**Truck Electrification and Effective Fuel Cost Analysis**

Scott Andermann, Sam Avis, Emily Yeacker

DBSA 6201

1. **Introduction**

In the past decade, the shift to electric powered vehicles has accelerated drastically. Advancements in battery and motor technology have reduced costs and increased efficiencies to the point where electric vehicles are nearly on par with their gasoline or diesel powered counterparts in both performance and cost of ownership. The future of electric vehicles (EV) is bright with General Motors and other Original Equipment Manufacturers (OEM) committing to fully electrify their fleet of vehicles by 2030. The truck market is one of the largest automotive segments with over 750,000 Ford F150 trucks sold in 2020 alone. (Cain) To this point, major manufacturers have not mass produced an electric pickup truck. As gasoline prices rise and regulations tighten, the logical next step in the pickup truck evolution is a fully electric option. Studying fuel economy trends over time can help guide future development activities and specifications. The purpose of this paper will be to analyze the underlying fuel economy trends of the sedan segment and extrapolate this to the truck segment to determine a competitive fuel economy target for an EV pickup truck. While fuel economy alone is a misleading way to classify EVs, it is an easily understood stand in for cost of ownership. While fuel is not burned in an EV, an effective fuel economy can be calculated based on the variable costs associated with owning and driving an EV.. The results from analyzing the sedan trends will be applied to the pickup truck segment to guide specification development for major OEMs such as Ford or Toyota. The dataset used is publicly available from fueleconomy.gov and is valuable for any manufacturer to aid in developing fuel economy requirements. Fuel economy is very important as it not only determines how far a car can travel, but allows a firm to calculate the difference in cost of ownership for an EV truck versus a gasoline truck. This analysis will attempt to emulate the decision making of a typical Original Equipment Manufacturer (OEM), that’s starting a ground up EV truck project. Typical automotive development processes are 2-3 years long for a ground up new product so the analysis will attempt to predict effective fuel economy for the 2024 model year. Knowing the cost of ownership impact, an MSRP premium can be determined as the fuel cost savings can be used to offset the increased component cost of an EV powertrain.

1. **Method**

To learn more about the variable costs between owning an electric car and gasoline car, all types of gasoline cars and all electric cars currently in the market were studied.

Combined Fuel Efficiency (fe\_comb) is the dependent variable. Combined fuel efficiency is a weighted average of City and Highway MPG values that is calculated by weighting the City value by 55% and the Highway value by 45%.

*fe\_comb* = .55 \* *fe\_city* + .45 \* *fe\_hw*

Equation 1: Combined fuel economy calculation

**Dataset**

The dataset for this analysis was gathered from fueleconomy.gov and consists of the EPA rated fuel economies of all makes and models produced during each year. It was necessary to merge yearly data to create a master dataset consisting of gasoline and electric cars produced from 2016-2021.

**Data Preparation**

The first step in the analysis process was to collect the data and merge yearly datasets for gas and electric cats.

It was determined that exotic car manufacturers should not be included in the dataset since fuel economy is not a major selling point for their customers. Therefore, Bentley, Ferrari, Porsche, and similar luxury and exotic manufacturer data was excluded from our dataset as outliers.

With outliers and extra noise removed from the dataset, the data was simplified. The original nine car classes were condensed into four classes that more simply describe each car type: Large Cars, Midsize Cars, Small Cars, and Trucks.

Null values were filled as required. One example being that the manufacturer field for all electric vehicles had all the information in one cell and the manufacturer column was null. A ‘left string and find’ function was used to pull the information into the manufacturer field.

A number of the dependent variables in this dataset are categorical so it was necessary to encode them in preparation for the algorithms. Manufacturer, fuel type, and car class were encoded using the OneHotEncoder function in Python’s sklearn package.

Lastly, the Year variable was normalized for the dataset with 2015 corresponding to year 0, 2016 to year 1, etc.

**Data Exploration**

To understand the data most clearly, there are three different categories of interest: gasoline cars, gasoline trucks, and EV cars. Investigating the underlying trends of each segment, the analysis will be able to draw conclusions for a fourth category, EV trucks.

The data can be explored to visualize underlying trends. Figure 1 shows the average fuel economy for each year for each vehicle segment. Also shown on Figure 1 is the yearly trend of EV fuel economy. The trend was found using a simple linear regression calculation on the EV segment yearly average. The linear regression showed an average increase of 0.8mpg each year.

