CHARLES UNIVERSITY
FACULTY OF SOCIAL SCIENCES
Institute of Economic Studies

The Relationship between Success in North American Sports and Birth Rates

Bachelor's thesis

Author: Tomáš Suchomel

Study program: Ekonomie a finance

Supervisor: PhDr. Radek Janhuba, M.A., Ph.D.

Year of defense: 2021

Declaration of Authorship

The author hereby declares that he compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain any other academic title.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis in whole or in part and agrees with the thesis being used for study and scientific purposes.

Liberec, July 27, 2021

Tomas Suchomel

Abstract

Sports are world phenomenon deeply ingrained in North American culture. The existing research suggests that sports outcomes have the potential to significantly influence the lives of emotionally interested fans. However, the question of relationship between sport outcomes and birth rates is not yet well-documented. Hence, the objective of this thesis is to establish whether there is a significant relationship between overall victories in major North American sport competitions and the number of births in the relevant territorial units. In general, no conclusive evidence was found that an overall victory is associated with any significant effect. However, more telling results were obtained for individual competitions. Interestingly, our novel method considering the natural variability in pregnancy length outperformed the usual approach adopted in similar studies as results show that R^2 and adjusted R^2 are both larger for the new method in all comparable regressions.

JEL Classification	F12, F21, F23, H25, H71, H87
Keywords	birth rates, sport outcomes, North America, unexpectedness

Title	The Relationship between Success in North American Sports and Birth Rates

Author's e-mail	36719043@fsv.cuni.cz
Supervisor's e-mail	janhuba@gmail.com

Abstrakt

Sport je světový fenomén hluboce zakořeněný v severoamerické kultuře. Stávající výzkum naznačuje, že sportovní výsledky mají potenciál významně ovlivnit život emocionálně zainteresovaných fanoušk?. Otázka vztahu mezi sportovními výsledky a porodností však dosud není dobře zdokumentována. Cílem této práce je proto zjistit, zda existuje významný vztah mezi celkovými vítězstvími v hlavních severoamerických sportovních soutěžích a počtem narozených v příslušných územních jednotkách. Obecně nebyly nalezeny žádné přesvědčivé d?kazy o tom, že by celkové vítězství bylo spojeno s jakýmkoli významným účinkem. Pro jednotlivé soutěže však byly dosaženy přesvědčivější výsledky. Je zajímavé, že naše nová metoda beroucí ohled na přirozenou variabilitu délky těhotenství překonala obvyklý přístup podobných studiích, neboť výsledky ukazují, že R^2 a adjusted R^2 jsou obě lepší při aplikaci nové metody, a to ve všech srovnatelných regresích.

Klasifikace JEL	F12, F21, F23, H25, H71, H87
Klíčová slova	porodnost, sportovní výsledky, Severní Amerika, nečekanost

Název práce	Vztah mezi úspěchy v hlavních severoamerický sportech a porodností
E-mail autora	36719043@fsv.cuni.cz
E-mail vedoucího práce	janhuba@gmail.com

Acknowledgments

The author is especially grateful to PhDr. Radek Janhuba Ph.D. for his patient guidance and consulting. All remaining errors are my own.

Typeset in FSV LaTeX template with great thanks to prof. Zuzana Havrankova and prof. Tomas Havranek of Institute of Economic Studies, Faculty of Social Sciences, Charles University.

Bibliographic Record

Suchomel, Tomas: *The Relationship between Success in North American Sports and Birth Rates*. Bachelor's thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies, Prague. 2021, pages 72. Advisor: PhDr. Radek Janhuba, M.A., Ph.D.

Contents

List of Tables

List of Figures

Acronyms

BT Baseline timing method

CDC Centers for Disease Control and Prevention

FE Fixed effects estimation

GAT Gestational age adjusted timing method

LMP Last normal menses

MLB Major League Baseball

MSA Metropolitan Statistical Area

NBA National Basketball Association

NCHS National Center for Health Statistics

NFL National Football League

NHL National Hockey League

OE Obstetric estimate of gestation at delivery

OLS Ordinary least squares estimation

RE Random effects estimation

US United States of America

Chapter 1

Introduction

Sports are ingrained in North American culture. Large sport events, like final matches of major competitions, entertain tens of thousands fans present at stadiums, as well as millions in front of their TV sets. Since fans tend to emotionally experience accomplishments of their favourite team, it is worth examining how the outcomes of these competitions affect the behaviour on population level. One of not yet well-documented fields of interest is the question of interplay between sport outcomes and birth rates.

The existing research reliably shows that sports results have the potential to significantly influence the lives of emotionally interested fans. For example, Inoue *et al.* (2017) claim that attending a game in person has a positive effect on life satisfaction of the spectator. Likewise, Janhuba (2019) suggests that unexpected wins of college football teams have positive effects on life satisfaction of surrounding communities. Lucas (2014) shows that the US counties with higher levels of life satisfaction grew at substantially faster rates than those with low life satisfaction.

On the other hand, only a few authors have so far directly addressed the relationship between sport successes and birth rates. Montesinos *et al.* (2013) observed a significant increase in births in Catalonia nine months after a local football team reached the Champions League Final in a dramatic fashion. Less conclusive findings were obtained in an article by Hayward & Rybińska (2017) who studied the number of newborns in counties of the Super Bowl winners nine months after winning the trophy. Other studies confirm the positive effect of a favourable outcome of a sport event on the level of testosterone of a fan (Bernhardt *et al.* 1998) which ought to, in result, increase their interest in sexual stimuli (Gorelik & Bjorklund 2015). Therefore, the published findings support the validity of the claim that large sporting events may have the potential to function as exogenous influences on births in the population of interest.

The objective of the thesis is to establish whether there is a significant relationship between overall victories in major North American sport competitions and the number of births in the relevant territorial units. To study this relationship, the US birth data provided by the National Center for Health Statistics (2020) was utilized together with known results of selected competitions. In this manner, two datasets using two different time of interest estimation methods were created. The effect in question is then estimated using the fixed effects estimation.

The rest of the thesis is structured as follows: *Literature Review* aims to explain the inference behind the research question, its relevance and context, and to provide all information necessary for understanding the methodology and obtained results. In *Data Description*, the datasets creation is explained. Firstly, the birth data are discussed. Methods of collection, relevant variables, and possible limitations are specifically addressed. Secondly, the major North American sport competitions of interest are introduced and the historical context behind the leagues and their relevant winners is provided. *Methodology* chapter describes how a specific overall winner was linked to the relevant territorial unit of appropriate time. Further, the model selection is debated, and a method of distinguishing between particular sport outcomes is introduced. In the chapter *Results*, the obtained findings are reviewed, interpreted, and discussed. Finally, the thesis is summarized by *Conclusion*.

Chapter 2

Literature Review

2.1 Sports and Economics

Sport is a world phenomenon, attracting millions of fans through all sorts of media, causing commercial revenue to reach record highs (Bryson *et al.* 2015). What may set sport events apart from other sources of entertainment is their characteristic uniqueness and unrepeatability experienced in a collective of fans (Fumarco & Principe 2021). The vast public interest in sports also manifests itself in the large amount of money people are willing to spend either by betting on sport outcomes, or on merchandise of their favourite teams[1].

Unsurprisingly, the branch of economic literature dealing with sports and its effects is growing. While prior research focused primarily on stadiums, or arenas, and their effects on the surrounding economies, nowadays the significant increase in data availability and accompanying technological progress in data processing allow for a broader scope of research. This development have attracted a number of scholars, covering a great variety of topics ranging from racial discrimination on sports labour markets to studying effects of sport participation on one's earnings (Bryson *et al.* 2015). An extensive review of literature in which sport data were used for economic research can be found in Bar-Eli *et al.* (2020).

This thesis builds on a branch of economic literature studying so-called unrelated effects of sports. In practice, researchers active in this field of interest attempt to find systematic changes in observed variables that could have been explained by seemingly unrelated sport events.

[1] If the current trend remains unchanged, The Economist forecasts that Americans will stake more than $60bn on sports events by 2025(The Economist 2019).

The potential extent of the non-intuitive influences that may be related to sport events can be nicely demonstrated on financial markets. Edmans *et al.* (2007) claim that unfavourable outcomes of international soccer matches have a negative effect on the next trading day's national stock market index returns. Likewise, Truong *et al.* (2021) suggest that the results of Vietnamese national soccer team are connected to the daily abnormal returns on the Vietnamese stock market. Apart from the sentiment effect, researchers also consider how investors' different attention allocation that could be contributed to a major sport competition taking place translates to the markets. For instance, Drake *et al.* (2015) find evidence that prolonged stock disruptions are more likely to be present in the market during, in the US, a popular basketball tournament called March Madness.

The notion that fans emotionally experience their favourite team performances has motivated researchers to study how various sport outcomes may influence fans' day-to-day behaviour.

Eren & Mocan (2018) investigate whether emotional shocks accompanying results of a favourite college football[2] team have any impact on professional behaviour of a highly educated group of individuals. The researchers conclude that unexpected losses may play some role in the tendency of higher severity of sentences ruled by the judges who were likely to be affiliated with the beaten team.

Recently the most publicly discussed topic is to what extent sport outcomes may contribute to family violence. In this regard, Card & Dahl (2011) examine the interplay between the incidence of reported domestic violence and emotional cues associated with wins and losses of local professional football teams. The authors suggest that upset losses of the local team translate to a roughly 10 % increase in the rate of at-home violence by men against their partners. More importantly for this thesis, the authors stress that upset losses in important matches, such as in the ones facing a rival or in play-off phase, have a bigger impact on the emotional cues associated with the increase, than losses in casual games.

A special stream of literature attempts to uplift the aggregated findings to a general discussion connecting sport outcomes to subjective well-being of cohorts of individuals. For example, Janhuba (2019) finds that unexpected wins of local college football teams have positive effects on life satisfaction of

[2]The word 'football' is used here in its American connotations, to address its European namesake the word 'soccer' will be used instead thorough this thesis.

surrounding communities. Interestingly, findings of Janhuba (2019) suggest that the impacts of emotional shocks caused by unexpected football wins are larger when the experience is shared with others. This dovetails well for this thesis with the discussion remark of Card & Dahl (2011) that couples seem likely to be together to watch football games.

At last but not least, the fact that fans deeply care about performances of their favourite teams can be also demonstrated on the increased cardiovascular death rates observed in the counties of defeated Super Bowl participants in the week after the loss (Schwartz *et al.* 2013). On a more positive note, findings of Fernquist (2000) suggest that the event of local sport team clinching play-off slightly lowers the suicide rate in the local population.

2.2 Natality

Before reviewing the branch of literature that deals with relating singular events to changes in childbirth incidence, a presentation of relevant facts concerning births, natality, and pregnancies in general is in order as it may be crucial for understanding the findings and conclusions of the discussed literature. Of special importance for the research question is to properly utilize the knowledge of the time aspect of pregnancies, since it is necessary to correctly determine the time periods in which the desired effect may be observed. In this regard, the thesis relies heavily on Joyce *et al.* (2015) and supplementarily on Jukic *et al.* (2013) and Spong (2013). In other aspects of pregnancies and births in general, the thesis draws on information presented in Foley & Kalro (2010).

