DOMESTIC VERSUS JAPANESE AUTOMOBILES

by

Frank Sirounis

B.A., University of Saskatchewan, 2000

PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS

In the Department of
Economics

SIMON FRASER UNIVERSITY

March 2003

APPROVAL

Name: **Frank Sirounis**

Degree: **M. A. (Economics)**

Title of Project: **Domestic Versus Japanese Automobiles**

Examining Committee:

Chair: **Steeve Mongrain**

Ken Kasa
Senior Supervisor

Brian Krauth
Supervisor

Clyde Reed
Internal Examiner

Date Approved: Friday March 28, 2003

Title of Project *Domestic Versus Japanese Automobiles*

Author: ______________________

Frank *Sirounis*

ABSTRACT

This paper examines the Canadian used car market in 2001 and looks at the difference between domestic and Japanese automobiles in terms of resale values. A cross section of vehicles ranging in age from 1987 to 1998 are looked at controlling for vehicle reliability, options, year of production, car class and wear and tear. The research shows that Japanese vehicles hold their value by a margin of four percent better than domestic brand vehicles.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

I. INTRODUCTION:

Some people swear by domestic automobiles while others are convinced that Japanese makes are the way to go. The purpose of this project is to attempt to see what the difference is between domestic and Japanese vehicles in terms of used car prices. Specifically, the question will be: "Do Japanese vehicles hold their value better than a domestic vehicle in Canada?" The vehicle categories looked at are small cars, medium sized cars, luxury, sports and minivans (i.e., there are no trucks, motorcycles nor commercial vehicles).

Theoretically, how much a car holds of its original value is a function of quality. That is to say, a car that is more reliable on average and needs fewer expensive repairs will command a greater share of its original price compared to other cars.

II. THE DATA:

The cross-sectional data for this project were obtained exclusively from Phil Edmonston's Lemon-Aid Guide for Used Cars 2001. Edmonston describes how he gathered the data for this book in a fairly rigorous way without giving specifics as to the exact formula for determining a used car's value. Despite some murkiness, Edmonston has put a lot of detail, research, and consistency in this book. Furthermore, the book is widely used in Canada with almost one million copies sold. This book and the data in it has proved to be a valuable tool when shopping for a used vehicle and thus it may be considered a legitimate source of data.

Each class of the variables used are briefly summarized in the following table. A more detailed discussion of each variable comes after the table.

Table 1: Variable Definitions

Variable	Definition
used price / new price	used price to new price ratio
JPN	=1 if Japanese or =0 if domestic
Y87 to Y98	dummy for each year from 1987 to 1998
rel1 to rel5	Dummy for each reliability category from 1 to 5
RELIABLITY	a variable that has five integer values (1 to 5) for the reliability rating of the car
SMALL, MEDIUM, LUXURY, SPORTS, MINIVAN	Dummy for each class of car

What follows now is detailed descriptions of each variable and the data that comprises them.

Used prices:

Used car prices are as of March 2000 for standard models. The following is a list of assumptions and points to be made about the used cars and their prices. First of all it is assumed that the vehicle is in good condition. That means there is no rust, tires have a 50% or better tread life left, brakes have 50% or better life left, all components are in good working order and so on. Second, there is a maximum of 20,000 km for each calendar year. Third, base or standard models are used for this study. The Lemon-Aid guide provides prices for all trim levels of vehicles, but only the base models were chosen to control for options such as air conditioning, power seat, power windows and so forth. This is discussed further in the 'Year and Car Classes' section.

Finally, region is controlled for. The prices are for Quebec and Ontario markets. Eastern Canadians should add ten percent, while Western Canadians should add 15 to 20 percent on the listed prices. Edmonston says the reasons for this is "less competition and inflated new-vehicle prices in these regions". There may be some truth to Edmonston's views, however, he should also consider the heavy salting of Quebec and Ontario roads compared to other regions. Also on average Quebec and Ontario cars are subject to more stop and go traffic and traffic jams than the other provinces which is also not very car friendly.

The following is an excerpt from Edmonston's guide discussing how he obtained the data:

> Why are *Lemon-Aid*'s prices lower than the prices found in dealer guides? The answer is simple: dealer guides inflate their prices knowing that you will bargain the price down, so you'll be convinced you made a great deal even if you didn't.
>
> I used newspaper classified ads from Quebec, Ontario, and B.C., as well as auction reports, for my used values. I then check these figures with the *Red Book* and *Black Book*; I don't start with the *Red Book*'s retail or wholesale figures (like real estate listings, vehicle prices are inflated about 10 percent for wholesale/private sales and almost 20 percent for retail/dealer sales – compare the two and you'll see what I mean). I then project what the value will be by year's end, and that further lowers my prices. I almost always fall way under the *Red Book*'s value, but not far under the *Black Book*'s price.

There is a top and bottom used price in the book that Edmonston says is a range that one uses for negotiation. Only the top values were used in the regressions which is reasonable since negotiation for things like degree of wear and tear can be held constant if the upper value of all the cars is used. The CPI was used to put both new and used car values into 1992 dollars.

New prices:

The manufacturer's suggested retail price (MSRP) is given as the new car price. This is the price that is advertised by makers and serves as a good reference point from which to compare used prices.

A minor sticking point is that warranty information is not given with these new prices. Chrysler boasts about how many cheaply made minivans they have out on the road which some say is only due to the generous standard warranty package of about 7 years/115,000 km (whichever comes first), while the industry standard lies somewhere around 3 years/60,000 km. However, who pays for the warranty but the consumer. The cost of the warranty is built into the car price and so it is reasonable to leave the warranty information out.

Japanese:

This variable is relatively self-explanatory except there is a twist. What do a Chrysler Colt, a Ford Probe, and a GM Metro have in common? They are all Japanese vehicles built for the North American brands. The engines, transmissions, suspensions and major electrical components are all sheathed under a North American nameplate. Since these vehicles are essentially Japanese made, they are coded as Japanese vehicles. Similarly, Mazda's pickup truck is essentially a Ford Ranger under Mazda's nameplate.

The previous point aside, Domestic vehicles are those made by the "Big Three" North American automobile manufacturers: Chrysler, Ford and General Motors which are companies that are all based in the United States. Japanese vehicles are all makes that have their company roots and headquarters in Japan such as Honda, Nissan, and Toyota. Some Japanese manufacturers have plants in Canada and the United States, but the automobiles they produce are still coded as Japanese as the headquarters of those manufacturers are in Japan.

