HEDONIC PRICE ANALYSIS OF THE INTERNET RECREATIONAL EQUINE MARKET

by

JENNIFER FREEBORN

B.S., Colorado State University, 2006

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Agricultural Economics
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2009

Approved by:

Major Professor
Dr. Ted Schroeder

Copyright

Abstract

Hedonic pricing models have long been used to evaluate prices in high-end segments of the equine industry. However, the lower end markets, including most of the recreational and pleasure horses, have yet to be studied in the economic literature. This study evaluates characteristics affecting the price of recreational horses advertised online, and provides a framework for future market studies on various segments of the equine industry.

Data were collected in both the spring and fall of 2008, in addition to a survey being conducted in the fall to collect more accurate pricing information. Three different functional forms were used and their outcomes discussed to determine which one best fit the data. Results show that the semi-log functional form appeared to best fit the data. Characteristics that consistently showed statistical significance included the horse being advertised using a photo ad, the age of the horse, and whether or not the horse was registered. Gender variables and the state which the horse was sold from showed statistical significance in most of the models; although the variables denoting breed were statistically significant as a group, no model consistently found statistical significance in any of the variables individually. Color characteristics did not demonstrate statistical significance consistently in any model.

Finally, suggestions for future research are discussed. Data issues could be avoided with larger or more specific data sets; various data sources could be examined or created such as live equine auctions; regions could be examined by show or rodeo circuit instead of by state. There are political issues in the industry that need to be addressed, but a lack of available data needs to be examined and corrected before many issues can be thoroughly examined. The equine industry is often overlooked in economic literature and is a multi-billion dollar agricultural industry which deals with legislative and taxation issues just like the rest of the agricultural world and is deserving of attention.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would first like to acknowledge my entire committee for their help through my entire career at Kansas State University. Dr. Ted Schroeder has dealt with me through many difficult days and great days, and I appreciate his patience with me through it all. Dr. Dhuyvetter has provided me with assistance in technical knowledge in class work and articles that I never could have learned without his gracious help. Dr. Hikaru Peterson has been an excellent instructor and mentor for me ever since I took her marketing class. You have all provided me with invaluable assistance, and I cannot thank you enough.

I would also like to acknowledge the help of my colleagues, especially Chris Crosby for his technical help with Excel and SAS and my cubicle buddy Robin Slattery for her help in getting me through this final semester and getting me out of the office. The entire 4th floor, both the students who are newly arrived and those who have left to start new lives, has been a great help to me and provided me with experiences I will never forget.

Finally, I would like to acknowledge the help of my family and friends. Without them, I would not be who I am today or where I am today. I am forever grateful.

Dedication

I would like to dedicate this to all of the people throughout my lifetime that have helped me develop my deep love and interest in the equine world. Without their coaching and support, I would never have looked into an arena that is likely to become my career.

CHAPTER 1 - Introduction and Objectives

Introduction

The equine industry, both in the United States and world-wide, is one that is unique in many aspects and characteristics, making it difficult to describe, classify, and research. The equine industry combines the world of companion animals with the livestock industry – many owners consider their horses to be their pets, while others use horses for business. In recent years, the equine industry has come under increased national scrutiny as issues such as infectious diseases and the National Animal Identification System have highlighted health and disease tracking concerns, in addition to equine slaughter, transportation, and abandonment being hot topics on the animal welfare front. Legislation affecting the industry has been brought forward and has become highly controversial, in no small part due to the lack of information and understanding of the equine industry as a whole. The statistical information available on the industry varies widely across sources and years. No estimates currently exist on supply or demand elasticities in the general equine marketplace, only in Thoroughbred yearlings, limiting the ability to create structural models of the industry to understand how the industry may react to legislative changes. This lack of reliable data limits the ability to research a multi-billion dollar industry that employs many people and contributes substantially to the economies of many states and counties. The following paragraphs contain background information on the equine industry and set the stage for the objectives of this study.

U.S. Equine Population

The actual number of equids in the United States is difficult to ascertain, due to the fact that the United States Department of Agriculture (USDA) Census, conducted every five years, only applies to "on-farm" livestock. As stated in the National Animal Health Monitoring System (NAHMS) Equine 2005 project, Part II (p. 18), "The U.S. equine population is difficult to enumerate because of the diversity of the equine industry, the geographic breadth of the equine population, and the suburban areas not included in the traditional livestock enumeration." Many horses are not located on "farms," as defined by the USDA as a property that can or has produced more than $1,000 of agricultural goods on an annual basisor (as of 1987) has five or more equids owned by the same owner, not including boarding stables (where horses are owned by multiple individuals) or commercial enterprises such as racetracks. The Equine 2005 Part II Booklet states (p. 7), “There is no accurate estimate of the current total number of equids in the United States because the number of equids on nonfarm operations does not exist.” This same source suggests that the on-farm estimation may only constitute 50-60% of the total equine population.

A summary of recent equine population estimates can be seen in Table 1. The only USDA estimates for the total number of equids in the United States come from the National Agricultural Statistics Service (NASS) Surveys in 1998-1999. The January 1, 1998 estimate for horse numbers by USDA was 5.25 million head, with 3.20 million on farms and 2.05 million on

Table 1: Equine Population Estimates

Year	Group	Estimate (mil. head)	Notes
1986	AHC*	5.25	All equids
1986	AVMA**	6.6	"Pet" equids only
1991	AVMA	4.9	"Pet" equids only
1996	AHC	6.9	All equids
1996	AVMA	4	"Pet" equids only
1997	USDA Census	3.02	On farms only
1998	USDA-NASS	5.25	All equids
1999	USDA-NASS	5.35	All equids
2001	AVMA	5.1	"Pet" equids only
2002	USDA Census	3.64	On farms only
2003	AHC	9.2	All equids, published 2005
2007	AQHA***	2.9	AQHA Registered Horses
2007	USDA Business Plan	5.8	All equids

*American Horse Council

**American Veterinary Medical Association

***American Quarter Horse Association

non-farms. If we assume the same percentage of horses are on or off farms in 2002 as in 1998, then we can take the estimated 3.64 million head of horses on farms as 60.95% of the total, and come up with a total number of horses in the U.S. in 2002 as 5.97 million head (which would mean there were approximately 2.33 million off-farm horses in 2002). The most recent USDA estimate comes from the USDA-Animal and Plant Health Inspection Service (APHIS) *Business Plan to Advance Animal Disease Traceability* which gives the June 2007 equine population estimate as 5.8 million horses in approximately 570,000 locations. The most recent American

Horse Council (AHC) commissioned survey, performed by Deloitte for the calendar year of 2003, published in 2005, suggested there were 9.2 million horses in the U.S. The American Veterinarian Medical Association (AVMA) horse number estimates were obtained from a 2007 Oklahoma State University Extension Publication (Freeman, 2007). The AVMA estimates horses that are specifically owned by "households" which do not include those owned by ranches, farms, or other operations – they specifically estimated horses that were treated by their owners more as "pets." The conclusion of the extension publication states that all the sources seem to agree that the U.S. horse population has had an annual growth rate of between 3 and 5% over the last decade, though the AVMA numbers indicated a decline prior to the last decade (Freeman, 2007). The AVMA studies seem to indicate there was a reduction in the number of "pet" equids between 1986 and 1996, and that the number was increasing again between 1996 and 2001. The USDA information and the AHC studies both indicate upward trends in horse numbers over the entire time period included in this table.

In the 2003 AHC survey, Quarter Horses are the most common breed of horses, with 3,288,302 in this survey, and the only other breed broken out was Thoroughbreds, with 1,291,807 horses. Other horses included other breed's registered horses and grade (non-registered, non-pedigreed) horses. The leading registry of the Quarter Horse breed is the American Quarter Horse Association (AQHA). The AQHA 2007 annual report states that there were 2,859,851 Quarter Horses in the United States. This total is down by just over 28,000 from the previous year, and new registrations were also down by over 26,000. This is the first time in recent history that AQHA registration numbers declined. Most people familiar with the equine industry agree that the market for horses is depressed in 2008, and this may be an explanation for the downturn in AQHA numbers. AQHA also keeps track of transfer numbers, and had a total of 188,907 ownership transfers in 2007, some being within the United States and some internationally (2008). There are 902,453 registered owners of AQHA Quarter Horses in the United States, including horses used for racing, showing, and recreational or other purposes (2008).

Economic Statistics on the U.S. Equine Industry

Data collected and published by the American Horse Council Foundation shows that the equine industry directly contributed approximately $38.8 billion to the U.S. economy in 2004;

including indirect impacts, the total economic impact jumped to $101.5 billion. Of this total, $26.1 billion was contributed by horse racing; the show horse industry accounted for a $28.8 billion impact, and recreational horses, which accounted for about 42% of U.S. horse numbers, had the largest share at $32.0 billion dollars. The recreational horse industry segment had the largest contribution to the direct impacts at $11.9 billion as well. In addition, the equine industry contributed $1.9 billion in annual taxes across all levels of government according to the American Horse Council Foundation (2005a).

The AHC also collected data on participation and employment supplied or created by the equine industry, and stated that approximately 4.66 million Americans were involved in the industry as owners, employees, or volunteers (2005). Just over 453,000 jobs were directly in the horse industry and another 1.41 million are related to or generated by the horse industry, including 435,100 jobs, or over 30%, in by the recreational horse industry. Table 2 shows the jobs generated by the horse industry broken down by category (American Horse Council Foundation 2005a).

Table 2: Employment in the Equine Industry by Segment

Industry Segment	Direct Jobs	% Direct Jobs	Total Jobs	% Total Jobs
Entire Industry	453,612	100.0%	1,411,333	100.0%
Racing	146,625	32.3%	383,826	27.2%
Showing	99,051	21.8%	380,416	27.0%
Recreation	128,324	28.3%	435,082	30.8%
Other	79,612	17.6%	212,010	15.0%
*Source: American Horse Council Foundation Economic Impact Survey (2005)				

Further financial information from the most recent AHC survey is also of interest. The majority of horse owners (56%) earned under $75,000 per year in household income, with 16% earning between $75,000 and $100,000, 15% between $100,000 and $150,000, 9% over $150,000, and 4% not reporting. The "average" horse in the industry earned $1,172 and costs the owner $2,882 annually – notably, however, the recreational industry only made $536 per horse at an annual cost of $2,319. The two largest expenses in all segments were "Feed, Bedding, and Grooming Supplies," and "Boarding and Training." Veterinary services cost an average across the industry of $251 annually per horse. AHC's survey also did individual state evaluations on 15 states. The top five states of the 15 breakout states specifically surveyed by the AHC for four different categories are shown in Table 3. Texas, California, and Florida were consistently in the top three states with Kentucky consistently in the top five across the four categories (2005a).

Table 3: Top Five Equine States

Rank	By Total Effect on GDP	By Number of Horses	By Number of Industry Participants	By Total Effect on Full-Time Equivalent Employment
1	California	Texas	Texas	California
2	Texas	California	Florida	Florida
3	Florida	Florida	California	Texas
4	Kentucky	Oklahoma	Kentucky	Kentucky
5	Louisiana	Kentucky	Ohio	Missouri
*Source: American Horse Council Foundation Economic Impact Survey (2005)				

Precise estimates of the average value of horses are difficult to obtain, but census information allows for a rough estimate. The 2002 Census of Agriculture reported that the number of horses and ponies sold from farms was 470,423, in addition to 17,385 mules, burros, and donkeys, for the total value of approximately $1.33 billion. Using these numbers, the average value for the horses sold is $2,724 (USDA, 2002). In the 1998-1999 equine reports, the total equine sales were approximately 539,600 head in 1997 with a value of $1.64 billion for an average of $3,042; in 1998 the equine sales numbers were approximately 557,600 head for $1.75 billion with an average sale amount of $3,146 (USDA, 1999). These values are summarized in

Table 4: Equine Sales and Average Value

Year	Number Sold (head)	Total Value	Average Value ($/head)
1997	539,600	$1,641,196,000	$3,041.50
1998	557,600	$1,753,996,000	$3,145.62
2002	487,808	$1,328,733,000	$2,723.89
*Sources: USDA Census Data and NASS Equine 1998 Survey			

Table 4. One important point from a financial standpoint is that of U.S. live animal exports, the equine industry consistently exports several times greater value than any other livestock industry. In 2005, the last estimate conducted showed that live equid exports were approximately $461,541,000, with the closest number to that being live poultry exports at $95,522,000 (USDA, 2006b). The equine industry clearly has a major economic impact across many levels of the U.S. economy.

Issues of Concern Regarding Horses

Horses are often in the news; recently, equine contagious diseases and equine welfare have been two major topics of concern internationally. From the NAIS (National Animal Identification System) and other health and tracking regulations, to the attempts to ban horse slaughter in the United States, to Wild Horse and Burro legislation, the industry is rife with controversy and often saddled with restrictive legislation. West Nile Virus in the United States, Avian Influenza (which is a zoonotic or cross-species disease), and Equine Influenza have led to new legislation, impacts to the bloodstock and insurance industries, and in some cases to stock market impacts (Barrett, 2006; Marrs and Mathews, 2003). In 2002, the state of North Dakota alone lost over $3 million in state and personal funds due to West Nile Virus (Ndiva Mongoh et al., 2007). The equine influenza outbreak in Australia last spring, which brought the equestrian world in that nation to a shuddering halt, has yet to be fully comprehended or quantified. Postponement or cancellation of some of the largest races in the thoroughbred industry, complete bans on equine commerce for anywhere from 72 hours to several weeks will have untold impacts on breeders, trainers, show circuits, and even recreational riders (McLennan, 2007; Wainwright and Moore, 2007; Maugeri, 2007; Foster, 2007; Eddy et al., 2007).

The issue of horse slaughter has turned the U.S. horse industry into a battleground over horse welfare versus economic reality, with only one economic study having been published, by North, Bailey, and Ward (2005) since this issue came to the forefront early in the 21st century. A front page article in the Denver Post recently highlighted equine welfare problems already occurring since the January 2007 closures of the Texas slaughterhouses and the September 2007 closure of the Dekalb, IL plant (Booth, 2008). A recent report from the Government Accountability Office (GAO) discusses the problems related to the BLM Wild Horse program, indicating that without changes to the program the number of wild horses kept and cared for in captivity will take too much of the program’s budget for it to function, and that disposal methods of unadoptable horses need to be evaluated and chosen carefully due to animal welfare concerns (GAO, 2008). Little economic research has been conducted on most of these issues, and what research is conducted struggles to find available and reliable data sources.

As mentioned above, one of the major topics of discussion regarding the horse industry in recent years has been horse welfare, especially as related to equine processing (Einhorn, 2008). The humane equine slaughter issue first arose in the early 21st century, when laws were instated and enforced by USDA on the manner that horses could be transported to slaughter plants. These plants already had the same inspection policies and humane treatment policies that we expect of our slaughter industry nationwide. Throughout the debate over the passage and upholding of equine processing bans in the states of Illinois and Texas, and the national introduction of numerous bills regarding horse slaughter and transportation regulations, legislators and members of the equine industry have searched for hard economic facts about the possible impacts of such legislation. The fact is, this research does not exist. One article by North, Bailey, and Ward (2005) exists on the possible direct effects of a nationwide horse processing ban. However, though this was admittedly a start, it does not include the impacts that may spread throughout the entire industry due to an increased supply of horses not sent to slaughter among other factors. Further explaining the impacts on the equine industry from any legislation will require more basic information on the equine industry. Current legislation has been introduced to prevent horses from crossing international borders for processing, an event which has occurred frequently since the three slaughter facilities in the United States are now closed due to the enforcement of a 1949 Texas law and new legislation in the state of Illinois.

The front page of the February 17, 2008 Denver Post contained a detailed article about the increase in neglect and abandonment of horses on the Front Range, including those by some rescue shelters (Booth, 2008). One of Colorado's most prominent papers, this article reported concerns about equine humane rescue facilities being packed full with nowhere for these horses to go (Booth, 2008). Auction yards have reported turning horses away because they literally have no value in the market. Calls have come into the offices of equine organizations reporting horses turned loose in pastures or tied to trailers where they do not belong, without evidence of who the owner was, indicating an increase in abandonment cases. A new coalition of numerous breed organizations was formed to work on a resolution for this issue. The Unwanted Horse Coalition reports that no statistical information is known about unwanted horses in the U.S., only that the current number of unwanted horses exceeds the ability to care for them at rescue facilities, in part due to the high costs of maintaining a healthy equid (Unwanted Horse Coalition, 2008). Thoroughbred racing has also recently seen some welfare issues with new track surfaces, and reports of increasing numbers of breakdowns and a decrease in the number of races they can run annually (Beyer, 2006).

Yet even considering all the mainstream media articles regarding the equine industry, all these economic, social, and political impacts the equine industry brings to the United States and the international community, no one knows the true economic impacts of any of the controversial issues above. In fact, estimates on the number of horses in the U.S. even vary widely and the actual number is impossible to derive from current data – probably not even to within a million head of horses! Economic impact studies on the equine industry are few and far between, mostly centered on thoroughbred racing, betting, and hedonic pricing models of high-end auctions. These studies are focused on the high-end, high-dollar, well-publicized side of the industry, and the lower end "recreational" horses, as the AHC survey terms them, are greatly ignored by USDA agencies and economists in general. To come up with feasible solutions for or effectively manage these issues, research needs to be conducted on the horse industry in general, not just individual high-end segments.

Importance of the Recreational Horse Segment

To breakdown the numbers of horses in the United States, the most recent AHC survey had a total population estimate of 9,222,847 horses, with 844,531 of those in racing, 2,718,954 in showing, 3,906,923 in recreation, and 1,752,439 horses used for other purposes. These data show that of the approximately 9.2 million horses in the United States, approximately 42% are recreational, compared to 29% in showing, 9% in racing, and 19% used for other purposes (American Horse Council Foundation, 2005a). This indicates that the largest percentage of horses are used for recreational and pleasure purposes as opposed to the serious competition industries or for business or breeding purposes. As clearly demonstrated in the above statistical information on the horse industry, the recreational segment of the horse industry is large both economically and in relation to the number of horses and people involved. Beattie et al. (2001) also reported that the pleasure horse industry in the state of Arizona was large and significant in relation to other industry segments.

The NAHMS Equine 2005 reports horses broken down into different (and more numerous) categories than the AHC survey. These reports only looked at operations with five or more equids (consistent with the definition of "farm" for the USDA Census). Small operations (5-9 head) had 36.1% of all equids, medium operations (10-19 head) had 34.2%, and large operations with 20 or more equids accounted for 29.7% of all equids in this survey. The percentage of equids on the property by primary purpose of the operation were as follows: Boarding/training, 5.9%; Breeding farm, 14.4%; Farm/ranch, 40.3%; Residence with equids for personal use 37.0%; and Other, including carriage services, guest ranches, and riding stables, 2.4%. Small operations were most likely to state their primary purpose was having equids for personal use, at 46.0%, while large operations said this was their primary function 10.4% of the time. Large operations were boarding/training facilities or breeding farms more often than medium and small operations. Broken down by primary use of equids, the percentages were: Pleasure, 45.7%; Lessons/school, 1.4%; Show/competition, 9.6%; Breeding, 15.9%; Racing, 1.4%; Farm/ranch work, 24.8%; and Other (such as horse trader, carriage or pony rides, etc.), 1.2%. The comparison of Equine 1998 to Equine 2005 indicated that more horses are on smaller operations than before; the percentage of large operations has decreased while medium and small operation numbers have increased, and 10% less horses were on the large operations in 2005 as compared to 1998.

Even with the recreational industry having a $32.0 billion annual impact on the U.S. economy, research including data on horses from this segment is rarely seen. Certainly, some recreational horses are at the lower end of the market place and would be included in economic analysis conducted in the North, Bailey and Ward (2005) study on horse slaughter. As documented in Maryland by a news source (Trejos, 2005), Arizona by economists (Beattie et al., 2001), and Florida and New Jersey by news sources (James, 2003; Lu, 2003), this industry is contributing to local economies on an enormous scale. Estimates from Maryland indicated that in Montgomery County, horses have a $196.2 million annual impact, surpassing all other agricultural goods (Trejos, 2005). Another news article indicates that no one seems to know the real impact of the equine industry on their area (James, 2003). This anecdotal evidence indicates that the recreational horse industry is important to agriculture at the local level at the very least; the AHC information indicates a significant economic impact at the national level from this low end market segment of the equine industry.

Summary

It becomes obvious in a study of the equine industry economics that there is an enormous lack of information across the board. There is no end of researchable questions – from the impacts of equine slaughter, to cost benefit analyses of NAIS and other registration systems, to modeling both the low- and high-end markets worldwide, to evaluation of breeding selection for various traits and the risk analysis of trade-offs (such as the trade-off for speed and agility as opposed to longevity and durability in thoroughbred racehorses), to the true short and long term economic impacts of disease outbreaks, such as the 2007 equine influenza outbreak in Australia. However, in my experience, what this industry needs the most is to compile an accurate data set on the industry targeting all horse owners nationwide and worldwide, to understand the decisions that horse owners make. Horses do not fit in with other traditional agricultural commodities – they are long-lived and are an emotional commodity to many owners as opposed to a business commodity. Personal utility and value models greatly differ, both from other agricultural goods and across the equine industry itself. Until we obtain better information on the industry, we will never be able to fully analyze the industry on a macro scale, nor correctly interpret signals we may see on a micro scale, such as at a horse auction.

Objectives

As health, welfare, and commerce concerns arise in the equine industry, one looks to economic research for answers or at least a framework with which to make decisions. However, this livestock sector is often overlooked in economic research, most notably the recreational segment of the industry. Information on this segment could be extremely useful in fully evaluating some of the hot topics in the equine political arena. To that end, this paper looks to address the question of what traits affect a horse's price in this low-end of the horse market, which could provide useful information to numerous individuals and businesses that own, buy, and sell horses in this segment. In addition, accurate information in this area could assist researchers in conducting economic impact studies of various political issues such as slaughter and animal identification.

The primary objective of this thesis is to discover and define traits and advertising modes that affect the value of horses throughout the lower end of the horse industry, primarily in the recreational segment; tied into this is the objective to analyze differences between the spring and the fall horse markets. Identifying an appropriate functional form to evaluate determinants of recreational horse prices is another goal of this research. Discussion and evaluation of issues with collinearity and heteroskedasticity will be conducted throughout results and suggestions on how to avoid issues with these data problems will be made. Topics and segments of the industry needing further research will also be suggested. The final objective, and perhaps the most important, is to provide a framework to quantify price determinants in the lower end of the equine industry including the recreational segment, a segment that contributes greatly to individual county's economies and billions in economic impacts to the United States annually. This study will hopefully demonstrate both the value of this research and the need for additional data sources in the equine market.

In the following pages, a literature review of the use and evolution of hedonic price models in horses will be considered. This literature provides the basis for the hedonic conceptual model described in the next section, followed by the information about the data and development of an empirical model. Next, the model results from three data sets using Ordinary Least Squares (OLS) and Generalized Method of Moments (GMM) regressions will be presented along with a discussion of the variables and their apparent relevance to the price of horses, and finally some concluding remarks will be stated along with suggestions for future research.

CHAPTER 2 - Literature Review

Hedonic pricing models have long been used to evaluate the characteristics that impact horse prices in other segments of the industry besides recreational horses, with the most common analysis being done on yearling racehorses at auction. Horse racing, the “Sport of Kings,” is the most intensely studied segment of the industry, with the most common analysis being hedonic pricing models of yearling racehorses at auction. Numerous articles including Robbins and Kennedy (2001) and Parsons and Smith (2008) cite Lawrence (1970) as being the first to use hedonic pricing models in the equine industry, including an article on thoroughbred pricing and a thesis on quarter horse prices. Since then, numerous similar articles have been published on the thoroughbred racing and breeding industries as well as segments of the quarter horse industry. Similar articles include Lansford et al. (1998) evaluating traits of racing quarter horse yearlings, Robbins and Kennedy (2001), who evaluated prices in a Canadian regional thoroughbred yearling market, and Parsons and Smith (2008) on traits affecting the price of thoroughbred yearlings in Great Britain. In addition, thoroughbred broodmares have come under scrutiny by Neibergs (2001), and by Stoeppel and Maynard (2006) who studied prices only of bred mares. The value of show horses at the American Quarter Horse Association World Show has been studied by Taylor et al. (2006). No research was found on the valuation of any other breeds (such as warmbloods or Arabians) or industry segments (such as recreation), and few nations appear to have conducted any market evaluations of their respective equine industry segments.

Buzby and Jessup (1994) examined factors influencing value at select thoroughbred yearling sales. They compared their results to three earlier articles evaluating thoroughbred yearling prices: Hastings (1987), Commer (1991) and Karungu et al. (1993). These three previous articles only looked at either yearling characteristics or macroeconomic variables; Buzby and Jessup simultaneously evaluated both sets of independent variables. They hypothesized correctly that both macroeconomic and yearling specific traits would be important, while the yearling specific traits would be more influential than macroeconomic variables. They ran three regressions with a semi-logarithmic model, one using each of the yearling specific characteristics and the macroeconomic factors and one using both, in addition to a linear model on the yearling specific characteristics alone to allow for direct comparison to Commer’s (1991) linear model. They used 1,027 yearlings from Keeneland select yearling sales sold between 1980 and 1990. Their models returned R-squared values of 0.05, 0.24, and 0.26, and F-tests indicated that the combined traits explained the price data better than either of the models with one set of variables alone. They found that the macroeconomic variables were nearly always all statistically significant and yearling specific traits such as sex, the racing success of the dam, and the stud fee of the sire were also statistically significant.

Chezum and Wimmer (1997) conducted a study on a phenomenon they call adverse selection in the thoroughbred yearling market, which is where buyers will avoid buying from sellers who both race and breed as opposed to those who just breed. They used a log-linear hedonic pricing model with the hammer price at the 1994 Keeneland September yearling sale as the dependent variable, and developed an independent variable for racing intensity for the seller that they used among other traits of the yearlings to evaluate whether that particular trait affected the sale price. They found that there was evidence for adverse selection in this market. Their models returned R-squared values of greater than 0.5 in all cases. Measures for racing success in the yearling's pedigree, sex, and age were included in the statistically significant variables.

Also in 1997, Neibergs and Thalheimer developed a recursive model to estimate supply and demand functions in the thoroughbred yearling market, and found that purses were the largest factor affecting the dependent variable price. The primary objective of their study was to develop a dynamic econometric model of the yearling market of the thoroughbred industry. They developed a supply model, based on lagged variables because of the breeding decisions being a year behind the birth of a foal, which is another year behind the sale of yearling thoroughbreds. They also developed a formula for foal to yearling transfer and finally an inverse demand function for yearlings. They used data on thoroughbred yearling sales in North America from 1960-1994. Supply and demand elasticities were developed. This was not a pricing model, but one looking at the thoroughbred industry as a whole over time and primarily macroeconomic factors. Based on their results, they suggest that purses can be a useful policy strategy within the thoroughbred industry to encourage growth. They additionally note that federal taxes have an economically significant impact on the thoroughbred industry.

Lansford et al. (1998) created a semi-log hedonic model evaluating traits of racing quarter horse yearlings using data from 5,295 sales for the time period of 1982-1992. Their hedonic model used genetic traits as well as macroeconomic variables to account for the valuation of quarter horse yearlings over this time period. They found and corrected for heteroskedasticity in their model and ended up with a model that had an adjusted R-squared value of 0.424. All but two of the variables included in their model were statistically significant, and supported the idea that older yearlings bring more money, as well as yearlings that are direct descendents of race winners or whose dam or sire have produced other winners. Sex was also statistically significant in their model supporting the idea that male yearlings were valued higher than females.

Neibergs authored a study in 2001 in *Agribusiness* entitled "A Hedonic Price Analysis of Thoroughbred Broodmare Characteristics." This was the first study of its kind, and the author justified it because broodmares are a primary capital input into the racehorse breeding industry. As they evaluated broodmares priced on their individual traits, they chose to use a semi-log hedonic pricing model similar to those used in previous literature evaluating horse and other livestock species. The objective of this research was to provide a framework for future research on broodmare values by finding traits that are important to broodmare prices with an end goal of helping breeders value their animals. The author felt that broodmares primarily experience a derived demand on their expected racehorse producing abilities, and therefore structured the model into four main categories; breeding, racing, genetic characteristics, and marketing factors that may impact demand directly. Niebergs came up with 19 total independent variables to evaluate across the four categories. The data were collected from the Keeneland November bloodstock sale in 1996, and all data points were collected on 1,602 of the mares sold. Due to the cross-sectional data, heteroskedasticity was tested for, found, and controlled. Only three of the variables and two of the days in the marketing binary variable for sale days were not statistically significant. The authors found that the broodmares breeding characteristics, such as whether the mare is currently barren or whether she has produced graded stakes winners, had the greatest marginal effects. The significant variables were as expected, but the authors emphasize clearly that this information was specific to the 1996 market and further research is necessary in evaluating long-run trends of price effects in the broodmare industry. They certainly achieved their objective of finding relevant variables to evaluate in future studies and communicating to the industry what traits to look at when valuing their breeding stock. Notably, there has only been one publication on broodmare prices since this initial article.

In the Fall 2001 *Journal of Agribusiness*, Vickner and Koch published "Hedonic Pricing, Information, and the Market for Thoroughbred Yearlings." They have two primary objectives; to extend the work of Chezum and Wimmer (1997) and test their theory of adverse selection, where sellers who both breed and race as opposed to only breeding thoroughbreds are discounted. They also provide information on what traits impact yearling prices, and provide marginal values for each trait, extending the contribution of Lansford et al. (1998). To complete these two objectives, they used a semi-log hedonic hammer price model which included many previously statistically significant variables and some never before tested. The model evaluating the marginal values of various traits appears to be a semi-log model though it is not explicitly stated in the text. They studied a 5% sample, or 212 yearlings, from the 1999 Keeneland September Yearling Sale, and gathered information on 20 independent traits, such as age, sex, and pedigree factors, on each. They justified their random sample with statistical tests showing it was similar to the population, but easier to work with. As this was strictly cross-sectional data, they did not include any macroeconomic factors. They failed to find adverse selection, contradicting what Chezum and Wimmer found in 1997, and the authors speculate that new variables or sampling method could have affected this outcome. After removing the two variables for adverse selection, they found eight of their independent variables and the intercept were statistically significant. They were able to attribute a marginal value in dollars to each result, which will allow sellers to better evaluate available data on yearlings. The authors contributed several new statistically significant variables in yearling thoroughbred pricing, and achieved both objectives of extending the literature. Though they found no adverse selection, they opened the door for further research on this issue.

