2023

Gambling on Growth: An Analysis of the Early Impact of Historical Horse Racing on Kentucky's Thoroughbred Industry

Barrett W. Kerr
University of Kentucky, bwkerr.2023@gmail.com
Author ORCID Identifier:
https://orcid.org/0009-0005-1107-6291
Digital Object Identifier: https://doi.org/10.13023/etd.2023.068

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Barrett W. Kerr, Student

Dr. C. Jill Stowe, Major Professor

Dr. Tyler Mark, Director of Graduate Studies

Gambling on Growth: An Analysis of the Early Impact of Historical Horse Racing on Kentucky's Thoroughbred Industry

---

THESIS

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in the College of Agriculture, Food and Environment at the University of Kentucky

By
Barrett W. Kerr
Lexington, Kentucky
Director: Dr. C. Jill Stowe, Professor of Agricultural Economics
Lexington, Kentucky
2023

https://orcid.org/0009-0005-1107-6291

ABSTRACT OF THESIS

Gambling on Growth: An Analysis of the Early
Impact of Historical Horse Racing on Kentucky's
Thoroughbred Industry

With a continuous decline in on-track wagering on Thoroughbred racing, racetrack operators are exploring alternative revenue sources to fund purses and create better quality racing. Historical Horse Racing (HHR), which represents one of the most recent advancements in parimutuel wagering, has become a popular system used by many states to offset the live wagering decline. Proponents of HHR suggest that this alternative wagering option will benefit the live racing industry. Using data from Kentucky racetracks from 2002 – 2021, this study examines the relationship between HHR wagering and key industry metrics, such as purses, on-track wagering, sales prices, and foal crop. Both OLS and staggered difference-in-differences models are used to explore the initial effects of HHR on Thoroughbred racing in Kentucky. While HHR was introduced in Kentucky fairly recently, preliminary analysis suggests that HHR enhances purses but, to this point, has not influenced other key metrics.

KEYWORDS: Thoroughbred Racing, Parimutuel Wagering, Historical Horse Racing, Difference-in-Differences

Barrett W. Kerr
(Name of Student)

04/25/2023
Date

Gambling on Growth: An Analysis of the Early Impact of Historical Horse Racing on Kentucky's Thoroughbred Industry

By

Barrett W. Kerr

C. Jill Stowe

Director of Thesis

Tyler Mark

Director of Graduate Studies

04/25/2023

Date

ACKNOWLEDGMENTS

I would like to start by thanking my thesis advisor, Dr. Jill Stowe, for this amazing research opportunity and the support she has given me. She has continued to believe in me, even when I didn't believe in myself, and pushed me to challenge myself in order to achieve my goals and aspirations. I would also like to thank my thesis committee, Dr. Alison Davis, Dr. Steven Buck, Dr. Nicholas Pates, and Laura Holoubek, J.D. This project would not have been possible without their continued support and the critiques they provided.

I would also like to thank my academic advisor, Dr. Roger Brown, for his continuous support in my studies and endeavors. Without his support, I would not be receiving either of my degrees, nor would I be able to attend law school this upcoming fall. I also owe a great deal of thanks to Dr. Tyler Mark who has supported me in my completion of the University Scholars Program. His support for myself and the rest of the graduate students does not go unnoticed, and our success would be nothing without him. Beyond that, I would like to thank the entire Agricultural Economics Department at the University of Kentucky. The department has become like a second family to me, and I know their doors are always open if I were to ever need help.

To my advisors and peers at the Stuckert Career Center, none of my success would be possible without the support and guidance each of you has given me. The time each of you has put into my success is greatly appreciated in ways words cannot describe.

To my friends and classmates, Amelia Floyd, Mason Hamilton, Olivia Gibson, Emily Hehn, Clare Martin, Brinley Gothard, and others not mentioned, thank you for standing by me and helping me navigate these past four years of my schooling. Each one of you has challenged me to grow as a person and supported me both inside and outside of the classroom, and for that, I am so extremely grateful. To the McKenzie Family: Hilary, Taylor, Mya, and Skylar, thank you for being my home away from home and for all the kindness you have showered me with over the years. Day or night, you and your family have always been there to aid me in ways I will not ever be able to repay, and I know you will continue to be in my life no matter where I may be.

Lastly, I would like to thank my family. To my parents, Chris and Tanya, thank you for helping to raise me into the man I am today and for encouraging me to chase my dreams and find new adventures. To my sister, Isabella, thank you for showing me what it looks like to find motivation and be strong even when things get hard and I feel like giving up. To my brother, Marshall, thank you for showing me how to continue to think outside the box and try new things, even if they are outside of my comfort zone. To my great-grandparents, my grandparents, my aunts, uncles, and cousins, thank you all for being the backbone of support I have continued to rely on for so many years and I hope I continue to make each of your proud as I move into my next stages of life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

1.1 Kentucky's Thoroughbred Industry

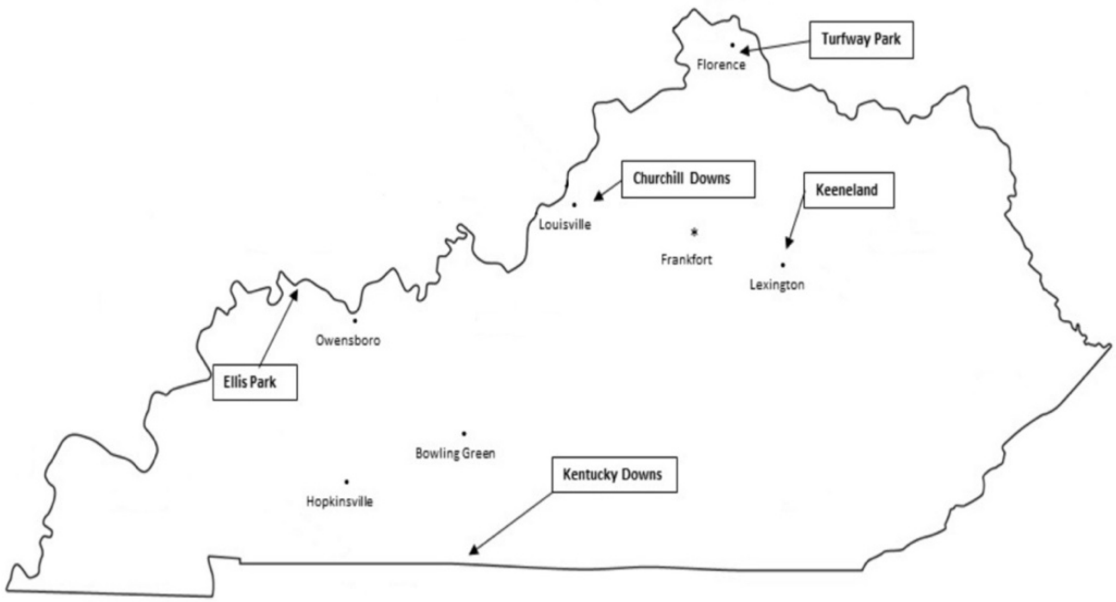
Kentucky's Thoroughbred racing industry is buried deep within the history of the state. Many historians attribute the first horses introduced to Kentucky to be those brought here by Daniel Boone in 1769, which were eventually stolen by Native Americans (International Museum of the Horse). However, more horses continued to move into the region as settlers arrived from the east to settle the new land. Eventually, the first circular track in the state was developed in Lincoln County, in which horses raced in a counterclockwise direction to oppose the established English method of horse racing, which was clockwise (Pendleton, 2023).

In 1789, Kentucky was credited with having its first Thoroughbred race. In October of 1789, the first long-distance race was run in Lexington, Kentucky; Thoroughbred horses were more suitable for competing over these longer distances. A stallion named Blaze is credited with being the first stallion to stand at stud in Kentucky, launching the foundation for what was to become the "Thoroughbred Nursery of the World" and establishing of one the state's signature industries.

On May 17th, 1875, the horse racing industry in Kentucky would forever be changed. On this date, the inaugural running of the Kentucky Derby took place, helping to change the course of Kentucky's Thoroughbred racing (Nicholson, 2012). The race was held at a facility known as the Louisville Jockey Club and Driving Park, which today is better known as Churchill Downs. From this historic event, Kentucky's racing industry boomed, helping to open tracks such as Ellis Park in Henderson (1922), Keeneland in Lexington (1936), Turfway Park in Florence (1959), and Kentucky Downs in Franklin

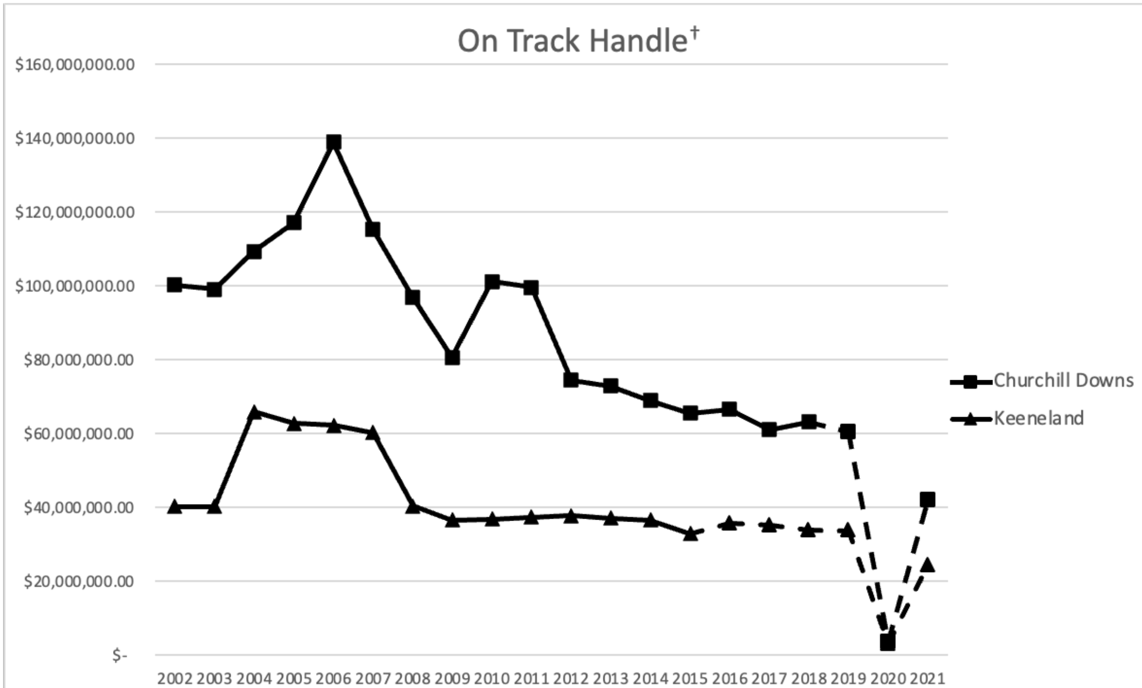
(1990) (Figure 1.1). Since then, the racing and closely related breeding industries contribute significantly to Kentucky's economy, with nearly $2 billion in additional revenue generated and over 22,000 jobs (*2012 Kentucky Equine Survey*, 2013).

Figure 1.1 Map of Thoroughbred Racetracks (Adapted from KHRC Biennial Report 2019 – 2020)



Although the Thoroughbred racing industry continues to be highly visible in KY, its long-run sustainability may be in question. In the past 30 years, on-track handle, which represents dollars wagered on-site on live racing, and the number of race days have decreased on a year-to-year basis. Purses, which is the prize money that is returned to the owner, trainer, and jockey of the winning horses, have increased only marginally over time. This decline in live wagering (on-track handle) is especially problematic. Race purses are primarily funded by on-track handle, so the decline in wagering has made it difficult for racetracks to offer attractive prize money. Figures 1.2 and 1.3 illustrate the trends in on-track handle and purses at the five Kentucky racetracks over the past 20 years. With the decline in expected returns to racing, the number of foals born annually decreased across both Kentucky and the nation. Given these downward trends, the racing industry is looking for alternative long-term solutions to counteract these patterns and add stability back to the industry.

Figure 1.2a Comparison of On-Track Handle between Churchill Downs and Keeneland



†The solid and dashed lines illustrate the trend pre- and post-HHR introduction at that racetrack, respectively.

