The Determination of Ski Lift Ticket Price in the Rocky Mountains

Ethan Spiro
Econometrics 272
Professor Studenmund
11/6/2011

Appendix Reference

- Appendix 1: Data Collection From Excel
- Appendix 2: First Regression
- Appendix 3: Second Regression
- Appendix 4: Testing with NewLFT Dropped
- Appendix 5: VIF Tests
- Appendix 6: Park Test
- Appendix 7: Heteroskedasticity Adjusted Standard Errors

Table of Contents

I. Abstract…………………………………………………………………3

II. The Dependent Variable……………………………………………………..5

III. Selecting the Dependent variable………………………………………5

IV. Literature Review……………………………………………………………8

V. Independent variables………………………………………………………10

a) First Tier…………………………………………………………………10

b) Second Tier……………………………………………………….....…12

c) Third Tier……………………………………………………..…….…14

d) Fourth Tier………………………………………………………………15

VI. Data Collection………………………………………………………....……16

VII. Specifying the Model……………………………………………………..…17

VIII. First Regression……………………………………………………..……18

IX. Second Regression…………………………………………………….....…21

X. Final Documented Regression……………………………………………30

XI. Conclusion…………………………………………..……………31

XII. Works Referenced……………………………………………………..34

XIII. Glossary of Variables……………………………………….………35

I. Abstract:

Skiing is an expensive hobby and requires a fair amount of investment for the user to actually make their way as gracefully as possible down a mountain. Not only is the gear expensive for this hobby, but lift ticket costs are a substantial investment. For example, the current price for a 1-day lift ticket at Deer Valley, has reached a staggering $97. This is no meager sum for the average American. It would only seem natural that the Ski Industry would be hurting with all the concern these days about global warming and the economy slumping in a recession. Quite the opposite has actually occurred over the past five to six years though, "Ski resorts in the United States have shown steady increases in revenue over the last five years, according to the research firm IbisWorld" and "In 2009, Ski resorts generated $2.6 billion in revenue with 57.8% of that revenue coming from lift and season ticket sale alone" (Korkki). The ski industry seems to be catering to the wealthy more than ever before, with resorts completely restructuring themselves to be more appealing to this richer clientele. This begs the question of why lift tickets are priced as they are. In other words, what does the lift ticket actually guarantee the skier or boarder. Physically, it is just a little slip of paper, but it allows access to everything the mountain has to offer. **This project seeks to explain and model the characteristics of a ski resort, that can either increase or decrease *the price* of a ski lift ticket by either increasing or decreasing the utility of the consumer who purchased said ticket. I hope to build a hedonic price model to describe the relationship between the price of the ticket and the independent variables affecting that price.** With a perfect equation, the model could be used to see the impacts of different variables on the price of the ticket, but also it could be adapted to make assumptions on what resorts might be overcharging or undercharging depending on what their lift ticket price is in relation to what they have to offer. This would be highly relevant to the

average consumer who is probably feeling the stress on their wallets from skiing more than ever before.

II. The Dependent Variable:

The dependent variable in this study is **the price of a one-day adult lift ticket at any given resort in the Rocky Mountains in the 2005/2006 ski season**. The price is in dollar terms and the forty chosen samples were randomly sampled from the full list of the 99 Ski Resorts in the Rocky Mountains. The Rocky Mountains States include Idaho, Montana, Wyoming, Colorado, Utah, Arizona and New Mexico. The adult ticket is defined as *however the mountain in sampling chose to define it*. Although there is no standard adult ticket, the adult ticket usually means around ages (16 –18) through ages (60 & up), but it should be noted that the mountains did not all encompass the same range of ages within their respective adult tickets. This hitch was viewed as negligible and later explained in the independent variable section (V.C.) but still should be taken into account by the reader. The 'one-day' part means that the ticket was in effect from when the mountain opened to when they closed, but it should also be noted that this value fluctuated slightly between different resorts and during different times of the season. College rates, child rates, senior citizen rates, season passes, half-day rates, afternoon rates, multi-mountain passes, early season rates, late season rates and discount tickets from non-mountain vendors were not taken into account.

III. Selecting the Dependent Variable Samples

For an unbiased selection of my samples, I numbered the full list of resorts in the Rocky Mountains from 1 – 99 and then used a random number generator with limits [1, 99] to select 40 resorts from that list. I did not want to select the resorts myself, because I would have been more inclined to select resorts I was familiar with, hence skewing the result. If I generated a number I had already generated, I simply discarded it and generated another. For my random number generator I used Random.org which can be found in the works cited.

Below is the full list of the 99 resorts with the ones selected in boxes as well as a list of the ones selected below with notes below on problems I encountered and resorts I dropped:

Arizona

1. Arizona Snowbowl
2. Elk Ridge Ski Area
3. Mount Lemmon Ski Valley
4. Sunrise Park Resort

Colorado

5. Arapahoe Basin
6. Aspen/Snowmass
7. Aspen Highlands
8. Aspen Mountain
9. Buttermilk
10. Snowmass
11. Beaver Creek Resort
12. Breckenridge Ski Resort
13. Copper Mountain Resort
14. Crested Butte Mountain
15. Durango Mountain Resort
16. Echo Mountain Park
17. Eldora Mountain Resort
18. Hesperus Ski Area
19. Howelsen Ski Area
20. Kendall Mountain
21. Keystone Resort
22. Loveland Ski Area
23. Monarch Ski Area
24. Powderhorn Resort
25. Silverton Mountain
26. Ski Cooper
27. SolVista Basin
28. Sunlight Mountain Resort
29. Steamboat Resort
30. Telluride Ski Resort
31. Vail Ski Resort
32. Winter Park Resort
33. Wolf Creek Ski Area

Idaho

34. Bald Mountain
35. Bogus Basin
36. Brundage Mountain
37. Cottonwood Butte
38. Kelly Canyon
39. Little Ski Hill
40. Lost Trail Powder Mountain
41. Lookout Pass
42. Magic Mountain Resort
43. Pebble Creek
44. Pomerelle
45. Rotarun Ski Area
46. Schweitzer Mountain
47. Silver Mountain Resort
48. Snowhaven
49. Soldier Mountain
50. Sun Valley

Montana

51. Bear Paw
52. Big Sky
53. Blacktail Mountain
54. Bridger Bowl
55. Discovery
56. Great Divide
57. Lookout Pass
58. Lost Trail Powder Mountain
59. Maverick Mountain
60. Montana Snowbowl
61. Moonlight Basin
62. Red Lodge Mountain Resort
63. Showdown
64. Teton Pass
65. Turner Mountain
66. Big Mountain
67. Yellowstone Club

New Mexico

68. Angel Fire Resort
69. Pajarito Mountain
70. Red River Ski Area
71. Ski Apache
72. Ski Santa Fe
73. Sipapu
74. Taos Ski Valley
75. Sandia Peak
76. Ski Cloudcroft