Robbins and Kennedy (2001) evaluated prices in a Canadian regional thoroughbred yearling market using a semi-log hedonic pricing model. Their data consisted of sale price and traits collected on yearlings sold in British Columbia from 1985 to 1997. This market is a lower-end market than the thoroughbred sales in Keeneland, and they tested to see if similar traits affected the price of thoroughbred yearlings there as elsewhere. In addition, they wanted to extend the research on the effect of dam's traits on the sale price of their offspring. They also note that thoroughbred yearlings appear to be consumer goods as has been previously thought, though they provide only indirect evidence to support this claim. They reached conclusions that supported previous work, that traits such as sire breeding fee capture the effect of quality of sire on the yearling's price. They also extended previous literature and showed that a dam's primary contribution to the yearling's price comes from having other offspring already running well.

More recently, in 2006, a study was published in the *Journal of Agricultural and Resource Economics* on the show horse industry by Taylor et al. entitled “Show Quality Quarter Horse Auctions: Price Determinants and Buy-Back Practices.” This study was the first in the show horse industry using data from the American Quarter Horse Association (AQHA) World Show form 1995-2002, though as with the other articles it notes hedonic price modeling has long been used in livestock pricing and the thoroughbred yearling industry. The two major objectives of this study were to determine what traits affect the price of Quarter Horse show horses and to determine if horses that do not sell systematically differ from those horses that sell. As with much of the prior literature on livestock price modeling, the authors chose a semi-log hedonic pricing model to evaluate the data. Fifty-five independent variables were evaluated, which could be grouped into the following categories: Genetic and Physical Traits, Individual Performance, Performance of Offspring, Quality of Pedigree, Sale Order, and Year. Initially, the authors estimated a Heckman model, the second step of which suggested that no-sale horses are no different than those sold. Second, the authors used a simple ordinary least squares regression for the hedonic pricing model, using only data from the horses that sold. The observations totaled 3,090 horses from the sale years 1995-2002, and were adjusted for inflation. The authors concluded, based on both the Heckman model results and their hedonic model, there were no systematic differences between the no-sale horses and those that sold. In addition, they found that knowing individual characteristics of the horses certainly impacted price, especially horses with distinguished show records, those that are eligible for certain special awards, have a strong pedigree, and have potential future breeding value. The authors also note that horses are considered a luxury good and therefore the annual economic conditions impacted the prices of sale horses as well. The article certainly met its two objectives of establishing what characteristics hold value in show horses and whether or not the no-sale horses differed from horses sold. They suggest future research into why the buy-back method is used at all, and additional variables about the horse sellers themselves which may impact price.

Stoeppel and Maynard (2006) studied prices of bred thoroughbred broodmares, extending the work of Neibergs (2001). A hedonic pricing model of the semi-log functional form was used. They utilized cross-sectional data from the 2005 Keeneland broodmare sale to evaluate what characteristics of the broodmares had the greatest effect on their value, finding that age, the broodmare sire’s stud fee and the stud fee of the sire of the foal she is carrying all had highly significant impacts on price. In addition, they found that they had similar results to Neibergs' study, and moved forward into forecasting out-of-sample. They also obtained a comparatively high adjusted R-squared value for cross sectional data of 0.83. They note that the accuracy was superior to naïve expectations, but that this model was not enough alone to predict the sale price of a thoroughbred broodmare.

Finally, Parsons and Smith (2008) conducted the most recent study on horse prices, evaluating traits affecting the price of thoroughbred yearlings in 2004 in Great Britain. They primarily wanted to evaluate new policies implemented affecting the British thoroughbred industry. They used 2,211 observations collected from 2004 yearling auctions in Great Britain. A log-linear functional form was used for hedonic pricing models which returned R-squared values of 0.62, 0.62, 0.63 and 0.63 across four models. They found results consistent with the North American studies for traits affecting thoroughbred yearling prices, and went on to extend the model to evaluate some of the British racing incentive programs and how effective their policies are in increasing yearling price. Due to a lack of evidence for strong statistical significance in the regression, they conclude that the incentive programs are ineffective.

Mean prices of the horses being evaluated in the literature mentioned above are shown in Table 5 to provide evidence that only the high end of the equine market has been previously studied. The average price for Taylor et al. (2006) was converted from the natural log form; for Parsons and Smith (2008) the average price had to be converted from 2004 British Pounds into 2004 U.S. dollars. The prices are given in nominal prices, not adjusted for inflation, but even so

Table 5: Average Prices in Previous Literature

Study	Year	Horse Type	Average Price
Chezum & Wimmer	1997	Thoroughbred Yearlings	$38,741.00
Lansford et al.	1998	Race-Bred Yearling Quarter Horses	$7,111.47
Neibergs	2001	Thoroughbred Broodmares	$71,271.00
Vickner & Koch	2001	Thoroughbred Yearlings	$77,140.00
Robbins & Kennedy	2001	Thoroughbred Yearlings: Summer	$11,213.00
Robbins & Kennedy	2001	Thoroughbred Yearlings: Fall	$2,967.00
Taylor et al.	2006	Show Quality Quarter Horses	$6,063.00
Stoeppel & Maynard	2006	Thoroughbred Broodmares in Foal	$169,735.23
Parsons & Smith	2008	Thoroughbred Yearlings	$69,750.84

are considerably greater than the sample average prices of this study, mentioned below, or the average sale prices of horses obtained using the USDA Census data which returned average sale prices ranging from $2,724 to $3,146. This further indicates the lack of research on the low end of the equine industry, the segment which includes the majority of horses in the industry.

No research was found on the valuation of any other breeds (such as Paints or Arabians) or industry segments (such as recreation), which is therefore the aim of this thesis. Only two breeds, Thoroughbreds and Quarter Horses, appear to have been evaluated in economic literature; numerous other breed organizations may be interested in what traits affect the value of their animals. In addition, only Great Britain, the United States, Australia, and Canadian markets have evaluated any horse market data in my research. Other countries that are vital to the international equine community could conduct research about their respective equine markets. The journal articles in the past have focused on horse sales at what is considered the "high end" of the industry. The low end of the horse industry has been greatly ignored in economic research, and this segment of the horse industry encompasses the greatest financial segment and the greatest number of people and horses (American Horse Council Foundation, 2005). As shown in Table 5, the average values of horses studied in several of the hedonic pricing model studies previously mentioned are clearly higher than the value of horses nationwide as indicated by the U.S. Census. The data sets used for the hedonic pricing models in this study have average prices for the horses that are much lower than those of previous studies, and therefore are focusing on the lower-valued segment of the industry.

In studying economic literature on the equine industry, it is clear that this is an internationally important, often over-looked industry that contributes several billion dollars to the worldwide economy annually. Gaps exist in equine economic literature from the top end horses in racing and show arenas to the lower end including recreational and slaughter horses. Some high-end horse markets, like the ones evaluated above, keep relatively good data on prices and market trends, but no research or data collection has been done on a regular basis on the low-end, recreational horses in the industry; those that have a $32.0 billion dollar annual impact to the United States economy and keep some local communities alive (American Horse Council, 2005; Lu, 2003; James, 2003). Many current major legislative issues, such as the National Animal Identification Program, horse slaughter, and animal welfare, could be greatly clarified through increased market evaluation of the horse industry. In addition, the timing for this research to begin is excellent due to the likely increase in horse sales online over the past several years, considering that internet access has spread out into the rural areas more so than in the past.

CHAPTER 3 - Data Source

The data for this analysis is cross-sectional across horses sold on the website www.dreamhorse.com in the spring and fall of 2008. It is important to note that in the horse industry, especially among the lower end where recreational riders are, much of the value of a horse is decided on factors that are hard or impossible to quantify – the look of the eye, apparent intelligence, riding level and discipline, ground manners, eye appeal of the horse (color, conformation, markings, etc.), and buyer's history with certain types of horses (e.g., they like grey horses more because their first horse was grey). These variables are difficult to quantify, if not impossible, and can only be partially captured even by adding every available variable to the equation; however, by adding too many variables multicollinearity problems can arise. Therefore, it cannot be expected that every aspect of a horse's price will be captured with the currently available data used in this thesis. This study is also different because all previous hedonic pricing analyses on horses have been within one breed, where in this model, breed is allowed to vary. Due to human error and a lack of control over the amount and accuracy of the data presented in the ads, the fit of the pricing model may be lower than seen in previous studies. It would be desirable to include more factors in the model, but a lack of quantifiable factors prevents this from being possible.

Justification of Website Choice

The website selected as a data source for this study is currently one of the largest horse sale websites online. As with other marketplaces, the internet opened new opportunities for the horse market to expand into, and that option has been primarily utilized by the lower end, or recreational, horse market. The author has personally used this and other websites to purchase and sell horses in the past, and has always had the greatest response and found the best options for horses to purchase through dreamhorse.com as compared to other websites. At two points in time, a search through Google was conducted on the phrase "horses for sale" to find and compare other horse sale websites to dreamhorse.com. Listed horse advertisement numbers are examined from the Prairie USA region, as defined in the drop down selection list on dreamhorse.com as the states of Kansas, Missouri, Nebraska, North Dakota, Oklahoma, South Dakota, and Texas, and each of these individual states. The total number of horses listed on each site was also included if available. As a comparison for the horse numbers listed on these websites, estimates from USDA suggest that around 500,000 horses have been sold annually based on Census information and NASS Equine (1999).

On June 17, 2008, a Google search for "horses for sale" was conducted and the top 10 websites resulting from this search, in order, were: 1) equinenow.com, 2) equinehits.com, 3) horsefinders.com, 4) horsetopia.com, 5) dreamhorse.com, 6) equine.com, 7) horseclicks.com, 8) horseville.com, 9) agdirect.com, and 10) 2buyhorses.com. A quick evaluation showed that horsefinders.com, horsetopia.com, dreamhorse.com, equine.com, horseville.com, and 2buyhorses.com would not allow for a search of all horses, either by limiting the searches to a maximum number of returns or not allowing the search to occur at all. However, by conducting searches by gender on dreamhorse.com, we were able to come up with a total number of advertised horses at that time of 55,675 head. In comparison, equinehits.com had a greater total of 61,218 head advertised in the U.S. and Canada, while equinenow.com returned 30,317 total ads in North America, horseclicks.com had 9,483 total ads, and agdirect.com 13,989 total ads. This indicates dreamhorse.com is one of the most frequently used horse advertising websites.

The same Google search was conducted on October 15, 2008, returning the top 10 horse sale websites as the following, in order: 1) horsetopia.com, 2) equinenow.com, 3) horsefinders.com, 4) equine.com, 5) equinehits.com, 6) horseville.com, 7) horseclicks.com, 8) horsesforsale.org, 9) myhorseforsale.com, and 10) dreamhorse.com. The total number of ads, if it was able to be determined and the ads in the Prairie USA region and per state are included in Table 6. These numbers and the techniques required to obtain each number are discussed in the following paragraphs.

Table 6: Internet Horse Advertising October 2008

Website	Total	KS	MO	NE	ND	OK	SD	TX	Prairie Region
Horsetopia	N/A	230	500	105	71	311	135	1,004	2,356
Equinenow	25,066	328	869	250	121	901	264	1,794	4,527
Horsefinders	4,950	60	150	49	17	72	14	286	648
Equine	N/A	345	756	394	223	538	264	1,882	4,402
Equinehits	69,840	797	2,007	621	432	1,406	477	4,971	10,711
Horseville	N/A	200	596	138	90	492	135	1,170	2,821
Horseclicks	15,460	179	481	147	63	392	94	1,014	2,370
Myhorseforsale	1,472	38	27	16	6	257	10	242	596
Dreamhorse	58,166	499	1,290	438	256	1,262	241	3,322	7,308

Horsetopia.com limits searches to a maximum of 500 results, but conducting a search for horses in Kansas returned 230 ads, Missouri returned 500 ads, Nebraska returned 105 ads, North Dakota returned 71 total ads, 311 ads from Oklahoma, South Dakota had 135 ads, and Texas had a total of 1,004 ads (searched by selecting mares (494), geldings (335), and stallions (175) separately), for a total of 2,356 ads in the prairie region as defined by dreamhorse.com. These numbers include horses for sale and for lease and include “sold ads” of horses which have already been sold for which the ads have not been removed.

Equinenow.com showed 25,066 ads on October 15, 2008. They allowed a search to display all ads. Numbers in the states in the prairie region are: Kansas, 328; Missouri, 869; Nebraska, 250; North Dakota, 121; Oklahoma, 901; South Dakota, 264; Texas, 1,794. This makes a total of 4,527 horse ads in the prairie region. These ads include horses for sale and standing at stud.

Horsefinders.com showed only 1,000 ads, but it appears to be limited to that number per search. The numbers of horses are listed by state, and the total number of ads if you total the state numbers and numbers from Canadian provinces is 4,950, with 4,801 from the U.S. states and 149 from Canada. Kansas has 60 ads, Missouri has 150 ads, Nebraska has 49 ads, North Dakota has 17 ads, Oklahoma has 72 ads, South Dakota has 14 ads, and Texas has 286 ads, for a total of 648 ads in the prairie region. These ads appear to include only unsold horses except for a few sold ads that are left up by the seller.

Equine.com limited search results to 1,000 as well. Searching by state for the prairie region, Kansas had 345 results, including sold ads, and horses for sale and for lease. Further, Missouri had 756 results; Nebraska had 394 results; North Dakota had 223 results; Oklahoma had 538 ads; South Dakota had 264 ads; and Texas, not surprisingly, had more than 1,000 ads. Therefore, I searched ads from $0 to $2,000 and from $2,001 to $5,000 and $5,001 and up. 762 horses were advertised in Texas for under $2,000. 588 results were found searching from $2,001 to $5,000, and 532 ads were found from $5,001 up, for a total of 1,882 ads from Texas. This comes up with a total of 4,402 horse ads in the prairie region on equine.com.

Equinehits.com had 65,683 ads for horses for sale, 2,298 ads for horses for stud, 1,402 ads for horses for lease, and 457 horses for trade. This gives a total of 69,840 horse ads posted on this website. By state, Kansas had 797 horses posted, Missouri had 2,007 ads, Nebraska had 621 ads, North Dakota had 432 ads, Oklahoma had 1,406 ads, South Dakota had 477 ads, and Texas had 4,971 ads. This comes to a total of 10,711 horse ads on this website in the prairie region, for horses for sale, lease, trade, and stud. Equinehits.com also states on one of its pages, “Thank you for using EquineHits, the world’s largest free photo classifieds.”

Horseville.com does not allow more than 3,000 ads to be shown. However, searching by state, we can find that Kansas has 200 ads, Missouri has 596 ads, Nebraska has 138 ads, North Dakota had 90 ads, Oklahoma had 492 ads, South Dakota had 135 ads, and Texas had 1,170 ads posted on this site. This included ads for sale, lease, trade, stud, and sold ads. This gives a total of 2,821 horse ads in the prairie region on this site. They state that their website has over 17,000 ads.

Horseclicks.com sells horses, properties, trailers, saddles, and more. If you go to the horses section, it states at the top of the page that it is the “World’s Largest 100% Free Equine Photo Classifieds. No charges, no fees, no commissions – it’s all free.” The horses are listed 10 per page, and there are a total of 1,546 pages, for a total of 15,460 ads on October 15th, 2008. Kansas had 179 ads, Missouri had 481, Nebraska had 147, North Dakota had 63, Oklahoma had 392, South Dakota had 94, and Texas had 1,014 total ads, for a total number of ads in the Prairie region of 2,370. These ads all appear to be for horses for sale only, and a few noted that the horses were sold. They allow horses for sale, stud, lease, or trade to be advertised.

Horsesforsale.org is very different than the other horse websites. They are listed by breed with advertiZers, it appears to all be linked through GlobalAdvertiZing.com. There are horses listed by breed, including 21 different breeds, some with multiple websites listed. They are also listed by discipline, with 20 different disciplines. Due to the way this website is arranged, it was impossible to get a count of total horses or horses per state in the prairie region.

Myhorseforsale.com had 1,472 total ads listed on October 15, 2008. Kansas had 38 ads posted, Missouri had 27 ads, Nebraska had 16 ads, North Dakota had 6 ads, Oklahoma had 257 ads, South Dakota had 10 ads, and Texas had 242 ads. This makes for a total of 596 horse ads in the Prairie region on this website.

Dreamhorse.com, the data source for this thesis, does not allow for a search of the total number of horse ads posted on the website. There were 29,643 ads for Colts, Geldings, and Stallions, and 28,523 ads posted for Mares, Mares in Foal, Fillies, and Unborn Foals, for an estimated total number of ads of 58,166. Kansas had 499 ads, Missouri had 1,290 ads, Nebraska had 438 ads, North Dakota had 256 ads, Oklahoma had 1,262 ads, South Dakota had 241 ads, and Texas had 3,322 ads. This made a total of 7,308 ads in the Prairie USA region, including horses for stud, sale, lease, or already sold ads. These totals make it the second largest advertiser of horses in the prairie region and of horses online out of the top 10 results through a Google search. As it is also the only website in the top 10 which allows for a search of only horses which have been sold, it is a justifiable choice as a data source for this thesis.

Information on Posting Ads on Dreamhorse.com

When a person posts an ad, they have numerous fields of information to fill in and numerous options for their type of ad. When a person is looking to buy a horse on dreamhorse.com, all of the traits entered by the seller can be searched for by the buyer, so a buyer can search for specific desirable traits in the horses and obtain a personalized set of search results. Therefore, it is important to understand the traits that a seller can post regarding their horse and what parameters the seller may place on their search to optimize the number of views your horse may receive from prospective buyers. To sell horses on dreamhorse.com, you first go to the website, www.dreamhorse.com. Then, a user account must be created. Once this is completed, a seller can post an advertisement on their horse for sale by clicking on “Create New Ad.” When this link is clicked, you are taken to a page with specific directions and a link to the dreamhorse.com Terms of Use. Following this information, there are numerous fields to choose from or fill with information on your horse.

The first field is the horse name, which can be any name the owner designates. Some individuals use this line for registered names, some for barn names, and some for descriptions of the horse. The next fields are the city/town, state, and zip code where the horse is located. This information can be used for sellers searching for horses in their area, as transportation is costly and horses at the lower end of the market are not often purchased from great distances away. Most people prefer to see the horse in person, try it out, and not have to spend a great deal in travel costs looking at horses they may not purchase or getting the horse home. Following the location fields, there is room for an Ad Title/Headline, which is a short blurb of information on your horse. Then there is a space for the seller to write whatever they wish about their horse. Free text ads are limited to 500 characters in this space, but if you purchase any level of other advertisement, this is increased to 1,500 characters. URLs for other websites are not allowed in this field. Following this, there are drop down lists to select the breed of your horse and an optional secondary breed. There are 195 possible selections for primary breed of the horse – but these include the options “Other,” “Goat,” “Zebra,” and a few other species and options that are not necessarily breeds of horses, but that people are allowed to post ads on. Then, there is a drop down list of 200 possible “secondary breeds,” which include different paint colors and the original breeds in the primary breeds drop down list.

After these lists, there are places to select if the horse is or is not gaited (has a set of gaits specific to its breed), a warmblood (a breed group of horses), or registered (with any association that records breeding or performance registries). Then, there are fields where you can fill in the registration organization and registration number of the horse for sale. After that, there is a section for foaling date, where you can select from a drop down list the month foaled if you choose and the year foaled, which is required. Next, the seller can select the horses’ gender from Colt, Filly, Gelding, Mare, Mare in Foal, Stallion, or Unborn Foal. Following this, there is a drop-down list to select the base color of the horse from 30 possible color options, followed by a field where you can fill in other colors or markings on the horse. Next, you can select the hands high the horse is, where a hand is the standard measurement for the height of a horse equating 4 inches. Following the number of hands is often a .1, .2, or .3, which means the horse is that number of inches plus one, two, or three more inches, before moving up to the next “hand” high. Then the seller can select the approximate weight of the animal, and a temperament score from a scale of 1-10 where horses ranked a 1 are extra calm and a 10 are “high-spirited.” As horses can be listed on this site for trade, stud, or other purposes, the next option is to click if the horse is or is not for sale, and is or is not a private treaty sale, and the asking price for the animal. Selecting private treaty means the horse’s price will be used for search purposes but will not be displayed in the ad. You can then select if the ad is for a “Horse Wanted,” “For Lease,” “May Trade,” or “Missing Horse.”

Following these selections, the rest of the information is optional. Next, there is a section where the person posting the ad can select if their horse is at stud, and fill in the stud fee, booking fee, if shipped semen is available or not, and the shipping fee. After that, there is an optional list of skills or potential that the horse you are selling may have. The seller may select up to five of these characteristics from a list of 141 possible options. Following this option list, the seller can fill in their horse’s pedigree with up to four generations of information, and in the first two generations (sire and dam and grandsires and grand-dams) there is also a field for a small amount of information such as show points, color, or foaling year. This is the last option on this page.

After filling in all of this information, especially the required fields marked with a red asterisk, you click submit, which moves you to the next page. On the next page, you are given your ad number and the options to add another horse, order photos, view your cart, or go back to your dreamhorse account homepage. If you select order photos, you are taken to a page describing the options for photo ads on dreamhorse.com. A 90-day photo ad is $20 per photo, a six-month photo ad is $25 per photo, a six-month video ad with two photos and a video is $50, or a six-month spotlight ad is $100. There are also renewal options for photo ads. You can renew one photo for 30 days for $5, renew or upgrade to a video ad for $45, renew or upgrade to a spotlight ad for $90, renew your photo ad for 90 days for $15, renew and replace your photo for $20 for 90 days, simply replace existing photos or videos for $10 per photo or video, order an additional photo for 90 days for $20 per photo, or order an extra six-month photo for $25. Free text ads are posted for 90 days, with the option for unlimited renewal every 90 days.

Data Collection Process

Data were collected from March 13 to March 28, 2008 for the spring data set and from October 8 to 11 for the fall data sets by running the same search every day with oldest ads first, which captured horses sold between these two dates with the aim to not miss many horses that had been sold within the 90 days prior to the data collection. In the spring data set, the initial data collection included horses that sold for greater than $5,000 and were being advertised in the ad as high-end show horses, which gave a total of 545 observations. However, the extremely high priced horses skewed the data drastically so they were eliminated from the data set to look more explicitly at recreational horses; this brought the total down to 470 observations. When initially evaluating the data, the observations ranged from $1 to $50,000 with an average of $3,181.28, which was notably greater than the median and mode of $1,500. Furthermore, the standard deviation of $4,704.85 also seemed high for the recreational industry. A distribution of the original prices is shown in Figure 1 on the following page. The value of $5,000, which is pointed out in Figure 1, was chosen due to a high number of horses being priced at that value and values greater than that being less common. It also is a reasonable range of prices for the recreational horse industry, based on the author's personal experience. In the fall, data were only collected on horses sold for $5,000 or less to save time in the data collection process. The following paragraphs describe the search criteria selected for this data set.

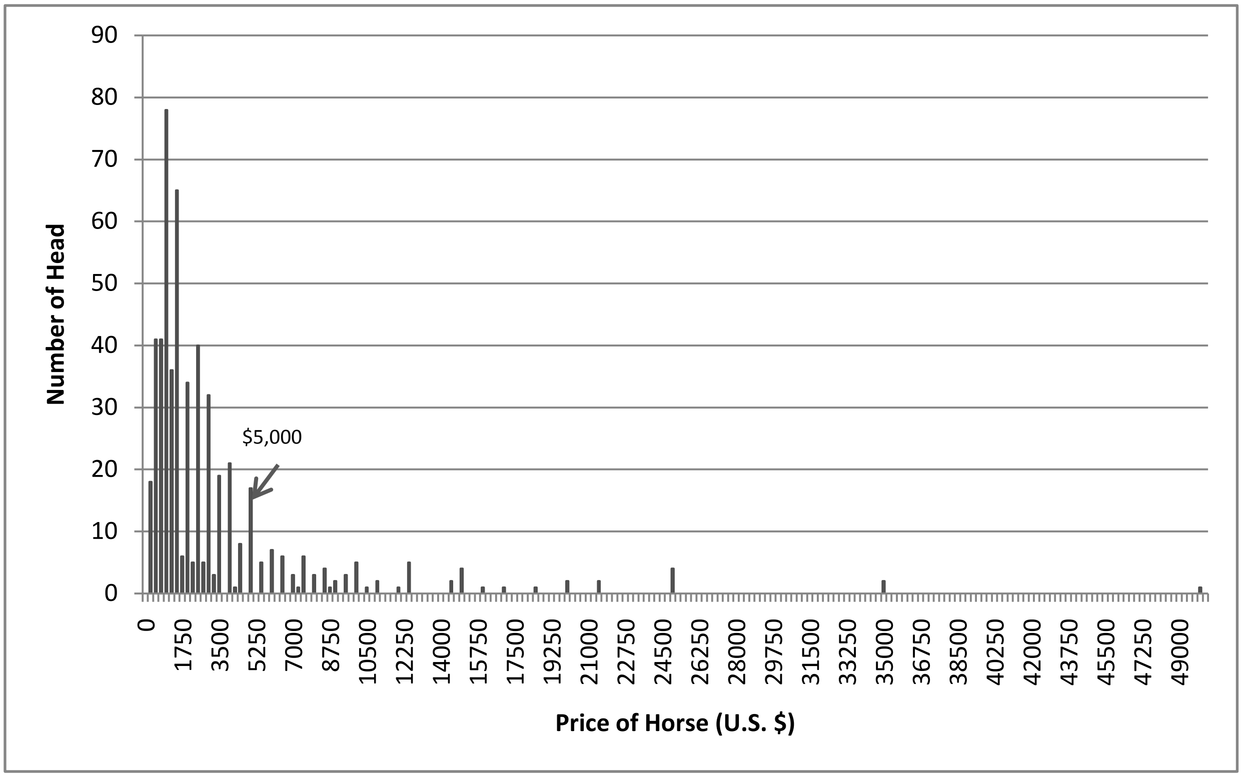

Figure 1: Spring Horse Price Distribution including Horses over $5,000

The author first went to the website, www.dreamhorse.com, and clicked on "Advanced Search." Next, the option to include all breeds with no exclusions was selected, and the fields where you can search by zip code or state were left blank. Alternatively, the region selection was made for the "Prairie USA" area, which includes the states of Kansas, Missouri, Nebraska, North Dakota, Oklahoma, South Dakota, and Texas. The fields of horse name, ad title, bloodline name/word, and generation were left blank in the search. "Any Skill" was selected from the skills drop down list, and the selections for "Warmbloods Only" or "Gaited Only" were not marked. "Last 90 Days" was selected from the drop down list titled "Newly added within," to only look at ads posted and horses sold within the last 90 days. "Any Gender" was also selected. Minimum and maximum age fields and year foaled fields were left with no minimum, maximum, or any year. "Any Color" was selected from the color list and the other color box was left blank. Minimum and maximum heights and weights were set to no minimums or maximums, respectively. For the initial spring data collection, price was also set to no minimum or maximum, but for the fall data set the maximum price was set to $5,000. "Exclude Private Treaty Ads" was selected as these ads do not list price information and therefore would not be useful to this study. Minimum and maximum temperament scores were set to no minimum or maximum. "Must be for Sale" and "SOLD Only" were both selected, to limit the data set to sold horses that were listed as for sale. "Must Be At Stud" was set to "Show All" to include any horses whether or not they were standing at stud at the time of their sale. None of the lowest nine check boxes were marked (Must Ship Semen, Must Be For Lease, May Trade, Horse Wanted, Only Show Spotlight Ads, Only Show Video Ads, Only Show Missing Horses, Only Show Photo Ads, Must Be Registered). "Sort by Oldest Ad" was selected in addition to "List 200 Per Page," and then the search was submitted.

The dates the spring search was conducted were March 13, 14, 15, 24, 25, and 26. Searches were conducted at multiple times each day. The total number of ads returned for each of these days was 482, 479, 485, 501, 515, and 507 respectively. These numbers include mostly ads that were present on previous days as well as some new ads, not all new ads every day. After eliminating the horses sold for over $5,000, 470 total data points were collected in the spring data set. The fall search was conducted in the same way, except a survey was sent through the website to each of the owners asking for the actual transaction price of the horse, and it was limited to horses with prices under $5,000 from the beginning to eliminate some of the labor. The search dates for the fall were October 8, 9, 10, and 11. On October 8, 396 ads were found with the given parameters in the early afternoon, and later in the evening two searches conducted returned 392 horses and 389 head. On October 9, multiple searches at different times returned 387, 388, 389, and 387 head. Searches conducted on October 10 returned 387 ads at one time and 376 ads each in two additional searches. Finally, on October 11, searches returned 373, 374, and 379 head at three different times. After totaling the observations recorded each day, a total of 407 useful data points were found for the fall data set. The survey responses and results are discussed in Chapter 7.

Issues with Data Source and Collection Process

There are admittedly some problems with the data source and data collection methods. Four primary issues are: 1) reliance on the accuracy of data provided by the seller; 2) use of the asking price rather than actual transaction price as the dependent variable; 3) the inability to know if everyone who sells a horse on dreamhorse.com posts the horse's ad as sold or simply deletes the ad, and whether or not there is a systematic component to the seller's choice; and 4) the inability to collect all of the data in an efficient fashion, as the collection took place over a number of days rather than in one day, at one time. These four issues are discussed below.

In relation to reliance on seller's data accuracy, in some fields of the horse's advertisement, the seller is allowed to choose or type in information that may or may not be accurate in regards to the horse. For example, if the horse is unregistered or has no verification of age, the year the horse was born, and therefore its age, could be approximations based on an examination of the teeth by the seller or a veterinarian or on a previous owner's hearsay. Breed could also be inaccurate in an unregistered horse if the seller was incorrectly told the breeding of the grade horse or simply guessed at the breed. Registration itself depends on the seller's data accuracy unless the prospective buyer is able to access the registry's database and check that the horse is actually registered. Color is a factor which is also often mistaken, due to differing definitions of colors in various publications. For example, sorrel is sometimes considered a form of chestnut. The color buckskin is often incorrectly called dun whether or not the horse carries the dun factor gene. In addition, at times a horse can express a color it genetically should not be, as color genetics in horses are as of yet an imperfect science. Temperament score is also a scale not based on any scientific definition, but a number chosen by the seller to "best describe" their horse's personality. Temperament is a big concern, as generally calmer horses are preferred in the recreational horse market for children and newer horse owners, and therefore would likely tend to be biased towards the low end of the 1-10 range, as is evidenced by the data. Any of these fields could be slightly erroneous if horse sellers systematically believe one or more breed, age, color, or temperament scores are better than others, and therefore inaccurately list their horse as having that characteristic rather than the factual one. The only real way to get accurate information for breed and color in most cases would be a genetic test, most veterinarians consider looking at a horse's teeth the best way to judge their age, and temperament scores in horses are not scientifically measurable without an expert spending some time with the animal or putting each animal through the exact same process (similar to scoring cattle out of a chute based on whether they walk, trot, or bolt out of the squeeze chute after procedures are performed). Therefore, we use these data and assume they are the most accurate we can obtain or at least assume they are the information given to the buyer of the horse, and that the buyer is purchasing the horse based partially or fully on the information given in the advertisement.

