Winning Streaks and Housing Peaks

An investigation in professional sports success and real estate sale prices

Master's Thesis Real Estate Studies
Rick Roodbergen
S3698157
Supervisor: Dr. Xiaolong Liu

Abstract

This thesis studies the relationship between the success of the local NFL team and if this success changes the amount of influence the local sports stadium has on the surrounding real estate market. The incorporation of sports team athletic success adding to the current available literature on sports stadiums and real estate markets. The study evaluates the effects of a Super Bowl victory on the local real estate market surrounding the team's home stadium in three cities which won a Super Bowl in the last 20 years. Utilizing a difference-in-difference model, the thesis explores the possible correlation between the proximity of a sports stadium and housing prices, while also investigating the link between athletic success and housing market outcomes in the vicinity of the stadium. The results show a negative relationship in the city of Denver, indicating that in the year following the Super Bowl, houses in the direct vicinity of the stadium sold at a slight discount. No relationship was found in the cities of Pittsburgh and Tampa. Despite limitations in data sources and variables, the findings contribute to the understanding of the economic influence of sports stadiums on local housing markets and provide valuable insights for policymakers considering public funding for stadium construction.

Disclaimer

Table of Contents

Preface

When starting this Master's Thesis process I knew I was getting myself into a tough position opting for a creative and very niche topic, especially considering doing research has never been my strongest area of expertise. However the fact I did this has given me the opportunity to learn only more both about a topic which I find very fascinating as well as how to set up a proper research model. Having found a topic I am sincerely passionate about has aided me in the completion of this Master's Thesis. I have learned an incredible amount from the long process I have taken for this thesis. This long process being both of my own doing as well as the result of starting fulltime employment during the process. Having always specialized in the methodological portions of academic research in my working experience I have had to learn how to correctly find and define a good research topic. This has proven to be quite a challenge for me and I am very grateful that my supervisor Dr. Xiaolong Liu has shown incredible patience in helping me go forward with this project.

I would like to thank my supervisor Dr. Xialong Liu for the large amounts of patience and feedback I have needed in order to get on my way with my current topic. The very timely responses of great quality have truly set me up to succeed in this difficult Master's Thesis process. I understand I have not always been the most timely myself, but always knowing I could get a very detailed response on very short notice has helped me whenever I needed it.

I would also like to thank the Geodienst, as for their assistance I would not have been able to find the datasets and geographical alterations necessary to fulfill the analysis portion of the research.

The following thesis is organized into 5 chapters. Starting with the introduction to the topic, following with the theoretical framework and methodology of the study. After the methodology the results will be presented and concluding with a discussion and research implications.

1 Introduction

1.1 Motivation

Last Super Bowl, Super Bowl LVII(57) was watched by around 200 million people(NFL, 2023), which is around 60 percent of the entire population of the United States. This amount of viewers is not new to American television records as, of the 30 most watched television broadcasts in the United States, 22 are all Super Bowl broadcasts. The Super Bowl is therefore one of the most important sporting events in the United States. Not only in the amount of viewers does the Super Bowl show large numbers. Also in economic traffic in both the region where the match happens as well as the football teams which participate or win the Super Bowl. This increased interest in the teams participating in the ultimate match of the NFL season leads to an increase in economic activity in the home region. This increased economic activity is among one of the factors leading to an increase in demand in real estate around the stadium.

The main way in which American football teams can differentiate from each other is through their success on the field. The connection between athletic success and economic activity is not new but it is rarely studied. Previous studies on per capita income(Coates & Humphreys, 2002) and the financial market (Payne, Tresl & Friesen, 2018) have shown that increased success athletically can increase economic factors in the surrounding area. The possible connection between athletic success and real estate prices is a territory which has yet to be studied in depth. Providing a currently missing in-depth analysis is what this thesis aims to provide.

A consistent factor in all academic literature on the topic is the lack of expansion on the possible differentiating effects which a sports franchise can have on the local market. Athletic success being the prime example of the way for sports franchises to stand out from the pack. In order to better understand the reasoning behind the prevalent differing results in the academic literature, this thesis aims to study the further workings of the relationship between sports stadium and housing prices by incorporating athletic success.

One of the main aspects of the local economy which is affected by the presence of sports stadiums is the real estate market. Both the positive and the negative amenities which a sports stadium can provide to the local real estate market have been discussed and researched extensively in existing literature. On the positive side, previous literature has suggested various slight positive effects on the local real estate market. From increased economic activity to a change in the mix of services provided in the neighborhood and an increase in local welfare(Humphreys & Zhou, 2015). Increased interest in the local team would lead to an increased exposure of these positive amenities which the stadiums provide, this increased exposure could lead to an increase of the perceived amenity of the stadium which would subsequently lead to an increase in demand for housing in the local region.

The presence of an American football stadium in the area can be convenient for fans of the local team and can be a source of great amenity. However it should not be understated that proven

connections have been found between sports stadiums and several disamenities which could hamper this price growth (Locke, 2019, Montolio & Planells-Struse, 2016). Additionally, recent literature has found significant correlations between the presence of sporting events with increased congestion and therefore CO_2 emissions(Humphreys & Pyun, 2018), crime (Montolio & Planells-Struse, 2016) and air pollution(Locke, 2019). It is quite likely and sometimes stated directly by other researchers that these increased negative externalities are the reason for lower housing transaction prices in the area around a sports stadium(Humphreys & Nowak, 2017. Joshi, Horn & Berrens, 2020). By the same consequences of an increased interest in the local teams being a driver for positive influences, this increased interest can lead to negative influences. Increased interest and especially increased traffic to and from games can be a major driver for increased experienced disamenity from the local NFL stadium. This increased disamenity would subsequently lead to a lowering of housing prices in the near vicinity of the stadium.

It is not only the increased interest and exposure of the home team that is able to fuel an increased housing price of properties surrounding a stadium. A victory in the Super Bowl has shown to be a driver for increased economic activity in the home region. This increased economic activity is also able to be a catalyst for increase in demand, especially considering that investment climate in the real estate industry increases in the year following a Super Bowl victory(Payne, Tresl & Friesen, 2018)

Gaining a deep understanding in the impacts of sports stadia is increasingly important in current times due to the increased amounts of public funds allocated towards new stadium construction. In Las Vegas(Akers, 2018) and Atlanta(deMause, 2017) new stadiums have been constructed both with sizable contributions from their respective governmental bodies. This thesis aims to add into the literature by gaining more understanding in the impacts of these sports stadiums in the local market. Additionally this thesis aims to find the impacts of athletic success of a local sports team on the local market.

Due to these links named before, the link between winning a super bowl and house pricing in the area is plausible and not often studied in current academic discourse on this topic. However, research on the precise workings of this particular link is missing and could add layers of information in multiple ways. Through the deeper understanding on the question if the area around the stadium is affected by athletic performance, will aid lawmakers and other governmental decision makers in further funding of construction of new sports stadiums. On the other hand, a deeper understanding of the effects of the super bowl could possibly add another layer of economic effect which the most watched yearly televised event in the United States has on the surrounding area and therefore host city proper.

1.2 Academic Relevance

A way in which the economic development of an area can be seen is through the sale prices of local real estate. Other studies have already taken a look at how the presence or departure of a sports stadium influences the local real estate market in the immediate vicinity, with varying results. The most common results of these studies will be provided in the chapter below.

Currently the academic literature on the topic of sports success and economic effects has only scratched the surface. With the most important article not focusing on the real estate market but on income. The seminal study by Coates & Humphreys(2002) finds a positive relationship between success of the local American football team and the income of workers in the following year. Finding that a victory in the super bowl temporarily increases the per capita income of local workers. This finding has proven to be the starting point for the build of literature of sports success and the local economic market.

While research has pointed both ways when it comes to the actual connection between a property's value and the presence of a professional sports team, very little research has been done when it comes to the more precise possible interactions within this connection. A particular research gap has developed between the athletic success of a sports team and the effect on the sports stadium amenity in the real estate market.

Studying this currently existing gap in the literature will give a deeper understanding of the effect that sports stadiums have on the surrounding neighborhood. Specifically the inclusion of athletic success into the research will improve on the existing academic literature on the relationship of sports stadiums and real estate as well as the literature on the relationship between athletic success of the local professional sports team and the local economy. Deeper insight into the effects of a sports stadium on the surrounding area should aid lawmakers in future decision-making when considering stadium construction, especially in a time where very substantial sums are allocated for the construction of these stadiums.

