Shifting Gears: *Does Transmission Technology Impact Fuel Efficiency?*

By James Lamb

Executive Summary

In this brief study, I examine the implications of transmission choice (automatic vs. manual) for fuel economy [1] in a cross section of 32 automobiles. This inquiry is undertaken using exploratory data analysis and multivariate regression on. The data are from the 1974 edition of Motor Trend magazine and are available in the "mtcars" dataset in the "'datasets"' package for R. The results of this research suggest that manual-transmission (so-called "stick shift") cars are expected to have better fuel economy than automatic-transmission cars with the same number of engine cylinders and weight. Given the small sample size, the fact that the data come from a single year, and my limited abilities in R, these results should only be interpreted as preliminary.

Evidence from the Literature

I began by reviewing recent academic research and articles in the popular press discussing the determinants of automobile fuel economy. The U.S. EPA (2012) finds a strong negative linear relationship between car weight and fuel economy. DriverSide (2014) asserts that higher displacement, the amount of air consumed by an engine in one rotation, should be associated with higher fuel efficiency and, by extension, fuel economy. In a recent release from *Motor Trend*, Kong (2013) suggests that the relationship between between cylinders and fuel economy for passenger vehicles is not categorically positive or negative; more cylinders means more power and lower fuel *efficiency*, opposing factors whose net impact is unclear. Kong (2013) also asserts that displacement and horsepower share a strongly positive linear relationship, suggesting that including both in the same regression would introduce multicollinearity and confound coefficient interpretation. Finally *Consumer Reports* (2014) asserts unequivocally that cars with manual transmisisons are expected to have better fuel economy, though this result comes from simplistic analysis that does not control for supervening factors such as those cited above.

Using this brief literature review as a guide, in this report I examine the relationship between transmission technology and fuel economy holding the effects of weight, displacement, and cylinders constant. In order to preserve the parsimony of the model, I restrict my analysis to this subset of variables.

Exploratory Analysis

To begin, I looked at unadjusted linear relationships between the variables of interest and both the dependent variable and their fellow regressors. Figure 1 gives scatterplots showing the *direction* of the relationship bewtween variable pairs, while Figure 2 gives correlation ellipses quantifying the *strength* of those relationships. Correlation ellipses which are tighter (less circular) suggest stronger correlation. I've included a vector of random pulls from the normal distribution ("noise") in Figure 2 as evidence of this. These figures provide some preliminary evidence that fuel economy varies negatively with changes in the number of cylinders, car weight, and displacement.

Model Selection

The evidence presented in figures 1 & 2 is in the form of unadjusted, simple linear correlation. To craft a more credible narrative, I next examined these relationships in the context of multivariable linear regressions. Parameters were estimated using Ordinary Least Squares (OLS). Figure 3 gives the results of this analysis. Equation 1 presents provides some evidence that cylinders and car wieght are negatively related to fuel economy. The estimated coefficient on displacement has the expected positive sign, but a t-test on that coefficient suggests that it is not statistically different from 0. Figure 2 suggests that a strong positive relationship exists between Weight and Displacement, so this low t-statistic could be the result of multicollinearity. Because the interest of this particular inquiry is the identification of individual relationships, not prediction, omitting one of these two variables is allowable and potentially useful. Equation 2 gives results when Weight is omitted and

[1] In this report, fuel refers to miles per gallon of gasoline ("mpg").

Equation 3 shows the outcome when Displacement is omitted. On the basis of Akaike's Information Criterion, Equation 3 is preferred.

A *t*-test on the estimated coefficient of Transmission in Equation 3 suggests that transmission technology has no impact on fuel economy. However, the observed negative relationship betweeen transmission technology and weight suggests that this result may be spurious, and the result of multicollinearity. In an attempt to correct this, I estimated Equation 4, which allows for interaction between transmission technology and weight. In confirmation of the multicollinearity story, this substantially improved the fit (on the basis of adjusted R^2 and AIC) and the significance of the individual coefficients.