The price data for the Fall and Spring 2008 models are the asking, not the actual transaction prices of the horses. They are the price the seller had listed on dreamhorse.com at the time the horse was marked sold. Whether this differs from the actual transaction price is questionable. To deal with this issue, a survey was conducted of all sellers in the Fall 2008 data. These survey data and the results are discussed in another section; as will be demonstrated in that section, a high correlation was found between the asking price and actual price, and the model run on survey data shared several similarities in its results with the data obtained online for the fall model.

If and when the horse is sold, or even if it is not sold, the seller has a choice of whether to mark the ad as sold, delete the ad, or leave the ad unchanged on dreamhorse.com. If there is a systematic component to the horses or sellers that leave their ads unchanged or delete their ads versus marking them as unsold, or if numerous horses are marked as sold when they are actually kept by the owner, it is not captured in these models. In the author's experience, it is unlikely that a seller would leave up a sold horse's ad and not mark it sold due to the desire to not get additional e-mails or phone calls from people interested in purchasing the horse.

It is uncertain how many people remove their horses as opposed to marking their ad as sold; however, a check was conducted on a sample of 800 ads from the Prairie region of horses posted in the last 90 days for sale or sold for under $5,000 over a 24-hour period from October 8 to 9 out of a total number of 3,942 ads on the 8 and 3,940 ads on the 9. The sample consisted of web results pages 1, 2, 9, and 10, with 200 horses per page, to capture both photo ads (listed first) and text ads (listed after photo ads). It was found that of these 800 ads only 2 were removed from the website during this time for no apparent reason, both were from Texas and both were Palomino, coincidently apparently, one born in 2004 and the other in 2007. This same type of sample was conducted on November 13, 2008 and a sample of 800 ads was taken from a total of 3,445 ads on the 13 and 3,472 ads on the 14. Pages 4, 5, 14, and 15 were selected for the sample of 800 ads this time. This time, only one was missing 24 hours later, and it was a sorrel paint gelding from Texas. Even though all three horses whose ads were removed for no apparent reason were from Texas, this is not surprising as most of the horses in the data sets are from the state of Texas. Based on these results, the author feels the phenomenon of removing horses without marking them sold or for no explanations is rare. The results of the survey, discussed below, returned one horse marked as sold which was kept by the owner, and several horses that were traded.

Finally, the last data issue is the fact that even though the search limited ads to the past 90 days, the search conducted in the spring occurred over a 13 day period and in the fall over a 4 day period. Therefore, some data would be lost by being on the borderline of 90 days in the beginning of the search and going past 90 days by the last day of the search. This is an issue which cannot be fixed and it simply must be realized that the collection time frame is slightly longer in the spring than the fall and is also slightly longer than the actual 90 days. The spring data set would actually reflect a 103-day ad posting period, and the fall data set a 94-day posting period. Inefficient data collection methods being available are to fault for this issue.

CHAPTER 4 - Model Development

Conceptual Model

In 1966, Lancaster published a journal article entitled, "A New Approach to Consumer Theory," which provided the basis for hedonic pricing models that have been used ever since on various consumer and capital goods. Lancaster's work provided a new way of thinking about consumer goods. Prior to this article being published, consumer theory suggested that, "…goods are what are thought of as goods" (Lancaster, 1966). Consumers simply wanted more goods, and the value was in the good and the number of goods a consumer had, not the characteristics of those goods. Rather than the good being the thing that provides utility to the consumer, the good actually possesses characteristics, and these characteristics are what provide utility, and hence value, to the consumer. He also added that many goods could share a certain characteristic and that a single good could have many characteristics which affected its value in the eyes of the consumer. Assuming a linear relationship in the characteristics and the good, we therefore can find that the value of a good is simply the sum of the value of its characteristics, and therefore a vector of characteristics can be used to discover the value of a good. Shortly after Lancaster's work was published was when the first hedonic pricing models were used in horses according to several of the papers in the literature review citing Lawrence (1970).

Besides the horse market, since the publication of the Lancaster article numerous markets have used hedonic pricing models to evaluate the value of their respective goods. Faminow and Gum (1986) used a pricing model based on the characteristics of feeder cattle to study auction price differentials. Chvosta, Rucker, and Watts (2001) estimated prices and considered transaction cost using an hedonic modeling approach when looking at bull auction sales. Dhuyvetter et al. (1996) used an hedonic approach to study beef bull price differentials. Mintert et al. (1990) also looked at characteristics of cows when studying cow auction price differentials. Outside of the livestock market, Kolodinsky, DeSisto, and Wang used hedonic modeling to determine the price premium consumers are paying for rBST-free milk. Espinosa and Goodwin (1991) apply this approach to wheat characteristics in panel data from the state of Kansas to study what characteristics, including conventional characteristics as well as milling and dough factors, influence wheat prices in the state. Finally, this approach was used outside of agricultural markets all together in the housing market, based on characteristics such as dwelling size, quality, and lot size, by Witte, Sumka, and Erekson in 1979. These articles all show that the hedonic pricing model is one which applies to many areas of consumerism, and has been used often since Lancaster's 1966 work was published.

Using the conceptual framework developed by previous authors, three different functional forms of hedonic pricing models will be used here to attempt to identify factors that affect the prices of horses at the lower, recreational end of the equine industry, one linear, one semi-log, and one log-linear. In an hedonic pricing model, the utility a consumer gets from the purchase, and therefore the price of the commodity, is based not on the quantity they purchase but on the characteristics of the item, or in this case, the horse. At the higher end of the market, in a much more controlled environment, the horse characteristics one can evaluate have more to do with bloodlines, money earners in shows and races, and proven value than a horse used for recreational purposes. However, the conceptual model is still the same, as shown below in equation (1). The vector of characteristics includes aspects such as the state the horse was sold from, breed, color, temperament, and registration. Therefore, the linear conceptual model appears as follows:

$$P_i = \beta * \mathbf{x}_i + e_i, \qquad (1)$$

where i stands for each individual horse, P stands for the sale price, β is the coefficient vector on characteristics, **x** is a vector of characteristics such as those listed above, and e is the error term. In words, the asking price or actual price of a given horse will be regressed against quantifiable factors about that horse known to the buyer at the time of the sale. As documented above, this approach has been used numerous times in the horse industry in the show, breeding, and racing arenas, as well as numerous other sectors of the livestock, agriculture, and other economic markets, but never attempted in the lower end market with recreational or pleasure type horses.

Variable Descriptions

Dependent Variables: Price

The models for Spring and Fall, 2008, dreamhorse.com data both use the dependent variable *PHorse*, which is the Asking Price of the horse as listed on the dreamhorse.com ad labeled as "Sold." The Fall Survey model uses the Actual Sale Price of the horse as reported by the seller through an e-mail survey done through the dreamhorse.com website, which is the variable called *APHorse*. The maximum value of this variable is set at $5,000 to limit the data set to prices with a fairly normal distribution and to focus more on the lower-end, small show, breeding, or trail horse market as opposed to major show circuit horses or racing horses which can be valued in the many thousands of dollars. This limit was set to avoid creating bias in the model based on a few high valued horses sold on dreamhorse.com, and to hopefully look more strictly at the recreational rather than show horse market. An example of one of the horses that was eliminated was an unregistered horse of an unknown breed that was capable of jumping at the highly competitive prelim level, and therefore its value was based not on characteristics included in this model but on its show potential. As we do not have the actual selling price of most of the horses in the fall sample or any of the horses in the spring sample, we assume a positively correlated relationship between the asking price and selling price. Only horses listed as sold were included in an attempt to capture prices closer to actual market price as opposed to a price the seller wishes their horse to achieve. Additionally, owners of horses not listed as sold could not be surveyed in the fall data set to obtain actual selling prices, as that price would not exist.

PHorse

This variable, as mentioned above, is the dependent variable in both the Fall and Spring dreamhorse.com models, taken as the Asking Price in dollars per horse listed in the advertisement.

APHorse

The actual transaction price of the horse was obtained through a survey of sellers conducted through dreamhorse.com's e-mail contact. This price was on average slightly lower than the asking prices denoted by the *PHorse* variable. This value is used only for the Fall Survey model.

Independent Variables

The explanatory variables here consist mostly of binary variables, with the notable exceptions of *Age*, *Age2*, and *Tmp*. Each variable is described and defined below as the industry has many complexities not easily understood by individuals not involved in the industry. Summary statistics on the variables are included in Chapters 5 for the Spring, 6 for the Fall, and 7 for the Fall Survey data.

Type of Ad

Sellers get to choose the type of ad they wish to advertise with when they post the ad for their horse on dreamhorse.com. Text ads are free, while ads with photos require a fee paid to dreamhorse.com. *Pht* is a dummy variable which takes the value of 1 if the ad has a photo in it – this can include 90-day or 6-month photo ads with one or multiple photos and possibly video ads. There were too few spotlight ads and video ads to make a category for them alone. This variable takes the value of 0 only if the ad is a free text ad with no photos of any kind, requiring no purchase from dreamhorse.com. The expected sign of this variable is positive, as photo ads are listed at the top of buyer's searches, are therefore likely seen more and provide more information through more allowed characters in the description and the photograph itself. Also worth noting is that if a seller initially values the horse at a greater value they may be more willing to spend the money to post a photo ad, as it is a lower percentage of what they expect in return than it would be for a more cheaply priced animal. Therefore, other traits that add value to the horse, such as training level or discipline, may be captured by this variable.

State

This study looked only at what dreamhorse.com called the "Prairie USA" region. The states of Kansas, Missouri, North Dakota, South Dakota, Nebraska, Oklahoma, and Texas were included in this region. The variables *KS*, *MO*, *NE*, *OK*, and *TX* assumed values of 1 if the horse was sold from (is located in prior to being sold) the states of Kansas, Missouri, Nebraska, Oklahoma, and Texas, respectively, and 0 if the horse was from a different state. The states of North Dakota and South Dakota had fewer horses for sale than the other states in the region, and therefore were combined into one variable, *DK*, which takes a value of 1 if the horse was listed in the ad as from North Dakota or South Dakota and 0 if it was from another state. These variables did not have expected signs, as there is no reason to assume that a horse from any state will sell better or worse than any other state. Some differences may be seen between states with greater supply or less supply due to travel costs to move horses between states, but since there is no measure for state's demand, it is difficult to make a prediction on expected signs. The variable *KS* is left out of the equation here to provide a basis for comparison of the other variables. As there is no previous research to compare the horse price markets across these states, one could only take an educated guess at the expected signs of these variables, and therefore no predictions will be made regarding sign.

Breed

As mentioned above, there are 195 possible breeds (including a few different species) and 200 possible secondary breeds for horses listed for sale on dreamhorse.com. Breeds are a required category for posting an ad on dreamhorse.com, therefore every horse in this sample has a breed. Some breeds are far more common and popular than others, and therefore have variables specific to that breed. Other breeds are much less common and therefore were aggregated into single breed variables by use or type of breed if possible. The variables listed and described below are all dummy variables in relation to the breed and/or secondary breed of the horses. Some individual horses in the data set may have more than one breed in this category due to crossbreeding, and another binary variable is included to capture the positive or negative effects of a horse being crossbred. None of the breed variables have expected signs, because once again there is no legitimate background or comparison between the values of these breeds. However, producer opinion suggests that the crossbred term could likely be negative due to the desire in some sectors of the equine industry to produce and keep purebred animals, due to the uncertainty of which breed's traits the animal would tend towards having, or in the event that a crossbred animal was an accidental breeding with uncertain potential.

Pnt

This variable takes on the value of 1 if either the primary or secondary breed of the horse was Paint, or any of the Paint color secondary breeds (which are Paint Overo, Paint Solid, Paint Tobiano, or Paint Tovero). These secondary breeds simply have to do with the specific color pattern of the horse's coat, and if Paint was the primary breed and one of these was the secondary breed, the horse was not considered a crossbred. If the primary breed was a different breed, such as Quarter Horse, and the secondary breed was one of these secondary paint breeds, then the horse was considered a crossbred between the first breed, such as Quarter Horse, and Paint.

QH

This variable takes a value of one if either the primary or secondary breed of the horse was listed as Quarter Horse. Quarter Horses are by far the most common breed on this website and are also the most common breed in the United States. They are known for their cow sense, quickness and athleticism over short distances, and versatility, and therefore are an incredibly popular breed.

Pony

This variable is an aggregate variable of all pony breeds and miniature horses. In the spring data, the following breeds were included in this category: the general category of "Pony," Shetland Pony, Miniature, American Paint Pony, Welsh Pony, POA, and Missouri Fox Trotter Pony. In the fall data, the following breeds were aggregated into this category: Miniature, POA, Quarter Pony, Welsh Pony, Pony, Shetland Pony, Welara Pony, and Welsh Cob. This variable takes on a value of 1 if any of the above breeds mentioned were the primary or secondary breeds of the horse. Occasionally, a pony will be a cross between two different breeds in this category, in which case the crossbred term is marked to take a value of 1.

TB

This variable takes on a value of one if either the primary breed or secondary breed of the horse is Thoroughbred. This variable also takes a value of 1 for Appendix Quarter Horses, which are horses that are registered or bred as Quarter Horses but are crossed with Thoroughbreds in recent generations, and carry an "X" in front of their AQHA registration number to denote this fact. An Appendix Quarter Horse is a first generation cross between either a Thoroughbred and a Quarter Horse or an Appendix Quarter Horse and a Quarter Horse (Wikipedia, 2008).

ClrBrd

This term stands for Color Breed, and is an aggregate of breeds which are named or identified by a specific color. The Paint breed could also be considered a color breed, but being a common breed it was not aggregated into this category because there were great enough numbers in the data set that it could be an independent breed category. The Palomino registry and Buckskin registry organizations were not included in this category unless they were the primary or only breed the horse had listed, as they are simply color breeds and this information is captured in the color category. It takes the value of 1 if any of the following breeds were listed as the primary or secondary breed of the horse for sale for the spring data set: Appaloosa, Pinto, Spotted Saddle Horse, American Crème and White (Albino), American Buckskin (ABRA), Champagne Horse, Warmbloods of Color. The fall data set had the following breeds in this category: Palomino Horse Breed Association (PHBA), Appaloosa, Pinto, and American Buckskin (ABRA).

Arab

This variable takes the value of 1 if either the primary or secondary breed of the horse for sale was listed as Arabian, Arabian-Cross, or Half-Arabian. This breed is not as common on dreamhorse.com as Quarter Horse and Paint, but Arabians cater to a different industry segment and have a strong reputation for certain traits, so it is important to separate them from other breeds and look at them as a stand-alone group.

Saddle

This variable is an aggregated group of a large number of other breeds. This category includes horses that are typically used as saddle horses, such as the American Saddlebred, or sport horses, such as Warmbloods. The following breeds were included in this category in the spring dreamhorse.com data: Haflinger, Missouri Fox Trotter Horse, Morgan, Tennessee Walker, Spotted Saddle Horse, Paso Fino, Friesian Cross, Trakehner, Dutch Warmblood, Hanoverian, Standardbred, American Saddlebred, Warmbloods of Color, Lipizzan, and Racking Horse. The following breeds were included in this category in the fall data set: Missouri Fox Trotter Horse, Morgan, Tennessee Walker, Friesian Cross, American Warmblood Society, Rocky Mountain Horse, Paso Fino, Colorado Rangerbred, American Saddlebred, International Sport Horse, Trakehner, Andalusian, Azteca, Lipizzan, Half Dutch, Oldenburg, and United Mountain Horse.

Other

This category captured all other breeds and species listed on dreamhorse.com for the two time periods data were collected. All of the animals sold during these time periods were equids, so no cattle, goats, alpacas, or other species are included in this category. The breeds/types of equids included in this category in the spring data were: Draft Cross, Mustang, Mule, Spanish Mustang, Other, Donkey, Shire, Draft, and Belgian. The fall data included the following breeds and types of equids: Donkey, Mustang, Draft Cross, Percheron, Miniature Donkey, and Mule.

Cross Bred Term

The variable *Xbred* is another binary variable that takes a value of 0 if the animal is a purebred and 1 if the animal is listed as a crossbred. In some cases, the horse will be crossbred between two of the categories above, while other times the breeds will both be in the same category (such as ponies or saddle horses) and still be crossbred, so this term will still take the value of 1.

Gender

This is another category of binary variables and is based on what the seller labeled their horse when posting it online. On dreamhorse.com, the possible selections for gender are Stallion, Mare, Mare In Foal, Filly, Gelding, Colt, or Unborn Foal. No unborn foals were posted for sale in these data sets, so that variable is not included. Here, they are categorized by fixed male horses (Geldings), male horses (Studs/Colts), and female horses (Mares/Mares in Foal/Fillies).

Geld

This variable denotes geldings, or fixed male horses, and only includes horses listed in their advertisements by their owners as geldings. It is much more common in the equine world, as in other livestock species, to geld male horses as opposed to spaying mares. If the horse is listed by the seller as a gelding, this variable takes on the value of 1.

Stud

This category includes horses listed by their owners as stallions or colts. Studs, or stallions, are adult male horses that are usually used for breeding purposes and also can be ridden and shown. If the sex of the horse for sale is listed by the seller as Stallion, this variable takes the value of 1. Colts are young male horses. Different sellers may have different ages where they consider a colt to become a stallion, or a young animal that is gelded could be a young male horse but still a gelding, not a colt. This variable takes the value of 1 if the seller listed the animal as a colt in the ad.

Mare

This category includes horses listed by their owner as mares, mares in foal, or fillies. Mares are adult female horses. Once again, the age at which a filly turns into a mare is not specifically defined in the equine world, so this is based on if the ad lists the horse as a mare. If so, this variable takes the value of 1. Mares in foal are also included, as they are simply female horses carrying a foal. The *InFoal* variable is added to account for this difference. Similar to a colt, a filly is a young female horse, and the age at which a horse turns from a filly to a mare is a matter of opinion and differs across the industry. Therefore, this variable is based on what gender the seller listed the animal as on dreamhorse.com. If the seller listed the animal as a filly, this variable takes the value of a 1.

Mares in Foal

The variable *InFoal* takes on a value of a 1 if the horse is listed as in foal, or currently pregnant. This is a binary variable designed to capture increases or decreases in value considering the fact that by purchasing this horse you are also purchasing the foal it carries. This is separate from the gender variables and accounts for if the ad is a "two-in-one" package, a mare in foal, or not.

Color

The variables related to the color of the horses are listed below. As with breed, color is a required category on dreamhorse.com and therefore every horse in this sample has a color. Also worth noting here is that though most breeds of horses can be a wide variety of colors, some particular breeds can only be a certain color or range of colors. Individuals in the equine industry often have a favorite color of horse, or alternately breeders try to breed for specific colors of horses if they are well known for a certain color or are going for a certain genetic factor, such as dilute or dun factor genes. The only concern with this group of variables is that it is up to the seller to correctly report the color of the horse, which can be difficult to do. If the horse has not been genetically tested, it is often difficult to correctly identify the true genetic color of the horse. Seal Brown and Black Bays, for example, are often called black by owners even if not genetically black. Some grays are also called roans by their sellers if the seller does not know how to correctly identify roan versus gray. Without genetically testing and visually inspecting every horse, we cannot be certain the colors are correct, so for the purpose of the data set we are using the color that the seller identified the horse as in the advertisement. Additionally, no variable was removed from this group of binary variables due to the fact that horses can be more than one color. The majority of horses that are included in more than one color category are in the *Paint* category and one other color. One horse in the fall data set was in both the dilute and dun factor color categories, as it was buckskin with dun factor markings. As the horse carried both genetic factors, it is assumed that the values associated with both variables applied to this horse, and therefore it was left in the data set. Color descriptions were obtained from

Wikipedia.com. Due to the fact that different individuals have different favorite colors, no expected values are assigned to the individual colors.

Bay

This variable takes the value of a 1 if the horse is a bay, which is a color where the horse's body color can range from a light reddish-brown to a dark brown with a black mane, tail, and legs, also called black points.

Sor

Sorrel is a dark or light red coat color, and is sometimes called chestnut. This variable is a 1 if the seller listed the horse as sorrel.

Blk

True black horses are relatively less common than other coat colors such as sorrel and bay. Dark brown and black bay horses are often mistaken as black horses, but it is impossible to tell if this is the case for any of the black horses in this data set. This variable takes the value of a 1 if the seller listed the horse as black.

Grey

Grey horses can be born any color and usually lighten with age. They have black skin with white or mixed hairs. If the horse in the data set was labeled as a grey by the seller, this variable takes on the value of a 1.

Pal

This variable stands for Palomino, which is a common dilute color in this data set. Due to being much more common than other dilute colors, it was aggregated into an individual category. The two colors included in this data are Palomino and Chocolate Palomino. Palomino horses genetically carry one copy of what is known as the dilution gene, and without this gene would be chestnut or sorrel but because of the gene turn a golden color with a flaxen mane and tail. If the horse was listed as a palomino or chocolate palomino by the seller, this variable takes the value of 1.

Paint

This is a variable denoting the color pattern pinto, which means the horse has large patches of multiple colors on its body. One color is white while the other can be any range of different colors and occasionally two other colors, such as black and brown. In this data set, this variable is used in combinations with other colors, with the base color (non-white) taking a value of one in addition to this color pattern variable taking a value of 1 if the horse was listed as a paint or pinto colored animal.

Dun

In this data set, dun is the variable used to describe horses listed as grulla, red dun, dun, or dun with black points. Dun is actually a genetic factor which causes the horse to have markings, such as a dorsal stripe, masking, zebra stripes on the legs, and other distinct "dun factor" markings. This gene is often bred for and found to be desirable in certain breeds. If the horse was listed as any of the mentioned colors, this variable takes the value of a 1.

Dilute

This variable includes any color, other than palomino, that is affected by a dilution gene. These colors include Buckskin, which is bay with one copy of the dilution gene, Champagne, Cremello, Smoky Black, Perlino, or Silver Dapple. The genetic factor causing these colors is not the same gene in all cases though they all cause the horse's body color to be lighter than the base color, but due to these horses being fairly uncommon they were aggregated into this category. This variable will take the value of a 1 if the horse is any coat color affected by a dilution gene and 0 if not, as listed by the seller.

Roan

Roans, which can often be confused with or listed as other colors, can be any base color with white hairs evenly dispersed throughout their coat, except for having darker, sometimes solid colored, heads. These horses do not lighten throughout their lifetime like grays do. If the horse is listed by the seller as any kind of a roan, including red, bay, or blue roan, it will take a value of a 1 for this variable. Blanketed appaloosas were also included in this color category as many of them express roan characteristics.

ChBr

This color category included several common colors which did not have enough animals to be in their own category. Chestnut, which is a common coat color where the horse's body is brown or yellowish, with no red tint, and Brown, which can be genetically bay or genetically chestnut horses but are indeterminate upon observation and, having not been DNA tested, are classified as brown. This category also includes the color "Chocolate" in the fall data set.

Method of Sale

After a horse is sold the seller has the option of marking the horse as "Sold" or "SOLD HERE," to denote which horses were sold using off-website methods and which sold using the dreamhorse.com website. *Online* is the variable that takes a value of a 1 if the horse was marked as being sold on the website and 0 if the horse is just marked as sold. This is to see if the horses sold online get higher or lower prices than horses sold through other methods, holding all else constant.

Age Terms

Horses have a long lifespan in comparison to other livestock species, and their useful or rideable life is limited in the early years by size, growth, and training level and later in life by injuries, remaining useful life, and health issues. Therefore, the expectation is that the value of a horse increase at a decreasing rate with age to a peak and then decrease, which is why an age squared term is added.

Age

This variable is the horse's age in years. If the horse was under a year of age, this took a value of 0.5, assuming the foal is 6 months of age, to allow for the log-log function to take the natural log of this value as the natural log of zero is undefined.

Age2

This is the squared age term to account for an increasing value in young horses and decreasing value with older horses. In all log-log models, this variable is actually the squared log of age, not the log of age squared as this would cause perfect multicollinearity in the age terms.

Temperament Score

The seller is allowed to choose a temperament score for their horse from a scale of 1 to 10, with 1 being calmer and 10 being more spirited. The temperament score is very subjective, based on the seller's opinion of how "hot" or "calm" their horse is. The variable *Tmp* is simply the temperament rank that the owner gave their horse. No horse had a higher temperament score than 8, and the averages in each data set were around 3, probably due to the idea that sellers would like buyers to think their horse is calm, so as temperament score increases the price would be expected to decrease. Typically horses with a lower temperament score would sell for a greater price due to being easier to work with and more useful for kids and inexperienced riders. If no temperament was listed by the owner of the horse, the sample average temperament of the horses that were listed was assigned to this horse to allow for the natural log to be taken in the log-log functional form and to prevent a bias from occurring in this variable resulting from these horses having the lowest temperament score of zero. The expected sign on this variable is negative, as more spirited animals tend to be more difficult to handle and therefore less desirable and less valuable than calmer horses. Squaring temperament was attempted to examine a non-linear relationship with the price of the horse, but was consistently insignificant and therefore was removed from the model.

Registration

A horse being registered means the horse's pedigree information, and sometimes DNA, show records, color, age, sex, offspring, and other important facts about the horse are recorded in a breed, color, or performance registry. Registries also commonly record ownership information and transfers throughout the horse's lifetime. The 2007 AHC Horse Industry Directory listed 125 Breed Organizations in the U.S., and 66 Show and Sport Organizations (some of which were also breed organizations). On dreamhorse.com, there is a field to select whether the horse is registered or not as a Yes or No (the default setting is No), and two additional fields to list the organization the horse is registered with, which the seller fills in, and the registration number. With certain organizations you can look up a horse's information by their registered name or number; however, most only give this privilege to members of the organization. Registration of a horse can be extremely valuable with some organizations, as it allows the horse to be shown on major show circuits and also allows the offspring of those horses to be registered with the organization in many cases. Parentage verification through registration or DNA tests also help owners verify the horse has the bloodlines listed, which is important to many buyers. If a horse is not registered, it is called a “grade” horse.

Another term for a horse being registered is for a horse to be “papered” or have “papers.” Therefore, the variable *Paper* is a binary variable taking the value of a 1 if the horse is registered with an organization (or in the cases of some young colts and fillies if the paperwork to register them is pending or has been filled out but not sent it). The expected sign of this variable is positive, due to the added value of parentage verification, ability to show in certain show circuits, and ability to register future offspring.

Empirical Models

Through a series of testing various regression models on the collected data, including OLS and Generalized Method of Moments techniques, the following final regression models were obtained. A Box-Cox transformation was attempted to distinguish whether a linear or semi-log functional form was more appropriate, which is a transformation applied to the equation and which returns a λ value between zero and one, with zero indicating a semi-logarithmic equation and one indicating a linear function is the best fit. If the value is in between the two, the best fit for the data is somewhere between linear and semi-log. In this case when the transformation was initially run a corner solution was found and the λ had to be built into an equation forcing it to stay between zero and one. Even after this was accomplished, due to apparent multicollinearity issues during the test and differing results for each of the data sets when the transformation successfully ran, the results were inconclusive. When *Age2* was included in the Box-Cox transformation, both the fall and spring data sets came back with biased results, while the fall survey returned a λ value of 1 indicating that a linear model would be the best. The biased results simply did not return a value for λ and indicated R^2 values that were negative or greater than one. By removing *Age2*, we eliminate the bias in the models and return λ values of 1 for the fall and fall survey data and $1.327*10^{-15}$, or nearly zero, for the spring data. This indicates that without age squared the two models using fall data are best represented using linear models and the spring data using a semi-logarithmic model. The author chose to present the results of three functional forms for each data set – a simple linear form, a semi-log form where the natural log is only taken for the dependent variable, *PHorse*, and a log-linear functional form where all of the non-binary variables in the data set are their natural logarithms – due to the issues experienced with the Box-Cox transformations. To avoid multicollinearity issues, two variables were removed from the regression to provide a basis for comparison for the other binary variables in that group. The state of *Kansas* is not included in the model, to provide comparison for the other states. Also, the gender variable *Gelding*, a fixed male horse, is removed from the data set so that the variables *Mare* and *Stud* can be compared to it. These two variables are removed to prevent perfect multicollinearity in their respective categories. The term *i*, for individual horses, is dropped here for the sake of convenience, but each equation is estimated using each horse's data, and *e* is the error term. Equations (2), (3), and (4) below represent the functional forms of the model to be estimated:

Linear Model:

$$PHorse = \beta_0 + \beta_1 \cdot Pht + \sum_{h=1}^{5} \beta_{1+h} \cdot State_h + \sum_{j=1}^{8} \beta_{6+j} \cdot Breed_j + \beta_{15} \cdot Xbred + \sum_{k=1}^{2} \beta_{15+k} \cdot Gender_k + \beta_{18} \cdot InFoal + \sum_{l=1}^{10} \beta_{18+l} \cdot Color_l + \beta_{29} \cdot Online + \beta_{30} \cdot Age + \beta_{31} \cdot Age2 + \beta_{32} \cdot Tmp + \beta_{33} \cdot Paper + e \quad (2)$$

Semi-Log Model:

$$lnPHorse = \beta_0 + \beta_1 \cdot Pht + \sum_{h=1}^{5} \beta_{1+h} \cdot State_h + \sum_{j=1}^{8} \beta_{6+j} \cdot Breed_j + \beta_{15} \cdot Xbred + \sum_{k=1}^{2} \beta_{15+k} \cdot Gender_k + \beta_{18} \cdot InFoal + \sum_{l=1}^{10} \beta_{18+l} \cdot Color_l + \beta_{29} \cdot Online + \beta_{30} \cdot Age + \beta_{31} \cdot Age2 + \beta_{32} \cdot Tmp + \beta_{33} \cdot Paper + e \quad (3)$$

Log-Log Model:

$$lnPHorse = \beta_0 + \beta_1 \cdot Pht + \sum_{h=1}^{5} \beta_{1+h} \times \cdot State_h + \sum_{j=1}^{8} \beta_{6+j} \cdot Breed_j + \beta_{15} \cdot Xbred + \sum_{k=1}^{2} \beta_{15+k} \cdot Gender_k + \beta_{18} \cdot InFoal + \sum_{l=1}^{10} \beta_{18+l} \cdot Color_l + \beta_{29} \cdot Online + \beta_{30} \cdot lnAge + \beta_{31} \cdot (lnAge)^2 + \beta_{32} \cdot lnTmp + \beta_{33} \cdot Paper + e \quad (4)$$

In the following chapters, the results of these models for each of three data sets (Spring, Fall, and Fall Survey) will be presented and compared using these variables and functional forms. An attempt will be made to discern what traits add to or subtract from the value of a recreational or pleasure horse, one at the lower end of the equine industry as compared to horses previously studied in pricing models. These Ordinary Least Squares (OLS) or Generalized Method of Moments (GMM) regressions will be performed using SAS 9.1, depending on the results of a White's Test for heteroskedasticity due to the cross-sectional nature of the data.