General Information

A typical pregnancy begins at conception and may end in a live birth, a spontaneous miscarriage, an induced abortion, or a stillbirth. The process leading from a conception to an eventual birth is called gestation, the variable measuring this process in time is called a *gestational age*. A common notion is that a pregnancy lasts 9 months. In practice, potential mothers are given a future date of delivery as counted by adding 280 days to the date of their last menstrual period, however, only 4% of births are actually delivered at the given date. In early human development research and sexology textbooks, gestational age is usually counted in units of weeks. The average gestational age at delivery usually ranges between 38.5 and 38.9 weeks, with 80 % of births taking place

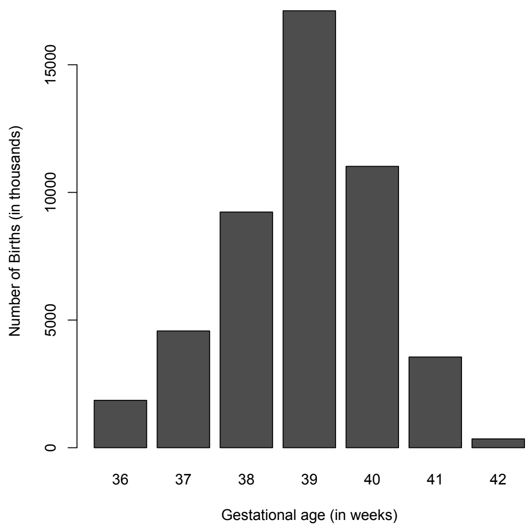

Figure 2.1: Births Distribution by Estimated Gestational Age
Source: the Author's own computation based on the CDC data

within 10 days of the week 39 of gestational age, as can be seen in Figure 2.1. A foetus is ideally developed for living on its own at $39 - 40$ weeks of gestational age. Births prior to 37 weeks of gestational age are considered pre-term, those born at 42 weeks and beyond post-term.

Time Trends in Births

Births are known to be monthly seasonal, but the exact trend can be culturally and geographically specific. In the US, the number of births usually peak in summer months and hit the nadir in winter season, as can be clearly seen in the Figure 2.2. A special birth phenomenon are medically induced births, of which an interesting corollary is so-called 'weekend effect' in births. This term encapsulates the observation that the number of births on weekends is significantly lower than during workdays. As show Gelman *et al.* (2013), similar effect also holds for 'special' days in the US, such as the 1st January, the Independence Day, the Thanksgiving, or the Christmas Day. As implied, the cause is not of biological origin but mostly as a consequence of the combination of labour contracts in healthcare and medically induced births.

In general, Vollset *et al.* (2020) compute that a large number of countries are facing long-term decline in fertility levels of their citizens, as can be clearly seen in Figure 2.2, hence potentially threatening the stability of pension schemes

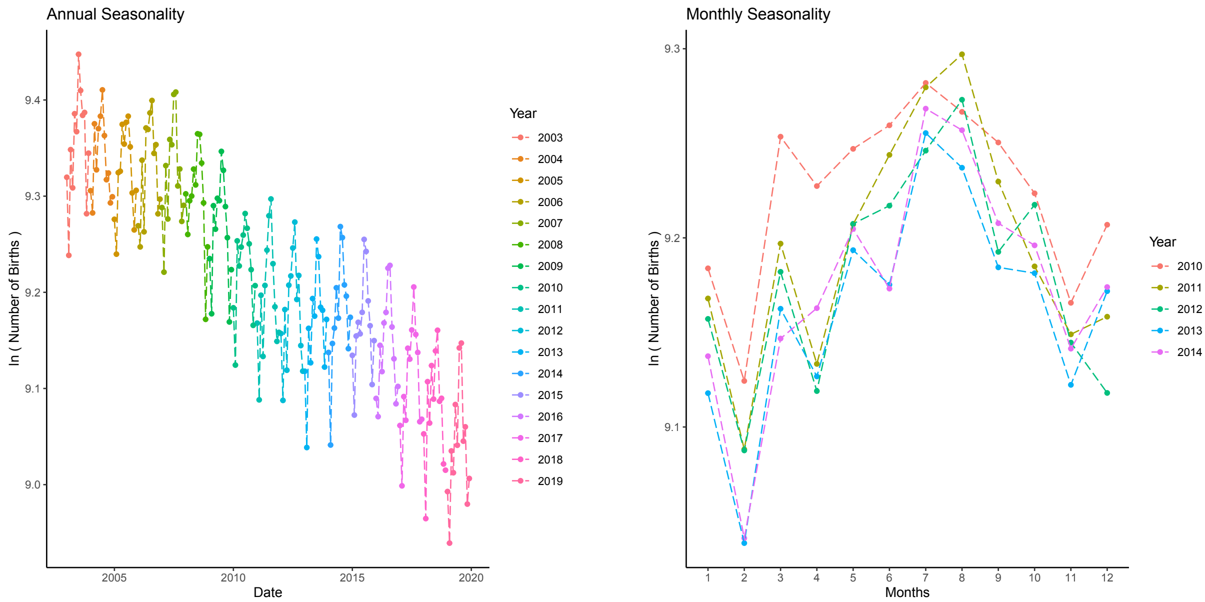

Figure 2.2: Seasonality in Births : Chicago-Naperville-Elgin area
Source: the Author's own creation based on the CDC data

worldwide. Observing the hypothesized positive effect of major sport outcomes on the number of births might thus have several contributions. First and foremost, given the context of the robust decline, government policies aiming to introduce positive exogenous shocks to the fertility levels of their citizens seem to play more significant role in the future of public planning. Naturally, findings confirming the hypothesized positive effect of sport outcomes on birth rates may contribute by providing rationale to municipal politicians that in order to boost local birth rates it may be of avail to invest public funds into local popular sport teams.

Births and Economics

Births and their determinants have always been a subject of special attention from economists, dating back as far as to the infamous "An Essay On The Principle of Population" by Thomas Robert Malthus, which was published for the first time in 1798. As a result, several economic variables have been found to be correlated with birth rates. Ermisch (1988) finds that higher women net real wages are related with lower birth rates, while the men wages with higher ones. The pro-cyclical pattern in fertility rates has been documented, for example, in Silver (1965) or Schaller *et al.* (2020). Dettling & Kearney (2014) examine the relationship between house prices and birth rates, specifically, whether a rise in house prices discourages families from having a child. Interestingly, the authors find that while house prices have indeed negative effects on birth rates of tenants, they seem to simultaneously have positive effects on birth rates of home-owners. Dehejia & Lleras-Muney (2004) claim that high unemployment rate at the time of conception may have various effects on corresponding birth characteristics, such as childbirth incidence, or birth weight. Lenhart (2021) shows that a dollar increase in minimum wage is associated with 3% decline in teenage birth rates. He (2018) connects births with monetary policy, finding a positive effect of an increase in the nominal interest rate on fertility, and also computing that 1% increase in annual inflation rate is likely associated with 0.06 % increase in annual population growth. A great review of additional literature connecting fertility levels to macroeconomic indicators is presented in Sobotka *et al.* (2011).

Exogenous Shocks to Births

Despite the relatively small amount of conducted research, the piquancy of connecting singular events to birth rates recurrently attracts media attention. Typically, a media outlet reports an occurrence of a baby boom some time after for a certain community seemingly important event, such as World Cup or a presidential election, had taken place. For instance, when Icelandic soccer players knocked out fancied Englishmen from EURO 2016 quarter finals thus achieving the greatest accomplishment in Icelandic soccer history, the British newspaper The Telegraph published an article reporting that the celebrations of the historic victory had translated into an occurrence of a baby boom in Iceland (Gibson 2017). Likewise, a morning show "Good Morning, America" aired a segment claiming that a release of an erotic novel called "The Fifty Shades of Grey" had similar positive effect on American population (Grech 2017). Neither report proved to reflect reality, as was shown in Grech & Masukume (2017), and Grech (2017), respectively. Longevity of this media tendency can be demonstrated on the front-page article published in the New York Times informing its readers that the New York City blackout of 1965 caused a spike in births nine month later (Tolchin 1966). Although this claim has also been debunked (Udry 1970), it inspired researchers to look for similar patterns of behaviour elsewhere, as the initially reported effect was later documented to have occurred in Columbia (Fetzer *et al.* 2018) and in Zanzibar (Burlando 2014).

Needless to say, such media reports tend to be backed up mostly by anecdotal evidence and/or urban legends, and as such are not subjected to peer-review. The common motivation behind the conducted research on the topic has been verification of such claims.

The question is well researched in the case of natural disasters that were established to have had an impact on the subsequent number of births in relevant regions. An increase in births was documented to have occurred in the US areas hit by Hurricane Hugo in 1989 (Cohan & Cole 2002). Interestingly, the effects of hurricanes along the Atlantic coasts of the US were found to depend on their severity, with low severity hurricanes being associated with increases in births and high severity ones with decreases (Evans *et al.* 2010). Similarly, Nandi *et al.* (2018) found that the 2001 Gujarat earthquake was associated with a rise in the incidence of childbirth and other birth characteristics in certain districts of Gujarat, India. A further review of natural disasters influencing fertility levels, may be found in de Oliveira *et al.* (2021).

Few events of anthropological origin were also found to have influenced the incidence of births. A decrease in fertility was observed among white women of Southern US states after an anti-segregationist ruling of the Supreme Court in 1954 (Rindfuss *et al.* 1978). Contrarily, evidence suggests that a positive effect on fertility levels occurred in the state of Oklahoma following the infamous Oklahoma City bombing of 1995 (Rodgers *et al.* 2005). A truly spectacular case of a state intervention in fertility levels of its citizens has been the China's One Child Policy established in 1979 (Liao 2013).

2.3 Relating Sport Success to Natality

To validate the research question, it is vital to expound on the inference connecting singular sport events to having a positive impact on the number of births. In this regard, the thesis adopts rationale as was laid out in Hayward & Rybińska (2017). The mainstay are findings of Bernhardt *et al.* (1998) which describe the observed positive effect of a favourable outcome of a sport event on the level of testosterone of the emotionally associated fan. This observation has allowed for a special stream of early development research studying so-called birth sex ratio and that specifically related to major sport events, as higher level of testosterone is suspected of being associated with higher prevalence of males to females in the number of births (Grech & Masukume 2017; van der Meij *et al.* 2012). But for this thesis, it is far more important to highlight the fact that higher testosterone levels are also well-documented to have a positive effect on the persons' interest in sexual stimuli (Rupp & Wallen 2008). Naturally, it is reasonable to suppose that the influence also effects the tendency to conceive children. Combining these two findings with the scenes that typically accompany celebrations of major sport successes, we arrive at the core of the rationale behind the research question. Supporting anecdotes of this actually happening can be occasionally found reported in the media, such as was the case of a couple of Boston Red Sox fans told by The Boston Globe (Grossfeld 2005).