To commence, the study of the factors which influence housing prices has quite the accumulated base of literature since the seminal work of Rosen (1974) introducing the hedonic pricing method.

The current academic consensus largely finds that a sports stadium does provide a positive amenity effect to the real estate pricing in the vicinity(Propheter, 2019. Tu, 2005. Bradbury, Coates & Humphreys, 2023. Keeler, Stephens & Humphreys. 2021). Current academic literature does not go deeper into the inner workings of this relationship or taking into effect other factors like athletic success or the other amenities a stadium could offer like educational facilities suggested by Tu(2005), some logical explanations for the existence of the effect are provided but not studied outright(Tu, 2005. Feng & Humphreys, 2007). It is exactly the explanation and dynamics of this relationship which provides the research gap that this thesis aims to append.

Additionally to finding out the research gap from the connection between athletic success and housing prices through the stadium amenity, this thesis also aims to explore the sports stadium amenity effect by studying multiple stadiums.

Current existing literature on the topic of sports stadiums and real estate prices are very similar on a methodological level through the study of a singular stadium. Multiple studies study singular moments in time, particularly at moments where sports teams either leave a stadium to move to a new location or when a new stadium is constructed and a professional sports team moves into this stadium for their home games(Tu, 2005. Humphreys & Nowak, 2015. Hyun, 2022). This very specific look at a stadium may give a skewed look in regards to local context, the studies all examine the newest amenity or the recent loss of an amenity which may provide differing results to the normal situation if a stadium is present for a longer period of time. These studies take a hedonic analysis approach for the valuation of the amenity of the stadium.

Some studies which look at multiple cities are present in the current body of literature, these studies often look at large scale samples of all stadiums in the United States and looking at the average price of real estate in a census block to reach conclusions(Feng & Humphreys, 2012. Huang & Humphreys, 2014). This approach does allow for a more complete look at the connection between housing and sports stadia, however through the usage of census block averages and the sizable scale of the analyses it is near impossible to take into account local context like sports team success.

1.3 Research Problem Statement

Most of the other studies in this field lack certain aspects which will aid in understanding the relationship between sports stadiums and housing sale prices. Some studies have quite the limited analysis size of only a single city(Tu, 2005. Propheter, 2019. Humphreys & Nowak, 2015. Hyun, 2022) while others which do take up larger analyses lack the per-unit approach(Huang & Humphreys, 2014. Feng & Humphreys, 2012).

Therefore, this thesis aims to study multiple cities on a per-unit level, which will attempt to give a more robust overview of the relationship between sports stadiums and real estate sale prices. The hypothesis of this study is that athletic success will enhance the positive effect that the sports stadium has on the real estate sale prices in the stadium's vicinity. The expectation is that the perceived amenity value of the stadium near the dwelling will increase in the year where the team has been highly successful athletically.

As the sports stadium amenity theory has been studied to great extent, often proving a positive correlation between distance to a stadium and housing prices. Additionally, the positive effects of a Super Bowl victory, the pinnacle of athletic success in American football, has been proven to have a positive effect on the local economy before(Sommers, 2000. Payne, Tresl & Friesen, 2018. Du & Zhang, 2022). The only link missing in this particular field of research is the link between athletic success and housing prices which is what this thesis aims to accomplish.

In the last multiple decades multiple new major league sports venues were being constructed, with the effects on the social and economic front still being debated in academic discourse. The same debate ongoing then is starting to pick back up in the current times. One of the main reasons why the study of sports stadiums and their effect on housing prices had been lacking in the past was the inability for researchers to separate the impact of the sports amenity with the surrounding housing prices. Oftentimes stadiums were constructed in new communities or very near to the previous sports venue, which proved empirically challenging to analyze the effect. In order to differentiate between the positive amenity effect of a newly constructed stadium and the dynamics of team success on the local housing market this study only focuses on stadiums which have been standing on their locations for several years. Expansion on the process of city selection is found in section 3.1.

Hence, this thesis aims to fill this research gap using the following main research question. This research design will allow the research to fully focus on the precise relationship between athletic success, measured in a victory in the Super Bowl, and the surrounding transaction prices of homes. The main research question of this thesis is as follows:

1. How does a victory in the Super Bowl influence the sports stadium amenity effect around the home stadium?

The rest of this study is ordered as follows; first the theoretical framework and theoretical history of this topic is outlined. Next the methodology for this thesis is expanded upon. Then the data and results of the analysis are outlined and discussed and at the end a final conclusion is drawn which aims to answer the outlined research question.

2. Theoretical Framework

Reasons given for the improvement in housing values in the surrounding area of the stadium include the improvement of infrastructure in the area and increased employment opportunities. The recreational accommodations and learning facilities in the stadium are also given as possible reasons for the decrease in negative effect from the location of the stadium, the stadium has more purpose in the surrounding neighborhoods than only the home field for the local NFL team(Tu, 2005). The overall amenity effect of the stadium is expected to increase once a team wins a Super Bowl. A home team reaching the highest point of athletic success in their sport will bring a lot of increased exposure to the team, this increased exposure could lead to an increased experience of the amenity factors which an American football stadium offers.

A victory in the Super Bowl is shown to be one of the best things to happen both to the local fanbase (Menefee, 2024) and the local economic market(Slobe, 2024). This increased interest in the team is able to be a driver throughout the economic market, including the real estate market.

Measuring athletic success and their economic consequences has been done before in the study of per capita income and postseason play in professional sports. A slightly surprising result within the study shows that in the year after a super bowl win the per capita income in the city that won the super bowl goes up. This finding is interesting especially due to the fact that this research shows that winning the super bowl, which due to the nature of the Super Bowl being often played in a neutral site, has economic benefits for the city of the home team(Coates & Humphreys, 2002). The connection that this study provides indicates that, even though the actual event is played elsewhere, the team winning the Super Bowl still experiences the positive economic effects in the home city.

A victory in the Super Bowl has proven to not only be positive for the per capita income in the winning city. A commonly known effect in the financial sector is the Super Bowl effect, where in anticipation of the biggest game of the American football season the amount of public investment in companies geographically connected to the participating teams will increase. The success of even reaching the super bowl, let alone winning this match, does not only keep to the sports teams associated with the match but the city entirely. This shows the fact that positive economic effects spillover into the cities of participating teams and increasingly so for victorious cities. (Payne, Tresl & Friesen, 2018). Du & Zhang (2021) particularly find an increase in investment in the FIRE(Finance, Insurance and Real Estate)sectors when a sports team is experiencing success on the field. The increase in investment in the home city is expected to particularly center around the calling card of the victorious NFL team which is their home stadium. Increased investment in the home market of the Super Bowl winning team will increase demand which will subsequently result in a temporary price increase. Both the studies which found a link between economic factors and athletic success have found a very temporary effect, which has informed this study in the temporary nature of athletic success and economic drivers. This is reflected in the methods of the analysis.

One of the earliest pieces of literature made on the connection between sports stadiums and housing prices was done by Tu (2005). Primarily, the study finds that the existence of a professional sports stadium is a negative influence on the transaction price of single-family homes. These results do come with an important additional angle considering the history of the case studied in the work by Tu(2005). Since the team responsible for stadium construction previously encountered fierce resistance from surrounding neighborhoods in areas with higher property values it could be concluded that the stadium construction location was based on these lower property values in the surrounding neighborhoods. This is also seen in the results with properties in the area near the newly constructed stadium to hold lower property values. By comparing the housing values before, during and after the construction of the stadium the study finds a reduction in the negative impact of the location of the property on the sale price.

Other research on a larger scale, where all professional sports stadiums and arenas were studied, performed by Feng & Humphreys (2012) shows the persistent nature of sports facilities and their positive effects on housing prices in the surrounding neighborhoods. This study as well as similar works (Huang & Humphreys, 2014) do not take into account the singular home sales, these studies focus on average home sales of neighborhoods. The study encompasses every home location of a professional sports franchise in the United States' largest four sports(American football, basketball, baseball and ice hockey). The study also finds a positive correlation between housing transaction prices and distance to a stadium. Meaning the further away from the stadium a property is located the lower the price of said property. This is an important finding since this research establishes a clear link between sports stadiums and housing prices.

