Readme File for Replication Package for

Vehicle and Fuel Taxation for Transport Demand Management: Learnings from the Literature through a Development Lens

He He, Chaeyoung Kim

Data source

The data used in this paper is sourced from numerous literatures which are publicly available. The details are as follows:

- **Weighted average elasticities**: Extracted from 40 literatures published between 1992 and 2022, encompassing data from the years 1929 to 2019.
- **Externality costs**: Extracted from 17 literatures published between 1995 and 2019, with data from 1989 to 2017.
- **The exchange rates** employed for calculating externality costs are obtained from the following sources:
    - Euro, yen, pound (Year 1999 to 2022): https://www.ofx.com/en-us/forex-news/historical-exchange-rates/eur/usd/
    - Euro (ECU) (Year 1989): https://aei.pitt.edu/79650/1/1990.4.pdf
    - Euro (Year 1998): The average of monthly exchange rates from January 1998 to December 1998, https://www.insee.fr/en/statistiques/serie/010002053

The complete list of papers, including author, publication years, and corresponding values, is available in Excel and the Appendix of the paper.

Weighted Average Elasticity

Weighted average elasticity_reproducibility.xlsx comprises three sheets: All, Table, and Calculation. The **"All" sheet** contains a comprehensive list of values gathered from various sources, with the variables elasticity and standard error representing data from the literature. The calculated value, weighted elasticity, is derived using these elasticities and standard errors. The ultimate goal is to obtain the inverse-variance weighted average elasticities. To do so, the initial step involves computing each weighted elasticity.

To compute weighted elasticities, certain values are excluded before the calculation. The exclusions are as follows:

1. Data lacking standard errors
2. Values from meta-analysis studies (to prevent double-counting and weighting distortion)
3. Values from the average of previous studies (for the same reasons mentioned above)

4. Granular data values (e.g., peak, off-peak, day, night) (to avoid redundancy, as total overall values are already counted in the dataset)

The weighted elasticity is calculated by multiplying the elasticity value by the inverse of the squared standard error.

The **"Calculation" sheet** illustrates the derivation of the weighted average elasticity and weighted average standard error based on tax instruments and independent variables. Three types of tax instruments (purchase, ownership, and use) and four independent variables (new vehicles, fuel economy, fuel demand, and traffic volume) result in 12 combinations, but cases with no data lead to 9 results.

The weighted average elasticity is computed as the sum of weighted elasticity divided by the sum of inverse of the squared standard error**.** *The standard error for the weighted average* is the inverse of the square root of the sum of the inverses of the squared standard errors. For additional details on formula, please refer to Appendix B of the paper.

The **"Table" sheet** displays the final results of the inverse-variance weighted average elasticities for new vehicles, fuel economy, fuel demand, and traffic volume with respect to purchase, ownership, and use taxes, with standard errors in parentheses. The 'total number of studies' table indicates the count of data from the studies used in these calculations.

Externality Costs

Externality costs_reproducibility.xlsx consists of seven sheets: Full data, exchange rate, data for calcs, per 1000vkm, sheet1, dollar per 1000vkm, and table 1. These sheets represent the step-by-step process of calculating average externality costs.

The **"Full data" sheet** provides a comprehensive list of values with geographical and temporal scopes obtained from various sources. It includes five external cost variables (air pollution, GHG emissions, congestion, crashes, and noise) and three scopes (spatial, temporal, and vehicle fuel types). Each data point has a different unit, which will be standardized in a later step.

The **"exchange rate" sheet** contains the exchange rates of the euro, yen, and pound, used to convert different currencies into the same unit, the US dollar.

The **"data for calcs" sheet** includes only data used for calculations, with values presented in currency per vehicle-kilometers (vkm). Data with units other than vkm are excluded. The 'year' column represents the data year, and the 'currency unit' column indicates the unit of currency. If the data year is specified in the source, that year's exchange rate is used; otherwise, the exchange rate is based on the publication year of the paper.

The **"per 1000vkm" sheet** converts values into currency per 1000 vkm. Values are simply multiplied by 1000 from the values from the "data for calcs" sheet, except for values already in the unit of currency per 1000vkm. The 'currency rate' column corresponds to the data year and is derived from the "exchange rate" sheet.

The **“sheet1” sheet** contains values converted into the dollars per 1000vkm using the exchange rate. To get this, values of external cost in dollars per 1000vkm are multiplied by the appropriate currency rate.

The **“dollar per 1000vkm” sheet** presents the final results as numerical values, intended for use in the calculations on the next step, specifically in the “table1” sheet.

The **“table1” sheet** includes a table with externalities (congestion, air pollution, GHG emissions, crashes, noise) in columns and scope (spatial and temporal) in rows. Each scope is further categorized: spatial scope into urban, suburban, metropolitan, and rural, and temporal scope into peak time and off-peak time. Each value represents the average external costs in dollars per 1000vkm. Values not matching the specified scope are excluded from the spatial and temporal breakdowns.

A box and whisker chart visually depicts the data distribution into quartiles, showing maximum, 75th percentile, median (50th percentile), mean, 25th percentile, and minimum statistics. The chart is created by calculating the percentile of externality cost variables, as shown in the second table, using the PERCENTILE.INC function. Please refer to the Excel file for detailed functions and calculation formulas used.