 371,309 for house repair/supplies (despite a mean of 2,530), so winsorizing by wave seems reasonable
  + Below, summary tables show Wave 1 before winsorization on left and Wave 1 after winsorization on right.
  + Assessing the sensitivity of selected key results (or unusual/surprising results) to the window could be important to follow up on



* Table 2 is made using the texsave package which saves the data set. The table is created by manually placing the collapse values into the desired format, so one should be careful when adding new rows or removing old ones. Given the current format, the process is automated.
* Financial planning horizon is not reported for waves 2, 3, 9, or 10