Spring 2021 **Firstname M. Lastname**

University of Central Florida
College of Business

QMB 6911
Capstone Project in Business Analytics

Solutions: Problem Set #11

Chapter 1

Introduction

1.1 Introduction

In this paper I analyze the rates of return from sales of used trucks. My aim is to produce a model that recommends whether a particular used truck that was purchased as a trade-in should be sold on the retail lot or at an auction.

In the following pages, I will investigate these questions and ultimately fit a model for the rates of return from sales of used trucks., allowing for endogeneity in the choice of the characteristics of trucks produced and sold by each dealer. To understand the dynamics of a sample selection model, it is worthwhile to understand the research related to these questions.

1.2 Economic Theory

1.2.1 Market for Lemons

Akerlof's description of the market for lemons is a long-standing contribution to the understanding of asymmetric information.

1.2.2 Characteristic Theory

The paper by Lancaster (1966) is a contribution to economic theory in which he makes the case for a sound theory of consumer choice, in which the consumer's preferences are defined on the characteristics of different goods, rather than the quantities of uniformly-defined goods.

1.2.3 Hedonic Pricing Models

Rosen (1974) builds on the work of Lancaster (1966) by providing an empirical framework for estimating the value of products based on their characteristics.

1.3 Empirical Framework

1.3.1 Tobit Models

Heckman (1979) described sample selection models as a model specification question.

Amemiya (1984) summarized the variaous types of Tobit models and provides a review of research with applications of these models.

Lee and Trost (1978) is a well-known application to housing markets, which incorporates the fact that the residents make a decision to either own or rent a home before buying. This is similar to the decision over whether to sell a used truck at auction or on the retail lot.

Chapter 2

Data Description

2.1 Data Description

A data engineer for several dealerships has gathered relevant and appropriate information, and organized a dataset concerning 9,861 sales involving a trade-in of a truck at nine dealerships (some having more than one location), around one-half of which were sold at retail, while the other one-half were sold at auction. These data are contained in the file `UsedTrucks.dat`, which is available in the `Data` folder. Each vehicle in the dataset is a row, while the columns correspond to the variables whose names and definitions are the following:

Variable	Definition
`type`	sale type (an integer); `Auction=1` versus `Retail=0`
`pauc`	price when sold at auction (an integer)
`pret`	price when sold retail (an integer)
`mileage`	odometer mileage (an integer)
`make`	make of vehicle (an integer)
`year`	model year of vehicle (an integer)
`damage`	an index of damage to vehicle; `1` little damage, `10` a lot (an integer)
`dealer`	dealer id (an integer)
`ror`	rate-of-return (a real number)
`cost`	net amount given to trade-in (an integer)

Some of the variables above warrant some description. The first set of variables describe the characteristics of the vehicles. The `damage` feature variable was constructed by workers at the various dealership, and is really just an ordinal ranking, at best. The `dealer` variable is an integer between `1` and `9`. The `make` variable is an integer between `1` and `9`, where `1` and `2` denote Ford, while `3` denotes Chevrolet, `4` Dodge, `5` General Motors, `6` Toyota, `7` Nissan, and `8` Subaru, and `9` others. The `cost` variable is how much the dealer gave the buyer of a new vehicle for the truck given in trade, plus any taxes and fees that might be associated with the truck, rounded to the nearest integer.

The variables listed above also include the dependent variables. The `ror` variable is the natural logarithm of the ratio either the price garnered at auction `pauc` or the price earned from retail to the variable `cost`; it is a real number. Suppose a particular vehicle `cost` $1,000 net in trade, and the vehicle sold $1,100, then the rate-of-return `ror` would be $\log(1.1) = 0.0953101\ldots$, or about 9.5 percent.

Chapter 3

Analysis of the Dependent Variable

3.1 Transforming the Dependent Variable

3.1.1 Probability Density Function of ROR on Used Trucks

Figure 3.1 shows the kernel-smoothed probability density function of ROR on trucks.

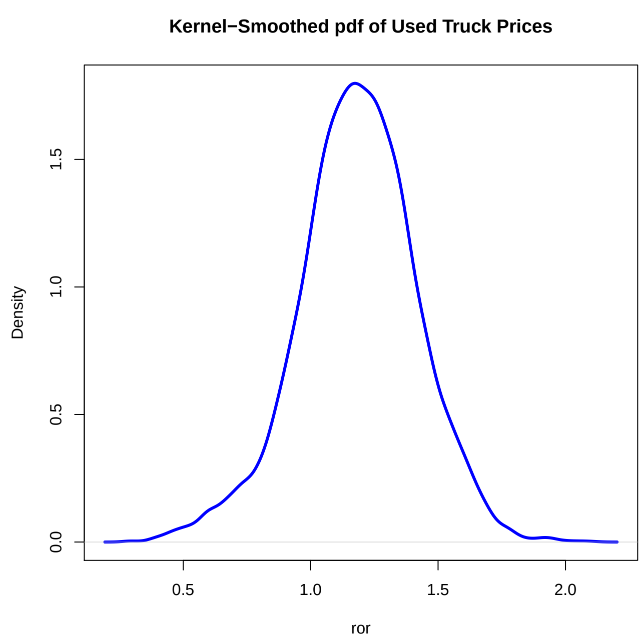


Figure 3.1: Probability Density Function of ROR on Used Trucks

As a comparison, Figure 3.2 shows the kernel-smoothed probability density function of the natural logarithm of the rate of return.

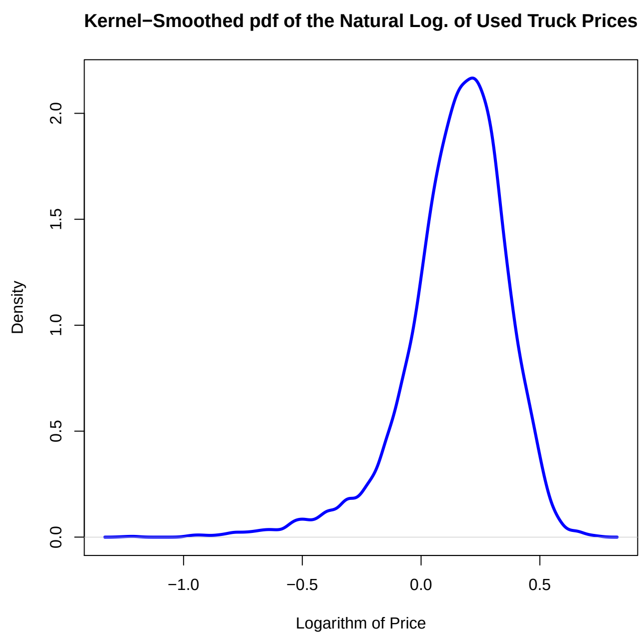
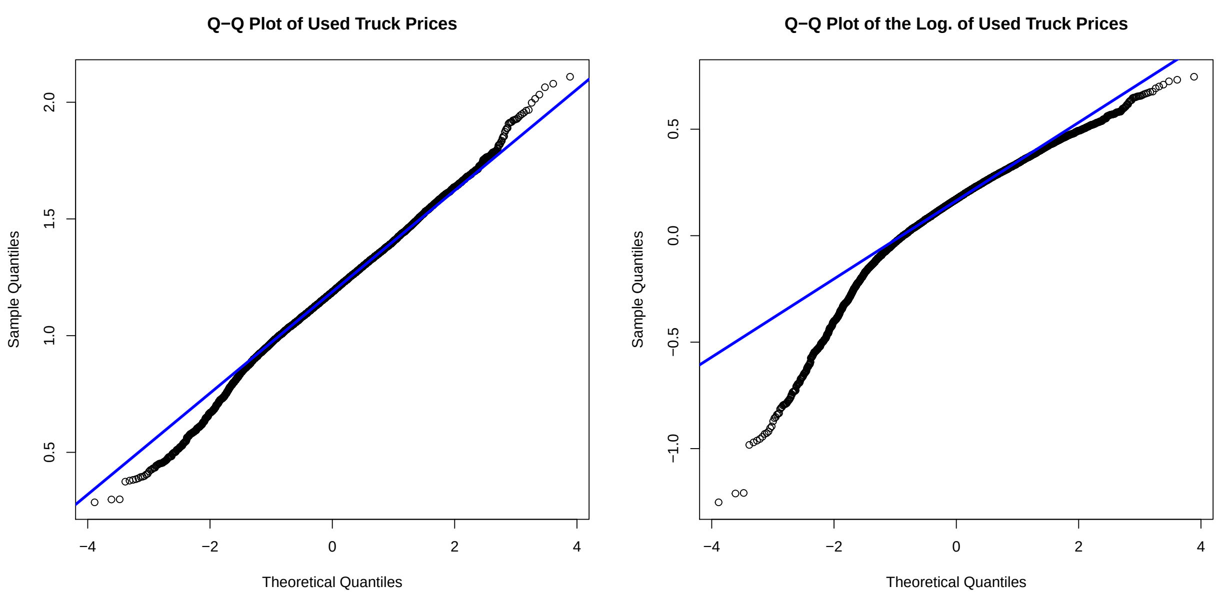


Figure 3.2: Probability Density Function of the Logarithm of ROR on Used Trucks

3.1.2 Normality of the Original and Transformed Variables

Figure 4.2 shows a pair of Q-Q plots, comparing quantiles of the empirical distribution against the quantiles of the normal distribution. In the left panel, Figure 4.2a shows this comparison for the original level of the ROR on trucks, without transformation. In the right panel, Figure 3.3b shows this comparison for the logarithmic transformation of ROR on trucks, without transformation. Consistent with the pair of distributions estimated above, each plot shows a divergence from a normal distribution, suggesting that an optimal transformation might lie somewhere else. The Box-Cox transformation allows for this possibility.