Utah

77. Alta
78. Beaver Mountain
79. Brian Head
80. Brighton
81. The Canyons
82. Deer Valley
83. Eagle Point Ski Resort
84. Park City Mountain Resort
85. Powder Mountain
86. Snowbasin
87. Snowbird
88. Solitude
89. Sundance
90. Wolf Mountain

Wyoming

91. Beartooth Pass
92. Big Horn
93. Grand Targhee
94. Hogadon
95. Jackson Hole
96. Ski Antelope
97. Snow King
98. Snowy Range
99. White Pine

The 40 Samples: Numbers Generated and Corresponding Resorts

1. (93) Grand Targhee, WY
2. (66) Big Mountain, MT
3. (65) Turner Mountain, MT
4. (43) Pebble Creek, ID
5. (1) Arizona Snowbowl, AZ
6. (31) Vail, CO
7. (22) Loveland, CO
8. (50) Sun Valley, ID
9. (96) Antelope, WY
10. (87) Snowbird, UT
11. (3) Mount Lemmon Ski Valley, AZ
12. (46) Schweitzer, ID
13. (69) Pajarito Mountain, NM
14. (29) Steamboat, CO
15. (95) Jackson Hole, WY
16. (94) Hogadon, WY
17. (82) Deer Valley, UT
18. (15) Durango Ski Resort, CO
19. (49) Soldier Mountain, ID
20. (53) Blacktail Mountain, MT
21. (5) Arapahoe Basin, CO
22. (61) Moonlight Basin, MT
23. (77) Alta, UT
24. (37) Cottonwood Butte, ID
25. (52) Big Sky, MT
26. (10) Snowmass, CO
27. (74) Taos Ski Valley, NM
28. (80) Brighton, UT
29. (9) Buttermilk, CO
30. (97) Snow King Ski Area, WY
31. (58) Lost Trail Ski Area, MT
32. (32) Winter Park Resort, CO
33. (71) Ski Apache, NM
34. (79) Brian Head, UT
35. (35) Bogus Basin, ID
36. (14) Crested Butte Mountain, CO
37. (57) Lookout Pass, MT
38. (81) The Canyons, UT
39. (88) Solitude, UT
40. (28) Sunlight Mountain Resort, CO

Despite being selected via the random number generator, certain samples had to be dropped due to lack of information for all independent variables during data collection. Resorts dropped due to lack of information:

Pine Creek, WY	Lookout Pass, MT	Keystone, CO
Beartooth Pass, WY	Little Ski Hill, ID	Howelson Hill, CO
White Pine, WY	Bald Mountain, ID	
Discovery, MT	Echo Mountain Park, CO	

I also dropped The Yellowstone Club, MT and Silverton, CO even though they were both selected by the random number generator because their ticket prices did not incorporate my definition assigned to the dependent variable in this study. Both of these resorts price their tickets based on membership fees. For example, Yellowstone Club charges members $20,000 annually to ski.

IV. Literature Review:

Before diving headfirst into the independent variable selection and specifying the model, it is important to do some research on the topic at hand. Even before that, it is important to fully understand and define the hedonic price model, since that is what I will be using for this regression.

The textbook, *Using Econometrics: A Practical Guide, Sixth Edition,* defines a Hedonic Price Model as one that uses measures of the quality of a product as independent variables instead of measures of the market for that product. The explanation continues by stating that Hedonic models are most useful when the product being analyzed is heterogeneous in nature because we need to analyze what causes products to be different and therefore have different prices. This very much is in tune with the Ski Industry, where the products or mountains are differentiated by their characteristics, and the price (of a lift ticket) is a reflection of these differentiated characteristics and based on the perceived utility of these characteristics.

There are generally two schools of thought for these mountain characteristics. Some economists in the past have stressed participatory characteristics as the central determinant of ski lift ticket prices. By participatory elements for ski resorts, I mean elements that are directly pinned to the product you pay for, which is skiing in this case. So for example, average snowfall is a participatory element because it is directly related to the activity you have purchased, but the hot tub you soak in after a day of skiing is not a participatory characteristic. A paper that exemplifies this school of thought is *Ski-Lift Pricing, with Applications to Labor and Other Markets* by Barro and Romer (1987). In this paper the authors study ski lift pricing and amusement park entrance fees, primarily by deconstructing the supply and demand components of these two markets. They come to the conclusions that the lift ticket is purchased by future utility judged by the consumer, and in the case of skiing, this utility is above all a function of the number of runs they receive. For them, number of runs is the most important factor, but they also include vertical drop and lift capacity as determinants as well. Near the end of their paper, they mention a final determinant as customer preference. They argue that mountains cater to either avid skiers or recreational skiers. Although I believe this exists at some level today, this model is slightly outdated and that resorts are trying to appeal to a broader crowd than they were in the past.

Another paper that approaches the question of lift ticket prices with the mindset that participatory characteristics are the sole determinant of price is *The Choice of Ski Areas: Estimation of a Generalized CES Preference Ordering with Characteristics* by Morey (1984). This paper takes a different approach than Barro and Romer (1987) and instead of analyzing supply and demand components, Morey analyzes customer behavior in choosing their ski resort. The main determinant of resort choice in this paper is from the amount of terrain available at a given resort and the types of that terrain. He also includes skier's ability as an element, which relates to the previous idea of avid skiers and recreational skiers. This study's downfall is a small sample size of only 163 students sampled.

A more relevant study dealing with participatory characteristics of ski resorts as determinants for their prices is *Market Segmentation and the Diffusion of Quality-Enhancing Innovations: The Case of Downhill Skiing* by (Mulligan and Llinares 2003). In this study they actually run a regression on lift ticket prices with chairlifts, vertical drop, location on U.S. National Forest land, the number of competitors within a 125-mile drive of the ski area, the population within 125 miles of the ski area, and a set of regional dichotomous variables. Again, the main determinant of utility for the skier deals with the number of runs for the consumer, and they make the case that detachable chairlifts significantly increase the number of runs because they are a more efficient transport system. They make an interesting point that the number of competitors is a proxy for demand in the equation, and that despite the belief that more competition leads to lower prices, price actually tends to increase when there are more competitors.

For the most part non-participatory characteristics are left out of the studies mentioned so far. Non-participatory characteristics are in this case, characteristics of the mountain that aren't directly related to the actual act of skiing, but relatable to the entire skiing experience. For example, the number of restaurants on the mountain, number lodges at the base and basically anything a ski resort encompasses. *The Effects of Non-Participatory Characteristics: The Case of Lift-Tickets* by (Koslow 2006) and *A Hedonic Price Model for Ski Lift Tickets* by (Falk 2006) are two studies that incorporate these non-participatory characteristics into their analysis on ski lift tickets, in the United States and Austria respectively. Both found the incorporation of such variables to be statistically significant in the direction they expected. I tend to agree that non-participatory elements are a theoretically valid point for increasing lift ticket prices. From general experience I have witnessed utility from skiing as more than a characteristic such as the number of runs, and I believe more than ever before, resorts are trying to improve these non-participatory elements and as a result their ticket prices are rising.