Reliability:

A discrete scale from 1 to 5 where 5 is the highest reliability rating for a vehicle. The rating is relative to all cars in every class and year of the sample. For example, one will rarely find a one or two year old car with a reliability rating of 1 since even a cheap car that is new will fare well against the others. Similarly, an initially high quality car with a reliability rating of 5 will eventually go down to a 1 given it is old enough. The following tables shows average reliability ratings for each car class, selected years, and whether the car is a domestic or Japanese vehicle.

Table 2: Small Cars:

	1998	1997	1996	1995
Japanese	4.6	4.4	4.3	4
Domestic	2.6	2.6	2	1.7
Overall	4.1	4	3.8	3.4

Table 3: Medium Cars

	1998	1997	1996	1995
Japanese	4.8	4.5	4.3	4
Domestic	3	2.7	2.3	2
Overall	3.9	3.6	3.5	2.9

Table 4: Luxury Cars:

	1998	1997	1996	1995
Japanese	4.8	4.7	4.8	4.5
Domestic	3.5	3.4	3.1	2.7
Overall	4.2	4.1	4.1	3.7

Table 5: Sports Cars:

	1998	1997	1996	1995
Japanese	5	4.8	4.6	4.1
Domestic	4	3.3	3.3	2.5
Overall	4.6	4.2	4.1	3.5

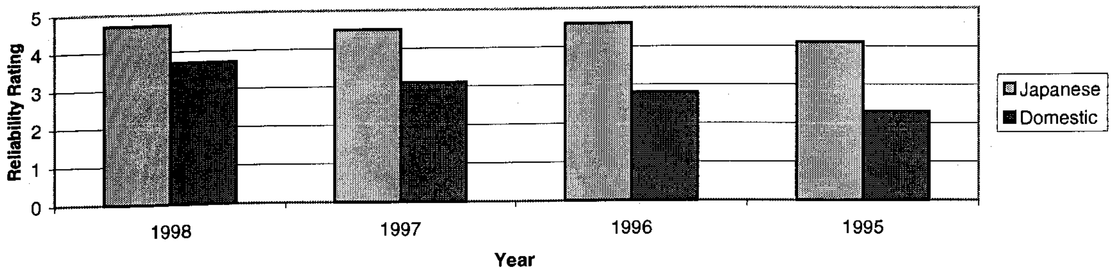
Table 6: Minivan Cars:

	1998	1997	1996	1995
Japanese	4.3	4	4.3	3.3
Domestic	3.5	3	2.8	2.2
Overall	3.8	3.6	3.2	2.6

Table 7: All Cars:

	1998	1997	1996	1995
Japanese	4.7	4.5	4.6	4.1
Domestic	3.7	3.1	2.8	2.3
Overall	4.1	4	3.7	3.3

Figure 1: Average Reliability Ratings for Domestic and Japanese Automobiles



The following legend shows a vehicle's relative degree of overall reliability as well as which mechanical and body parts are subject to premature failure:

Table 8: Reliability Numbers

Number	Rating
1	Unacceptable
2	Below Average
3	Average
4	Above Average
5	Excellent

It is assumed that these overall reliability ratings are generated from compiling the ratings on specific car components (e.g. engine, electrical system etc) and giving them some sort of weighted average. A compromise was giving up the idea of having the average maintenance cost per year of a car. Instead, I used the "reliability rating" as a proxy for the cost to maintain the vehicle because the data on maintenance costs are few and far between. As it stands, it is assumed that repairs cost the same for all vehicles. However, this may not be too damaging to the results since cars are grouped into different classes. Repairs amongst luxury cars will cost about the same as will repairs among small cars. The reliability data presented in Lemon-Aid was composed of consumers of the actual cars that filled out surveys on their vehicles. Edmonston does not reveal sample sizes for any or all cars and only says that there are "thousands".

Years and Car Classes:

The 2001 edition of Lemon-Aid focuses on the years from 1991 to 1998. Most of the observations fell within the eight year span, but there were a few observations that went back as far as 1987.

Five vehicle classes were considered: small, medium, luxury, sports, and minivan. Incidentally, no Japanese cars fell into the large car class so it was not included in the analysis. Since the question of concern is "Domestic versus Japanese" the data from each class were stacked and dummy variables added to show the difference for each class. Of the 473 vehicles included in the cross sectional data 268 are Japanese.

The following table breaks down the sample by the number of observations for each category per year:

Table 9: Number of Observations per Year and Category

Year	Small	Medium	Luxury	Sports	Minivan	All
1987	0	1	0	0	0	1
1988	1	1	1	0	0	3
1989	1	2	1	0	0	4
1990	1	2	2	0	1	6
1991	10	11	13	8	6	48
1992	12	11	15	9	6	53
1993	12	12	17	9	7	56
1994	12	12	16	8	7	55
1995	12	13	20	11	8	64
1996	12	13	20	11	9	65
1997	12	12	18	10	9	61
1998	11	12	17	9	7	57
Total	96	102	140	75	60	473

The question ("Domestic or Japanese?") was formulated to leave out European, Russian, Korean and any other makes only for simplicity. In addition, the data only pertains to automobiles sold in Canada.

One should also note that Lemon-Aid misses some observations in two ways. Used car prices and or reliability ratings for specific cars in specific years are not presented. By assumption, not enough data was collected on those specific vehicles. Also, some cars were left out of the guide altogether. At a casual glance, omitting vehicles from the data is especially prevalent in the sports car class. For example, the Toyota MR2 and Acura NSX were both missing from the line-up. Obviously, but explicitly, observations that were missing a used car price (i.e. the dependent variable) were left out.