The study by Feng & Humphreys represents the other portion of research performed in this field of study, where the effect is investigated through higher level aggregation groups instead of unit-level analysis. This level of analysis allows for a more general understanding of the effect between sports stadiums and housing prices, with the tradeoff being that much of the local context and specific workings of this relationship are disregarded or out of scope of the analysis.

Matheson (2019) comes to similar conclusions as Feng & Humphreys(2012). Matheson(2019) finds while reviewing most of the available literature on the topic of housing prices and sports stadiums that there is more than just the stadium which provides an increased amenity for the surrounding neighborhood. Important notes that are made in this comprehensive essay is the neighborhood development aspect of stadium manufacturing. Professional sports stadiums and arenas are able to function as an anchor for local economic development. This local development could be one of the main causes for the housing sale price increase. This finding informed the thesis that, to eliminate bias from this neighborhood development aspect, only stadiums which have been standing for a sizable amount of time will be considered in this research.

It is important to note that not all research agrees with the finding of Matheson (2019). Hyun (2021), for example, uses a case study of a baseball stadium to find that the positive externalities generated by the economic development or increased investment in infrastructure do not always outperform the negative externalities generated by the stadium. These negative externalities are identified as increased congestion and noise or light pollution. Humphreys & Nowak (2015) find

similar results in the excess disamenities on the housing market generated by the outgoing Seattle basketball team. The overall conclusions of these papers show that the connection between sports stadiums and the local housing market is increasingly complex and warrants careful consideration and study. Especially considering the wider implications which these sports stadiums hold on the local real estate market.

Considering the relationship which has been found often between the sports stadiums and housing sale prices. It is increasingly interesting for research as the exact workings of the relationship are still largely unknown. Considering the available literature suggests mostly positive but also negative or an absence of a relationship this indicates not all is known about this relationship. Gaining understanding on this relationship is increasingly important as the amounts of subsidies for the developments of these home stadiums are only growing. In order to understand what drives the amenity effect of a sports stadium on local housing markets new angles need to be taken into consideration. Athletic success of the local professional sports team has been the driver for other economic factors, as will be discussed in the next portion of this chapter. Considering the impact of athletic success has the potential to give insight into the precise connection between sports stadiums and the local housing market.

3. Methodology

3.1 Data source

The criteria used for the cities chosen to be included in this research are as follows; The local professional NFL team of the city needs to have won the Super Bowl in the last two decades, The sports stadium in question needed to have sufficient housing in the vicinity. For example, this meant that Gillette Stadium in Foxborough was excluded from the research as very little housing exists in the vicinity of the stadium. Lastly, in order to not introduce unwanted variance in the research the stadium needs to be the home stadium of a single NFL team, as the amenity effect on housing sales of hosting two teams should be considered quite different opposed to stadiums hosting a single NFL team. This consideration excluded both New York and Los Angeles stadiums as one of the teams in both stadiums have won a Super Bowl(New York Giants and Los Angeles Rams resp.) and host two teams in their stadium. Public data availability played the largest role after these considerations.
The three cities that apply for all these considerations and therefore will be used in this study are the cities of Pittsburgh, Tampa & Denver.

The Pittsburgh Steelers, having won their most recent Super Bowl in 2008 and their second most recent in 2006, were established in 1933. The stadium of the Steelers is the Acrisure stadium, which was formerly known as Heinz Field. Acrisure Stadium was built in 2001 and has a seating capacity of around 65 thousand(Steelers, 2008). The stadium is located in the North Shore neighborhood of Pittsburgh, on the northern shore of the Ohio and Allegheny river. Both Super Bowl victories are included in the analysis of the research.

The Tampa Bay Buccaneers won their most recent Super Bowl in 2021. The team joined the NFL in 1976 as an expansion team. The Buccaneers play their matches in Raymond James stadium, which was built in 1998. The stadium has a seating capacity of around 69 thousand(Raymond James Stadium, date unknown). The stadium is located in the West Tampa neighborhood.

The Denver Broncos are the third team included in the analysis. The Broncos won their most recent Super Bowl in 2015. The Broncos were established in 1959(Pro Football Hall of Fame, date unknown). The home games of the Denver Broncos are played at Empower Field At Mile High Stadium. The stadium was built in 2001 and has a seating capacity for 76 thousand people (Empower Field at Mile High, date unknown).

Data of housing sales and housing characteristics is sourced from various regional data centers across the three cities included in this research. Data from Denver is sourced from the Denver Open Data Catalog. The owner of the data is the City and County of Denver Assessment Division which has published the data under an Open Data License. Data from Allegheny County, the county of Pittsburgh, is distributed by the Western Pennsylvania Regional Data Center. Where the data is published by Allegheny County under an Open Data License. Data from the city of Tampa is published by the Hillsborough County Property Appraiser and available for use by the

public. Additionally, information on locational characteristics is sourced nationwide from census data sourced by the Esri Living Atlas portal, which is licensed by the University of Groningen.

3.2 Model Specification

This thesis will be using a variation on the hedonic pricing model introduced by Rosen (1974). The model will allow for the quantification of previously non-priced amenities into the housing sale price, through this method it becomes possible to find the attached value of the sports stadium being in the vicinity of a dwelling while also accounting for other locational attributes which the dwelling possesses .

Surprisingly absent in the current literature on the sports stadium amenity theory which focuses on more than one stadium within the research framework while also taking a unit-level sales data approach. This research aims to fill this methodological gap in the research, looking at multiple cities while also taking a unit-level approach.

The variation of the hedonic model which this thesis will use is the difference-in-difference model. Through the use of the difference-in-difference model a quasi-experimental research method is made possible. This method allows the researcher to do very specific research into the effects of an intervention, or in this case an event, in situations where ordinary experimentation is not. The method originates from the field of public health research as a way of researching causal relationships where a randomized trial is not possible or ethical(Wing, Simon & Bello-Gomez, 2018). This kind of situation where a randomized experimental approach is not available is also present in the research on real estate, and especially in the research on amenity effects on housing prices as a lot of property values are very context dependent. The usage of difference-in-difference modeling has been quite frequent in recent years in the real estate research sphere(Fang, 2021. Chun-Chang, Chi-Ming & Hui-Chuan, 2020) .

Considering the already existing literature and the goals of this research the model of this thesis is defined as follows:

$$Y_{it} = \alpha + \beta 1 Treated + \beta 2 Post + \beta 3 (Treated * Post) + \beta 4 \theta_{it} + \lambda_i + \Delta_t + \varepsilon_{it}$$

Equation 1. Regression equation of the difference-in-difference model

Y is the natural logarithm of the sale price of a property *i* in year *t*. Treated is defined as a binary variable being 1 if the property is located within 2.5 kilometers of the home stadium and 0 if the property is not located within this buffer zone. The control area for the analysis is 5 kilometers. These area sizes were chosen as a result of extensive analysis of several area sizes. Due to the very sizable parking areas surrounding all three stadiums the amount of housing in the direct vicinity is quite low. This has prompted this research to expand the treatment area compared to other research

$\beta2Post$ is the variable in the model defining the time after the Super Bowl victory. This variable is equally defined as binary variable, being 1 if the property is sold in the year after the Super Bowl victory and 0 if the property is not sold within that specific year. To take the Tampa Super Bowl victory as an example, for the full year after the Super Bowl victory (February 7th 2021 until February 7th 2022) the Post variable is defined as 1. Any transaction outside of this timeframe is defined as 0. The time after the Super Bowl which is decided due to multiple factors, previous research uses this treatment time in similar work performed by Coates & Humphreys (2002) who equally use a year as the treatment period. Additionally as discussed in the previous chapter the economic effects of athletic success have shown to be quite temporary phenomena, which is why such a temporary treatment effect is assumed within this model. The usage of a year also makes logical sense as the frequency of the Super Bowl is annual, making a team only the winning team for a single year after which a newer Super Bowl champion is crowned. The increased interest for the Super Bowl winner which is assumed to be a driver for the real estate price increase will therefore presumably move to the newer Super Bowl champion.