(a) ROR on Used Trucks

(b) Transformed ROR on Used Trucks

Figure 3.3: Q-QPlots of the Log. and Levels of ROR on Used Trucks

Chapter 4

Transforming the Dependent Variable

4.0.1 Box-Cox Transformation of ROR on Used Trucks

Under the Box–Cox transformation of R_n, the price of truck n is calculated as follows,

$$\Lambda(R_n) \equiv \begin{cases} \frac{R_n^{\lambda}-1}{\lambda} & \text{if } \lambda > 0 \\ \log R_n & \text{if } \lambda = 0. \end{cases}$$

The following code block defines a function that performs a Box-Cox transformation. Note that the function that we used for tractor and fly reel prices works just as well for the dependent variable rate of return.

```
# Box-Cox transformation.
Lambda_Price <- function(price, lambda) {
  if (lambda == 0) {
    return(log(price))
  } else {
    return((price^lambda - 1)/lambda)
  }
}
```

Log-likelihood Function

Under the Box-Cox transformation, the ROR on tractors can be decomposed into a location parameter μ^0 and an error U, so

$$\Lambda(R_n) = \mu^0(\lambda) + U_n,$$

where the U_ns are independent, mean-zero, constant-variance $\sigma^2(\lambda)$, Gaussian (normal) errors. In the above equation, for clarity, the dependence of μ^0 and $\sigma^2(\lambda)$ on λ is made explicit.

The next code block defines a likelihood function for the normal distribution of the errors as a function of the parameter λ.

```
log_like_uni <- function(price, lambda) {

  # Calculate maximum likelighood estimates of the parameters.
  n <- length(price)
  lambda_price <- Lambda_Price(price, lambda)
  mu_0_lambda <- mean(lambda_price)
  sigma_2_lambda <- sum((lambda_price - mu_0_lambda)^2)/n

  # Calculate the log-likelihood from the sum of the logarithms
  # of the density of the normal distribution.
  like <- - n/2*log(2*pi*sigma_2_lambda)
  like <- like - 1/2/sigma_2_lambda*sum((lambda_price - mu_0_lambda)^2)
  like <- like + (lambda - 1)*sum(log(price))
  return(like)
}
```

As a first approximation, One can calculate the value of the log-likelihood function on a grid of values to find an optimal value of λ. The plot of this likelihood function is shown in Figure 4.1. The red points represent the values of the log-likelihood at the optimum $\hat{\lambda}$ and at $\lambda = 0$ and $\lambda = 1$.

```
# Calculate values of the log-likelihood function.
lambda_grid <- seq(-1, 2.5, by = 0.001)
like_grid <- 0*lambda_grid
for (lambda_num in 1:length(lambda_grid)) {
  like_grid[lambda_num] <- log_like_uni(price = trucks[, 'ror'],
                                        lambda = lambda_grid[lambda_num])
}
```

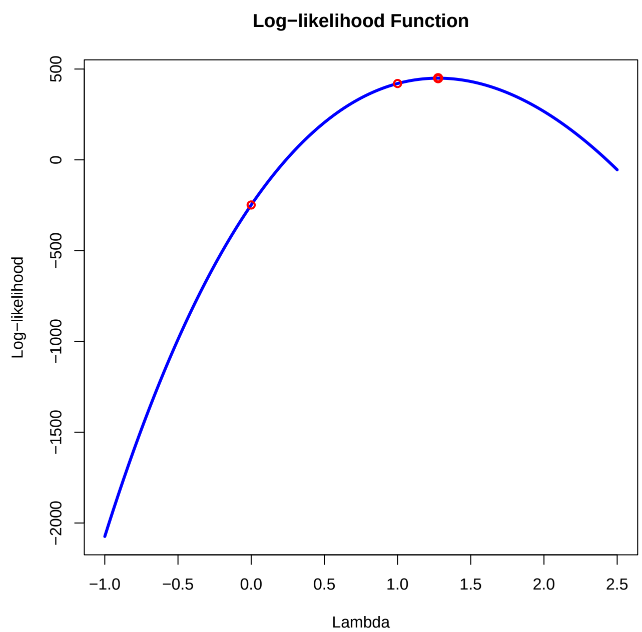


Figure 4.1: Log-likelihood Function for Box-Cox Transformation

Testing for an Appropriate Transformation

Now we consider the statistical properties of these estimates by calculating a likelihood ratio statistic.

```
> # Calculate likelihood ratio statistics.
> LR_stat_0 <- - 2*(like_mu_0 - like_MLE)
> print(LR_stat_0)
[1] 10.26956
> LR_stat_1 <- - 2*(like_mu_1 - like_MLE)
> print(LR_stat_1)
[1] 483.2539
>
>
> # Compare to quantile of chi-squared distribution with 1 degree of freedom.
> LR_cv_5 <- qchisq(p = 0.95, df = 1)
> print(LR_cv_5)
[1] 3.841459
>
> # Calculate p-values for these tests.
> p_value_0 <- 1 - pchisq(q = LR_stat_0, df = 1)
> print(p_value_0)
[1] 0.001352434
> p_value_1 <- 1 - pchisq(q = LR_stat_1, df = 1)
> print(p_value_1)
[1] 0
>
```

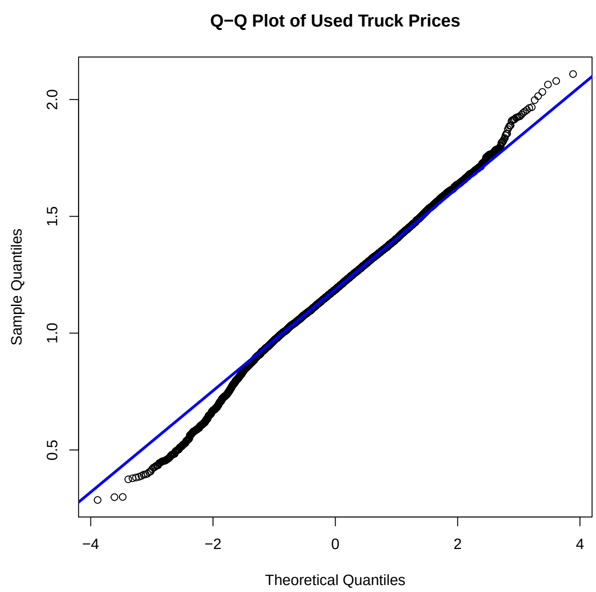
Statistically, this is evidence to reject them both. This suggests using the transformation at the MLE. However, one may want to investigate further to find out whether it is worth transforming the data, since the Box-Cox transformation at the MLE offers only a marginal improvement over the original variable. There exists a trade-of between interpretability and the accuracy of the statistical specification, and, perhaps, the original variable is close enough for practical purposes.

4.0.2 Normality of the Transformed Variable

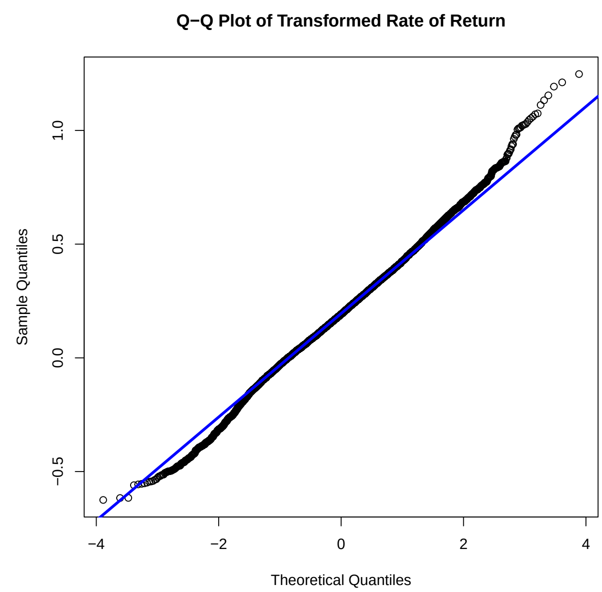
Now compare the quantiles of the distribution of the transformed variable with the original. We already plotted normal QQ plot for ROR on tractors when considering the log transformation. Now we can generate a new dependent variable with the results from the estimates above.

```
# Generate new dependent variable with results from estimates above.
trucks[, 'trans_ror'] <- Lambda_Price(price = trucks[, 'ror'],
                                     lambda = lambda_hat)
```

Figure 4.2 shows this comparison and the panel on the right, Figure 4.2b, shows that the quantiles of the distribution of the transformed variable nearly overlap with those of the normal distribution. From a purely statistical perspective, this provides evidence that the prices are best modeled with the transformation at the optimal λ just above one. From a practical point of view, however, the added complexity is not warranted when the original variable is close enough to normally distributed.



(a) ROR on Used Trucks



(b) Transformed ROR on Used Trucks

Figure 4.2: Q-QPlots of the Transformed ROR on Used Trucks

Chapter 5

Preliminary Graphical Analysis

5.1 Transforming the Dependent Variable

5.1.1 Introduction

This note summarizes the findings in the script `UsedTruck_Data_Vis.R`, which analyzes the prices of used trucks, the dependent variable in the `UsedTrucks.dat` dataset. The output includes several plots of the dependent variable against the explanatory variables.

The primary goal is to determine the relative value of trucks in both the auction and retail market, taking into account the fact that the dealers decided to send each truck to a particular market. A consequence of this analysis is a recommendation engine to determine whether used trucks with particular characteristics should be sent to either the auction or retail market.