Naturally, other factors, such as positive effects to subjective well-being, might play a role too, as for example Lucas (2014) shows that the US counties with higher levels of life satisfaction grew at substantially faster rates than those with low life satisfaction, which, in tandem with the findings of Janhuba (2019), could indicate supporting rationale for the possible connection. Lastly, as shown, the public seemed to be genuinely interested in the topic and rather intuitively agreed with the existence of such relationship.

As mentioned, there are so far only four publications that conceptually address the success-in-sport-births relation. These are now reviewed in the order of their date of publishing, as to show how the approach to the issue in question has developed.

The pioneering work on the topic in question is an article titled 'Barcelona Baby Boom: Does Sport Success affect Birth rate?' which has been published by the British Medical Journal. In it, Montesinos *et al.* (2013) test the anecdotal evidence claiming that the dramatical way in which a local soccer team clinched to the Champions League Final 2009 led to an increase in birth events in Catalonia nine months later. Employing an ARIMA model on birth data from three Catalonian hospitals located in a strongly pro-Barcelona county, the group of researchers confirm the story, finding a 16 % increase in births nine month after the match and 11 % increase ten month after the match. As mentioned, this article has been first of its kind and as such is always cited as an inspiration for authors of similar studies. Nevertheless, the researchers conduct the study in rather light-hearted manner, which was then accordingly published in a special Christmas edition of the British Medical Journal.

Attempting to expand Montesinos *et al.* (2013) findings to the American settings, Hayward & Rybińska (2017) examine the number of births in the US counties of Super Bowl winners nine months after the trophy gain. The article's chief motivation is to ascertain an official NFL commercial claiming that a Super Bowl victory brings a baby boom to the communities of supporters of the winning team. However, Hayward & Rybińska (2017) fail to find evidence for systematic changes in births that would be associated with Super Bowl victories, losses, or even with clinching play-offs. Although the findings are inconclusive, the article greatly expands Montesinos *et al.* (2013) by adding possible rationale for connecting favourable sport outcomes to eventual births. On the other hand, the article includes a number of possible limitations. Firstly, it is the fixation of time dimension solely on the following October after the Super Bowl. This determination is based on the claim from Jukic *et al.* (2013) stating that 269 days is the median length of gestation. Additionally, the authors do not apply a logarithmic transformation on the birth data, do not include any control variables to their estimation, and in general do not adhere to practices common in econometrics, such as a robustness check, or treating endogeneity. Finally, the article focuses only on the NFL and, moreover, on a rather narrow time period of $2003 - 2012$.

In a more empirically oriented article, Bernardi & Cozzani (2021) study the interplay between sport outcomes of local teams competing in Spanish soccer league and the corresponding number of births nine months later. The article aims to generalize findings of Montesinos *et al.* (2013) by focusing on a broad spectrum of soccer matches. The authors conclude that unexpected losses of local teams may be the cause of a small decrease in the number of births nine months thereafter, which they explain in terms of the Prospect Theory[3]. An interesting contribution of the article is the inclusion of unexpectedness to the question, and thus connecting it to a branch of literature represented by Card & Dahl (2011) or Janhuba (2019). As in the mentioned papers, the unexpectedness of a sport event outcome is modelled utilizing available betting information, odds in particular.

While authors of previous articles are demographers, sociologists, or health statisticians, the final paper has been written by economists. Fumarco & Principe (2021) explore the association between European soccer national teams' performances in international cups and the corresponding number of births after the events. Their findings suggest that an improvement in national team performance in international soccer competitions may explain a slight decrease in births in respective countries nine months after the events. The authors hypothesize that the drop may be a consequence of changes in time alloca-

[3]Specifically, they attribute the negative results to loss aversion of fans, i.e., that possible losses resonate with people more than possible same sized gains (Bernardi & Cozzani 2021).

tion, with the sport events being the cause of disruptions in the fans' reserved intimacy time.

Altogether, no common conclusion has been reached, since the conducted studies produce quite different findings, but all authors concur that the evidence supporting the validity of the effect in question is rather week. The most challenging task for in fertility interested conceptual studies seems to be the construction of an appropriate dataset, with the methodology then being rather monotonous in employing some kind of fixed effects estimation. Interestingly, all studies have quite rigid rules about determining the time dimension of interest by simply fixing it to roughly 9 months after the investigated event. In any case, given the small number of rather limited articles devoted to similar research question, the thesis could contribute greatly to the research of the issue in question.

Chapter 3

Data Description

To study the interplay between births and favourable sport outcomes, birth data and suitable performers are needed. For each competition, an original dataset of deciding matches was constructed using R Studio.

The time period of interest, for most of the competitions, is $2002 - 2018$, meaning that the thesis considers overall winners of the studied competitions whose deciding match is dated within the years 2002 and 2018. In the case of the NFL, the time period of interest is $2003 - 2019$, with the exception being due to the distinct timetable of the competition. The main reason for the specific time ranges is the availability of birth data, specifics of which are described in detail in the following section.

3.1 Birth Data

The sole source of birth data is a federal agency called the Centers for Disease Control and Prevention (CDC), specifically the database CDC WONDER, which is administered by a sub agency of the CDC named the National Center for Health Statistics (NCHS).

The CDC WONDER database is a great source of public health data for the United States of America (US). While some competitions of interest have several participating teams of Canadian origin[4], the empirical part of the thesis does not need to account for them, as not a single overall winner of interest has its home stadium located in Canada. Hence, only data concerning births

[4]MLB: Toronto Blue Jays; NBA: Toronto Raptors; NHL: Edmonton Oilers, Calgary Flames, Montreal Canadiens, Ottawa Senators, Toronto Maple Leafs, Vancouver Canucks, Winnipeg Jets.

in the US were collected. A list of home stadium locations for all appearing teams can be found in the Appendix B.

Before describing the particularities of the births data, a brief summary of the US statistical geography is in order. Since 1959, the United States of America comprises 50 states[5]. Each state, except for Alaska and Louisiana, is composed of counties. Alaska is of no concern, since no relevant team origins in it. Louisiana, which teams appear, is composed of so called parishes[6]. Another special case concerns particular cities, of those which teams appear are Baltimore and St. Louis. Both cities are recorded separately from their surrounding counties, Baltimore County and St. Louis County, respectively. At last, Washington D.C. is reported independently as a single county, with the short cut D.C used to denote its 'State' of belonging. Henceforth, the word 'county' will be used to refer to all statistically equivalent units of counties, such as parishes or particular cities.

Sets of counties are divided into various higher statistical units. Of interest for this thesis are so-called metropolitan statistical areas (MSA). MSAs are combinations of counties, not necessarily within a state, as delineated by the US Office of Management and Budget for statistical purposes. Typical MSA is intended as a region that consists of a city and surrounding communities that are linked by either social or economic factors. A list of currently defined MSAs was downloaded from the Bureau of Economic Analysis website.

The thesis relies on two subsets of the part of the CDC WONDER database entitled "Live births in the United States, 1995 − 2019". Specifically, two datasets covering time periods 2003 − 2006 and 2007 − 2019 respectively are used[7]. In the used datasets, births data are collected on monthly basis separately for each county which population exceeds 100,000 persons. While weekly data would be even more suitable for the analysis, such are not publicly available. The main benefit of monthly data could be that it may absorb the pregnancy length variability, such as is described in Section 2.2. Counties with population less than 100,000 are for every combination of state, year, and month aggregated into a single observation labelled Unidentified Counties. Populations of counties in the US are determined decennially in Censuses conducted by the US Census Bureau. The 2003 − 2013 data show counties with a

[5]Not only, but the so-called dependent territories are of no interest for the thesis.

[6]Parishes are in practice used equivalently to counties with the different denotation being of traditional rather than statistical nature.

[7]The eldest dataset of 1995 − 2002 period does not record births on monthly basis, but only annually, which renders it unusable in the context of the research question.

population of 100,000 persons or more as reported by the 2000 Census. Analogously, the 2014 − 2019 data identifies counties with a population of 100,000 persons or more using figures from the 2010 Census. As a result, there are 56 more counties named in the 2014 − 2019 data than in the 2003 − 2013 data. To achieve comparability between the datasets, these counties will be added back to their respective Unidentified Counties observations. Apart from the inclusion base, the two datasets differ mainly in collected variables concerning early human development, and as such, for purposes of this thesis, can be combined into one dataset. The amalgamated dataset covers years between and including 2003 and 2019, with the earliest observation being recorded for January 2003 and the latest one for December 2019. Given the research question, the accessible time range allows to study the effects of winners from the time period 2002 − 2018, and 2003 − 2019 in the case of the NFL. This stems from the fact that for every winner, the birth data at least 36 weeks post the deciding match are necessary for the analysis. An interested reader may see Sections 2.2 and 4.1 for more context behind this inference.

CDC WONDER collects a large number of variables, almost all of which concern characteristics used for early human development research, e.g., maternal risk factors or a delivery method, and as such are not reviewed here. Of interest for this thesis are births and their gestational age.

A variable Births counts a number of recorded births, which corresponds to the number of issued birth certificates in hospitals located in the observed county in a particular combination of year and month. Notably, this practice includes only liveborns[8] of the US citizens, thus excluding stillbirths and liveborns of non-citizens, while including those born prematurely and post-termly. In the case of interest, a reader may see Section 2.2 for more information concerning the biological background of birth events.

Each birth is characterised by its gestational age. A variable Gestational Age at Birth is defined as the duration of the pregnancy at the time of birth, and is used to control for premature and post-term births. In other words, the variable measures the time period between a conception and an eventual birth. CDC WONDER offers three groupings by gestational age which differ in the number of recorded age groups. This thesis will use the most thorough

[8]The NCHS record a birth as 'liveborn' when the complete expulsion or extraction from its mother of a product of human conception, irrespective of the duration of pregnancy, which, after such expulsion or extraction, breathes, or shows any other evidence of life such as beating of the heart, pulsation of the umbilical cord, or definite movement of voluntary muscles.

one, with 30 age groups, ranging from 17 weeks to 47 weeks by a week[9], with only 0.2 % of births occurring outside of this range (Joyce *et al.* 2015). Births with unknown gestational age can be discounted from the analysis without significant consequences, as they comprise, according to Joyce *et al.* (2015), less than 1 % of the total number of births. Left untreated, they may cause an unnecessary diffusion in results. There are two methods used in the dataset to estimate the gestational age, the last normal menses (LMP) and the obsteric estimate of gestation at delivery (OE). Technical details of each method are beyond the focus of this thesis[10], for which it is sufficient to acknowledge in estimation results that the OE is more precise than the LMP (Jukic *et al.* 2013) and as such will be preferred whenever possible. In practice, this means that the LMP will be used on $2003 - 2006$ data, as it is the only option, and the OE on the rest. The details of the application of gestational age in the estimation process are described in Section 4.1.

The contours of the original dataset may imply several limitations. Firstly, it may fail to consider changing hospitals' location distribution throughout the observed years, which could negatively influence the comparing attempts. Secondly, the definition of a liveborn should, in the context of the research question, be accounted for, as it could diffuse the crucial time dimension. And finally, the frequency of birth data collection does not allow such flexibility that would be optimal in the context of the research question.

