# The factors that determine retail fuel prices

Econometrics Empirical Groupe Assignment

Created by: Levente Attila Kiss, Marco Romanelli, Milán Szabolcs Sztojka

Introduction:

In this assignment our task was to prepare an essay for submission within the framework of the Econometrics subject, in which we identify the factors that determine retail fuel prices at the station level on data containing fuel prices for all Hungarian stations for a randomly selected week. The assignment was written and presented in a group of three. During the research, we first identify the relevant factors that influence fuel prices, then analyse the data and variables describing these factors, which, for the sake of easier utilization, are put trough on several data cleaning processes and transformations. To make it easier to create a model that best describes the problem, we tried different models and select based on several critical factors. Finally, we presented the main results of the submission in the conclusion.

At the beginning of our research, we first examined the data for the week assigned to us (Dataset of 2024.04.11-04.17) to define the main factors behind gas and diesel prices all around Hungary. We identified two main components, the company and the location of the fuel station. The first one contains more than 100 fuel station owner companies, from multinational big firms like Mol or Shell, to small family businesses with only one station. Therefore, we assume that this huge different could lead to difference in prices. The second factor, the location is more complicated. We think this is the main component behind the fuel prices. On one hand because it is important what features a fuel station have, like based on our own experience a station on a highway or in Budapest is more expensive. On the other hand, the location of the city also contains important economical factors like wage, what can impact prices, or car density, what can impact demand in the region.

Data/Data clearing:

The attached dataset contains cross-sectional data, therefore in the external data collection we focused on searching the closest time period to the original data, which is about fuel prices in the period of 11.-17. of April 2025. As additional data we obtained three datasets from the Hungarian Central Statistical Office. First data is about passenger cars per thousand capita, so the car density, for every Hungarian county from 2024. It is from the TIMEA app. Than we used the settlement data to connect city names and counties. At last, we import data about average gross earnings in Hungarian counties from 2024, it is measured in HUF.

The original dataset contains variables about the name of the company, the city and the diesel and gas prices of the day in HUF/liter. As it is asked, we prepare the weekly average price per station, to minimised potential data problems, in both kind of fuels. As data cleaning we rename variables and correct misspelled city names to be able to attach the datasets. After filtering for not available data the final dataset contains 1207 observations. For the used variables we did some variable transformations. From the company variable we created categories as a factor variable. After analysing the distribution of the fuel station owner firms we created five categories, four for the biggest companies with more than 100 fuel stations, they are Mol, Shell, Orlen and OMV. The fifth category contains every other smaller firms, as control category. Finally, from the address variable we created two dummy variables, as important features of the station, one if it is on a highway and one if it is in the capital city.

Modelling:

Estimate different regression models and discuss the parameter estimates and the explanatory power of the model. Carefully interpret the results and draw conclusions about which factors are important in determining fuel prices and which are less important. Which is the best model? model 3 for both?

Conclusion:

Differences observed between gasoline and diesel, main results, summary of what we did

GitHub repository:

Link: <https://github.com/Lecsak014/Econ_empirical_assignment.git>

Name: Econ\_empirical\_assignment (Public) created by Lecsak014 (Levente Kiss)