5.1.2 Histogram and Density of ROR on Truck Sales

All Trucks Together

Plot the histogram and density together with the following code.

```
hist(trucks[, 'ror'],
    main = 'Histogram and Density of ROR on Truck Sales',
    xlab = 'Price', col = 'red',
    probability = TRUE)
rug(trucks[, 'ror'])
lines(density(trucks[, 'ror']),
     col = 'blue', lwd = 3)
```

Figure 5.1 is a histogram of the logarithm of ROR on trucks, along with a rug plot and a kernel density estimate, generated by the code block above. Without taking logs, we can see that the

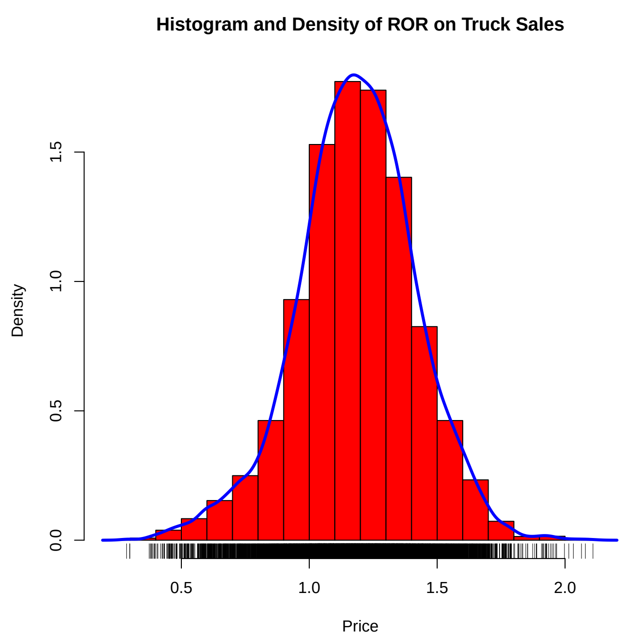


Figure 5.1: Relative Histogram of ROR on Used Trucks

distribution is approximately symmetric, with some bunching in the upper tail but most of the mass centered about the mean.

Comparison By Sale Type

Now we investigate the value of trucks sold in the different markets. Figure 5.2 shows the kernel density estimate of the rates of return of trucks sold at auction in green and those sold in the retail market in red. The distribution of rates of return is nearly symmetric for retail sales. The distribution for sales at auction has some asymmetry in that a higher fraction of trucks are sold at a loss when sold at auction.

```
library(sm)
sm.density.compare(trucks[, 'ror'],
                   trucks[, 'type'],
                   xlab = "ROR on Truck Sales",
                   lwd = 3,
                   col = c('red','green'))
title(main = 'ROR on Truck Sales by Sale Type')
legend('topright', c('Retail', 'Auction'),
       fill = c('red','green'),
       cex = 0.75)
```

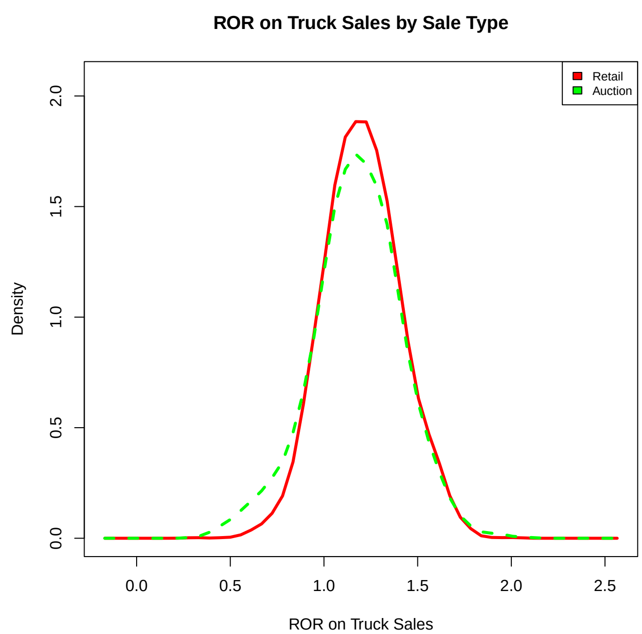


Figure 5.2: Densities of ROR on Truck Sales by Sale Type

Comparison By Make

Figure 5.3 shows the densities of the rate of return, separated by the name of the manufacturer. The figure was generated by the following code.

```
sm.density.compare(trucks[, 'ror'],
                   trucks[, 'make_f'],
                   xlab = "ROR on Truck Sales",
                   lwd = 3,
                   col = rainbow(length(levels(trucks[, 'make_f']))),
                   xlim = c(0.85, 1.25))
title(main = 'ROR on Truck Sales by Make')
legend('topright', unique(make_list_f),
       fill = rainbow(length(levels(trucks[, 'make_f']))),
       cex = 0.5)
```

We see that the distribution of rates of return differs very little by make.

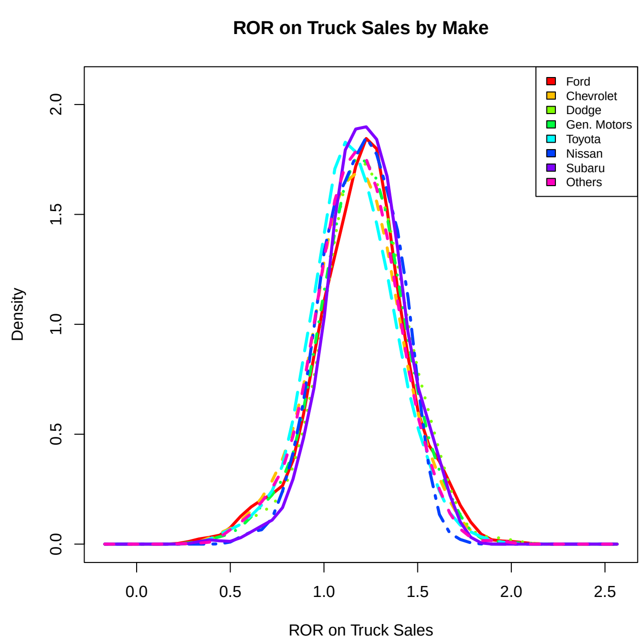


Figure 5.3: Densities of ROR on Truck Sales by Manufacturer

5.1.3 Sales Volume By Make and Type of Sale

Figure 5.4 shows a spinogram of the number of truck sales by manufacturer and the type of market in which the truck is sold. It is generated by the following code.

```
# Create a table and plot it in a spinogram.
counts <- table(trucks[, 'make_f'],
            trucks[, 'type_f'])

# Plot the spinogram.
spine(counts,
     main = 'Spinogram of Sales by Make and Type of Sale')
```

The code block first tabulates the number of sales by manufacturer and whether or not the truck was sold in auction or retail. These counts are shown in Table 5.1.

	Auction	Retail
Chevrolet	532	215
Dodge	1491	602
Ford	1180	621
Gen. Motors	1027	501
Nissan	916	367
Others	159	113
Subaru	206	137
Toyota	1273	521

Table 5.1: Sales Volume by Make and Type of Sale

It appears that sales of used trucks are tilted toward the auction market for Chevrolet and Dodge, among the domestic makes, and for Toyota and Nissan, among the imports.

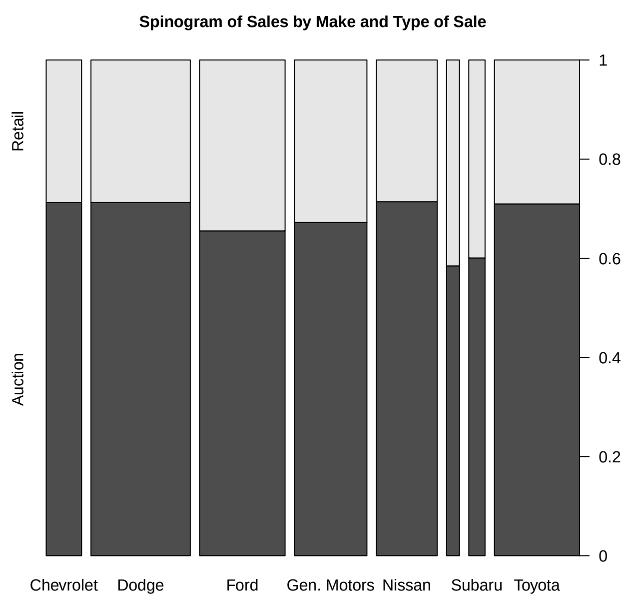


Figure 5.4: Sales Volume of ROR on Used Trucks by Make and Type of Sale

5.1.4 Scatterplot Matrix of Numeric Variables

Figure 5.5 shows a scatterplot of the numeric variables in the dataset, which include the rate of return, along with the year of manufacture of the vehicle, the mileage, and the level of damage of the vehicle. The `gclus` package was used to cluster the correlation matrix to color code by strength of correlation. The correlation matrix is shown in Table 5.2.

	ROR	Mileage	Year	Damage
ROR	1.000	-0.416	0.382	-0.587
Mileage	-0.416	1.000	-0.906	0.560
Year	0.382	-0.906	1.000	-0.510
Damage	-0.587	0.560	-0.510	1.000

Table 5.2: Correlation Matrix of Numeric Variables

The correlation matrix amd the scatterplot matrix were generated by the following code.

```
# Select some numerical variables.
colnames(trucks)
col_sel_num <- c(9, 4, 6, 7)
colnames(trucks)[col_sel_num]
col_sel_text <- c('ROR', 'Mileage', 'Year', 'Damage')

# Create a covariance matrix and determine
# parameters for scattergraph matrix.
mydata <- trucks[, col_sel_num]
mydata.corr <- abs(cor(mydata))
mycolors <- dmat.color(mydata.corr)
# Order by magnitude of correlation.
myorder <- order.single(mydata.corr)

# Plot the scatterplot matrix.
cpairs(mydata,
      myorder,
      panel.colors = mycolors,
      gap = 0.5,
      main = c('Scatterplot Matrix Colored by Correlation')
)
```

We see some clear correlation between the variables damage, year and mileage. Older vehicles are more likely to be damaged and to have high mileage.