$\beta4\theta_{it}$ stands for the sum of products of structural control variables and their respective coefficients. See Table 1 for the full list and summary statistics of the variables used in the model

λ_i stands for the census tract level location fixed effects included in the model.

Δ_t stands for the year fixed effects which are included in the model.

Table 1. Summary Statistics empirical data of housing prices around Denver, Pittsburgh and Tampa.

City	Denver (N = 63,213)		Pittsburgh (N = 18,954)		Tampa (N = 37,960)	
	Mean	St.Dev	Mean	St.Dev	Mean	St.Dev
Log of sale	12.91838	.8641998	11.02913	1.194631	12.19707	.8713606
Property sold within year of Super Bowl	.0535808		.0656853		.016333	
Property within treated area	.3235727		.1637121		.1623815	
Interaction Post*Treated						
0 | 0	.6396944		.7794133		.8236301	
0 | 1	.3067249		.1549013		.1600369	
1 | 0	.036733		.0568745		.0139884	
1 | 1	.0168478		.0088108		.0023446	
Property area(Ft2)	1317.178	760.4613	1524.536	542.9903	.110409	.0737039
Bedrooms	2.191242	.9059335	2.791126	.7628692	2.564594	1.178408
Bathrooms	1.8198	.8079755	1.189353	.4414432	1.855532	.9983949
Stories	1.543369	.7825783	1.860125	.4435889	1.348815	.9100362
Property distance to CBD	3046.837	1428.289	3854.224	1413.793	5357.318	2231.9

Notes: The sample used in this research includes data from the time period of 2004 to 2024. Price is measured as the log of the sale price. The treatment groups are defined as being properties located within 2.5 kilometers of their respective stadium. The control group is defined as properties located between 2.5 and 5 kilometers of the respective stadium. The time period for the treatment is defined as the exact year after the victory in the Super Bowl.
Square footage is defined in square feet. Notable exception is the square footage in Tampa which is formatted differently due to data availability limitations.
The table depicts mean values for all variables and standard deviations for all numerical variables.

3.3 Empirical data

Noted in Table 1 above are the summary statistics for the data used in the analysis. The variables picked within the study are mostly dependent on previous hedonic models and data availability. The study period of this thesis is defined as the past 20 years. This allows for a large enough time

period to select multiple cities in the analysis using the stringent conditions while also reserving a sizable dataset for the analysis portion. Time periods of more than 10 years are more common in existing academic literature, being employed by Humphreys & Nowak (2015) and Chikish, Humphreys & Nowak (2018).

The summary statistics provided in Table 1 above give an overview of the data used in the difference in difference analysis of this thesis. The outcome variable appears normal with no significant outliers, the independent variable of this analysis is the log of housing sale prices. Notable in Table 1 is the very equal means and standard deviations across the three cities in this analysis, which is a testament to the consistent nature of the real estate market across the three cities.

The number of rows of data covered by both the treatment and post groups are 1,065(Denver), 77(Pittsburgh) and 89(Tampa). A surprising result is the high amount of identifying observations in the city of Denver. Further analysis of the data shows that this difference in identifying observations is due to a significantly higher housing density around the stadium in all directions and a lack of other objects which obstruct the amount of housing possible in the area. For example the amount of housing in the area of Raymond James stadium in Tampa is limited due to the location of Tampa International Airport being very close to the stadium. A similar situation in Pittsburgh occurs where the Allegheny Center and the Ohio river form obstacles for housing and therefore there are less housing transactions in the immediate vicinity of this stadium.

Housing in Denver generally shows to be smaller in floor space and amount of bedrooms while the housing in Pittsburgh shows to be relatively larger. Property areas in Tampa show a very different property area, while documentation from the data source states that the property size is measured in square feet this is clearly not the case. In order to not introduce any unwanted errors the decision was made to leave the variable as is without further alterations apart from general data cleaning as clear documentation of the conversion is lacking from the data source. The distance to CBD variable show significant differences, however due to the different placements of the stadiums across the different cities this is expected and of no specific concern for the analysis.

The approach for data cleaning is characterized by the attempt to attain uniformity across the three different data sources in order to achieve comparable results. This included only looking at real estate units with a primary housing attribute and using equal measurement levels where possible. This meant all real estate properties which did not solely hold a residential purpose were cut. Additionally as much as possible equal measurement levels were obtained, this has been successful apart from the property size variable in Tampa. Due to insufficient documentation there was no possibility to correctly convert this data to be similar to the other property size data.

The data was cleaned on all specifications of hedonic pricing methods in order to attain the best results possible from the empirical analysis. In order to prevent outliers all measurable variables in the models were cleaned by shaving off the bottom and top percentage of the total dataset. This meant the top and bottom percentage of lot size, bed- and bathrooms, stories, sale price

and distance to the CBD. This method was preferred as most data entry and calculation errors were handled in a single repeatable action. Two additional measures of data cleaning were taken. The log of sale price displayed large numbers between 0 and 5, all these data entries were deleted as further investigation showed that all these entries represented non-market transactions with sale prices between 0 and 100.000. Deleting these allowed the dependent variable to be normally distributed. The second specific data cleaning measure taken deals with the property size where instead of the top 1% of properties were shaved off instead the top 3% were deleted. This is done due to the large amount of outlying data points which skewed the results of the analysis.

4. Results and Discussion

4.1 Empirical model results

Empirical results of the various models among the three cities are shown below. Within the results, the different models are representative of the different cities used in the analysis(Denver, Pittsburgh and Tampa resp.). Additionally, next to the variables shown in the tables below, year fixed effects and location fixed effects based on census tract of the 2020 census are included in the model. The identifying observations show the amount of housing units which are both within the 1.5 kilometer zone and sold within the year after the Super Bowl victory for each city.

Table 2. Results of the difference-in-difference model in all three cities(Made by author).

	Difference-in-Difference Results		
	Denver	Pittsburgh	Tampa
Variable			
Treated	.0848829 ***	.2398317 ***	.1692832 ***
	(.0138129)	(0.0445654)	(.0210917)
Post	-.0350416	.1264309	.0083354
	(.0335941)	(.074904)	.0288585)
Treated x Post	-.0956891***	-.1078007	-.0487759
	(.0242992)	(.078864)	(.0668516)
Identifying observations	1,065	167	89
Housing characteristics	Yes	Yes	Yes
Distance to CBD	Yes	Yes	Yes
Census Tract FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	101,089	18,954	37,956
R^2	0.4701	0.4110	0.5595
Adj. R^2	0.4691	0.4079	0.5586

Statistical significance is noted by: ***$p <.01$, **$p < 0.05$, *$p < 0.1$

Notes: The table depicts the results of the difference-in-difference model performed. The dependent variable is the logarithm of housing price. The DiD estimate is captured by the interaction term Treated x Post. The identifying observations are all identified observations where the property lies within 2.5 kilometers of the stadium and is sold within 1 year of the Super Bowl victory. Robust standard errors are reported in parentheses. Control variables include: property size, number of bedrooms, number of full bathrooms, amount of stories.

The results for the Treated variable show a consistent pattern, the Treated variable in the model is statistically significant for Denver, Pittsburgh and Tampa on the 1% level. The coefficient of the Treatment variable is positive indicating a positive relationship between this treatment area and sale price of the property in all three cities.

Another more consistent result within the models presented in Table 1 is the non-significance of the "post" variable. In all three models the binary variable indicates if a property is sold within the first year of a Super Bowl victory is not shown to be statistically significant to the log of sale price. This means that specifically in the year after the home NFL team won the Super Bowl, other than the normal yearly effects of the real estate market, no increased 'boost' or temporary increase in prices happened in the neighborhood around each stadium.

The third and last main variable to specifically analyze in these empirical results is the interaction Post*Treatment. The variable is aimed to investigate the effect of a Super Bowl victory specifically within the 1.5 kilometer buffer area around each of the three investigated stadiums on housing prices. The interaction variable of Treated*Post is shown to not be statistically significant when compared to the log of sale price in Pittsburgh and Tampa. While in the city of Denver the Treated*Post interaction shows a statistically significant negative relationship with the log of sale price. This would indicate that in Denver in the year after the Super Bowl housing in the area around the stadium sold at a slight discount compared to other years in that specific area.