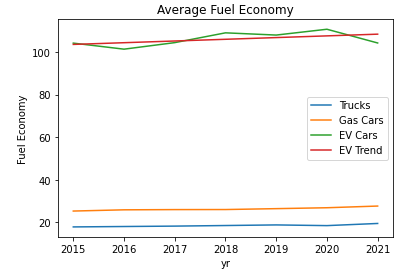


Figure 1: Year over year average fuel economy

Figure 2 is a subset of Figure 1 but only displays data from gasoline cars and trucks. The EV data is of a much higher magnitude so variations in the gas data cannot be seen on a combined plot. A linear regression was performed on each category to reveal a trend of increasing fuel economy of 0.33mpg and 0.22mpg for gasoline powered cars and trucks respectively.

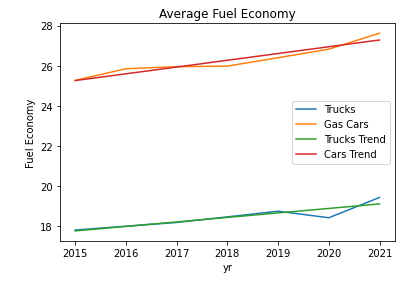


Figure 2: Year over year average fuel economy - Gasoline only

Figures 3 and 4 are also only related to gasoline powered vehicles. Looking at the data it seems that both of these features contribute to total fuel economy. These variables will be included in the predictor models.

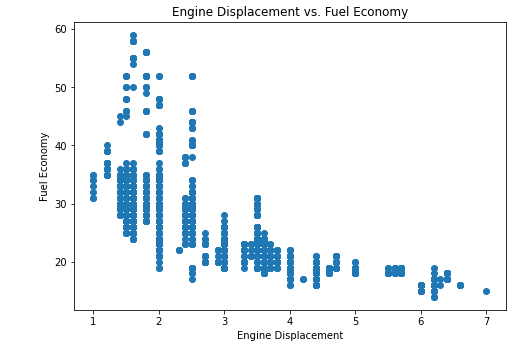


Figure 3: Engine displacement versus fuel economy

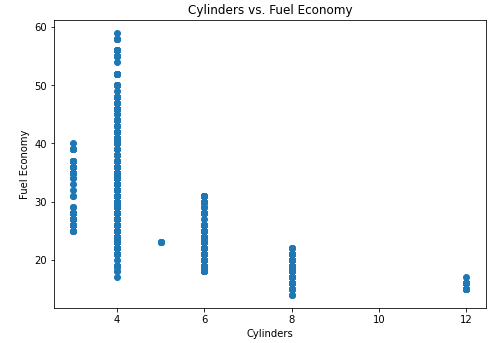


Figure 4: Engine Cylinders versus fuel economy

Another interesting plot compares total populations, EV to gasoline to see the underlying relationship. Figure 5 shows that there is a significant amount of variability in the data, especially the EV data. There are some outliers in the Gas category which when looking back at the data are hybrid vehicles such as the Toyota Prius and Hyundai Ioniq which have a fuel economy of greater than 50mpg. These data points were not excluded since regulations are continually tightening and hybrid vehicles are becoming more and more common. Excluding these data points would bias the models toward the past since they would be excluding the most advanced technologies for fuel efficiency.

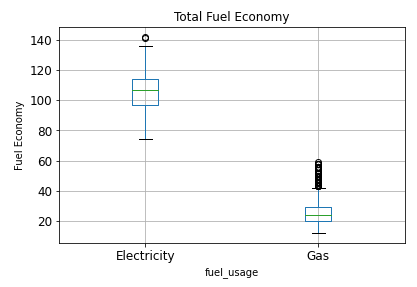


Figure 5: EV versus gas fuel economy

Figure 6 looks at the variability of fuel economy over time for each category. This is a deeper dive to the data presented in Figure 1, there are clearly a number of variables that contribute to the total fuel economy number.

