

Readme file for Stata codes and data files

Gasoline Taxes and Consumer Behavior, AEJ: Economic Policy
Shanjun Li, Joshua Linn, and Erich Muehlegger

1. Aggregate Analysis – Annual Data

The Stata do file for the annual analysis graphed in Figures 1-6 and presented in Tables 1-4 and 6 is `annual_analysis_replication.do`.

There are two data files: (1) `annual_regression_data.dta`, and (2) `Us-Coordinates.dta`.

`Annual_regression_data.dta`: Quantities and prices are from publicly available data from the Energy Information Administration. Tax, travel and highway infrastructure data was collected from annual issues of Highway Statistics Annual published by the Federal Highway Administration. Demographic data is from publicly available sources from Census Bureau and the Bureau of Economic Analysis. Political variables were collected from publicly available information.

`Us-Coordinates.dta` contains geolocational shape files outlining state borders and used for the maps in Figures 2 and 3.

2. Aggregate Analysis – Monthly Data

The Stata do file for the monthly regressions of gasoline quantities on gasoline prices (Table 5) is `monthly_regressions_replication.do`.

The data file for the monthly regressions is `monthly_regression_data.dta`.

Quantities and prices for the regressions are from publicly available data from the Energy Information Administration. Tax data was collected from annual issues of Highway Statistics Annual published by the Federal Highway Administration.

3. Household Analysis

Stata do files for household-level MPG and VMT analysis are: `mpg_regression_AEJ.do`, and `vmt_regression_AEM.do`, respectively.

There are two data files for the MPG analysis: `mpg_forall.dta`, and `mpg_fornew.dta`, corresponding to the two samples used in Table 11.

There are two data files for the VMT analysis: vmt_odometer.dta, and vmt_selfreport.dta, corresponding to the two samples used in Table 12.

The data sources are from National Household Travel Survey (1995, 2001 and 2009 waves) combined with the EPA fuel economy database for fuel economy data of different vehicle models.

4. Media Salience Analysis

Stata do file for media salience analysis: media_replication.do

There are two data files for the media salience analysis: newspaper_data.dta and tvnews_data.dta, corresponding to the analyses of print media coverage and tv nightly news coverage.

Data on print newspaper coverage is from Lexis Nexis. Word counts are aggregate monthly counts of the number of words from articles satisfying two criteria: (1) published in the LexisNexis category “Major U.S. Newspapers” between 1985 and 2008, excluding the “Journal of Commerce” and “The Christian Science Monitor, AND (b) indexed by LexisNexis as being “highly relevant” (i.e., receiving a relevance score greater than 85%) to either Gasoline Prices, or Fuel Taxes.

Data on nightly news coverage was collected from the Vanderbilt Television News Archive and includes all 3,926 nightly news segments on ABC, NBC, CBS, FOX and CNN between 1983 and 2010 in which the word “gasoline” appears in the segment abstract. Tvnews_data.dta aggregates the raw segment level data up to the month level and reports the total number of seconds of nightly news coverage each month.

Data on gasoline prices are based on publicly reported tax-inclusive, monthly average retail gasoline prices from the Energy Information Administration.