In the next chapter the results found in the analysis will be examined through a multitude of sensitivity analyses. After which a wider discussion on the causes and consequences of the found results will be given.

4.2 Sensitivity Analysis

In this section a set of sensitivity analyses are performed to assess the robustness of the DID-model performed in this thesis. The analyses include variations in: Time window, treatment area and fixed effects variations.

Time window variations to using different treatment periods through the "Post" variable were tested by using longer and shorter periods for the "post" variable. The results remained consistent indicating that the results are not sensitive to the definition of the given time period.

Treatment area variations were also performed to test if the results kept consistent when treatment area grew larger and smaller. Results did not remain consistent when analyzing smaller and larger treatment areas. The inconsistent results show in the larger treatment area(3km). This is particularly the case in the results out of the city of Denver. The cause for this inconsistency could lie in the fact that the whole dataset is comprised of all points within 5 kilometers of the stadium. Making the treatment area larger will include the majority of the real estate properties in the dataset which will make the results inconsistent.

Fixed effect variations were performed to find if the results of the analysis stayed consistent when using smaller location fixed effects. Results stayed consistent when using smaller location fixed effects showing that the results of the analysis are not skewed by the location fixed effects used in the analysis.

4.3 Discussion of results

This research studies the connection between sports stadia and housing prices in the year after a Super Bowl victory by the local team in order to find if athletic success from the local team influences the amenity value of the stadium on housing sale prices. Using a difference-in-difference model the sale price of houses within 2.5 kilometers of the American Football stadium is analyzed. The study finds a consistent positive effect from the sports stadium on the housing price while not finding a positive effect from a victory in the Super Bowl on the housing sale price. The effect of other housing structural attributes are consistent with the hedonic housing literature.

Continue on the results in each city and how they stack up against each other. The results show that in both Pittsburgh and Tampa no relation was found between the Treatment*Post interaction and sale price. This result would indicate that home-buyers do not show increased optimism on the neighborhood or consider the stadium to be more of a local amenity in a year after the local team wins the championship compared to other years in these two cities. While in Denver a statistically significant negative relationship was found between a victory in the Super Bowl and the housing prices in the direct vicinity of the stadium. This would indicate that the houses around the stadium would sell at a slight discount. Various reasons can be found for this particular finding.

Negative effects caused by professional sports stadiums are already proven by previous academic study. Particularly negative effects correlated by the presence of a sports stadium are found to be increased crime, air pollution and congestion(Humphreys & Pyun, 2018. Montolio & Planells-Struse, 2016. Locke, 2019). The latter two of these causes could be influenced by an increase in popularity of the home team after winning a Super Bowl. In the short term these increases in disamenities could lead to less demand for housing in the vicinity of the stadium. A possible cause for the negative relationship found in the analysis could therefore be the increase in disamenities outweighing the positive experienced amenity from the stadium.

Another possible cause for the negative relationship found in the analysis could be a shift in focus from government following the large victory for the area. After the local team has succeeded athletically there could be a possibility that both local government start to look elsewhere for the development of the city. Demand might change due to changing development patterns by the local lawmakers which could cause this temporary downward shift in housing demand and subsequently housing sale prices(Drobne & Bogataj, 2017. Gyourko & Molloy, 2015).

It is important to acknowledge the inconsistent nature of these results across the different cities, as this is one of the first works of research where multiple cities are investigated in a difference-in-difference approach. These results, which show that in Denver a negative relationship is found while no relationship is found in the other two cities, show the exact reason why local context and situation holds a very large influence on the results of research. The current body of literature available on this topic is mainly based on case studies of single stadiums, the results of this analysis show that these results can be inconsistent and context is key when it comes to understanding the role of stadiums in the real estate market.

4.4 Limitations

The extensive size of the study coupled with using various data sources does provide multiple challenges to the analysis. These mainly pertain to the analysis portion of the research.

Due to the multiple data sources used in the various cities in the analysis, data is not available in equal formats or accuracies. For every city additional formatting, geocoding and additional measurements had to be taken to equalize the data in order to perform data analysis. The non-equality of the number of variables present also led to a low amount of uniformly available variables included in the datasets for hedonic analysis, this is particularly in the difference in structural attribute information. For this research, priority was given to the interaction of treatment period and treatment area over achieving the highest proportion of explained variance, as the main focus of the analysis laid on this specific interaction. This has led to relatively low R-squared values. However the study by Humphreys & Nowak(2017) has achieved a similar R-squared value of around 0.5. This lower explained variance could lead to a distorted representation of reality in the results and therefore inaccurate results. Especially considering other hedonic analyses on the real estate market which achieve much higher levels of R-squared.

The non-uniformity of the different data is particularly evident in the lot size variable in Tampa. Where unfortunately due to insufficient documentation no formatting could be done to equalize the data without making errors.

One of the main questions that come as a result of this analysis is why the city of Denver does find a decisive result and the cities of Pittsburgh and Tampa does not. When discussing this question it is important to note the difference in identifying variables in the dataset. Due to the urban structure around Mile High Stadium in Denver being more residential compared to Tampa and Pittsburgh it is possible that the small amount of identifying observations have a significant impact on the results in Pittsburgh and Tampa. The difference in identifying observations between the cities is seen as one of the main limitations of the study.

Another important limitation to note is the limitations found in the sensitivity analyses, particularly in Denver. Especially when variations in the treatment area are applied the results of the analysis turn inconsistent. Further research considering the effect of sports stadiums should be considered to fully grasp the extent of the negative relationship found in the city of Denver.

5. Conclusion

This thesis aimed to find a relationship between a victory in the Super Bowl and real estate sales prices. The hypothesis for this thesis which was derived from previous literature in that this thesis would find a positive correlation between sports stadium vicinity and housing prices with an increase in this effect in the year after the local team has won the championship. As previous literature has found a positive correlation between real estate sales prices and a professional sports stadium in the vicinity in single case studies(Tu, 2005. Propheter, 2019) or on larger statistical aggregations(Feng & Humphreys, 2012. Huang & Humphreys, 2014).

The results of the analysis find the opposite of the initial hypothesis. In the city of Denver a negative relation is found between the interaction Treated*Post(indicating all properties within 2.5 kilometers of the stadium sold in the year after the Super Bowl victory) and housing sale prices. The negative relationship found indicates that a victory in the Super Bowl causes the houses in the near vicinity of the stadium to temporarily drop in price. Reasons given for this possible relationship is the increase in interest being a cause for an increase in disamenities tied to the sports stadium, ranging from congestion to air pollution, or a shift in focus of public investment by local government after the Super Bowl victory.

In Pittsburgh and Tampa no relation was found. Several reasons for this are equally given, from lower residential density to research design limitations.

Multiple limitations to the analysis apply, especially when considering the analysis design and data availability. Since data is not available in equal formats and reporting on several characteristics which are of influence to housing sale prices is inconsistent the overall analyses suffer.

The results of this thesis should aid policy makers in determining the contribution of taxpayer money for the construction of new American Football stadiums. Additionally, these results contribute to the overall body of work on real estate and sports stadiums through its innovative method and incorporation of athletic success to the study of sports stadiums and their connection to real estate.

Future research is encouraged to look into the causes of disamenities around sports stadiums and their response to athletic success. As this will aid in the understanding of the effects which professional sports stadiums have on local neighborhoods and their respective real estate market. Additionally, future research is encouraged to focus on the overall impact of sports stadiums across economic and financial realms in order to give a more complete picture of the economic effects of sports stadiums in neighborhoods. Avenues like job creation and economic activity should be combined with the results in the real estate market to enhance the knowledge of the economic externalities of sports stadiums on local neighborhoods.

Sources

Ackers, Mick (2018). Final cost of Raiders stadium in Las Vegas pegged at $1.8 billion. LasVegasSun.com. Retrieved 05/03/2024

Bradbury, J. C., Coates, D., & Humphreys, B. R. (2023). The impact of professional sports franchises and venues on local economies: A comprehensive survey. *Journal of Economic Surveys*, *37*(4), 1389-1431.

Chikish, Y., Humphreys, B. R., & Nowak, A. (2018). Sports arenas, teams and property values: Temporary and permanent shocks to local amenity flows. *Available at SSRN 3254241*.