V. Independent Variables

* expected signs on the coefficients are to my interpretation of the variable and the theorized functional forms are to the best of my ability

A. 1st Tier: High Theoretical Validity
To Include in Initial Regression w/ Expected Coefficients and Theorized Functional Forms

- (SNW) Snowfall (inches/year): The more snow a resort receives per year, the better the skiing tends to be. Snowfall is a good indicator of how much utility the skier gets out of his/her individual runs.
  - Expected Sign on the Coefficient: (+)
    - Functional Form: Linear

- (VRT) Total Vertical Drop (feet): More vertical drop leads to longer runs and higher enjoyment. There tend to be fewer people on the slopes and more to patrol for the ski patrol. Mulligan and Llinares (2003) cite vertical drop as an important participatory characteristic of a mountain in determining lift ticket price.
  - Expected Sign on the Coefficient: (+)
    - Functional Form: Linear

- (ACR) Total Skiable Acres (Acres): Barro and Romer (1987) stress the more area for skiers, the more runs they have to choose from. Total skiable acres is a better measurement than *total runs* because runs are very different in size, especially in the Rockies where the terrain is varied with bowls, glades and much more open slopes. It should be noted that Total Skiable Acres might be fairly susceptible to some omitted variable bias because of lack of information. For example, information on the size of a ski patrol was unavailable for the samples chosen. The larger the ski area, the more patrol they tend to employ, and more ski patrol leads to a higher ticket price, so there could be a positive bias on Total Skiable Acres.
  - Expected Sign on the Coefficient: (+)
    - Functional Form: Semi-log right. I picked Semi-log right for skiable acres, because it does not make sense that acres will linearly affect price as they increase. At a certain point, the skier cannot access all the acres if there are too many. This does not mean they get less utility from more acres, but simply their utility increases at a decreasing rate with more and more acres. The ticket prices should reflect this. This is expressed using the semi-log right form.

- (PLS) Plush Rating (1 – 5 with 5 being the highest): The plush rating is an indicator of the strength of mountains non-participatory characteristics that Falk (2006) and Koslow (2006) cite in their papers. For example, how many restaurants are on the mountain, how many accommodations, presence of childcare centers, shuttles for the skiers etc… These characteristics are not directly related to skiing, but very theoretically valid in terms of affecting ticket prices by enhancing the overall skiing experience by pampering the users. The rating is also somewhat hard to determine. The process of determination of these ratings is explained in the data collection section.

- Expected Sign on the Coefficient: (+)
  - Functional Form: Linear

- (AAS) Total Acres of Artificial Snowmaking (Acres): Similar reasoning to why snowfall was included. Also, snowmaking is expensive for the resort, and this expense is probably reflected in higher ticket prices. My theory is that by omitting this variable, it would induce a negative variable bias on the coefficient of snowfall. This is because snowfall and snowmaking are negatively correlated and the expected sign for Total Acres of Artificial Snowmaking is positive.
  - Expected sign on the Coefficient: (+)
    - Functional Form: Linear

- (LFT) Lift Capacity (people/hour): Lift capacity per hour means the number of people who can ride up the mountain on a lift in one hour at full capacity. The greater a mountains lift capacity, the greater their ability to break up lift line and actually get people skiing, leading to a greater utility for the skier. Mulligan and Llinares (2003) say this is the most important factor for determining lift ticket price.
  - Expect sign on the Coefficient: (+)
    - Functional Form: Linear

B. 2nd Tier: Medium Theoretical Validity

Not to Include in Initial Regression but w/ Expected Signs of Coefficients and Theorized Functional Forms (think about adding these variables if regression shows signs of severe omitted variable bias)

- (SKI) Ski Only Resort (Dummy Variable → 0 – otherwise, 1 – ski only): Some resorts ban snowboarders and only allow skiers. Theoretically this would increase the price of the ticket, because the skiers utility would increase with no boarders crowding the slopes, and the mountain itself would be losing out on the profits from boarders and in response increase their ticket price.

- ➢ Expected sign on the Coefficient: (+)
  - ▪ Functional Form: Intercept Dummy

- ❖ (CMP) Number of other resorts in the state (number): When there are very few resorts in a State, the idea of skiing is somewhat of a novelty (example: Arizona) and the price of a lift ticket might be higher because of that. One would also expect that higher competition leads to lower prices of all the resorts, but Mulligan and Llinares (2003) note that "…we expect, ceteris paribus, that higher lift ticket prices at a particular ski resort would lead to higher prices at neighbouring ski resort". There are arguments both ways, and it could be that states with very few (Arizona) and many ski resorts (Colorado) have higher than normal prices. I will leave it out initially, but might be worth adding to the equation in some type of parabolic form if the equation shows serious signs of omitted variable bias.
  - ➢ Expected sign on the coefficient (+/ –)
    - ▪ Functional Form: Possible Quadratic

- ❖ (DVS) Diversified Runs/Difficulty (Dummy Variable → 0 – otherwise, 1 – At minimum 15% beginner, 35% Intermediate and 35% Advanced): The more diversified a mountain, is then the more a person would be willing to pay. For example, an entire family can go to the same mountain even if they have different skill levels. More diversity in runs also means better adaptability in all types of snow conditions. For example, when there hasn't been snowfall in a while, skiers can still enjoy the large percentage of intermediate groomed runs. Problem with this variable is that I'd have to come up with the definition of 'diversified', but still worth considering because of theory.
  - ➢ Expected sign on the Coefficient: (+)
    - ▪ Functional Form: Intercept Dummy

- ❖ (BCA) Variable for Backcountry Access (Dummy → 0 – otherwise, 1 – has backcountry access): Some mountains have fairly small inbound acreage, but have access to massive backcountry skiing areas through controlled gates that lead out of the resort. The backcountry is not included in total skiable acres, but basically acts in the same way by increasing the skier's utility. There isn't a clear correlation between backcountry access and any of the other variables, so it is hard to determine what the bias with omission would be. It is also hard to determine what exactly is defined as a backcountry, since they vary mountain to mountain. It is worth considering though if the equation has serious problems.

- Expected Sign on the Coefficient: (+)
  - Functional Form: Intercept Dummy

C. 3rd Tier: Low Theoretical Validity or Redundancy Expected
Dropped from Consideration

- (POP) Proximity to large population centers (Dummy → 0 – not within 30 miles, 1 – within 30 miles): Differing views of the effects on lift ticket prices. One argument is that the closer the ski resort is to a large population center, the more expensive the ticket is because the utility of the skier is increased from the convenience of the drive. Other arguments say this has little effect, and the longer the drive, the more "resort like" the mountain usually is. I'm unsure of what the sign might be, but still is worth considering if the equation shows omitted variable bias.

- (RUN) Number of Runs (total number of runs): Redundancy with Total Skiable Acres.