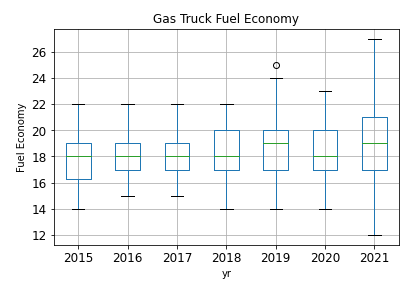
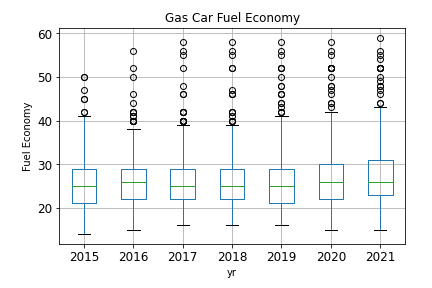
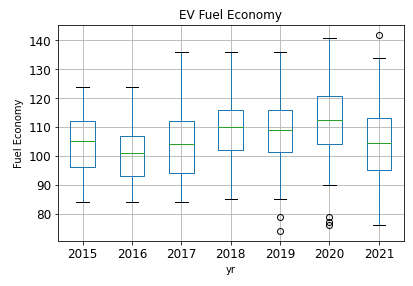


Figure 6: Fuel economy over time for each category

Investigation to this point revealed that it will not be possible to directly predict an accurate fuel economy for an EV truck based on the dataset used. The only dependent variables included in the dataset related to EV fuel economy are Year, Manufacturer, and Fuel Usage. As can be seen in Figure 6, there is a significant amount of variability to the EV fuel economies, thus for an accurate prediction, more independent variables would be needed. Since there are multiple EV models produced by each manufacturer each year, a prediction based on these variables will not be accurate. In order to draw any conclusions a different method of predicting fuel economy will be used. Luckily, there are more variables available for the gas data points: engine displacement and cylinder count. Therefore the method of prediction was adjusted. First, a prediction will be made for the 2024 model year truck and then the relationship between gas and EV cars will be used to calculate a viable fuel economy for an EV truck. Figure 7 is a flow chart showing the new analysis plan. Figure 8 is the relationship between the fuel economy of an average gasoline car and an average EV car for each year of data. Equation 2 shows the ratio calculation. Another linear regression was performed on these results to see if there is any change to the relationship over time. The result showed a decline in the ratio of 0.02 each year, which was determined to be a negligible change. A ratio of 4:1, EV FE to gas FE will be used for the final EV truck calculation.

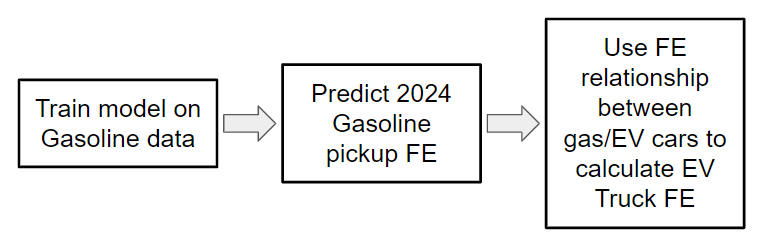


Figure 7: Analysis flow chart

*FE Ratio = EV FE / Gas FE*

Equation 2: EV to Gas Fuel economy ratio calculation

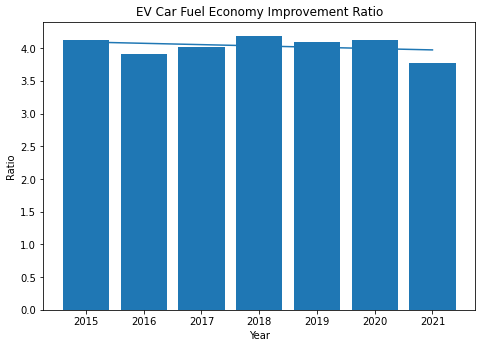


Figure 8: Fuel economy improvement for EV cars

**Analytical Models & Evaluation Metric**

With data exploration complete, and a method for predicting finalized, the prediction models were chosen. A number of different predictors were used with the performance of each algorithm compared using a Mean Squared Error calculation. Linear regression, Lasso, Ridge, Logistic regression, XGBoost, SVR, and LGBM were used to train and test algorithms. The dataset was randomly split into training and test sets by a 67/33 split. Table 1 shows the MSE for each algorithm. Also included in Table 1 is the MSE for models trained with the EV data included. Since the EV data is low featured, the MSE in each model is much higher than when training only gas cars and trucks only. This validates the decision to exclude EV car data from the prediction models. The algorithm is underfit when the EV data points are included. Table 1 also shows that most of the models investigated have a similar performance based on the MSE metric with Logistic Regression being an especially poor predictor and LGBM being a better than average predictor.