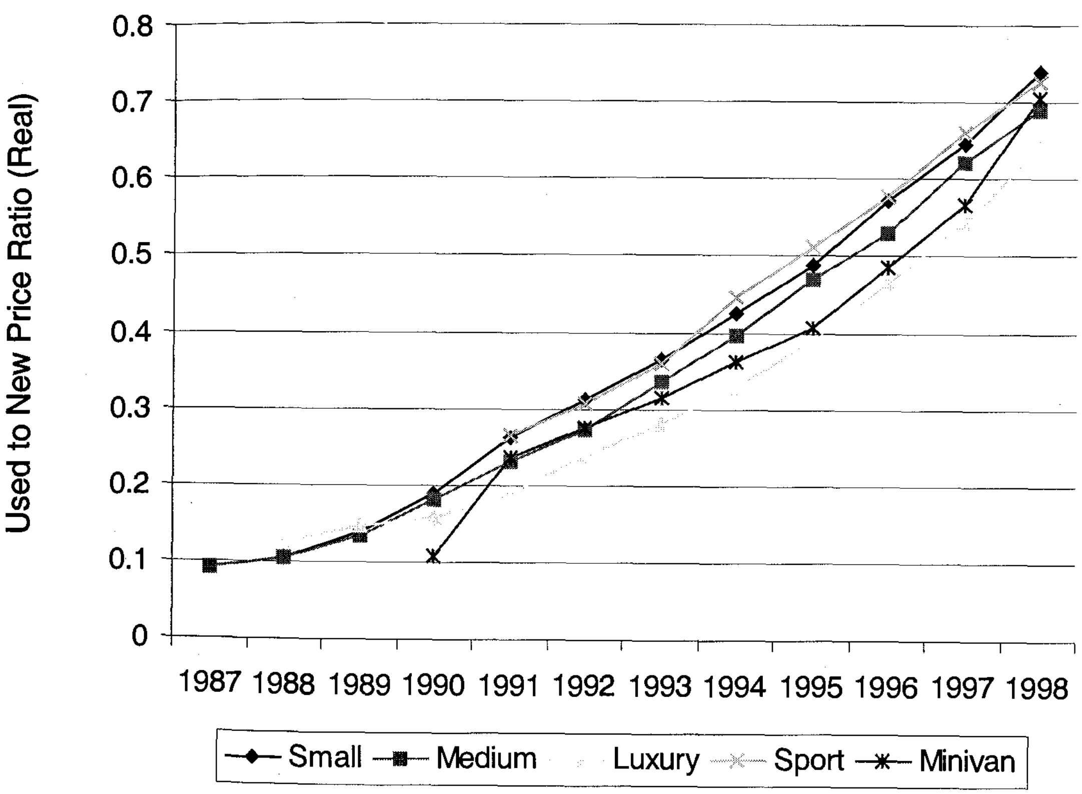
The small, medium, minivan classes seemed to be the most homogenous. These classes were easiest to determine whether the vehicle was a base model or not. Standard options seem to be most consistent across these three classes of automobiles and are assumed to be so in this analysis. The assumption is tall, but also one that may yield useful results in this analysis if it holds true since one can compare apples to apples.

The luxury models posed a problem because there is a lot more variability among the models for what are considered standard options. The previous assumption is a lot harder to fit in this case and may cause some errors in this analysis. The sports car class is also potentially troubling because of what is considered a sports car. For example a base 1991 Toyota Celica and a base 1991 GM Corvette were $16,418 and $42,798 new respectively. It is comparing a mildly sporty car with something approaching an exotic sports car in terms of horsepower, the difference between 150 to 350 horsepower. These types of situations have the ability to cause heteroskedastic errors when the data from all the car classes is stacked. Thus, results from a test for heteroskedasticity is presented later.

The Lemon-Aid guide listing for luxury automobiles poses another problem when it lists the automobiles. The North American cars are put into more detailed subsections for each of the cars, while the Japanese luxury cars are grouped together by make. For example, the Ford sports coupe Lincoln Mark VIII is put in one section while the comparable Lexus sports coupe SC 400 is grouped together with all the rest of the Lexus automobiles. The result is that all the Lexus cars are given the same reliability ratings per year, while North American luxury vehicles are given their own reliability ratings. It seems reasonable that all of the Lexus cars will have the same level of quality and so the reliability numbers are not troubling. What is of concern is that the Japanese brands will not have their cars duly represented in the sample. A sensitivity analysis will be conducted whereby two regressions are run on luxury automobiles alone. The first regression will have the sample the way the book lists the cars and the other regression will have the Japanese cars split up into separate observations for each type of luxury car. The results are presented later.

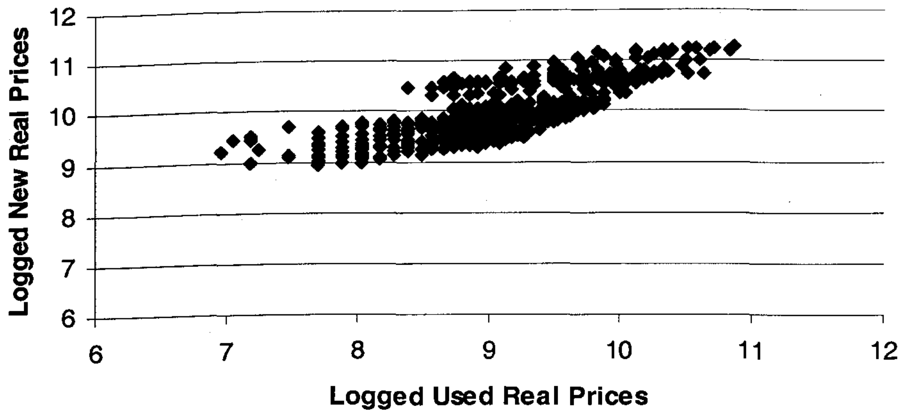
What follows are a couple of graphs that aid in the visual inspection of the data.

Figure 2: Averages of Price Ratio by Year by Category:



The smooth lines in the above graph indicate that there does not seem to be any outliers among the averages of used/new prices per year per car class.

Figure 3: Scatter Plot of Logged New Real Prices to Logged Used Real Prices



The scatter plot does not visually indicate the presence of any outliers. Indeed, no outliers were detected in the data.

III. PROCEDURE AND RESULTS:

To reiterate, the question at hand is whether there is a difference between Japanese cars and Domestic cars in terms of how well they hold their value over time. The exact specification for the regression was formulated keeping this question in mind.