- (CHR) Number of Chairlifts (total number of chairlifts): Redundancy with Lift Capacity.

- (SUN) Days of Sunshine (days out of average season): Theory is weak because preference is not always sun while skiing; it can ruin the snow conditions sometimes.

- (SKL) Ski school (Dummy Variable → 0 – otherwise, 1 – has a school): Not directly relatable to ticket price, and redundancy with plush rating.

- (TMP) Temperature (average temperature during months of season): Theory is weak because preference is not always on the same temperature.

- (AGE) How many Ages the Pass accounts for (range of total ages): Very minor changes between resorts and virtually identical if mountains are close to each other. Viewed as irrelevant.

- (FST) Located on US Forest Land (Dummy Variable → 0 – otherwise, 1 – yes): The land is actually cheaper to rent, but these resorts have high constraints on parking and expansion. Unclear final effect on prices.

D. 4th Tier: Varying Degrees of Theoretical Validity
Insufficient Data Available

* Note: Although these variables did not have sufficient data for collection, the theoretically strong ones should not be cast off without inspection. If my equation does not perform as hypothesized, it might be important to find a proxy for some of these to add in, or at the very least mention them in analysis of the problems encountered and conclusion.

- (ADV) Advertising (dollars/year): Resorts don't publish this.

- (GRM) Number of Ski Grooming machines the Resort Owns (number of groomers): Insufficient Data.

- (DMF) Variable for Distinct Mountain Feature (Dummy variable → 0 – otherwise, 1 – at least one): Example – Corbett's Couloir, Jackson Hole, WY. Problems judging these features.

- (COR) Owned by a Larger Corporation (Dummy variable → 0 – otherwise, 1 – yes)

- (PRK) Terrain Park (Dummy variable → 0 – otherwise, 1 – at least one): They change significantly throughout each season making data unavailable.

- (RCE) Race course (Dummy variable → 0 – otherwise, 1 – at least one): Same reason as Terrain Park.

- (PAT) Ski Patrol (number of permanent staff): Not published for most resorts.

- (WRK) Lift workers (number of permanent staff): Fluctuates too much and not published.

- (HRS) Hours of Operation (hours/day): Fluctuates throughout the season, not conclusive data.

VI. Data Collection

* Data can be referenced in **Appendix 1**:

Description of Data: Almost all data is taken from the 2005/2006 ski season and is from Edition 13 of The White Book® referenced at the end of this paper. I chose the 2005/2006 season despite current data being available for most of the variables, because Lift Capacity, a highly important theorized variable, was only available for the 2005/2006 season. I made sure to collect the other variables from the 2005/2006 seasons for accurate effects. For example, if I chose all my independent variables from 2005/2006 except for say acres, which I used 2006/2007 values, I would be using incorrect data because the resort could have opened up a new area for skiing in 2006/2007, and that expansion wouldn't be accounted for in the price of the ticket in 2005/2006. I found no need to lag any variables since skiers tend to look at the upcoming/current year for their determination in buying tickets, and resorts are great at publishing this information. Once the data had been collected I cleaned and inspected it to check for mistakes or outliers. More information on this regarding the dependent samples is mentioned earlier in Section III.

My only exception for data collection was annual snowfall. I could not find annual snowfall statistics in the White Book for all my samples, so I found it in other places, most notably the resorts individual website or the website OntheSnow.com which is referenced at the end of this paper. I judged this as appropriate because annual snowfall is the ten year average of snowfall, so there is no reason for this to change significantly from year to year.

Determination of the Plush Rating: All of my data is compiled of hard statistics except for the plush rating. To determine the plush rating, or a rating of the non-participatory factors influencing a lift ticket price, I did three things:

i. Assessed the user ratings of the resorts for factors I deemed non-participatory on OntheSnow.com.
ii. Perused the individual resort websites myself, checking out accommodations, other activities on the mountains and other services offered by the resort.
iii. Read the descriptions of lodging etc… on the White Book for each resort

After assessing the non-participatory factors for the resorts, I rated them on a scale of 1 – 5 with 5 being the "most plush". I did this before I collected any of the other data for the other variables so I would not be influenced in my rating. Most notably I tried not to look at the price of a lift ticket, even though this did occasionally pop up during my research. The process turned out to be rather inefficient, as I ended up rating about 10 resorts that I later had to drop from the regression because of insufficient data for other variables. My ratings should be fairly unbiased, but in the end they still have an aspect of human error to them and should be analyzed accordingly.

VI. Specifying the Model

1) The Independent Variables and how they should be measured:

* all chosen because of reasons mentioned in Section V.A.

a) Dependent: (Y) Price of an Adult Lift Ticket (dollars)
b) (SNW) Snowfall (inches/year)
c) (VRT) Total Vertical Drop (feet)
d) (ACR) Total Skiable Acres (acres)
e) (PLS) Plush Rating (1 – 5 with 5 being the highest)
f) (AAS) Total Acres of Artificial Snowmaking (acres)
g) (LFT) Lift Capacity (people/hour)

2) The Functional Form of the Variables:

a) Functional forms for Y, β_1SNW, β_2VRT, β_4PLS, β_5AAS, β_6LFT are all linear to the best of my knowledge.

b) Functional form for β_3ACR is hypothesized as being semi-log right → β_3 lnACR for reasons mentioned in the description of this variable in Section V.A.

Initial Model:

$$\hat{Y}_i = \hat{\beta}_o + \hat{\beta}_1(SNW)_i + \hat{\beta}_2(VRT)_i + \hat{\beta}_3(lnACR)_i + \hat{\beta}_4(PLS)_i + \hat{\beta}_5(AAS)_i + \hat{\beta}_6(LFT)_i + e$$

VII. First Regression

* First Regression results from Eviews can be referenced in **Appendix 2**

Results of Regression 1:

$$\hat{Y}_i = 0.753633 + 0.012933(SNW)_i - 0.000381(VRT)_i + 2.675122(lnACR)_i + 6.781682(PLS)_i + 0.017640(AAS)_i - 0.0000031(LFT)_i$$

(0.011762) (0.001842) (1.797669) (1.371713) (0.007906) (0.000143)

t = 1.099600 t = –0.207057 t = 1.488107 t = 4.943951 t = 2.231213 t = –0.216392

N = 40 $\bar{R}^2$ = 0.840645

Analysis of Regression 1:

There are some serious problems with the results of the first regression run. The most alarming problems are the two unexpected negative signs on the coefficients for Vertical Drop and Lift Capacity. The estimated equation is therefore implying that a one foot increase in the Vertical Drop is resulting in a .000381 decrease in the lift ticket price for the average Ski Resort in this sampling. It is also implying that an increased capacity of 1 more person/hour for a ski resort in terms of Lift Capacity is resulting in a .0000031 decrease in the Lift Ticket Price for that resort. Although both of these impacts are miniscule and the respective t-scores are both low at t = –0.207057 and t = –0.216392, the unexpected signs cause some worry because of the strong theoretical backing of variables increasing the price of the ticket. Out of the full list of problems facing a regression:

1. **Omitted Variable**
2. Irrelevant Variable
3. **Incorrect Functional Form**
4. Multicollinearity
5. Serial Correlation
6. Heteroskedasticity

I looked at Omitted Variables and played around the idea that I was using an Incorrect Functional Form or had specified my variables incorrectly. I did this before running any tests for irrelevant variables, multicollinearity, serial correlation or heteroskedasticity because only (1) Omitted Variables and (3) Incorrect Functional Form could have caused the unexpected negative signs on VRT and LFT, because these problems cause bias in the coefficient estimates of the included variables.