Additionally, the age at which a horse is most valuable will hopefully be obtained. Also, the level of impact will be discerned through the marginal values on the statistically significant variables; the factors most significantly impacting value of a recreational horse will be identified and quantified.

CHAPTER 5 - Spring 2008 Data, Results and Interpretation

Summary Statistics

Table 7 on the next two pages contains the summary statistics on the data collected on sold horses in March of 2008 from dreamhorse.com. The means of the binary variables are the percentage of horses in the survey which were in that category, and the sums of the binary variables are the number of horses in that category. The total number of observations collected in the spring was 470, as denoted by the column N. There were 192 ads that were photo, video, or spotlight ads in this data set, or about 41% of the ads. This variable has a positive expected sign as the added information and exposure in a photo ad should allow the horse to be seen by a greater market and therefore sold for a higher price. In the state category, the majority of horses, 54%, were sold from the state of Texas, while only 12 horses, or 2.5%, were from the state of Nebraska and 19 from North and South Dakota combined. This is not surprising as several major horse organizations, including AQHA and APHA, are based in Texas and the state has a long tradition of horsemanship. None of the states have expected signs, due to no previous research on recreational horse prices in these various states, and Kansas is removed from the model as a base variable to prevent perfect multicollinearity and therefore no sign prediction is applicable.

Quarter Horses, being the most common breed in the United States, represent the majority of breeds at 51% of the animals, while Arabians, being a less common breed and not being aggregated with other breeds, have the lowest numbers in this data set at 19 horses, or 4% of the sample. As with states, no previous research presents a reason to project signs on this group of variables. Eleven percent of the horses in this sample are crossbred, which constitutes 54 horses, and this variable has a negative expected sign due to producer opinion that purebred, pedigreed horses generally being preferred to crossbred horses in the equine world. Like the state Kansas, geldings are removed from the gender variable group to provide comparison. The sample has a majority of female horses, with over 50% being mares or fillies. Uncut male horses constitute the lowest portion of horses at 11%. In comparison to geldings, mares and stallions both have a negative expected value in the pleasure horse world, due to stallions being more difficult to handle and mares being known as moody due to heat cycles. Geldings have a reputation for being more solid and having more even temperaments, and therefore appeal to a wider variety of recreational horse riders. Only 5.3%, or 25, of the mares in the sample are in

Table 7: Summary Statistics of Horses Sold in the Spring

Group	Label	Description	Mean	Std. Dev.	Min	Max	Sum	Sign
	Independent Variables:							
	Pht	Photo Ad	0.40851	0.49208	0	1	192	+
State	*KS*	Kansas	0.07660	0.26623	0	1	36	N/A
	MO	Missouri	0.15745	0.36461	0	1	74	?
	DK	The Dakotas	0.04043	0.19717	0	1	19	?
	NE	Nebraska	0.02553	0.15790	0	1	12	?
	OK	Oklahoma	0.15745	0.36461	0	1	74	?
	TX	Texas	0.54255	0.49872	0	1	255	?
Breed	*Pnt*	Paint	0.21489	0.41119	0	1	101	?
	QH	Quarter Horse	0.51277	0.50037	0	1	241	?
	Pony	Pony Breeds	0.08085	0.27290	0	1	38	?
	TB	Thoroughbred	0.06383	0.24471	0	1	30	?
	ClrBrd	Color Breeds	0.05106	0.22036	0	1	24	?
	Arab	Arabian	0.04043	0.19717	0	1	19	?
	Saddle	Saddle Breeds	0.08511	0.27934	0	1	40	?
	Other	Other Breeds and Equids	0.04043	0.19717	0	1	19	?
	Xbred	Crossbreds	0.11489	0.31923	0	1	54	-
Gender	*Geld*	Geldings	0.36809	0.48280	0	1	173	N/A
	Mare	Female Gender	0.51702	0.50024	0	1	243	-
	Stud	Male Gender	0.11489	0.31923	0	1	54	-
	InFoal	Mares in Foal	0.05319	0.22465	0	1	25	-
Color	*Bay*	Bays	0.22553	0.41838	0	1	106	?
	Sor	Sorrel	0.22979	0.42114	0	1	108	?
	Blk	Black	0.10213	0.30314	0	1	48	?
	Grey	Grey	0.09362	0.29161	0	1	44	?
	Pal	Palomino	0.06596	0.24847	0	1	31	?

Color	*Paint*	Pinto	0.14681	0.35429	0	1	69	?
	Dun	Dun Factor	0.05745	0.23294	0	1	27	?
	Dilute	Dilute Gene	0.05957	0.23695	0	1	28	?
	Roan	Roan	0.06809	0.25216	0	1	32	?
	ChBr	Chestnut/Brown	0.09574	0.29455	0	1	45	?
	Online	Sold Online	0.42128	0.49429	0	1	198	?
	Age	Age	7.32021	5.35331	0.50	26	3,441	+
	Age2	Age Squared	82.18245	112.98345	0.25	676	38,626	-
	Tmp	Temperament	3.04590	1.52216	1	8	1,432	-
	lnAge	Natural Log of Age	1.67884	0.86122	-0.69315	3.25810	789.05682	+
	(lnAge) ^{2}	Natural Log of Age Squared	3.55864	2.64122	0	10.61519	1,673	-
	lnTmp	Natural Log of Temperament	0.97080	0.56320	0	2.07944	456.27572	-
	Paper	Registration	0.72766	0.44564	0	1	342	+
	Dependent Variables:							
	PHorse	Asking Price of the Horse	1,784.14	1,240.46	1.00	5,000.00	838,547.00	N/A
	lnPHorse	Natural Log of the Asking Price	7.18434	0.96465	0.00000	8.51719	3,376.63931	N/A

foal, and this variable has a negative expectation as most people at the pleasure end of the industry are not looking for a foal as a project. Also, they have to consider the added expenses of caring for two horses instead of just one, which at this end of the market detracts from the value of the horse. In the color category, there are once again no expectations of sign, though Taylor et al. had a positive sign on the color black and a negative sign on chestnut (as compared to sorrel). However, they analyzed a different market segment (show horses) and therefore there is uncertainty as to whether the same predictions will hold. Bays and sorrels are the most common colors in the data set, both having just over 22% of the horses. Dun factor horses and horses carrying the dilute gene, excluding palominos, are the least common colors in the set with less than 6% each. Less than 15%, or 69 horses, are in the paint color category, and most of these will be paint and a base color in a different category. Forty-two percent of the horses were listed as sold on dreamhorse.com, and the variable *Online* has an unknown expected sign.

Age and *Age2* are two of the continuous variables in the data set. The age distribution is shown in Figure 2, and clearly shows a skewed distribution with the majority of horses between

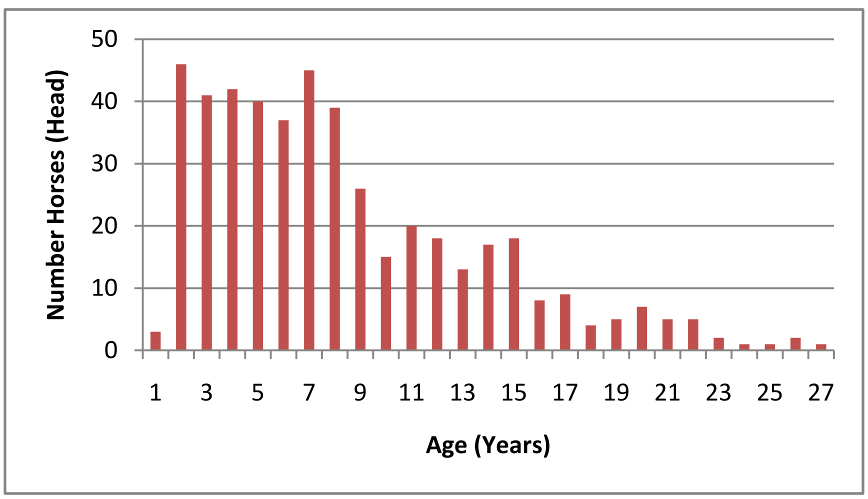

Figure 2: Spring Age Distribution

the ages of 2 and 7 and less horses out to the right going up into the 20+ year age range. The average age of a horse in this data set is 7.32 years of age. The value of a horse is expected to increase with age at a decreasing rate, reaching a peak and finally declining, which indicates the need for the inclusion of an age squared term. Temperament scores are also a continuous variable in this data set, and though the seller can select a temperament score from 1 to 10, in this data set the range is only from 1 to 8 and is skewed right like the age distribution, with the majority grouped around scores of 2 and 3. The average temperament score in this data set is 3.05, and the expected sign is negative. The natural logs of each of the continuous variables are taken for the log-log model, and their summary statistics are also presented in Table 7. Finally, almost 75% of the horses, or 342 out of the 470 in this sample, were registered with some organization. The expected sign on the *Paper* variable is positive as registration can allow participation in organizations and events that grade horses are not allowed in.

The dependent variable in the linear model is *PHorse* and for the semi-log and log-log models is *lnPHorse*. A distribution of *PHorse*, which is simply the asking price of the horses in this sample, is shown in Figure 3. The average asking price for a horse ranged from $1 (basically

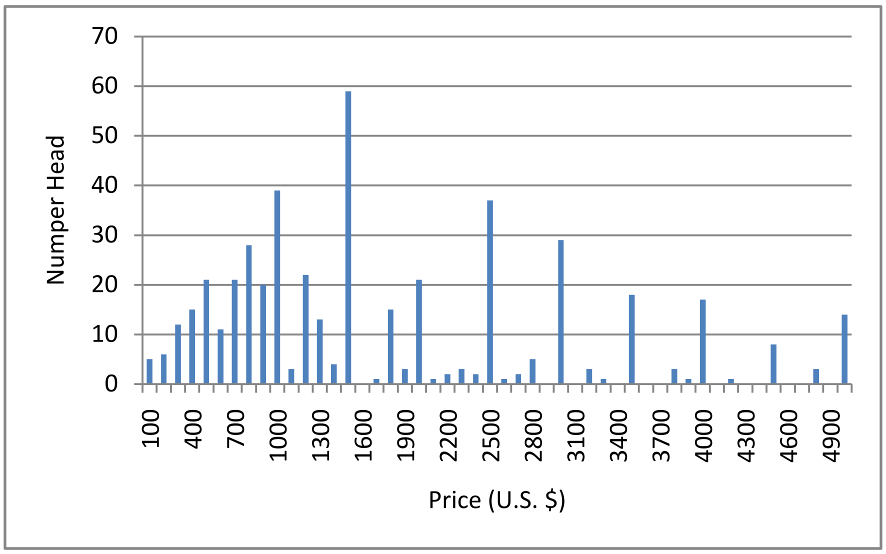

Figure 3: Spring Asking Price Distribution

free) to $5,000 with an average of $1,784 and a median of $1,500. The graph shows that this variable is slightly skewed right, though not as extreme as the graph for age. This dependent variable will be used in the following regression to analyze which traits of horses in this sample affected the asking price of that horse. We are also assuming a relationship between the asking price posted on dreamhorse.com and the final selling price of the horse.

White's Test Results and Correction

Due to the cross sectional nature of the data, a White's Test was run on each of the three functional forms examined with this data set. The results for the spring data set are shown in Table 8. In this case, we cannot reject heteroskedasticity at the 10% level for the linear model and at the 1% level for each of the semi-log and log-log models, and therefore we cannot be certain the standard errors and t-statistics are accurate and they may indicate statistical

Table 8: Spring White's Test Results

Linear Model Heteroscedasticity Test, Spring Data					
Equation	Test	Statistic	DF	Pr > ChiSq	Variables
Phorse	White's Test	429.2	387	0.0685	Cross of all vars
Semi-Log Model Heteroscedasticity Test, Spring Data					
Equation	Test	Statistic	DF	Pr > ChiSq	Variables
lnPHorse	White's Test	464.3	387	0.0042	Cross of all vars
Log-Log Model Heteroscedasticity Test, Spring Data					
Equation	Test	Statistic	DF	Pr > ChiSq	Variables
lnPHorse	White's Test	462.4	387	0.0050	Cross of all vars

significance where none exists. Therefore, each of the three models in this data set were transformed using the generalized method of moments (GMM) modeling method to correct the significance statistics in the results. As OLS assumes constant variance and the GMM model does not, this model can be used to correct for heteroskedasticity issues. None of the coefficients were changed from the OLS model to the GMM, as expected, and therefore direct comparisons can still be made between this and the results of the other two data sets. Additionally, though there were minor changes in the statistical significance of individual variables, none of the variables that were statistically significant with the OLS model became insignificant in the GMM, and no additional variables became significant in GMM that were not under the OLS method.

Group F-Test Results

F-Tests were run on individual groups of binary variables to check for statistical significance as a group in relation to the price of horses. The F-tests were run for each functional form, though with only the binary variables being evaluated in the F-tests the semi-log and log-log functional forms have the same results. The four groups of binary variables which were tested were: States, with Kansas excluded to prevent perfect collinearity; Breeds, where all were included due to crossbreeding preventing perfect collinearity; Colors, where all were included due to the *Paint* variable allowing for multiple colors per horse and preventing multicollinearity; and Genders, where geldings are left out to prevent multicollinearity.

Table 9 shows the results of these F-Tests. Each of the groups State, Breed, and Gender show statistical significance at less than the 1% level in most cases, and at just over the 1% level

Table 9: Spring F-Test Results

Spring Group F-Tests		
States	F-Value	F-Signficance
Linear	3.05**	0.0101
Semi-Log	3.33***	0.0057
Log-Log	3.33***	0.0057
Breeds	F-Value	F-Signficance
Linear	3.18***	0.0016
Semi-Log	2.77***	0.0053
Log-Log	2.77***	0.0053
Colors	F-Value	F-Signficance
Linear	0.99	0.4516
Semi-Log	0.70	0.7288
Log-Log	0.70	0.7288
Genders	F-Value	F-Signficance
Linear	6.45***	0.0017
Semi-Log	4.43**	0.0124
Log-Log	4.43**	0.0124

in the linear form for the States and the semi-log forms for the Genders. However, the group of Color variables shows no statistical significance for any of the functional forms represented here. This is interesting, as in one previous study (Taylor et al., 2004) some of the color variables were statistically significant, but apparently with multiple breeds included color is not as significant. This lack of statistical significance could be due to different colors being more or less common across different breeds and other factors due to the highly heterogeneous nature of the data set. We will see in the overall model if these results hold and if any of the individual state, breed, or gender variables are statistically significant while none of the color variables are, or if these indications do not hold.

Overall Model Results

The results of the GMM model using three different functional forms, linear, semi-log, and log-log, are reported in the following subsections. Statistics of overall fit will be reported and discussed in addition to individual statistically significant variables. Interpretations for this data set will also be reported in this section; comparisons between sample models and conclusions will be drawn in later chapters.

Linear Spring Model Results and Discussion

Table 10 shows the results of the linear model for the spring data. Overall measures of fit suggest a good statistical fit of the model. The R-Squared value is 0.3113, while the Adjusted R-Square is 0.2591, suggesting that the model explains approximately 25.91% of the linear variation in the prices of the horses. The overall F-value is highly statistically significant, suggesting that at least some of the coefficients in the model are significantly different from zero. Only six of the variables are statistically significant; none of the state variables are statistically significant, which seems counterintuitive as compared to the group F-test on states shown in a previous section. As expected, the variable *Pht* is highly statistically significant and positive, indicating that a horse advertised with a photo ad as comparison to a text ad is worth $689.06 more, ceteris paribus. This result could be due to the fact that horses advertised through photo ads draw more attention and get more views, or because sellers advertising for higher prices are not held back by the cost of posting a photo ad, as compared to cheaper horses the cost is a lower percentage of a higher priced horse's value. With photo ads costing from $25-$100 on dreamhorse.com, the additional return appears to be well worth the cost. Saddle horses are valued higher than other breeds, ceteris paribus, by over $600 at the 5% significance level. There are many possible explanations for this result, one of which is the type of activities that certain breeds of saddle horses can participate in, such as eventing, and the rareness of such breeds. Another explanation is that the model is simply incorrect and a linear model is not the right one to explain the relationship between breeds and price of horse. The variable mare is highly statistically significant and suggests that a mare will bring $360.02 less than a gelding, ceteris paribus. This is as expected, as mares traditionally have a reputation for being moodier than geldings and therefore may be less desirable in the recreational industry, where breeding is not a major consideration. Uncut male horses were not statistically different than geldings in this

Table 10: Spring Linear Model Results

Spring Dreamhorse.com Linear Model Results (Dependent Variable = *PHorse*)				
N	R-Square	Adj R-Sq	F-Value	F-Significance
470	0.3113	0.2591	5.97	<.0001
Variable	Parameter Estimate	Standard Error	t-Value	P-statistic
Intercept	328.51020	425.60	0.77	0.4406
Pht	**689.05510*** **	**126.90**	**5.43**	**<.0001**
MO	-181.99800	199.30	-0.91	0.3616
DK	309.02240	325.00	0.95	0.3422
NE	607.71260	501.20	1.21	0.2259
OK	83.35292	202.60	0.41	0.6810
TX	167.05380	175.10	0.95	0.3405
Pnt	-107.78400	280.10	-0.38	0.7006
QH	212.12410	257.00	0.83	0.4095
Pony	-233.15000	331.60	-0.70	0.4823
TB	-89.79030	329.20	-0.27	0.7852
ClrBrd	-313.99200	248.70	-1.26	0.2075
Arab	-39.54520	369.00	-0.11	0.9147
Saddle	**629.10940** **	**316.10**	**1.99**	**0.0472**
Other	30.54188	286.80	0.11	0.9152
Xbred	165.52130	211.60	0.78	0.4346
Mare	**-360.02300*** **	**120.90**	**-2.98**	**0.0031**
Stud	-140.93200	185.00	-0.76	0.4467
InFoal	-304.25200	228.20	-1.33	0.1831
Bay	13.38589	247.10	0.05	0.9568
Sor	-48.23380	225.30	-0.21	0.8306
Blk	147.28200	189.90	0.78	0.4385
Grey	19.90138	291.50	0.07	0.9456
Pal	139.89600	295.30	0.47	0.6359
Paint	132.42780	156.70	0.84	0.3986
Dun	159.17220	355.40	0.45	0.6545
Dilute	55.41991	288.60	0.19	0.8478
Roan	392.24000	284.70	1.38	0.1690
ChBr	269.50000	235.70	1.14	0.2535
Online	-11.20550	98.75	-0.11	0.9097
Age	**158.72030*** **	**41.06**	**3.87**	**0.0001**
Age2	**-7.02259*** **	**2.17**	**-3.24**	**0.0013**
Tmp	-42.27490	36.60	-1.16	0.2487
Paper	**850.19310*** **	**129.20**	**6.58**	**<.0001**

regression. As expected, none of the color variables were statistically significant in this model. Being sold online also did not carry any statistical significance in relation to the asking price of the horse. Registration was highly statistically significant and also had the greatest impact on the price of horses, with registered horses being worth $850.19 more than grade horses, ceteris paribus.

The age variables are also of interest. Both were highly statistically significant and had the expected sign. Therefore, we can use these variables to figure out the age at which a horse is at its highest value in this data set:

$$\frac{\partial PHorse}{\partial Age} = 158.7203 + 2(-7.02259) \text{ x } Age = 0$$

$$158.7203 - 14.04518 \text{ x } Age = 0$$

$$158.7203 = 14.04518 \text{ x } Age$$

$$Age^{*} = 11.3007 \text{ years.}$$

These results indicate that a horse is at its highest value at just over 11 years of age. This is not surprising as typically in the equine world producers and owners consider a horse at their prime between 8 and 12 years of age. This age is a good balance between years of experience making the animal well broke and useful for any level of rider, and still having a great deal of longevity left before the animal's usable lifespan runs out.

Semi-Log Spring Model Results and Discussion

The results of the semi-log model, where the dependent variable is *lnPHorse* but the right hand side variables are all still in linear form, are shown in Table 11. Ten variables in this regression showed statistical significance at the 15% level, as well as the intercept. This model has overall fit statistics similar to those of the linear model, though the R-squared values cannot be directly compared due to the change of the dependent variable. The R-squared in this case is 0.2671 while the adjusted R-squared value is 0.2116, indicating that 21.16% of the linear variation in the dependent variable is explained by the data. The F-statistic of 4.81 is highly significant as well. In this model, the state variables indicated much greater statistical significance and at least one breed and gender variable were also significant, while no color variables were statistically significant, as expected based on the earlier F-tests.

Table 11: Spring Semi-Log Model Results

Spring Dreamhorse.com Semi-Log Model Results (Dependent Variable = *lnPHorse*)				
N	R-Square	Adj R-Sq	F-Value	F-Significance
470	0.2671	0.2116	4.81	<.0001
Variable	Parameter Estimate	Standard Error	t-Value	P-statistic
Intercept	**6.42170***	**0.4549**	**14.12**	**<.0001**
Pht	**0.35544***	**0.0858**	**4.14**	**<.0001**
MO	-0.02817	0.1688	-0.17	0.8675
DK	**0.31354***	**0.2132**	**1.47**	**0.1422**
NE	**0.43899****	**0.2337**	**1.88**	**0.0610**
OK	**0.25753****	**0.1513**	**1.70**	**0.0895**
TX	**0.29247****	**0.1460**	**2.00**	**0.0457**
Pnt	-0.43512	0.4289	-1.01	0.3109
QH	-0.16949	0.3595	-0.47	0.6376
Pony	**-0.52898***	**0.3524**	**-1.50**	**0.1340**
TB	-0.39213	0.4320	-0.91	0.3646
ClrBrd	-0.30706	0.2416	-1.27	0.2045
Arab	-0.13740	0.2857	-0.48	0.6308
Saddle	0.09038	0.4136	0.22	0.8271
Other	-0.29899	0.3127	-0.96	0.3395
Xbred	0.02015	0.1910	0.11	0.9160
Mare	**-0.28041*****	**0.0738**	**-3.80**	**0.0002**
Stud	-0.14472	0.1234	-1.17	0.2415
InFoal	-0.11668	0.1424	-0.82	0.4132
Bay	-0.12071	0.2833	-0.43	0.6702
Sor	-0.05330	0.2759	-0.19	0.8469
Blk	-0.06547	0.2668	-0.25	0.8063
Grey	-0.02580	0.2896	-0.09	0.9290
Pal	0.10433	0.2847	0.37	0.7142
Paint	0.18667	0.1447	1.29	0.1977
Dun	0.01967	0.2973	0.07	0.9473
Dilute	0.01124	0.2862	0.04	0.9687
Roan	0.25117	0.2802	0.90	0.3706
ChBr	-0.05283	0.2868	-0.18	0.8539
Online	-0.03694	0.0699	-0.53	0.5973
Age	**0.12468*****	**0.0280**	**4.46**	**<.0001**
Age2	**-0.00589*****	**0.0015**	**-3.99**	**<.0001**
Tmp	-0.03194	0.0261	-1.22	0.2225
Paper	**0.71816*****	**0.1886**	**3.81**	**0.0002**

Using this functional form, the intercept became statistically significant which it was not in the linear model. Photo ads are also once again highly statistically significant, and indicate that if a horse is advertised using a photo ad its price will be 35.5% higher than a horse sold using a text advertisement. As mentioned previously, this may be due to higher quality horses being advertised using photo ads or due to the increased exposure garnished, but 35.5% of even a horse priced at $100 is still a greater return than the $25 cost of a basic photo ad. Three out of five state variables in this regression showed statistical significance at the 10% level, and a fourth at the 15% level. Of these, the state of Nebraska had the greatest positive impact on price as compared to the state of Kansas, as it suggests that a horse's price if sold from Nebraska would be 43.9% higher, ceteris paribus. Texas and Oklahoma were 29.2% and 25.8% higher than Kansas, respectively; and the Dakotas, which were only statistically significant at the 15% level, indicated a 31.4% increase in price over horses sold in the state of Kansas, ceteris paribus. *Pony* was the only breed variable to express statistical significance, and it was only at the 15% level, but it showed that a pony would have a reduction in price of 52.9%, ceteris paribus. *Mare* was once again highly statistically significant and negative, indicating a reduction in price of 28.0% as compared to a gelding. Not surprisingly, based on the results of the linear model, the horse being papered had the largest impact on price in this data set and was highly statistically significant, indicating that a horse would be worth 71.8% more by being registered as opposed to grade. Finally, both of the age terms are highly statistically significant and have the expected signs; using these coefficients we can determine the age at which a horse is most valuable according to this functional form, as follows:

$$\frac{\partial lnPHorse}{\partial Age} = 0.124677 + 2(-0.00589) \text{ x } Age = 0$$

$$0.124677 – 0.01178 \text{ x } Age = 0$$

$$0.124677 = 0.01178 \text{ x } Age$$

$$Age^* = 10.5838 \text{ years}$$

Once again, this maximum age value is reinforced by the popular industry belief that a horse is at its prime between the ages of 8 and 12. This age is slightly lower than that indicated by the linear model but the two estimates vary by less than 10% from each other and therefore do not appear to differ with any statistical significance.

Table 12: Spring Log-Log Model Results

Spring Dreamhorse.com Log-Log Model Results (Dependent Variable = *lnPHorse*)				
N	R-Square	Adj R-Sq	F-Value	F-Significance
470	0.2672	0.2118	4.82	<.0001
Variable	Parameter Estimate	Standard Error	t-Value	P-statistic
Intercept	**6.10498***	**0.4367**	**13.98**	**<.0001**
Pht	**0.31018***	**0.0900**	**3.45**	**0.0006**
MO	-0.05632	0.1683	-0.33	0.7380
DK	0.29320	0.2100	1.40	0.1633
NE	**0.43667**	**0.2376**	**1.84**	**0.0668**
OK	**0.26518**	**0.1470**	**1.80**	**0.0719**
TX	**0.29279**	**0.1413**	**2.07**	**0.0388**
Pnt	-0.35463	0.4280	-0.83	0.4078
QH	-0.08669	0.3566	-0.24	0.8080
Pony	-0.47833	0.3519	-1.36	0.1748
TB	-0.32842	0.4220	-0.78	0.4368
ClrBrd	-0.26606	0.2493	-1.07	0.2866
Arab	-0.05952	0.2904	-0.20	0.8377
Saddle	0.15062	0.4116	0.37	0.7146
Other	-0.23505	0.3239	-0.73	0.4684
Xbred	-0.01171	0.1895	-0.06	0.9507
Mare	**-0.26622***	**0.0760**	**-3.50**	**0.0005**
Stud	0.00915	0.1335	0.07	0.9454
InFoal	-0.14950	0.1453	-1.03	0.3040
Bay	-0.05623	0.2633	-0.21	0.8310
Sor	-0.01966	0.2534	-0.08	0.9382
Blk	0.01648	0.2398	0.07	0.9453
Grey	0.00452	0.2716	0.02	0.9867
Pal	0.18765	0.2694	0.70	0.4864
Paint	**0.21354***	**0.1445**	**1.48**	**0.1402**
Dun	0.11354	0.2822	0.40	0.6876
Dilute	0.02278	0.2632	0.09	0.9311
Roan	**0.42558***	**0.2661**	**1.60**	**0.1104**
ChBr	0.06347	0.2633	0.24	0.8097
Online	-0.01555	0.0703	-0.22	0.8249
lnAge	**0.75333***	**0.2252**	**3.35**	**0.0009**
$(lnAge)^2$	**-0.20040**	**0.0792**	**-2.53**	**0.0118**
Tmp	-0.04842	0.0724	-0.67	0.5041
Paper	**0.69701***	**0.1850**	**3.77**	**0.0002**

Log-Log Spring Model Results and Discussion

Table 12 contains the results of the log-log model for the spring data set. This is the final functional form, where the natural log is taken of the dependent variable and continuous independent variables, age and temperament. The log of age was also squared in this model to allow for an age of maximum value. Like the semi-log model, 10 of the independent variables were statistically significant at the 15% level, in addition to the intercept being statistically significant. However, unlike either of the previous models and as opposed to the F-tests earlier on color, two of the color variables are statistically significant at greater than the 15% level and none of the breed variables are statistically significant. The statistics of overall fit indicate that 21.18% of the variation in price is explained by the data. The F-value is also highly significant.

The intercept is highly significant, providing a base price for the hedonic model when coupled with the averages of the significant continuous variables. Having a photo advertisement is also highly statistically significant, indicating that a horse sold using a photo ad would sell for 31.0% greater than a horse sold with a text ad. Three of the state variables, *NE*, *OK*, and *TX* exhibited statistical significance at greater than the 10% level, with Nebraska again having the greatest increase in price over Kansas at 43.67%, while Texas and Oklahoma also both had greater prices than Kansas by 29.28% and 26.52%, respectively, ceteris paribus. As mentioned before, none of the breed variables were statistically significant, which goes against the earlier F-tests indicating that at least one of this group had a coefficient significantly different from zero. Female horses were once again at a discount to geldings by 26.62%, ceteris paribus.

Interestingly, two of the color variables showed statistical significance at greater than the 15% level, being *Paint* and *Roan*. These results say that if a horse has some sort of pinto coloration, its price will increase by 21.35%, and if it carries the roan gene or phenotypically exhibits roaning, it will go for a price 42.56% greater than other colors of horses, ceteris paribus. *Paper* had the greatest economic impact of any of the binary variables, indicating with a highly statistically significant coefficient that a horse's price value in this data set would increase by 69.70% if it was registered versus being grade. *lnAge* was highly statistically significant, while *lnAge2* was statistically significant at just less than the 1% level. The peak age of a horse's value indicated by this model requires us to take the derivative with respect to the log of age, which results in *lnAge** being 1.8796. We then take the exponential of this value to get an age of 6.5506 years as the peak value. This age is much lower than those indicated by the two other models, and therefore raises some concerns about how this model is handling the age variables.