Chun-Chang, L., Chi-Ming, L., & Hui-Chuan, H. (2020). The impact of a mass rapid transit system on neighborhood housing prices: an application of difference-in-difference and spatial econometrics. Real Estate Management and Valuation, 28(1), 28-40.

Coates, D., & Humphreys, B. R. (2002). The economic impact of postseason play in professional sports. *Journal of Sports Economics*, *3*(3), 291-299.

Davenport, T. (2022). Titans, Nashville reach deal for domed, $2.1B stadium. *ESPN*. URL: https://www.espn.com/nfl/story/_/id/34816198/titans-nashville-reach-deal-domed-22b-stadium Visited on: 23/05/2024

Davies, L. E. (2006). Sporting a new role? Stadia and the real estate market. *Managing Leisure*, *11*(4), 231-244.

deMause, Neil (2017). "Why are Georgia taxpayers paying $700m for a new NFL stadium?". The Guardian. Retrieved 05/03/2024

Du, R., & Zhang, J. (2022). Super bowl participation and the local economy: Evidence from the stock market. *Growth and Change*, *53*(4), 1513-1545.

Empower Field At Mile High. (Date Unknown). About us. URL: https://www.empowerfieldatmilehigh.com/stadium/about-us Visited on: 25/05/2024

Fang, J. (2021). Impacts of high-speed rail on urban smog pollution in China: A spatial difference-in-difference approach. *Science of The Total Environment*, *777*, 146153.

Feng, X., & Humphreys, B. R. (2012). The impact of professional sports facilities on housing values: Evidence from census block group data. *City, Culture and Society*, *3*(3), 189-200.

Gyourko, J., & Molloy, R. (2015). Regulation and housing supply. In Handbook of regional and urban economics (Vol. 5, pp. 1289-1337). Elsevier.

Huang, H., & Humphreys, B. R. (2014). New sports facilities and residential housing markets. *Journal of Regional Science*, *54*(4), 629-663.

Humphreys, B. R., & Nowak, A. (2015). *Professional Sports Facilities, Teams and Property Values: Evidence from Seattle's Key Arena*. West Virginia Univ., Department of Economics.

Humphreys, B. R., & Nowak, A. (2017). Professional sports facilities, teams and property values: Evidence from NBA team departures. *Regional Science and Urban Economics*, *66*, 39-51.

Humphreys, B. R., & Pyun, H. (2018). Professional sporting events and traffic: Evidence from US cities. *Journal of Regional Science*, *58*(5), 869-886.

Humphreys, B. R., & Zhou, L. (2015). Sports facilities, agglomeration, and public subsidies. *Regional Science and Urban Economics*, *54*, 60-73.

Hyun, D. (2022). Proud of, but too close: the negative externalities of a new sports stadium in an urban residential area. *The Annals of Regional Science*, *68*(3), 615-633.

Joshi, A., Horn, B. P., & Berrens, R. P. (2020). Major league soccer expansion and property values: do sports franchises generate amenities or disamenities?. *Applied Economics*, *52*(44), 4881-4899.

Kaplan, D. (2023). Nashville on cusp of raising the public subsidies bar with new Titans stadium. *The Athletic*. URL: https://www.nytimes.com/athletic/4449753/2023/04/25/titans-stadium-nashville-public-funds/ Visited on: 23/05/2024

Keeler, Z. T., Stephens, H. M., & Humphreys, B. R. (2021). The amenity value of sports facilities: evidence from the Staples Center in Los Angeles. *Journal of Sports Economics*, *22*(7), 799-822.

Locke, S. L. (2019). Estimating the impact of Major League Baseball games on local air pollution. *Contemporary Economic Policy*, *37*(2), 236-244.

Menefee, C. (2024). The Impact of a Super Bowl Win on Building a Team's Fan Base. *SSRS*. URL: https://ssrs.com/insights/the-impact-of-a-superbowl-win-on-building-a-teams-fan-base/ Visited on 22/06/2024

Montolio, D., & Planells-Struse, S. (2016). How time shapes crime: The temporal impacts of football matches on crime. *Regional Science and Urban Economics*, *61*, 99-113.

NFL. 2023. Visited on: 14/12/2023 URL: https://www.nfl.com/news/super-bowl-lvii-total-viewing-audience-estimated-at-200-million.

Payne, B. C., Tresl, J., & Friesen, G. C. (2018). Sentiment and stock returns: Anticipating a major sporting event. *Journal of Sports Economics*, *19*(6), 843-872.

Propheter, G. (2019). Estimating the effect of sports facilities on local area commercial rents: Evidence from Brooklyn's Barclays Center. *Journal of Sports Economics*, *20*(1), 91-114.

Propheter, G. (2023). Sports Facilities as a Housing Amenity: Do Prices Follow Facilities?. *Journal of Sports Economics*, *24*(4), 443-474.

Raymond James Stadium(date unknown). About Raymond James Stadium. URL: https://raymondjamesstadium.com/about-us Visited on 25/05/2024 Visited on: 25/05/2024

Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of political economy*, *82*(1), 34-55.

Slobe, J. (2024). Beyond Touchdowns: How Chiefs Wins Drive Economic Prosperity in Kansas City. *KcSourcelink*. URL: https://www.kcsourcelink.com/2024/02/06/beyond-touchdowns-how-chiefs-wins-drive-economic-prosperity-in-kansas-city/ Visited on 22/06/2024

Sommers, P. M. (2000). The Super Bowl theory: fourth and long. *The College Mathematics Journal*, *31*(3), 189-192.

Steelers, (Date unknown). Steelers Standings | Pittsburgh steelers steelers.com. URL: https://www.steelers.com/team/standings/ visited on 05/04/2024

Steelers.com.(2008). Heinz Field Stadium Fact Sheet. URL: https://web.archive.org/web/20080417185453/http:/www.steelers.com/heinzfield/stadiumfacts/ Visited on 25/05/2024

Stephenson, C. (2022). Tennessee Titans, Nashville mayor unveil details of $2.1B stadium deal. URL: visited on 23/05/2024

Tu, C. C. (2005). How does a new sports stadium affect housing values? The case of FedEx field. *Land Economics*, *81*(3), 379-395.

Wing, C., Simon, K., & Bello-Gomez, R. A. (2018). Designing difference in difference studies: best practices for public health policy research. Annual review of public health, 39, 453-469.

Appendix A. Stadium selection criteria

1) Local professional NFL team needs to have won the Super Bowl in the last two decades
   a) 13 teams included.
2) The sports stadium in the vicinity needs to have sufficient housing the vicinity.
   a) New England Patriots and Kansas City Chiefs are dropped.
   b) 12 possible years remain due to repeat Super Bowl wins by these teams
3) The home stadium needs to house a single NFL team.
   a) Los Angeles Rams and New York Giants are dropped.
   b) 9 possible years remain due to repeat Super Bowl wins by these teams
4) Public data availability
   a) Philadelphia Eagles, Seattle Seahawks, Baltimore Ravens, Green Bay Packers, New Orleans Saints & Indianapolis Colts are dropped
   b) 3 teams remain. Denver Broncos, Pittsburgh Steelers and Tampa Bay Buccaneers.