My first response was to examine my full list of variables to determine if I had omitted some variable that was positively correlated with either LFT or VRT and had an expected negative impact on the price of Lift Tickets or alternatively was negatively correlated with either LFT or VRT and had an expected positive impact on the price of Lift Tickets. This would have created a negative bias on LFT or VRT. These four options mathematically this looks like:

$$\text{Expected bias in } \overset{(-)}{\beta VRT} = \overset{(+)}{\beta(\text{Variable Omitted})} \times \overset{(-)}{f(r_{vrt,ov})}$$

$$\text{Expected bias in } \overset{(-)}{\beta VRT} = \overset{(-)}{\beta(\text{Variable Omitted})} \times \overset{(+)}{f(r_{vrt,ov})}$$

$$\text{Expected bias in } \overset{(-)}{\beta LFT} = \overset{(+)}{\beta(\text{Variable Omitted})} \times \overset{(-)}{f(r_{lft,ov})}$$

$$\text{Expected bias in } \overset{(-)}{\beta LFT} = \overset{(-)}{\beta(\text{Variable Omitted})} \times \overset{(+)}{f(r_{lft,ov})}$$

Out of full list of variables that I had thought of previously, but omitted from the equation due to varying degrees of theoretical validity and data availability [(SKI) (CMP) (DVS) (BCA) (POP) (RUN) (CHR) (SUN) (SKL) (TMP) (AGE) (FST) (ADV) (GRM) (DMF) (PRK) (RCE) (PAT) (WRK) (HRS)] I could not determine a single variable that was correlated strongly (either negatively or positively) with either VRT or LFT that would also have the effect of causing a negative bias in either of these two variables OR had already been explained by one of the variables or another variable already in the equation.

For example, the omitted variable CHR describes the total number of lifts at a resort, which would be positively correlated with LFT, but also has a theoretically positive impact on the price of the ticket, resulting in a positive bias:

$$(+) \times (+) = (+)$$

CHR is also basically explained already by LFT, so adding it would probably cause severe multicollinearity between the variables. I went back to hypothesizing more variables that could have possibly caused the negatives biases, but eventually moved on to examining the variables that I had included in the initial regression.

I looked specifically at the two variables showing the unexpected negative signs – VRT and LFT. I came to the conclusion that I had misinterpreted how these variables actually increased the lift ticket price while reexamining my definition for each. I realized that LFT, which means *the number of people who can ride up the mountain on a lift in one hour at full capacity* is not a conclusive variable. A lift capacity only really determines how many people can get up the mountain in an hour, but what really matters is the efficiency of this for the mountain. In order to see the true capability of a mountains lift system, mountain stats need to be incorporated into the variable itself. For example, a small mountain and a large mountain could have identical lift capacities, but the large mountain is transporting people a greater distance at the same speed, and therefore have a better lift system. I realized that LFT needed to incorporate vertical drop, so I devised a new variable:

- (NewLFT) New Lift Capacity (Lift Capacity/Vertical Drop)
  - Expected Sign of the Coefficient: (+)

I also dropped VRT and LFT from the equation, because it no longer made sense theoretically to include them when they were already included in another variable. Also, VRT and lnACR had a correlation coefficient of 0.74901 anyways which is fairly high, and should come as no surprise since they both describe to some extent the size of the resort. The new equation looked like:

$$\hat{Y}_i = \hat{\beta}_o + \hat{\beta}_1(SNW)_i + \hat{\beta}_2(lnACR)_i + \hat{\beta}_3(PLS)_i + \hat{\beta}_4(AAS)_i + \hat{\beta}_5(NewLFT)_i + e$$

After determining this new equation and assembling the data for NewLFT which was easy since I had already collected data for both aspects of it, I was ready to run my second regression.

VII. Second Regression

* Second Regression results from Eviews can be referenced in **Appendix 3**

Results of Regression 2:

$$\hat{Y}_i = 2.271316 + 0.013964(SNW)_i + 2.287808(lnACR)_i + 6.344445(PLS)_i + 0.015763(AAS)_i + 0.201469(NewLFT)_i$$

(0.011348)	(1.570625)	(1.183390)	(0.007807)	(0.408775)
t = 1.230515	t = 1.456623	t = 5.361247	t = 2.019122	t = 0.492861

N = 40 $\bar{R}^2$ = 0.845962

Analysis of Regression 2:

The change to NewLFT and the removal of VRT and LFT as a specification change seems to have worked, and all the variables now have the expected positive coefficients. The $\bar{R}^2$ is a solid 0.84592, but some of the t-scores are still quite low. Now that the equation does not have any unexpected coefficients for the $\hat{\beta}$'s and is more 'workable' it is time to review the full list of econometric diseases to determine the strength of Regression 2:

1. Omitted Variable Bias: The omission of a relevant independent variable.

Although there is always possibility of a relevant variable still being omitted, my regression is not exhibiting signs of omitted variable bias, after the specification change from Regression 1. The main reason for my conclusion is that all the expected signs for the coefficients are in the direction that I hypothesized. The main sign of an omitted variable bias is the coefficients of the included X's are in the wrong direction, and in this case, they are not. Another detection method is that there is a surprisingly poor fit for the equation. This is also not the case, with an Adjusted R^2 of 0.84592. Finally, all my high theory variables are included in the equation to some level.

Conclusion: Low possibility of omitted variable bias and no remedies should be performed.

2. Irrelevant Variable: The inclusion of a variable that does not belong in the equation.

An irrelevant variable would result in decreased precision in the form of higher standard errors and lower t-scores. There is the possibility of an irrelevant variable, but my initial opinion is no since the five variables in the equation were chosen for high theory. That being said, the t-score for NewLFT is very low at t = 0.492861. There is a chance that NewLFT is irrelevant so it is important to look at the four specification criteria in determining whether a variable is irrelevant.

1. **Theory:** The theory for NewLFT is very strong. This has already been discussed in the analysis of Regression 1 and in the descriptions of LFT and VRT in the independent variable section V.A.

2. **t-test on $\hat{\boldsymbol{\beta}}$:** for a one-sided test at a 5% level of significance

   Hypotheses:
   $H_o = \hat{\beta}_5 \leq 0$
   $H_A = \hat{\beta}_5 > 0$

   Degrees of Freedom = N – K – 1 = 40 – 5 – 1 = 34
   $t_c = 1.6892$ (Stockburger 2011)
   $t_k = 0.492861$

   $|t_k| < t_c \quad \rightarrow \quad 0.492861 < 1.6892$

   So we fail to reject H_o at the 5% level, despite the sign being the same as H_A.