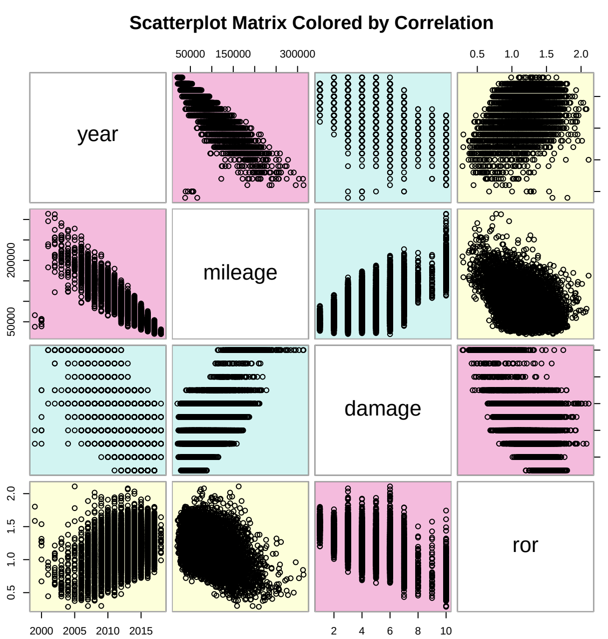


Figure 5.5: Scatterplot Matrix Colored by Strength of Correlation

5.1.5 Relationship between ROR, mileage and damage

Figure 5.6 shows a bubble plot, which is a form of scatterplot in which the size of the dots (the "bubbles") represent another variable. This analysis is based on the above investigation of the numeric variables in the dataset, which include the rate of return on the sale, the mileage of the vehicle and the level of damage of the vehicle. The rates of return are shown in the vertical axis, the mileage of the truck is on the horizontal axis, and the area of each bubble is proportional to the damage of the vehicle. The bubbleplot is generated by the following code.

```
# Calculate the radius of the bubbles
# so that the area represents horsepowwer.
r <- sqrt(trucks[, 'damage']/pi)

# Plot the bubble plot.
fig_file_name <- 'bubble_plot.pdf'
out_file_name <- sprintf('%s/%s', fig_dir, fig_file_name)
pdf(out_file_name)
symbols(trucks[, 'mileage'],
      trucks[, 'ror'],
      r,
      inches=0.30, fg="white", bg="lightblue",
      main = "Bubble Plot with point size proportional to damage",
      ylab = "ROR",
      xlab = "Mileage")
```

There appears to be lower returns on trucks with high mileage and more damage on those vehicles as well.

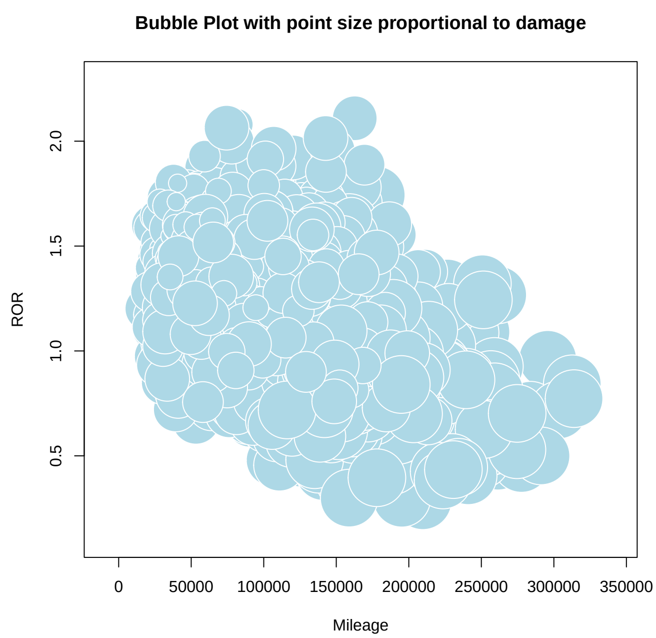


Figure 5.6: Bubble Plot with Point Size Proportional to Damage

Analysis by Sale Type

Figure 5.7 shows a bubble plot, similar to that in Figure 5.6, except that only the sales in the retail market are shown. The negative relationship between ROR and mileage still holds, and is perhaps more defined. Also, there exists quite a few more trucks with low damage, shown by the smaller bubbles.

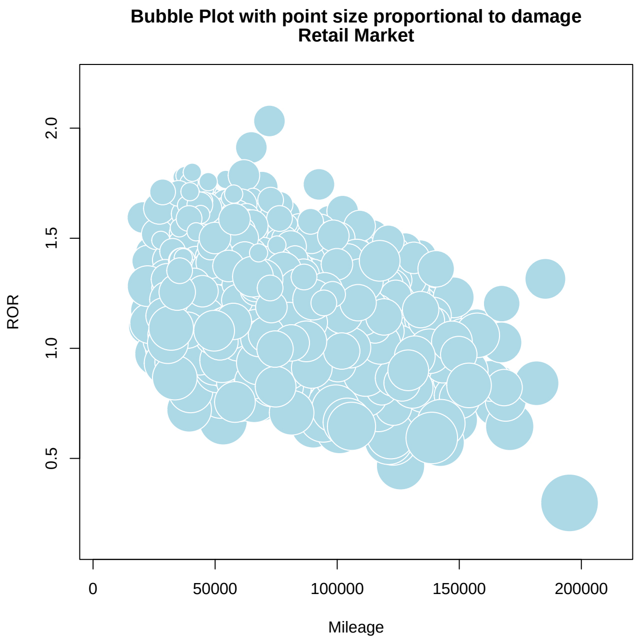


Figure 5.7: Bubble Plot with Point Size Proportional to Damage: Retail Market

Figure 5.8 shows another bubble plot, with the sales in the auction market. In contrast to the retail market, the lower right quadrant is more densely populated, indicating that many damages trucks with high mileage are sold at auction. Also, the overall size of the bubbles is larger, indicating that few vehicles with low damage are sold through auctions.

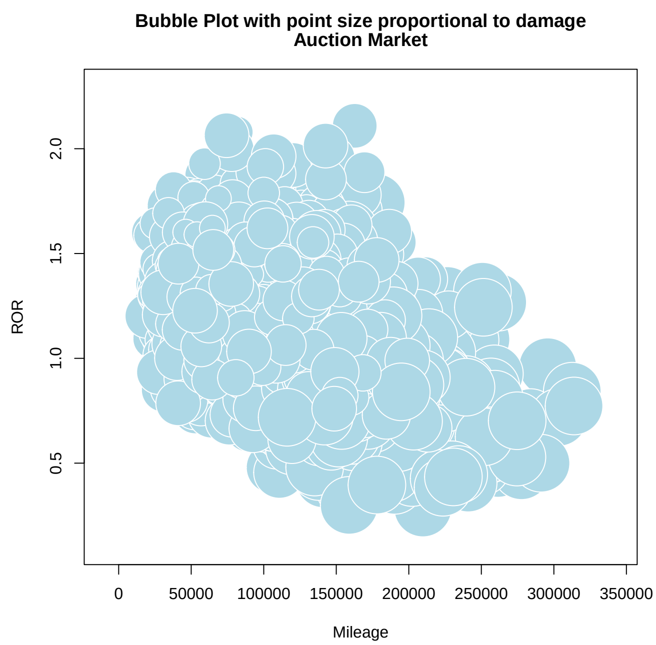


Figure 5.8: Bubble Plot with Point Size Proportional to Damage: Auction Market

Chapter 6

Preliminary Tabular Analysis

I analyze the data in subsets, according to whether the truck was sold in the auction or retail market, calculating the summary statistics for each subset and present these statistics in the LaTeX tables that follow.

6.1 Summary by Type of Sale

Table 6.1 lists summary statistics for numeric variables in separate columns for subsamples defined by the type of sale.

	Retail	Auction
Min. ror	0.2858	0.2979
Mean ror	1.1737	1.2049
Max. ror	2.1092	2.0326
Min. log_ror	-1.25233	-1.21090
Mean log_ror	0.13609	0.17146
Max. log_ror	0.74630	0.70931
Min. age	2.000	2.000
Mean age	7.493	5.908
Max. age	21.000	21.000
Min. mileage_K	19.78	20.40
Mean mileage_K	95.09	71.47
Max. mileage_K	313.90	195.16
Min. damage_num	1.000	1.000
Mean damage_num	4.865	3.614
Max. damage_num	10.000	10.000

Table 6.1: Summary by Type of Sale

In this table, we can see what a difference the sale type is worth: ROR is slightly higher for retail sales. In terms of the continuous explanatory variables, Auctioned vehicles are older, have higher mileage and have more damage. Together, this indicates the potential for sample selection and for a reason to prefer to sell used trucks on the retail lot, at least for some vehicles. An analysis of categorical variables might offer further insight into the differences between sales..

6.2 Type of Sale by Make of Vehicle

Table 6.2 lists the frequencies of observations of each make of vehicle by type of sale. Some makes of vehicle are more likely to go to auction, indicating that this may be an important variable for selection into type of sale.

	Auction	Retail	Totals
Chevrolet	532	215	747
Dodge	1491	602	2093
Ford	1180	621	1801
Gen. Motors	1027	501	1528
Nissan	916	367	1283
Others	159	113	272
Subaru	206	137	343
Toyota	1273	521	1794
Total	6784	3077	9861

Table 6.2: Type of Sale by Make of Vehicle

6.3 Type of Sale by Dealer

Table 6.3 lists the frequencies of observations of each each dealer by type of sale. Some dealers have more active retail lot for used trucks, indicating that this may also be an important variable for selection into type of sale and perhaps the return from each market.

	Auction	Retail	Totals
1	584	305	889
2	275	127	402
3	952	209	1161
4	1099	530	1629
5	338	135	473
6	1293	586	1879
7	519	234	753
8	1413	806	2219
9	311	145	456
Total	6784	3077	9861

Table 6.3: Type of Sale by Dealer

Chapter 7

Regression Modelling with Hedonic Price Theory

7.1 Pooled Regression Models

The results in Table 7.1 shows the effect of the variables on the ROR and log. ROR of cars, each with a single pooled model for both the auction and retail markets. Model 1 is the model for ROR and Model 2 is the model for the logarithm of ROR. All of the variables have statistically significant coefficients Both models have similar results but the variable `log_ror` had a distribution that was skewed left and the original variable `ror` had a symmetric distribution that appeared nearly normal. Without a large difference in model fit, the original variable is the better choice.