Figure 1.2b Comparison of On-Track Handle between Ellis Park, Kentucky Downs, and Turfway

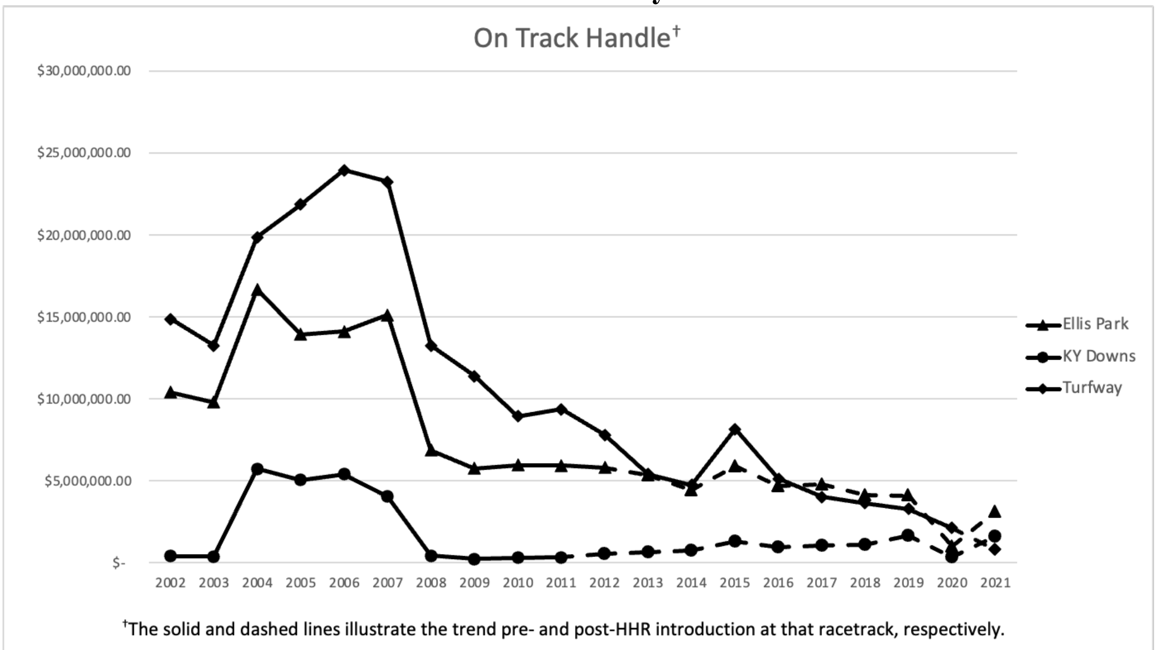


Figure 1.3a Comparison of Purses between Churchill Downs and Keeneland

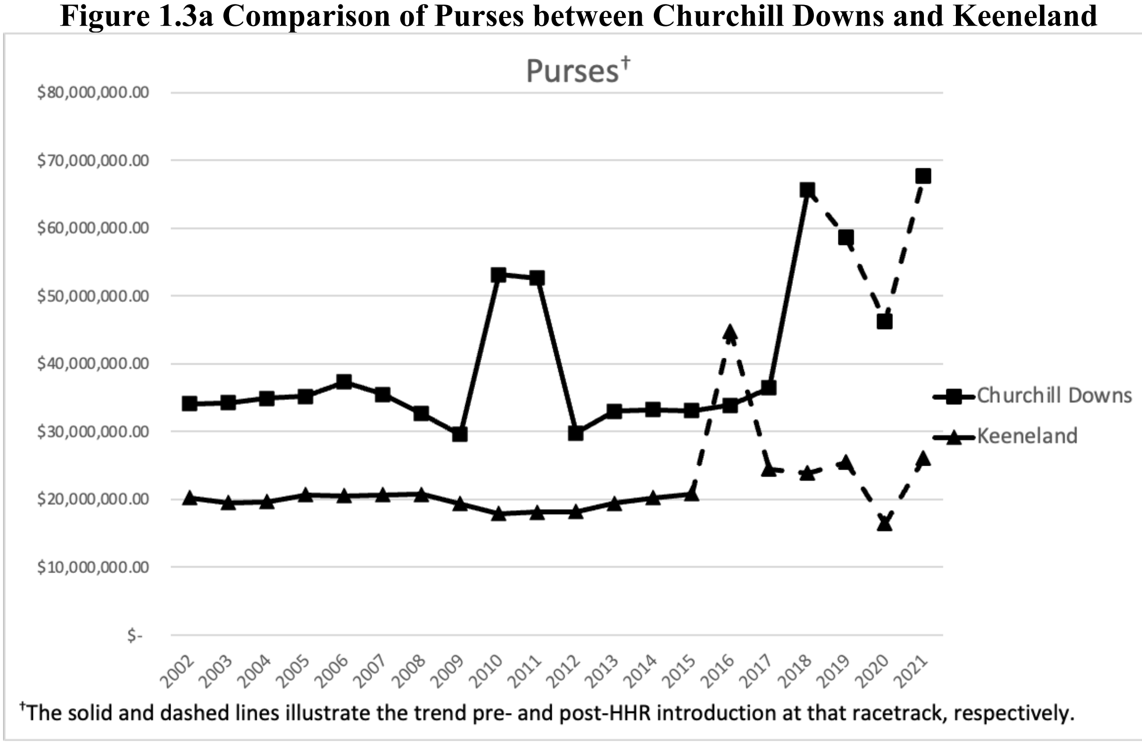
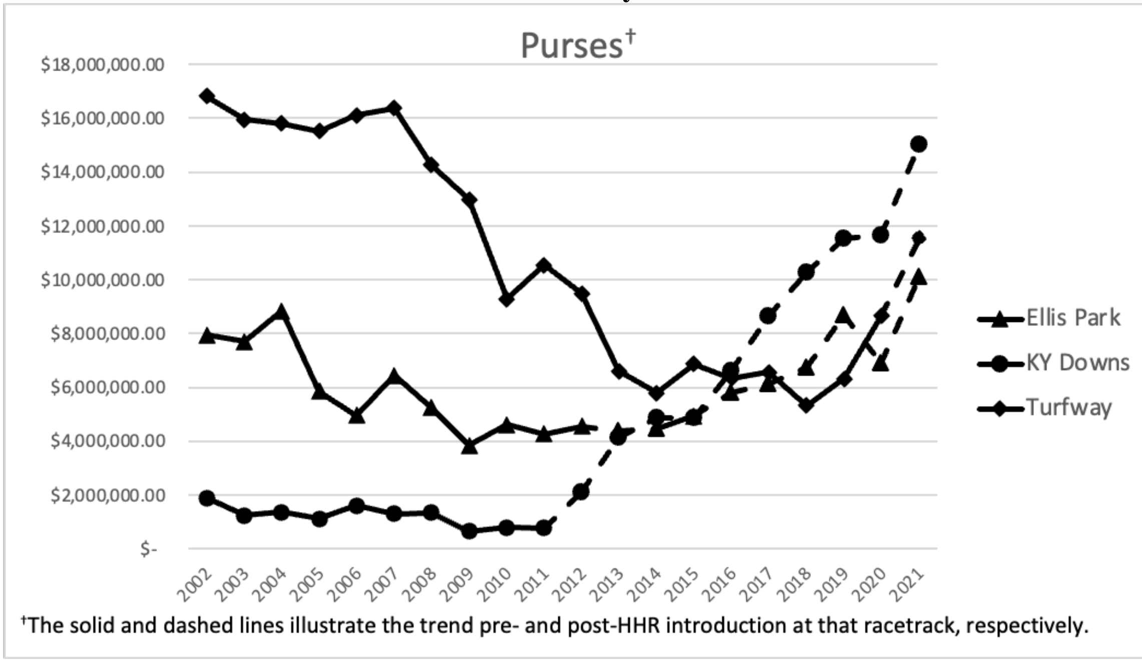


Figure 1.3b Comparison of Purses between Ellis Park, Kentucky Downs, and Turfway



1.2 Historical Horse Racing

To counteract downward trends in the racing industry, other states introduced alternative forms of gambling at racetracks. A portion of the proceeds from these sources were used to enhance prize money for racing, leaving Kentucky racetracks at a competitive disadvantage. Given the pressure from nearby states, such as Pennsylvania and West Virginia, Kentucky racetracks began expanding their gambling options by including historical horse racing. Historical horse racing (HHR), also known as instant racing, is a machine-based betting system that allows players to place bets on replays of past horse races. The earliest form of HHR appeared at Oaklawn Park in Hot Springs, Arkansas, in the 1990s and was modeled after slot machines to attract bettors to race wagering. Modern-day HHR machines often still look similar to slot-machines but can vary greatly in their appearance and theme. However, despite these visual differences, the machines are all based on actual randomized past races. Sometimes, the odds of these races are known to the players, and they can make bets accordingly. Other times, there is minimal indication that the machine is related to horse racing, and players simply press a button to begin gameplay (Kobin, 2021).

The introduction of these machines has not been without controversy, as their legality has been debated by lawmakers, industry leaders, and the public since they were first introduced to the state. Although the first machines were introduced at Kentucky Downs in September 2011, the Kentucky Horse Racing Commission (KHRC) began working to legalize instant racing machines in 2010 by modifying the definition of parimutuel wagering. This decision faced major backlash, and anti-gambling activists took the decision to court where it eventually reached the Kentucky Supreme Court in 2014. Their ruling, that the KHRC could legalize parimutuel wagering on historical races but sent the case back to another court to decide if HHR met the classification for parimutuel wagering, was eventually appealed multiple times between the lower courts and the Kentucky Supreme Court. Eventually, the Kentucky Supreme Court in 2020 found that HHR was deemed to not be considered parimutuel wagering, seeming to end the surge of instant racing. However, in 2021 state legislators in Kentucky amended the definition of parimutuel wagering to include wagering on HHR. This decision paved the way for HHR to continue, which industry supporters applauded because of its importance in supplementing income for the racing industry.

With the legalization of HHR, new rules and regulations were implemented to govern who could build these facilities, where they could be built, and what their association with racing would be. Since these facilities' affiliation with racetracks was key to the success of defining these machines as parimutuel wagering, each racetrack was allowed to operate two HHR facilities. KRS 230.380, Section 1 states that all tracks licensed by the racing commission are allowed to build a simulcast facility after applying and being approved by the state. A simulcast facility is a facility where individuals can eat, drink, and watch simulcast races, as well as wager on HHR machines, in a single building (Kruse, 2002).

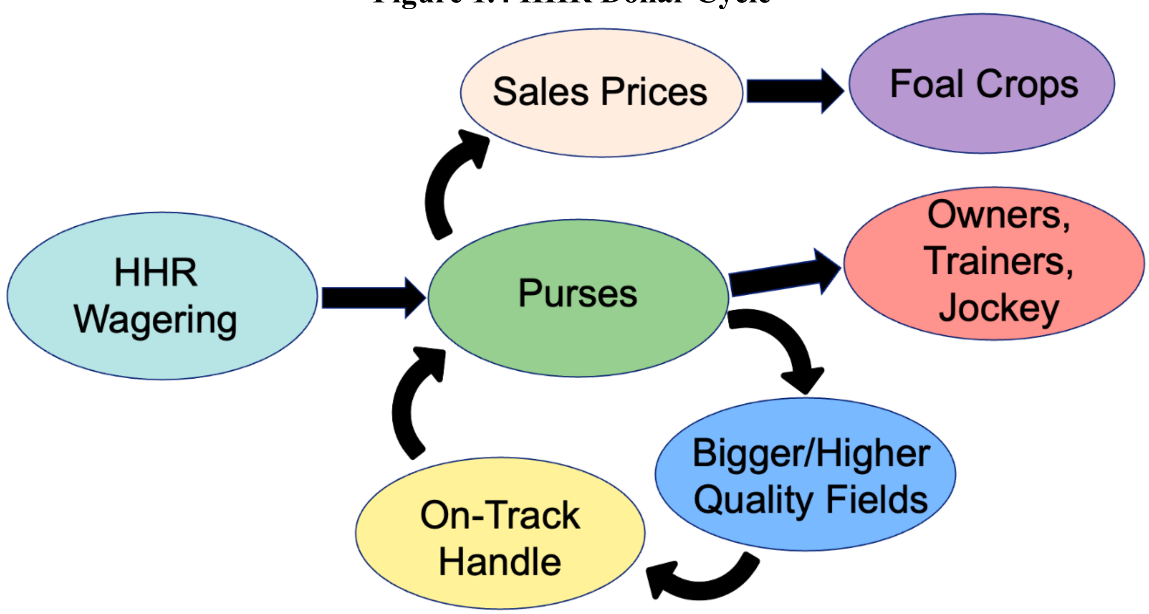
Because the term “racetrack” might be ambiguous, KRS 230.210 clarifies the term by classifying a track according to one of three descriptions:

1. For facilities operating as of 2010, the location and physical plant described in the “Commonwealth of Kentucky Initial/Renewal Application for License to Conduct Live Horse Racing, Simulcasting, and Pari-Mutuel Wagering,” filed for racing to be conducted in 2010.
2. Real property of an association, if the association received or receives approval from the racing commission after 2010 for a location at which live racing is to be conducted.
3. One (1) facility or real property that is:
    a. Owned, leased, or purchased by an association within a sixty (60) mile radius of the association’s racetrack but not contiguous to racetrack premises, upon racing commission approval; and
    b. Not within a sixty (60) mile radius of another licensed track premise where live racing is conducted and not within a forty (40) mile radius of a simulcast facility, unless any affected track or

simulcast facility agrees in writing to permit a noncontiguous facility within the protected geographic area.

It is this third description that has granted racetracks, such as Churchill Downs, Keeneland, and Kentucky Downs, to build a second facility housing HHR terminals and having no physical ties to a racetrack. Thus, new standalone simulcast facilities are appearing in the state, with facilities like The Mint – Corbin being opened in the past year. Additionally, although there is no statutory mandate, there exists an informal agreement that a percentage of the monies wagered on HHR will supplement purses at each track operating HHR machines, which enhances Kentucky's live racing industry. However, as on-track handle continues to decrease and new HHR facilities like The Mint have started to appear, some are questioning whether the introduction of HHR is truly benefiting the horse industry.

