AUTUMN 2003

IDENTIFYING FACTORS THAT DETERMINE THE PRICE OF USED CARS

TOM JAFARI
tdjafari@u.washington.edu
Economics 482
Professor Startz

EXECUTIVE SUMMARY

This paper explores the issue of determining and analyzing the factors associated with the pricing of used cars in order to give potential consumers greater insight in their search for a reasonably priced means of transportation. These factors include the type of transmission, location of sale, mileage, and the existence of power features. Running a regression against a data set that was limited to two and four door cars commonly driven by college students, the results show a positive correlation between a car having power features and being sold at a dealership with the price. Having an automatic transmission and a mileage that was above the blue book average, on the other hand, lowered the price of used cars. The existence of large standard errors from the regression yields large confidence intervals for some variables thus limiting the ability to quantify the degree to which the variables affect price. In terms of general behavior, though, the analysis does an effective job and clearly provides consumers in the market for a used car with useful information.

INTRODUCTION

When I began to look for a topic for my paper, I knew I wanted to explore something meaningful. And by meaningful, I don't mean something that would provide far reaching economic implications; being limited by the confines of a quarter requires a small, focused topic that can be reasonably answered within a relatively short time frame. I wanted to answer a question that had meaning to *me*. Having decided this, I was forced to determine what, if anything, was meaningful to a twenty-two year old senior in college. Being a senior, I've found that many of my activities are occurring outside the university community (job interviews, part-time work, social activities, etc.). Consequently, the need for a solid means of transportation has become increasingly important. Being blessed (or cursed) with an economic mindset, I wanted to make sure that I got the best price for my car, and, more importantly, that the price of the used car accurately reflected its quality.

Project Statement

This paper seeks to determine, through regression analysis, the factors that affect the pricing of used automobiles and their relative magnitudes to each other in order to give potential consumers greater insight in their search for a reasonably priced means of transportation.

MODEL AND METHODS

Anyone who's gone shopping for a used car has quickly discovered that there can be large variations in price. Obviously, there are a multitude of factors that go into the pricing of a used car. From my own research and experience, I believe that mileage, the existence of power features, whether the car has manual or automatic transmission, and whether or not it

is sold at a dealership all play a major role when determining what the price of a used car should be. The model used for regression analysis, then, becomes:

Price = B1(blue book value) + B2(automatic or manual) + B3(dealership or private party) + B4(power features) + B5(mileage – 12,000(2003 – year))

By incorporating the blue book value into the regression model, variations in the brand and year of a car will be taken into account. Moreover, the blue book value assumes a car that is in good condition, has a manual transmission, is sold by a private party, has no power features, and accrues an average of 12,000 miles per year. By doing so, the model should then gauge the significance of the other variables in explaining price variation.

The sheer size of the car market and large number of different models and brands required that several simplifications be made to get a reasonably representative data set. As I am interested in using this model to explain the pricing of cars that I, a college student, would be interested in, the observations were limited to two and four door cars. In other words, trucks and minivans were excluded from the data set. For both the two and four door groups, moreover, I examined the six most popular brands and models for college students. Determining the most popular brands and models for college students was accomplished by a small survey of my friends and acquaintances. The data set was then collected from usedcars.com, which is an internet forum where both dealerships and private parties bring their used car to market.

RESULTS AND ANALYSIS

From the regression (see Appendix I), it appears that every dollar increase in the blue book value results in a $1.18 increase in the market value for a used car. Intuitively, the relation between the blue book value and the market price of a used car would be such that a dollar increase in the blue book value would result in a dollar increase in the price of the used car by the definition of the blue book value. Variations in the other factors, then, would cause the fluctuation in price. In the model, the marginal value for the blue book price is relatively close to a one with a relatively small standard error, so it is acceptable for the purposes of this paper.

Buying a used car from a dealership, furthermore, adds around $70 to the price tag. This increase in price could be reflected in the increased confidence consumers may have in purchasing a quality car. With private parties, consumers often have little to no prior knowledge of the seller or of the condition of the car. At a used car lot, consumers are purchasing from a company that specializes in selling used cars. Consequently, they are able to give a more accurate assessment of the condition of the vehicle. Of course there is something to be said about the fairness and quality of information when dealing with a used car salesman, but for most intents and purposes, used car lots are more reliable and readily accessible to the consumer than private parties, which are reflected in its effect on the price of the used car.