Brief Comparison of the Three Spring Forms

Table 13, on page 63, shows a direct comparison between the results of the three functional forms for the spring data. The intercept was statistically significant in the two functional forms using *lnPHorse* as the dependent variable. Using a photo advertisement was highly statistically significant in all three cases, indicating that no matter which functional form is best, this variable is positive and significant in this data set. Comparing the linear model value at the average price to the semi-log model result for this variable, the linear model indicates a price increase of 38.62%, which is slightly higher than the 35.54% indicated in the semi-log model.

Some of the state variables showed statistical significance in the logarithmic functional forms, indicating that if there are differences in price across states it is probably in a percentage relationship to the price of the horse rather than a linear one, which is understandable. It is unlikely that a low priced horse in Texas would simply be worthless in the state of Kansas, for example, it would just have a lower price based on other market factors. It is reasonable to think that the value would be reduced by a respective percentage of the horse's value given the market situation in the different states, due to the influence of show circuits, large or small, and the availability or lack thereof of world-class trainers, exhibitions, and recreational riding areas or programs. All statistically significant states showed increases in price over the state of Kansas, which does not appear to bode well for the Kansas horse industry, though Missouri was negative but statistically insignificant in all three models.

The linear and semi-log models each had one breed variable statistically significant at greater than the 15% levels, but they were different breeds and were not highly significant. There may be some issues with the grouping of breeds or multicollinearity issues with the colors that have not been accounted for in this model. *Mare* was highly statistically significant and negative in all three models, indicating that in this data set, female horses were less valuable than male horses, particularly geldings, ceteris paribus. Only the log-log model had statistically significant color variables, and neither color was highly significant. All of the age variables across the functional forms were highly significant, with the exception of *Age2* in the log-log form which was significant at greater than the 2% level, indicating that age is also an important factor in this market no matter what functional form is used. Finally, the binary variable having the greatest impact in all three functional forms on the value of the horse was whether or not the horse was registered, indicating that registration of a horse does matter even to lower end clientele in the horse market.

Table 13: Comparison of Spring Models

Spring Comparison of Coefficients for Linear, Semi-Log, and Log-Log Parameter Estimates and Significance						
	Linear		Semi-Log		Log-Log[a]	
Variable	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Intercept	328.5102	0.4406	**6.4217*****	<.0001	**6.1050*****	<.0001
Pht	**689.0551*****	<.0001	**0.3554*****	<.0001	**0.3102*****	0.0006
MO	-181.9980	0.3616	-0.0282	0.8675	-0.0563	0.7380
DK	309.0224	0.3422	**0.3135***	0.1422	0.2932	0.1633
NE	607.7126	0.2259	**0.4390****	0.0610	**0.4367****	0.0668
OK	83.3529	0.6810	**0.2575****	0.0895	**0.2652****	0.0719
TX	167.0538	0.3405	**0.2925****	0.0457	**0.2928****	0.0388
Pnt	-107.7840	0.7006	-0.4351	0.3109	-0.3546	0.4078
QH	212.1241	0.4095	-0.1695	0.6376	-0.0867	0.8080
Pony	-233.1500	0.4823	**-0.5290***	0.1340	-0.4783	0.1748
TB	-89.7903	0.7852	-0.3921	0.3646	-0.3284	0.4368
ClrBrd	-313.9920	0.2075	-0.3071	0.2045	-0.2661	0.2866
Arab	-39.5452	0.9147	-0.1374	0.6308	-0.0595	0.8377
Saddle	**629.1094****	0.0472	0.0904	0.8271	0.1506	0.7146
Other	30.5419	0.9152	-0.2990	0.3395	-0.2351	0.4684
Xbred	165.5213	0.4346	0.0202	0.9160	-0.0117	0.9507
Mare	**-360.0230*****	0.0031	**-0.2804*****	0.0002	**-0.2662*****	0.0005
Stud	-140.9320	0.4467	-0.1447	0.2415	0.0091	0.9454
InFoal	-304.2520	0.1831	-0.1167	0.4132	-0.1495	0.3040
Bay	13.3859	0.9568	-0.1207	0.6702	-0.0562	0.8310
Sor	-48.2338	0.8306	-0.0533	0.8469	-0.0197	0.9382
Blk	147.2820	0.4385	-0.0655	0.8063	0.0165	0.9453
Grey	19.9014	0.9456	-0.0258	0.9290	0.0045	0.9867
Pal	139.8960	0.6359	0.1043	0.7142	0.1876	0.4864
Paint	132.4278	0.3986	0.1867	0.1977	**0.2135***	0.1402
Dun	159.1722	0.6545	0.0197	0.9473	0.1135	0.6876
Dilute	55.4199	0.8478	0.0112	0.9687	0.0228	0.9311
Roan	392.2400	0.1690	0.2512	0.3706	**0.4256***	0.1104
ChBr	269.5000	0.2535	-0.0528	0.8539	0.0635	0.8097
Online	-11.2055	0.9097	-0.0369	0.5973	-0.0156	0.8249
Age/lnAge	**158.7203*****	0.0001	**0.1247*****	<.0001	**0.7533*****	0.0009
Age2/(lnAge)^2	**-7.0226*****	0.0013	**-0.0059*****	<.0001	**-0.2004****	0.0118
Tmp/lnTmp	-42.2749	0.2487	-0.0319	0.2225	-0.0484	0.5041
Paper	**850.1931*****	<.0001	**0.7182*****	0.0002	**0.6970*****	0.0002
*,**,*** Denotes Significance at the 15%, 10%, and 1% levels, respectively						
[a] Log-log is the only format that uses *lnAge*, *lnAge2*, and *lnTmp*						

CHAPTER 6 - Fall 2008 Data, Results and Interpretation

Summary Statistics

Table 14 on the following pages contains the summary statistics for the Fall 2008 data taken directly off of dreamhorse.com. The variables in this data set are the same as described above, so this summary will briefly run through key points. See Chapter 5 for further discussion of particular variables and expected signs and Chapter 4 for full definitions of each variable. The total number of observations collected in the fall was 407. In this data set, 39.5% of horses were advertised using photo ads. Texas is the state with the greatest horse numbers in the sample at 42%, while Kansas and the combined Dakotas share the lowest numbers at 6.6%, or 27 horses each. The majority of horses, nearly 55%, are Quarter Horses, with Arabians again being the lowest breed category with only 3.9% of the sample or 16 observations. Ten percent of the sample horses were crossbreds. Horses of the female gender were again the most common constituting 48% of the sample, while uncut male horses constitute less than 15%. Twenty-two mares were carrying a foal at the time of sale, which is 5.4% of the observations. Bays and sorrels were once again the most common horse coat colors in the sample, followed closely by the chestnut and brown category, each having 17-18% of the total observations. Dun factor horses were the least common, constituting only 5.4% of the observed sale horses. Close to 36% of this sample's horses were listed as sold online. The age distribution, in Figure 4, shows a right skewed distribution with a high number of horses less than one year of age. The average age

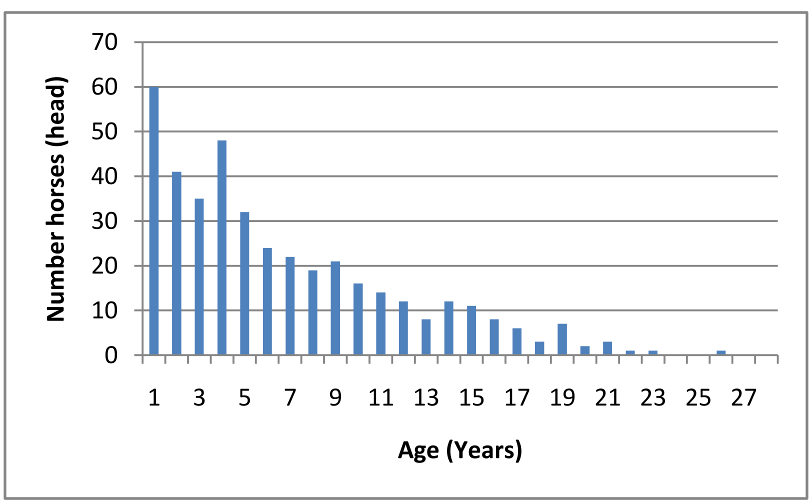

Figure 4: Fall Age Distribution

recorded in this sample was 5.77 years. Temperament scores ranged from 1 to 8 with an average of 3.25 and a low standard deviation of 1.61, indicating a tight grouping around the 2-3 range for temperament score. Slightly more than three quarters of the horses in this sample were listed as registered or 306 horses out of the 407 in the data set.

PHorse and *lnPHorse* are the dependent variables again in this case and are shown at the bottom of Table 14. The average price in this data set was $1,713.64 with a standard deviation of just over $1,200. Prices ranged from $1 to a maximum of $5,000, as this is the upper limit chosen for this data set by the author. The price distribution is shown in Figure 5, showing a slightly skewed but fairly even distribution with a peak at its median and mode of $1,500. These data will be used in the following models to ascertain which traits affected the prices of horses sold during the fall time period.

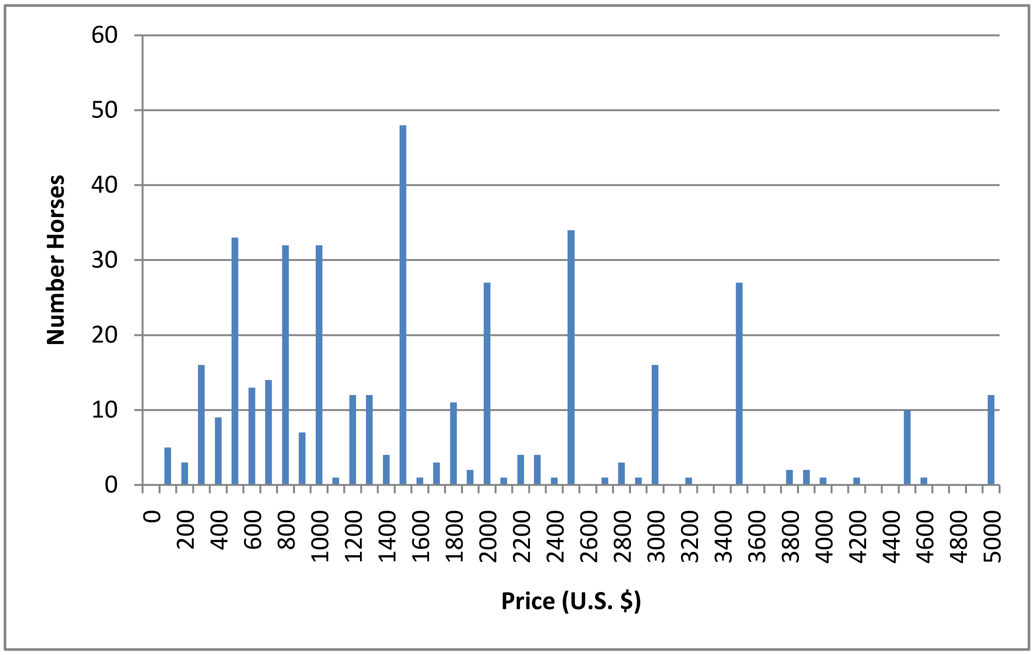

Figure 5: Fall Asking Price Distribution

Table 14: Summary Statistics of Horses Sold in the Fall

Group	Label	Description	Mean	Std. Dev.	Min	Max	Sum	Sign
	Independent Variables:							
	Pht	Photo Ad	0.39558	0.48958	0	1	161	+
	KS	Kansas	0.06634	0.24918	0	1	27	N/A
	MO	Missouri	0.18919	0.39214	0	1	77	?
State	*DK*	The Dakotas	0.06634	0.24918	0	1	27	?
	NE	Nebraska	0.07862	0.26948	0	1	32	?
	OK	Oklahoma	0.17936	0.38413	0	1	73	?
	TX	Texas	0.42015	0.49419	0	1	171	?
	Pnt	Paint	0.15233	0.35979	0	1	62	?
	QH	Quarter Horse	0.54791	0.49831	0	1	223	?
	Pony	Pony Breeds	0.11057	0.31398	0	1	45	?
Breed	*TB*	Thoroughbred	0.07125	0.25756	0	1	29	?
	ClrBrd	Color Breeds	0.04914	0.21643	0	1	20	?
	Arab	Arabian	0.03931	0.19458	0	1	16	?
	Saddle	Saddle Breeds	0.07125	0.25756	0	1	29	?
	Other	Other Breeds and Equids	0.04177	0.20031	0	1	17	?
	Xbred	Crossbreds	0.10319	0.30459	0	1	42	-
	Geld	Geldings	0.37101	0.48367	0	1	151	N/A
Gender	*Mare*	Female Gender	0.47912	0.50018	0	1	195	-
	Stud	Male Gender	0.14988	0.35739	0	1	61	-
	InFoal	Mares in Foal	0.05405	0.22640	0	1	22	-
	Bay	Bays	0.18182	0.38617	0	1	74	?
	Sor	Sorrel	0.17936	0.38413	0	1	73	?
Color	*Blk*	Black	0.07371	0.26162	0	1	30	?
	Grey	Grey	0.08600	0.28070	0	1	35	?
	Pal	Palomino	0.09091	0.28783	0	1	37	?

Color	*Paint*	Pinto	0.12285	0.32867	0	1	50	?
	Dun	Dun Factor	0.05405	0.22640	0	1	22	?
	Dilute	Dilute Gene	0.07862	0.26948	0	1	32	?
	Roan	Roan	0.08108	0.27330	0	1	33	?
	ChBr	Chestnut/Brown	0.17199	0.37784	0	1	70	?
	Online	Sold Online	0.35627	0.47948	0	1	145	?
	Age	Age	5.76658	5.13364	0.50	25	2,347	+
	Age2	Age Squared	59.54300	92.05302	0.25	625	24,234	-
	Tmp	Temperament	3.25483	1.61458	1	8	1,325	-
	lnAge	Natural Log of Age	1.24219	1.13045	-0.69315	3.21888	505.56941	+
	(lnAge) ^{2}	Natural Log of Age Squared	2.81781	2.57673	0	10.36116	1,147	-
	lnTmp	Natural Log of Temperament	1.04751	0.53731	0	2.07944	426.33687	-
	Paper	Registration	0.75184	0.43248	0	1	306	+
	Dependent Variables:							
	PHorse	Asking Price of the Horse	1,713.64	1,214.53	1.00	5,000.00	697,453.00	N/A
	lnPHorse	Natural Log of the Asking Price	7.12228	1.04216	0.00000	8.51719	2,898.76610	N/A

White’s Test Results

The White’s Test results for this model indicated that we can reject heteroskedasticity and assume homoskedasticity in all three functional forms, and therefore can move ahead with the OLS regression using SAS 9.1. We can reject the null at the 42.09% level for the linear model and at the 12.53% and 13.12% levels for the semi-log and log-log models, respectively, as indicated by the results in Table 15.

Table 15: Fall White's Tests Results

Linear Model Heteroscedasticity Test, Fall Data					
Equation	Test	Statistic	DF	Pr > ChiSq	Variables
Phorse	White's Test	372.8	368	0.4209	Cross of all vars
Semi-Log Model Heteroscedasticity Test, Fall Data					
Equation	Test	Statistic	DF	Pr > ChiSq	Variables
lnPHorse	White's Test	399.4	368	0.1253	Cross of all vars
Log-Log Model Heteroscedasticity Test, Fall Data					
Equation	Test	Statistic	DF	Pr > ChiSq	Variables
lnPHorse	White's Test	398.6	368	0.1312	Cross of all vars

Group F-Test Results

The F-Test results for the groups State, Breed, Color, and Gender for this data set are shown in Table 16. Note that the F-tests for the semi-log and log-log functional forms here are the same due to the binary nature of the independent variables in question. Each of the three functional forms indicate that at least one of the coefficients on the State variables is not equal to zero at better than a 5% significance level. All three functional forms returned highly significant F-statistics for breeds, indicating that at least one of these coefficients is not zero when regressed against the price of the horse. The same is true for the Gender group, as all three F-tests indicate significance at better than 1%. However, the Color group F-test indicates that none of the coefficients are significantly different from zero, indicating this group of coefficients is likely to be statistically insignificant. These results will be useful when evaluating the results received from the full models.

Table 16: Fall F-Test Results

Fall Group F-Tests		
States	F-Value	F-Signficance
Linear	2.32**	0.043
Semi-Log	2.40**	0.0365
Log-Log	2.40**	0.0365
Breeds	F-Value	F-Signficance
Linear	3.77***	0.0003
Semi-Log	2.71***	0.0064
Log-Log	2.71***	0.0064
Colors	F-Value	F-Signficance
Linear	0.58	0.8287
Semi-Log	0.63	0.7839
Log-Log	0.63	0.7839
Genders	F-Value	F-Signficance
Linear	14.72***	<.0001
Semi-Log	7.78***	0.0005
Log-Log	7.78***	0.0005

Overall Model Results

The OLS model results using linear, semi-log, and log-log functional forms are reported in the following subsections. The overall model fit and statistical significance will be discussed for each functional form. Statistically significant variables and some interesting insignificant variables will also be discussed in these sections; some attempt at explanations for signs and significance will be made, as well.

Linear Fall Model Results and Discussion

The statistics of overall fit and significance are shown in Table 17 along with the results of individual variables. The R-squared value is 0.2979, and the adjusted R-squared is 0.2358, indicating that 23.58% of the linear variability in horse price is explained by the data. The F-value of 4.8 is also highly significant; indicating that at least one of the coefficients in this model is statistically significant, which is not surprising since the earlier group F-tests indicated the same thing for the State, Breed, and Gender variable groups. Eight of the independent variables and the intercept are statistically significant in the overall model at greater than the 15% level. As expected, none of the color categories showed statistical significance, but unexpectedly none of the breed variables show statistical significance. Once again, there may be some issues in the model with color and breed variables being correlated or the linear model may not be the best fit to capture the statistical significance in the breed differences.

Photo ads were highly statistically significant as opposed to text ads, indicating that horses advertised online using photo ads had sale prices $524.21 higher than horses in text ads, ceteris paribus. This magnitude once again is much greater than the range of costs of the photo ads, indicating possibly a much greater return for your money if you advertise using a $25-$100 photo ad. Texas was the only statistically significant state in this data set, and only at the 15% level. As compared to Kansas, a horse sold from the state of Texas brought $349.83 more, ceteris paribus. Both of the gender variables were statistically significant and negative in this data set as compared to geldings. Female horses brought $462.45 less and uncut male horses $507.79 less than geldings, ceteris paribus, both at greater than the 1% significance level. All three of the continuous independent variables in this data set are highly statistically significant and have the expected signs. Temperament scores had a negative relationship with the price of horses, indicating that an increase of one on the temperament scale causes the horse's price to decrease by $92.95. The age at which a horse is at its maximum value can be achieved with the following calculations:

$$\frac{\partial PHorse}{\partial Age} = 179.34777 + 2(-8.67411) \times Age = 0$$

$$179.34777 - 17.34822 \times Age = 0$$

$$179.34777 = 17.34822 \times Age$$

$$Age^{*} = 10.3381 \text{ years.}$$

This model indicates that a horse is at its maximum value at just over 10 years of age, ceteris paribus, which fits into the typically accepted age range of the horse world where a horse is well broke and experienced but not yet having health issues associated with old age, and still having a good amount remaining it its usable life. Finally, whether or not the horse is registered is statistically significant and has the expected positive sign, and also is the variable having the greatest impact on the value of the horse. The coefficient indicates that a horse is worth $602.25 more if it is registered versus grade, ceteris paribus.

Table 17: Fall Linear Model Results

Fall Dreamhorse.com Linear Model Results (Dependent Variable = *PHorse*)				
N	R-Square	Adj R-Sq	F-Value	F-Significance
407	0.2979	0.2358	4.80	<.0001
Variable	Parameter Estimate	Standard Error	t-Value	P-statistic
Intercept	**1491.68408****	**819.8154**	**1.82**	**0.0696**
Pht	**524.20733*****	**119.0200**	**4.40**	**<.0001**
MO	-17.66168	252.6489	-0.07	0.9443
DK	282.18254	303.2745	0.93	0.3527
NE	285.36263	292.0403	0.98	0.3291
OK	99.80376	259.1630	0.39	0.7004
TX	**349.83126***	**234.4196**	**1.49**	**0.1365**
Pnt	35.17467	438.4757	0.08	0.9361
QH	133.23305	404.2610	0.33	0.7419
Pony	-230.12426	471.5918	-0.49	0.6259
TB	-319.29869	418.2124	-0.76	0.4457
ClrBrd	-289.22078	449.1140	-0.64	0.5200
Arab	-44.33020	484.4735	-0.09	0.9271
Saddle	530.90726	460.9964	1.15	0.2502
Other	-113.96319	482.2828	-0.24	0.8133
Xbred	329.23941	376.4700	0.87	0.3824
Mare	**-462.45225*****	**127.0110**	**-3.64**	**0.0003**
Stud	**-507.78610*****	**195.7684**	**-2.59**	**0.0099**
InFoal	-358.70052	259.8721	-1.38	0.1683
Bay	-585.49932	670.7917	-0.87	0.3833
Sor	-731.19333	670.3932	-1.09	0.2761
Blk	-737.14641	683.9163	-1.08	0.2818
Grey	-490.69134	693.4428	-0.71	0.4796
Pal	-659.07467	684.2025	-0.96	0.3360
Paint	79.76645	212.5012	0.38	0.7076
Dun	-720.24363	677.3462	-1.06	0.2883
Dilute	-313.61185	669.9611	-0.47	0.6400
Roan	-639.28233	681.2234	-0.94	0.3486
ChBr	-693.38408	673.4887	-1.03	0.3039
Online	-3.79516	119.2143	-0.03	0.9746
Age	**179.34777*****	**37.5242**	**4.78**	**<.0001**
Age2	**-8.67411*****	**1.9786**	**-4.38**	**<.0001**
Tmp	**-92.95406*****	**34.8293**	**-2.67**	**0.0079**
Paper	**602.25124*****	**154.9294**	**3.89**	**0.0001**

Semi-Log Fall Model Results and Discussion

The semi-log functional form, shown in Table 18, on the fall data set has a highly significant F-value of 3.83. In addition, the adjusted R-squared value of 0.1868 indicates that 18.68% of the variation in the price of horses is explained by the model. The intercept is statistically significant in addition to 12 of the independent variables being statistically significant at greater than the 15% level, including four of the states. Only one breed variable was statistically significant at the 15% level, and one gender variable at greater than the 10% level. As expected, none of the color variables were statistically significant.

The highly statistically significant coefficient for photo ads indicated that a horse advertised through a photo ad brought a price 38.14% higher than a text ad, ceteris paribus. The state of Texas had a highly statistically significant coefficient, indicating that horses sold in Texas as opposed to Kansas sold for a 53.79% higher price, ceteris paribus. The Dakotas and Nebraska also had coefficients statistically significant at better than the 10% level indicating that horses sold from those states were priced 56.70% and 43.40% higher than horses from Kansas, ceteris paribus. Oklahoma was statistically significant at the 15% level and its coefficient indicated that horses sold from there as opposed to Kansas would bring a price 33.45% higher, ceteris paribus. The breed group for saddle horses is the only one showing statistical significance at greater than the 15% level and indicated that a horse in this category would bring a price 59.39% higher than other horses, all else constant. In the gender variables, female horses had no statistical difference from geldings in this model, but the coefficient for uncut male horses is significant at the 10% level and indicates that these horses would bring a price 39.04% lower than geldings, ceteris paribus. All three of the continuous variables in this functional form were highly significant and had the expected signs. For each additional increase in temperament score level, the price of a horse would be reduced by 12.31%, ceteris paribus.

The age at which a horse achieves maximum value is indicated by the age and age squared variables, as calculated here:

$$\frac{\partial lnPHorse}{\partial Age} = 0.11898 + 2(-0.00623) \text{ x } Age = 0$$

$$0.11898 - 0.01246 \text{ x } Age = 0$$

$$0.11898 = 0.01246 \text{ x } Age$$

$$Age^{*} = 9.5490 \text{ years.}$$

Table 18: Fall Semi-Log Model Results

Fall Dreamhorse.com Semi-Log Model Results (Dependent Variable = *lnPHorse*)				
N	R-Square	Adj R-Sq	F-Value	F-Significance
407	0.2529	0.1868	3.83	<.0001
Variable	Parameter Estimate	Standard Error	t-Value	P-statistic
Intercept	**6.63883*** **	**0.7257**	**9.15**	**<.0001**
Pht	**0.38138*** **	**0.1054**	**3.62**	**0.0003**
MO	0.20967	0.2236	0.94	0.3491
DK	**0.56698** **	**0.2684**	**2.11**	**0.0353**
NE	**0.43403** **	**0.2585**	**1.68**	**0.0940**
OK	**0.33447* **	**0.2294**	**1.46**	**0.1457**
TX	**0.53787*** **	**0.2075**	**2.59**	**0.0099**
Pnt	0.18910	0.3881	0.49	0.6264
QH	0.26463	0.3578	0.74	0.4600
Pony	0.14877	0.4174	0.36	0.7217
TB	-0.19597	0.3702	-0.53	0.5968
ClrBrd	0.08260	0.3975	0.21	0.8355
Arab	0.20401	0.4288	0.48	0.6345
Saddle	**0.59390* **	**0.4081**	**1.46**	**0.1464**
Other	0.05333	0.4269	0.12	0.9006
Xbred	0.12534	0.3332	0.38	0.7070
Mare	-0.15403	0.1124	-1.37	0.1715
Stud	**-0.39043** **	**0.1733**	**-2.25**	**0.0248**
InFoal	-0.21444	0.2300	-0.93	0.3518
Bay	-0.48000	0.5937	-0.81	0.4194
Sor	-0.42638	0.5934	-0.72	0.4729
Blk	-0.48073	0.6054	-0.79	0.4276
Grey	-0.29946	0.6138	-0.49	0.6259
Pal	-0.46685	0.6056	-0.77	0.4413
Paint	-0.11475	0.1881	-0.61	0.5422
Dun	-0.42338	0.5996	-0.71	0.4805
Dilute	-0.60964	0.5930	-1.03	0.3046
Roan	-0.38598	0.6030	-0.64	0.5225
ChBr	-0.51361	0.5961	-0.86	0.3895
Online	-0.15011	0.1055	-1.42	0.1557
Age	**0.11898*** **	**0.0332**	**3.58**	**0.0004**
Age2	**-0.00623*** **	**0.0018**	**-3.56**	**0.0004**
Tmp	**-0.12305*** **	**0.0308**	**-3.99**	**<.0001**
Paper	**0.59057*** **	**0.1371**	**4.31**	**<.0001**

This age falls in the age of 8-12 years, where a horse is at its prime. Finally, the registration of a horse had a highly significant impact on the price of a horse, indicating that a horse that is registered will bring a price 59.06% higher than a grade horse, ceteris paribus.

Log-Log Fall Model Results and Discussion

The results of this regression are shown in Table 19. The R-squared value is 0.2395, and the adjusted R-square is 0.1722, indicating that 17.22% of the variability in price is explained by the independent variables. The F-value was also significant at greater than the 1% level. Nine of the independent variables were statistically significant at the 15% level. None of the breed or color variables were significant. Only two of the state variables were statistically significant at the 10% level and one of the gender variables at the 15% level.

The coefficients on *Pht* and *Paper* are both highly statistically significant. A horse being advertised in a photo ad brings a price 39.91% higher than a horse advertised in a text ad, and a registered horse brings 59.41% more than a grade horse, ceteris paribus. Two of the state variables, the Dakotas and Texas, had coefficients significant at the 10% level. Both indicated that they brought higher prices than Kansas, 55.30% and 46.40% respectively, ceteris paribus. *Stud* was the only statistically significant gender variable, meaning that there was no difference between female horses and geldings, but uncut male horses brought prices 31.11% lower than geldings, ceteris paribus. Horses being sold on dreamhorse.com versus other methods brought 15.60% less, ceteris paribus, at the 15% significance level. Of the three continuous independent variables, *lnAge* and *lnTmp* are both statistically significant at the 1% level and *lnAge2* is significant at the 10% level. The temperament score indicates that for every 1% increase in temperament score, price decreases by 33.18%. For example, moving from a score of 2 to 3 on the temperament scale would be a change of 12.5%, which would indicate a price decrease of over 400%. This magnitude seems unreasonably high as compared to the linear and semi-logarithmic models. To get the age of maximum value we can take the derivative with respect to *lnAge*, set it equal to zero and solve for *lnAge*, then take the exponential of that value. The derivative *lnAge** = 1.7534. After taking the exponential of the derivative, the result is an age of 5.7739 years, which is much lower than the typical industry standard of 8-12 years being the prime of a horse's life. This tends towards having a longer useful life remaining and less about the experience and training a horse may have adding to its value.

Table 19: Fall Log-Log Model Results

Fall Dreamhorse.com Log-Log Model Results (Dependent Variable = *lnPHorse*)				
N	R-Square	Adj R-Sq	F-Value	F-Significance
407	0.2395	0.1722	3.56	<.0001
Variable	Parameter Estimate	Standard Error	t-Value	P-statistic
Intercept	**6.83340*** **	**0.7252**	**9.42**	**<.0001**
Pht	**0.39908*** **	**0.1063**	**3.75**	**0.0002**
MO	0.15980	0.2250	0.71	0.4781
DK	**0.55295** **	**0.2706**	**2.04**	**0.0417**
NE	0.29789	0.2613	1.14	0.2550
OK	0.23463	0.2320	1.01	0.3124
TX	**0.46403** **	**0.2088**	**2.22**	**0.0268**
Pnt	0.14815	0.3926	0.38	0.7061
QH	0.20887	0.3617	0.58	0.5639
Pony	0.12349	0.4233	0.29	0.7706
TB	-0.23792	0.3738	-0.64	0.5248
ClrBrd	0.01073	0.4027	0.03	0.9788
Arab	0.11217	0.4326	0.26	0.7956
Saddle	0.57073	0.4122	1.38	0.1670
Other	-0.01381	0.4321	-0.03	0.9745
Xbred	0.15962	0.3377	0.47	0.6367
Mare	-0.15317	0.1152	-1.33	0.1843
Stud	**-0.31105***	**0.1941**	**-1.60**	**0.1100**
InFoal	-0.16080	0.2314	-0.69	0.4875
Bay	-0.42902	0.6011	-0.71	0.4759
Sor	-0.40516	0.6004	-0.67	0.5002
Blk	-0.40729	0.6128	-0.66	0.5067
Grey	-0.26958	0.6208	-0.43	0.6644
Pal	-0.39480	0.6132	-0.64	0.5201
Paint	-0.13065	0.1897	-0.69	0.4915
Dun	-0.34098	0.6071	-0.56	0.5747
Dilute	-0.64025	0.5995	-1.07	0.2862
Roan	-0.33629	0.6108	-0.55	0.5823
ChBr	-0.53618	0.6024	-0.89	0.3740
Online	**-0.15604***	**0.1073**	**-1.45**	**0.1465**
lnAge	**0.31392*** **	**0.1162**	**2.70**	**0.0072**
$(lnAge)^2$	**-0.08952** **	**0.0464**	**-1.93**	**0.0543**
lnTmp	**-0.33184*** **	**0.0940**	**-3.53**	**0.0005**
Paper	**0.59414*** **	**0.1384**	**4.29**	**<.0001**

Brief Comparison of the Three Functional Forms

Finally, Table 20 provides for direct comparison between the coefficients and statistical significance of each independent variable across the three functional forms in this data set. The photo ad coefficient is highly statistically significant and positive in all three functional forms, indicating a significant economic impact on price no matter what the form is. Registration had the same impact across all three forms, being positive and highly statistically significant in all cases. Of the state variables, Texas was positive and statistically significant at greater than the 15% level across all three of the models, while the Dakotas were positive and significant at the 10% level in the two logarithmic models. Additionally, the semi-log model showed statistical significance in Nebraska and Oklahoma to the 10% and 15% levels, respectively. Based on the F-tests showing that at least one of the state's coefficients should be significantly different from zero, all three of these models seemed to capture this element but the semi-log model seemed to capture the most statistical significance of the group of state variables.