Diagnostic Testing

Before proceeding to interpretation, I conducted several diagnostic tests on the estimates from Equation 4. Given the small sample size (only 32 observations), I did not hold these diagnostic tests to strict standards and used them only as a means of identifying extremely serious issues. Figure 5 shows some characteristics of the residuals. A histogram of standardized residuals suggests that they follow a somewhat-normal distribution (though with some positive skewness). The plot of residuals versus predicted values is also encouraging, suggesting no systematic relationship between the level of the dependent variable and the residuals. Note that the red line in this figure is not an arbitrarily appended horizontal line at 0, but a regression line fitted through the data in the plot.

In an attempt to assess the stability of regression coefficients (i.e. their sensitivity to the inclusion of new data), I created the leverage plots presented in Figure 4. Observations on this plots which appear further from the main grouping of data can be interepreted as exerting strong influence on the coefficient estimates. The plots look good (for such a small sample), and do not show any observations with sufficiently large influence to warrant omission.

Interpretation and Key Findings

The diagnostic tests provided strong evidence that Equation 4 was a reliable specification for answering the question at hand. Interpretation of the estimated coefficients from this model provided some useful preliminary insights into the relationship between fuel economy and the included car characteristics. Namely:

- A given car with average weight, manual transmission, and a 4-cylinder engine is expected to get 2.71 mpg better fuel economy than a similar car with a 6-cylinder engine, and 4.78 mpg better fuel economy than a similar car with an 8-cylinder engine.
- For a car with a given number of cylinders and manual transmission, a 1,000-pound increase in its weight is expected to reduce fuel economy by 2.40 mpg. For a similar car with *automatic* transmission, the same change in weight is expected to reduce fuel economy by 6.47 mpg.
- Most importantly, the results suggest that transmission technology does matter for fuel economy. A car with a given weight, number of cylinders, and automatic transmission is expected to have better fuel economy than a car with similar characteristics but manual transmission. This advantage is expected to be 11.57 mpg minus 4.07 mpg times the car's weight (in thousands of pounds).
  - Note, though, that the average weight of cars in the sample was 3,218 pounds. This implies that, on average, a car with automatic transmission is expected to have worse fuel economy than one with manual transmission.[2]

Conclusion

The results of this study provide preliminary evidence that, for all but the smallest autos, manual-transmission cars are expected to have better fuel economy than similar automatic-transmission cars.

The R markdown document with the underlying code used to generate this report can be found at this GitHub repo.

[2] The coefficients on Weight and the Weight-Transmission interaction lead to this interpretation. They suggest that automatic transmission implies worse fuel economy for cars with weight greater than 11.57/4.07 = 2,843 lbs.