3.2 Major North American Sport Competitions

The focus of this thesis is narrowed to four major North American sport competitions, namely, the National Basketball Association (NBA), the National Football League (NFL), the National Hockey League (NHL), and the Major League Baseball (MLB). Henceforth, the competitions will be referred to by their commonly used acronyms.

First of all, the reasoning behind the narrowing in the selection of sport competitions should be explained. Although it goes without saying that the research question could concern any major sport competition, there are relevant reasons for the particular selection. Most importantly, the necessary birth data for North America are well prepared and easily accessible, with especially

[9]Given the gestational age of a birth being n weeks and m days, if m < 4, then the birth is assigned to the week n, otherwise to the week $n + 1$.

[10]A thorough description of each method can be found in Joyce *et al.* (2015).

the US having a long tradition of data collection. Secondly, the selected competitions are all well-established, with the appearing teams being in existence, and moreover, in their current location, for more than 8 years prior to their respective trophy win[11]. Hence, it is not unreasonable to infer that fan bases in local communities may have been established as well. Furthermore, the major North American sport competitions are characteristic of their play-off style conclusion, allowing the *deciding match*, which is for the purposes of this thesis defined as the match assuring the champion status, to be easily traced in time, as it is, in any given season, the last match of the competition taking place. That is unlike the case of many European major sport competitions, leagues in particular. See the English Premier League, for which it is common to have a team being assured of becoming the overall winner many weeks before the league officially finishes. For an illustration consider Liverpool FC in the 2017/18 season, when it managed to assure its champion status while still having seven league matches ahead to play (Johnson 2020).

Another reason for the made selection is of methodological nature. Based on similar studies, the desired effect cannot reasonably be expected to be "large". In the light of this observation, it is sound to narrow the focus of the thesis to the most popular sport competitions in North America, and specifically to their most important matches. According to the Gallup Historical Sport Trends report, the most popular sports for North Americans to watch are football, baseball, basketball and ice-hockey[12], in that order. Notably, hockey is significantly less popular than the other sports, the inclusion of the NHL to the selection is motivated mostly by increasing the sample of studied counties, and is also supported by the NHL's great historical tradition[13], due to which it is usually mentioned in league with the more popular major sport competitions.

Notable exemptions from the selection are the Major League Soccer (MLS), the Canadian Football League (CFL), and competitions governed by the National Collegiate Athletic Association (NCAA), such as is so-called March Madness.

In the following subsections, the selected competitions are separately presented in the order of their respective popularity as reported in Gallup (2021). The presentations are organized in the following manner: (1) A competition is

[11] The youngest team that appear are Baltimore Ravens which moved in from Cleveland prior to the 1996 season, i.e., 17 seasons before their overall win of interest.

[12] Actually, ice-hockey was surpassed by soccer in 2017, but for most of the observed period, it is the fourth favourite major sport league to watch.

[13] Having been established in November 1917 translates roughly into 103 years of tradition.

briefly introduced, (2) the process of becoming an overall winner is described, (3) the timetable of the deciding matches of interest is presented, (4) specifics and summaries of the relevant winners are discussed. The aim is to list all information regarding the teams and competitions in question that may be of relevance to the research question. It should be mentioned that individual rules of respective sports in question are not reviewed.

National Football League

(1) The professional football league is by far the most popular sport competition in the US, with the Super Bowl being regularly among the most watched TV events of the year. The most watched Super Bowl in history is Super Bowl XLIX with a record 114.44 million viewers (Sports Media Watch 2021b).

(2) The structure of the NFL season consists of two parts. First, a potential winner plays 16 regular season games, then 2-3 post-season matches before reaching the finale, famously named the Super Bowl. A salient distinction of the NFL from other studied competitions is its single match play-off structure. Additionally, Super Bowls are always played at predetermined[14] stadiums. As a result, the venue is independent of the finale's participants, making it possible for one team to play the final at its home stadium. Nevertheless, such match did not happen in the studied time period[15], meaning that all deciding matches of interest were played at neutral venues. Super Bowls are traditionally denoted, and referred to, with a Roman numeral. A list of all Super Bowls with corresponding numerals can be found in Appendix A.

(3) The Super Bowl has been customarily played on the first Sunday in February[16].

(4) There are 11 different winners in the sample. The most accoladed, by far, are the New England Patriots, who under the tandem Bill Belichick – Tom Brady managed a spectacular record of five overall victories in the sample period[17]. Also noteworthy, the Seattle Seahawks, Tampa Bay Buccaneers,

[14]The decision is made by the NFL and is announced about 4 years prior to the match.

[15]Interestingly, such match took place only once in the history of the NFL, when Tampa Bay Buccaneers played the Super Bowl LV at its home stadium.

[16]Since 2004, therefore there is one Super Bowl in the sample that was played in January.

[17]Moreover, they claimed the Super Bowl trophy in 2002, and were the beaten finalist in 2008, 2012, and 2018.

Philadelphia Eagles[18] and New Orleans Saints celebrated their first Super Bowl victory just in the time period of interest.

Team						n
New England Patriots	2004	2005	2015	2017	2019	5
New York Giants	2008	2012				2
Pittsburgh Steelers	2006	2009				2
Baltimore Ravens	2013					1
Denver Broncos	2016					1
Green Bay Packers	2011					1
New Orleans Saints	2010					1
Indianapolis Colts	2007					1
Philadelphia Eagles	2018					1
Seattle Seahawks	2014					1
Tampa Bay Buccaneers	2003					1

Table 3.1: The NFL's overall winners of interest

All information used in the regard of the NFL results history, were obtained from the website *Pro-football-reference.com*[19].

Major League Baseball

(1) The professional baseball league is the most traditional major sport competition in North America, dating its origins back as far as to the late 19th century[20].

(2) A potential winner plays firstly 162 regular season games, based on which the team clinches to a four-round best-of-seven post-season tournament that culminates in the World Series.

(3) The World Series is customarily played in the Fall, with specific dates being dependent on the number of games played in the leading play-off phase.

(4) There are 11 distinct winners in the sample. The most successful team in are the Boston Red Sox, which after breaking the infamous 'Curse of the Bambino' [21] in 2004, managed to secure three additional triumphs. Interestingly, two winners, the Chicago Cubs and Chicago White Sox, origin in the

[18] Only Super Bowl victories are considered, thus their overall victory from 1960 is not counted.

[19] Pro-Football Reference (2021)

[20] Actually, the MLB was formed in 1903, but its predecessors the National League and the American League were established as early as 1876 and 1893, respectively.

[21] The 85 seasons long trophy drought experienced by the Boston Red Sox franchise that was allegedly caused by their decision to trade Babe "The Bambino" Ruth, arguably the best player in the baseball history, to New York Yankees in 1918.

same city[22]. The former triumph was especially historic, as it meant winning the World Series after record 107 seasons[23] long pause since their previous victory. Also worth mentioning is that the Anaheim Angels[24] and Houston Astros won their historically first Commissioner's Trophy in the studied time period.

Team					n
Boston Red Sox	2004	2007	2013	2018	4
San Francisco Giants	2010	2012	2014		3
St. Louis Cardinals	2006	2011			2
Anaheim Angels	2002				1
Chicago Cubs	2016				1
Chicago White Sox	2005				1
Florida Marlins	2003				1
Houston Astros	2017				1
Kansas City Royals	2015				1
New York Yankees	2009				1
Philadelphia Phillies	2008				1

Table 3.2: The MLB's overall winners of interest

All information concerning the MLB results history were obtained from the website *Baseball-reference.com*[25].

National Basketball Association

(1) The professional basketball league finals are commonly the most watched sport series in North America, in 2016 the NBA finals were watched by an average of 20 million viewers only in the United States (Sports Media Watch 2021a).

(2) The regular season consists of 82 games[26] then there are for each advanced team up to four rounds[27] of best-of-seven post-season games that culminate in the NBA Finals.

(3) The final series is usually played in June, with the exact date being dependent on the number of matches in previous series leading to the finale.

[22]This is the only instance of two teams competing in the same league sharing a single city while not being located in either Los Angeles, or New York City.

[23]Interestingly, the previous record-holder were their neighbours the Chicago White Sox which win in 2005 ended their 87 seasons long waiting for a trophy.

[24]Nowadays named the Los Angeles Angels.

[25]Baseball Reference (2021)

[26]In the NBA, a team plays with every other team during the regular season, this is unlike in the case of the NFL or MLB.

[27]Namely, the First Round, Conference Semifinals, Conference Finals and NBA Finals

In some cases the date of the finale was delayed due to pre-seasonal prolonged negotiations between the NBA and the players' association.

(4) There are 8 different winners in the sample. The most successful franchise among the winners of interest are the San Antonio Spurs, which under the leadership of couch Gregg Popovich managed to reap four NBA Championships just in the studied time period[28]. Interestingly, more than 80 % of the overall victories of interest are distributed just among four teams. This may have something to do with the relatively small number of players necessary to have under contract in order to be competitive. Notably, the Cleveland Cavaliers, Dallas Mavericks, and Miami Heat achieved their first NBA Championship trophy within the years of interest[29].

Team					n
San Antonio Spurs	2003	2005	2007	2014	4
Golden State Warriors	2015	2017	2018		3
Los Angeles Lakers	2002	2009	2010		3
Miami Heat	2006	2012	2013		3
Boston Celtics	2008				1
Cleveland Cavaliers	2016				1
Dallas Mavericks	2011				1
Detroit Pistons	2004				1

Table 3.3: The NBA's overall winners of interest

All information used in the regard of the NBA results history, were obtained from the website *Basketball-reference.com*[30].

National Hockey League

(1) Despite being less popular than its contemporaries, the professional hockey league has been expanding rapidly[31].

(2) The teams firstly play 82 regular season games, after which follows the play-off best-of-seven elimination tournament that determines the Stanley Cup champion.

[28] Actually, the Spurs dynasty begun by winning the Larry O'Brien Championship Trophy for the first time in the 1998/99 season.

[29] The Cleveland Cavaliers 4-3 victory over the Golden States Warriors from the 2016 finale is considered especially historic, and as such it is commonly presented as the most entertaining NBA Championship of all time.

[30] Basketball Reference (2021)

[31] The Las Vegas Golden Knights joined the competition in 2017, the Seattle Kraken are about to have their inaugural season, and in Québec locals cherish the prospect of renewing the now defunct Nordiques.

(3) The dates of the Stanley Cup finale series depend on a number of conditions, but in general, the final series usually concludes in June.

(4) There are 10 distinct winners in the sample, with the Chicago Blackhawks and Pittsburgh Penguins being the most successful with three overall victories. Notably, a winner of the 2004/2005 season is missing. The cause of this absence is the NHL lock-down of 2005, which did not resolve in time resulting in a suspension of the entire season. The 2005 winner is thus not included in the analysis. Finally, the Anaheim Ducks[32], Carolina Hurricanes, Los Angeles Kings, Tampa Bay Lightning and Washington Capitals won their first Stanley Cup just in the studied time period.