Only the semi-log model captured any statistical significance in the breed variables, and that was a positive coefficient on the saddle horse breed category with significance at the 15% level. The F-test results indicated that at least one of these coefficients should have been significantly different from zero, so the other two models did not seem to model the data as well as the semi-log in this case. The gender variable *Stud* had a statistically significant coefficient in all three models to at least the 15% level. In addition, the linear model showed statistical significance in the *Mare* variable's coefficient at greater than the 1% level. None of the models showed any statistically significant coefficients in the color variables, and the log-log model was the only one to have any statistical significance in the *Online* variable. The temperament score of a horse was highly statistically significant and had the expected negative sign in all three models. Finally, the age terms were highly statistically significant in the linear and semi-log models and statistically significant to at least the 10% level in the log-log model. The ages of maximum value were 10.3381, 9.5490, and 5.7739 years respectively. The linear and semi-log models ages were within a range accepted as typical for the industry, while the log-log model showed a much lower age than expected. The semi-log model overall appears to be the best at modeling this data set, most closely matching up with industry expectations and the indications of significance given by the F-tests.

Table 20: Comparison of Fall Models

Fall Comparison of Coefficients for Linear, Semi-Log, and Log-Log Parameter Estimates and Significance						
	Linear		Semi-Log		Log-Log[a]	
Variable	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Intercept	**1491.6841****	0.0696	**6.6388*****	<.0001	**6.8334*****	<.0001
Pht	**524.2073*****	<.0001	**0.3814*****	0.0003	**0.3991*****	0.0002
MO	-17.6617	0.9443	0.2097	0.3491	0.1598	0.4781
DK	282.1825	0.3527	**0.5670****	0.0353	**0.5530****	0.0417
NE	285.3626	0.3291	**0.4340****	0.0940	0.2979	0.2550
OK	99.8038	0.7004	**0.3345***	0.1457	0.2346	0.3124
TX	**349.8313***	0.1365	**0.5379*****	0.0099	**0.4640****	0.0268
Pnt	35.1747	0.9361	0.1891	0.6264	0.1482	0.7061
QH	133.2331	0.7419	0.2646	0.4600	0.2089	0.5639
Pony	-230.1243	0.6259	0.1488	0.7217	0.1235	0.7706
TB	-319.2987	0.4457	-0.1960	0.5968	-0.2379	0.5248
ClrBrd	-289.2208	0.5200	0.0826	0.8355	0.0107	0.9788
Arab	-44.3302	0.9271	0.2040	0.6345	0.1122	0.7956
Saddle	530.9073	0.2502	**0.5939***	0.1464	0.5707	0.1670
Other	-113.9632	0.8133	0.0533	0.9006	-0.0138	0.9745
Xbred	329.2394	0.3824	0.1253	0.7070	0.1596	0.6367
Mare	**-462.4523*****	0.0003	-0.1540	0.1715	-0.1532	0.1843
Stud	**-507.7861*****	0.0099	**-0.3904****	0.0248	**-0.3111***	0.1100
InFoal	-358.7005	0.1683	-0.2144	0.3518	-0.1608	0.4875
Bay	-585.4993	0.3833	-0.4800	0.4194	-0.4290	0.4759
Sor	-731.1933	0.2761	-0.4264	0.4729	-0.4052	0.5002
Blk	-737.1464	0.2818	-0.4807	0.4276	-0.4073	0.5067
Grey	-490.6913	0.4796	-0.2995	0.6259	-0.2696	0.6644
Pal	-659.0747	0.3360	-0.4669	0.4413	-0.3948	0.5201
Paint	79.7665	0.7076	-0.1148	0.5422	-0.1307	0.4915
Dun	-720.2436	0.2883	-0.4234	0.4805	-0.3410	0.5747
Dilute	-313.6119	0.6400	-0.6096	0.3046	-0.6403	0.2862
Roan	-639.2823	0.3486	-0.3860	0.5225	-0.3363	0.5823
ChBr	-693.3841	0.3039	-0.5136	0.3895	-0.5362	0.3740
Online	-3.7952	0.9746	-0.1501	0.1557	**-0.1560***	0.1465
Age/lnAge	**179.3478*****	<.0001	**0.1190*****	0.0004	**0.3139*****	0.0072
Age2/(lnAge)^2	**-8.6741*****	<.0001	**-0.0062*****	0.0004	**-0.0895****	0.0543
Tmp/lnTmp	**-92.9541*****	0.0079	**-0.1231*****	<.0001	**-0.3318*****	0.0005
Paper	**602.2512*****	0.0001	**0.5906*****	<.0001	**0.5941*****	<.0001

*,**,*** Denotes Significance at the 15%, 10%, and 1% levels, respectively

[a] Log-log is the only format that uses *lnAge*, *lnAge2*, and *lnTmp*

CHAPTER 7 - Fall 2008 Survey Data, Results and Interpretation

Survey Method and Response

During the collection of data for the Fall 2008 survey from dreamhorse.com, an additional step was taken to consider the problem of the prices listed on the website being asking prices, not final market transaction prices. Using the e-mail system set up on the dreamhorse.com ads, a one question survey was sent to each seller asking what the final sale price of their horse was. This survey is shown in Appendix A, and is short due to the character limit allowed on the dreamhorse.com website. In future surveys of horse owners, it may be better to send a link to an online survey, but for only one question the author felt this was the simplest and most effective method.

The survey was sent out 407 total times. Some of the horses on the sale list had the same contact information and/or the same sellers, but the survey was sent for each individual ad, not per seller. The response rate on this survey was excellent, prompting the use of the data in a model of its own, with 213 total responses for a response rate of 52.33%, and 162 of these reporting usable price data. There is a high correlation, 93.31%, between the asking price of the horses on dreamhorse.com and the reported actual price from the usable survey responses. Some survey responses included information on how the sellers were feeling about the current horse industry. Several sellers mentioned the industry has been extremely low in recent months and selling their horses was difficult. One individual had given away three horses in the last several months, stating no one was buying horses. Nine different responses expressed that the horse market is extremely tough right now as compared to previous years. Some individuals mentioned that due to high feed and transportation costs, they reduced the price or intentionally set a low price to get the horses out of their possession and decrease costs. Several sellers stated that if they reduced the price on horses, they generally reduced it by 5-10% as a rule. Eight individuals asked to see the results of this thesis when it was finished, as they were very interested in learning all they could about the horse market. Numerous sellers expressed that they traded or partially traded horses for other horses, hay, saddles, cattle, or various other items, and one seller admitted that they had not actually sold their horse but instead had kept it and changed the ad to a sold status to prevent e-mails and phone calls on the animal.

Of the 162 usable responses, 87 horses were sold for their asking prices, or 53.70% of the horses; so the majority of horses did sell for their asking price. In three cases, the different price was actually higher than that listed on the website. Therefore, 72 horses were sold for less than their asking prices, or 44.44% of the horses. On average, the horses sold for $156.01 less than their asking prices, which is an average of a 10.00% price reduction, while the 72 horses that were sold at a discount had an average price reduction of 24.02%. The distribution of these price reductions is shown in Figure 6, where we can see that the price reductions are rightward skewed and therefore most horse's prices were not drastically reduced from the asking price. This fact

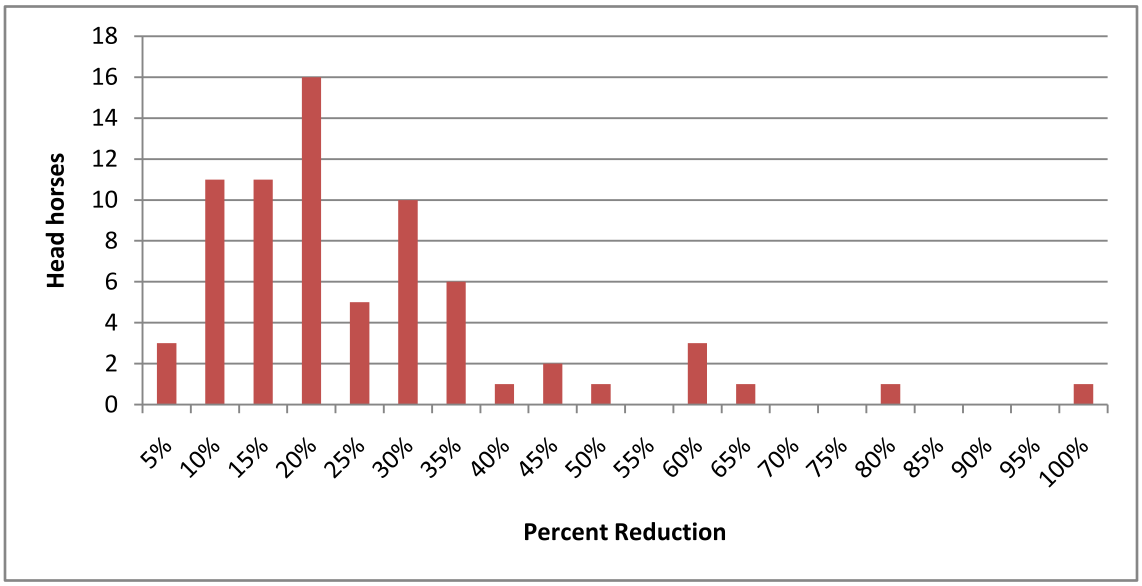

Figure 6: Distribution of the Price Reduction between Asking and Actual Prices

coupled with the high correlation between asking and actual price indicates that the use of asking price in the spring and fall data sets in Chapters 5 and 6 is appropriate for determining traits that affect the value of horses sold online.

In addition to this direct comparison, a linear OLS regression was performed on the independent variables in the data set with the dependent variable being the price difference in dollars between the asking price and the sale price. This model had an insignificant F-value of 0.99 as well as an adjusted R-squared value of -0.0026, which differs drastically from the R-squared value of 0.2029 and indicates near collinearity issues. Five variables expressed significance in this regression, being *Pht*, *Online*, *Age*, *Age2*, and *Tmp*. *Pht* and *Tmp* are positive, indicating that if these results could be trusted the price decrease would be greater if these variables were increased or were 1's in the case of the binary variables. *Online* is negative indicating prices would not decrease as much if this variable took on the value of 1. *Age* is positive and a large magnitude of over 40 while *Age2* is negative and small, indicating an increase in price deviations with increasing age at a slightly decreasing rate. However, as previously mentioned the overall statistics of fit indicate that the results of this model do not appear to be statistically significant; we will go ahead with the assumption that none of the individual factors about an equine affect whether or not its price is reduced from the asking price.

Summary Statistics

Table 21 contains the Fall Survey summary statistics on the following pages. The variables used in the model are the same as those used in both the spring and fall models using asking prices reported in Chapters 5 and 6. Therefore, for variable descriptions and discussion of the expected signs, see Chapter 5 Summary Statistics. There were 72 horses in this data set sold using photo ads, or 42% of the survey responses. Not surprisingly, the state having the most sold horses is Texas, at 38%, while the Dakotas, Kansas, and Nebraska all had low numbers at 13, 14, and 15 horses, respectively. Over half, 54%, of the horses are Quarter Horse, while only 6 horses were Arabians, and 9 each were in the Color Breed and Saddle Breed categories. Sixteen horses, just less than 10% of the sample, are crossbred. The majority of horses in the survey sample are female, 48%, while only 18% were uncut males and the rest are geldings. Only 9 mares, approximately 5.5% of the sample, were carrying foals at the time of their sale. The sorrel color category had the greatest number of horses, at 30 horses or 18.5%, with the combined category of chestnut and browns a close second at 17.9%. Only 9 horses, 5.5%, were dun factor horses. Approximately 12% of the horses were pinto colored, meaning that these horses likely were included in this color category and another base color. Sixty-four horses were listed as sold on dreamhorse.com, which constitutes 39.5% of the sample. The ages reported in this data set were 0.5 years to 25 years of age with an average age of 5.75 years, and as shown in Figure 7 were skewed right with a high number of horses under a year of age reported. The temperament scores ranged from 1 to 8, as with the other two data sets, and averaged 2.98 with a standard deviation of 1.51, indicating a tight distribution of temperaments around the mean. Finally, 116 of the horses in this data set were listed as registered or 71.6% of the sample.

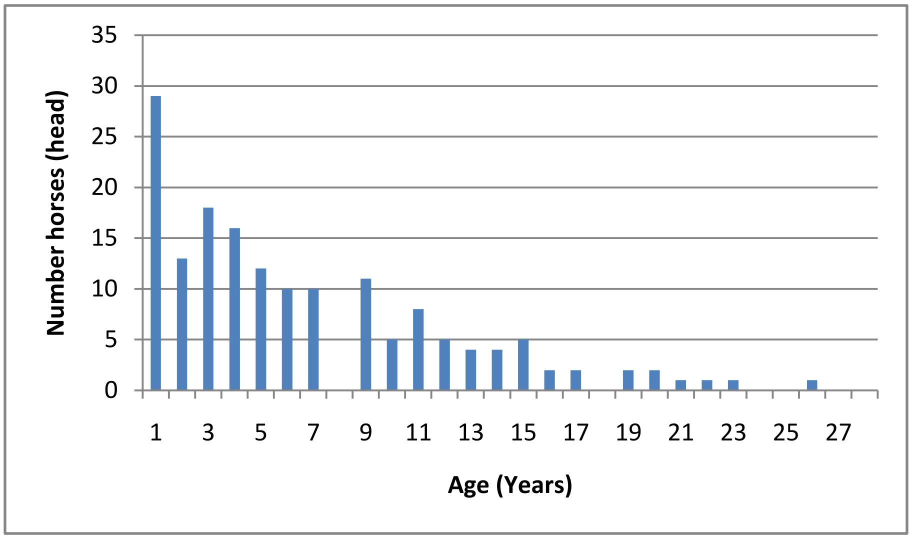

Figure 7: Fall Survey Age Distribution

The price variables here are *APHorse* and *lnAPHorse*, for actual price of the horse as opposed to asking price. The median price in this data set is $1,250, while the mode is $1,500 and the average is $1,526.41. The distribution of prices is show in Figure 8, which seems to be skewed right though not as drastically as the spring and fall data sets discussed above. These prices will be used as the dependent variable in the following models. Though some of the variable categories have very few observations and therefore the results of these models will be open to question, the purpose of running these models with exactly the same variable categories is to provide a comparison to show that the results of the spring and fall data sets are viable regardless of the use of asking prices instead of actual prices.

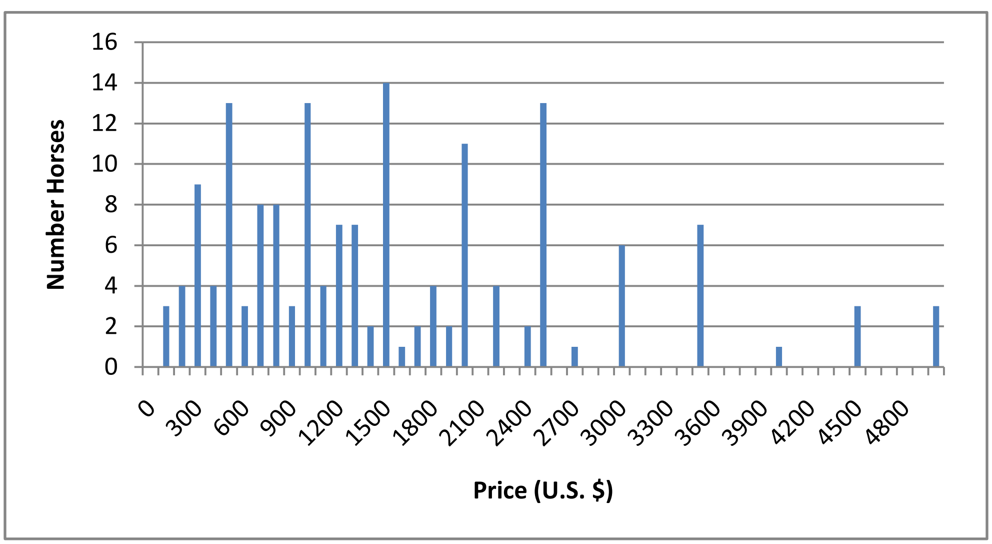

Figure 8: Fall Survey Actual Price Distribution

Table 21: Summary Statistics of Horses in the Fall Survey

	Label	Description	Mean	Std. Dev.	Min	Max	Sum	Sign
	Independent Variables:							
	Pht	Photo Ad	0.44444	0.49844	0	1	72	+
State	*KS*	Kansas	0.08642	0.28185	0	1	14	N/A
	MO	Missouri	0.22840	0.42110	0	1	37	?
	DK	The Dakotas	0.08025	0.27252	0	1	13	?
	NE	Nebraska	0.09259	0.29076	0	1	15	?
	OK	Oklahoma	0.12346	0.32998	0	1	20	?
	TX	Texas	0.38889	0.48901	0	1	63	?
Breed	*Pnt*	Paint	0.12963	0.33694	0	1	21	?
	QH	Quarter Horse	0.53704	0.50017	0	1	87	?
	Pony	Pony Breeds	0.14198	0.35011	0	1	23	?
	TB	Thoroughbred	0.06790	0.25236	0	1	11	?
	ClrBrd	Color Breeds	0.05556	0.22977	0	1	9	?
	Arab	Arabian	0.03704	0.18944	0	1	6	?
	Saddle	Saddle Breeds	0.05556	0.22977	0	1	9	?
	Other	Other Breeds and Equids	0.05556	0.22977	0	1	16	?
	Xbred	Crossbreds	0.09877	0.29927	0	1	16	-
Gender	*Geld*	Geldings	0.33951	0.47501	0	1	55	N/A
	Mare	Female Gender	0.48148	0.50121	0	1	78	-
	Stud	Male Gender	0.17901	0.38455	0	1	29	-
	InFoal	Mare in Foal	0.05556	0.22977	0	1	9	-
Color	*Bay*	Bays	0.14198	0.35011	0	1	23	?
	Sor	Sorrel	0.18519	0.38965	0	1	30	?
	Blk	Black	0.08642	0.28185	0	1	14	?
	Grey	Grey	0.06790	0.25236	0	1	11	?
	Pal	Palomino	0.07407	0.26270	0	1	12	?

Color	*Paint*	Pinto	0.12346	0.32998	0	1	20	?
	Dun	Dun Factor	0.05556	0.22977	0	1	9	?
	Dilute	Dilute Gene	0.11111	0.31524	0	1	18	?
	Roan	Roan	0.09259	0.29076	0	1	15	?
	ChBr	Chestnut/Brown	0.17901	0.38455	0	1	29	?
	Online	Sold Online	0.39506	0.49038	0	1	64	?
	Age	Age	5.75000	5.40883	0.50	25	932	+
	Age2	Age Squared	62.13735	102.78733	0.25	625	10,066	-
	Tmp	Temperament	2.98413	1.51043	1	8	483	-
	lnAge	Natural Log of Age	1.19288	1.17795	-0.69315	3.21888	193.24660	+
	(lnAge)^2	Natural Log of Age Squared	2.80195	2.65878	0	10.36116	454	-
	lnTmp	Natural Log of Temperament	0.96390	0.52696	0	2.07944	156.15158	-
	Paper	Registration	0.71605	0.45231	0	1	116	+
	Dependent Variables:							
	APHorse	Actual Price of the Horse	1,526.41	1,105.00	1.00	5,000.00	247,278.00	N/A
	lnAPHorse	Natural Log of the Actual Price	6.94957	1.23312	0.00000	8.51719	1,125.82966	N/A

White's Test Results

Due to the cross sectional nature of this sample, a check for heteroskedasticity was performed. A White's Test was performed on all three functional forms to be run on this data set, and oddly all three forms came out with exactly the same result. The result is that we can reject the problem of heteroskedasticity at the 46.3% level and use OLS estimation rather than the GMM or another correction for heteroskedasticity. The results of each White's Test are presented in Table 22.

Table 22: Fall Survey White's Test Results

Linear Model Heteroscedasticity Test, Fall Survey Data					
Equation	Test	Statistic	DF	Pr > ChiSq	Variables
APHorse	White's Test	162	161	0.4630	Cross of all vars
Semi-Log Model Heteroscedasticity Test, Fall Survey Data					
Equation	Test	Statistic	DF	Pr > ChiSq	Variables
lnAPHorse	White's Test	162	161	0.4630	Cross of all vars
Log-Log Model Heteroscedasticity Test, Fall Survey Data					
Equation	Test	Statistic	DF	Pr > ChiSq	Variables
lnAPHorse	White's Test	162	161	0.4630	Cross of all vars

Group F-Test Results

The F-tests were performed on four groups of binary variables: States, Breeds, Colors, and Genders. These results are shown in Table 23. None of the F-tests indicated statistical significance in the State or Color categories, meaning that none of the coefficients for these variables in this data set should be significantly different from zero. The linear model F-tests suggested that at greater than the 1% level the Breed variables should have at least one statistically significant coefficient, and at the 5% level the Gender category should have at least one significant coefficient. The logarithmic models did not indicate statistical significance in the F-tests for the Gender category, but in the Breed category the test suggested significance at greater than the 5% level. Therefore, in the following model results for the linear model we hope to see some statistically significant variables in the Gender and Breed groups, while we may only see significant variables in the Breed group of the logarithmic models.

Table 23: Fall Survey F-Test Results

Fall Survey Group F-Tests		
States	F-Value	F-Signficance
Linear	0.95	0.4481
Semi-Log	1.15	0.3356
Log-Log	1.15	0.3356
Breeds	F-Value	F-Signficance
Linear	2.99***	0.0039
Semi-Log	2.55**	0.0125
Log-Log	2.55**	0.0125
Colors	F-Value	F-Signficance
Linear	0.69	0.7309
Semi-Log	0.88	0.5572
Log-Log	0.88	0.5572
Genders	F-Value	F-Signficance
Linear	3.09**	0.0480
Semi-Log	0.31	0.7375
Log-Log	0.31	0.7375

Overall Model Results

Due to the lack of observations in numerous categories of this sample, it is highly likely that the results of this model are unreliable in comparison to a model with a greater number of observations, especially given the heterogeneous nature of the data set. However, the results were run to provide a comparison between this model and the other two data sets, especially the fall data sets, to see if the actual sale prices of the horses impacted which variables were statistically significant to any great extent. Since the survey data are admittedly unreliable, these results cannot be taken as definite, but should at least provide for some interesting comparison. All three models are OLS regressions, one linear, one semi-log, and one log-log.

Linear Fall Survey Model Results and Discussion

Table 24 shows the results for the linear functional form. The linear model results exhibited an R-squared value of 0.3338 and a drastically lower adjusted R-squared value of

0.1621, indicating that only 16.21% of the linear variation in the dependent variable is explained by the model. This lower adjusted R-squared value also indicates possible near collinearity in the model, which due to the low number of observations is not surprising. The F-value is significant at the 1% level, indicating that as a whole at least one of the coefficients is significantly different from zero and the model explains some of the variation in the price of horses.

Six coefficients were statistically significant, including one each in the Breed, Gender, and Color categories. *Pnt* was statistically significant in the Breed category at the 10% level, and this result indicated that a Paint horse would bring $1,335.45 more than any other horse, ceteris paribus. Notably, though except for *Pnt* they are statistically insignificant, all of the breed categories had positive coefficients, and the crossbred category had a negative coefficient as would be expected. Of the gender variables, *Mare* was negative and statistically significant, indicating that female horses would bring $478.57 less than geldings, ceteris paribus. Uncut male horses were not priced statistically different than geldings.

Dun was the color which showed statistical significance at the 15% level and indicated that a horse carrying dun factor would bring $1,804.01 less than another colored horse, ceteris paribus. Notably, all of the color variables are negative though statistically insignificant except for *Dun*. This coefficient is counterintuitive as it is an exceptionally large value, greater than the average price of a horse in this model, and because typically in stock-type horses (Quarter Horse and Paint breeds, primarily) which constitute the majority of this data set, dun-factor is a desirable and unique color. This coefficient could be the result of this category having extremely low numbers, only nine observations. A horse being registered was statistically significant at the 10% level and means the horse brought $526.98 more than a grade horse, ceteris paribus. Finally, the continuous variables *Age* and *Age2* were both significant at the 1% level. By taking the derivative of the price of the horse with respect to age, we can discover that the age where a horse is at its maximum value, ceteris paribus, is 10.3146 years of age. This age is within a range commonly used in the equine world of 8-12 years of age being the prime of a horse’s life.

Table 24: Fall Survey Linear Model Results

Fall Survey Dreamhorse.com Linear Model Results (Dependent Variable = *APHorse*)				
N	R-Square	Adj R-Sq	F-Value	F-Significance
162	0.3338	0.1621	1.94	0.0046
Variable	Parameter Estimate	Standard Error	t-Value	P-statistic
Intercept	1220.18578	1396.4273	0.87	0.3839
Pht	275.58527	193.9855	1.42	0.1579
MO	100.71872	349.9929	0.29	0.7740
DK	178.54986	444.1744	0.40	0.6884
NE	97.69265	423.8903	0.23	0.8181
OK	-174.09465	387.1584	-0.45	0.6537
TX	349.86742	328.2513	1.07	0.2885
Pnt	**1335.44817****	**790.6629**	**1.69**	**0.0936**
QH	899.71007	751.2803	1.20	0.2333
Pony	307.66758	851.8943	0.36	0.7186
TB	422.45413	735.8730	0.57	0.5669
ClrBrd	317.03785	795.1448	0.40	0.6908
Arab	422.97608	818.6252	0.52	0.6063
Saddle	698.12962	797.3132	0.88	0.3829
Other	990.88319	869.9926	1.14	0.2568
Xbred	-105.27181	641.5404	-0.16	0.8699
Mare	**-478.56876****	**207.8253**	**-2.30**	**0.0229**
Stud	-175.48109	287.9934	-0.61	0.5434
InFoal	-227.18194	435.8562	-0.52	0.6031
Bay	-1258.19707	1118.0391	-1.13	0.2625
Sor	-1363.63153	1112.2339	-1.23	0.2224
Blk	-1185.77590	1131.0404	-1.05	0.2964
Grey	-1389.17476	1134.7980	-1.22	0.2231
Pal	-1450.12186	1144.2932	-1.27	0.2074
Paint	-217.63217	337.1458	-0.65	0.5197
Dun	**-1804.01155***	**1182.9209**	**-1.53**	**0.1297**
Dilute	-823.28947	1141.5709	-0.72	0.4721
Roan	-1449.75907	1142.9934	-1.27	0.2070
ChBr	-1478.46964	1123.4020	-1.32	0.1905
Online	66.19809	187.9075	0.35	0.7252
Age	**196.70518*****	**54.7351**	**3.59**	**0.0005**
Age2	**-9.53525*****	**2.6645**	**-3.58**	**0.0005**
Tmp	-46.02038	62.6647	-0.73	0.4641
Paper	**526.97586****	**256.9489**	**2.05**	**0.0423**

Semi-Log Fall Survey Model Results and Discussion

The results of the semi-log functional form are shown in Table 25. The adjusted R-squared value of 0.1374 indicates that 13.74% of the variation in the dependent variable is explained by the model. The F-value of 1.78 is significant at just greater than the 1% level, indicating that as a group the variables explain some of the variation in the price of a horse.

Nine of the independent variables are statistically significant at greater than the 15% level, in addition to the intercept being highly significant. *Pht* and *Paper* are both statistically significant at the 10% level and positively impact the price of horses. A horse advertised in a photo ad will get 49.75% more than a horse in a text ad, and a registered horse will receive a price 58.2% higher than a grade horse, ceteris paribus. Though the F-tests indicated that the group States were statistically insignificant in this data set, three of the five state variables were significant and positive at greater than the 15% level. Missouri was statistically significant at the 15% level and had a coefficient of 0.58561, indicating that a horse sold from Missouri would bring a price 58.56% higher than one sold in Kansas. The Dakotas and Texas were both statistically significant at the 10% level and indicated that horses sold from those states would bring 84.12% and 81.88% higher prices, respectively, than horses sold from Kansas, ceteris paribus. Also not as indicated by the F-tests, none of the Breed variables were statistically significant in this model. However, as expected, none of the Gender variables were either.

One of the color variables, *Grey*, was statistically significant at the 15% level, indicating that a horse that is grey in color would receive a 191.40% premium over horses of other colors, ceteris paribus. However, due to the low number of observations in this category (11 head) this coefficient is likely biased as its magnitude seems unreasonable, especially considering that the F-test on the group of Color variables did not indicate statistical significance. All three continuous variables were statistically significant at the 10% level in this model and had the expected signs. The age at which a horse reaches its maximum value can be determined by taking the derivative of *APHorse* with respect to age, and this calculation results in an age of 7.7819 years, which is just less than the range of 8 to 12 years considered the prime of a horse's life, but would allow for a greater remaining useful life. Finally, the *Tmp* coefficient indicates that with each 1 unit increase in the score, the value of the horse will drop 15.61%.