Similarly, the existence of power features (defined here as power steering, braking, air conditioning, etc.) adds $358.47 to the price of the car. Having mileage above the blue book average resulted in decreasing the price by $.02 per extra mile above the average, which can be quite significant for cars that have very high (or low) mileages. A used car that has a

mileage that is 10,000 miles above average, for instance, would result in a $200 dollar decrease in price. This makes intuitive sense as increased mileage can be associated with greater wear and tear on the vehicle and the potential for costly mechanical problems. Lastly, buying a car that has an automatic transmission decreases the price of the car $476.17. In general, however, a car that has manual transmission is cheaper than having a car that is an automatic; so the results for this variable are somewhat disheartening.

STRENGTHS/WEAKNESSES

Although the regression yields results that are similar to expectations, the quality of the model is still subject to question. The first factor that becomes readily apparent is the size of the standard errors for the coefficients, which range from $300-450 for variables concerning transmission, power features, and location of sale. A 95% confidence interval for the effect on price when the car was sold at a dealership gives a range of (-590.57, 728.31). That is, we can conclude that the effect on price of selling a car from a dealership will be between a $600 decrease in price and a $730 increase in 95% of the samples we examine. Such a large range doesn't really provide much useful information since the effect on price can be both positive and negative. This suggests that more data points need to be gathered to tighten the interval and provide a more meaningful analysis. Secondly, the R-squared value of .7743 shows that the regression did a moderate job by explaining 77% of the price variation. As a result, there may still be factors that are unaccounted for in this model that play a significant role in determining the price of a used car.

CONCLUSION

The moderate success of the model suggests that several improvements can be made. First, there is the possibility to delve deeper into the effects of the variables on price variation between different types of cars—namely coupes and sedans (which, for all intents and purposes, are the most popular types of two and four door cars). It may prove interesting to run regressions against only coupes, then only sedans, and finally both to determine if there are any significant differences between them. Such an analysis may discover that two-door cars are more price sensitive to being sold in a dealership. Or perhaps four-door cars are more likely to have power features and thus have a lower influence on the price of the car as a result. Lastly, the broad definition of power features as a variable could be more descriptive to determine what types of power features have the largest effect on the price of the car.

But back to the real question at hand: Does this regression produce any useful results? If someone was looking for a specific dollar value associated with buying from a dealership versus a private party, for instance, the significance of the standard errors suggest that the regression may not be that useful. In terms of identifying general behavior, however, the regression effectively shows the potential effect (whether that be positive or negative) and magnitude of each variable on price. From the regression, moreover, we can see that that mileage, power features, and transmission are the most significant factors in determining the price of a used car. So in terms of supplying consumers with greater insight on identifying and measuring the factors that affect the pricing of used cars, the regression, although suffering from a limited data set, was nonetheless a modest success.

Appendix I: Regression Analysis of Used Car Pricing Model (from EViews)

Dependent Variable: PRICE
Method: Least Squares
Date: 12/05/03 Time: 12:48
Sample: 1 59
Included observations: 59

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	895.0324	454.4527	1.969473	0.0541
BBVAL	1.179891	0.096615	12.21235	0.0000
AUTO	-476.1697	331.4747	-1.436519	0.1567
DEALER	68.87011	336.4512	0.204696	0.8386
POWER	358.4650	412.0698	0.869913	0.3883
MILEAGE	-0.019212	0.004013	-4.787356	0.0000

R-squared	0.779074	Mean dependent var	4337.898
Adjusted R-squared	0.758232	S.D. dependent var	2540.459
S.E. of regression	1249.141	Akaike info criterion	17.19444
Sum squared resid	82698739	Schwarz criterion	17.40572
Log likelihood	-501.2361	F-statistic	37.37989
Durbin-Watson stat	2.280721	Prob(F-statistic)	0.000000