Appendix B Maps of Study Areas
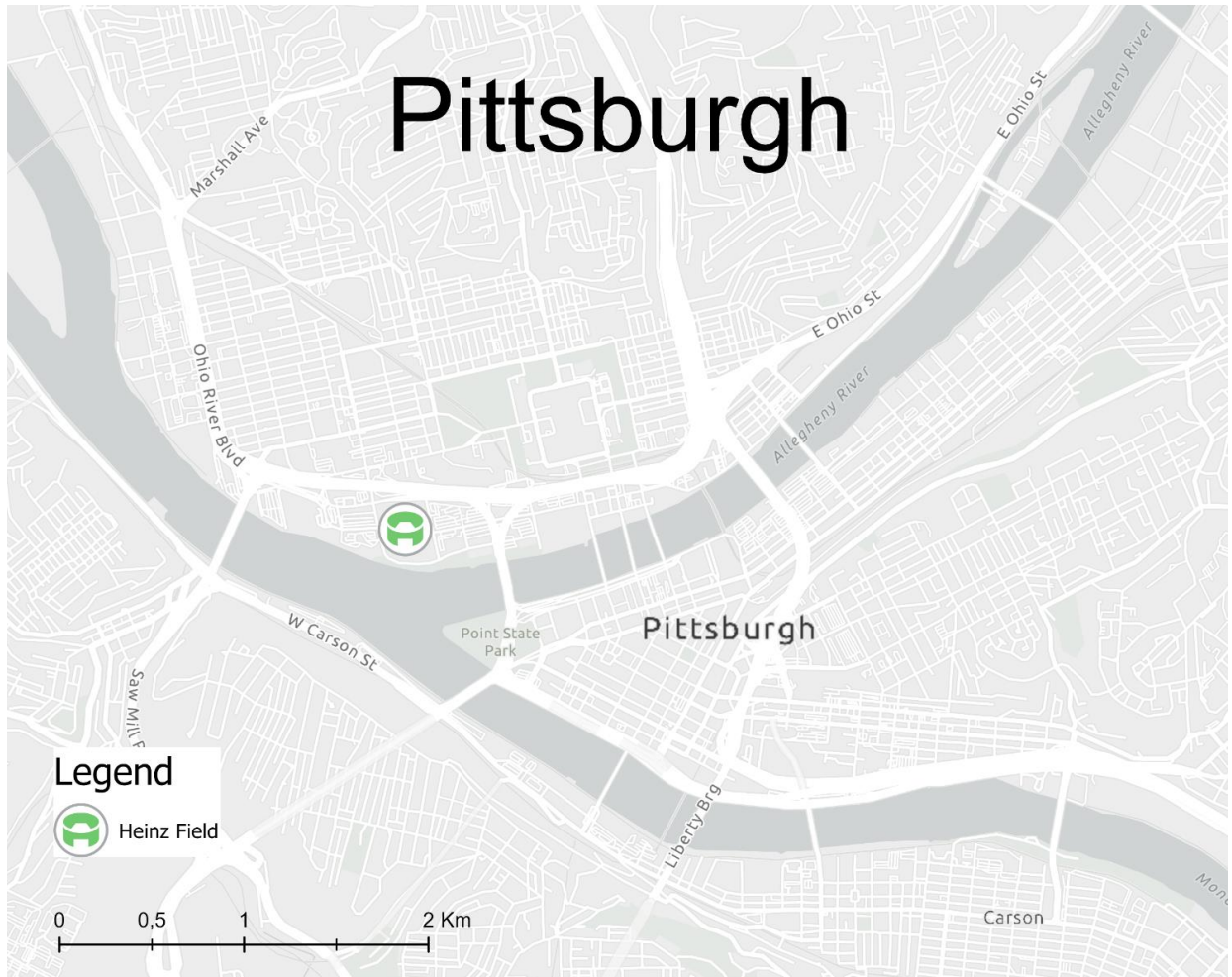

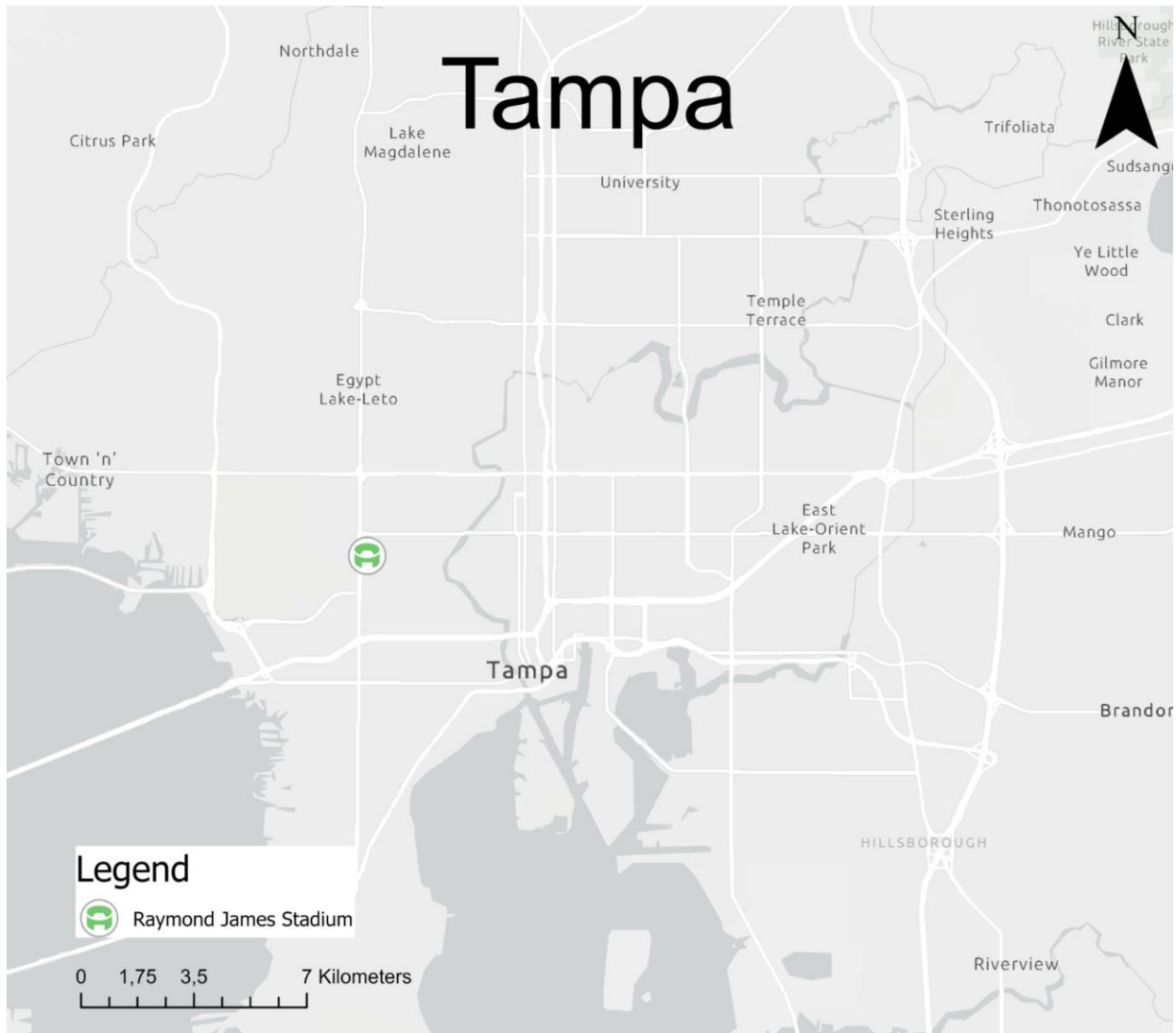Figure 1. Map of Denver study area

Figure 2. Map of Pittsburgh study area

Figure 3. Map of Tampa study area

Appendix C. Stata Syntax files.

Denver

*This is the data that has been linked together in GIS without any form of data cleaning yet
use "D:\RUG\Scriptie\Real_Estate_Transfer_Data\Denver\Final_Data_Denver.dta"

*Cutting data that did not join right in the beginning. Still seeing if it makes a huge impact to drop those cases
keep if _merge == "Matched (3)"

*making saleprice variable
drop if sale_price_1 == 0
drop if sale_price_1 == 1
keep if inrange(sale_price_1, r(p1), r(p99))
keep if tot_units <= 2

*A large chunk of cases was apparently low enough and large enough to mess with the normality assumption so I cut
it. There were a lot of cases with a sale price of 10 which I think I now cut through this way.
drop if lnsale <= 5

keep if Distance_Stadium < 5000

*Calculate the transaction which happened within a year of the super bowl for the "intime" variable
gen saledatereal = date(sale_date,"DMY hms")
format saledatereal %td

generate intime = inrange(saledatereal , td(07feb2015) , td(07feb2016))

*Generating the treated variable
drop if near_fid == -1
destring near_dist, generate(Distance_Stadium) dpcomma
generate treated = (Distance_Stadium < 1500)