Table 25: Fall Survey Semi-Log Model Results

Fall Survey Dreamhorse.com Semi-Log Model Results (Dependent Variable = *lnAPHorse*)				
N	R-Square	Adj R-Sq	F-Value	F-Significance
162	0.3142	0.1374	1.78	0.0125
Variable	Parameter Estimate	Standard Error	t-Value	P-statistic
Intercept	**6.47941*** **	**1.5812**	**4.10**	**<.0001**
Pht	**0.49750** **	**0.2197**	**2.27**	**0.0252**
MO	**0.58561***	**0.3963**	**1.48**	**0.1419**
DK	**0.84123** **	**0.5029**	**1.67**	**0.0968**
NE	0.60707	0.4800	1.26	0.2082
OK	0.58517	0.4384	1.33	0.1843
TX	**0.81881** **	**0.3717**	**2.20**	**0.0294**
Pnt	1.13265	0.8953	1.27	0.2081
QH	0.99432	0.8507	1.17	0.2446
Pony	0.45204	0.9646	0.47	0.6401
TB	0.08586	0.8332	0.10	0.9181
ClrBrd	0.51009	0.9003	0.57	0.5720
Arab	0.79937	0.9269	0.86	0.3901
Saddle	0.41455	0.9028	0.46	0.6469
Other	1.08287	0.9851	1.10	0.2737
Xbred	-0.04071	0.7264	-0.06	0.9554
Mare	-0.19767	0.2353	-0.84	0.4025
Stud	-0.25785	0.3261	-0.79	0.4306
InFoal	-0.09475	0.4935	-0.19	0.8481
Bay	-1.42866	1.2659	-1.13	0.2612
Sor	-1.24675	1.2594	-0.99	0.3240
Blk	-1.14126	1.2807	-0.89	0.3745
Grey	**-1.91395***	**1.2849**	**-1.49**	**0.1388**
Pal	-1.20757	1.2957	-0.93	0.3531
Paint	-0.10786	0.3817	-0.28	0.7780
Dun	-1.88856	1.3394	-1.41	0.1610
Dilute	-0.92796	1.2926	-0.72	0.4741
Roan	-1.03926	1.2942	-0.80	0.4235
ChBr	-1.21639	1.2720	-0.96	0.3407
Online	0.04023	0.2128	0.19	0.8503
Age	**0.10241** **	**0.0620**	**1.65**	**0.1009**
Age2	**-0.00658** **	**0.0030**	**-2.18**	**0.0309**
Tmp	**-0.15605** **	**0.0710**	**-2.20**	**0.0296**
Paper	**0.58200** **	**0.2909**	**2.00**	**0.0476**

Log-Log Fall Survey Model Results and Discussion

The final functional form used on the survey data is the log-log form where the continuous variables on the right hand side are natural logs as well as the dependent variable, actual price of the horse, and is shown in Table 26. The R-squared value is 0.2924, and the adjusted R-squared is 0.1100, indicating that 11% of the variation in the dependent variable is explained by the model. The F-value of 1.6 is significant at the 5% level, indicating that at least some of the independent variable's coefficients are significantly different from zero.

The intercept and seven of the coefficients are statistically significant at greater than the 15% level. *Pht* and *Paper* are both statistically significant at the 10% level and positive, as expected. Advertising using a photo ad made the price of a horse increase by 50.25% as compared to a text ad, and the horse being registered increases the value by 56.53% versus grade horses, ceteris paribus. Unlike the F-tests indicated for this model, two of the state variables are statistically significant at greater than the 15% level. The Dakotas' coefficient of 0.7519 is significant at the 15% level and indicates that a horse sold from the Dakotas will bring a price 75.19% higher than a horse sold from Kansas, ceteris paribus. The coefficient on Texas was also positive and significant at the 10% level, indicating that a horse sold from Kansas would bring 71.88% less than a horse sold from the state of Texas, ceteris paribus. The Breed, Gender, and Color variables were not statistically significant in this model.

The coefficient for temperament scores indicates that a 1% increase in temperament score results in a 41.89% decrease in the actual price of the horse, which is the expected sign at a significance level of 10%. This indicates that an increase in temperament score from a 2 to a 3 would decrease price by over 500%, which seems to be an unusually large magnitude as compared to the other functional form results and may indicate the log-log model is not the appropriate functional form. *lnAge* and *lnAge2* are significant at the 15% and 10% levels respectively, and *lnAge** = 1.0700, which when the exponential is taken results in an age of maximum value of 2.9155 years, which is extremely young and seems unlikely. This indicates that this functional form is probably not the best fit for this data set.

Table 26: Fall Survey Log-Log Model Results

Fall Survey Dreamhorse.com Log-Log Model Results (Dependent Variable = *lnAPHorse*)				
N	R-Square	Adj R-Sq	F-Value	F-Significance
162	0.2924	0.1100	1.60	0.0334
Variable	Parameter Estimate	Standard Error	t-Value	P-statistic
Intercept	**6.83264*** **	**1.5956**	**4.28**	**<.0001**
Pht	**0.50250** **	**0.2236**	**2.25**	**0.0264**
MO	0.49258	0.4030	1.22	0.2238
DK	**0.75188***	**0.5107**	**1.47**	**0.1434**
NE	0.36140	0.4830	0.75	0.4557
OK	0.42499	0.4613	0.92	0.3587
TX	**0.71882** **	**0.3768**	**1.91**	**0.0587**
Pnt	0.94848	0.9114	1.04	0.3000
QH	0.78958	0.8659	0.91	0.3636
Pony	0.23284	0.9865	0.24	0.8138
TB	-0.17846	0.8458	-0.21	0.8332
ClrBrd	0.18161	0.9169	0.20	0.8433
Arab	0.46438	0.9404	0.49	0.6223
Saddle	0.28496	0.9169	0.31	0.7565
Other	0.75673	0.9996	0.76	0.4504
Xbred	0.18477	0.7444	0.25	0.8044
Mare	-0.15184	0.2420	-0.63	0.5315
Stud	-0.19002	0.3716	-0.51	0.6100
InFoal	0.04143	0.4958	0.08	0.9335
Bay	-1.32437	1.2900	-1.03	0.3065
Sor	-1.12456	1.2812	-0.88	0.3817
Blk	-0.92097	1.2998	-0.71	0.4799
Grey	-1.80194	1.3096	-1.38	0.1712
Pal	-1.04345	1.3189	-0.79	0.4303
Paint	-0.13646	0.3892	-0.35	0.7264
Dun	-1.73445	1.3583	-1.28	0.2040
Dilute	-0.89936	1.3188	-0.68	0.4965
Roan	-0.88384	1.3202	-0.67	0.5044
ChBr	-1.23375	1.2983	-0.95	0.3438
Online	0.04585	0.2165	0.21	0.8326
lnAge	**0.33509***	**0.2280**	**1.47**	**0.1442**
$(lnAge)^2$	**-0.15658** **	**0.0907**	**-1.73**	**0.0869**
Tmp	**-0.41889** **	**0.2043**	**-2.05**	**0.0424**
Paper	**0.56526** **	**0.2938**	**1.92**	**0.0566**

Brief Comparison of the Three Functional Forms

A comparison between the three functional forms in Table 27 shows that the only variables that were consistently statistically significant were the age terms and the registration of a horse. The two logarithmic functions both had statistically significant intercepts, and the variable *Pht* was significant in both at the 10% level and was close to the same in magnitude. The two logarithmic functions also shared statistical significance in the states of the Dakotas and Texas, while the state of Missouri was only positive and statistically significant in the semi-log model. The only model having a statistically significant coefficient on breed was the linear model, which matches the results of the F-tests but seems to be unusually large in magnitude and is only significant at the 10% level. The linear model is also the only model having a statistically significant gender coefficient, as was indicated by the F-tests. Both the linear and semi-log models had statistically significant coefficients in the Color group, but in both cases they had unusually large magnitudes and were coefficients on colors with few observations. The two logarithmic functions also shared statistical significance in the temperament coefficient and both had the expected negative sign. As being registered and the age variables are consistently statistical significant across functional forms, they exhibit definite economic significance to the price of horses in this data set.

The semi-log model appears to capture information in the data better than the log-log model. However, between the linear and semi-log model, it is difficult to distinguish if one is clearly a better fit than the other. Though this data set clearly had near collinearity issues due to the low number of observations, it still showed similar results to larger data sets using asking prices. None of the coefficients in these models are trustworthy due to the data problems, but the comparison to the other two models in Chapter 8 gives further evidence that the use of asking price is a good proxy for the actual sale price of the horse. However, to avoid collinearity issues, a larger sample size with actual sale prices may exhibit further market information not captured with the asking price proxy.

Table 27: Comparison of Fall Survey Models

Survey Comparison of Coefficients for Linear, Semi-Log, and Log-Log Parameter Estimates and Significance						
	Linear		Semi-Log		Log-Log[a]	
Variable	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Intercept	1220.1858	0.3839	**6.4794*** **	<.0001	**6.8326*** **	<.0001
Pht	275.5853	0.1579	**0.4975** **	0.0252	**0.5025** **	0.0264
MO	100.7187	0.7740	**0.5856***	0.1419	0.4926	0.2238
DK	178.5499	0.6884	**0.8412** **	0.0968	**0.7519***	0.1434
NE	97.6927	0.8181	0.6071	0.2082	0.3614	0.4557
OK	-174.0947	0.6537	0.5852	0.1843	0.4250	0.3587
TX	349.8674	0.2885	**0.8188** **	0.0294	**0.7188** **	0.0587
Pnt	**1335.4482** **	0.0936	1.1327	0.2081	0.9485	0.3000
QH	899.7101	0.2333	0.9943	0.2446	0.7896	0.3636
Pony	307.6676	0.7186	0.4520	0.6401	0.2328	0.8138
TB	422.4541	0.5669	0.0859	0.9181	-0.1785	0.8332
ClrBrd	317.0379	0.6908	0.5101	0.5720	0.1816	0.8433
Arab	422.9761	0.6063	0.7994	0.3901	0.4644	0.6223
Saddle	698.1296	0.3829	0.4146	0.6469	0.2850	0.7565
Other	990.8832	0.2568	1.0829	0.2737	0.7567	0.4504
Xbred	-105.2718	0.8699	-0.0407	0.9554	0.1848	0.8044
Mare	**-478.5688** **	0.0229	-0.1977	0.4025	-0.1518	0.5315
Stud	-175.4811	0.5434	-0.2579	0.4306	-0.1900	0.6100
InFoal	-227.1819	0.6031	-0.0948	0.8481	0.0414	0.9335
Bay	-1258.1971	0.2625	-1.4287	0.2612	-1.3244	0.3065
Sor	-1363.6315	0.2224	-1.2468	0.3240	-1.1246	0.3817
Blk	-1185.7759	0.2964	-1.1413	0.3745	-0.9210	0.4799
Grey	-1389.1748	0.2231	**-1.9140***	0.1388	-1.8019	0.1712
Pal	-1450.1219	0.2074	-1.2076	0.3531	-1.0435	0.4303
Paint	-217.6322	0.5197	-0.1079	0.7780	-0.1365	0.7264
Dun	**-1804.0116** **	0.1297	-1.8886	0.1610	-1.7345	0.2040
Dilute	-823.2895	0.4721	-0.9280	0.4741	-0.8994	0.4965
Roan	-1449.7591	0.2070	-1.0393	0.4235	-0.8838	0.5044
ChBr	-1478.4696	0.1905	-1.2164	0.3407	-1.2338	0.3438
Online	66.1981	0.7252	0.0402	0.8503	0.0459	0.8326
Age/lnAge	**196.7052*** **	0.0005	**0.1024***	0.1009	**0.3351***	0.1442
Age2/(lnAge) ^{2}	**-9.5353*** **	0.0005	**-0.0066** **	0.0309	**-0.1566** **	0.0869
Tmp/lnTmp	-46.0204	0.4641	**-0.1561** **	0.0296	**-0.4189** **	0.0424
Paper	**526.9759** **	0.0423	**0.5820** **	0.0476	**0.5653** **	0.0566

*,**,*** Denotes Significance at the 15%, 10%, and 1% levels, respectively

[a] Log-log is the only format that uses *lnAge*, *lnAge2*, and *lnTmp*

CHAPTER 8 - Comparison of Models

This chapter will compare the data and results across the three data sets, primarily comparing spring to fall to examine changes in the market across at two different points in time; and fall to the fall survey data to show that similar distributions of breeds, colors, states, ages, and prices were obtained and used in the two models. A comparison of summary statistics will be first, followed by the White's test and F-test results, and finally compare the results of the regressions for each functional form across the data sets. Lastly, we will summarize which variables seem to be statistically and economically significant in determining a horse's price regardless of the functional form, season, or use of actual versus asking price.

Comparison of Summary Statistics

Referring to Table 28 on the following pages, which present data from the summary statistics of the variables for each model, we can compare the summary statistics between the models. This shows that though the average price of horses in the survey data set is lower than the asking prices, the distribution of types of horses as far as breed, gender, and color are similar and therefore comparisons can be drawn between the model results. It also allows us to study differences between the spring and fall markets and examine any traits of the horses (or their ads) that played a role in their response to the survey through dreamhorse.com.

Between the spring and fall data set, 63 less horses were sold in the fall than the spring, which could be due to a weakening market if some of the survey participants and recent news is correct that the horse market is severely depressed. Similar percentages were advertised using photo ads between spring and fall, approximately 40% in both cases. Only 73% of horses were advertised as registered in the spring compared to 75% in the fall, which is not a large difference, but could be accounted for by young, newly foaled horses who were not registered in the spring becoming registered by the fall, though this is also unlikely considering the small number of equines under a year of age listed in the spring data set. It is interesting that almost 75% of horses are registered throughout the entire data set, as the NAHMS 2005 Equine report reports that only 47.8% of horses are registered in the U.S. (USDA, 2006b). A higher percentage of horses from the spring data set were listed as sold online than the fall data, 42% versus less than 36%. This could be due to a weakening horse market, as mentioned by several of the survey

Table 28: Comparison of Summary Statistics

					Count	Spring	Fall	Survey							
					N	470	407	162							
	Mean			**Std. Dev.**			**Minimum/Maximum**						**Sum**		
Variable	Spring	Fall	Survey	Spring	Fall	Survey	Spring		Fall		Survey		Spring	Fall	Survey
Pht	0.4085	0.3956	0.4444	0.4921	0.4896	0.4984	0	1	0	1	0	1	192	161	72
KS	0.0766	0.0663	0.0864	0.2662	0.2492	0.2819	0	1	0	1	0	1	36	27	14
MO	0.1574	0.1892	0.2284	0.3646	0.3921	0.4211	0	1	0	1	0	1	74	77	37
DK	0.0404	0.0663	0.0802	0.1972	0.2492	0.2725	0	1	0	1	0	1	19	27	13
NE	0.0255	0.0786	0.0926	0.1579	0.2695	0.2908	0	1	0	1	0	1	12	32	15
OK	0.1574	0.1794	0.1235	0.3646	0.3841	0.3300	0	1	0	1	0	1	74	73	20
TX	0.5426	0.4201	0.3889	0.4987	0.4942	0.4890	0	1	0	1	0	1	255	171	63
Pnt	0.2149	0.1523	0.1296	0.4112	0.3598	0.3369	0	1	0	1	0	1	101	62	21
QH	0.5128	0.5479	0.5370	0.5004	0.4983	0.5002	0	1	0	1	0	1	241	223	87
Pony	0.0809	0.1106	0.1420	0.2729	0.3140	0.3501	0	1	0	1	0	1	38	45	23
TB	0.0638	0.0713	0.0679	0.2447	0.2576	0.2524	0	1	0	1	0	1	30	29	11
Color	0.0511	0.0491	0.0556	0.2204	0.2164	0.2298	0	1	0	1	0	1	24	20	9
Arab	0.0404	0.0393	0.0370	0.1972	0.1946	0.1894	0	1	0	1	0	1	19	16	6
Saddle	0.0851	0.0713	0.0556	0.2793	0.2576	0.2298	0	1	0	1	0	1	40	29	9
Other	0.0404	0.0418	0.0556	0.1972	0.2003	0.2298	0	1	0	1	0	1	19	17	16
Xbred	0.1149	0.1032	0.0988	0.3192	0.3046	0.2993	0	1	0	1	0	1	54	42	16
Geld	0.3681	0.3710	0.3395	0.4828	0.4837	0.4750	0	1	0	1	0	1	173	151	55
Mare	0.5170	0.4791	0.4815	0.5002	0.5002	0.5012	0	1	0	1	0	1	243	195	78
Stud	0.1149	0.1499	0.1790	0.3192	0.3574	0.3846	0	1	0	1	0	1	54	61	29
InFoal	0.0532	0.0541	0.0556	0.2247	0.2264	0.2298	0	1	0	1	0	1	25	22	9
Bay	0.2255	0.1818	0.1420	0.4184	0.3862	0.3501	0	1	0	1	0	1	106	74	23
Sor	0.2298	0.1794	0.1852	0.4211	0.3841	0.3897	0	1	0	1	0	1	108	73	30
Blk	0.1021	0.0737	0.0864	0.3031	0.2616	0.2819	0	1	0	1	0	1	48	30	14
Grey	0.0936	0.0860	0.0679	0.2916	0.2807	0.2524	0	1	0	1	0	1	44	35	11

Pal	0.0660	0.0909	0.0741	0.2485	0.2878	0.2627	0	1	0	1	0	1	31	37	12
Paint	0.1468	0.1229	0.1235	0.3543	0.3287	0.3300	0	1	0	1	0	1	69	50	20
Dun	0.0574	0.0541	0.0556	0.2329	0.2264	0.2298	0	1	0	1	0	1	27	22	9
Dilute	0.0596	0.0786	0.1111	0.2369	0.2695	0.3152	0	1	0	1	0	1	28	32	18
Roan	0.0681	0.0811	0.0926	0.2522	0.2733	0.2908	0	1	0	1	0	1	32	33	15
ChBr	0.0957	0.1720	0.1790	0.2946	0.3778	0.3846	0	1	0	1	0	1	45	70	29
Online	0.4213	0.3563	0.3951	0.4943	0.4795	0.4904	0	1	0	1	0	1	198	145	64
Age	7.3202	5.7666	5.7500	5.3533	5.1336	5.4088	0.5	26	0.5	25	0.5	25	3441	2347	931.5
Age2	82.1824	59.5430	62.1373	112.9834	92.0530	102.7873	0.25	676	0.25	625	0.25	625	38626	24234	10066
Tmp	3.0459	3.2548	2.9841	1.5222	1.6146	1.5104	1	8	1	8	1	8	1432	1325	483.4286
lnAge	1.6788	1.2422	1.1929	0.8612	1.1305	1.1779	-0.6931	3.2581	-0.6931	3.2189	-0.6931	3.2189	789.0568	505.5694	193.2466
(lnAge)^2	3.5586	2.8178	2.8020	2.6412	2.5767	2.6588	0	10.6152	0	10.3612	0	10.3612	1673	1147	453.9164
lnTmp	0.9708	1.0475	0.9639	0.5632	0.5373	0.5270	0	2.0794	0	2.0794	0	2.0794	456.2757	426.3369	156.1516
Paper	0.7277	0.7518	0.7160	0.4456	0.4325	0.4523	0	1	0	1	0	1	342	306	116
PHorse	1784.14	1713.64	1526.41	1240.46	1214.53	1105.00	1	5000	1	5000	1	5000	838547	697453	247278
lnPHorse	7.1843	7.1223	6.9496	0.9647	1.0422	1.2331	0	8.5172	0	8.5172	0	8.5172	3,376.6393	2,898.7661	1,125.8297

participants, or that less people are looking online to buy horses and more horses are being sold by word of mouth or through other means. Texas had a lower percentage of horses advertised online in the fall data set than spring, while the states of Oklahoma, Nebraska, the Dakotas, and Missouri all increased in percent of horses sold from their state. Quarter Horses dominated the breeds in both data sets, with no other breed showing large differences in percentages between the spring and the fall data sets. The percentage of crossbred horses is also close to the same, just slightly lower in the fall which could be the reason a higher percentage of horses in the fall data set were registered, as registering pedigreed purebred horses is much more common than registering grade horses based on color or performance traits. Distribution of gender was also similar, with just over a third being geldings, approximately half being female horses, and the remainder being uncut male horses. This makes sense if we assume that half of all horses are born male and half born female, so if you combine uncut males and geldings it would be approximately 50% of the observations. The percentage of mares in foal is approximately equal across the seasons, which is somewhat surprising if the market is down because you would have expected fewer mares to be rebred this spring than the previous year, and the percentage of mares in foal is slightly higher, 5.41% versus 5.32%, in the fall than in the spring.

Of the Color categories, where bays and sorrels dominated in the spring data set, those two categories remain high in the fall but the colors are more evenly distributed, and a much higher percentage of chestnut and brown horses is found in the fall data set. Pinto horses, those with multiple colors, are slightly less common in the fall data than the spring. The average temperament scores are similar, though slightly higher in the fall than the spring. However, the average age is notably lower in the fall than in the spring. If we compare the age distributions between the two data sets, seen previously in Figures 2 and 4, we can see that this shift in the average is likely due to a very high number of horses under a year old entering the market in the fall. These horses are quite likely weanlings born the previous spring who are now old enough to be weaned from their dam and sold by horse breeders. An interesting note for breeders is that it could be a positive or negative to market your weanlings in the fall market; there is a flood of young horses in the market this time of year, but also that is when buyers will be looking to buy young horses and the breeder could avoid feeding and caring for the horse through the winter by selling in the fall.

In comparing the dependent variable, asking price of the horse in this case, we see that the averages are similar, though slightly lower in the fall. Once again, referring back to figures in earlier chapters, Figures 3 and 5 show the price distributions for these data sets and show similarly skewed distributions. The spring and fall data sets showed similar standard deviations across all variables, nothing notable stood out. Overall, the minimums and maximums were extremely similar, with the ages ranging from a minimum of 0.5 to maximums of 25 years in the fall and to 26 in the spring, temperaments ranging from 1 to 8, and all the binary variables being confined to 0's and 1's and price limited to a maximum of $5,000. The number of observations for the states of the Dakotas and Nebraska increased from 19 to 27 and from 12 to 32, respectively, between the spring and fall data sets. Arabians and the group of other breeds and equine species both had low counts in the fall and spring data sets, at 19 each in the fall data set and 16 and 17, respectively, in the spring data sets. With such a low number of observations, the results on these categories could be questionable in the models.

Between the fall data and the fall survey data, the number of observations and sums are obviously lower since the survey data are taken from a survey of the horse sellers from the fall data set. The minimums and maximums are also precisely the same between the two data sets. However, a comparison of the averages is still of interest. Owners of registered horses did not report back as often in the survey, as only 71.6% of the horses were registered in the survey data versus 75.2% of the fall data. Horse owners who sold their horses on dreamhorse.com were also slightly more likely to report back, with 39.5% of the survey respondents having sold the horse online versus 35.6% in the fall data. Those using photo ads were also slightly more likely to respond, as 44.4% of the survey horses had photo ads as compared to 39.6% in the fall data. Interestingly, lower percentages of horse sellers from the southern states of Texas and Oklahoma responded to the survey than from all of the other states. Quarter Horses again dominated both data sets, and there were no major differences in percentages of the other breeds. Additionally, approximately 10% of each data set was crossbred horses. Colt and stallion owners seemed more apt to respond to the survey than did gelding owners, while mare and filly owner responses were approximately the same, as were the responses of sellers of mares in foal. A lower percentage of sellers of bay, grey, and palomino horse sellers responded to the survey than the other colors and a higher percentage of dilute gene horse sellers responded to the survey. The average age is close between the two data sets, but the average temperament score is slightly lower on average in the survey data than the online fall data. Finally, the dependent variable price is much lower in the survey data, which is not surprising as it is the variable that the survey was focused on and the information reported in the survey found that prices were on average 10% lower than those online. Looking at Figures 5 and 8 we can see that the price distribution takes on a similar shape in both of the data sets but around a slightly lower mean and median in the survey data.

Comparison of White Tests and F-Tests

Referring back to Tables 8, 15, and 22 we can compare the results of the White's Tests across the three models. The spring data set was the only one that exhibited a concerning degree of heteroskedasticity requiring the use of the GMM model to correct for this issue. The two logarithmic models in the fall data have slightly more significant White's Test statistics than the linear model for that data set or the survey data, but none of these were statistically of concern.

Comparing Tables 9, 16, and 23, we can examine the F-Statistic differences between the three models. The F-tests are very similar between the fall and spring data sets, with the Color group showing no statistical significance in either one and the other three groups of binary variables all showing statistical significance at the 5% level or greater. The survey data set F-tests are drastically different, likely in part due to the much smaller number of observations in this model. The Breed category is the only one still expressing statistical significance in all three models, while Gender shows statistical significance in only the linear model. Like the spring and fall data sets, the Color variables do not show statistical significance, but unlike the other two data sets, neither do the States.

Comparison of Model Results

Using Tables 29, 30, and 31 shown below, comparisons will be given of the three functional forms between the spring and fall data sets using the pricing data obtained online, and between the fall and survey data sets.

Spring to Fall Comparison

In the linear model between the spring and fall, the similarities are that the coefficients for *Pht*, *Mare*, *Age*, *Age2*, and *Paper* all share the same signs and are highly statistically significant, indicating that the significance of these variables are not affected by any seasonal difference between the two samples. The age at which maximum value is reached is also similar between the two linear models. Additionally, the fall model has highly statistically significant coefficients for *Tmp* and *Stud*, which the spring model does not. This could mean that stallions and colts are worth less in the fall than the spring, or just at these two particular points in time as one sample from each season does not allow for much seasonal comparison. Breeding stallions would be less desirable in the fall if you were looking to breed, because you would have to over-winter them before use, where as a spring purchase of a stallion could be used for matings right away. The statistical significance of temperament could be the result of buyers getting pickier in a weaker horse market or some other unknown factor. Other differences between the spring and fall linear models were that the state of Texas was statistically significant in the fall model but not in the spring, saddle horses were statistically significantly more valuable in the spring but not in the fall, and the intercept held statistical significance in the fall model. Additionally, *Pht* and *Paper* both expressed lower magnitudes in the fall than the spring, suggesting there is less value obtained from these traits in the later months of the year.

Comparing the semi-log models between the fall and the spring data sets, we find many similarities but also differences. The intercept is highly statistically significant in each, as well as the variables *Pht*, *Age*, *Age2*, and *Paper*. Once again, these variables prove to stand out in their statistical significance from the rest across time and data sets. Additionally, the same four states of the Dakotas, Nebraska, Oklahoma, and Texas exhibited statistical significance between the spring and the fall, indicating these markets saw no seasonal or market differences except in magnitude from spring to fall, but they were all positive and statistically significant as compared to horses sold from the Kansas market. Of these four, Texas, Oklahoma, and the Dakotas' coefficients all increased in magnitude from spring to fall. The age terms were similar and the age of maximum value was also similar between these two models. Being registered seemed to carry a slightly lower magnitude effect in the fall than the spring, while having a photo ad showed a slightly higher magnitude impact. Once again, the differences included *Tmp* being statistically significant in the fall model but not in the spring, and this could be due to the reasons mentioned above or due to other unknown causes. In the spring model, ponies and female horses saw significantly decreased values as compared to their counterparts, ceteris paribus, while the fall data set saw significantly increased value in saddle horses and decreased value in stallions and colts as compared to the spring model. Further examination and data would be required to explain these differences.