Appendix: Plots and Regression Results

Figure 1. Plot of MPG vs. Explanatory Variables

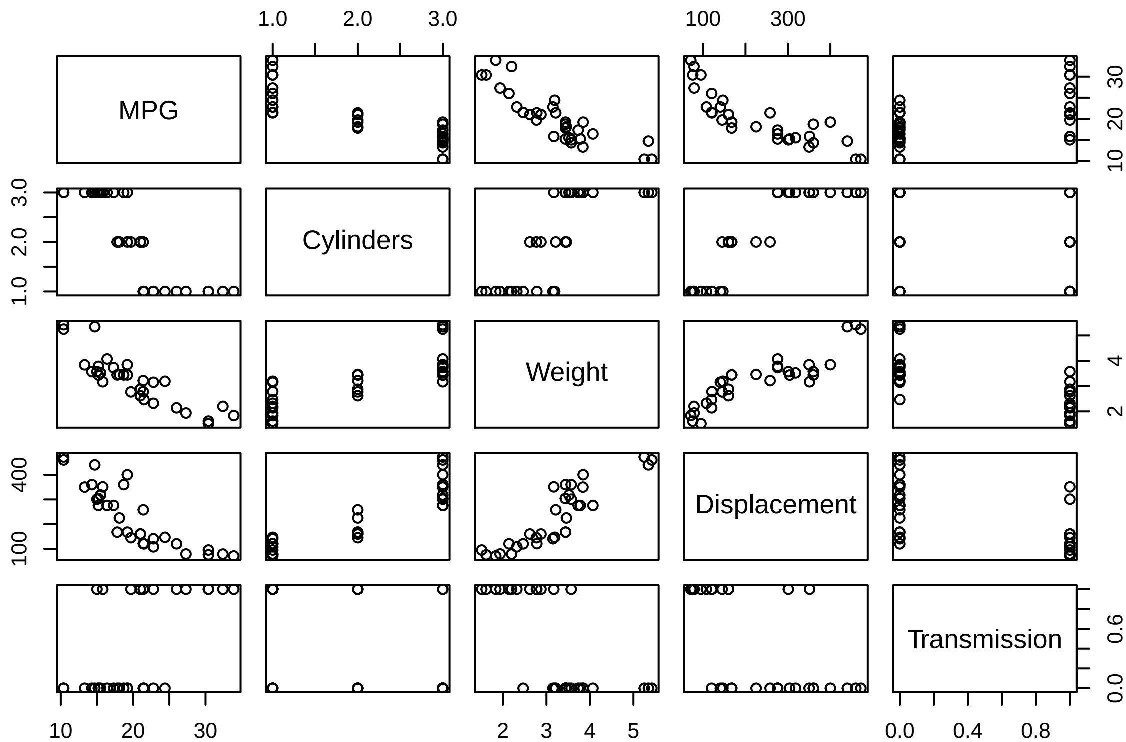


Figure 2. Simple Linear Correlations

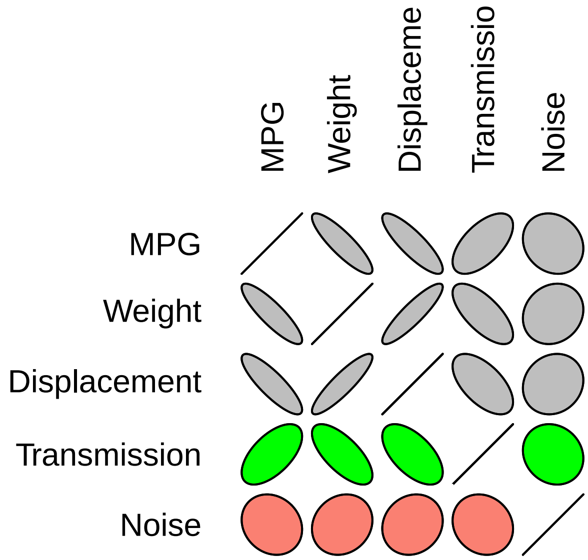


Figure 3. Regression Results

```
##
## Regression Results
## ==========================================================================================
##                                               Dependent variable:
##                      ---------------------------------------------------------------------
##                                                      MPG
##                              (1)               (2)               (3)               (4)
## ------------------------------------------------------------------------------------------
## Constant             33.820*** (2.914) 33.750*** (2.813) 27.830*** (1.970) 29.770*** (2.840)
## Cylinders6           -4.305*** (1.492) -4.257*** (1.411) -4.680*** (1.636)  -2.710* (1.357)
## Cylinders8           -6.318** (2.648)  -6.079*** (1.684)  -5.050* (2.867)  -4.776*** (1.556)
## Weight               -3.249** (1.249)  -3.150*** (0.908)                   -2.399*** (0.844)
## Displacement           0.002 (0.014)                      -0.022* (0.011)
## Transmission           0.141 (1.327)     0.150 (1.300)     1.636 (1.317)   11.570*** (4.088)
## Weight:Transmission                                                        -4.068*** (1.397)
## ------------------------------------------------------------------------------------------
## AIC                       160.589           158.607           165.991           151.579
## Observations                32                32                32                32
## R2                        0.838             0.838             0.795             0.877
## Adjusted R2               0.806             0.813             0.765             0.854
## ==========================================================================================
## Note:                                                        *p<0.1; **p<0.05; ***p<0.01
```