The following general specification is used:

$$\text{used price / new price} = \mu + \nu JPN + \psi_1 Y88 + \psi_2 Y89 + \psi_3 Y90 + \psi_4 Y91 + \psi_5 Y92 + \psi_6 Y93 + \psi_7 Y94 + \psi_8 Y95 + \psi_9 Y96 + \psi_{10} Y97 + \psi_{11} Y98 + o_2 REL2 + o_3 REL3 + o_4 REL4 + o_5 REL5 + \pi_1 MEDIUM + \pi_2 LUXURY + \pi_3 SPORTS + \pi_4 MINIVAN + \varepsilon$$

So, the model is a function of the used/new price ratio regressed on all intercept dummies for whether or not the vehicle is a Japanese make, the year of the car, the class of the car, and an intercept. The base year in this model is

1987. Regressions were run for each of the individual car classes. The output is presented in the Appendix I.

Table 10: Summary Of Regression Results For All Classes Of Automobiles Together:

Dependent Variable: USED_TO_NEW
Method: Least Squares
Included observations: 473

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Prob.
JPN	0.041	0.000
Y88	-0.616	0.000
Y89	-0.581	0.000
Y90	-0.542	0.000
Y91	-0.501	0.000
Y92	-0.453	0.000
Y93	-0.406	0.000
Y94	-0.360	0.000
Y95	-0.290	0.000
Y96	-0.222	0.000
Y97	-0.152	0.001
Y98	-0.060	0.166
REL2	0.028	0.114
REL3	0.047	0.002
REL4	0.030	0.058
REL5	0.051	0.001
MEDIUM	-0.016	0.056
LUXURY	-0.075	0.000
SPORTS	-0.036	0.009
MINIVAN	-0.105	0.000
C	0.706	0.000
Adjusted R^2	0.858426	

(Note: the complete regression output can be found in the Appendix I.)

<u>Interpretation of Results:</u>

Domestic vs Japanese:

For all the cars, the Japanese automobiles held their value 4.1 percentage points better than the domestic brands. So, if a domestic vehicle has a used to new value of 60%, then its Japanese counterpart will have a used to new value of 64.1%. This is the main point of the project and the coefficient is statistically significant.

Reliability Dummy Coefficients:

The numbers on reliability are a little puzzling in that that the dummy on a reliability rating of 3 is given a higher coefficient value than the higher reliability rating of 4.

Table 11: Number of Observations per Reliability Rating

Rating	No. of Obs
Rel 1	23
Rel 2	141
Rel 3	100
Rel 4	102
Rel 5	106

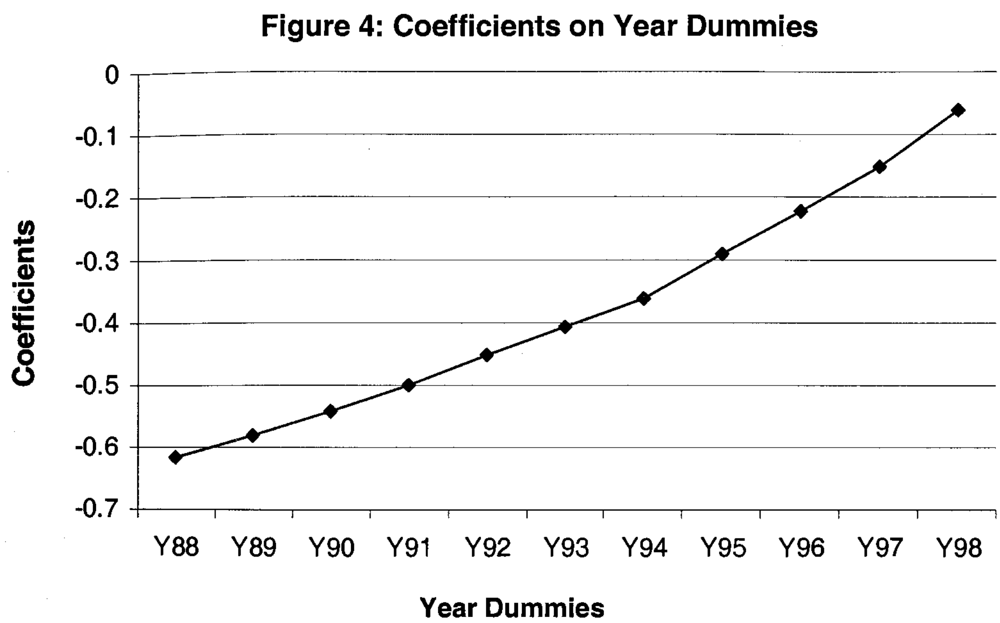
As shown in the table, the number of observations for each reliability rating is almost the same for 'rel 3' and 'rel 4'. Given this, one would have to conclude that there may be some missing information that causes the coefficient on 'rel 3' to be higher than the coefficient on 'rel 4'.

Class Dummy Coefficients:

Recall that the base category for class of cars is 'small'. It seems as if all the other classes of cars do worse than the small cars when it comes to holding their value. As a reminder, the small car class has the lowest initial prices out of any of the classes. This may coincide with economic theory in that having a car is preferred to having no car at all and that the small car gives one more bang for the buck – that is, demand for small cars is more inelastic compared to the demand for the four other car classes. The minivan has the biggest loss compared to all the other vehicles – this may be explained by the overall poor quality of the vehicles in the sample of minivans. The next to lowest coefficient is on luxury automobiles. People buy luxury cars as a status symbol, thus they do not want to buy used, nor cheap brand luxury cars.

Year Dummy Coefficients:

As a car gets older it will hold less of its original value. The dummy coefficients for the years ranging from 1998 back through to 1988 are becoming more and more negative displaying a drop in price as one looks at older and older vehicles.



A Test for Heteroskedasticity:

There may be heteroskedastic errors in the regression because of the luxury and sports car observations. White's test for heteroskedasticity is appropriate here. The null hypothesis of homoskedastic errors can be rejected with a p-value of 0.013, however when using White's heteroskedastic-consistent matrix the results did not change. That is to say that the significance of all the variables, except the year dummy on 1998, did not change and is still consistent with the story.

Table 12: Test for Heteroskedasticity

Car Class	P-Value	Heteroskedasticity
All	0.013	Present
Small	0.168	Not present
Medium	0.498	Not present
Luxury	0.319	Not present
Sport	0.001	Present
Minivan	0.396	Not present

Comparison of individual car class regressions:

The following table shows similar results as before except this time the results are from separate regressions run on each car class.