Team				n
Chicago Blackhawks	2010	2013	2015	3
Pittsburgh Penguins	2009	2016	2017	3
Detroit Red Wings	2002	2008		2
Los Angeles Kings	2012	2014		2
Anaheim Ducks	2007			1
Boston Bruins	2011			1
Carolina Hurricanes	2006			1
New Jersey Devils	2003			1
Tampa Bay Lightning	2004			1
Washington Capitals	2018			1

Table 3.4: The NHL's overall winners of interest

All information used in the regard of the NHL results history, were obtained from the website *Hockey-reference.com*[33].

Summary

There are in total 40 different winners, 14 of which succeeded multiple times. 14 teams also became champions for the first time just in the time period of interest. Interestingly, Boston is the only place that have a winner in all competitions[34]. A complete list of all deciding matches for each competition can be found in Appendix A. A table of all individual winners with their stadium locations is presented in Appendix B. Finally, a distribution of individual win-

[32] At that time known as the Mighty Ducks of Anaheim, with the different name being directly connected to then owner: The Walt Disney Company, which named the team after their film 'The Mighty Ducks' from 1992.

[33] Hockey Reference (2021)

[34] Namely, the New England Patriots, Boston Red Sox, Boston Celtics, and Boston Bruins.

ners by competition and the MSA of their location can be found in Appendix C.

Chapter 4

Methodology

4.1 Time and Location Identification

We begin by presenting this work's answer to the crucial question of determining the link between a specific overall winner and the relevant territorial unit of appropriate time. However, it is important to state that it is unavoidable for any assignment, that the resulting estimation will take into account a great number of births that were in no way associated with the overall victory of a local team. So, we look for disruptions in general trends that may be attributable to the held celebrations.

Lets begin with the time dimension. The underlying assumption of the assignment is that the celebratory intercourse leading to an eventual birth took place in the night of the deciding match, or in a few following days. The crucial task now is to select for each deciding match such combination of months that catches the largest number of births conceived during the celebrations. To achieve this, two methods will be employed to determine the months of interest.

Firstly, a method of similar studies is replicated, meaning that the treatment month is determined simply by adding 9 months to the date of the deciding match. Such approach is in line with a common notion concerning the length of gestation, and as such was adopted, for example, in Fumarco & Principe (2021). For the sake of convenience, this method will be called the '*Baseline timing*' method (BT).

Second method is novel in its utilization of the richness of data provided by the CDC WONDER database, specifically the variable 'Gestational Age at Birth'. Employed correctly, this method should allow to control for a number of births which were certainly unrelated to the deciding match, and hence, more precise estimation results ought to be obtained. As stated in Section 3.1, two methods were used in the initial database to estimate gestational age, the LMP for births data from the $2003-2006$ period and the more precise OE for the rest. Based on Joyce *et al.* (2015), it is valid to assume that more than 93 % of all pregnancies result in birth within 36 and 42 weeks of gestational age as estimated by the OE, and 85 % as estimated by the LMP. We use this gestational age estimates to count back and determine the likely month of the conception. Specifically, we use the 15th day of each month as a reference point. From this point we then subtract the number of weeks of the given gestational age, and denote the month of the resulting date, as the month when the births of the subtracted gestational age were likely to be conceived. We repeat this for all relevant gestational age categories. As a result, we have a dataset which instead of monthly births shows a number of births likely conceived in the given month. Notably, this moves boundaries of the dataset approximately by 9 months further to the past. The treatment months are then simply the months of the dates of the deciding matches. This method of time dimension assignment will be referred to as to '*Gestational age adjusted timing*' (GAT).

Having determined the treatment months for each winner, it is now necessary to connect it to a proper set of counties that would best reflect the winners' respective fan bases. There are various options available. For a baseline estimation, we assign each winner to the MSA that encompasses the county where the team's home stadium is located[35]. We make one exception, in the case of the MSA "Los Angeles-Long Beach-Anaheim", we stick instead with the original county-level locations: "Los Angeles County" and "Orange County". The rationale is, firstly, the relatively large population of the combined MSA[36], and, secondly, because of the convenience of not combining winners from Los Angeles and Anaheim into one observation[37]. A list of all relevant home stadiums with their surrounding MSAs can be found in Appendix B.

This approach entails a number of benefits. Firstly, it accommodates the problem of teams that have their home stadium located in a different state than their name may indicate[38]. Due to the appearance of such teams in the sample of winners, assigning teams only to the state, or county, of their home

[35]Similar approach was adopted, for example, in Card & Dahl (2011).

[36]Combined MSA has, according to the Census 2010, roughly 13 million residents, separately, Los Angeles County has 9 million and Orange County 3 million residents.

[37]Namely, the Anaheim Angels and Ducks from Orange County, and the Los Angeles Kings and Lakers from Los Angeles County.

[38]For example, the stadium of New York Giants is located in one of the neighbouring states of New York, specifically in New Jersey.

stadium[39] location is insufficient, as it may alter the estimation to omit a great number of fans affiliated with the concerned team. Another crucial benefit of the assignment is that the resulting dataset can be readily connected with a number of statistical variables that are regularly collected using MSA as a base.

On the other hand, several limitations must be taken into consideration. First and foremost, the approach surely entails a great overlap of various fan bases and non-fans, causing a large number of births unrelated to the overall victory to be included in the estimation. Since it is, at least to the author's knowledge, impossible to identify whether a particular birth was initiated by a couple of the winning team's supporters, it is unavoidable but to introduce a measurement error to the estimation, and that by linking all births in a given area to one winner, regardless of whether the teams is even supported by all fans living in that area. As in the case of Janhuba (2019), this assignment means that the estimation process will produce intention-to-treat estimates with the estimated effect probably being biased downward. For a better comparison of results, a subset of unique MSAs[40] is subsetted, Another major consequence of the assignment is that the observations labelled 'Unidentified Counties' are discounted from the estimation, since they are unable to be correctly assigned to their proper MSA.

To recapitulate, the two dimensions necessary for any panel econometric method were determined. The 40 overall winners have translated into 25 MSAs of 17 years, i.e of 204 months, constituting two balanced longitudinal datasets of 5100 observations. A list of all assigned MSAs with the count and distribution of the located winners can be found in Appendix C.

4.2 Model Specification

Based on the information listed above, it is likely that some kind of an unobserved effects model for panel data would be most suitable for the laid-out setting. Such conclusion allows for various estimation procedures. For our study, we choose the *fixed effect estimation* (FE) to be the primary method, since it is common in fertility related research, as it was employed, for example, in Bernardi & Cozzani (2021), Dettling & Kearney (2014), Schaller *et al.* (2020), and others.

Before getting to the particularities of the regression analysis, it may be

[39]See, for example, the location of Dallas Cowboys.
[40]These are those that have $n = 1$ in Appendix C.

beneficial to explain why the FE is considered to be the optimal method given the contours of the dataset, and the context of the research question.

In econometrics, it is usual to adhere to the FE, when a researcher is faced with potential unobserved heterogeneity in the elements of his/her panel data sample, which is, moreover, suspected to be correlated with the explanatory variables.

In our case, we deal with a panel sample of MSAs. Each MSA represents a differently structured area, may it be in the size of current population, or in other socio-economical aspects, such as its specific social strata, aggregated preferences, real GDP per capita, urban and rural ratios, etc. and also, in a geographical sense, meaning, for instance, a specific climate, distance from the sea, and so on. These and other differences can be expected to be accumulated in the form of the individual unobserved heterogeneity, either time-variant or invariant, of each MSA.

Now, with the presence of individual unobserved effects being established, it is time to decide whether we suspect the unobserved heterogeneity to be correlated with the independent variables that we will employ to explain the birth data. By making this decision, we essentially choose between the FE and the *random effect estimation* (RE). Given the structure of the dataset and being cognizant of the information laid out in the last paragraph of Section 2.3, it is most reasonable to assume that the correlation is present and proceed primarily with the FE.

Combining the FE with the BT gives us an equation of the following form:

$$ln\,(births_{a\,t}) = \lambda_{BT}\,Win_{a\,t-9} \;+\; \boldsymbol{X}_{a\,t-9}\,\boldsymbol{\beta} \;+\; Month \;+\; Year \;+\; \gamma_a \;+\; \delta_t \;+\; \varepsilon_{a\,t} \tag{4.1}$$

where $births_{a\,t}$ is the number of births that took place in the MSA a at the time t, with $t \in \{1, 2, 3, \ldots, 204\}$[41]. A logarithmic transformation is applied on $births_{a\,t}$, as the variable takes on only positive values, and the entailing interpretation is suitable for the research question. $\boldsymbol{X}_{a\,t-9}$[42] is a vector of control variables characterising the MSA[43] a at the time of the supposed conception. $Month$ and $Year$ are dummy variables accounting for the monthly and annual seasonality in births, respectively. γ_a is the time-invariant specific heterogene-

[41]Number 1 corresponds to January 2003, number 2 to February 2003, etc.

[42]In the case of $t - 9 < 1$, we consider dates prior to the birth dataset.

[43]In the case of Los Angeles County and Orange County, we assign to each the characteristics of their overarching MSA: "Los Angeles-Long Beach-Anaheim".

ity of the MSA a, δ_t is the time-varying fixed effect, $\varepsilon_{a\,t}$ is an idiosyncratic error with 0 mean and variance σ^2_ε, $Win_{a\,t-9}$ is a dummy variable taking value 1 if the overall winner from 9 months back origins in the MSA a, and 0 otherwise. Finally, λ_{BT} is the parameter of interest.

Transferring to the GAT method slightly changes the estimated equation:

$$ln\,(births^{(\boldsymbol{w})}_{a\,d}) = \lambda_{GAT}\, Win_{a\,d} \;+\; \boldsymbol{X}_{a\,d}\,\boldsymbol{\beta} \;+\; Month \;+\; Year \;+\; \gamma_a \;+\; \delta_d \;+\; \varepsilon_{a\,d} \tag{4.2}$$

where $births^{(\boldsymbol{w})}_{a\,d}$ is the number of births that were likely conceived in the MSA a at the date d, as computed by the gestational age $\boldsymbol{w} = \langle 36, 37, \ldots, 42\rangle$, with $d \in \{1, 2, 3, \ldots, 204\}$[44]. The variables $\boldsymbol{X}_{a\,d}$, γ_a, δ_d, $Year$, $Month$ and $\varepsilon_{a\,d}$ have the analogous meaning to the ones explained in the previous equation. $Win_{a\,d}$ is a a dummy variable taking value 1 if the overall winner from the date d origins in the MSA a, and 0 otherwise. λ_{GAT} is again the parameter of interest.

The desired outcome confirming our hypothesis that sport outcomes may have the potential to serve as exogenous shocks on birth levels is a statistically significant $\boldsymbol{\lambda} > 0$. Although the theory suggests that the FE is the appropriate technique for the particular dataset, it is still useful to re-estimate the equations using pooled *ordinary least squares estimation* (OLS) and the RE. Then, it is possible to test for the presence of fixed effects using an F-test, and employ the Hausman's specification test to verify whether the FE is indeed the correct choice over the RE (Wooldridge 2003). And finally, the proper tests for serial correlation and cross-sections dependency must also be run.