3. $\bar{\boldsymbol{R}}^{\mathbf{2}}$: To test whether $\bar{R}^2$ increased or decreased when the irrelevant variable was dropped I ran a 3rd regression without NewLFT included that looked like:

   $$\hat{Y}_i = \hat{\beta}_o + \hat{\beta}_1(SNW)_i + \hat{\beta}_2(lnACR)_i + \hat{\beta}_3(PLS)_i + \hat{\beta}_4(AAS)_i + e$$

   * This Fourth Regression can be referenced in **Appendix 4**

   The new $\bar{R}^2$ with NewLFT dropped is 0.849294, so the $\bar{R}^2$ increased by roughly 0.005

4. **Impact on other coefficients if X is dropped**:

Variables:	Previous Coefficients:	New Coefficients:
SNW	(0.013964)	(0.013544)
lnACR	(2.287808)	(2.405532)
PLS	(6.344445)	(6.537838)
AAS	(0.015763)	(0.017010)

The results of the 4 criteria for irrelevant variables, is in this case, not entirely conclusive. We could not reject the null hypothesis for a one-sided t-test at a 5% level of significance, and the removal of the variable did not cause the other coefficients to change significantly. Both of these results imply the variable NewLFT is irrelevant. $\bar{R}^2$ increased, but only by 0.005, which also implies the variable does not belong but it is not very convincing. Strong theories, on the other hand, call for the inclusion of NewLFT. Theory is the most important factor in determining whether the variable is irrelevant or not, so in this case I would keep the variable in the equation.

* It should also be noted that the t-scores for SNW and lnACR would have also both failed to reject the respective null hypotheses for both of these variables, but only by a little bit. I determined it extraneous to complete the four specification criteria and run new regressions for each of these variables though, because the theory is solid for both of their inclusions, and I would have come to the same conclusion that I did for NewLFT.

Conclusion: Possibility of irrelevant variables, but no remedies should be performed

3. Incorrect Functional Form: The Functional Form is inappropriate.

Although an incorrect functional form is possible, my regression does not exhibit the signs of this. The signs of an incorrect functional form are biased estimates, a poor fit and a difficult interpretation. In my case, the signs on the coefficients are in the expected direction, the fit is solid with $\bar{R}^2$ = 0.84592, and the interpretation is not difficult.

Conclusion: Low Possibility of an incorrect functional form and no remedies should be performed.

4. Multicollinearity: Some of the independent variables are (imperfectly) correlated.

Multicollinearity would lead to unreliable estimates of the separate effects of the X's, high standard errors and low t-scores. One way to detect multicollinearity is to examine the simple correlation coefficients between the explanatory variables. The simple correlation coefficients can be seen in **Appendix 3**, in the simple correlation coefficients table. The highest correlation between two variables in the table is 0.696700 which is between PLS and NewLFT. The somewhat high correlation is expected between these two variables though. PLS basically describes how new and fancy a resort is with its non-participatory factors. Although a lift is not non-participatory variable, it makes sense the fancier a resort tends to be, the more high speed and expensive lifts they have. The correlation coefficient of 0.696700 is not alarming though. According to *Using Econometrics, A Practical Guide*, "Some researchers pick an arbitrary number, such as 0.80, and become concerned about multicollinearity any time the absolute value of the simple correlation coefficient exceeds that value." Since I am a novice econometrician, I will take this arbitrary value of 0.80 as an indicator of severe multicollinearity if the simple correlation exceeds it. In my case, it does not because $0.696700 < 0.80$ implying there is a low possibility of severe multicollinearity

Then again, there is a major limitation to this method in that the lack of a high simple correlation coefficient between two variables doesn't erase the possibility of a group of independent variables acting together to cause multicollinearity. Due to this, I will run a VIF test for each of the independent variables to see if there is any multicollinearity.

An example of what the VIF test looks like for variable NewLFT is shown below:

$$\widehat{NewLFT} = \hat{\alpha}_1 + \hat{\alpha}_2 SNW + \hat{\alpha}_3 lnACR + \hat{\alpha}_4 PLS + \hat{\alpha}_5 AAS + v$$

*Results of the VIF tests can be found in **Appendix 5**

The second step of the VIF test is shown below:

$$VIF(\hat{\beta}_i) = \frac{1}{(1 - R_i^2)}$$

The resulting R^2's from each independent variable acting as a dependent variable are formulated into the second part of the VIF test below:

SNW:

$$VIF(\hat{\beta}_i) = \frac{1}{(1 - 0.452052)} = 1.82499$$

lnACR:

$$VIF(\hat{\beta}_i) = \frac{1}{(1 - 0.664098)} = 2.97706$$

PLS:

$$VIF(\hat{\beta}_i) = \frac{1}{(1 - 0.664578)} = 2.98131$$

AAS:

$$VIF(\hat{\beta}_i) = \frac{1}{(1 - 0.548686)} = 2.21575$$

NewLFT:

$$VIF(\hat{\beta}_i) = \frac{1}{(1 - 0.557186)} = 2.25828$$

According to *Using Econometrics, A Practical Guide*, "…a common rule of thumb is that if $VIF(\hat{\beta}_i)$ is greater than 5, the multicollinearity is severe. In my case none of the $VIF(\hat{\beta}_i)$ are greater than 5, so I can conclude the possibility is low that severe multicollinearity exists in the equation.

Conclusion: Low Possibility of multicollinearity and no remedies should be performed.

5. Serial Correlation: Observations of the error term are correlated, as in: $\varepsilon_t = p\ \varepsilon_{t-1} + u_t$

Although my model is a cross sectional model, and not a time series, I will still test for serial correlation. Serial correlation could result with OLS no longer being the minimum variance estimator and with unreliable hypothesis testing.

To test for serial correlation I will use the Durbin-Watson test using the d statistic, using a 5% one sided level of significance test.

$$d = \sum_{2}^{T}(e_t - e_{t-1})^2 / \sum_{1}^{T} e_t^2 = 2.331466\ \ (from\ Regression\ 2, Appendix\ 3)$$

Sample Size, N = 40
K = 5 observations

Consulted Statistical Tables B-4, B-5, or B-6 in Appendix B of *Using Econometrics, A Practical Guide* to find upper critical d value, d_u, and the lower critical d value, d_L, respectively.

d_u value – 1.79
d_L value – 1.23

Given:
H_o: $\rho \leq 0$ $(no\ positive\ serial\ correlation)$
$H_{A:}$ $\rho > 0$ $(positive\ serial\ correlation)$

And the appropriate decision rule:
If $d < d_L$ $\quad$ $Reject\ H_o$
If $d > d_U$ $\quad$ $Do\ not\ reject\ H_o$
If $d_L \leq d \leq d_U$ $\quad$ $Inconclusive$

Decision Rule calculated from with our d-statistic: $2.331466 > 1.79$, so therefore we fail to reject the null hypothesis meaning there is evidence of **no** positive Serial Correlation.