Table 13: Regression Results for Each Car Class

Coefficient	Small	Medium	Luxury	Sports	Minivan
JPN (prob)	0.077 (0.002)	0.034 (0.001)	0.029 (0.003)	-0.041 (0.031)	0.093 (0.006)
Rel2 (prob)	-0.041 (0.217)	0.045 (0.002)	0.029 (0.457)	0.067 (0.020)	-0.055 (0.321)
Rel3 (prob)	0.005 (0.885)	0.047 (0.007)	0.043 (0.273)	0.109 (0.000)	-0.139 (0.030)
Rel4 (prob)	-0.066 (0.114)	0.064 (0.001)	0.049 (0.225)	0.157 (0.000)	-0.112 (0.119)
Rel5 (prob)	-0.073 (0.107)	0.096 (0.000)	0.060 (0.154)	0.132 (0.000)	-0.142 (0.075)

(Note: the complete regression results are found in the Appendix I.)

What is evident from the above table is that coefficients on the reliability ratings did not make any significant difference except for in the 'Medium' car class. Furthermore, every car class except 'Sports' has a positive coefficient on the Japan dummy variable. All of the JPN coefficients are significant. What may make the difference for the 'Sports' class is that the sample of Japanese sports cars was very slim and not representative of the class. All of the Japanese cars in this class are considered mild attempts at sports cars rather than full-out street-legal racing machines. Japanese sports cars like the Acura NSX and the Toyota Supra our left out of the sample. The North American sports cars in the sample are, on the other hand, considered full-out sports cars like Corvettes and Mustangs with the horsepower to back them up.

Note that different base years are chosen for each regression because of what the data permitted.

A General Specification Test:

The RESET is used to see if the regression model is appropriate. Three fitted terms (i.e. (used/new)2, (used/new)3, and (used/new)4) are used as extra regressors for each regression. The results are as follows:

Table 14: General Specification Test Results

Regression	Probability	H0: Specification is true
All	0.025443	Reject
Small	0.781465	Do not reject
Medium	0.898401	Do not reject
Luxury	0.809155	Do not reject
Sports	0.095409	Do not reject
Minivan	0.900694	Do not reject

The form of misspecification in the "All" category may be that the coefficients vary with car type.

IV. CONCLUSIONS:

Used Japanese vehicles hold more of their initial value than do domestic automobiles. As expected, cars hold less of their initial value the older they are. Evidence of differences among car classes is believable. For example, luxury cars depreciate more quickly than do small cars. Also, small cars depreciate least out of any car class.

A conjecture as to why Japanese vehicles hold their value about 4% better than domestic vehicles (given that reliability and some hedonic characteristics are held constant) could be that consumers are accurate in their pricing and those rating reliability have underestimated Japanese vehicles. It could also be the case that consumers have overpriced Japanese vehicles relative to domestics and that the reliability ratings are accurate.

Another reason why domestic cars hold their value so well even though Japanese vehicles have such higher reliability ratings may be that there is something inherently attractive about domestic vehicles. Styling, dealer characteristics, safety and honouring warranties are but a few more things that are not included in regressions here, but that may tip the scale more toward domestic vehicles given that the reliability of Japanese vehicles is high relative to domestic vehicles.

Appendix II contains some further discussion about specification.

Appendix 1: Regression Results

Table 15: All

Dependent Variable: USED_TO_NEW
Method: Least Squares
Date: 08/14/01 Time: 21:41
Sample: 1 473
Included observations: 473

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
JPN	0.041455	0.008	5.264	0.000
Y88	-0.616293	0.051	-11.987	0.000
Y89	-0.580710	0.054	-10.804	0.000
Y90	-0.541684	0.048	-11.294	0.000
Y91	-0.501229	0.044	-11.455	0.000
Y92	-0.452654	0.042	-10.814	0.000
Y93	-0.406239	0.042	-9.783	0.000
Y94	-0.360153	0.041	-8.679	0.000
Y95	-0.289963	0.042	-6.915	0.000
Y96	-0.222381	0.042	-5.283	0.000
Y97	-0.151537	0.044	-3.484	0.001
Y98	-0.059501	0.043	-1.388	0.166
REL2	0.027959	0.018	1.582	0.114
REL3	0.046714	0.015	3.165	0.002
REL4	0.030459	0.016	1.898	0.058
REL5	0.051104	0.016	3.212	0.001
MEDIUM	-0.016011	0.008	-1.913	0.056
LUXURY	-0.074655	0.007	-10.558	0.000
SPORTS	-0.036472	0.014	-2.633	0.009
MINIVAN	-0.105019	0.016	-6.724	0.000
C	0.705616	0.055	12.872	0.000

R-squared	0.864425	Mean dependent var	0.433923
Adjusted R-squared	0.858426	S.D. dependent var	0.169237
S.E. of regression	0.063678	Akaike info criterion	-2.626581
Sum squared resid	1.832796	Schwarz criterion	-2.441927
Log likelihood	642.1863	F-statistic	144.0971
Durbin-Watson stat	0.915780	Prob(F-statistic)	0.000000

Table 16: Small

Dependent Variable: USED_TO_NEW
Method: Least Squares
Date: 08/14/01 Time: 21:26
Sample: 1 96
Included observations: 96

Variable	Coefficient	Std. Error	t-Statistic	Prob.
JPN	0.077394	0.023468	3.297842	0.0015
Y89	0.032745	0.079077	0.414088	0.6799
Y90	0.084425	0.079077	1.067636	0.2889
Y91	0.145830	0.059277	2.460157	0.0160
Y92	0.186562	0.061166	3.050080	0.0031
Y93	0.262166	0.062040	4.225722	0.0001
Y94	0.327347	0.061588	5.315117	0.0000
Y95	0.396406	0.062197	6.373393	0.0000
Y96	0.487951	0.063341	7.703569	0.0000
Y97	0.558198	0.064106	8.707366	0.0000
Y98	0.665332	0.064497	10.31568	0.0000
REL2	-0.041026	0.032960	-1.244697	0.2169
REL3	0.005447	0.037549	0.145054	0.8850
REL4	-0.066308	0.041450	-1.599712	0.1136
REL5	-0.073271	0.044879	-1.632615	0.1065
C	0.068788	0.065622	1.048240	0.2977