4.3 Heterogeneity of Sport Successes

As mentioned in Section 2.1, the unrepeatability of the sport event experience is considered to be one of the main lures of the sport related entertainment (Fumarco & Principe 2021). Therefore, treating all deciding matches homogeneously may negatively affect the conclusions of the conducted estimations. A re-estimation of the above-listed regressions while only considering a strictly defined subset of the deciding matches may improve the quality of results, and help us to understand the driving factors behind the effect. Based on the reviewed literature, three features of a deciding match have been identi-

[44]Number 1 corresponds to April 2002, number 2 to May 2002, …, number 204 to March 2019.

fied as possible means of distinguishing between favourable sport outcomes: unexpectedness, drama, and uniqueness of the victories.

Unexpectedness

At first, we incorporate a distinction often made in sport economic literature, specifically, the element of unexpectedness of sport outcomes. Janhuba (2019) even argues that when analysing shocks induced by sport events, the inclusion of the unexpected element of the results into the empirical methodology is necessary. The potential relevance of unexpectedness to the issue in question can be shown on its inclusion in Bernardi & Cozzani (2021), where it led to statistically significant estimates.

To infer the expected outcome of each game, it is usual for interested studies to utilize the existence of a well-organized betting market, with the rationale being that betting odds ought to reflect all relevant available information about the upcoming game (Card & Dahl 2011).

Two kinds of odds were collected for each winner from the website *sportoddshistory.com*[45], specifically, odds offered prior to the finale[46], and those offered before the particular season started. Pre-final odds may tell us whether the eventual winner was considered to be the underdog, or the favourite of the finale. Pre-seasonal ones then should control for the beaten finalist, and also reflect how the winner had been expected to perform in the season. All, except the pre-Super Bowl odds, are of so-called American Betting Odds kind, in which the odd number represents how much one must bet to win \$100[47].

We define a victory as an *unexpected* one, if either the pre-final odd is positive[48], or the pre-seasonal one is higher than the average value of the other winners' pre-seasonal odds in the sample[49]. In the case of the pre-Super Bowl odds, a victory is considered *unexpected*, if the pre-final spread against is +7 points[50] or more.

[45]SportsOddsHistory.com (2021)

[46]In the case of final series, the last possible odds before the series begun were collected.

[47]For example, if team A has an odd of -300 to win, then \$1 of potential profit is made for every \$3 risked, meaning that in order to make profit of \$100, we would have to bet \$300.

[48]i.e. one has to bet less than \$100 to eventually win \$100, meaning that prior to the finale the winner was considered to be the underdog.

[49]Therefore, we control for how the winner's expected performance stood in comparison to other winners.

[50]i.e. the winner was expected to lose by a touchdown.

As a result, the estimated equations change to:

$$ln\,(births_{a\,t}) = \Lambda_{BT}\; unexpWin_{a\,t-9} \;+\; \boldsymbol{X}_{a\,t-9}\,\boldsymbol{\beta} \;+\; Month \;+\; Year \;+\; \gamma_a \;+\; \delta_t \;+\; \varepsilon_{a\,t}$$

$$ln\,(births^{(\boldsymbol{w})}_{a\,d}) = \Lambda_{GAT}\; unexpWin_{a\,d} \;+\; \boldsymbol{X}_{a\,d}\,\boldsymbol{\beta} \;+\; Month \;+\; Year \;+\; \gamma_a \;+\; \delta_d \;+\; \varepsilon_{a\,d} \tag{4.3}$$

where $\boldsymbol{\Lambda}$[51] are the parameters of interest, and all other variables have analogous meaning to those used in the equations (4.1) and (4.2), respectively.

Drama

Findings of Montesinos *et al.* (2013) and a discussion remark of Hayward & Rybińska (2017) suggest that significant changes in births following a sport event outcome may depend on the dramatic manner in which the game was played out[52]. Therefore, it may be reasonable to take the dramatic aspect of a sport game into consideration. But, as all sports quite differ, the concept of a dramatic deciding match must be defined for each competition separately.

A deciding match is considered *dramatic*, if either:

- **NFL**
  i) the final score difference was less than 9 points[53].

- **MLB**
  i) the deciding match was resolved in the over time;
  ii) the last game was decided by points scored in the 9th innings;
  iii) the finale took seven games to conclude.

- **NBA**
  i) the final score difference was 4 points or less[54];
  ii) the finale took seven games to conclude.

- **NHL**
  i) the score difference within the last 2 minute of the deciding match was

[51] $\boldsymbol{\Lambda} = \langle \Lambda_{BT}, \Lambda_{GAT} \rangle$.

[52] Compare, for example, the Super Bowl XLIX with the Super Bowl LIII. While both are technically victories of New England Patriots, the difference in the accompanying emotional experience could not, arguably, be larger.

[53] i.e. it was theoretically possible for the opposing team to equalize in the last possession by scoring a touchdown with the 2 points conversion.

[54] i.e. the opposing team could potentially equalize in just one successful play.

at some point one goal[55];
ii) the deciding match was resolved in the over time;
iii) the finale took seven games to conclude.

Again, we employ a slight alteration of the afore-defined equations:

$$ln\,(births_{a\,t}) = \kappa_{BT}\;dramWin_{a\,t-9} \;+\; \boldsymbol{X}_{a\,t-9}\,\boldsymbol{\beta} \;+\; Month \;+\; Year \;+\; \gamma_a \;+\; \delta_t \;+\; \varepsilon_{a\,t}$$
$$ln\,(births^{(\boldsymbol{w})}_{a\,d}) = \kappa_{GAT}\;dramWin_{a\,d} \;+\; \boldsymbol{X}_{a\,d}\,\boldsymbol{\beta} \;+\; Month \;+\; Year \;+\; \gamma_a \;+\; \delta_d \;+\; \varepsilon_{a\,d} \tag{4.4}$$

where $\boldsymbol{\kappa}$ are the parameters of interest.

Uniqueness

As can be seen in the subsections of Section 3.2, a number of teams that appear in the sample have succeeded repeatedly in their respective competitions. Considering this, it is not unreasonable to surmise that fans might get emotionally accustomed to experiencing overall victories of their favourite team, and thus the overarching positive effect could have been worn out. Based on this conjecture, it may be worth restricting the analysis to only those victories that may be considered momentous for the communities of fans. This approach allows us to essentially control the estimates for so-called sport dynasties[56].

A victory is defined as *unique*, if it is the first victory for the team given four years prior. Admittedly, the bound is determined rather arbitrarily, but we justify it by the quite narrow time range of $2002 - 2019$, and the fact that victories which are 2-3 years apart are usually thought of as belonging to the same dynasty[57]. In the course of this determination, winners outside the time period of interest are also taken into consideration[58].

[55]i.e. we control for matches in which a power play tactic was unsuccessfully used by the opponent.
[56]The Merriam-Webster dictionary defines a dynasty as a "sports franchise which has a prolonged run of successful seasons".
[57]A typical example is the Chicago Bulls dynasty which claimed the NBA Championship in 1991, 1992, 1993 and 1996, 1997, 1998.
[58]This affects the victories of New England Patriots (2004), Los Angeles Lakers (2002), San Antonio Spurs (2003), Detroit Red Wings (2002), and New Jersey Devils (2003).

At last, lightly adjusted versions of (4.1) and (4.2) are estimated:

$$
\begin{aligned}
ln\,(births_{a\,t}) &= \xi_{BT}\; uniqWin_{a\,t-9} \;+\; \boldsymbol{X}_{a\,t-9}\,\boldsymbol{\beta} \;+\; Month \;+\; Year \;+\; \gamma_a \;+\; \delta_t \;+\; \varepsilon_{a\,t} \\
ln\,(births^{(\boldsymbol{w})}_{a\,d}) &= \xi_{GAT}\; uniqWin_{a\,d} \;+\; \boldsymbol{X}_{a\,d}\,\boldsymbol{\beta} \;+\; Month \;+\; Year \;+\; \gamma_a \;+\; \delta_d \;+\; \varepsilon_{a\,d}
\end{aligned}
\tag{4.5}
$$

where $\boldsymbol{\xi}$ are the parameters of interest.

Chapter 5

Results

We apply the fixed effect estimation on two panel datasets of 5100 observations, with the datasets being distinct in the method of determining the dependent variable, in the case of the BT we study the number of births, with the GAT we deal with the number of likely conceived births, i.e. controlling for pre-term and post-term births, in the particular combination of MSA and month. We hypothesized that the GAT should lead to a more precise estimation.

To better evaluate the results, we have also subsetted a sample of only unique MSAs[59]. The unique MSAs subset represents areas where only a single overall winner origins.

Our control variables include house prices and the unemployment rate in the area, two variables that are, according to information listed in the last paragraph of Section 2.3, strongly related with birth rates. The unemployment rate is collected on monthly basis for each MSA, and as such is in the optimal format. On the other hand, the house prices, reflected in so-called house price index (HPI)[60], are, while using MSA as base, collected quarterly. We overcome this hurdle, by assigning respective quarters to the months they represent, with the justification being that the important overall trend in house prices is thus captured. A number of other control variables come to mind, that may be appropriate given the context of the research question, such as a fraction of college educated population, a share of the age group of 18-44 among residents, or the GDP trends. Unfortunately, these are, at least to the author's knowledge, not available in the proper format[61] and/or for the entire time period of interest. Furthermore, it is important to state that while the control variables reflect the

[59]These are such MSAs that in Appendix C have n = 1.

[60]January 2000 = 100.

[61]Collected monthly on the MSA level.

whole MSA, we do not work with the births for the entire MSA, as we had to, in the process of location determination, discard a number of counties.

In general, it is recommended to treat the results with considerable caution, as the assignment of winners to their encompassing areas and proper time periods is, while to the author's knowledge the best possible one, still rather fragile, as we rely on data of various frequencies. Also, as mentioned, all findings are intention-to-treat estimates, and as such are likely to be downward-biased. Full regressions result are located in Appendix E.

All computations were conducted using statistical software R version 4.0.4, with an extensive use of the **plm** package and its accompanying methodology, as laid out in Croissant & Millo (2008).

5.1 Baseline Regressions

Table 5.1 shows the results of regressions using the BT and GAT method, respectively. Table 5.2 expands the results by distinguishing between particular sport competitions. Models denoted with the numbers (1) and (3) are regressed using a full sample (N = 25) of MSAs, while models (2) and (4) only on the subset of unique MSAs (N = 14).

Before reviewing the obtained results, a number of tests to ascertain the validity of using the FE are in order. We applied the following tests: poolability test, test for individual and time effects, Hausmann test, serial correlation test, and finally the Pesaran's test for cross-section dependency. As a result, we shall proceed using the FE with both time and individual effects. Also, all standard errors have been adjusted for clustering at the MSA level.