Table 29: Comparison of Linear Models across Data Sets

Comparison Across Data Sets for Linear Models Parameter Estimates and Significance						
	Spring Linear		Fall Linear		Survey Linear	
Variable	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Intercept	328.5102	0.4406	**1491.6841****	0.0696	1220.1858	0.3839
Pht	**689.0551*****	<.0001	**524.2073*****	<.0001	275.5853	0.1579
MO	-181.9980	0.3616	-17.6617	0.9443	100.7187	0.7740
DK	309.0224	0.3422	282.1825	0.3527	178.5499	0.6884
NE	607.7126	0.2259	285.3626	0.3291	97.6927	0.8181
OK	83.3529	0.6810	99.8038	0.7004	-174.0947	0.6537
TX	167.0538	0.3405	**349.8313***	0.1365	349.8674	0.2885
Pnt	-107.7840	0.7006	35.1747	0.9361	**1335.4482****	0.0936
QH	212.1241	0.4095	133.2331	0.7419	899.7101	0.2333
Pony	-233.1500	0.4823	-230.1243	0.6259	307.6676	0.7186
TB	-89.7903	0.7852	-319.2987	0.4457	422.4541	0.5669
ClrBrd	-313.9920	0.2075	-289.2208	0.5200	317.0379	0.6908
Arab	-39.5452	0.9147	-44.3302	0.9271	422.9761	0.6063
Saddle	**629.1094****	0.0472	530.9073	0.2502	698.1296	0.3829
Other	30.5419	0.9152	-113.9632	0.8133	990.8832	0.2568
Xbred	165.5213	0.4346	329.2394	0.3824	-105.2718	0.8699
Mare	**-360.0230*****	0.0031	**-462.4523*****	0.0003	**-478.5688****	0.0229
Stud	-140.9320	0.4467	**-507.7861*****	0.0099	-175.4811	0.5434
InFoal	-304.2520	0.1831	-358.7005	0.1683	-227.1819	0.6031
Bay	13.3859	0.9568	-585.4993	0.3833	-1258.1971	0.2625
Sor	-48.2338	0.8306	-731.1933	0.2761	-1363.6315	0.2224
Blk	147.2820	0.4385	-737.1464	0.2818	-1185.7759	0.2964
Grey	19.9014	0.9456	-490.6913	0.4796	-1389.1748	0.2231
Pal	139.8960	0.6359	-659.0747	0.3360	-1450.1219	0.2074
Paint	132.4278	0.3986	79.7665	0.7076	-217.6322	0.5197
Dun	159.1722	0.6545	-720.2436	0.2883	**-1804.0116****	0.1297
Dilute	55.4199	0.8478	-313.6119	0.6400	-823.2895	0.4721
Roan	392.2400	0.1690	-639.2823	0.3486	-1449.7591	0.2070
ChBr	269.5000	0.2535	-693.3841	0.3039	-1478.4696	0.1905
Online	-11.2055	0.9097	-3.7952	0.9746	66.1981	0.7252
Age	**158.7203*****	0.0001	**179.3478*****	<.0001	**196.7052*****	0.0005
Age2	**-7.0226*****	0.0013	**-8.6741*****	<.0001	**-9.5353*****	0.0005
Tmp	-42.2749	0.2487	**-92.9541*****	0.0079	-46.0204	0.4641
Paper	**850.1931*****	<.0001	**602.2512*****	0.0001	**526.9759****	0.0423
*,**,*** Denotes Significance at the 15%, 10%, and 1% levels, respectively						

Table 30: Comparisons of Semi-Log Models across Data Sets

Comparison Across Data Sets for Semi-Log Models						
Parameter Estimates and Significance						
	Spring Semi-Log		Fall Semi-Log		Survey Semi-Log	
Variable	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Intercept	**6.4217*** **	<.0001	**6.6389*** **	<.0001	**6.4794*** **	<.0001
Pht	**0.3554*** **	<.0001	**0.3814*** **	0.0003	**0.4975** **	0.0252
MO	-0.0282	0.8675	0.2097	0.3491	**0.5856* **	0.1419
DK	**0.3135* **	0.1422	**0.5670** **	0.0353	**0.8412** **	0.0968
NE	**0.4390** **	0.0610	**0.4340** **	0.0940	0.6071	0.2082
OK	**0.2575** **	0.0895	**0.3345* **	0.1457	0.5852	0.1843
TX	**0.2925** **	0.0457	**0.5379*** **	0.0099	**0.8188** **	0.0294
Pnt	-0.4351	0.3109	0.1891	0.6264	1.1327	0.2081
QH	-0.1695	0.6376	0.2646	0.4600	0.9943	0.2446
Pony	**-0.5290* **	0.1340	0.1488	0.7217	0.4520	0.6401
TB	-0.3921	0.3646	-0.1960	0.5968	0.0859	0.9181
ClrBrd	-0.3071	0.2045	0.0826	0.8355	0.5101	0.5720
Arab	-0.1374	0.6308	0.2040	0.6345	0.7994	0.3901
Saddle	0.0904	0.8271	**0.5939* **	0.1464	0.4146	0.6469
Other	-0.2990	0.3395	0.0533	0.9006	1.0829	0.2737
Xbred	0.0202	0.9160	0.1253	0.7070	-0.0407	0.9554
Mare	**-0.2804*** **	0.0002	-0.1540	0.1715	-0.1977	0.4025
Stud	-0.1447	0.2415	**-0.3904** **	0.0248	-0.2579	0.4306
InFoal	-0.1167	0.4132	-0.2144	0.3518	-0.0948	0.8481
Bay	-0.1207	0.6702	-0.4800	0.4194	-1.4287	0.2612
Sor	-0.0533	0.8469	-0.4264	0.4729	-1.2468	0.3240
Blk	-0.0655	0.8063	-0.4807	0.4276	-1.1413	0.3745
Grey	-0.0258	0.9290	-0.2995	0.6259	**-1.9140* **	0.1388
Pal	0.1043	0.7142	-0.4669	0.4413	-1.2076	0.3531
Paint	0.1867	0.1977	-0.1148	0.5422	-0.1079	0.7780
Dun	0.0197	0.9473	-0.4234	0.4805	-1.8886	0.1610
Dilute	0.0112	0.9687	-0.6096	0.3046	-0.9280	0.4741
Roan	0.2512	0.3706	-0.3860	0.5225	-1.0393	0.4235
ChBr	-0.0528	0.8539	-0.5136	0.3895	-1.2164	0.3407
Online	-0.0369	0.5973	-0.1501	0.1557	0.0402	0.8503
Age	**0.1247*** **	<.0001	**0.1190*** **	0.0004	**0.1024* **	0.1009
Age2	**-0.0059*** **	<.0001	**-0.0063*** **	0.0004	**-0.0066** **	0.0309
Tmp	-0.0319	0.2225	**-0.1231*** **	<.0001	**-0.1561** **	0.0296
Paper	**0.7182*** **	0.0002	**0.5906*** **	<.0001	**0.5820** **	0.0476
*,**,*** Denotes Significance at the 15%, 10%, and 1% levels, respectively						

Table 31: Comparisons of Log-Log Models across Data Sets

Comparison Across Data Sets for Log-Log Models						
Parameter Estimates and Significance						
	Spring Log-Log		Fall Log-Log		Survey Log-Log	
Variable	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Intercept	**6.1050*** **	<.0001	**6.8334*** **	<.0001	**6.8326*** **	<.0001
Pht	**0.3102*** **	0.0006	**0.3991*** **	0.0002	**0.5025** **	0.0264
MO	-0.0563	0.7380	0.1598	0.4781	0.4926	0.2238
DK	0.2932	0.1633	**0.5530** **	0.0417	**0.7519* **	0.1434
NE	**0.4367** **	0.0668	0.2979	0.2550	0.3614	0.4557
OK	**0.2652** **	0.0719	0.2346	0.3124	0.4250	0.3587
TX	**0.2928** **	0.0388	**0.4640** **	0.0268	**0.7188** **	0.0587
Pnt	-0.3546	0.4078	0.1482	0.7061	0.9485	0.3000
QH	-0.0867	0.8080	0.2089	0.5639	0.7896	0.3636
Pony	-0.4783	0.1748	0.1235	0.7706	0.2328	0.8138
TB	-0.3284	0.4368	-0.2379	0.5248	-0.1785	0.8332
ClrBrd	-0.2661	0.2866	0.0107	0.9788	0.1816	0.8433
Arab	-0.0595	0.8377	0.1122	0.7956	0.4644	0.6223
Saddle	0.1506	0.7146	0.5707	0.1670	0.2850	0.7565
Other	-0.2351	0.4684	-0.0138	0.9745	0.7567	0.4504
Xbred	-0.0117	0.9507	0.1596	0.6367	0.1848	0.8044
Mare	**-0.2662*** **	0.0005	-0.1532	0.1843	-0.1518	0.5315
Stud	0.0091	0.9454	**-0.3111* **	0.1100	-0.1900	0.6100
InFoal	-0.1495	0.3040	-0.1608	0.4875	0.0414	0.9335
Bay	-0.0562	0.8310	-0.4290	0.4759	-1.3244	0.3065
Sor	-0.0197	0.9382	-0.4052	0.5002	-1.1246	0.3817
Blk	0.0165	0.9453	-0.4073	0.5067	-0.9210	0.4799
Grey	0.0045	0.9867	-0.2696	0.6644	-1.8019	0.1712
Pal	0.1876	0.4864	-0.3948	0.5201	-1.0435	0.4303
Paint	**0.2135* **	0.1402	-0.1307	0.4915	-0.1365	0.7264
Dun	0.1135	0.6876	-0.3410	0.5747	-1.7345	0.2040
Dilute	0.0228	0.9311	-0.6403	0.2862	-0.8994	0.4965
Roan	**0.4256* **	0.1104	-0.3363	0.5823	-0.8838	0.5044
ChBr	0.0635	0.8097	-0.5362	0.3740	-1.2338	0.3438
Online	-0.0156	0.8249	**-0.1560* **	0.1465	0.0459	0.8326
lnAge	**0.7533*** **	0.0009	**0.3139*** **	0.0072	**0.3351* **	0.1442
$(lnAge)^2$	**-0.2004** **	0.0118	**-0.0895*** **	0.0543	**-0.1566** **	0.0869
lnTmp	-0.0484	0.5041	**-0.3318*** **	0.0005	**-0.4189** **	0.0424
Paper	**0.6970*** **	0.0002	**0.5941*** **	<.0001	**0.5653** **	0.0566
*,**,*** Denotes Significance at the 15%, 10%, and 1% levels, respectively						

The log-log models for the spring and fall data sets all have highly statistically significant intercepts and coefficients for photo ads, the natural log of age, and whether or not the horse is registered. The magnitude for *Paper* decreased from the spring to the fall, giving the possible indication that a horse's registration is worth less in the fall. The natural log of age squared is highly significant in the fall data set and statistically significant at greater than the 2% level in the spring data set, so these are similar results. Once again, temperament is statistically significant to the fall data set but not the spring. Horses being sold online and uncut male horses are also negative and statistically significant at the 15% level in the fall data set but not significant in the spring data set. The spring model shows statistical significance in two color categories, *Roan* and *Paint*, which does not occur in any other models. Texas is statistically significant at the 10% level in both models and positive, while Nebraska and Oklahoma are positive and significant in the spring data set and the Dakotas are positive and significant in the fall data set. This could indicate some difference in the markets between the spring and the fall.

Fall to Survey Data Comparison

Due to the weak nature of the data from the survey, likely due to a lack of observations, many coefficients that were expected to be statistically significant and were in the fall and spring data sets were not in the survey data. Some statistics also indicated significance where it intuitively or based on the group F-tests did not make sense. The linear models had similarities in the coefficient on mare, being statistically significant, negative and of similar magnitude. In addition, the age variables were similar and indicated similar ages of maximum value, and the value of a horse being registered was clearly shown. However, the color and breed characteristics which indicated statistical significance in the survey data did not have significance in the fall data; and the intercept, temperament scores, stallion and colt variable, state of Texas, and photo ad variables were also not statistically significant in the survey data though they were in the fall data set.

In the semi-log data set, many more coefficients expressed statistical significance in the survey data using this functional form than the linear functional form. Photo ads, the states of the Dakotas and Texas, the age terms, temperament scores, and registration binary variable are all statistically significant in both models at the 15% level. However, the survey data did not indicate any statistical significance in the gender terms, while the fall data set showed significance with a negative coefficient on the variable *Stud*. The survey also did not capture the statistical significance of the states of Nebraska and Oklahoma as the fall data model did, nor the positive statistical significance of saddle horses over other breed types. Additionally, the survey data indicated statistical significance in the state of Missouri, which was not indicated in any previous model, and in the color grey, which was a variable with a low number of observations and therefore this determination of statistical significance is doubtful.

Finally, comparing the log-log model between the fall and survey data sets, we see similar levels of statistical significance in the intercept, the states of the Dakotas and Texas, having a photo ad versus a text ad, being registered versus being a grade horse, and in the natural logs of age, age squared, and temperament. The only two differences between these models besides magnitude differences in these coefficients is that the variables *Stud* and *Online* are both statistically significant and negative at the 15% level, and this significance is not seen in the survey data. However, it is clear to see that there are many similarities in the statistically significant variables across these two data sets, especially in the logarithmic models.

Summary

The semi-log model appears to have more statistically significant variables in both the spring and fall data sets than the other functional forms and appears to be the best fit for the data, as was found to be true in the majority of previous equine related hedonic pricing models. The semi-logarithmic functional form clearly captures more significant variables and is consistently a better fit than the log-log form. The linear form seems comparable to the semi-log form in goodness of fit, but does not capture some expected significance in individual variables, particularly in the State group. Changes in price based on characteristics of the horses appear to occur on a percentage basis as opposed to the set dollar amounts of the linear model, and the additional curvature of the log-log model does not seem to add to the statistical fit of the model; if anything, it seems to detract from the statistical significance of the model and the variables. The survey data also seems to reflect this judgment, that the semi-log model best represents the data, even though there are admittedly data problems in that model due to a lack of observations.

CHAPTER 9 - Discussion and Conclusions

Across all three data sets and all nine models discussed in this thesis, we can clearly see similarities in the statistically significant variables across time and functional form. In the logarithmic models, the intercept tended to be statistically significant. Only one of the linear models, for the fall data set, had a statistically significant intercept. Posting a photo ad of your horse clearly gives added exposure, in addition to the possibility that sellers using photo ads are already advertising their horses at higher prices, and using photo advertisements online appears to be worth the $25, $50, or $100 you might spend to purchase this ad, depending on the asking price of your horse. If you are selling a cheaply priced horse, the increase in price of 35-40% indicated by the semi-log functional forms would cover the cost of a $25 ad at a selling price of less than $100. A horse priced over $300 could see a price increase of over $100 by using a photo ad, indicating that a spotlight ad would be affordable if these magnitudes are correct. This is assuming that the price increase indicated by the *Pht* coefficient is solely due to the use of a photo ad. If, as suspected, horses using photo ads have higher initial values based on training level, discipline, or other factors, the gains seen by advertising using a photo ad may not be as great for every horse as indicated by the magnitudes in the models.

The states clearly had some economic significance as indicated by the F-tests and picked up by the logarithmic models, though mostly not in the linear models, indicating that the differences in horse prices between states are on a percentage basis, not a set dollar amount. Though the F-tests indicated that breeds should have some statistical significance, none of the models exhibited this in the breed variables nor was any detectable statistical significance consistent across time or functional forms. The breed category for saddle horses was the only one which consistently showed some level of statistical significance and was consistently positive. Perhaps a data set with less heterogeneity in the data or a larger sample size could more correctly capture the differences in prices associated with different breeds. Throughout the models, gender clearly had some economic impact, though sometimes the coefficient for female horses was statistically significant, sometimes uncut male horses, and sometimes both. However, it is safe to say that gender definitely has an impact on price, though it may depend on the location, time of year, and market you are looking to sell to. If there is a seasonal difference in demand for genders, it will take more years of data to accurately determine, but this data suggests that female horses are less valuable than male horses in the spring, and the opposite is true in the fall. Also, geldings seemed to be consistently more valuable than mares and uncut males at this end of the market.

Though color variables were occasionally statistically significant, the F-tests indicate that in this data set color did not have a statistical impact. It is possible that due to the heterogeneity of the data or of the importance placed on other factors such as age, temperament, and use of the horse, color does not play an economically significant role in the recreational market, even though Taylor et al. (2004) found some statistically significant color variables in the show Quarter Horse marketplace. A horse being registered definitely impacted the horse's price, in most cases having the greatest impact of any of the binary variables, and this economic significance was shown across time and functional forms. Horse temperament seemed to matter more in the fall than in the spring, though repeated trials would be necessary to indicate if this was actually true or just a data fluke in 2008. Finally, as expected, age clearly economically impacts the value of a horse, as it is an indicator for training and experience level as well as the remaining useful life. It increases at a decreasing rate, reaching a peak around 9-10 years of age as indicated by most models and expected by the industry, and decreasing after those years as useful life decreases, and did not show any differences across the seasons.

Though when the Box-Cox transformations they were all either biased or had different results from one another, it seems clear based on the model results that the semi-log functional form was the most appropriate for the data. As mentioned previously, this functional form did the best overall job of capturing statistical significance in the variables. This is consistent with previous literature, as almost all previous pricing models in the equine industry used the log-linear functional form, the only exception being Commer (1991) as noted in Buzby and Jessup (1994). The linear functional form was not clearly worse in statistics of overall fit, but did not always capture significance where it was expected. The linear models often resulted with similar results to the semi-logarithmic models on the photo ad, registration, and age variables, while usually showing more statistical significance in the gender variables and less in the state group. There did not seem to be many differences between the models across the data sets; the semi-log model did not appear to do a better job on the spring data than the fall data sets, and the linear models did not appear to do better on the fall data models than the spring data. Therefore, the indications given by the Box-Cox transformation in this case did not prove to be valuable.

Overall, the models developed and discussed here are in line with previous price evaluations on the equine industry as far as the types of variables examined and the functional forms used; only it is looking at a different and much more diverse segment. The use of three different functional forms provided a comparison which indicated that the state market differences are not linear but percentage differences, and that the value associated with gender may tend to be more linear. The horses studied here were extremely diverse, with different breeds, colors, ages, and genders, which differ from all previous hedonic pricing models; all previous models either focused on one breed, an age group within a year, a specific gender, or a combination of all three. The markets studied in previous literature are drastically different from that evaluated here being high end and comparatively homogenous segments. Therefore it is a major leap as an extension of previous literature. However, similar to previous literature, ages and genders were statistically significant in this model, and therefore likely have economic impacts. Differing breeds and states the horses were sold from had never been previously evaluated, and internet horse sales had never been studied. States appeared to have an economically significant impact on prices; though the F-tests indicated that breeds should have had statistical significance, none appeared consistently across the three data sets. Colors were evaluated in Taylor et al. (2004) and came up with some statistical significance, which was not found in this study in either the F-tests or the overall models with any consistency.

The information found in this study may be useful to numerous members of the equine industry, especially those using or interested in using online methods to advertise horses for sale. It is clear that registered horses bring higher prices than grade horses, therefore, registering foals when they are young and registration fees are low or finding a registration organization that will accept an older, non-pedigreed horses (such as certain performance or color organizations) may be valuable for individuals looking to sell horses from breeding or through horse trading. Any buyers or sellers of horses could use the semi-logarithmic functional form for the time of year they are planning to enter the equine marketplace. By applying the traits of the horse they are looking to buy or sell, they could find the best deals on horses to buy or find an appropriate price range in which to sell their horse. For individuals interested in buying and selling horses as a business, known as horse traders, the age, gender, and temperament significance may be of particular importance to their businesses in selecting horses that they can make the greatest profits on. They could also use the model to find horses that are underpriced based on their characteristics, which are likely horses they could make a profit by selling instantly. Individuals interested in advertising online may find it useful to post a low temperament score for their horse to get the horse more search exposure when prospective buyers are entering their search criteria; this is due to the fact that individuals looking to buy horses can conduct searches for temperaments and as horses with lower temperaments seem to be more valuable, they would likely be searched for more often. Finally, anyone interested in the horse market in general may be interested in studying these results and conducting further research in regards to the breeds, colors, and regional market differences across states and expanding to across the nation or internationally. These models provide a starting point for future research in the general low-end recreational horse market, as opposed to the high valued marketplaces of Thoroughbred racehorses and Quarter Horse show and race horses.

This starting point is only just that, a starting point. It provides a basic process with which to model data and establishes some expectations of economically significant variables, such as age and registration. However, due to possible collinearity issues or a lack of data, it does not establish differences in the market between breeds and does not clearly establish whether color is of interest or not. The high degree of heterogeneity in the data combined with a relatively low number of observations causes some of the results to be unclear and therefore the significance to be questionable. Additional information with a greater number of data points could provide a clearer picture of the significance, if any, of these variables. Alternatively, more closely studying individual breeds at the recreational market level, such as Quarter Horses, Paints, or Arabians, could establish differences in the market between breeds not indicated by this model. It is possible that colors would be statistically significant within a breed, but not across breeds, as different breeds commonly have different sets of coat colors genetically present in the breed. Clear examples of this is the Fresian horse, which is widely recognized for its black coat color, and the Lippizan breed, which are nearly always born black or dark brown and turn nearly white by maturity.

Information accuracy could be improved by the use of auction data at an open auction market, where any individual can sell their horse, and seller information can be verified. Comparisons could be made across sales through auctions versus the internet versus any number of other sources, if a route to data collection could be established. Alternatively, phone or e-mail surveys to collect further information on horse sales could be used to improve data accuracy.

Regional variations could be established through the use of show or rodeo circuits, which are often similar across organizations, as opposed to simply state regions as used by dreamhorse.com. There are many options available to further examine these markets for horses, if data sources can be found and utilized. Even internet horse sale sites could be utilized to further expand data sets, as it is not uncommon for large numbers of horses to be marketed and sold online, assuming that data accuracy could be achieved and a faster way to collect the data could be established, rather than individually recording all the records as was done with these data sets.

In studying economic literature on the equine industry it is quite clear that this is an internationally important, often over-looked industry that contributes several billion dollars to the worldwide economy annually. Gaps exist in equine economic literature throughout the different industry segments. Some high-end horse markets, like the ones evaluated in the literature review, keep relatively good data on prices and market trends, but no research or data collection has been done on a regular basis on the low-end, recreational horses in the industry; those that have a $32.0 billion dollar annual impact to the United States economy and keep some local communities alive (American Horse Council Foundation, 2005; Lu, 2003, James, 2003). Only two breeds are evaluated anywhere in economic literature, Thoroughbreds and Quarter Horses, and numerous other breed organizations may be interested in what traits affect the value of their animals. In addition, only a few nations including Great Britain, the United States, and Canadian markets have evaluated any horse market data. Numerous other countries that are vital to the international equine community could conduct research about the values of their horses and horse market. Many current major legislative issues, such as the National Animal Identification Program, horse slaughter, wild horse and burro legislation, and animal welfare, could be greatly clarified through increased market evaluation of the horse industry. Alone, the models evaluated in this thesis do not provide much information to help with these considerable equine industry issues; but hopefully, the objective of providing a starting point for further research on the recreational equine industry has been achieved.

References

2buyhorses.com. 2008. Accessed Jun. 2008. Online: www.2buyhorses.com.

Agdirect.com. 2008. *Horses for Sale, Stud Services*. Accessed Jun. 2008. Online: www.agdirect.com.

American Horse Council Foundation. 2005a. *The Economic Impact of the Horse Industry on the United States*. Deloitte Consulting LLP, United States.

American Horse Council Foundation. 2005b. *The Economic Impact of the Horse Industry on the United States: Technical Appendix*. Deloitte Consulting LLP, United States.

American Horse Council Foundation. 2008. *Legislative Issues and Policies*. American Horse Council Foundation, Washington, D.C. Accessed Mar. 2008. Online: http://horsecouncil.org/legislation.html.

American Quarter Horse Association. 2008. *2007 AQHA Annual Report*. Amarillo, TX.

Barrett, S. 2006. "Bloodstock – Neck and Neck." *Post Magazine*, 25th May, p. 19.

Beattie, B., T. Teegerstrom, J. Mortensen, and E. Monke. 2001. "A Partial Economic Impact Analysis of Arizona's Horse Industry." Dept. of Agri and Res Econ, University of Arizona, Tucson.

Beyer, A. 2006. "For Run of Fatal Breakdowns, No Easy Place to Fix Blame." *The Washington Post*, 2nd Aug., Sports pp. E03.

Booth, M. 2008. "No Room At The Pen." *The Denver Post*, 17th Feb., Sec. 1 pp. 1.

Buzby, J., and E. Jessup. 1994. "The relative impact of macroeconomic and yearling-specific variables in determining thoroughbred yearling price." *Applied Economics* 26:1-8.

Chezum, B., and B. Wimmer. 1997. "Roses or Lemons: Adverse Selection in the Market for Thoroughbred Yearlings." *The Review of Economics and Statistics* 19(3):521-526.

Chvosta, J., R. Rucker, and M. Watts. 2001. "Transaction Costs and Cattle Marketing: The Information Content of Seller-Provided Presale Data at Bull Auctions." *American Journal of Agricultural Economics* 83(2):281-301.

Dhuyvetter, K., T. Schroeder, D. Simms, R. Bolze Jr., and J. Geske. 1996. "Determinants of Purebred Beef Bull Price Differentials." *Journal of Agricultural and Resource Economics* 21(2):396-410.

Dreamhorse. *Horses for sale*. 2008. Accessed Mar. 2008. Online: www.dreamhorse.com.

Eddy, A., J. Elder, C. Egan, J. Schell, P. Bartley, J. Hannaford, and R. Wells. 2007. "All bets are off as $8 billion horse racing industry goes to the dogs; The flu that stopped a nation- The Equine Influenza Outbreak." *Sunday Age*, Melbourne, Australia, 26th Aug., pp. 4.

Einhorn, C. 2008. "Death Across the Border Awaits Horses Spared It in the U.S." *The New York Times*, 11th Jan., Sec. A, Col. 0.

Equine.com. 2008. Accessed Jun. 2008. Online: www.equine.com.

EquineHits.com. 2008. Accessed Jun. 2008. Online: www.equinehits.com.

Equine Now. *Horses for Sale*. 2008. Accessed Jun. 2008. Online: www.equinenow.com.

Espinosa, J., and B. Goodwin. 1991. "Hedonic Price Estimation for Kansas Wheat Characteristics." *Western Journal of Agricultural Economics* 16(1):72-85.

Faminow, M., and R. Gum. 1986. "Feeder Cattle Price Differentials in Arizona Auction Markets." *Western Journal of Agricultural Economics* 11(2):156-163.

Foster, G. 2007. "Aussie flu could hit Sportingbet; Market Report." *Daily Mail*, London, England, 29th Aug., pp. 68.

Freeman, D.W. 2007. *Oklahoma Horse Industry Trends*. Oklahoma Cooperative Extension Service, June. Accessed Jul. 2008. Online: http://osufacts.okstate.edu/docushare/dsweb/Get/Document-2087/CR-3987web.pdf.

Government Accountability Office. 2008. *Bureau of Land Management: Effective Long-Term Options Needed to Manage Unadoptable Wild Horses*. Washington, DC.

HorseClicks. 2008. Accessed Jun. 2008. Online: www.horseclicks.com.

Horsefinders.com. 2008. Accessed Jun. 2008. Online: www.horsefinders.com.

Horsesforsale.org. 2008. Accessed Oct. 2008. Online: www.horsesforsale.org.

Horsetopia.com. *Horses for Sale*. 2008. Accessed Jun. 2008. Online: www.horsetopia.com.

Horseville. 2008. Accessed Jun. 2008. Online: www.horseville.com.

James, B. 2003. "Equine Industry Creates a Conundrum." *St. Petersburg Times*, Florida, 7th Apr., Sec. Citrus Times, pp. 1.

Kolodinsky, J., T.P. DeSisto, and Q. Wang. 2003. "Willingness-to-pay for GE-free Products: A Hedonic Model for Milk with the rBST-free Characteristic." *Consumer Interests Annual* 49:1-9.

Lancaster, K. J. 1966. "A New Approach to Consumer Theory." *The Journal of Political Economy* 74(2):132-157.

Lansford, N. H. Jr., D. W. Freeman, D. R. Topliff, and O. L. Walker. 1998. "Hedonic Pricing of Race-Bred Yearling Quarter Horses Produced by Quarter Horse Sires and Dams." *Journal of Agribusiness* 16(2):169-185.

Lawrence, R. G. 1970. "An Econometric Analysis of Changes in Quarter Horse Prices, 1950 through 1967." PhD dissertation, Texas A&M University, College Station.

Lu, S. 2003. "Tough Times for the State Animal." *The New York Times*, 23rd Feb., Sec. 14NJ Col. 1.

Marrs, D., and C. Mathews. 2003. "In Brief – Company News." *Business Day*, South Africa, 12th Dec., Economy, Business & Finance, pp. 8.

Maugeri, M. 2007. "Flu to slow racing stocks." *Herald Sun*, Australia, 27th Aug., pp. 29.

McLennan, L. 2007. "Horse flu, poll butt… Retail Strong for DJ's." *The Courier Mail*, Australia, 1st Oct., Finance, pp. 33.

Mintert, J., J. Blair, T. Schroeder, and F. Brazle. 1990. "Analysis of Factors Affecting Cow Auction Price Differentials." *Southern Journal of Agricultural Economics,* Dec.: pp. 23-30.

MyHorseForSale.com. 2008. Accessed Oct. 2008. Online: www.myhorseforsale.com.

Ndiva Mongoh, M., R. Hearne, N.W. Dyer, M.L. Khaitsa. 2008. "The economic impact of West Nile virus infection in horses in the North Dakota equine industry in 2002." *Tropical Animal Health and Production* 40:69-76.

Neibergs, J.S. 2001. "A Hedonic Price Analysis of Thoroughbred Broodmare Characteristics." *Agribusiness* 17(2):299-314.

Neibergs, J.S, and R. Thalheimer. 1997. "Price Expectations and Supply Response in the Thoroughbred Yearling Market." *Journal of Agricultural and Applied Economics* 29(2):419-435.

North, M. S., D. Bailey, and R. A. Ward. 2005. "The Potential Impact of a Proposed Ban on the Sale of U.S. Horses for Slaughter and Human Consumption." *Journal of Agribusiness* 23(1):1-17.

Parsons, C., and I. Smith. 2008. "The Price of Thoroughbred Yearlings in Britain." *Journal of Sports Economics* 9(1):43-66.

Robbins, M. and P. E. Kennedy. 2001. "Buyer Behaviour in a Regional Thoroughbred Yearling Market."*Applied Economics* 33:969-977.

Stoeppel, K. M., and L. J. Maynard. 2006. "Hedonic Price Analysis of Thoroughbred Broodmares in Foal." Working paper, Dept. of Agri. Econ., University of Kentucky.

Taylor, M. R., K.C. Dhuyvetter, T. L. Kastens, M. Douthit, and T. L. Marsh. 2006. "Show Quality Quarter Horse Auctions: Price Determinants and Buy-Back Prices." *Journal of Agricultural and Resource Economics* 31(3):595-615.

Trejos, N. 2005. "Making Hay in a Horse-Based Economy; Recreational Riders Spur Growth on Md. Farms, Other Businesses." *The Washington Post*, 13th Jun., Metro pp. B01.

Unwanted Horse Coalition. *The Issue*. 2008. Accessed Jun. 2008. Online: http://www.unwantedhorsecoalition.org/?id=2.

U.S. Department of Agriculture. 1999. *Equine*. USDA-NASS. Accessed Feb. 2008. Online: http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1466.

U.S. Department of Agriculture. 2002. *The Census of Agriculture*. USDA-NASS. Washington DC, Accessed Feb. 2008. Online: http://www.agcensus.usda.gov/Publications/2002/index.asp.

U.S. Department of Agriculture. 2006a. *Baseline Reference of Equine Health Management Strategies at Equine Events in Six States, 2005*. USDA-APHIS-VS, CEAH, National Animal Health Monitoring System. Fort Collins, CO #N456.0607.

U.S. Department of Agriculture. 2006b. *Equine 2005, Part I: Baseline Reference of Equine Health and Management, 2005*. USDA:APHIS:VS, CEAH. Fort Collins, CO #N451-1006.

U.S. Department of Agriculture. 2006c. *Equine 2005, Part II: Changes in the U.S. Equine Industry, 1998-2005*. USDA-APHIS-VS, CEAH. Fort Collins, CO #N452-0307.

U.S. Department of Agriculture. 2008. *A Business Plan to Advance Animal Disease Traceability, Version 1.0*. Washington, DC, June.

Vickner, S. S., and S. I. Koch. 2001. "Hedonic Pricing, Information, and the Market for Thoroughbred Yearlings." *Journal of Agribusiness* 19(2):173-189.

Wainwright, R., and M. Moore. 2007. "The Cough That Stopped A Nation." *Sydney Morning Herald*, Australia, 3rd Nov., Sec. News and Features pp. 25.

Wikipedia, the free encyclopedia. 2008. Accessed Mar. 2008. Online: http://en.wikipedia.org/wiki/Main_Page.

Witte, A., H. Sumka and H. Erekson. 1979. "An Estimate of a Structural Hedonic Price Model of the Housing Market: An Application of Rosen's Theory of Implicit Markets." *Econometrica* 47(5):1151-1173.

Appendix A - Dreamhorse.com Price Survey

Letter to dreamhorse.com sellers for Fall Survey Data:

Hi, my name is Jenny Freeborn, and I am a graduate student in agricultural economics at Kansas State University. I am doing a pricing model on horses sold on dreamhorse.com for my master's thesis, and would appreciate your help by answering the following question on the horse you recently sold: What was the actual price the horse sold for if different than listed in the ad? If it was the same as the ad please tell me that.

Please know that I am not asking for or using any personal or specific information on you or your horse, it will simply be aggregated into a data set. Feel free to contact me if you have any concerns. Thank you so much for your time and help with my thesis!

Sincerely,

Jenny Freeborn