R-squared	0.913166	Mean dependent var	0.467396
Adjusted R-squared	0.896884	S.D. dependent var	0.174129
S.E. of regression	0.055916	Akaike info criterion	-2.778929
Sum squared resid	0.250126	Schwarz criterion	-2.351537
Log likelihood	149.3886	F-statistic	56.08629
Durbin-Watson stat	0.805256	Prob(F-statistic)	0.000000

Table 17: Medium

Dependent Variable: USED_TO_NEW
Method: Least Squares
Date: 08/14/01 Time: 21:37
Sample: 1 102
Included observations: 102

Variable	Coefficient	Std. Error	t-Statistic	Prob.
JPN	0.033700	0.009446	3.567680	0.0006
Y88	0.012225	0.041896	0.291797	0.7712
Y89	0.040245	0.036283	1.109186	0.2705
Y90	0.066644	0.036960	1.803141	0.0749
Y91	0.079143	0.032911	2.404784	0.0184
Y92	0.119161	0.032918	3.619988	0.0005
Y93	0.173084	0.033312	5.195851	0.0000
Y94	0.227967	0.033759	6.752847	0.0000
Y95	0.302682	0.033669	8.989986	0.0000
Y96	0.362377	0.033763	10.73293	0.0000
Y97	0.446540	0.033992	13.13645	0.0000
Y98	0.511204	0.034173	14.95940	0.0000
REL2	0.045280	0.014077	3.216673	0.0018
REL3	0.047484	0.017027	2.788795	0.0065
REL4	0.063838	0.018949	3.368884	0.0011
REL5	0.095718	0.018751	5.104744	0.0000
C	0.092543	0.029625	3.123780	0.0024

R-squared	0.974765	Mean dependent var	0.430168
Adjusted R-squared	0.970015	S.D. dependent var	0.171085
S.E. of regression	0.029625	Akaike info criterion	-4.049365
Sum squared resid	0.074601	Schwarz criterion	-3.611869
Log likelihood	223.5176	F-statistic	205.2109
Durbin-Watson stat	0.902718	Prob(F-statistic)	0.000000

Before the general regression is run on all the stacked data (i.e. all the classes), a similar regression will be run on the two samples of the luxury car data. The specification used is the same as above except there are no dummy variables for car class. This is to perform a kind of sensitivity analysis. On the one hand there is the smaller luxury car sample that the guide publishes, and on the other there is the expanded luxury car sample that includes more observations for Japanese cars.

Table 18: Specification Test for Two Luxury Car Samples

Sample:	Luxury (Smaller Sample)		Luxury (Larger Sample)	
Number of Observations	106		140	
Coefficient on JPN	0.025556	Prob. 0.0262	0.029211	Prob. 0.0030
P-value of REL2 coeff.	0.3637		0.4568	
P-value of REL3 coeff	0.2397		0.2734	
P-value of REL4 coeff	0.2259		0.2246	
P-value of REL5 coeff	0.0968		0.1541	

So, both of the Japanese dummy coefficients are significant and in fact gained some significance with the larger sample. The difference on the Japanese dummy coefficient between the two regressions is 0.003655. The dummy variables on every reliability rating are insignificant in both situations, however, it should not be surprising that coefficients were even more insignificant with the larger sample. Recall that many of the Japanese vehicles were given identical reliability ratings.

Regardless of these findings, common sense says that one should count all the Japanese luxury vehicles as separate observations just as the domestic luxury vehicles were reported. Thus, the larger of the two luxury samples was chosen to include in the general regression for all the cars.

Table 19: Luxury

Dependent Variable: USED_TO_NEW
Method: Least Squares
Date: 08/14/01 Time: 21:36
Sample: 1 140
Included observations: 140

Variable	Coefficient	Std. Error	t-Statistic	Prob.
JPN	0.029211	0.009636	3.031468	0.0030
Y89	0.022089	0.051922	0.425433	0.6713
Y90	0.048317	0.045223	1.068419	0.2874
Y91	0.093877	0.039296	2.388974	0.0184
Y92	0.135337	0.039237	3.449228	0.0008
Y93	0.175581	0.039049	4.496485	0.0000
Y94	0.214277	0.039399	5.438616	0.0000
Y95	0.274694	0.039348	6.981145	0.0000
Y96	0.346254	0.039478	8.770815	0.0000
Y97	0.423531	0.039459	10.73332	0.0000
Y98	0.520438	0.039720	13.10254	0.0000
REL2	0.028720	0.038472	0.746515	0.4568
REL3	0.043088	0.039170	1.100042	0.2734
REL4	0.048897	0.040063	1.220495	0.2246
REL5	0.059653	0.041598	1.434060	0.1541
C	0.052661	0.053778	0.979225	0.3294

R-squared	0.947900	Mean dependent var	0.388998
Adjusted R-squared	0.941598	S.D. dependent var	0.151922
S.E. of regression	0.036714	Akaike info criterion	-3.664091
Sum squared resid	0.167144	Schwarz criterion	-3.327903
Log likelihood	272.4863	F-statistic	150.4026
Durbin-Watson stat	0.687494	Prob(F-statistic)	0.000000

Table 20: Luxury Alt (i.e. with condensed Lexus and Infiniti categories – not used in the 'All' category)

Dependent Variable: USED_TO_NEW
Method: Least Squares
Date: 08/15/01 Time: 22:48
Sample: 1 106
Included observations: 106

Variable	Coefficient	Std. Error	t-Statistic	Prob.
JPN	0.025556	0.011306	2.260500	0.0262
Y89	0.022089	0.051661	0.427580	0.6700
Y90	0.046490	0.045096	1.030920	0.3053
Y91	0.093528	0.039342	2.377289	0.0196
Y92	0.125709	0.039529	3.180198	0.0020
Y93	0.164689	0.039529	4.166296	0.0001
Y94	0.203664	0.039877	5.107327	0.0000
Y95	0.269959	0.039706	6.798963	0.0000
Y96	0.349090	0.039886	8.752296	0.0000
Y97	0.422822	0.040143	10.53284	0.0000
Y98	0.517497	0.040356	12.82319	0.0000
REL2	0.035075	0.038421	0.912903	0.3637
REL3	0.046394	0.039200	1.183512	0.2397
REL4	0.049124	0.040284	1.219444	0.2259
REL5	0.072122	0.042981	1.677977	0.0968
C	0.053010	0.053687	0.987399	0.3261