As can be seen in Table 5.1, λ is positively signed but short of statistical significance for both timing methods when regressed on the full sample. Therefore, given the cluster robust standard errors, we do not have sufficient evidence to reject the null hypothesis, that major North American sport competitions have no impact on the corresponding number of births, and that at every conceivable significance level.

More interesting results are obtained in Table 5.2, where we distinguish between particular competitions. As in the general case, when we regress on the full sample of MSAs, no statistically significant results are obtained, either for

[62] Controls variable include unemployment rate and house price index (HPI), HPI is statistically significant in most of the estimations, while the unemployment rate is not, as can be seen in the Appendix E.

	Dependent variable:			
	ln(Births)			
	BT		GAT	
	(1)	(2)	(3)	(4)
$\boldsymbol{\lambda}$: *Win*	0.002	−0.0004	0.004	0.011
	(0.007)	(0.012)	(0.007)	(0.011)
Controls[62]	Yes	Yes	Yes	Yes
Month dummy	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes
Fixed effects	Time and MSA	Time and MSA	Time and MSA	Time and MSA
Observations	5,100	2,856	5,100	2,856
R^2	0.532	0.485	0.696	0.703
Adjusted R^2	0.527	0.477	0.693	0.699
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 5.1: Regression Results - General

the GAT, or BT method. On the other hand, we have a number of statistically significant estimates in the case of sample of unique MSAs. Namely, our results show that a victory in the MLB, NBA and NHL may be associated with having an impact on the corresponding number of births nine months later, all at the significance level of $\alpha = 0.05$. However, the GAT method narrows the statistically significant estimates to only the Stanley Cup victory, but that is more significant with the p-value of less than 0.01.

The accompanying interpretation is that an overall victory in the NHL may be associated with a roughly 2% increase in corresponding births, while holding all other factor fixed and controlling for individual and time fixed effects.

Notably, both signs are present among the statistically significant estimates, with the MLB and NHL showing a positive effect, and the NBA with the opposite. It is possible that the different findings are driven by the specific demographic strata favouring the given sport.

However, it is important to bear in mind, that these significant findings were obtained while comparing among unique MSAs, and none are confirmed by regressions on the full sample.

Interestingly, the NFL, which is, by far, the most popular sport competition in North America, is the only one without a statistically significant estimate, and that either for the GAT or BT method. This may be explained, for example, by the locations of the particular Super Bowl champions in the sample.

	Dependent variable:			
	ln(Births)			
	BT		GAT	
	(1)	(2)	(3)	(4)
λ:				
NFL	−0.015	−0.005	−0.010	0.002
	(0.009)	(0.024)	(0.013)	(0.031)
MLB	0.017	0.024**	0.014	0.025
	(0.011)	(0.012)	(0.011)	(0.020)
NBA	−0.007	−0.024**	−0.001	−0.003
	(0.013)	(0.012)	(0.012)	(0.015)
NHL	−0.006	0.019**	−0.003	0.025***
	(0.010)	(0.009)	(0.011)	(0.006)
Controls	Yes	Yes	Yes	Yes
Month dummy	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes
Fixed effects	Time and MSA	Time and MSA	Time and MSA	Time and MSA
Observations	5,100	2,856	5,100	2,856
R^2	0.532	0.485	0.696	0.703
Adjusted R^2	0.527	0.477	0.692	0.698
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 5.2: Regression Results - Individual Leagues

5.2 Effect Heterogeneity

In the previous estimations, all observed victories were treated homogeneously, despite being undoubtedly heterogeneous. As described in Section 4.3, we have attempted to overcome this obstacle by distinguishing individual results based on their unexpectedness, drama, and uniqueness. With this distinction, we re-run the regressions, again using both the GAT and BT.

Table 5.3 shows aggravated results for all leagues. In Table 5.4 and Table 5.5, estimates for individual competitions are presented, with the division being for the sake of convenience.

Similarly to the previous estimations, the general results include few statistically significant estimates. The only instance where we can reject the null hypothesis of no effect, is in the case of dramatic victories regressed on the full sample of MSAs using the BT method. However, this result is supported neither by the unique subset regression, nor by the GAT method, but it is considerably closer to statistical significance than the other results.

Again, far more interesting results are obtained when we distinguish between individual leagues. As can be seen in Table 5.4, the previous statistically significant estimate is mainly driven by the dramatic Super Bowl victories, which has been estimated with a particularly strong effect, and that moreover in almost all of our estimates, with the only exception being the estimation on the full sample while using the BT method. But even this estimate is of the same sign as the other estimates. Altogether, our findings indicate that a dramatic Super Bowl victory is associated with a 8-10% decrease in the corresponding number of births. We may attempt to explain this by using the logic as laid out in Fumarco & Principe (2021), who attributes the possible drop in births to a different than usual time allocation. In our case, one possible explanation is that a significant portion of the winners' area population spent the time usually reserved for intimacy by other activities, some of which may be attributable to the overall victory, such as participating in celebrations taking place in the streets. As the effect is especially strong when regressed only on a sample of unique MSAs, which is mostly composed by less populated areas, such as Green Bay, Wisconsin, this may be very well the case. Notably, the effect is not so persistently present in results for other competitions.

Interestingly, unique victories seem not to have any effect at all, as not a single statistically significant estimate has been obtained.

At last, we examine the effects of the inclusion of unexpectedness. We have especially strong estimates for unexpected NBA victories, as is the only instance of statistically significant estimates being obtained for all applied methods in the thesis. Moreover, all estimates are of the same negative sign. The accompanying interpretation is that an unexpected gain of the NBA Championship is associated with roughly 3% to 5% reduction in the number of corresponding births. Again, this may be explained by the different time allocation. Less conclusive findings are obtained for the unexpected Stanley Cup victories, as there is just one estimate short of statistical significance. Interestingly, the sign is positive, and therefore have the opposite meaning to the one estimated for the unexpected NBA victories. Specifically, our findings suggest that an unexpected triumph in the Stanley Cup Finals is associated with 2-3% increase in the number of corresponding births, while holding all other factors fixed. The impact of unexpectedness is not that surprising, as have been shown in Section 2.1, it is a strong factor in effects studied by sports economic literature, see for example Eren & Mocan (2018). In general, our findings indicate that the aspect of unexpectedness seems to play a significant role even in our particular sport-related research. Therefore, in this regard, we can state that our conclusions are in line with the previous research.

	Dependent variable:			
	ln(Births)			
	BT		GAT	
	(1)	(2)	(3)	(4)
Λ: *Unexp*	0.010	−0.001	0.013	−0.0003
	(0.010)	(0.028)	(0.010)	(0.034)
ξ: *Unique*	0.004	0.018	−0.00003	0.028
	(0.009)	(0.028)	(0.010)	(0.034)
κ: *Dram*	−0.019***	−0.021	−0.011	−0.028
	(0.007)	(0.015)	(0.008)	(0.017
Controls	Yes	Yes	Yes	Yes
Month dummy	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes
Fixed effects	Time and MSA	Time and MSA	Time and MSA	Time and MSA
Observations	5,100	2,856	5,100	2,856
R^2	0.532	0.485	0.696	0.704
Adjusted R^2	0.527	0.477	0.692	0.699
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 5.3: Heterogeneity Effect - General

	Dependent variable:			
	ln(Births)			
	BT		GAT	
	(1)	(2)	(3)	(4)
NFL:				
Λ: *Unexp*	0.014	0.007	0.011	−0.008
	(0.013)	(0.022)	(0.017)	(0.029)
ξ: *Unique*	−0.005	0.018	−0.014	0.041
	(0.017)	(0.034)	(0.022)	(0.042)
κ: *Dram*	−0.039***	−0.076***	−0.021	−0.110***
	(0.011)	(0.024)	(0.014)	(0.033)
MLB:				
Λ: *Unexp*	0.009	0.036***	0.012	0.064***
	(0.023)	(0.008)	(0.023)	(0.010)
ξ: *Unique*	0.013		0.012	
	(0.014)		(0.013)	
κ: *Dram*	0.007	−0.015	0.0001	−0.052**
	(0.024)	(0.016)	(0.023)	(0.021)
Controls	Yes	Yes	Yes	Yes
Month dummy	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes
Fixed effects	Time and MSA	Time and MSA	Time and MSA	Time and MSA
Observations	5,100	2,856	5,100	2,856
R^2	0.533	0.486	0.696	0.704
Adjusted R^2	0.527	0.476	0.692	0.699
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 5.4: Heterogeneity Effect - NFL and MLB

5.2.1 Limitations

Several explanations for the unconvincing results come to mind. First, it may be that the emotional cues accompanying overall victories are not as strong as hypothesized. Moreover, it might be more telling to look at the shocks that are experienced by the fans of the losing team, as hypothesized in Bernardi & Cozzani (2021). Additionally, it is probable that our time range is just too narrow for any effect to accumulate, as mentioned in Section 3.1, the data availability only allows for the time range of $2002 - 2019$, while, for example, Fumarco & Principe (2021) worked with a dataset of multiple decades. It is also possible, that the studied effect is something that cannot be quantitatively generalized, since the particularities of each victory, and its following celebrations, are so distinct, that they may be worth examining only on individual level, as Montesinos *et al.* (2013) did. This would be in line with our findings, as statistically significant estimates are mostly obtained when areas with single winners are considered. Other possibility is that in the North American sport culture fans are not as emotional attached to their local professional teams, as are fans in the European setting, where previous studies found statistically significant relationships. This is not unreasonable, since the sport structures in the mentioned continents differ greatly. Especially, in the US, it is more than reasonable to surmise that the fans affections are distributed among larger number of teams than their European counterparts, as American sport teams below professional status also enjoy widespread support.

	Dependent variable:			
	ln(Births)			
	BT		GAT	
	(1)	(2)	(3)	(4)
NBA:				
Λ: *Unexp*	−0.044*	−0.056***	−0.036***	−0.026*
	(0.025)	(0.019)	(0.011)	(0.018)
ξ: *Unique*	0.046	0.059*	0.043***	0.036
	(0.032)	(0.033)	(0.016)	(0.037)
κ: *Dram*	−0.017	−0.022	0.0003	−0.004
	(0.013)	(0.019)	(0.011)	(0.018)
NHL:				
Λ: *Unexp*	0.018	0.019**	0.029**	0.026***
	(0.013)	(0.009)	(0.015)	(0.006)
ξ: *Unique*	−0.003		−0.010	
	(0.012)		(0.014)	
κ: *Dram*	−0.018		−0.013	
	(0.017)		(0.019)	
Controls	Yes	Yes	Yes	Yes
Month dummy	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes
Fixed Effects	Time and MSA	Time and MSA	Time and MSA	Time and MSA
Observations	5,100	2,856	5,100	2,856
R^2	0.533	0.486	0.696	0.704
Adjusted R^2	0.527	0.476	0.692	0.699
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 5.5: Heterogeneity Effect - NBA and NHL

Chapter 6

Conclusion

The objective of this thesis has been to establish whether there is a significant relationship between overall victories in major North American sport competitions and the subsequent number of births in the relevant territorial units.