*cleaning up the variables for the regression analysis
drop lnsale
generate lnsale =ln(sale_price_1)
drop if missing(lnsale)
*This should delete 132 observations
*Now check normal distribution of the variable
hist lnsale

summarize lnsale intime treated land_area bed_rms

*Have to drop the missing variables for this one to ensure equal groups
drop if missing(bed_rms)

destring story, replace dpcomma
destring cbd_distance, replace dpcomma

*Dropping outliers
destring bed_rms, replace dpcomma
summarize bed_rms, detail

```
local cutoff = r(p99)
drop if bed_rms > `cutoff'

destring full_b, replace dpcomma
summarize full_b, detail
local cutoff = r(p99)
drop if full_b > `cutoff'

destring story, replace dpcomma
summarize story, detail
local cutoff = r(p99)
drop if story > `cutoff'

destring area_abg, replace dpcomma
summarize area_abg, detail
local cutoff = r(p99)
drop if area_abg > `cutoff'


*fixed effects
tabulate fips, generate(tract)
tabulate sale_year_1, generate(year)
drop tract_fips

*original regression

sum lnsale intime##treatedv2 area_abg full_b bed_rms story cbd_distance

reg lnsale intime##treatedv2 area_abg full_b bed_rms story cbd_distance tract* year*


*sensitivity analyses
gen intimekort = inrange(saledatereal, td(07feb2015), td(8aug2015))
gen intimelang = inrange(saledatereal, td(08feb2015), td(8feb2017))
gen treatedkort = Distance_Stadium < 2000
gen treatedlang = Distance_Stadium < 3000
tabulate nbhd_1, generate(neighborhood)

*time period
reg lnsale intimekort##treatedv2 area_abg full_b bed_rms story cbd_distance tract* year*
reg lnsale intimelang##treatedv2 area_abg full_b bed_rms story cbd_distance tract* year*
*treatment area
reg lnsale intime##treatedkort area_abg full_b bed_rms story cbd_distance tract* year*
reg lnsale intime##treatedlang area_abg full_b bed_rms story cbd_distance tract* year*
*fixed effects check
reg lnsale intime##treatedv2 area_abg full_b bed_rms story cbd_distance neighborhood* year*
```

Pittsburgh

```
import delimited "D:\RUG\Scriptie\Real_Estate_Transfer_Data\pittsburgh\Data_Pittsburgh.csv"

*make intime and treated variable
destring stadium_distance, replace dpcomma
```

```
gen treatedv2 = stadium_distance < 2500
drop if stadium_distance > 5000
tab treatedv2 intime
*intime
generate saledatereal = date(user_saledate,"DMY hms" )
format saledatereal %td

generate intime = inrange(saledatereal, td(01feb2009) , td(01feb2010)) | inrange(saledatereal, td(05feb2006),
td(05feb2007))

*saleprice generation
generate lnsale =ln(user_saleprice)
drop if missing(lnsale)

destring user_stories, replace dpcomma
drop if missing(user_stories)

*fixed effects
tabulate user_yearofsale, generate(YEAR)

*Dropping outliers
destring user_bedrooms, replace dpcomma
summarize user_bedrooms, detail
local cutoff = r(p99)
drop if user_bedrooms > `cutoff'

destring user_fullbaths, replace dpcomma
summarize user_fullbaths, detail
local cutoff = r(p99)
drop if user_fullbaths > `cutoff'

destring user_stories, replace dpcomma
summarize user_stories, detail
local cutoff = r(p99)
drop if user_stories > `cutoff'

destring user_finishedlivingarea, replace dpcomma
summarize user_finishedlivingarea, detail
local cutoff = r(p99)
drop if user_finishedlivingarea > `cutoff'


destring user_stories, replace dpcomma
drop if missing(user_stories)
tabulate tract_fips, generate(TRACT)
destring cbd_distance, replace dpcomma
drop if missing(user_bedrooms)
drop if lnsale == 0
drop if lnsale < 5

sum lnsale intime##treatedv2 user_bedrooms user_finishedlivingarea user_fullbaths user_stories cbd_distance
reg lnsale intime##treatedv2 user_bedrooms user_finishedlivingarea user_fullbaths user_stories cbd_distance
YEAR* TRACT*
```

```
*sensitivity analysis
gen intimekort = inrange(saledatereal, td(01feb2009), td(1aug2009))
gen intimelang = inrange(saledatereal, td(01feb2009), td(1feb2011))
gen treatedkort = stadium_distance < 2000
gen treatedlang = stadium_distance < 3000
tabulate nbhd_1, generate(neighborhood)
*time period
reg lnsale intimekort##treatedv2 user_bedrooms user_finishedlivingarea user_fullbaths user_stories cbd_distance
YEAR* TRACT*
reg lnsale intimelang##treatedv2 user_bedrooms user_finishedlivingarea user_fullbaths user_stories cbd_distance
YEAR* TRACT*
*treatment area
reg lnsale intime##treatedkort user_bedrooms user_finishedlivingarea user_fullbaths user_stories cbd_distance
YEAR* TRACT*
reg lnsale intime##treatedlang user_bedrooms user_finishedlivingarea user_fullbaths user_stories cbd_distance
YEAR* TRACT*
*fixed effects check
reg lnsale intime##treatedv2 user_bedrooms user_finishedlivingarea user_fullbaths user_stories cbd_distance
neighborhood* year*
```

Tampa

```
import delimited "Complete_Data.csv"

*These are the codes for all normal market transaction. So no foreclosures or bank sales etc.
keep if rea_cd == "1" | rea_cd == "2" | rea_cd == "2A" | rea_cd =="2B"

*Dropping outliers
destring tunits, replace dpcomma
summarize tunits, detail
local cutoff = r(p99)
drop if tunits > `cutoff'

destring tbeds, replace dpcomma
summarize tbeds, detail
local cutoff = r(p99)
drop if tbeds > `cutoff'

destring lotsize, replace dpcomma
summarize lotsize, detail
local cutoff = r(p99)
drop if lotsize > `cutoff'
summarize lotsize, detail
local cutoff = r(p99)
drop if lotsize > `cutoff'
summarize lotsize, detail
local cutoff = r(p99)
drop if lotsize > `cutoff'


destring tstories, replace dpcomma
local cutoff = r(p99)
```

```
drop if tstories > `cutoff'

destring tbaths, replace dpcomma
local cutoff = r(p99)
drop if tbaths > `cutoff'

drop lotsize
destring acreage, generate(lotsize) dpcomma
destring freeway_distance, replace dpcomma
drop lnsale2
generate lnsaleV2 = ln(s_amt)

destring cbd_distance, replace dpcomma


gen saledate1 = date(s_date, "DMY")
format saledate1 %td
drop intime
gen intime = inrange( saledate , td(8feb2021), td(8feb2022))

*location and time fixed effects
tabulate tract_fips_1, generate(Tracteffect)
tabulate saleyear, generate(Year)

drop if saledate1 < date("01jan2004", "DMY")


sum lnsaleV2 intime##treated lotsizeV2 tbaths tbeds tstories cbd_distance
reg lnsaleV2 intime##treated lotsize tbaths tbeds tstories cbd_distance Tracteffect* Year*

gen treatedv2 = distance_stadium < 2500
drop if distance_stadium > 5000
tab intime treatedv2
sum lnsaleV2 intime##treatedv2 lotsizeV2 tbaths tbeds tstories cbd_distance
reg lnsaleV2 intime##treatedv2 lotsize tbaths tbeds tstories cbd_distance tracteffect* year*

*sensitivity tests variable definitions
gen intimekort = inrange( saledate, td(8feb2021), td(8aug2021))
gen intimelang = inrange( saledate, td(8feb2021), td(8feb2023))
gen treatedkort = distance_stadium < 2000
gen treatedlang = distance_stadium < 3000
tabulate nbhc_1, generate(neighborhood)

*time period tests
reg lnsaleV2 intimekort##treatedv2 lotsize tbaths tbeds tstories cbd_distance tracteffect* year*
reg lnsaleV2 intimelang##treatedv2 lotsize tbaths tbeds tstories cbd_distance tracteffect* year*

*treatment area tests
reg lnsaleV2 intime##treatedkort lotsize tbaths tbeds tstories cbd_distance tracteffect* year*
reg lnsaleV2 intime##treatedlang lotsize tbaths tbeds tstories cbd_distance tracteffect* year*

*fixed effects variations
reg lnsaleV2 intime##treatedv2 lotsize tbaths tbeds tstories cbd_distance neighborhood* year*
```