My Paper

Julian Reif*

March 24, 2020

Abstract

This paper provides an example of a document with tables and figures that were automated using Stata.

*University of Illinois and NBER.

1 Summary

This example paper includes tables and figures that were created using Stata and outputted into the **/analysis/results** project folder. Copy the contents of that folder to **/paper** to update the tables and figures in this document.

Table 1 reports summary statistics for Stata's **auto.dta** dataset.[1] The average price of automobiles in this dataset is $6,165. The price distribution, iillustrated in Figure 1, is skewed right.

I estimate the association between automobile prices and fuel efficiency using the following linear model:

$$PRICE_i = \alpha + \beta X_i + \varepsilon \tag{1}$$

The outcome variable, $PRICE_i$, is the price of automobile i. The parameter of interest is β, a vector of coefficients. In my first specification ("spec 1"), the vector X_i includes miles per gallon. The second specification ("spec 2") also includes the car's weight. I estimate this model using ordinary least squares and report standard errors that are robust to heteroskedasticity. The analysis is performed first using Stata, and then repeated using R.

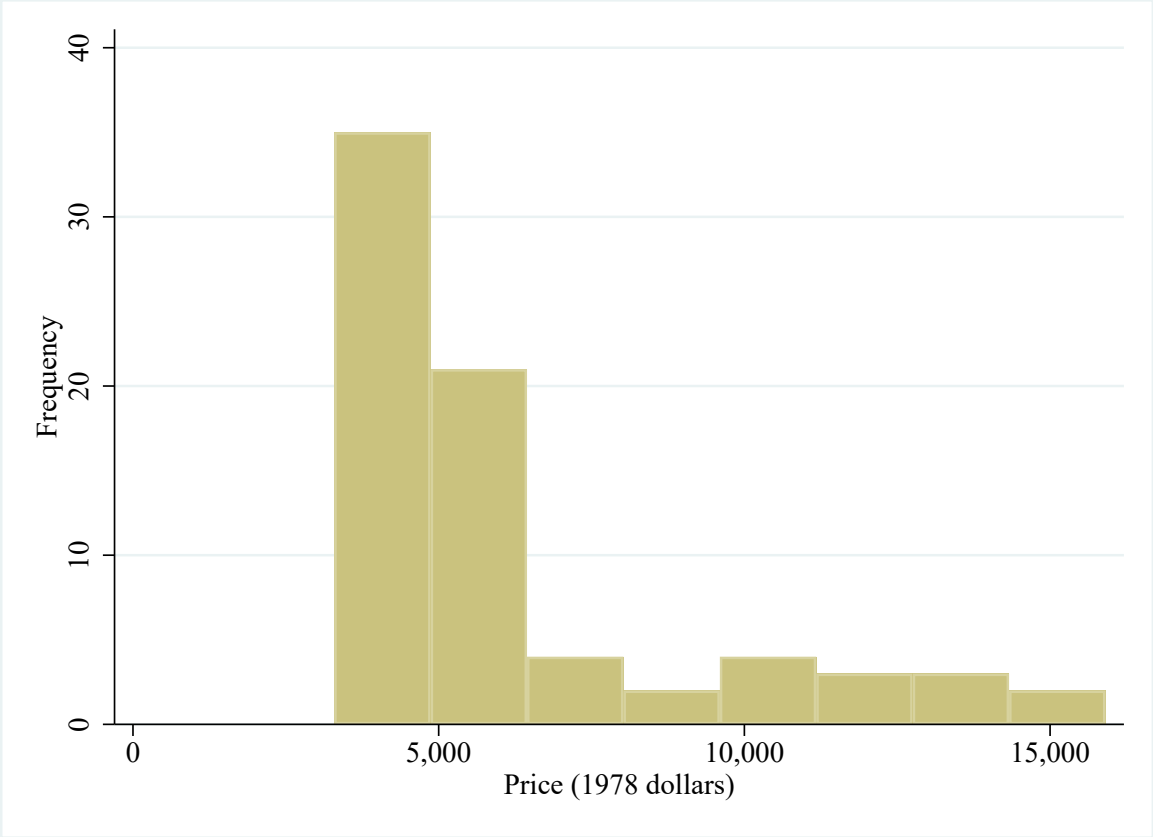
Table 2 reports my Stata estimates, separately for domestic and foreign cars. Column (1) reports that an increase in fuel efficiency of 1 mile per gallon is associated with a $329 reduction in the price of domestic automobiles. Column (2) shows that this association becomes positive and insignificant when I also include weight as a regressor. Columns (3) and (4) show that these associations are similar for foreign automobiles.

Table 3 compares these Stata estimates to estimates from R. Panel A reproduces the Stata estimates that were presented in Table 2. Panel B of Table 3 reports estimates when I repeat this analysis in R using the **lm_robust** command from the *estimatr* package. The point estimates and the standard errors are identical across both software packages.

[1] Type **sysuse auto, clear** at the Stata prompt to load this dataset.

2 Figures and Tables

Figure 1: Automobile prices



Notes: Data were obtained from Stata's built-in auto dataset.

Table 1: Summary statistics

	Mean	Stdev.	Min	Max	Count
Weight (pounds)	3,019	777	1,760	4,840	74
Miles per gallon	21.3	5.79	12	41	74
Price (1978 dollars)	6,165	2,949	3,291	15,906	74

Notes: Count reports the number of non-missing values for the variable.

Table 2: Association between automobile price and fuel efficiency

	(1)	(2)	(3)	(4)
	Domestic		Foreign	
	Spec 1	Spec 2	Spec 1	Spec 2
Miles per gallon	-329***	238	-250**	-19.8
	(81.2)	(203)	(88.2)	(51.7)
Weight (pounds)		4.42***		5.16***
		(1.34)		(0.770)
N	52	52	22	22
R^2	0.254	0.483	0.399	0.785

Notes: Outcome variable is price (1978 dollars). Columns (1) and (2) report estimates of β from equation (1) for domestic automobiles. Columns (3) and (4) report estimates for foreign automobiles. Robust standard errors are reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table 3: Association between automobile price and fuel efficiency, Stata and R

	(1)	(2)	(3)	(4)
	Domestic		Foreign	
	Spec 1	Spec 2	Spec 1	Spec 2
A. Stata output (regress)				
Miles per gallon	-329***	238	-250**	-19.8
	(81.2)	(203)	(88.2)	(51.7)
Weight (pounds)		4.42***		5.16***
		(1.34)		(0.770)
B. R output (lm_robust)				
Miles per gallon	-329***	238	-250**	-19.8
	(81.2)	(203)	(88.2)	(51.7)
Weight (pounds)		4.42***		5.16***
		(1.34)		(0.770)

Notes: Outcome variable is price (1978 dollars). Columns (1) and (2) report estimates of β from equation (1) for domestic automobiles. Columns (3) and (4) report estimates for foreign automobiles. Robust standard errors are reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.