|  | **Gas & EV** | | **Gas Only** | |
| --- | --- | --- | --- | --- |
|  | Training MSE | Testing MSE | Training MSE | Testing MSE |
| Linear Regression | 20.98 | 3.22E+28 | 14.63 | 16.25 |
| Lasso | 25.85 | 28.91 | 14.96 | 16.41 |
| Ridge | 20.38 | 23.53 | 14.63 | 16.24 |
| Logistic Regression | 23.94 | 24.00 | 17.32 | 24.70 |
| XGBoost | 19.14 | 22.00 | 11.00 | 15.33 |
| SVR | 135.87 | 180.63 | 13.98 | 18.80 |
| LGBM | 19.53 | 20.26 | 11.51 | 10.49 |

Table 1: Algorithm MSE

1. **Results**

With the model development completed, the model can be applied to the business problem. A new data point was created to represent the vehicle of interest, namely a 2024 gas powered truck. The test data point was chosen to be a 2024 Ford Gas Truck with a 3.5L engine and 6 cylinders. These variables were chosen since the Ford F150 3.5L truck is the most popular and best selling model in 2021. This makes it a reasonable baseline for future developments. Each algorithm was then used to predict fuel economy and results are shown in Table 2. Table 2 also details the calculated EV truck FE based on Equation 3. Equation 3 is based on the ratio collected from Equation 2 and investigated in Figure 8.

*EV FE = 4 \* Gas FE*

Equation 3: EV Truck FE Calculation

|  | **Predicted Gas FE** | **Calculated EV FE** |
| --- | --- | --- |
| Linear Regression | 25.0 | 100.0 |
| Lasso | 25.0 | 99.9 |
| Ridge | 23.2 | 92.7 |
| Logistic Regression | 19.0 | 76.0 |
| XGBoost | 21.5 | 86.0 |
| SVR | 21.5 | 86.0 |
| LGBM | 19.6 | 78.4 |
| Trend analysis | 19.8 | 79.1 |

Table 2: Predicted FE of 2024 gas and EV truck FE

The seven models were applied and the results showed a target fuel economy of 80-100mpg. An eighth model was added as a comparison that was based on the linear regression shown in Figure 9, the same linear regression from Figure 2. This linear regression was calculated from the average year over year fuel economy of gas trucks and then extrapolated to the 2024 model year. Of the models studied, LGBM had the lowest MSE and thus was rated as the most accurate. The LGBM model also mimics the intuitive linear regression model. The MSE of the remaining models was significantly higher than the LGBM model. Sources of high MSE for these models may be the low number of features or including the hybrid outliers as discussed previously. To further reduce the MSE of any of the models, more features would need to be added to the dataset. The trend analysis method is more influenced by past developments than the latest advancements in fuel economy in the sedan segment.

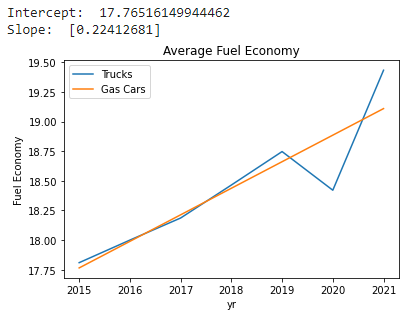


Figure 9: Linear regression of mean gas truck fuel economy

1. **Business Applications**

Results from this analysis can be used throughout any OEM to concentrate efforts on valuable projects. Engineering and marketing can both benefit from this single piece of data.

Sales and marketing can use the data obtained from this analysis to properly position products in development in the marketplace by not over or under pricing products and ensuring the effective fuel economy spec is aligned with market wants. Developing a pricing premium based on product features is essential to stay competitive in the marketplace today. Pricing too low leads to reduced profits and pricing too high results in lost sales to competitors. Using the algorithms outlined above, effective fuel economy for the pickup segment was predicted between 80-100mpg depending on the algorithm used. For the final analysis, the results from the LGBM model will be used since the LGBM model minimized the MSE. Assuming a user will own their car for 5 years, the length of a typical auto loan, the real and Net Present Value of cost savings due to electrification can be calculated. A discount rate of 5% was used as it’s the typical rate for an auto loan in 2021. NPV calculations using the results from the LGBM algorithm are shown in Table 3. A discounted cost savings of over $7,000 will be passed to the customer over a 5 year period. Since most consumers finance auto purchases, this $7,000 can be added to the selling price of the EV without impacting the actual ownership cost of an EV truck when compared to a similar gasoline truck.