Conclusion: Low Possibility of Serial Correlation and no remedies should be performed.

6. Heteroskedasticity: The variance of the error term is not constant for all observations as in: $VAR(\varepsilon_i) = \sigma^2 Z_i^2$

If Heteroskedasticity is present there will not be a bias in the $\hat{\beta}$'s, but OLS will no longer be the minimum variance estimator and hypothesis testing will be unreliable. To test for Heteroskedasticity I will use a Park Test. Shown below is the equation for running the Park Test:

$$\ln(e_i^2) = \alpha_o + \alpha_1 lnZ_i + u_1$$

$e_i = the\ residual\ from\ the\ ith\ observation$
$Z_i = your\ best\ choice\ as\ to\ the\ possibile\ proportionality\ factor\ (Z)$
$u_i = a\ classical\ (homoskedastic)\ error\ term$

For my proportionality factor Z, I chose lnNewLFT because it is a good indicator of size which is important for a cross sectional model. lnACR was also a good choice, but already in the natural log form so I left it alone. I ran the equation (ls ln(Resid)^2 c lnNewLFT on Eviews:

*The results of the this equation can be found in **Appendix 6**

Results of Eviews Park Test:

$$\ln(e_i^2) = 7.073003 + 0.452411(lnNewLFT)_i$$

(0.562035)
t = 0.804951
N = 40 $R^2 = 0.016765$

Next I tested the significance of $\hat{\alpha}_1$ with hypotheses using a two-tailed 1% test:
$H_o: \alpha_1 = 0$ (Homoskedasticity)
$H_A: \alpha_1 \neq 0$ (Heteroskedasticity)

Degrees of freedom = 40 – 1 – 1 = 38
$t_c = 2.7132$ (Stockburger 2011)
$t_k = 0.804951$

$|0.804951| < 2.7132$, so we fail to reject the null hypothesis that there is Homoskedasticity and we must assume there is Heteroskedasticity, or might I say Hetero-ski-dasticity, present in the equation.

Conclusion: There is a high possibility for Heteroskedasticity and Heteroskedasticity-Corrected Standard Errors should be performed as a remedy.

* I chose to use Heteroskedasticity-Corrected Standard Errors instead of redefining the variables, because I had already re-specified the equation earlier, and could think of no other theoretical underpinnings for another change.

Heteroskedasticity-Corrected Standard Errors adjust the estimation of the $SE(\hat{\beta})$'s for heteroskedasticity while still using the OLS estimates of the slope coefficients.

*The results of the HCSE equation can be found in **Appendix 7**

Results of the HCSE Equation:

$$\hat{Y}_i = 2.271316 + 0.013964(SNW)_i + 2.287808(lnACR)_i + 6.344445(PLS)_i + 0.015763(AAS)_i + 0.201469(NewLFT)_i$$

(0.009462)	(1.168632)	(0.985976)	(0.005556)	(0.476373)
t = 1.475751	t = 1.957681	t = 6.434683	t = 2.837043	t = 0.422923

N = 40 $\bar{R}^2 = 0.845962$

The results of the standard errors and respective t-scores from Regression 2 without the HCSE's is shown below:

(0.011348)	(1.570625)	(1.183390)	(0.007807)	(0.408775)
t = 1.230515	t = 1.456623	t = 5.361247	t = 2.019122	t = 0.492861

N = 40 $\bar{R}^2 = 0.845962$

The standard errors increased for every variable with the addition of HCSE's, but I don't believe they affected the t-scores enough for me to re-question whether any of the variables are irrelevant. This is because the critical t-value for this regression, assuming a 5% one-sided test with 34 degrees of freedom is $t_c = 1.6892$, and I had previously determined that the coefficients for SNW, lnACR, and NewLFT would all have failed to reject their respective null hypotheses. PLS and AAS on the other hand were able to reject the null hypotheses in favor of the alternative. With the new HCSE's the same conclusions regarding the decision rule for these t-tests would be reached, because the t-scores weren't negatively affected enough for PLS and AAS to fail to reject the null hypotheses, while SNW, lnACR and NewLFT would all fail to reject the null hypothesis again. In a sense, the t-scores decreased, but not enough to play into my decision about t-scores for determining if the variable was irrelevant, while the theory behind each variable has not changed. Thus, with the HCSE's I would not change any of the previous conclusions I had reached regarding the other five tests for diseases in my regression.

IX. Final Documented Regression

* Same as the HCSE Eviews equation that can be fully referenced in **Appendix 7**

$$\hat{Y}_i = 2.271316 + 0.013964(SNW)_i + 2.287808(lnACR)_i + 6.344445(PLS)_i + 0.015763(AAS)_i + 0.201469(NewLFT)_i$$

(0.009462)	(1.168632)	(0.985976)	(0.005556)	(0.476373)
t = 1.475751	t = 1.957681	t = 6.434683	t = 2.837043	t = 0.422923

N = 40 $\bar{R}^2 = 0.845962$

X. Conclusion

The results of my first regression were not what I expected. Despite an initial high $\bar{R}^2$ of 0.840645, two of my variables VRT and LFT had unexpected signs for their coefficients. After an arduous process of testing for and theorizing omitted variables, I realized I possibly had a specification error so I went back and changed the specification of the equation to incorporate a new variable NewLFT and omitted both VRT and LFT. The addition of the new variable seemed to correct for the bias I had been encountering and my Second Regression showed no signs of significant omitted variable bias. I finally had an equation that I could work with, and I went on to test the Second Regression for the six different diseases that we have studied that can affect a regression. manipulation

Examining the equation for omitted variable bias, an incorrect functional form, multicollinearity and serial correlation I came to the conclusion that there was a low possibility for any of these diseases. The actual determination of this is explained thoroughly in Section VII. When I tested for irrelevant variables, the results were not so clear. Specifically, the specification criteria were implying to *keep and drop* the variable NewLFT. I ultimately decided to keep the variable because of theory. It should be noted though, that three of my variables – NewLFT, SNW and lnACR all had low t-scores, but I kept them all because of theory. My final test for Heteroskedasticity using the Park Test, showed that I had a high possibility of Heteroskedasticity within my equation, so I ran it again with HCSE's. The result of this last regression, is my final regression.