Figure 1.4 HHR Dollar Cycle



It is possible that HHR could be beneficial for the horse racing industry as a whole. Primarily, HHR wagering, if used in ways similar to other states, may be used to enhance race purses. Higher purses benefit the industry in several ways. First, higher purses mean that owners, trainers, and jockeys have higher earnings potential. Additionally, this could result in trickledown effects to supporting businesses such as feed, fencing, transportation, etc. This supports the idea that enhancing unrestricted purses is the most effective use of incentive funds for the entirety of the industry (Neibergs and Thalheimer, 1999). Second, higher purses should attract more horses and better-quality horses in each race, leading to better-quality races. In turn, better quality races with larger fields should attract more bettors, helping to increase the on-track handle, which in turn increases future purses. Additionally, the increase in purses increases a horse's earning potential, which places upward pressure on sales prices. Finally, higher sales prices may influence production decisions for breeders (Karungu et al., 1993). Therefore, HHR wagering may benefit the racing industry when a portion of wagering revenue is allocated back to the industry.

1.3 Research Questions and Objectives

The current debate regarding the impact of HHR on the Thoroughbred racing industry serves as the primary motivation of this study. The purpose of this study is to investigate the preliminary effect that the introduction of HHR has had on the Thoroughbred racing industry in the state of Kentucky. The results from this study will inform industry professionals and lawmakers by quantifying the relationship between HHR and established industry metrics, such as on-track handle, purses, sales prices, and foal crops.

The relationship between HHR and industry metrics is investigated using two approaches. Simple correlations between HHR and industry metrics will be explored using ordinary least squares (OLS) regression models. This first step helps us understand whether HHR has any statistically significant relationships with standard industry metrics but does not suggest any causal relationship.

The second part of the study attempts to quantify the effect that the introduction of HHR has on purses and on-track handle at Kentucky Thoroughbred racetracks using a staggered difference-in-differences analysis. By comparing treated groups in a specific time period to untreated groups, the difference-in-differences estimate will quantify the initial impact of HHR for an already established racetrack [and what effect HHR growth has] on the success of live racing.

1.4 Thesis Structure

Chapter 2 provides the literature review of related background papers. Chapter 3 outlines the data and its sources, as well as the empirical methods used to complete the ordinary least squares regressions and the staggered difference-in-differences analysis. Results for both analyses are presented in Chapter 4. Lastly, Chapter 5 addresses the conclusions and implications of the study.

CHAPTER 2. LITERATURE REVIEW

2.1 The Importance of Wagering in Thoroughbred Racing

Following a flurry of research on betting on horse racing in the 1970s, 80s, and 90s, the topic has since received little academic attention. However, the research that has been conducted sets up an interesting narrative to allow for current research into HHR to take place. Recent research suggests that the equine industry, in the years before HHR's introduction, was considered by many to be stagnant and outdated, not adapting to current markets and interests that surround it (Lambert, 2022). Meeker (1989) introduces these concerns and describes new attitudes that have led lawmakers to adopt new gambling practices such as HHR (1989).

To compensate for lower attendance across all racetracks, tracks must offer attendees new betting options that have higher payouts. With the advent of the state lottery, racetracks had to appeal to attendees' changing risk preferences towards higher payoffs, leading to new on-track gambling combinations (called exotics), such as exactas, trifectas, the pick six, and the pick three (Meeker, 1989). However, Meeker posits that any alternative wagering systems, such as HHR, must be controlled by the racetracks in order for them to benefit the industry.

This idea of racetracks controlling wagering systems such as HHR has become a common theme across the Thoroughbred industry. Researchers conducting an economic analysis of the entire racing industry in California were able to quantify how state control of wagering systems can be important to the industry's survival (Carter et al., 1991). The state determines the takeout rate[1], number of race days, tax rate, and other regulatory decisions that affect the racing industry. With the Thoroughbred industry contributing almost $1 billion, both directly and indirectly, to California's state economy, it is evident that the industry as a whole is of economic importance as a state revenue source. Additionally, with $2.3 billion being wagered on Thoroughbred races at the time of the study in 1989, the tax revenue for the state was enough to generate over $150 million for local governments (Carter et al., 1991). However, if takeout rates are set too high, the price of wagering for bettors increases, and betting across all wagering systems decreases, resulting in less revenue for the state. Therefore, Carter et al. (1991) suggest that the takeout rate, which is determined by the state, plays a large role in the success of statewide racing wagering systems and should be regulated in a way that benefits both the state and the racing industry.

Carter et al. (1991) also raises one of the strongest arguments in support of HHR gambling years before it ever became a point of conversation within the industry. Since race purses, breeder incentive awards, and the sale of horses represent the three primary sources of revenue for the racehorse sector, these three items must be prioritized in order for racing to survive. Specifically, purses, which are controlled by the racetracks and racing associations, need to be heavily prioritized in order to keep races attractive. Higher purses have been shown to lead to higher quality races, which in turn should lead to more dollars wagered on live racing (Carter et al., 1991). Thus, the need to increase purses suggests that systems like HHR could help the industry survive if on-track wagering continues to decline.

[1] The takeout rate is defined as the percentage of every wagering dollar that is extracted before payouts are determined. The money that is taken out is typically used to pay taxes or help fund racing operations.

2.2 Gambling Options and Effects

Wagering is an important component of horse racing. A percentage of each dollar wagered is used to fund purses, which attract horses and bettors alike. And while there are many sources of wagering, live on-track wagering is the largest contributor to purses. Decreasing attendance at the races has led to lower on-track wagering. Without another source of revenue, this puts downward pressure on purses. And, with more legal gambling options, there has been an increase in the competition for the gambling dollar. As more forms of wagering become available, the gambling industry, which is supported by a substantial majority of Americans, continues to evolve to fit the preferences of the bettors (Bishop, 2000).

Suits (1979) investigates the elasticity of demand for on-track wagering in Nevada. Results from this study show that demand for live on-track wagering is positively correlated with the number of race days. As the number of race days has fallen, so too has demand for on-track wagering. This lack of demand helps to explain one of the overarching reasons for the decrease in live racing on-track handle. Additionally, demand for on-track wagering is relatively elastic; an increase in the price of wagering, as measured by track takeout, has a proportionally larger negative effect on live racing handle. Therefore, raising taxes on live wagering or tracks taking a larger cut of revenues will not supply the industry with more funding for purses and operations, meaning other forms of wagering are likely the better solution (Morgan and Vasche, 1979).

Simmons and Sharp (1987) looked at the impact of state lotteries on the Thoroughbred racing industry. The authors identified variables potentially influencing average daily handle, including the price of wagering, seating capacity, number of stakes races held, market population, per capita income, poverty level, unemployment rate, competition from casinos, and competition from lotteries. All explanatory variables were significant except for poverty level and competition from casinos. More recent research, however, suggests that competition from casinos negatively impacts on-track wagering, which is discussed in a later section. Similar to Suits (1979), the price of wagering and the number of races are major determinants of on-track handle.

Another issue of interest that this study examines is the relationship between on-track wagering and lotteries, which is important to consider since Kentucky has a state lottery system. According to Simmons and Sharp (1987), the state lottery has a negative effect on the live racing handle for Thoroughbred races. Race meets held in states without lotteries resulted in an on-track handle that was $324,000 greater than meets run in states with lotteries, which was conditioned on the population of the state and the number of seats at the racetrack. It is important to understand that other wagering systems, which have risen in popularity throughout Kentucky in the past few years, have offered bettors substitutes for on-track wagering. Thus, if bettors view HHR as a substitute to on-track wagering the same way they view the lottery, this could exacerbate the continuous decrease in on-track handle.

2.3 Bettor Behavior

One strand of research focused on bettor preferences and behavior related to wagering on horse races. The number of races across North America decreased by over 20 percent between 1985 and 2002. Citing earlier studies which found the price of a wager (the track takeout) is elastic, Gramm, McKinney, Owens, and Ryan (2007) provide evidence that there are two types of bettors, informed and uninformed, that wager on horse racing. For their study, the informed bettor was defined as the information-seeking, risk-averse individual whose demand for wagering on races is dependent on both returns and the quantity/quality of the information available about a particular race. The uniformed bettor was defined as an individual who is risk-loving or risk-neutral, who has high costs for gathering information, and is unresponsive to reduced information about a race. Based on an analysis of 2,957 races at all major racetracks in the fall of 2002, they found that higher-quality races and larger field sizes resulted in more wagering activity.

Additionally, their study further found evidence of informed bettors in horse racing who choose to place bets according to both risk and return. This result confirms the findings of two earlier studies (Thaler and Ziemba (1988) and Ali (1977)). Ali (1977) characterized bettors' behavior to explain the observed subjective-objective probability relationship using subjective and estimated objective winning probabilities from 20,247 harness horse races. The motivation for this study came from laboratory experiments which suggested that when making a decision under uncertain conditions, low probability events are overbet and high probability events are underbet, which Ali confirmed using wagering data. Moreover, the study suggested that the more capital the bettor has, the less he tends to be risk loving.

Thaler and Ziemba (1988) also found evidence that bettors overestimate longshots and underestimate favorites (called "the favorite-longshot bias"), but also contributed to the existing research by identifying market efficiencies and anomalies. Their research was focused on categorizing racetracks and lotteries as anomalies in economic markets. The authors categorized the racetrack betting market as a market that convenes for roughly 20-30 minutes where bettors wager on any number of horses that are in the field for the upcoming race. They assumed that all bettors in this market are expected value maximizers with rational expectations. This gave the following two market efficiency conditions:

1.) No bets should have positive expected values (weak)
2.) All bets should have expected values equal to (1 - t) times the amount bet (strong)

where t is the transaction cost percent, or the fee for gambling. They concluded that even though racetracks appear to be efficient in their gambling practices, both of these principles are actually violated. This violation mainly arises from the favorite-longshot bias. This bias, which has been heavily studied in other papers examining the efficiency of parimutuel betting markets, suggests that Thoroughbred betting markets continue to inefficiently price outcomes (Gramm and Owens, 2006). The favorite-longshot bias is based on the idea that expected returns per one-dollar bet increase monotonically with the probability of the horse winning. Thus, favorites are expected to win, and longshots are not expected to win, with bets on longshot horses often having a negative expected value.

That bettors make wagers that are known to have negative expected value implies that they must be "logically" risk-seeking. Furthermore, Thaler and Ziemba believe that fans ultimately go to racetracks to bet, and watching races is an entertaining way to support their real interest in gambling. Thus, beyond the potential for financial returns, bettors also earn some utility from the racing experience. Ultimately, they conclude that modeling gambling behavior is complicated and relies on numerous factors that are hard to track, such as success in earlier races (Thaler and Ziemba, 1988).

It is important to understand the behaviors and preferences of bettors on-track to determine if these same relationships will exist for HHR wagering. As time goes on, it will be important to understand the behaviors of bettors in HHR and simulcast facilities to see if these bettors operate in two completely different markets, or if their attitudes and behaviors are suitable for both markets. This may also help determine the extent to which on-track wagering and HHR are viewed as substitutes.

2.4 Historical Horse Racing

As the number of races, size of purses, attendance, and on-track handle have continued to decline nationwide, states with racing have started considering systems like simulcasting and HHR as ways to save the sport. Rudd, Mills, and Flanegin (2009) claim that the addition of video machines provides an additional form of entertainment for the tracks, offering an enticing form of wagering that could lead to a long-term appeal to horse racing. Between 1995 and 1999, live handle at racetracks without video machines declined by almost 38%, whereas those with a so-called "racinos" reported an increase in handle of 7%.

However, they also suggest that these machines are not without their own issues. Ultimately, the success of video terminals comes down to the preferences of the bettor. If these terminals are seen as a complement good to on-track wagering, the machines could potentially add new bettors to live racing. On the other hand, if bettors see the machines as a substitute good, the racetracks may use these machines as their main source of revenue and limit the promotion and growth of their core business, horse racing.

This concern was further explained by Thalheimer (2012), who defined the interrelationship of the demand for slot machine and horse race wagering at new racinos.

Using Prairie Meadows Racetrack and Casino (1993-2006) in Iowa, Thalheimer studied the effects of video machine wagering, which was introduced in 1995. He found that video machine wagering demand was positively related to parimutuel race wagering at the racino, but not vice versa. More specifically, in the presence of live races, video machine wagering increased by 13% for a typical five-day week. However, the introduction of video machine wagering had the opposite effect on live racing handle, decreasing handle by 22% on live races. The two variables that increased parimutuel wagering on live races were an increase in race days and an increase in simulcast coverage. This confirmed earlier studies (Thalheimer, (1998), Thalheimer (2008)) showing that live racing benefited casino-style games, but the introduction of new games did not increase live racing handle.

This lack of increase in on-track handle may not deter more video machines from being introduced to the racing industry. As seen by studies conducted in New Jersey, New York, and Pennsylvania, which allow video machine wagering, the effects have been too large to justify the removal of video wagering. For example, over a five-year period in Pennsylvania, the addition of video terminals increased purses by 100%, taking purses from an average of $120,000 per day to over $240,000 per day (Malinowski and Avenatti, 2009). Based on these findings, and other findings supporting the increase of the simulcast handle, racing days, and eventual increases in live racing handle, the authors conclude that parimutuel wagering alone can no longer support or ensure the sustainability of the racing industry. The ever-growing costs of maintenance and promotion to introduce racing to a new generation must be supported by large-scale revenues, like the ones brought in by video machine wagering.