	ROR	Log. ROR
(Intercept)	1.6349^{***}	0.5841^{***}
	(0.0104)	(0.0092)
auction	0.0863^{***}	0.0793^{***}
	(0.0041)	(0.0036)
dealer2	-0.0588^{***}	-0.0530^{***}
	(0.0105)	(0.0094)
dealer3	0.0763^{***}	0.0626^{***}
	(0.0079)	(0.0070)
dealer4	-0.0661^{***}	-0.0607^{***}
	(0.0073)	(0.0065)
dealer5	0.0664^{***}	0.0524^{***}
	(0.0100)	(0.0089)
dealer6	0.0177^{*}	0.0129^{*}
	(0.0071)	(0.0063)
dealer7	0.0420^{***}	0.0350^{***}
	(0.0087)	(0.0077)
dealer8	-0.0622^{***}	-0.0582^{***}
	(0.0070)	(0.0062)
dealer9	-0.0920^{***}	-0.0812^{***}
	(0.0101)	(0.0090)
makeDodge	-0.0119	-0.0072
	(0.0074)	(0.0066)
makeFord	-0.0080	-0.0075
	(0.0076)	(0.0068)
makeGen. Motors	-0.0197^{*}	-0.0162^{*}
	(0.0078)	(0.0069)
makeNissan	-0.0657^{***}	-0.0555^{***}
	(0.0080)	(0.0072)
makeOthers	-0.0778^{***}	-0.0662^{***}
	(0.0124)	(0.0111)
makeSubaru	-0.0552^{***}	-0.0494^{***}
	(0.0115)	(0.0102)
makeToyota	-0.0578^{***}	-0.0490^{***}
	(0.0076)	(0.0068)
damage_num	-0.0777^{***}	-0.0737^{***}
	(0.0012)	(0.0011)
mileage_K	-0.0008^{***}	-0.0009^{***}
	(0.0001)	(0.0001)
age	-0.0070^{***}	-0.0059^{***}
	(0.0015)	(0.0013)
R^2	0.4346	0.4756
Adj. R^2	0.4335	0.4746
Num. obs.	9861	9861

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 7.1: Pooled Regression Models: ROR and Log. ROR

7.2 Separate Models by Sale Type

To test for many possible differences in models by brand of tractor, Table 7.2 shows the estimates for two separate models by brand of tractor. Model 1 shows the estimates for the full sample, Model 2 shows the estimates from the full model for trucks sold at auction and Model 3 represents trucks sold on the retail lot.

We can also test for all of the differences at the same time by using an F-test. In this case, the full, unrestricted model has $K = 2 \times 19 = 18$ parameters, one for each variable in two models. The test that all of the coefficients are the same has $M = 19 - 1 = 18$ restrictions. The one restriction fewer accounts for the auction indicator in the full model, which allows for two separate intercepts. The F-statistic has a value of

$$\frac{(RSS_M - RSS)/M}{RSS/(N-K-1)} = \frac{(299.7855 - 289.0844)/18}{289.0844/9822} = 20.19907.$$

This is a very high value for the F-statistic. A value around 20 is more than 10 times the critical value of the F-statistic at the 1% level of significance. Conclude that ROR differs by sale type, so interaction terms should be considered, if not two separate models.

	Full Sample	Auction	Retail
(Intercept)	1.6349^{***}	1.6799^{***}	1.8294^{***}
	(0.0104)	(0.0138)	(0.0141)
auction	0.0863^{***}		
	(0.0041)		
dealer2	-0.0588^{***}	-0.0478^{***}	-0.0786^{***}
	(0.0105)	(0.0139)	(0.0130)
dealer3	0.0763^{***}	0.0740^{***}	0.0663^{***}
	(0.0079)	(0.0101)	(0.0111)
dealer4	-0.0661^{***}	-0.0672^{***}	-0.0708^{***}
	(0.0073)	(0.0098)	(0.0089)
dealer5	0.0664^{***}	0.0547^{***}	0.0953^{***}
	(0.0100)	(0.0130)	(0.0127)
dealer6	0.0177^{*}	0.0132	0.0309^{***}
	(0.0071)	(0.0095)	(0.0087)
dealer7	0.0420^{***}	0.0331^{**}	0.0696^{***}
	(0.0087)	(0.0115)	(0.0107)
dealer8	-0.0622^{***}	-0.0618^{***}	-0.0788^{***}
	(0.0070)	(0.0094)	(0.0083)
dealer9	-0.0920^{***}	-0.0957^{***}	-0.0970^{***}
	(0.0101)	(0.0133)	(0.0124)
makeDodge	-0.0119	-0.0107	-0.0231^{*}
	(0.0074)	(0.0096)	(0.0097)
makeFord	-0.0080	-0.0094	-0.0176
	(0.0076)	(0.0099)	(0.0097)
makeGen. Motors	-0.0197^{*}	-0.0231^{*}	-0.0134
	(0.0078)	(0.0101)	(0.0100)
makeNissan	-0.0657^{***}	-0.0627^{***}	-0.0800^{***}
	(0.0080)	(0.0104)	(0.0105)
makeOthers	-0.0778^{***}	-0.0849^{***}	-0.0871^{***}
	(0.0124)	(0.0173)	(0.0143)
makeSubaru	-0.0552^{***}	-0.0629^{***}	-0.0686^{***}
	(0.0115)	(0.0157)	(0.0135)
makeToyota	-0.0578^{***}	-0.0564^{***}	-0.0736^{***}
	(0.0076)	(0.0098)	(0.0100)
damage_num	-0.0777^{***}	-0.0706^{***}	-0.1083^{***}
	(0.0012)	(0.0015)	(0.0018)
mileage_K	-0.0008^{***}	-0.0008^{***}	-0.0021^{***}
	(0.0001)	(0.0001)	(0.0002)
age	-0.0070^{***}	-0.0065^{***}	-0.0043^{*}
	(0.0015)	(0.0019)	(0.0020)
R^2	0.4346	0.3916	0.6429
Adj. R^2	0.4335	0.3900	0.6408
Num. obs.	9861	6784	3077

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 7.2: Separate Regression Models: ROR for Auction vs. Retail Sales

7.3 Interaction Terms

In this section, I will explore interactions between the continuous variables and the market in which the vehicles are sold. Although there appears to be differences in the make and dealer categories, the differences do not appear to be as large as those for the continuous variables. In addition, separate terms for these categories by sale type would introduce many degrees of freedom and risk overfitting the model.

Differences in Depreciation by Sale Type

The columns of Table 7.3 show the results of tests for interactions between the sale type on the effects of age, mileage and damage. Each interaction term is significant, along with the separate intercpt for the type of sale. When all three interactions are included, however, the age interaction is no longer significant. This suggests a single slope coefficient for age but separate slopes for mileage and damage by type of sale, along with a separate intercept by type of sale. The difference in slope coefficients are sizeable fractions of the slope coefficients for mileage and damage.

	Age	Mileage	Damage	Miles and Damage	All Interactions
(Intercept)	1.7260^{***}	1.7416^{***}	1.7691^{***}	1.8148^{***}	1.8128^{***}
	(0.0133)	(0.0132)	(0.0131)	(0.0143)	(0.0145)
auction	-0.0241^{*}	-0.0381^{***}	-0.0772^{***}	-0.1292^{***}	-0.1263^{***}
	(0.0110)	(0.0105)	(0.0107)	(0.0126)	(0.0130)
dealer2	-0.0596^{***}	-0.0594^{***}	-0.0570^{***}	-0.0576^{***}	-0.0576^{***}
	(0.0105)	(0.0104)	(0.0104)	(0.0103)	(0.0103)
dealer3	0.0757^{***}	0.0752^{***}	0.0737^{***}	0.0734^{***}	0.0734^{***}
	(0.0079)	(0.0078)	(0.0078)	(0.0078)	(0.0078)
dealer4	-0.0671^{***}	-0.0676^{***}	-0.0678^{***}	-0.0684^{***}	-0.0684^{***}
	(0.0073)	(0.0073)	(0.0072)	(0.0072)	(0.0072)
dealer5	0.0681^{***}	0.0680^{***}	0.0650^{***}	0.0663^{***}	0.0662^{***}
	(0.0099)	(0.0099)	(0.0098)	(0.0098)	(0.0098)
dealer6	0.0179^{*}	0.0180^{*}	0.0187^{**}	0.0187^{**}	0.0187^{**}
	(0.0071)	(0.0071)	(0.0070)	(0.0070)	(0.0070)
dealer7	0.0430^{***}	0.0439^{***}	0.0435^{***}	0.0445^{***}	0.0446^{***}
	(0.0087)	(0.0086)	(0.0086)	(0.0086)	(0.0086)
dealer8	-0.0645^{***}	-0.0654^{***}	-0.0667^{***}	-0.0680^{***}	-0.0680^{***}
	(0.0069)	(0.0069)	(0.0069)	(0.0069)	(0.0069)
dealer9	-0.0934^{***}	-0.0943^{***}	-0.0950^{***}	-0.0960^{***}	-0.0960^{***}
	(0.0100)	(0.0100)	(0.0099)	(0.0099)	(0.0099)
makeDodge	-0.0138	-0.0140	-0.0138	-0.0149^{*}	-0.0148^{*}
	(0.0074)	(0.0074)	(0.0073)	(0.0073)	(0.0073)
makeFord	-0.0098	-0.0098	-0.0113	-0.0119	-0.0118
	(0.0076)	(0.0076)	(0.0075)	(0.0075)	(0.0075)
makeGen. Motors	-0.0204^{**}	-0.0202^{**}	-0.0201^{**}	-0.0204^{**}	-0.0203^{**}
	(0.0078)	(0.0077)	(0.0077)	(0.0077)	(0.0077)
makeNissan	-0.0678^{***}	-0.0679^{***}	-0.0672^{***}	-0.0683^{***}	-0.0682^{***}
	(0.0080)	(0.0080)	(0.0079)	(0.0079)	(0.0079)
makeOthers	-0.0825^{***}	-0.0828^{***}	-0.0830^{***}	-0.0853^{***}	-0.0852^{***}
	(0.0124)	(0.0124)	(0.0123)	(0.0123)	(0.0123)
makeSubaru	-0.0613^{***}	-0.0620^{***}	-0.0623^{***}	-0.0655^{***}	-0.0653^{***}
	(0.0114)	(0.0114)	(0.0113)	(0.0113)	(0.0113)
makeToyota	-0.0599^{***}	-0.0601^{***}	-0.0607^{***}	-0.0617^{***}	-0.0616^{***}
	(0.0076)	(0.0076)	(0.0075)	(0.0075)	(0.0075)
damage_num	-0.0789^{***}	-0.0795^{***}	-0.1112^{***}	-0.1066^{***}	-0.1067^{***}
	(0.0012)	(0.0012)	(0.0024)	(0.0024)	(0.0024)
mileage_K	-0.0010^{***}	-0.0022^{***}	-0.0010^{***}	-0.0019^{***}	-0.0020^{***}
	(0.0001)	(0.0002)	(0.0001)	(0.0002)	(0.0002)
age	-0.0194^{***}	-0.0063^{***}	-0.0063^{***}	-0.0060^{***}	-0.0039
	(0.0019)	(0.0015)	(0.0015)	(0.0015)	(0.0028)
auction:age	0.0180^{***}				-0.0029
	(0.0017)				(0.0033)
auction:mileage_K		0.0017^{***}		0.0011^{***}	0.0013^{***}
		(0.0001)		(0.0001)	(0.0003)
auction:damage_num			0.0429^{***}	0.0355^{***}	0.0356^{***}
			(0.0026)	(0.0028)	(0.0028)
R^2	0.4412	0.4439	0.4496	0.4531	0.4531
Adj. R^2	0.4401	0.4428	0.4485	0.4519	0.4519
Num. obs.	9861	9861	9861	9861	9861