R-squared	0.953347	Mean dependent var	0.375005
Adjusted R-squared	0.945571	S.D. dependent var	0.156579
S.E. of regression	0.036530	Akaike info criterion	-3.643111
Sum squared resid	0.120099	Schwarz criterion	-3.241082
Log likelihood	209.0849	F-statistic	122.6082
Durbin-Watson stat	0.697516	Prob(F-statistic)	0.000000

Table 21: Sports

Dependent Variable: USED_TO_NEW
Method: Least Squares
Date: 08/14/01 Time: 21:39
Sample: 1 75
Included observations: 75

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
JPN	-0.040760	0.018235	-2.235304	0.0308
Y92	0.041354	0.027317	1.513893	0.0807
Y93	0.074163	0.028082	2.640917	0.0083
Y94	0.150276	0.029015	5.179249	0.0000
Y95	0.218405	0.027229	8.021060	0.0000
Y96	0.277123	0.028652	9.671871	0.0000
Y97	0.353167	0.029548	11.95250	0.0000
Y98	0.412731	0.031680	13.02795	0.0000
REL2	0.067363	0.044366	1.518362	0.0203
REL3	0.109228	0.045576	2.396600	0.0000
REL4	0.157400	0.047883	3.287193	0.0000
REL5	0.131570	0.051182	2.570626	0.0002
C	0.202048	0.041959	4.815378	0.0000

R-squared	0.903659	Mean dependent var	0.490669
Adjusted R-squared	0.885013	S.D. dependent var	0.165460
S.E. of regression	0.056107	Akaike info criterion	-2.766796
Sum squared resid	0.195176	Schwarz criterion	-2.365098
Log likelihood	116.7548	F-statistic	48.46253
Durbin-Watson stat	1.082028	Prob(F-statistic)	0.000000

Table 22: Minivan

Dependent Variable: USED_TO_NEW
Method: Least Squares
Date: 08/14/01 Time: 21:38
Sample: 1 60
Included observations: 60

Variable	Coefficient	Std. Error	t-Statistic	Prob.
JPN	0.093345	0.032116	2.906471	0.0056
Y91	0.139075	0.105305	1.320693	0.1931
Y92	0.197481	0.111060	1.778141	0.0820
Y93	0.243149	0.114020	2.132509	0.0383
Y94	0.305283	0.114039	2.676999	0.0103
Y95	0.355353	0.110569	3.213870	0.0024
Y96	0.412366	0.114147	3.612597	0.0007
Y97	0.528606	0.114553	4.614510	0.0000
Y98	0.689977	0.116951	5.899732	0.0000
REL2	-0.055316	0.055225	-1.001642	0.3218
REL3	-0.139229	0.062145	-2.240384	0.0299
REL4	-0.112163	0.070601	-1.588702	0.1190
REL5	-0.142640	0.078301	-1.821674	0.0750
C	0.108026	0.094358	1.144860	0.2582

R-squared	0.772999	Mean dependent var	0.420643
Adjusted R-squared	0.708846	S.D. dependent var	0.174870
S.E. of regression	0.094358	Akaike info criterion	-1.682485
Sum squared resid	0.409555	Schwarz criterion	-1.193804
Log likelihood	64.47454	F-statistic	12.04940
Durbin-Watson stat	0.667168	Prob(F-statistic)	0.000000

Appendix 2: Alternative Regression Specifications

Part 1:

Originally, the following regression was chosen:

$$(1)\ \ln\_used\_price = \alpha + \beta(\ln\_new\_price) + \gamma JPN + \delta_1 Y91 + \delta_2 Y92 + \delta_3 Y93 + \delta_4 Y94 + \delta_5 Y95 + \delta_6 Y96 + \delta_7 Y97 + \delta_8 Y98 + \zeta_1 rel2 + \zeta_2 rel3 + \zeta_3 rel4 + \zeta_4 rel5 + \varepsilon_1$$

where

ln_used_price:	Is the natural log of the used price in 1992 dollars
ln_new_price:	Is the natural log of the new price in 1992 dollars

After looking at the regression output of the 'All' category, a mentor pointed out that the coefficient estimate on ln_new_price looked suspiciously close to one. A Wald test was performed on this coefficient and the similar coefficient on the separate car class regressions to see if the null held.

Table 23: Wald Test Results

Car class	**H_0: $\beta = 1$ (by Wald test)**	
	p-value	**Conclusion**
Small	0.4822	Do not reject
Medium	0.0000	Reject
Luxury	0.0000	Reject
Sports	0.0154	Reject
Minivan	0.2461	Do not reject
All	0.4774	Do not reject

Since the null was not rejected for 'All' cars, the assumption was made that this applied to all cars. Thus, the specification was changed to bring the ln_new_price variable to the left-hand side. The dependent variable became (used price)/(new price). This makes sense since what is being investigated is the change in a used car price relative to its new car price.

Before this specification change was implemented, another test was performed to determine if the influence of reliability jumps by equal amounts as one goes from category to adjacent category. If the dummies are equal to each other then the influence of reliability differs for the base category but is the same for all other categories. The way this was accomplished was by having another variable called RELIABILITY that takes on integer values from 1 to 5. RELIABILITY takes the place of the four reliability dummies in creating a restricted version of regression (1).

$$(2)\ ln\_used\_price = \eta + \theta(ln\_new\_price) + \iota JPN + \kappa_1 Y91 + \kappa_2 Y92 + \kappa_3 Y93 + \kappa_4 Y94 + \kappa_5 Y95 + \kappa_6 Y96 + \kappa_7 Y97 + \kappa_8 Y98 + \lambda RELIABILITY + \varepsilon_2$$

An F-test was performed to test if the restriction that all of the slope coefficients on the reliability dummy coefficients are equal to each other using the restricted (2) and unrestricted (1) sum of squares from the regressions. The p-value was given by Eviews after calculating the F-statistic. A 5% level of significance was used.

Table 24: F-Test for Reliability Variable Restriction

Car class	$H_0: \zeta_1=\zeta_2=\zeta_3=\zeta_4$	
	p-value	Conclusion
Small	0.0012	Reject
Medium	0.0616	Do not reject
Luxury	0.1791	Do not reject
Sports	0.0801	Do not reject
Minivan	0.3868	Do not reject
All	0.0794	Reject

Again, since the null is rejected for the 'All' cars category, the categorical variable was not included in any of the regressions.