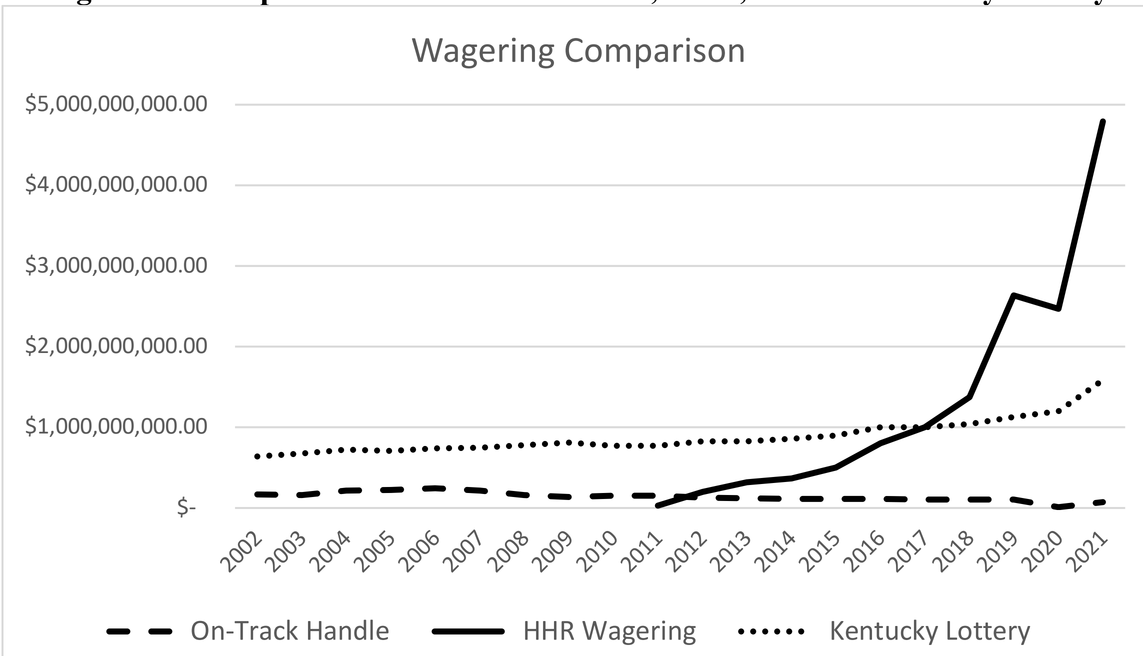
In Kentucky, although the large-scale adoption of HHR machines is quite recent, these same effects are already being observed. While Kentucky's live on-track handle continues to dwindle, the HHR handle continues to grow. As of the end of 2022, it was estimated that each machine in Kentucky contributes roughly $143,000 per month in total gross revenue to the state, which equates to a yearly total of $6.5 billion with the 3,763 machines that can be found throughout Kentucky (Lambert, 2022). This yields roughly $51 million in tax revenue for Kentucky each year. The state government is motivated to find ways to support this new form of wagering. It is important to determine whether these machines are continuing to possess a one-sided relationship with live on-track wagering and if these facilities' expansions will result in an accelerated decline for the live racing industry.

Table 2.1 shows the number of HHR terminals at the racetracks in Kentucky in June 2022. Figure 2.1 illustrates trends in on-track handle, sales of Kentucky Lottery tickets, and HHR wagering. In 2021, on-track handle was just under $72 million, Kentucky Lottery sales were almost $1.6 billion, and HHR wagering was almost $4.8 billion.

Table 2.1 HHR Terminals by Racetrack

	Number of Terminals in June 2022
Kentucky Downs	1049
Ellis Park	307
Keeneland	848
Churchill Downs	1102
Turfway	457
Total:	*3763*

Figure 2.1 Comparison of On-Track Handle, HHR, and the Kentucky Lottery



CHAPTER 3. DATA AND EMPIRICAL MODEL

3.1 Data

3.1.1 Sample

The sample for this study was comprised of the five Thoroughbred racetracks in Kentucky: Churchill Downs (Louisville), Keenland (Lexington), Ellis Park (Henderson), Kentucky Downs (Franklin), and Turfway Park (Florence). All of the tracks introduced HHR during the time of observation.[2]

3.1.2 Data Collection

3.1.2.1 Time Frame

Data were collected for calendar years 2002-2021. Financial reports were available electronically for all the active racetracks in Kentucky starting in 2002. The most recent figures available across all five tracks came from 2021 reports. For the OLS analysis, data was used through 2021. For the difference-in-differences analysis, data until 2019 was utilized. The difference in data ranges exists for two reasons. First, 2019 was the final year before the COVID shutdown, which canceled or delayed races, limited attendance, and affected purses and wagering numbers across the state. Second, for Turfway to remain a control for the staggered difference-in-differences model, data collection ceased when Turfway Park opened its HHR facility and lacks its own counterfactual untreated park to benchmark its HHR facility opening. Although Turfway Park's racetrack gaming facility did not open until 2022, Newport Racing and Gaming

[2] The Red Mile, Bluegrass Downs, Oak Grove, and Thunder Ridge were excluded from the sample because they are standardbred tracks that participate in harness racing. Additionally, Oak Grove and Thunder Ridge were not in operation for the entire duration of the study.

was opened in September 2020 under the extension of Turfway's racing license. Thus, for Turfway to be considered a control, there needed to be no presence of HHR, making 2019 the last year this was possible.

3.1.2.2 Sources

Data for this study come from reports on publicly-available sites and services. The main data source is the Kentucky Horse Racing Commission reports, and other sources were used to supplement missing values. These supplemental sources include BloodHorse, Equibase, and The Jockey Club. Each source was utilized to collect certain variables within the data set, which are explained in more detail below.

3.1.2.3 Kentucky Horse Racing Commission

The Kentucky Horse Racing Commission (KHRC) is an independent agency of the Kentucky state government that is tasked with regulating the horse racing industry. This includes both animal/human welfare within the racing industry as well as all parimutuel wagering on the racing industry that occurs across the state. They publish three types of reports that were used to assemble the dataset for this study: biennial reports, annual reports, and monthly parimutuel wagering reports.

The biennial reports and annual reports are very similar in composition and structure. Both are reported based on a calendar year rather than a fiscal year. These reports were first made available electronically in 2002, which was available in a biennial report with 2003. Biennial reports were available for 2002-2003, 2004-2005, 2006-2007, 2008-2009, 2010-2011, 2012-2013, 2015-2016, 2017-2018, and 2019-2020. There are only four annual reports available: 2006, 2008, 2010, and 2014. The contents of the reports include a financial analysis for the entire state, veterinary reports, division of funds, and statistics for both Thoroughbred and Standardbred tracks. These reports provided yearly totals for the annual race days, annual purses, and on-track handle for each racetrack.

In June 2015, KHRC began publishing monthly reports, which are organized online according to fiscal year. These reports include less legislative and administrative information than the previously mentioned reports but include more detailed financial information. More specifically, they break down, on a per-track basis, how much of the total parimutuel handle comes from each wagering source (live on-track, simulcast, advanced deposit, etc.), and how parimutuel tax revenue is allocated across government recipients. Additionally, these reports give the total number of dollars wagered on historical horse racing per gaming facility and the number of terminals within each gaming facility each month. These reports provided values for the total amount wagered on historical horse racing per year. Although these reports are published according to the fiscal years, the yearly totals reported in each data set reflect calendar year totals, which were computed by adding values for the twelve relevant reported months for each year.

3.1.2.4 Supplemental Sources

Although the Kentucky Horse Racing Commission provided most of the data for the study, additional data were needed. Supplemental sources include The Jockey Club, BloodHorse, and Equibase.

The Jockey Club is a national organization responsible for breed integrity through their upkeep of *The American Stud Book.* Every year, The Jockey Club releases a yearly fact book that serves as a guide for Thoroughbred racing across all of North America. The fact books were originally published in paper form, but now are available online through The Jockey Club website. Each fact book reports results and statistics from the previous year. For example, the 2022 Fact Book reports statistics from the 2021 racing season. These fact books report statistics from three main enterprise categories: breeding, racing, and sales. In addition, statistics are reported in aggregate for North America as well as by state/territory/province to provide statistics that are relevant to specific regions. The Kentucky fact book, alongside the national fact book, provide the number of Kentucky foals born, national mean yearling sales price, national median yearling sales price, and Kentucky average purse per race.

BloodHorse and Equibase also helped fill gaps in the data. BloodHorse is a news organization that covers racing, breeding, and sales across the United States. Equibase, in a similar manner, serves as the sole collector of North American Thoroughbred racing statistics and data. It reports handicapping products, as well as video replay and statistics for races, to better inform people of the success of different horses, jockeys, trainers, and owners throughout the industry. These two sources were accessed to obtain values for the number of race days or purses when they did not appear in the KHRC reports.

3.1.3 Interpolated Values

Despite having multiple data sources, there were still two specific instances when missing values had to be estimated to complete the analysis: on-track handle for 2015 and total purses for 2021. Eventually, the former will be obtained using an open records request, while the latter will be available when the KHRC 2021-2022 biennial report is published.

On-track handle for 2015 was calculated by taking an average of the on-track handle from 2014 and 2016. Since on-track handle generally follows a constant pattern over time, this is a reasonable approach in estimating this figure. However, the Breeders' Cup was held at Keeneland in 2015. This world championship event is run across two days and is a collection of the world's best horses in the Thoroughbred industry competing against one another. With such high-quality horses and significantly higher purses, the on-track handle for this event likely results in a yearly on-track handle that is higher than the value interpolated.

To interpolate a value for 2021 purses, a different approach was used. OwnerView, a website collaboration between The Jockey Club and the Thoroughbred Owners and Breeders Association, posted average daily purses for each track in the sample for 2021. Thus, total purses for each track were determined by multiplying the average daily purses by the number of race days.

While the estimated value of 2021 purses should be fairly accurate, rounding errors may exist. This value can easily be updated later when the Kentucky Horse Racing Commission releases the 2021-2022 biennial report, which is expected by the end of 2023.

3.2 Motivation for Empirical Modeling

In this study, the relationship between HHR and industry metrics is investigated using two approaches. It is important to understand the purpose of each of these approaches as well as their limitations. The OLS regressions were conducted as a means to illustrate the descriptive relationship between two variables. More specifically, the OLS regressions utilize key industry metrics as the dependent variables and HHR wagering as the independent variable. The results from the OLS models describe how HHR wagering and industry metrics move together, but they do not have any causal interpretations. In addition, it is important to note that due to limited data and degrees of freedom, only univariate regressions were performed, which means there is potential for omitted variable bias. Omitted variable bias arises when a variable is left out of an OLS regression but has a relationship with both the independent and dependent variables. This results in an inaccurate estimate of the slope coefficient, which captures the effect of the independent variable on the dependent variable, because some of the effect should be attributed to the missing, or omitted, variable.

One way to avoid the potential for omitted variable bias is to utilize an approach that can control for other factors, which is accomplished using a staggered difference-in-differences model. This type of approach takes into account the non-uniform rollout and growth in HHR wagering for different tracks across years and can also address baseline differences and temporal shocks to the Kentucky horse racing industry. For example, if there were large economic events or statewide changes, such as the 2008-2009 recession or new gambling laws, these changes would be accounted for in the model in the pre-treatment period. The model, as the name suggests, can be conceived as a simple difference in the change in outcomes variables between 2011 and 2021 for tracks with and without HHR. The double "difference" comes from the fact that you are examining differences in changes of the outcome variable as opposed to differences in levels. This approach constructs a clearer picture of the counterfactual evolution of these key metrics over time had HHR never been introduced (as seen in Figures 1.2 and 1.3). The idea of the parallel trends assumption, which is explained in more detail in Section 3.4, helps to address this issue of omitted variable bias within the model by looking at how the racetracks behaved both before and after treatment. Thus, the staggered difference-in-differences model goes one step further in its analysis and could potentially result in a more accurate overall estimate for the effect of HHR wagering on purses and on-track handle.

3.3 OLS Regression Models

3.3.1 Model

To understand the descriptive relationships between HHR wagering and key metrics that are used in the racing industry, ordinary least squares (OLS) regressions were utilized. OLS regressions are used to explain linear relationships between two or more variables. Regressions can be completed using one or more independent variables to explain variation in the dependent variable. OLS does this by estimating the equation for a line (intercept and slope coefficients) that minimizes the sum of squared residuals, where the residuals are the differences between the observed values of the dependent variable and the corresponding predicted values.

The standard equation for OLS regression is as follows:

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_n + \varepsilon_0$$

where y is the dependent variable, β_0 is the constant (where the line crosses the y-axis), β_i is the slope coefficient for independent variable x_i,and ε_0 is the error term (Gross & Grobb, 2004). While ε is unobservable, this random error term is included in the equation to account for a collection of small errors that will occur in the regression and is assumed to be normally distributed with mean 0 and variance σ^2.

3.3.2 Postestimation Tests

3.3.2.1 Shapiro-Wilk W Test

First proposed in 1965, the Shapiro-Wilk test is a test for normality to determine if the sample comes from a normal distribution (Shapiro and Wilk, 1965). Their model can be described as:

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

where *W* is the calculated test statistic, $x_{(i)}$ are the ordered sample values from the data reported, and a_i are the constants that are generated by the expression:

$$(a_1, a_2, \ldots, a_n) = \frac{m^T V^{-1}}{(m^T V^{-1} m)^{1/2}}$$

In this equation, $m = (m_1, m_2, \ldots, m_n)^T$ represent the expected values of the ordered statistics, which are independent and identically distributed random variables following the standard normal distribution of *N*(0,1). *V* is the covariance matrix of the order statistics for the data (Ramachandran & Tsokos, 2021).