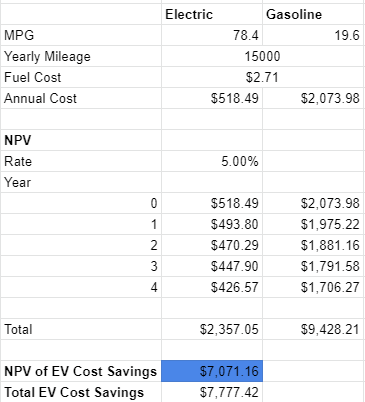


Table 3: NPV Analysis of cost of ownership of EV Truck

Similarly engineering can use this data to guide early development and ensure efforts are directed toward a product that will be successful. Fuel economy can guide decision making surrounding battery capacity and powertrain efficiency, two of the largest costs in building any EV. Having a concrete number to target can allow a narrower focus for the engineering team to weigh the benefits and drawbacks of different ideas and relate decisions back to the consumer in a meaningful way. The data transformation above discussing MSRP premium is also important to engineering teams as it can ease manufacturing cost requirements. Battery packs and powertrains are more expensive than a similar gasoline powertrain and as such the MSRP premium will allow the OEM to maintain acceptable margins in the cutthroat automotive market.

1. **Further Analysis**

Fuel economy is a challenging metric to predict as it is affected by many aspects of the car. As seen in the analysis the dataset used had some challenges in making accurate predictions. There are a number of ways that this analysis can and should be continued in order to make a more informed business decision. First and foremost, more independent variables are needed for predicting EV fuel economy beyond the scope of what is reported to the EPA. Variables that could be included to help improve accuracy of the EV algorithms are battery voltage, motor power rating, and aerodynamic drag coefficient. Similarly for the gas data points, engine power and aerodynamic drag would be helpful in making a more accurate prediction. These variables can be easily incorporated into these models, the challenge is collecting the data for six years of product lines.

1. **Conclusions**

Fuel economy is one of hundreds of features that are important to consumers. Unlike most, it will have a monetary impact on the consumer for the life of the product. Therefore it’s important that OEMs take great care in setting fuel economy targets. For established segments a year over year trend is sufficient to predict where competitors will be, for new segments the prediction requires more attention. This analysis shows that EVs enjoy a significant effective fuel economy advantage over their gasoline counterparts, this leads to a lower cost of ownership. However, since the technology is more expensive to produce an EV the upfront cost is necessarily higher. The pickup truck segment, when electrified, should target an effective fuel economy of at least 80mpg and a pricing premium of less than $7,000 in order to match the trends in the established EV car segments. Using this analysis, OEMs can push automobiles into the future while maintaining profitability.

1. **References**

Cain, T. (n.d.). Ford F-Series sales figures. Retrieved March 28, 2021, from <https://www.goodcarbadcar.net/ford-f-series-sales-figures/>

Diaz, M. N. (2020, February 14). *Electric Vehicles: A Primer on Technology and Selected Policy Issues* (Rep. No. R46231). Retrieved March 28, 2021, from Congressional Research Service website: <https://crsreports.congress.gov/product/pdf/R/R46231>

Download fuel economy data. (n.d.). Retrieved February/March, 2021, from https://www.fueleconomy.gov/feg/download.shtml

Gao Z et al. Energy Consumption and Cost Savings of Truck Electrification for Heavy-Duty Vehicle Applications. Transportation Research Record: Journal of the Transportation Research Board, No. 2628, Transportation Research Board, Washington, D.C., 2017, pp. 99–109. <https://dx.doi.org/10.3141/2628-11>

Loveday, S. (2018, July 5). What is mpge? Retrieved March 28, 2021, from <https://cars.usnews.com/cars-trucks/what-is-mpge>

Text version of the gasoline label. (2018, May 01). Retrieved April 25, 2021, from https://www.epa.gov/fueleconomy/text-version-gasoline-label#:~:text=Combined%20fuel%20economy%20is%20a,the%20Highway%20value%20by%2045%25.

Van Vliet, Oscar, et al. “Energy Use, Cost and CO2 Emissions of Electric Cars.” *Journal of Power Sources*, vol. 196, no. 4, 2011, pp. 2298–2310., doi:10.1016/j.jpowsour.2010.09.119.