Now, we have the specification that is used in the body of this paper.

$$\text{used price / new price} = \mu + \nu JPN + \psi_1 Y88 + \psi_2 Y89 + \psi_3 Y90 + \psi_4 Y91 + \psi_5 Y92 + \psi_6 Y93 + \psi_7 Y94 + \psi_8 Y95 + \psi_9 Y96 + \psi_{10} Y97 + \psi_{11} Y98 + o_2 REL2 + o_3 REL3 + o_4 REL4 + o_5 REL5 + \pi_1 MEDIUM + \pi_2 LUXURY + \pi_3 SPORTS + \pi_4 MINIVAN + \varepsilon$$

After all this, one more test was implemented - this time, to check to see if the coefficients on the reliability dummy variables are jointly significantly different from zero.

The Wald test was used for the restriction on each of the regressions.

Table 25: Wald Test Results for Reliability Variable Jointly Different From Zero

Car class	$H_0: \zeta_1=\zeta_2=\zeta_3=\zeta_4=0$	
	p-value	**Conclusion**
Small	0.000272	Reject
Medium	0.000018	Reject
Luxury	0.275203	Do not reject
Sports	0.000815	Reject
Minivan	0.145573	Do not reject
All	0.000170	Reject

Again, the general specification chosen to apply all the regressions was because it fit best with the 'All' regression results.

Appendix 2 Part 2:

Another idea to make the project richer in information was to include an interaction variable. The interaction is between the reliability dummies and the dummies for each year. The thought behind this is to determine what kind of effect a reliability rating has for different aged cars.

Year dummy variables 1992 through 1998 (with 1991 as the base category) are used to interact with dummy variables 1 to 4 (with 5 as the base category). These interaction variables are included in a general regression with an intercept, a dummy for Japan, year dummies and reliability rating dummies.

One is careful to omit the 'year1997-rel1' and 'year1998-rel2' interaction variables since no observations exist with these combinations and thus will create a matrix that is not of full rank when trying to perform the regression.

What one finds is that no coefficients are significant except those on the year dummies from 1995 to 1998 and the dummy on Japan. The full regression results follow this discussion (on the next page).

So what happened? The answer lies in the reliability ratings. Each of the cars in the Lemon-Aid guide is given a reliability rating relative to ALL the other cars in the guide across years of make and class. No matter how cheap a newer car (i.e. 1997 and 1998 models) is in terms of quality, it will never have a reliability of 1. Conversely, now matter how high quality a new car is it will never have a reliability rating of 5 when its about seven years old or older.

This is why an interaction dummy between year and reliability does not make sense in the context of the present data.

Table 26: Interaction Variable Regression Results

Dependent Variable: USED_TO_NEW'
Method: Least Squares
Date: 08/19/01 Time: 20:48
Sample: 1 459
Included observations: 459

Variable	Coefficient	Std. Error	t-Statistic	Prob.
JPN	0.030822	0.007751	3.976575	0.0001
Y92	0.019159	0.080496	0.238010	0.8120
Y93	0.038512	0.063681	0.604769	0.5457
Y94	0.088411	0.059684	1.481328	0.1393
Y95	0.163320	0.059067	2.764976	0.0059
Y96	0.228393	0.058526	3.902418	0.0001
Y97	0.307192	0.058641	5.238548	0.0000
Y98	0.394354	0.058461	6.745583	0.0000
REL1	-0.092018	0.060226	-1.527881	0.1273
REL2	-0.075750	0.058562	-1.293494	0.1966
REL3	-0.023122	0.061198	-0.377832	0.7057
REL4	-0.047480	0.061230	-0.775434	0.4385
MEDIUM	-0.012594	0.009005	-1.398538	0.1627
LUXURY	-0.074970	0.008258	-9.078846	0.0000
SPORTS	0.016743	0.009526	1.757602	0.0796
MINIVAN	-0.033259	0.010785	-3.083875	0.0022
Y2R1	0.021568	0.084469	0.255340	0.7986
Y2R2	0.028833	0.082525	0.349381	0.7270
Y2R3	-0.003341	0.085105	-0.039252	0.9687
Y2R4	0.022310	0.084723	0.263333	0.7924
Y3R1	0.067359	0.068998	0.976248	0.3295
Y3R2	0.054225	0.066034	0.821165	0.4120
Y3R3	0.054050	0.070191	0.770039	0.4417
Y3R4	0.038138	0.068915	0.553413	0.5803
Y4R1	0.068127	0.066423	1.025664	0.3056
Y4R2	0.056461	0.062526	0.903000	0.3670
Y4R3	0.063627	0.066654	0.954587	0.3403
Y4R4	0.048685	0.065603	0.742115	0.4584
Y5R1	0.065255	0.066552	0.980509	0.3274
Y5R2	0.050733	0.061979	0.818559	0.4135
Y5R3	0.015032	0.065301	0.230200	0.8180
Y5R4	0.037166	0.064686	0.574565	0.5659
Y6R1	0.116862	0.069215	1.688395	0.0921
Y6R2	0.069238	0.062727	1.103799	0.2703
Y6R3	-0.028707	0.064417	-0.445647	0.6561
Y6R4	0.047956	0.064215	0.746803	0.4556
Y7R2	0.100462	0.063967	1.570514	0.1171
Y7R3	-0.012958	0.064572	-0.200679	0.8410
Y7R4	0.049374	0.063917	0.772470	0.4403
Y8R2	0.108732	0.065154	1.668842	0.0959
Y8R3	0.019357	0.064764	0.298894	0.7652
Y8R4	0.045488	0.063898	0.711887	0.4769
C	0.306239	0.058317	5.251247	0.0000

R-squared	0.890013	Mean dependent var	0.442896
Adjusted R-squared	0.878908	S.D. dependent var	0.163570
S.E. of regression	0.056919	Akaike info criterion	-2.805359
Sum squared resid	1.347769	Schwarz criterion	-2.418541
Log likelihood	686.8298	F-statistic	80.14889

Reference

Edmonston, Phil. 2000. *Lemon-Aid Used Cars 2001 28th ed*. Toronto: Stoddart Publishing Co. Limited.