California Tobacco Synthetic Control – Variable Key

**state** – this variable contains the state identifier. Although input as a number into the synth program, when browsing the data (command: “browse” or “Ctrl + 8”), the program will display blue text that contains the state’s name. The blue text indicates that the variable is “labeled data”.

**year** – this represents the year of the observation. This is an important variable which must be identified as the time variable to Stata. We manually code it as the time variable in our panel data by using the tsset command.

**cigsale** – the dependent variable of our analysis, this variables describes the number of cigarette packs sold per person in each given state-year.

**lnincome** – is the log state per capita GDP. This is used as an indicator variable in our analysis. We want the synthetic to resemble the true California in other dimensions than just cigarette sales, hence why we estimate the synthetic weights by matching on this information as well.

**beer** – beer consumption per capita for each state-year observation. Again, this is an indicator variable in the analysis.

**age15to24** – another indicator variable, this tells what percentage of the state’s population falls between the ages of 15 and 24.

**retprice** – the final indicator, this variable lists the average retail price of cigarettes for each state-year.