Figure 4. Equation 4 Leverage Plots

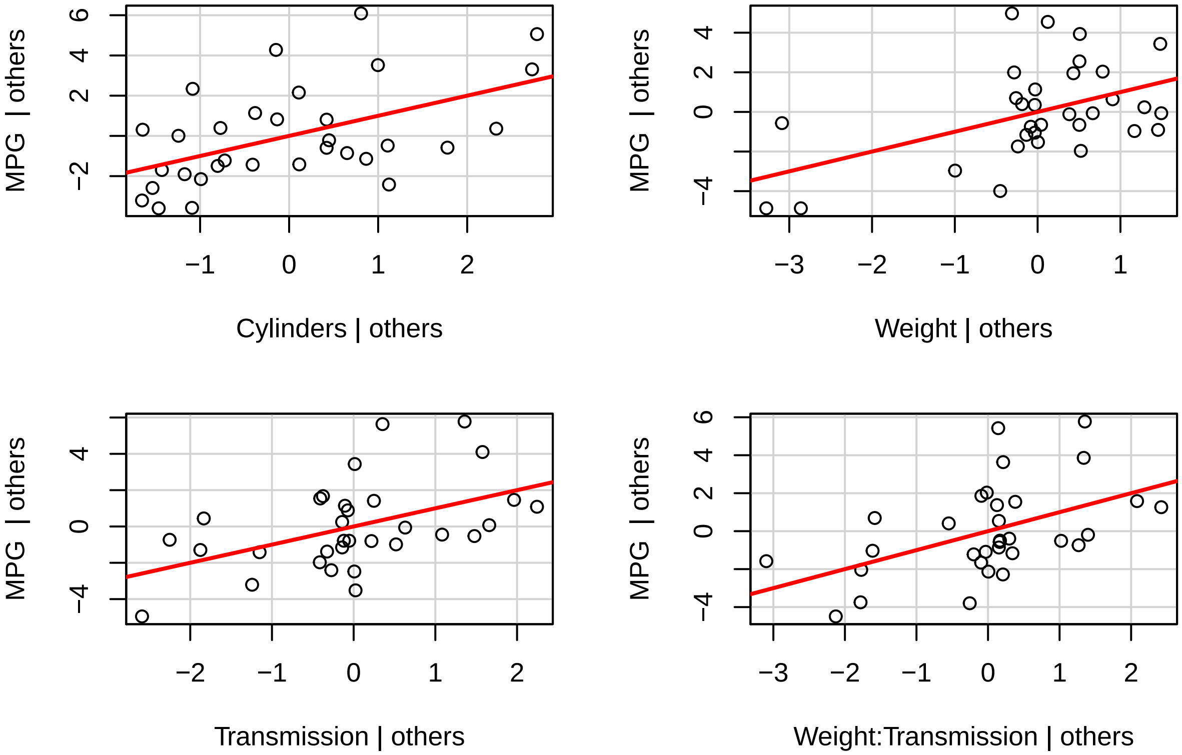
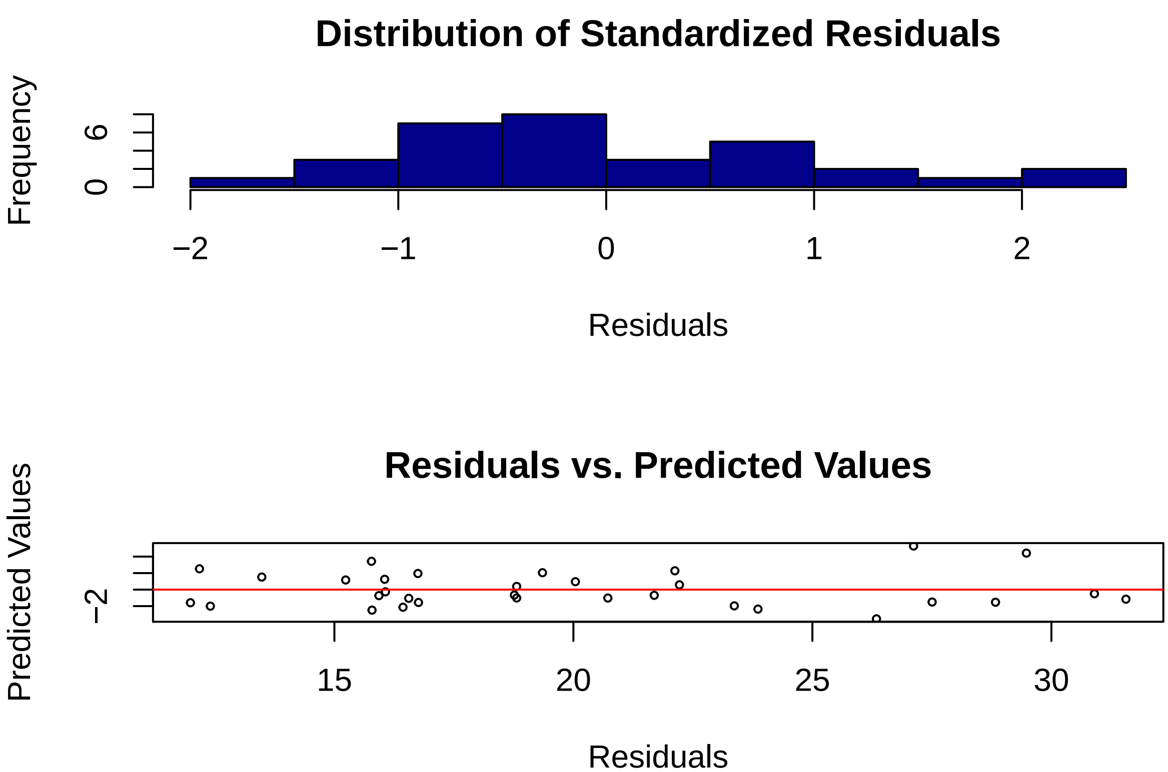