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 7.3: Regression Models with Interaction Terms

Chapter 8

Nonparametric Regression Models

8.1 Linear Regression Model

A natural staring point is the recommended linear model from Problem Set #7.

The results of this regression specification are shown in Table 8.1.

Next, I will attempt to improve on this specification.

	Model 1
(Intercept)	1.8148^{***}
	(0.0143)
auction	-0.1292^{***}
	(0.0126)
dealer2	-0.0576^{***}
	(0.0103)
dealer3	0.0734^{***}
	(0.0078)
dealer4	-0.0684^{***}
	(0.0072)
dealer5	0.0663^{***}
	(0.0098)
dealer6	0.0187^{**}
	(0.0070)
dealer7	0.0445^{***}
	(0.0086)
dealer8	-0.0680^{***}
	(0.0069)
dealer9	-0.0960^{***}
	(0.0099)
makeDodge	-0.0149^{*}
	(0.0073)
makeFord	-0.0119
	(0.0075)
makeGen. Motors	-0.0204^{**}
	(0.0077)
makeNissan	-0.0683^{***}
	(0.0079)
makeOthers	-0.0853^{***}
	(0.0123)
makeSubaru	-0.0655^{***}
	(0.0113)
makeToyota	-0.0617^{***}
	(0.0075)
damage_num	-0.1066^{***}
	(0.0024)
mileage_K	-0.0019^{***}
	(0.0002)
age	-0.0060^{***}
	(0.0015)
auction:damage_num	0.0355^{***}
	(0.0028)
auction:mileage_K	0.0011^{***}
	(0.0001)
R^2	0.4531
Adj. R^2	0.4519
Num. obs.	9861

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 8.1: Recommended Regression Model

8.2 Nonlinear Specifications

8.2.1 Nonparametric Specification for Age

First, consider a set of nonparametric specifications for the relationship between ROR on used trucks and age. Figure 8.1 depicts a plot of the deviations from predicted ROR against the residuals from the regression for age: the "excess age" compared to what would be expected given the other characteristics of a used truck. Figure 8.1 also overlays the nonparametric estimate (shown in green).

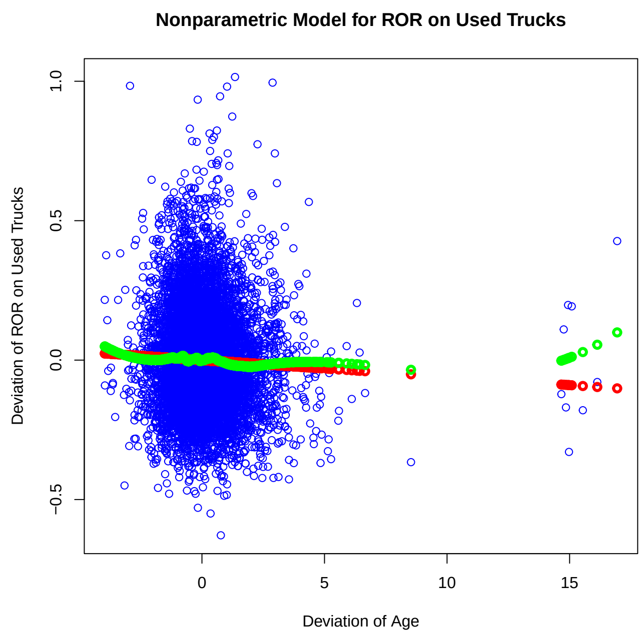


Figure 8.1: Nonparametric Model for ROR on Used Trucks: Excess Age

8.2.2 Nonparametric Specification for Mileage

First, consider a set of nonparametric specifications for the relationship between ROR on used trucks and mileage. Figure 8.2 depicts a plot of the deviations from predicted ROR against the residuals from the regression for mileage: the "excess mileage" compared to what would be expected given the other characteristics of a used truck. Figure 8.2 also overlays the nonparametric estimate (shown in green).

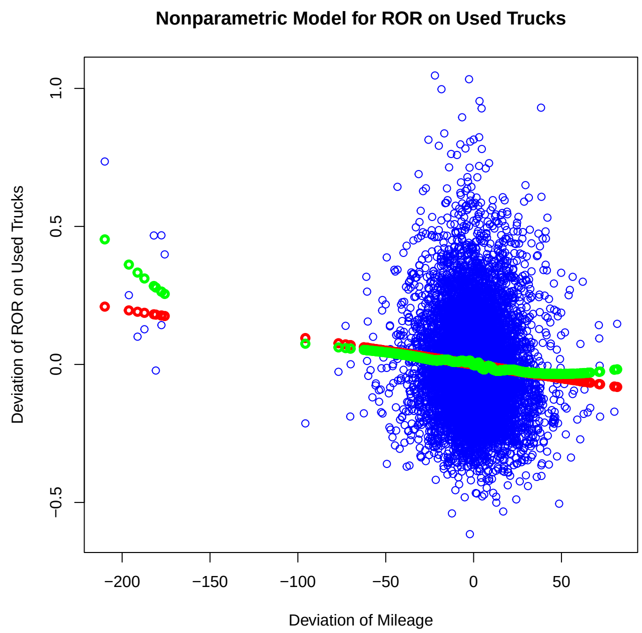


Figure 8.2: Nonparametric Model for ROR on Used Trucks: Excess Mileage

8.2.3 Nonparametric Specification for Damage

First, consider a set of nonparametric specifications for the relationship between ROR on used trucks and damage. Figure 8.3 depicts a plot of the deviations from predicted ROR against the residuals from the regression for damage: the “excess damage” compared to what would be expected given the other characteristics of a used truck. Figure 8.3 also overlays the nonparametric estimate (shown in green).

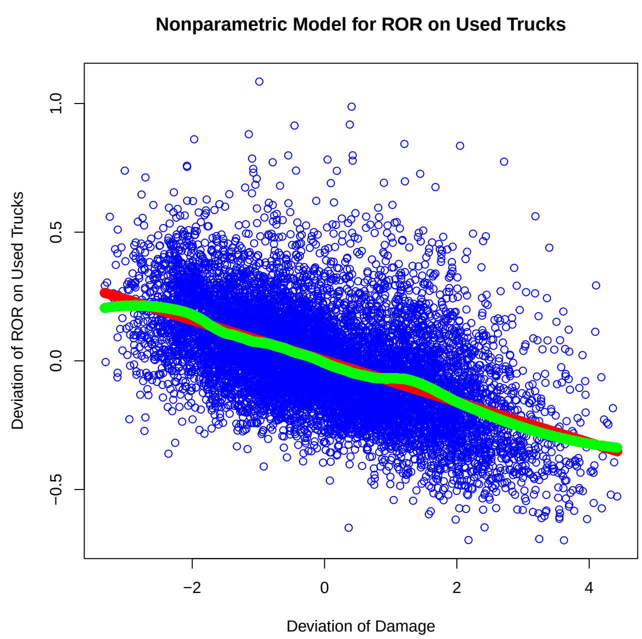


Figure 8.3: Nonparametric Model for ROR on Used Trucks: Excess Damage

8.3 Semiparametric Estimates

As I was building the above nonparametric models, I stored the predictions and will now use them as variables in linear models. Table 8.2 shows the estimates from a set of models. Model 1 is the benchmark linear model in Table 8.1. Model 2 is a semi-parametric model with a nonparametric fit on age substituted in for the age variable. Models 3 and 4 are semi-parametric models with nonparametric fits on mileage and damage, respectively. Model 5 is a maximally semiparametric model, with nonparametric fits for all continuous variables. For each of the single-variable semiparametric models, the coefficients are near one and the fits are similar to the linear model. Even with maximal flexibility, the fit of Model 5 is actually worse than the benchmark linear model. Across all models, the adjusted $\bar{R}^2$ values are all hovering around 0.45. All things considered, these are comparable models and the linear model is sufficient.