However, this is not the model Stata 17 uses. Stata 17 and the command *swilk* was modeled using Patrick Royston's approach, (Royston (1982), Royston (1992), Royston (1993)). Where Shapiro-Wilk used a sample size of $n = 50$ for their model, Royston used $n = 2000$ for his model. This was based on the suggested approximation from Shapiro and Wilk, which was unable to be evaluated at the time due to it being an extremely data-hungry measurement that exceeded computational abilities. Thus, Royston's method was created as a simpler way to approximate the test statistic, *W*, for any sample size given the current technology available. Royston's equation is as follows:

$$W = \left(\sum \boldsymbol{a}_i y_i\right)^2 / \sum (y_i - \bar{y})^2 \ .$$

In this model, $\boldsymbol{a} = (a_1, \ldots, a_n)^T$ is such that $(n-1)^{-1/2}\sum a_i y_i$ is the best linear unbiased estimate for the standard deviation of y_i, if normality for the data is assumed. Beyond that, the exact value for $\boldsymbol{a}$ can be found using the equation:

$$\boldsymbol{a} = (m^T V^{-1} V^{-1} m)^{-\frac{1}{2}} m^T V^{-1} \,.$$

In this equation, V is the covariance matrix of the order statistics of a sample of n standard normal random variables of $N(0,1)$. These variables all also have an expectation vector m (Royston, 1992).

Using these equations, Stata was used to calculate the test statistic, W, for the OLS regressions. The command *swilk* was used and the W test statistic was recorded for each model. For this test, the null hypothesis, or H_0, is that the residuals are normally distributed. An alpha of 0.05 was chosen as the relevant significance level. If the *p*-value for the W test statistic fell below this 0.05 value, the null hypothesis was rejected.

3.3.2.2 Breusch-Pagan/Cook-Weisberg Test

The Breusch-Pagan test is a Lagrange multiplier test for heteroskedasticity of the residuals. This test was first proposed by Breusch and Pagan (1980) and can be derived from the function:

$$\sigma_i^2 = \sigma^2 h(x_i' \alpha)$$

where h is an unknown, continuously differential function that does not depend on i. Additionally, $h(.) > 0$ and $h(0) = 1$. The null hypothesis is that the error variances are all equal, with the alternative stating that the variances are a multiplicative function of one or more variables found within the equation. In the simplest way, the test can be computed by taking the number of observations within the data set multiplied by the R^2 of an auxiliary regression. This auxiliary equation can be found by regressing ε_i^2, or the squared residuals from the OLS equation, on z_i and a constant value (Verbeek, 2017). This test was completed using the Stata 17 command *hettest*. A critical value of 0.05 was used to determine whether or not the null hypothesis was rejected.

3.3.3 Key Industry Metrics

Key industry metrics used in this analysis are defined in Table 3.1. These metrics include on-track handle, purses, mean and median yearling sales price, Kentucky foal crop size, average purse per race, average on-track handle per race, and HHR handle.

Table 3.1 Table of Key Variables used by the Thoroughbred Industry

Variable Name	Description
OTH	Aggregate yearly value of On-Track Handle from all tracks in the sample
Purses	Aggregate yearly value of purses from all tracks in the sample
Mean Price	National mean yearling price of yearlings sold in auction
Median Price	National median yearling price of yearlings sold in auction
KY Foal Crop	Number of foals born in Kentucky in a calendar year
Avg Purse Per Race	Average purse per race for all tracks in the sample
Avg Handle Per Race	Average on-track handle per race for all tracks in the sample
HHR	Aggregate yearly value of HHR wagering from all tracks in the sample

Using these variables, the purpose of the OLS regressions was to develop the "best" model to describe the relationship between two variables, using significance of coefficient estimates, adjusted R^2, and post-estimation tests as guides. In general, it was expected that HHR wagering would have a lagged effect on most of these variables since dollars wagered in the current year could not contribute to purses and other metrics in the current year; rather these funds would be contributed in the following years.

3.3.4 Specific Models to be Estimated

Data used for this portion of analysis was from 2011-2021, since HHR was introduced in 2011.

3.3.4.1 HHR and On-Track Handle

Model 1 explores the relationship between HHR and the on-track handle at the racetracks. Model 1b is a *ln-ln* transformation of Model 1a. The two models are as follows:

Model 1a: $OTH_t = \beta_0 + \beta_1 HHR_{t-1} + \varepsilon_t$

Model 1b: $\ln(OTH_t) = \beta_0 + \beta_1 ln(HHR_{t-1}) + \varepsilon_t$

3.3.4.2 HHR and Purses

Model 2 analyzes the relationship between HHR wagering and purses. The two models are as follows:

Model 3a: $Purses_t = \beta_0 + \beta_1 HHR_{t-2} + \varepsilon_t$

Model 3b: $\ln(Purses_t) = \beta_0 + \beta_1 \ln(HHR_{t-2}) + \varepsilon_t$

3.3.4.3 HHR and Sales Price

Models 3 and 4 explore the relationship between HHR wagering and the national mean yearling sales price. This model was motivated by the idea that HHR wagering might be reflected in higher sales prices. The equations used in the models are as follows:

Model 3a: $Mean\ Price_t = \beta_0 + \beta_1 HHR_{t-1} + \varepsilon_t$

Model 3b: $\ln(Mean\ Price_t) = \beta_0 + \beta_1 \ln(HHR_{t-1}) + \varepsilon_t$

Model 4a: $Mean\ Price_t = \beta_0 + \beta_1 HHR_{t-2} + \varepsilon_t$

Model 4b: $\ln(Mean\ Price_t) = \beta_0 + \beta_1 \ln(HHR_{t-2}) + \varepsilon_t$

Models 5 and 6 analyze the relationship between HHR wagering and the national median yearling sales price. The equations used in the models can be found below:

Model 5a: $Median\ Price_t = \beta_0 + \beta_1 HHR_{t-1} + \varepsilon_t$

Model 5b: $\ln(Median\ Price_t) = \beta_0 + \beta_1 \ln(HHR_{t-1}) + \varepsilon_t$

Model 6a: $Median\ Price_t = \beta_0 + \beta_1 HHR_{t-2} + \varepsilon_t$

Model 6b: $\ln(Median\ Price_t) = \beta_0 + \beta_1 \ln(HHR_{t-2}) + \varepsilon_t$

3.3.4.4 HHR and Kentucky Foal Crop

The last model, Model 7, explores the relationship between HHR wagering and the number of foals produced in Kentucky on an annual basis. Here, a lagged effect is expected for a few reasons. First, biology and breeder decisions cannot react immediately to increases in HHR. There is likely at least a 2-year process from the time higher purses are observed to the time a foal is born. Secondly, this effect could be even longer if breeders want to that higher purses are sustained over an extended period of time. The models for Kentucky foal crop were:

Model 7a: $KY\ Foal\ Crop_t = \beta_0 + \beta_1 HHR_{t-1} + \varepsilon_t$

Model 7b: $\ln(KY\ Foal\ Crop_t) = \beta_0 + \beta_1 \ln(HHR_{t-1}) + \varepsilon_t$

Model 8a: $KY\ Foal\ Crop_t = \beta_0 + \beta_1 HHR_{t-2} + \varepsilon_t$

Model 8b: $\ln(KY\ Foal\ Crop_t) = \beta_0 + \beta_1 \ln(HHR_{t-2}) + \varepsilon_t$

3.4 Staggered Difference-in-Differences

3.4.1 Preferred Model

The model for this analysis was based on staggered adoption difference-in-differences (DID) estimator put forward by De Chaisemartin and D'Haultfœuille (2020). Difference-in-differences is a technique used by researchers to compare the changes in outcomes between a group that receives some form of treatment and a control group and was first used by Snow in 1855 (Goodman-Bacon, 2021). Originally, difference-in-differences models were developed using one treatment time period, where a single group received treatment at the same time. However, this model was inappropriate in situations where they were analyzing a treatment that was implemented at different times, or was staggered, throughout the treatment group.

Goodman-Bacon (2021) began the analysis of two-way fixed effect estimators when looking at a staggered adoption model. He decomposed the two-way fixed effects estimator by showing that it equals a weighted average of all possible two-group/two-period DID estimators in the data. The decomposed model was referenced in De Chaisemartin and D'Haultfœuille (2020) using the following model:

$$w_{g,t} = \frac{\varepsilon_{g,t}}{\sum_{(g,t):D_{g,t}=1} \frac{N_{g,t}}{N_1} \varepsilon_{g,t}}$$

where w equals the weight, g is the group, t is the time of treatment, and ε is equal to the residual of observations in cell (g,t) using a two-way fixed effect DID model. The denominator is equal to the expectation of a weighted average of the treated cells, where $D_{g,t}$ is a treatment indicator = 1 if the group was treated in t and is 0 otherwise, $N_{g,t}$ is the number of observations in group g at time t, and N_1 is the number of treated observations over all groups and time. However, this decomposition gave way to an interesting result that the weights sum to one but are non-convex and can be negative. Since weights across all groups and time being studied must sum up to one, some of the weights are bound to be negative. This negative weight becomes an issue when a group is treated across multiple time periods and another is not treated, which could allow for the treatment group to actually enter the analysis as a control. This issue matters if there appears to be treatment effect heterogeneity because it could result in bias for the final DID estimator. Thus, negative weights are a concern when treatment effects differ across periods where a large fraction of groups are treated, and to groups treated for many periods (De Chaisemartin and D'Haultfœuille, 2020).

Since negative weights only arise when newly treated groups are compared to established treated groups, researchers such as De Chaisemartin and D'Haultfœuille (2020), Callaway and Sant'Anna (2021), and Sun and Abraham (2021) developed staggered adoption DID methods to eliminate the issue of negative weights. For De Chaisemartin and D'Haultfœuille (2020), they eliminated this concern by only comparing non-treated groups in a specific time period to switching groups within that same time period. This change in approach gave way to the two-way fixed effect estimator, which can be modeled by the following equation:

$$DID_M = \sum_{t=2}^{T} \left(\frac{N_{1,0,t}}{N_s} DID_{+,t} + \frac{N_{0,1,t}}{N_s} DID_{-,t} \right)$$

where DID_M is the weighted average of the DID estimators, $N_{1,0,t}$ is the number of observations that were newly treated in period t, $N_{0,0,t}$ is the number of observations that were not treated in period t or period t-1, N_S is the number of switching observations, and $N_{g,t}$ is the number of observations in group g at time t. However, for this specific analysis, no group ever loses its treated status. Thus, for this analysis, the equation simplifies to:

$$DID_M = \sum_{t=2}^{T} \left(\frac{N_{1,0,t}}{N_s} DID_{+,t} \right)$$

which is a weighted average of only the $DID_{+,t}$ estimators. The $DID_{+,t}$ estimator compares the evolution of the mean outcome between $t - 1$ and t for the switchers and those who remain untreated. This is the estimator that was reported as the "Dynamic" estimator for this analysis. The equation for the $DID_{+,t}$ estimator is modeled by the equation:

$$DID_{+,t} = \sum_{g:D_{g,t}=1,D_{g,t-1}=0} \frac{N_{g,t}}{N_{1,0,t}} \left(Y_{g,t} - Y_{g,t-1}\right) - \sum_{g:D_{g,t}=1,D_{g,t-1}=0} \frac{N_{g,t}}{N_{0,0,t}} \left(Y_{g,t} - Y_{g,t-1}\right)$$

The parallel trends assumption is one of the critical identification assumptions underlying the DID design. It states that absent the treatment, the treated groups' outcomes would identically trend in the same way as the control group. If this was true, outcome trends in the control group can be used as a counterfactual baseline for the treated group. Because the counterfactual treated group trends are unobservable, this assumption cannot be completely verified. However, researchers can assess its plausibility by comparing outcome trends in the treated group and control group before treatment. If, prior to treatment, these trends are statistically identical, it would provide support for the parallel trends assumption.

This "parallel trends assumption" is tested by De Chaisemartin and D'Haultfœuille (2020) using a placebo estimator. This placebo estimator (*pl*) is modeled by the equation:

$$DID_{+,t}^{pl} = \sum_{g:D_{g,t}=1, D_{g,t-1}= D_{g,t-2}=0} \frac{N_{g,t}}{N_{1,0,0,t}} \left(Y_{g,t-1} - Y_{g,t-2}\right) - \sum_{g:D_{g,t}= D_{g,t-1}=D_{g,t-2}=0} \frac{N_{g,t}}{N_{0,0,0,t}} \left(Y_{g,t-1} - Y_{g,t-2}\right)$$

This estimator compares the evolution of the mean outcome from $t-2$ to $t-1$ in two sets of groups: those untreated at $t-2$ and $t-1$ but treated at t, and those untreated at $t-2$, $t-1$, and t. For this effect, the desired outcome is a result that is not statistically different from 0, because this means that the two groups of racetracks were behaving in a similar manner prior to the introduction of treatment.

To calculate the standard error, this method utilizes a technique called block bootstrap sampling, which assumes that errors across racetracks are independent, but are dependent within racetracks. In this method, racetracks are randomly sampled, using replacement, and the effects are re-estimated using the randomly sampled data. The purpose of this is to get an estimate of the uncertainty around the treatment effect being put in place. For this analysis, 500 replications were used to determine the value of standard error.

To complete this analysis, the *did_multiplegt* package was used in Stata 17. Unlike the OLS regression models which used data from 2011 to 2021, data for this portion of the analysis will use data from 2002-2019 from the five Thoroughbred racetracks.