Granted there are still significant breakdowns within my equation that should be taken into account when analyzing the entire regression. First of all, there could still be an omitted variable from the equation. It is almost unrealistic that there is not considering the list of variables that I did not include and those I could not find data for – (SKI) (CMP) (DVS) (BCA) (POP) (RUN) (CHR) (SUN) (SKL) (TMP) (AGE) (FST) (ADV) (GRM) (DMF) (PRK) (COR) (RCE) (PAT) (WRK) (HRS)]. Especially two of these variables, (ADV) advertising and (COR) Owned by a Larger Corporation, which I could not find data for, might be affecting lift ticket price and are not included in the equation. Resorts spend a lot of money on advertising and it would be careless not to take this into account when thinking about lift ticket pricing. To some extent the Plush Rating might have served as a proxy for Advertising, but Advertising should not be ignored. In reality, a lot of ski resorts are also owned by larger parent companies. The data was available for determining the ownership, but the pricing schemes for these large parent companies was not available. The different parent companies surely price their resorts differently and it probably is largely dependent on what other resorts they own in the area and across the United States. This should also be taken into account and recognized that it is not part of the equation. Finally, the equation certainly has some human error to it. This human error was discussed in Section VI in the determination of the plush rating. Taking all these potential drawbacks into account, and the statistical drawbacks within the regression (such as low t-scores) we can analyze the effects of the different variables for my final regression:

$$\hat{Y}_i = 2.271316 + 0.013964(SNW)_i + 2.287808(lnACR)_i + 6.344445(PLS)_i + 0.015763(AAS)_i + 0.201469(NewLFT)_i$$

- An inch more of snow/year results in a 0.013964 increase in the ticket price in dollars.
- The increase of one natural log of total skiable acres, increases the lift ticket price by 2.287808 dollars.
- A one unit increase in the plush rating, results in a 6.344445 dollar increase in ticket price.
- One more acre of artificial snowmaking results in a 0.015763 dollar increase in ticket price.
- A one unit increase in Lift Capacity/Vertical Drop results in a 0.201469 increase in ticket price.

The most statistically significant variable with the highest t-score was PLS, which reinforces Koslow's (2006) and Falk's (2006) theory that non-participatory elements are very important in the determination of lift-ticket prices.

If adapted, this regression can hopefully be used by both consumers looking to assess the true value of a ski resort they plan on visiting, to see if it overvalued or undervalued with respect to the resort characteristics mentioned in the regression. The model could also be used for resorts themselves to determine an accurate price for their own lift tickets, considering their particular features.

XI. Works Referenced

Barro, Robert J., and Paul M. Romer. "Ski-Lift Pricing, with Applications to Labor and Other Markets." *The American Economic Review, Vol. 77, No. 5*. American Economic Association, Dec. 1987. Web. 03 Dec. 2011. <http://0-www.jstor.org.oasys.lib.oxy.edu/stable/pdfplus/1810214.pdf?acceptTC=true>.

Enzel, Robert G. "The White Book of Ski Areas." *Inter-Ski, Inc*. 2007. Web. 03 Dec. 2011. <http://www.inter-ski.com/whitebook30/whitebook.htm>.

Falk, Martin. "A Hedonic Price Model for Ski Lift Tickets." *WIFO*. Austrian Institute of Economic Research, Sept. 2006. Web. <http://www.webmeets.com/files/papers/EARIE/2007/495/gpaper12_vfs.pdf>.

Korkki, Phyllis. "Ski Resorts Thrive, Despite Recession Pain." *The New York Times*. 20 Feb. 2010. Web. 03 Dec. 2011. <http://www.nytimes.com/2010/02/21/business/economy/21count.html?adxnnl=1>.

Koslow, Michael. "The Effects of Non-Participatory Characteristics: The Case of Lift-Tickets." *Issues in Political Economy, Vol. 15*. Elon University, Aug. 2006. Web. <http://org.elon.edu/ipe/koslow.pdf>.

Morey, Edward R. "The Choice of Ski Areas: Estimation of a Generalized CES Preference Ordering with Characteristics." *MIT Press*. The Review of Economics and Statistics, Vol. 66, No. 4, Nov. 1984. Web. 03 Dec. 2011. <http://0-www.jstor.org.oasys.lib.oxy.edu/stable/pdfplus/1935982.pdf>.

Mulligan, James G., and Emmanuel Llinares. "Market Segmentation and the Diffusion of Quality-Enhancing Innovations: The Case of Downhill Skiing." *MIT Press*. The Review of Economics and Statistics, Vol. 85, No. 3, Aug. 2003. Web. 03 Dec. 2011. <http://0-www.jstor.org.oasys.lib.oxy.edu/stable/pdfplus/3211692.pdf>.

"Resorts." *On the Snow*. Web. 3 Dec. 2011. <http://www.onthesnow.com/>.

Stockburger, David W. "Critical T Values for the T-distribution." *Psychological Statistics at Missouri State University*. Web. 05 Dec. 2011. <http://www.psychstat.missouristate.edu/introbook/tdist.htm>.

Studenmund, A.H. *Using Econometrics: a Practical Guide*. 6th ed. Boston: Addison-Wesley, 2011. Print.

XII. Glossary of Variables

- (SNW) Snowfall (inches/year
- (VRT) Total Vertical Drop (feet
- (ACR) Total Skiable Acres (Acres
- (PLS) Plush Rating (1 – 5 with 5 being the highest)
- (AAS) Total Acres of Artificial Snowmaking (Acres)
- (LFT) Lift Capacity (people/hour)
- (NewLFT) New Lift Capacity (Lift Capacity/Vertical Drop)
- (SKI) Ski Only Resort (Dummy Variable → 0 – otherwise, 1 – ski only)
- (CMP) Number of other resorts in the state (number)
- (DVS) Diversified Runs/Difficulty (Dummy Variable → 0 – otherwise, 1 – At minimum 15% beginner, 35% Intermediate and 35% Advanced)
- (BCA) Variable for Backcountry Access (Dummy → 0 – otherwise, 1 – has backcountry access)
- (POP) Proximity to large population centers (Dummy → 0 – not within 30 miles, 1 – within 30 miles)
- (RUN) Number of Runs (total number of runs)
- (CHR) Number of Chairlifts (total number of chairlifts
- (SUN) Days of Sunshine (days out of average season)
- (SKL) Ski school (Dummy Variable → 0 – otherwise, 1 – has a school)
- (TMP) Temperature (average temperature during months of season)
- (AGE) How many Ages the Pass accounts for (range of total ages)
- (FST) Located on US Forest Land (Dummy Variable → 0 – otherwise, 1 – yes)
- (ADV) Advertising (dollars/year)

- ❖ (GRM) Number of Ski Grooming machines the Resort Owns (number of groomers)
- ❖ (DMF) Variable for Distinct Mountain Feature (Dummy variable → 0 – otherwise, 1 – at least one)
- ❖ (COR) Owned by a Larger Corporation (Dummy variable → 0 – otherwise, 1 – yes)
- ❖ (PRK) Terrain Park (Dummy variable → 0 – otherwise, 1 – at least one)
- ❖ (RCE) Race course (Dummy variable → 0 – otherwise, 1 – at least one)
- ❖ (PAT) Ski Patrol (number of permanent staff)
- ❖ (WRK) Lift workers (number of permanent staff)
- ❖ (HRS) Hours of Operation (hours/day)