	Model 1	Model 2	Model 3	Model 4	Model 5
(Intercept)	1.8148***	1.8049***	1.7686***	1.4954***	1.2196***
	(0.0143)	(0.0142)	(0.0131)	(0.0125)	(0.0097)
auction	−0.1292***	−0.1303***	−0.0687***	−0.0175	−0.0244***
	(0.0126)	(0.0126)	(0.0107)	(0.0103)	(0.0042)
dealer2	−0.0576***	−0.0570***	−0.0558***	−0.0360***	−0.0068
	(0.0103)	(0.0103)	(0.0103)	(0.0103)	(0.0115)
dealer3	0.0734***	0.0716***	0.0741***	0.0346***	−0.0208*
	(0.0078)	(0.0078)	(0.0078)	(0.0077)	(0.0085)
dealer4	−0.0684***	−0.0678***	−0.0678***	−0.0425***	−0.0052
	(0.0072)	(0.0072)	(0.0072)	(0.0071)	(0.0080)
dealer5	0.0663***	0.0658***	0.0637***	0.0429***	0.0057
	(0.0098)	(0.0098)	(0.0098)	(0.0097)	(0.0109)
dealer6	0.0187**	0.0184**	0.0186**	0.0104	−0.0069
	(0.0070)	(0.0070)	(0.0070)	(0.0069)	(0.0078)
dealer7	0.0445***	0.0433***	0.0439***	0.0040	−0.0384***
	(0.0086)	(0.0085)	(0.0086)	(0.0085)	(0.0095)
dealer8	−0.0680***	−0.0674***	−0.0675***	−0.0307***	0.0055
	(0.0069)	(0.0069)	(0.0069)	(0.0068)	(0.0076)
dealer9	−0.0960***	−0.0956***	−0.0950***	−0.0720***	−0.0472***
	(0.0099)	(0.0099)	(0.0099)	(0.0098)	(0.0110)
makeDodge	−0.0149*	−0.0150*	−0.0148*	−0.0167*	−0.0287***
	(0.0073)	(0.0073)	(0.0073)	(0.0073)	(0.0081)
makeFord	−0.0119	−0.0115	−0.0109	−0.0047	0.0154
	(0.0075)	(0.0075)	(0.0075)	(0.0074)	(0.0083)
makeGen. Motors	−0.0204**	−0.0196*	−0.0197*	−0.0148	0.0003
	(0.0077)	(0.0077)	(0.0077)	(0.0076)	(0.0085)
makeNissan	−0.0683***	−0.0676***	−0.0667***	−0.0580***	−0.0391***
	(0.0079)	(0.0079)	(0.0079)	(0.0079)	(0.0088)
makeOthers	−0.0853***	−0.0832***	−0.0808***	−0.0858***	−0.0104
	(0.0123)	(0.0122)	(0.0122)	(0.0122)	(0.0135)
makeSubaru	−0.0655***	−0.0640***	−0.0610***	−0.0494***	0.0181
	(0.0113)	(0.0113)	(0.0113)	(0.0112)	(0.0125)
makeToyota	−0.0617***	−0.0611***	−0.0608***	−0.0477***	−0.0285***
	(0.0075)	(0.0075)	(0.0075)	(0.0074)	(0.0083)
damage_num	−0.1066***	−0.1068***	−0.1117***		
	(0.0024)	(0.0024)	(0.0024)		
mileage_K	−0.0019***	−0.0022***		−0.0031***	
	(0.0002)	(0.0001)		(0.0002)	
age	−0.0060***		−0.0181***	−0.0047**	
	(0.0015)		(0.0008)	(0.0015)	
auction:damage_num	0.0355***	0.0357***	0.0403***		
	(0.0028)	(0.0028)	(0.0026)		
auction:mileage_K	0.0011***	0.0011***		0.0007***	
	(0.0001)	(0.0001)		(0.0001)	
age_np		1.1187***			3.5782***
		(0.1865)			(0.2203)
mileage_np_ret			1.0144***		1.9099***
			(0.1107)		(0.1266)
mileage_np_auc			1.0422***		3.0366***
			(0.1360)		(0.1608)
damage_np_ret				1.0085***	1.0543***
				(0.0224)	(0.0253)
damage_np_auc				0.9956***	1.0986***
				(0.0191)	(0.0220)
R^2	0.4531	0.4542	0.4531	0.4613	0.3237
Adj. R^2	0.4519	0.4530	0.4519	0.4601	0.3223
Num. obs.	9861	9861	9861	9861	9861

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 8.2: Semiparametric Models for ROP on Used Trucks

Chapter 9

Sample Selection Models

I will revisit the recommended linear model from Problem Set #7, which was supported in Problem Sets #8 and #9 by considering other nonlinear specifications within a Generalized Additive Model.

Then I will further investigate this nonlinear relationship by considering the issue of sample selection: car dealers will sell their vehicles in either the auction or retail market, according to which market offers the highest price or rate of return.

9.1 Linear Regression Model

A natural staring point is the recommended linear model from Problem Set #7.

9.1.1 Linear Models by Sale Type

Last week I investigated whether the functional form should include different specifications by the market in which the vehicle was sold. The model included the continuous variables age, mileage, and damage, as well as categorical variables for the dealer and make of vehicle. In addition to the indicator for the auction or retail market, the model included an interaction between the the auction market indicator and the continuous variables mileage and damage. The dependent variable was chosen as the rate of return, since the results were similar to those from the model with the optimal Box-Cox transformation, without the added complexity. The results of this regression specification are shown in Table 9.1.

Next, I will attempt to improve on this specification, using Tobit models for sample selection.

	Full Sample	Auction	Retail
(Intercept)	1.8148^{***}	1.6521^{***}	1.6521^{***}
	(0.0143)	(0.0106)	(0.0106)
auction	-0.1292^{***}		
	(0.0126)		
dealer2	-0.0576^{***}	-0.0527^{***}	-0.0527^{***}
	(0.0103)	(0.0108)	(0.0108)
dealer3	0.0734^{***}	0.0807^{***}	0.0807^{***}
	(0.0078)	(0.0081)	(0.0081)
dealer4	-0.0684^{***}	-0.0592^{***}	-0.0592^{***}
	(0.0072)	(0.0075)	(0.0075)
dealer5	0.0663^{***}	0.0660^{***}	0.0660^{***}
	(0.0098)	(0.0102)	(0.0102)
dealer6	0.0187^{**}	0.0178^{*}	0.0178^{*}
	(0.0070)	(0.0073)	(0.0073)
dealer7	0.0445^{***}	0.0378^{***}	0.0378^{***}
	(0.0086)	(0.0089)	(0.0089)
dealer8	-0.0680^{***}	-0.0577^{***}	-0.0577^{***}
	(0.0069)	(0.0071)	(0.0071)
dealer9	-0.0960^{***}	-0.0861^{***}	-0.0861^{***}
	(0.0099)	(0.0103)	(0.0103)
makeDodge	-0.0149^{*}	-0.0132	-0.0132
	(0.0073)	(0.0076)	(0.0076)
makeFord	-0.0119	-0.0101	-0.0101
	(0.0075)	(0.0078)	(0.0078)
makeGen. Motors	-0.0204^{**}	-0.0211^{**}	-0.0211^{**}
	(0.0077)	(0.0080)	(0.0080)
makeNissan	-0.0683^{***}	-0.0634^{***}	-0.0634^{***}
	(0.0079)	(0.0082)	(0.0082)
makeOthers	-0.0853^{***}	-0.0806^{***}	-0.0806^{***}
	(0.0123)	(0.0127)	(0.0127)
makeSubaru	-0.0655^{***}	-0.0565^{***}	-0.0565^{***}
	(0.0113)	(0.0117)	(0.0117)
makeToyota	-0.0617^{***}	-0.0553^{***}	-0.0553^{***}
	(0.0075)	(0.0078)	(0.0078)
damage_num	-0.1066^{***}	-0.0723^{***}	-0.0723^{***}
	(0.0024)	(0.0012)	(0.0012)
mileage_K	-0.0019^{***}	-0.0007^{***}	-0.0007^{***}
	(0.0002)	(0.0001)	(0.0001)
age	-0.0060^{***}	-0.0065^{***}	-0.0065^{***}
	(0.0015)	(0.0015)	(0.0015)
auction:damage_num	0.0355^{***}		
	(0.0028)		
auction:mileage_K	0.0011^{***}		
	(0.0001)		
R^2	0.4531	0.4086	0.4086
Adj. R^2	0.4519	0.4075	0.4075
Num. obs.	9861	9861	9861

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 9.1: Linear Regression Models: ROR for Auction vs. Retail Sales

9.2 Sample Selection

9.2.1 Predicting the Selection into Samples

The specification in Model 1 Table 9.1 assumes a linear functional form for the relationship between characteristics of used trucks and the rate of return, without selecting into samples by type of market, but including only some interaction variables. To investigate this relationship further, consider the set of variables that are related to whether or not a dealer decides to sell the used trucks in the retail or auction market, with the characteristics observed in the dataset.

Table 9.2 shows the estimates for a probit model to predict the selection into samples by type of sale. Model 1 in Table 9.2 shows a preliminary probit model to predict the selection indicator, with all the other explanatory variables in the model. Damaged cars are more likely to go to auction, as well as cars with high mileage. Dealers vary in terms of how many cars they sell at auction. The age and make of vehicles did not seem as important, given the other variables. Model 2 shows the result of a variable-reduction exercise to eliminate variables that are not statistically significant. These estimates provide a concise but useful model to indicate the types of cars that dealers would prefer to sell at auction.

This model is used to specify the selection equation of the sample selection estimates discussed next.