Figure 5. Regression Diagnostics: Residuals



References

[1] Benson Kong, "More Power and More MPG: How Automakers are Improving Both By Optimizing Today's Gasoline Engines" *Motor Trend*, April 2013 Retrieved from: http://bit.ly/1uq0tcf

[2] Consumer Reports, "Save gas and money with a manual transmission", February 2014. Retrieved from: http://bit.ly/1wBfKsd

[3] DriverSide, "Top 10 Factors Contributing to Fuel Economy". Retrieved from: http://bit.ly/1tIL3BK

[4] Frank McCown, "Producing Simple Graphs with R", 2007. Retrieved from: http://bit.ly/Zy8Jxt

[5] Friedrich Leisch and R-core, "Sweave User Manual" April 2014 Retrieved from: http://bit.ly/1sdZJdI

[6] Marek Hlavac, "stargazer: LaTeX/HTML code and ASCII text for well-formatted regression and summary statistics tables." April 2014 Retrieved from: http://bit.ly/1AYItrv

[7] Paul E. Johnson, "Outreg rockchalk", 2014. Retrieved from: http://bit.ly/Z3vSrN

[8] Robert I. Kabacoff, "Quick-R: Regression Diagnostics", 2014. Retrieved from: http://bit.ly/1BYs8FV

[9] —. "Quick-R: Combining Plots", 2014. Retrieved from: http://bit.ly/1wTQUaa

[10] Suman Gautam, "What Factors Affect Average Fuel Economy of U.S. Passenger Vehicles?", *Honors Projects*, Illinois Wesleyan University. 2010. Retrieved from: http://bit.ly/1rbm0oT

[11] Winston Chang, "Correlation Matrix", *Cookbook for R*, O'Reilly Media, December 2012. Retrieved from: http://bit.ly/1msTHn8