3.4.2 Robustness Check

The robustness check was completed using the same methods from above, with the addition of one more placebo and dynamic effect. The additional placebo effect was completed to provide further evidence of the parallel trends assumption, which is crucial for the DID analysis. The additional dynamic effect was estimated to further quantify the effects of the treatment, HHR wagering, after its implementation.

In order to estimate the additional dynamic effect, one of the tracks had to be dropped. Since the data being analyzed ended in 2019, Churchill Downs is only able to be analyzed for a dynamic effect for 1 year after treatment is introduced. Therefore, this additional dynamic effect is only using three treated groups instead of four. Although this is not ideal due to the already limited data, the number of dynamic effects being estimated must be equal to or greater than the number of placebo effects being estimated. Therefore, due to the importance of the additional placebo effect and the fact that this is a preliminary analysis, this elimination of one group was necessary to achieve a second placebo effect.

CHAPTER 4. RESULTS

4.1 Descriptive Statistics

To better understand the data, descriptive statistics for each variable utilized in the both the OLS and DID models were compiled. These statistics were completed using data from 2002-2021, resulting in *n*=20.

Table 4.1 Descriptive Statistics for Annual Purses, On-Track Handle, and HHR

Descriptive Statistics				
	Mean	Std. Dev.	Minimum	Maximum
Annual Purses	$ 84,200,000	$ 16,900,000	$ 64,100,000	$ 130,000,000
On-Track Handle	$ 141,000,000	$ 55,800,000	$ 10,200,000	$ 245,000,000
HHR Wagering	$ 1,320,000,000	$ 1,450,000,000	$ 29,300,000	$ 4,790,000,000

Table 4.1 displays the descriptive statistics for annual purses, on-track handle, and HHR wagering. The average value of annual purses across all tracks was $84,200,000, with a minimum of $64,100,000 and a maximum of $130,000,000. For on-track handle, the average value over the 20-year period was $141,000,000. The lowest annual on-track handle was $10,200,000, while the highest annual on-track handle was $245,000,000. For HHR wagering, which was first introduced in 2011, the average annual wagering amount was $1,320,000,000, with a minimum of $29,300,000 and a maximum of $4,790,000,000.

Table 4.2 Descriptive Statistics for HHR per Track

HHR Wagering by Track		
	Mean	Std. Dev
Churchill Downs	$ 1,030,628,385.80	$ 687,342,122.99
Keeneland	$ 335,074,078.63	$ 170,814,997.09
Ellis Park	$ 85,002,982.73	$ 60,388,357.29
KY Downs	$ 698,496,271.55	$ 487,586,400.50

Table 4.2 displays the descriptive statistics for HHR wagering for each track. This is presented to illustrate the size and impact each track has on the statewide HHR

wagering total. On average, Churchill Downs contributes the most to the statewide HHR total, averaging over $1 billion per year. Kentucky Downs contributes the next highest amount, averaging nearly $700,000 in HHR wagering per year. Keeneland contributes, on average, a yearly amount of $335,074,078.63 towards the statewide total. Ellis Park contributes an average of $85,002,982.73 annually. These figures are directly related to the number of terminals at each track, as was illustrated in Table 2.1.

Table 4.3 Comparison of Purses Before and After the Introduction of HHR

	Average Yearly Purses before HHR Introduction	Average Yearly Purses after HHR Introduction
Churchill Downs	$ 36,164,969.69	$ 62,106,447.50
Ellis Park	$ 5,969,714.00	$ 5,708,930.29
Keeneland	$ 19,639,412.38	$ 27,859,176.20
KY Downs	$ 1,257,794.44	$ 5,976,057.00
Turway Park	$ 10,940,695.11	--

Table 4.4 Comparison of On-Track Handle Before and After the Introduction of HHR

	Average On-Track Handle before HHR Introduction	Average On-Track Handle after HHR Introduction
Churchill Downs	$ 91,748,157.52	$ 61,838,655.00
Ellis Park	$ 10,452,171.20	$ 4,894,046.29
Keeneland	$ 45,701,949.82	$ 34,302,644.54
KY Downs	$ 2,437,527.67	$ 918,094.77
Turway Park	$ 11,221,461.30	--

Tables 4.3 and 4.4 provide a first look at the impact HHR has had on both purses and on-track handle. Table 4.3 reports the average annual purse at each of the five racetracks before and after the introduction of HHR.[3] For the four other tracks, all of the average yearly purses increased following the introduction of HHR except for Ellis Park. Table 4.4 displays the average yearly on-track handle before and after the introduction of

[3] Turfway Park was excluded from the "after" introduction period since it was only studied prior to the introduction of any HHR machines at Turfway Park

HHR at each of the five racetracks. Across the four racetracks that had an “after” period, each of the racetracks exhibit a decrease in average yearly on-track handle following the introduction of HHR. However, this downward trend in on-track handle was already something that was occurring in the absence of HHR and should not be interpreted as a causal effect.

Table 4.5 Descriptive Statistics for Mean Price, Median Price, and KY Foal Crop

Descriptive Statistics				
	Mean	Std. Dev.	Minimum	Maximum
Mean Price	$ 61,109.90	$ 13,468.34	$ 42,273	$ 86,907
Median Price	$ 18,568.15	$ 6,100.26	$ 10,023	$ 32,000
KY Foal Crop	8,701.21	1,092.07	7,183	10,517

Lastly, Table 4.5 presents the descriptive statistics for the national mean yearling price, national median yearling price, and the size of the Kentucky foal crop. The average yearly national mean sales price was $61,09.90, with a minimum and maximum of $42,273 and $86,907, respectively. The median sales price averaged $18,568.15 over the 20-year period. The lowest median sales price was $10,023 and the highest was $32,000. Note that the large difference between the mean and median foal prices suggests that sales prices are skewed. In this data, the median is found to be lower than the mean, meaning the data possesses positive skewness. This is due to the fact that every year, a few extreme sales prices drive up the mean value. For the last variable, Kentucky foal crop, the mean value was 8,701.21 foals born in a calendar year. The highest reported foal crop was 10,517 foals, and the smallest foal crop had 7,183 foals.

4.2 OLS Regression Results

This section presents the results of the OLS regressions in Models 1 – 9 as well as postestimation test results.

4.2.1 On-Track Handle and Lagged HHR Wagering

Table 4.6 Regression Results of the Determinants of On-Track Handle

DV: On-Track Handle (a) and *ln* (On-Track Handle) (b)		
	Model 1a	**Model 1b**
HHRLag1	-0.033***	-0.310
Constant	131,000,000***	24.496***
n	10	10
Adj. R^2	0.777	0.212
F-Stat	32.44	3.42
Prob > F	0.001	0.102
SW *p*-value	0.214	0.001
BP *p*-value	0.001	0.061

The best model explaining the relationship between on-track handle (dependent variable) and HHR wagering (independent variable) is presented in Table 4.6. In Model 1a, on-track handle decreases by an average of $0.033 for each additional dollar wagered on HHR the previous year, ceteris paribus. This coefficient estimate is significant at the 1% level. Additionally, the adjusted R^2 value indicates that 77.70% of the variation in on-track handle is described by the variation in HHR_{t-1}. Post estimation tests were conducted to test for normality and heteroskedasticity of residuals. The *p*-value for the Shapiro-Wilk test statistic, which tests for normality in the data, was 0.214. At this value, the null hypothesis that the residuals are normally distributed cannot be rejected at the 5% level. Additionally, the Breusch-Pagan test, which tests for heteroskedasticity, resulted in a *p*-value of .001, which is smaller than 0.05. Thus, the null hypothesis is rejected in this case, meaning that heteroscedasticity cannot be ruled out.

Model 1b investigates the same relationship with a *ln-ln* transformation of on-track handle and HHRLag1. In this model, on-track handle decreases by an average of 0.310% for each additional increase of 1% on HHR wagering from the previous year, ceteris paribus. However, this value is not statistically different from 0. Model 1b has an adjusted R^2 of 0.212. The value cannot be compared to the adjusted R^2 in Model 1a because the dependent variables are different. Postestimation tests resulted in a reversal of the test statistics. In Model 1b, the Shapiro-Wilk test produced a p-value of 0.001; therefore, the null hypothesis that the residuals are normally distributed is rejected. Conversely, the Breusch-Pagan test had a p-value of .061, which means that the null hypothesis of constant variance of residuals fails to be rejected at the 5% level.

4.2.2 Annual Purses and Lagged HHR Wagering

Table 4.7 Regression Results of the Determinants of Annual Purses

DV: Annual Purses (a) and ln(Annual Purses) (b)		
	Model 2a	**Model 2b**
HHR Lag2	0.022***	0.147***
Constant	74,100,000***	15.380***
n	9	9
Adj. R^2	0.615	0.594
F-Stat	13.78	12.69
Prob > F	0.008	0.009
SW *p*-value	0.143	0.012
BP *p*-value	0.877	0.961

The best model explaining the relationship between HHR wagering and purses is presented in Table 4.7. Models 2a and 2b utilize annual purses as the dependent variable but introduce HHR_{t-2} as the independent variable. The slope coefficients for HHR_{t-2} in Model 2a was 0.022 and ln(HHR_{t-2}) in Model 2b was 0.147, respectively. Both estimates were significant at the 1% level. In Model 2a, annual purses increase by an average of $0.022 for each additional dollar wagered on HHR from two years prior, ceteris paribus. This means that for every $1 million increase in HHR_{t-2}, annual purses increase by $22,000. In Model 2b, annual purses increase by 0.147% for each additional 1% of HHR wagering that was made two years earlier, ceteris paribus.

The postestimation tests for Model 2a fail to reject normality of residuals and homoscedasticity of residuals. For Model 2b, the Shapiro-Wilk test had a *p*-value of 0.012, meaning the null hypothesis of normality of residuals is rejected. For the Breusch-Pagan test, Model 2b had a *p*-value of 0.961, meaning that the postestimation test fails to reject homoskedasticity of residuals in this model.

4.2.3 National Mean Yearling Price and Lagged HHR Wagering

Table 4.8 Regression Results of the Determinants of National Mean Yearling Price

DV: National Mean Yearling Price (a) and *ln* (National Mean Yearling Price) (b)	Model 3		Model 4	
	Model 3a	**Model 3b**	**Model 4a**	**Model 4b**
HHRLag1	0.000007*	.084***	--	--
HHRLag2	--	--	0.000008**	.067**
Constant	65010.240***	9.471***	66619.670***	9.854***
n	10	10	9	9
Adj. R^2	0.254	0.555	0.389	0.396
F-Stat	4.07	12.21	6.09	6.24
Prob > F	0.078	0.008	0.043	0.041
SW *p*-value	0.984	0.767	0.732	0.570
BP *p*-value	0.139	0.069	0.723	0.503

Table 4.8 presents the OLS regression results utilizing national mean yearling price as the dependent variable and HHR_{t-1} and HHR_{t-2} as the independent variables in Models 3 & 4, respectively. The relationship being analyzed was whether or not HHR wagering might be reflected in sales prices. First looking at Model 3a, the slope coefficient estimate for HHR_{t-1} is 0.000007 and is significant at the 10% level. This means that holding all else constant, a $1 billion increase of HHR wagering from the previous year has, on average, a $7,000 increase in the national mean yearling price. For Model 3b, the slope coefficient estimate is 0.084, which is significant at the 1% level.

Holding all else constant, this means that on average a 1% increase in HHR wagering from the previous year results in a 0.084% increase in the national mean yearling price.

Based on the postestimation tests, the null hypotheses for the both the Shapiro-Wilk test and the Breusch Pagan test fail to be rejected in both Models 3a and 3b. This means that both models fail to reject normality of residuals and homoscedasticity of residuals.

Looking at Models 4a and 4b, the slope coefficient estimate for Model 5a is equal to 0.000008, and is significant at the 5% level. This means that, on average, an additional $1 billion spent on HHR wagering from two years ago leads to an $8,000 increase of the national mean yearling price of the current year, ceteris paribus. The slope coefficient estimate for Model 4b is equal to 0.067, and is significant at the 5% level. This indicates that a 1% increase in HHR wagering from two years earlier results in a 0.067% increase in the national mean yearling price, ceteris paribus.

Looking at the adjusted R^2 values, Model 4a has an adjusted R^2 equal to 0.396, and Model 4b has an adjusted R^2 equal 0.389. This means that 39.60% of the variation in the national mean yearling price is described in Model 4b, and 38.90% in Model 4a.

For the postestimation tests, the null hypotheses for the both the Shapiro-Wilk test and the Breusch Pagan test fail to be rejected for both Models 4a and 4b. This means that both models fail to reject normality of residuals and homoscedasticity of residuals.

When comparing Models 3a and 4a, Model 4a appears to be the better model. Not only does the coefficient estimate have a higher level of significance, the model also possesses a higher R^2 value. However, when comparing Models 3b and 4b, Model 3b is found to be the better model. The slope coefficient estimate was found to have a higher level of significance (although both are significant at the 5% level) and the model, overall, describes more of the variation in the natural log of national mean yearling price.