	Full Model	Reduced Model
(Intercept)	-1.4894^{***}	-1.4477^{***}
	(0.0900)	(0.0709)
dealer2	0.2309^{**}	0.2232^{**}
	(0.0824)	(0.0822)
dealer3	0.2177^{**}	0.2157^{**}
	(0.0662)	(0.0660)
dealer4	0.2881^{***}	0.2848^{***}
	(0.0575)	(0.0573)
dealer5	-0.0181	-0.0206
	(0.0807)	(0.0805)
dealer6	-0.0011	-0.0019
	(0.0562)	(0.0561)
dealer7	-0.1792^{*}	-0.1720^{*}
	(0.0696)	(0.0694)
dealer8	0.2202^{***}	0.2171^{***}
	(0.0547)	(0.0545)
dealer9	0.2629^{***}	0.2528^{**}
	(0.0795)	(0.0793)
makeDodge	-0.0251	
	(0.0606)	
makeFord	-0.0393	
	(0.0612)	
makeGen. Motors	-0.0346	
	(0.0627)	
makeNissan	0.1167	
	(0.0652)	
makeOthers	-0.0266	
	(0.0955)	
makeSubaru	0.0384	
	(0.0890)	
makeToyota	0.1321^{*}	
	(0.0617)	
damage_num	0.2562^{***}	0.2543^{***}
	(0.0109)	(0.0109)
mileage_K	0.0089^{***}	0.0091^{***}
	(0.0010)	(0.0005)
age	0.0045	
	(0.0126)	
AIC	10665.0246	10674.8535
BIC	10801.7551	10754.0132
Log Likelihood	-5313.5123	-5326.4267
Deviance	10627.0246	10652.8535
Num. obs.	9861	9861

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 9.2: Probit Models for Type of Sale

9.2.2 Estimating a Sample Selection Model

Table 9.3 shows the estimates from a model that accounts for sample selection. The models are estimates from the Tobit model of type 5, which is a model specification that allows for switching of the observations in the sample into two models: one for the rate of return on used trucks sold at auction and one for those sold on the retail lot.

Each column shows the estimates from a separate model and the pair of models is the result of a downward selection procedure in which a statistically insignificant variable was removed from Model 1. In each model, the estimates are grouped into three categories. The first block of coefficients describe the selection model to determine whether a used truck would be sold at auction. These coefficients are denoted by the prefix"S:". Below these lies two blocks of coefficients for the observation equations. The notation "O: ¡name of variable¿ (i)" indicates the coefficient for the particular variable in the observation equation for sample i. In this specification, the first observation equation represents trucks sold on the retail lot (`auction == 0`), while equation 2 represents the trucks sold at auction (`auction == 1`).

Model 1 shows the estimates from the full model. All of the coefficients are significant for the selection equation and the observation equation for cars sold on the retail lot. Only mileage is insignificant in the observation equation for auctions. The goal will be to obtain a final model that has well-defined standard errors for all variables and, ideally, all coefficients statistically significant.

Model 2 shows the estimates from a reduced model, after eliminating `mileage_K` from the auction equation. Now all coefficients are significant in all three equations. This is a luxury of having a large sample size and a dataset that conforms to the model specification. I leave it to the reader to compare the coefficients from the selection model, in Table 9.3, with those from the linear model in Table 9.1.[1]

[1] As an aside, with so many coefficients in the regression models, the tables often extend out of the page under the default settings. For several of the tables in this note, I have used the texreg argument `fontsize = 'small'`. For Table 9.3, I have used an even smaller front by setting `fontsize = 'footnotesize'` and placed the standard errors on the same line as the coefficients by adding the argument `single.row = TRUE`. See the help file for texreg for more information, as well as the R vignette for the texreg package, in a file called `texreg.pdf`. Search for this keyword and you should find a document with the title *texreg: Conversion of Statistical Model Output in R to LaTeXand HTML Tables* by the package creator, Philip Leifeld.

	Model 1	Model 2
S: (Intercept)	$-1.31492\ (0.06207)^{***}$	$-1.31775\ (0.06194)^{***}$
S: dealer2	$0.18807\ (0.07444)^{*}$	$0.18814\ (0.07443)^{*}$
S: dealer3	$0.20976\ (0.05889)^{***}$	$0.20891\ (0.05887)^{***}$
S: dealer4	$0.23663\ (0.05217)^{***}$	$0.23693\ (0.05216)^{***}$
S: dealer5	$-0.02198\ (0.07212)$	$-0.02300\ (0.07210)$
S: dealer6	$0.00717\ (0.05092)$	$0.00663\ (0.05091)$
S: dealer7	$-0.16815\ (0.06261)^{**}$	$-0.16967\ (0.06257)^{**}$
S: dealer8	$0.18541\ (0.04967)^{***}$	$0.18562\ (0.04967)^{***}$
S: dealer9	$0.22978\ (0.07190)^{**}$	$0.22930\ (0.07189)^{**}$
S: damage_num	$0.23468\ (0.00969)^{***}$	$0.23321\ (0.00952)^{***}$
S: mileage_K	$0.00846\ (0.00045)^{***}$	$0.00857\ (0.00043)^{***}$
O: (Intercept) (1)	$1.84560\ (0.01507)^{***}$	$1.84588\ (0.01506)^{***}$
O: dealer2 (1)	$-0.09666\ (0.01465)^{***}$	$-0.09666\ (0.01464)^{***}$
O: dealer3 (1)	$0.04671\ (0.01226)^{***}$	$0.04681\ (0.01225)^{***}$
O: dealer4 (1)	$-0.08826\ (0.01008)^{***}$	$-0.08827\ (0.01007)^{***}$
O: dealer5 (1)	$0.09488\ (0.01425)^{***}$	$0.09498\ (0.01425)^{***}$
O: dealer6 (1)	$0.03177\ (0.00983)^{**}$	$0.03182\ (0.00983)^{**}$
O: dealer7 (1)	$0.08193\ (0.01216)^{***}$	$0.08208\ (0.01215)^{***}$
O: dealer8 (1)	$-0.09272\ (0.00950)^{***}$	$-0.09273\ (0.00950)^{***}$
O: dealer9 (1)	$-0.11305\ (0.01400)^{***}$	$-0.11299\ (0.01400)^{***}$
O: makeDodge (1)	$-0.02206\ (0.00940)^{*}$	$-0.02206\ (0.00940)^{*}$
O: makeFord (1)	$-0.01846\ (0.00943)$	$-0.01846\ (0.00943)$
O: makeGen. Motors (1)	$-0.01236\ (0.00966)$	$-0.01236\ (0.00966)$
O: makeNissan (1)	$-0.07772\ (0.01021)^{***}$	$-0.07772\ (0.01021)^{***}$
O: makeOthers (1)	$-0.08460\ (0.01412)^{***}$	$-0.08461\ (0.01412)^{***}$
O: makeSubaru (1)	$-0.06546\ (0.01332)^{***}$	$-0.06547\ (0.01332)^{***}$
O: makeToyota (1)	$-0.07072\ (0.00967)^{***}$	$-0.07072\ (0.00967)^{***}$
O: damage_num (1)	$-0.13096\ (0.00250)^{***}$	$-0.13081\ (0.00249)^{***}$
O: mileage_K (1)	$-0.00286\ (0.00017)^{***}$	
O: age (1)	$-0.00525\ (0.00194)^{**}$	$-0.00525\ (0.00194)^{**}$
O: (Intercept) (2)	$1.40767\ (0.01466)^{***}$	$1.40741\ (0.01467)^{***}$
O: dealer2 (2)	$-0.02510\ (0.01575)$	$-0.02488\ (0.01575)$
O: dealer3 (2)	$0.09131\ (0.01164)^{***}$	$0.09119\ (0.01164)^{***}$
O: dealer4 (2)	$-0.03604\ (0.01105)^{**}$	$-0.03586\ (0.01105)^{**}$
O: dealer5 (2)	$0.05096\ (0.01486)^{***}$	$0.05073\ (0.01486)^{***}$
O: dealer6 (2)	$0.01349\ (0.01072)$	$0.01342\ (0.01072)$
O: dealer7 (2)	$0.01461\ (0.01308)$	$0.01463\ (0.01309)$
O: dealer8 (2)	$-0.04045\ (0.01058)^{***}$	$-0.04034\ (0.01059)^{***}$
O: dealer9 (2)	$-0.06902\ (0.01512)^{***}$	$-0.06889\ (0.01512)^{***}$
O: makeDodge (2)	$-0.00853\ (0.00841)$	$-0.00856\ (0.00841)$
O: makeFord (2)	$-0.01082\ (0.00861)$	$-0.01076\ (0.00861)$
O: makeGen. Motors (2)	$-0.01707\ (0.00880)$	$-0.01705\ (0.00880)$
O: makeNissan (2)	$-0.05993\ (0.00899)^{***}$	$-0.05983\ (0.00899)^{***}$
O: makeOthers (2)	$-0.08232\ (0.01439)^{***}$	$-0.08201\ (0.01438)^{***}$
O: makeSubaru (2)	$-0.06210\ (0.01297)^{***}$	$-0.06187\ (0.01297)^{***}$
O: makeToyota (2)	$-0.05417\ (0.00850)^{***}$	$-0.05410\ (0.00850)^{***}$
O: damage_num (2)	$-0.05090\ (0.00179)^{***}$	$-0.05130\ (0.00172)^{***}$
O: mileage_K (2)	$-0.00011\ (0.00014)$	
O: age (2)	$-0.00643\ (0.00176)^{***}$	$-0.00757\ (0.00104)^{***}$
sigma1	$0.16656\ (0.00542)^{***}$	$0.16654\ (0.00542)^{***}$
sigma2	$0.23968\ (0.00255)^{***}$	$0.23982\ (0.00254)^{***}$
rho1	$-0.82670\ (0.02572)^{***}$	$-0.82665\ (0.02572)^{***}$
rho2	$0.94414\ (0.00450)^{***}$	$0.94439\ (0.00447)^{***}$
O: mileage_K		$-0.00287\ (0.00017)^{***}$
AIC	2411.22894	2409.86618
BIC	2792.63511	2784.07601
Log Likelihood	-1152.61447	-1152.93309
Num. obs.	9861	9861
Censored	3077	3077
Observed	6784	6784

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 9.3: Selection Models for ROR on Used Trucks