4.2.4 National Median Yearling Price and Lagged HHR Wagering

Table 4.9 Regression Results of the Determinants of National Median Yearling Price

DV: National Median Yearling Price (a) and *ln* (National Median Yearling Price) (b)				
	Model 5		**Model 6**	
	Model 5a	**Model 5b**	**Model 6a**	**Model 6b**
HHRLag1	0.000	0.016	--	--
HHRLag2	--	--	0.000	0.010
Constant	22918.120***	9.739***	22045.450***	9.868***
n	10	10	9	9
Adj. R^2	-0.058	-0.096	0.196	-0.132
F-Stat	0.5	0.21	2.95	0.06
Prob > F	0.498	0.657	0.130	0.806
SW *p*-value	0.758	0.599	0.495	0.606
BP *p*-value	0.009	0.075	0.373	0.120

Table 4.9 presents the OLS regression results using national median yearling price as the dependent variable (instead of mean), and HHR_{t-1} and HHR_{t-2} as independent variables. Model 5a had a slope coefficient estimate for HHR_{t-1} of 0, which was not statistically significant, and a negative adjusted R^2 value. For Model 5b, the slope coefficient estimate for HHR_{t-1} was equal to 0.016, which was again not significant. In this model, national median yearling price increase by 0.016% for each additional 1% of HHR wagering from a year ago.

For the Shapiro-Wilk Test, Model 5a was found to have a p-value of 0.758, meaning that the null hypothesis fails to be rejected for the normality of residuals. The same was found for Model 5b, which had a *p*-value of 0.599. For the Breusch-Pagan test, Model 5a p-value of 0.009, which rejects the null hypothesis of homoscedasticity of the residuals. For Model 5b, which had a *p*-value of 0.075, the null hypothesis of homoscedasticity fails to be rejected at the 5% level.

Looking at the Models 6a and 6b, which were between national median yearling price and HHR_{t-2}, the slope coefficient estimate for Model 6a was found to be not statistically different from 0, with an adjusted R^2 of 0.196. This means that, on average, a one dollar increase in HHR wagering from two years prior has no effect on the national median yearling price. For Model 6b, the slope coefficient estimate was equal to 0.010 and was not significant, and the adjusted R^2 was negative.

For the Shapiro-Wilk postestimation test, Model 6a was found to have a *p*-value of 0.495, and Model 6b was found to have a p value of 0.606. Both models fail to reject the null hypothesis of normality of the residuals. For the Breusch-Pagan test, both models also fail to reject the null hypothesis of homoscedasticity of the residuals.

4.2.5 Kentucky Foal Crop and Lagged HHR Wagering

Table 4.10 Regression Results of the Determinants of Kentucky Foal Crop

DV: KY Foal Crop (a) and *ln* (KY Foal Crop) (b)				
	Model 7a	**Model 7b**	**Model 8a**	**Model 8b**
HHRLag1	0.000	0.033**	--	--
HHRLag2	--	--	0.000	0.027**
Constant	7820.236***	8.322***	7916.552***	8.456***
n	9	9	8	8
Adj. R^2	-0.005	0.487	0.118	0.509
F-Stat	0.96	8.6	0.8	6.23
Prob > F	0.360	0.022	0.406	0.047
SW *p*-value	0.282	0.829	0.413	0.949
BP *p*-value	0.607	0.499	0.486	0.564

Table 4.10 presents the OLS regression results using Kentucky foal crop as the dependent variable and using both HHR_{t-1} and HHR_{t-2} as the independent variables. In Model 7a, which uses HHR_{t-1} as the independent variable, the slope coefficient estimate

is not statistically different from 0 and has a negative adjusted R^2 value. For model 7b, the slope coefficient estimate for HHR_{t-1} is equal to 0.033 and is significant at the 5% level. This means that Kentucky foal crop increases by 0.033 percent for every 1% HHR wagering increases from the previous year, ceteris paribus. The adjusted R^2 value indicates that 48.70% of the variation in Kentucky foal crop is described by the variation in HHR_{t-1}.

In Model 8a, the slope coefficient estimate for HHR_{t-2} is not statistically different from 0 but did have an adjusted R^2 value of 0.118. For Model 8b, the slope coefficient estimate for HHR_{t-2} was found to be 0.027 and is significant at the 5% level. This means that Kentucky foal crop increases by 0.027 percent for an increase of 1% in HHR wagering from two years ago, ceteris paribus. The adjusted R^2 value for this model is 0.509, which indicates that 50.90% of the variation in Kentucky foal crop is described by the variation in HHR_{t-2}.

The postestimation tests for both Models 7a and 7b, as well as Models 8a and 8b, fail to reject normality of residuals and homoscedasticity of residuals.

When comparing the two models, Models 8a and 8b are more likely to be the better models between Kentucky foal crop and HHR wagering. Given the biological timeline for equine reproduction, along with breeding decisions, it is likely that wagering dollars will not have an immediate effect on breeding decisions. Thus, it is more likely that it will take at least two years, if not longer, for wagering dollars to effect breeder’s decisions.

4.3 Staggered Difference-in-Difference Results

A staggered difference-in-differences model was used utilizing the De Chaisemartin and D'Haulfoeuille (2020) approach to estimate the impact of HHR wagering on both annual purses and on-track handle while trying to control for the effects of omitted variables that were not captured in the OLS section.

4.3.1 Purses

Table 4.11 Staggered Difference-in-Differences Results for Purses

Outcome Variable: Purses				
	Placebo 1	**Effect 0**	**Effect 1**	**Average E**
Estimate	493,439.40	9,647,848	15,700,000	12,700,000
SE	3,544,712	7,470,644	6,921,222	6,318,506
LB CI	-6,454,197	-4,994,615	2,135,898	290,399
UB CI	7,441,076	24,300,000	29,300,000	25,100,000
N	14	14	13	27
Switchers	4	4	4	8

Table 4.11 displays the results of the difference-in-differences model using annual purses per track as the outcome variable. Placebo 1, which is a way to demonstrate the parallel trends of the different tracks, resulted in an estimate of $493,439.40. However, the confidence interval of (-6,454,197, 7,441,076) contains 0, which means this estimate is not significantly different from 0. This is ideal, because it means the tracks behaved in a similar manner the year before treatment is introduced.

Looking at Effect 0, which is the effect at the time treatment is implemented, the estimated effect is equal to 9,647,848. This means that when a track implements HHR, the annual purses are expected to rise $9,647,848. However, since the confidence interval of (-4,994,615, 24,300,000) contains 0, this effect estimate is not statistically significant from 0. This estimate was run using 14 years of data, with 4 out of 5 tracks receiving treatment.

Effect 1 has a similar interpretation to Effect 0, except it captures the effect 1 year after the treatment is implemented. This value is equal to 15,700,000. This means that one year after HHR is implemented at a track, purses are expected to rise $15,700,000. With a confidence interval of (2,135,898, 29,300,000), this effect is statistically significant. This estimate, like Effect 0, included 4 tracks receiving treatment, but was now based on 13 years of data. This effect would be expected to be significant due to how purses are based on the previous year of wagering, meaning HHR dollars wagered in the current year should have an effect on the purses of the following year.

The Average column (the third column in Table 5.7) represents the average overall effect HHR introduction has on purses. Based on this model, HHR implementation averages an increase of $12,700,000 per year after treatment. Additionally, since the confidence interval of (290,399, 25,100,000) does not contain 0, this estimation value is statistically significant from 0. This average was based on the information from 8 tracks (4 tracks from Effect 0 and 4 tracks from Effect 1) and 27 years (14 years from Effect 0 and 13 years from Effect 1) of data.

4.3.2 *ln* Purses

Table 4.12 Staggered Difference-in-Differences Results for the Natural Log of Purses

Outcome Variable: *ln(* Purses)				
	Placebo 1	**Effect 0**	**Effect 1**	**Average E**
Estimate	-0.025	0.250	0.682	0.466
SE	0.114	0.213	0.220	0.138
LB CI	-0.248	-0.167	0.251	0.196
UB CI	0.199	0.667	1.112	0.736
N	14	14	13	27
Switchers	4	4	4	8

Table 4.12 displays the difference-in-difference results using the natural log of annual purses per track as the outcome variable. In this model, elasticity was being measure in response to the treatment, HHR, being applied at the 4 tracks. Placebo 1 had an estimate of -0.025 and a confidence interval of (-0.248, 0.199). This effect is insignificant, indicating that the tracks seem to behave in a parallel manner pre-treatment. Effect 0 for the natural log of purses was estimated as 0.250. However, this effect is found to be statistically insignificant, since 0 is contained within the confidence interval of (-0.185, 0.685). The estimate for Effect 1 was 0.682 after the first year of HHR treatment, and this result is statistically significant. Looking at the Average effect on the treated, the estimated value reported is 0.466. This value is found to be statistically significant based on the confidence interval of (0.199, 0.733). This means that, on average, the yearly effect on the treated is equal to a 46.6 percentage point increase in annual purses after treatment.

4.3.3 On-Track Handle

Table 4.13 Staggered Difference-in-Differences Results for On-Track Handle

Outcome Variable: On-Track Handle				
	Placebo 1	Effect 0	Effect 1	Average E
Estimate	1,743,184	1,884,915	4,826,826	3,355,870
SE	3,157,232	4,505,946	6,891,692	5,413,325
LB CI	-4,444,991	-6,946,739	-8,680,889	-7,254,247
UB CI	7,931,359	10,700,000	18,300,000	14,000,000
N	14	14	13	27
Switchers	4	4	4	8

Table 4.13 displays the difference-in-differences result when using on-track handle as the outcome variable. The estimate for Placebo 1 was found to be $1,743,184. However, the confidence interval of (-4,444,991, 7,931,359) contains 0, meaning the

estimate is not statistically different from 0. This indicates that the tracks possessed parallel trends for the year before treatment.

For Effect 0, the effect estimate is equal to 1,884,915. This means by the end of the period where treatment is first introduced, on-track handle is expected to increase by $1,884,915. However, this result is statistically insignificant, given that the confidence interval of (-6,946,739, 10,700,000) contains 0.

For Effect 1, or the effect estimated one year after the introduction of the treatment, the effect is estimated to be equal to 4,826,826. This means that one year after the initial introduction of the treatment, the on-track handle is expected to rise by $4,826,826. However, just like Effect 0, this value is statistically insignificant.

When looking at the average effect on the treated, the effect estimate is equal to 3,355,870. This means that, on average, after the treatment is implemented, on-track handle is expected to rise by $3,355,870. Like Effect 0 and Effect 1, this result is statistically insignificant.

4.3.4 *ln* On-Track Handle

Table 4.14 Staggered Difference-in-Differences Results for the Natural Log of On-Track Handle

Outcome Variable: *ln(* On-Track Handle)				
	Placebo 1	Effect 0	Effect 1	Average E
Estimate	-0.119	-0.009	0.274	0.132
SE	0.076	0.151	0.208	0.156
LB CI	-0.268	-0.305	-0.134	-0.174
UB CI	0.030	0.288	0.681	0.439
N	14	14	13	27
Switchers	4	4	4	8

Table 4.14 shows the difference-in-differences results using the natural log of on-track handle as the outcome variable. The estimate for Placebo 1 was found to equal an

estimate of -0.119. However, this estimate is statistically insignificant from 0, due to the fact that the interval of (-0.268, 0.030) contains 0. For Effect 0, the estimated effect was equal to -0.009 and was statistically insignificant. For Effect 1, or the effect one year after treatment is implemented, the effect is estimated to be 0.274 but is not statistically significant. The average treatment effect was estimated as 0.132 but is not statistically different from 0.

4.4 Robustness Check Results

4.4.1 Purses

Table 4.15 Robustness Results for Purses

Purses						
	Placebo 2	Placebo 1	Effect 0	Effect 1	Effect 2	Average E
Estimate	1,250,499	493,439.40	9,647,848	15,700,000	6,738,671	11,100,000
SE	6,575,872	3,544,712	7,470,644	6,921,222	5,695,317	5,627,838
LB CI	-11,600,000	-6,454,197	-4,994,615	2,135,898	-4,424,150	25,199.27
UB CI	14,100,000	7,441,076	24,300,000	29,300,000	17,900,000	22,100,000
N	13	14	14	13	11	38
Switchers	4	4	4	4	3	11

Table 4.15 provides the DID effect estimation with the inclusion of an additional placebo and dynamic effect. Placebo 2, which is the estimation for purses for the treated groups two years prior to the introduction of treatment, was found to have an estimate of $1,250,499. However, this estimate is statistically insignificant, due to the confidence interval of (-11,600,000, 14,100,000) contains the value of 0. Given that the estimate was statistically insignificant from 0, this insinuates that the tracks were behaving in a parallel manner two years prior to the introduction of the treatment, helping to further support the potential for the parallel trends theory.

Effect 2, which estimates the effect on purses two years after the introduction of the treatment, was estimated to be $6,738,671. This effect was found to be statistically insignificant from 0 due to the confidence interval of (-4,424,150, 17,900,000). The average effect, however, was still found to be significant. The estimate for the average effect was estimated as $11,100,000. This can be interpreted by saying that purses increase by, on average, $11,100,000 in the treatment groups due to the introduction of HHR. Due to the confidence interval of (25,199.27, 22,100,000) not containing 0, this estimate is found to be significantly different that 0.

4.4.2 *ln* Purses

Table 4.16 Robustness Results for the Natural Log of Purses

ln of Purses						
	Placebo 2	**Placebo 1**	**Effect 0**	**Effect 1**	**Effect 2**	**Average E**
Estimate	0.017	-0.025	0.250	0.682	0.783	0.552
SE	0.297	0.114	0.213	0.220	0.581	0.232
LB CI	-0.565	-0.248	-0.167	0.251	-0.356	0.098
UB CI	0.599	0.199	0.667	1.112	1.922	1.007
N	13	14	14	13	11	38
Switchers	4	4	4	4	3	11

Table 4.16 shows the difference-in-differences results with the addition of another placebo and dynamic effect. For Placebo 2, the estimate was found to be 0.017. However, Placebo 2 also had a confidence interval of (-0.565, 0.599) which contained 0 between its upper and lower bounds. Thus, this estimate is statistically insignificant from 0.

Effect 2 was found to have an estimate of 0.783, which indicates the elasticity of purses two years after the introduction of the treatment. However, this estimation for elasticity is statistically insignificant from 0, with a confidence interval of (-0.356, 1.922). The average effect was found to equal 0.552. The estimate is found to be statistically significant due to the fact that the confidence interval of (0.098, 1.007) does not contain 0 within its upper and lower bounds.

4.4.3 On-Track Handle

Table 4.17 Robustness Results for On-Track Handle

On-Track Handle						
	Placebo 2	Placebo 1	Effect 0	Effect 1	Effect 2	Average E
Estimate	472,808.20	1,743,184	1,884,915	4,826,826	8,124,504	4,656,407
SE	3,652,069	3,157,232	4,505,946	6,891,692	8,128,865	5,985,708
LB CI	-6,685,247	-4,444,991	-6,946,739	-8,680,889	-7,808,071	-7,075,580
UB CI	7,630,863	7,931,359	10,700,000	18,300,000	24,100,000	16,400,000
N	13	14	14	13	11	38
Switchers	4	4	4	4	3	11

Table 4.17 presents the difference-in-difference results using on-track handle as the outcome variable, as well as the addition of an extra placebo and dynamic effect. The estimate for Placebo 2 was found to equal $472,808.20. This estimate is not statistically different from 0 due to the confidence interval (-6,685,247, 7,630,863) not containing 0. Thus, this insignificant effect helps to show that live wagering at the tracks behaved in a parallel manner prior to the introduction of the treatment.

Effect 2 was found to equal $8,124,504. The confidence interval for this effect was found to be (-7,808,071, 24,100,000), which contains 0. Thus, the estimate for Effect 2 was found to be statistically insignificant. The average effect was estimated to be $4,656,407. This estimate was also statistically insignificant, due to the confidence interval (-7,075,580, 16,400,000) containing 0 within its bounds.

4.4.4 *ln* On-Track Handle

Table 4.18 Robustness Results for the Natural Log of On-Track Handle

ln of On-Track Handle						
	Placebo 2	Placebo 1	Effect 0	Effect 1	Effect 2	Average E
Estimate	-0.167	-0.119	-0.009	0.274	0.410	0.208
SE	0.231	0.076	0.151	0.208	0.353	0.204
LB CI	-0.620	-0.268	-0.305	-0.134	-0.282	-0.191
UB CI	0.285	0.030	0.288	0.681	1.102	0.607
N	13	14	14	13	11	38
Switchers	4	4	4	4	3	11

Table 4.18 presents the difference-in-differences results using the natural log of on-track handle as the outcome variable, with the addition of one more placebo and dynamic effect. Placebo 2 was found to have an estimate equal to -0.167, which was statistically insignificant. This was due to the fact that the confidence interval of (-0.620, 0.285) contained 0 within its upper and lower bounds. Thus, the tracks appear to behave in a similar manner in the two years prior to a treatment being introduced.

The estimate for effect 2 was found to be equal to 0.410. This again was found to be insignificant based on the confidence interval of (-0.282, 1.102). The average effect was found to be equal to 0.208. The confidence interval was found to equal (-0.191, 0.607), which contained 0. Therefore, this estimate is statistically insignificant.

CHAPTER 5. DISCUSSIONS AND CONCLUSIONS

This research offers an introductory view into the impact HHR has on the Thoroughbred racing industry in Kentucky. The analysis confirmed relationships that were expected to exist based on agreements on where HHR money is allocated throughout the state, but also leaves some relationships unanswered due to the early nature of this analysis.

The main objective of the study was to look at the relationship between HHR and established industry metrics. This was accomplished using OLS linear regressions to determine whether a linear relationship existed between HHR and other variables. In addition, a difference-in-differences approach was utilized to quantify the treatment effect, using HHR as the treatment and purses and on-track handle as the outcome variables of interest. This was pursued to begin to understand the overall effect HHR has on a track after it is implemented at a simulcast facility.

In the OLS analysis, HHR wagering and on-track handle are negatively related with no causal interpretation. The analysis also suggests a positive linear relationship between HHR_{t-1} and purses. This result was expected due to the nature of how HHR is allowed to operate within the state of Kentucky. Although not in legal writing, there is an informal “agreement” that a certain percentage of the total HHR wagered at each facility is allocated to purses at racetrack owning the HHR facility. In the long run, higher purses should result in larger field sizes and better-quality races, which could be beneficial in potentially increasing on-track wagering.

At this point, there is no statistically significant relationship between HHR wagering and the national mean and median yearling prices. Yearling prices are a function of expected earnings, so if HHR wagering increases purses, then an increase in expected earnings may results in an eventual increase in sales prices. It is too early to tell, based on this analysis, if this will be the case.

Additionally, it is too early to tell how HHR affects the yearly foal crop in Kentucky. If higher sale prices are realized, breeders may alter their production decisions, but this would be expected to have a lagged effect due to the biological constraints involved in breeding.

In the difference-in-differences analysis, only 4 out of twelve effect estimators are statistically significant. Given the lack of observations, only a broad interpretation should be made using these results until more data becomes available. All of the significant relationships were found to be related to purses.

According to this preliminary analysis, HHR wagering does appear to have a positive effect on purses in tracks where it is implemented when compared to tracks that have no HHR. This is expected, since a portion of a track's HHR revenue is supposed to be allocated to purses. However, it is not yet clear whether this will also result in an increase in on-track handle.

There is one major limitation to this study that must be understood. Only limited data are available. Many difference-in-difference models rely on large amounts of data to be able to complete their analysis. However, since there is only roughly 12 years of data available for the first track that introduced HHR, many of these effects have yet to fully be realized. With tracks like Churchill Downs introducing HHR beginning in 2018, the post treatment period of 1 year does not leave much room for analysis for treatment effects. Thus, it is important to remember that this study represents an initial analysis of the early impacts of HHR, as it is still continuing to grow and establish itself in Kentucky.

Further research should focus on revisiting this analysis once more data becomes available. In another 5 or 10 years, additional data would allow for more meaningful postestimation effects to be estimated. This would be useful in determining whether these initial effects were accurate in their estimations as well as understanding the long-term effects HHR will have on the Thoroughbred industry. Another extension of this project could include looking at HHR activities in other states to investigate whether similar relationships exist. Additionally, other metrics could be explored in the future, including field size, all sources handle, and average purse per race. By continuing to explore this relationship between HHR and the Thoroughbred industry, researchers may be able to understand if HHR will stall the decline in racing by enhancing the industry and bring a new audience to the racing world.

REFERENCES

2012 Kentucky Equine Survey (2013). Available online: https://equine.ca.uky.edu/files/2012_equine_survey_report_final_4.pdf.

Early Kentucky horse industry. (n.d.). International Museum of the Horse. Retrieved February 2, 2023, from https://imh.org/exhibits/past/legacy-of-the-horse/early-kentucky-horse-industry/#:~:text=In%201780%2C%20the%20first%20circular,English%20system%20of%20clock%2Dwise.

Ali, M. M. (1977). Probability and utility estimates for racetrack bettors. *Journal of Political Economy*, *85*(4), 803–815. https://about.jstor.org/terms

Bishop, M. (2000). And they're off: The legality of interstate pari-mutuel wagering and its impact on the thoroughbred horse industry. *Kentucky Law Journal*, *89*(3), 711–742. www.msnbo.com/news/472179.asp.

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001

Carter, H. O., Shepard, L. E., & Whitney, M. (1991). *An economic analysis of the California thoroughbred racing industry*.

de Chaisemartin, C., & D'Haultfœuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, *110*(9), 2964–2996. https://doi.org/https://doi.org/10.1257/aer.20181169

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277. https://doi.org/10.1016/j.jeconom.2021.03.014

Gramm, M., McKinney, C. N., Owens, D. H., & Ryan, M. E. (2007). What do bettors want? Determinants of pari-mutuel betting preference. *American Journal of Economics and Sociology*, *66*(3), 465–491.

Gramm, M., & Owens, D. H. (2006). Efficiency in pari-mutuel betting markets across wagering pools in the simulcast era. *Southern Economic Journal*, *72*(4), 926–937.

Gross, J., & Grob, J. (2004). *Linear regression*. Springer Science and Business Media.

Karungu, P., Reed, M., & Tvedt, D. (1993). Macroeconomic Factors and the Thoroughbred Industry. *Journal of Agricultural and Applied Economics*, *25*(1), 165–173. https://doi.org/10.1017/S1074070800018721

Kobin, B. (2021, September 30). What is historical horse racing and how do the betting machines work? *Courier Journal*. https://www.courier-journal.com/story/news/local/2021/09/30/churchill-downs-louisville-facility-what-historical-horse-racing/5929393001/

Kruse, H. (2002). Narrowcast technology, interactivity, and the economic relations of space: The case of horse race simulcasting. *New Media and Society*, *4*(3), 385–404.

Lambert, T. E. (2022). Kentucky and the thoroughbred industries: Prospects and challenges as gambling stagnates. *International Journal of Sport Policy and Politics*, *14*(1), 177–189. https://doi.org/10.1080/19406940.2021.1987292

Lambert, T. E. (2022). The economic impact of horse racing tracks and historical horse racing in Kentucky. *Faculty Scholarship*. https://ir.library.louisville.edu/faculty

Malinowski, K., & Avenatti, R. (2009). *Impact of slot machines/video lottery terminals (VLTs) on the economy, horse racing and breeding industry, agriculture and open space in states/provinces where they exist: Why is this important for New Jersey?*

Meeker, T. H. (1989). Thoroughbred racing-getting back on track. *Kentucky Law Journal*, *78*(3), 435–446.

Morgan, W. D., & Vasche, J. D. (1979). Horseracing demand, parimutuel taxation and state revenue potential. *National Tax Journal*, *32*(2), 185–194.

Neibergs, J. S., & Thalheimer, R. (1999). An economic analysis of the effectiveness of thoroughbred breeder/owner incentive policies. *Journal of Agricultural and Applied Economics*, *3*(3), 58–59.

Nicholson, J. C. (2012). *The Kentucky Derby: How the run for the roses became America's premier sporting event* (7th ed.). Sports Studies. https://uknowledge.uky.edu/upk_sports_studies/7

Pendleton, P. (2023, March 11). *The Lincoln county connection to counter clock racing*. WKYT. https://www.wkyt.com/2023/03/12/lincoln-county-connection-counter-clock-racing/

Ramachandran, K. M., & Tsokos, C. P. (2021). Categorical data analysis and goodness-of-fit tests and applications. In *Mathematical Statistics with Applications in R* (pp. 461–490). Academic Press. https://doi.org/10.1016/B978-0-12-817815-7.00011-7

Royston, P. (1982). An extension of Shapiro and Wilk's W test for normality to large samples. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, *31*(2), 115–124.

Royston, P. (1992). Approximating the Shapiro-Wilk W-test for non-normality. *Statistics and Computing*, *2*, 117–119.

Royston, P. (1993). A toolkit for testing for non-normality in complete and censored samples. *Journal of the Royal Statistical Society. Series D (The Statistician)*, *42*(1), 37–43. https://www.jstor.org/stable/2348109?seq=1&cid=pdf-

Rudd, D., Mills, R., & Flanegin, F. (2009). Racinos-the marriage of horse racetracks and casino/slots-style gambling-friends or foes? *Journal of Case Research in Business and Economics*, 28–35. https://www.researchgate.net/publication/237240992

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*(3), 591–611. https://www.jstor.org/stable/2333709

Simmons, S. A., & Sharp, R. (1987). State lotteries' effects on thoroughbred horse racing. *Journal of Policy Analysis and Management*, *6*(3), 446–448. https://www.jstor.org/stable/3324856

Suits, D. B. (1979). The elasticity of demand for gambling. *The Quarterly Journal of Economics*, *93*(1), 155–162. https://www.jstor.org/stable/1882605

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175–199. https://doi.org/10.1016/j.jeconom.2020.09.006

Thaler, R. H., & Ziemba, W. T. (1988). Anomalies parimutuel betting markets: Racetracks and lotteries. *Journal of Economic Perspectives*, *2*(2), 161–174.

Thalheimer, R. (2008). Government restrictions and the demand for casino parimutuel wagering. *Applied Economics*, *40*(6), 773–791. https://doi.org/10.1080/00036840600749672

Thalheimer, R., & Ali, M. M. (1995). The demand for parimutuel horse race wagering and attendance. *Management Science*, *41*(1), 129–143. https://www.jstor.org/stable/2632907

Thalheimer, R. (2012). The demand for slot machine and pari-mutuel horse race wagering at a racetrack-casino. *Applied Economics*, *44*(9), 1177–1191. https://doi.org/10.1080/00036846.2010.537645

Verbeek, M. (2017). Heteroskedasticity and autocorrelation. In *A Guide to Modern Econometrics* (5th ed., pp. 97–138). John Wiley and Sons, Inc.

VITA

Barrett Kerr

San Antonio, Texas

<u>Education</u>

B.S. in Agricultural Economics, May 2023 – University of Kentucky

<u>Professional Positions</u>

Student Assistant, October 2021 – May 2023
Stuckert Career Center, Lexington, KY