In general, no sufficient evidence was found that such victory is associated with any significant effect. In this regard, the findings are in line with those obtained by Hayward & Rybińska (2017). However, more statistically significant results were found, when only so-called unique areas with just one successful team were taken into consideration. This might be caused by overlap of fan bases of individual sports within a particular area, possibly causing the emotional cues accompanying the euphoria of victory to be somehow less exclusive. This may be particularly true in extremely successful areas such as Boston.

If heterogeneity of the victories are considered, a relatively strong evidence for the effect of unexpected victories were found, namely in the NBA and NHL, which is also in line with the previous research. Moreover, unique victories do not seem to matter at all, which is surprising given the Law of Diminishing Marginal Utility, which interpretation in our particular setting would be that rare victories should matter more, than the ones experienced repeatedly. Concerning the aspect of drama, our results are inconclusive.

Interestingly, the comparison of the BT and GAT results show that R^2 and adjusted R^2 are both larger for the GAT method in all comparable regressions. In other words, more variation in births is always explained when different pregnancy lengths are considered, which suggests that the GAT method using the full richness of data encompassed in the CDC WONDER database was designed properly. Transferring from the BT to GAT changes the estimate sign in only one instance, which may be interpreted as a sign of more precision taking place. Thus, it is reasonable to conclude, that the existing and mostly indecisive articles would have also benefit from similar approach which might be considered a real contribution of this thesis.

There are several conceivable extensions for future research. Firstly, the research question could be expanded to other significant North American competitions. Secondly, it may change its focus to study the effects of losses instead of successes, with the rationale being as outlined in Bernardi & Cozzani (2021). Also, it could be interesting to increase the scope of research by including all games of respective teams in their competitions, not only those resulting into a trophy gain. This is in the fashion of Fumarco & Principe (2021) and Bernardi & Cozzani (2021). In other words, to alter the focus from overall victories to just victories.

Bibliography

Bar-Eli, M., A. Krumer, & E. Morgulev (2020): "Ask not what economics can do for sports - Ask what sports can do for economics." *Journal of Behavioral and Experimental Economics* **89**: p. 101597.

Baseball Reference (2021): "World Series and MLB Playoffs." `https://www.baseball-reference.com/`, Last accessed on 2021-6-25.

Basketball Reference (2021): "NBA & ABA Champions." `https://www.basketball-reference.com/`, Last accessed on 2021-6-25.

Bernardi, F. & M. Cozzani (2021): "Soccer Scores, Short-Term Mood and Fertility." *European Journal of Population* pp. 1–17.

Bernhardt, P. C., J. M. Dabbs Jr, J. A. Fielden, & C. D. Lutter (1998): "Testosterone Changes during Vicarious Experiences of Winning and Losing among Fans at Sporting Events." *Physiology & Behavior* **65(1)**: pp. 59–62.

Bryson, A., B. Frick, & R. Simmons (2015): "Sport Economics: It May be Fun but What's the Point?" *National Institute Economic Review* **232**: pp. R1–R3.

Burlando, A. (2014): "Power Outages, Power Externalities, and Baby Booms." *Demography* **51**.

Card, D. & G. B. Dahl (2011): "Family Violence and Football: The Effect of Unexpected Emotional Cues on Violent Behaviour." *The Quartely Journal of Economics* **126(1)**: pp. 103–143.

Cohan, C. & S. Cole (2002): "Life Course Transitions and Natural Disaster: Marriage, Birth, and Divorce following Hurricane Hugo." *Journal of Family Psychology* **16**: pp. 14–25.

Croissant, Y. & G. Millo (2008): "Panel Data Econometrics in R: The plm Package." *Journal of Statistical Software* **27**: pp. 1–43.

Dehejia, R. & A. Lleras-Muney (2004): "Booms, Busts, and Babies' Health." *The Quarterly Journal of Economics* **119**: pp. 1091–1130.

Dettling, L. J. & M. S. Kearney (2014): "House Prices and Birth Rates: The Impact of the Real Estate Market on the Decision to have a Baby." *Journal of Public Economics* **110**: pp. 82–100.

Drake, M., K. Gee, & J. Thornock (2015): "March Market Madness: The Impact of Value-irrelevant Events on the Market Pricing of Earnings News." *Contemporary Accounting Research* **33**.

Edmans, A., D. Garcia, & O. Norli (2007): "Sport Sentiment and Stock Returns." *Journal of Finance* **62**: pp. 1967–1998.

Eren, O. & N. Mocan (2018): "Emotional Judges and Unlucky Juveniles." *American Economic Journal: Applied Economics* **10(3)**: pp. 171–205.

Ermisch, J. (1988): "Economic Influences on Birth Rates." *National Institute Economic Review* **126**: pp. 71–81.

Evans, R., Y. Hu, & Z. Zhao (2010): "The Fertility Effect of Catastrophe: U.S. Hurricane Births." *Journal of Population Economics* **23**: pp. 1–36.

Fernquist, R. (2000): "An Aggregate Analysis of Professional Sports, Suicide, and Homicide Rsates: 30 U.S. Metropolitan Areas, 1971-1990." *Aggression & Violent Behavior* **5**: pp. 329–341.

Fetzer, T., O. Pardo, & A. Shanghavi (2018): "More than an Urban Legend: The Short- and Long-run Effects of Unplanned Fertility Shocks." *Journal of Population Economics* **31**.

Foley, C. & B. Kalro (2010): "The Johns Hopkins Manual of Gynecology and Obstetrics." *Journal of Pediatric & Adolescent Gynecology* **23**.

Fumarco, L. & F. Principe (2021): "More Goals, Fewer Babies? On National Teams Performance and Birth rates." *Working Papers 2109*, Tulane University, Department of Economics.

Gallup (2021): "Sports." `https://news.gallup.com/poll/4735/sports.aspx`, Last accessed on 2021-6-25.

Gelman, A., J. Carlin, H. Stern, D. Dunson, A. Vehtari, & D. Rubin (2013): *Bayesian Data Analysis*. Taylor & Francis Group.

Gibson, S. (2017): "Birth Record in Iceland Hospital Nine Months to the Day since Win over England." *The Daily Telegraph* https://www.telegraph.co.uk/football/2017/03/28/birth-record-iceland-hospital-nine-months-day-since-win-england/, Last accessed on 2021-7-16.

Gorelik, G. & D. Bjorklund (2015): "The Effect of Competition on Men's Self-Reported Sexual Interest." *Evolutionary Psychological Science* **1**.

Grech, V. (2017): "No Baby Booms or Birth Sex Ratio Changes following Fifty Shades of Grey in the United States." *Early Human Development* **110**: pp. 16–20.

Grech, V. & G. Masukume (2017): "Fake News of Baby Booms 9 months after Major Sporting Events distorts the Public's Understanding of Early Human Development Science." *Early Human Development* **115**: pp. 16–17.

Grossfeld, S. (2005): "Birth of a New Nation." *The Boston Globe* http://archive.boston.com/sports/baseball/redsox/articles/2005/07/01/birth_of_a_new_nation/?page=, Last accessed on 2021-7-16.

Hayward, G. M. & A. Rybińska (2017): "Super Bowl Babies: Do Counties with Super Bowl Winning Teams experience Increases in Births Nine Months Later?" *Socius* **3**: pp. 1–14.

He, Q. (2018): "Inflation and Fertility in a Schumpeterian Growth model: Theory and Evidence." *International Review of Economics and Finance* **58**: pp. 113–126.

Hockey Reference (2021): "NHL/WHA Playoffs." https://www.hockey-reference.com/, Last accessed on 2021-6-25.

Inoue, Y., M. Sato, K. Filo, J. Du, & D. Funk (2017): "Sport Spectatorship and Life Satisfaction: A Multi-Country Investigation." *Journal of Sport Management* **31**: pp. 1–40.

Janhuba, R. (2019): "Do Victories and Losses Matter? Effects of Football on Life Satisfaction." *Journal of Economic Psychology* **75**: p. 102102.

Johnson, D. (2020): "Record-breaking Liverpool set new standards with Premier League title win." *ESPN Online* https://www.espn.com/soccer/liverpool-engliverpool/story/4121943/record-breaking-liverpool-set-new-standards-with-premier-league-title-win, Last accessed on 2021-7-16.

Joyce, M., O. Michelle, K. Sharon, & G. Elizabeth (2015): "Measuring Gestational Age in Vital Statistics Data: Transitioning to the Obstetric Estimate." *National Vital Statistics Reports: the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System* **64**: pp. 1–20.

Jukic, A., D. Baird, C. Weinberg, D. Mcconnaughey, & A. Wilcox (2013): "Length of Human Pregnancy and Contributors to its Natural Variation." *Human Reproduction (Oxford, England)* **28**.

Lenhart, O. (2021): "The Effects of Minimum Wages on Teenage Birth rates." *Economics Letters* **198**: p. 109670.

Liao, P.-J. (2013): "The One-child Policy: A Macroeconomic Analysis." *Journal of Development Economics* **101**: pp. 49–62.

Lucas, R. (2014): "Life Satisfaction of US Counties Predicts Population Growth." *Social Psychological & Personality Science* **5**: pp. 383–388.

van der Meij, L., M. Almela, V. Hidalgo, C. Villada, H. Ijzerman, P. Lange, & A. Salvador (2012): "Testosterone and Cortisol Release among Spanish Soccer Fans watching the 2010 World Cup Final." *PloS one* **7**: p. e34814.

Montesinos, J., J. Cortes, A. Arnau, J. A. Sanchez, M. Elmore, N. Macia, & J. Bosch (2013): "Barcelona Baby Boom: Does Sporting Success Affect Birth rate?" *BMJ* **347**: p. 7938.

Nandi, A., S. Mazumdar, & Behrman (2018): "The Effect of Natural Disaster on Fertility, Birth Spacing, and Child Sex Ratio: Evidence from a Major Earthquake in India." *Journal of Population Economics* **31(1)**: pp. 267–293.

National Center for Health Statistics (2020): "Natality 1995-2019." https://wonder.cdc.gov/natality.html, Last accessed on 2021-6-28.

de Oliveira, V. H., I. Lee, & C. Quintana-Domeque (2021): "Natural Disasters and Early Human Development: Hurricane Catarina and Infant Health in Brazil." *Journal of Human Resources* pp. 0816–8144R1.

Pro-Football Reference (2021): "Super Bowl History." https://www.pro-football-reference.com/, Last accessed on 2021-6-25.

Rindfuss, R., J. Reed, & C. John (1978): "A Fertility Reaction to a Historical Event: Southern White Birth rates and The 1954 Desegregation Ruling." *Science (New York, N.Y.)* **201**: pp. 178–180.

Rodgers, J., C. John, & R. Coleman (2005): "Did Fertility go up after The Oklahoma City Bombing? An Analysis of Births in Metropolitan Counties in Oklahoma, 1990-1999." *Demography* **42**: pp. 675–92.

Rupp, H. & K. Wallen (2008): "Relationship between Testosterone and Interest in Sexual Stimuli: The Effect of Experience." *Hormones & Behavior* **52**: pp. 581–589.

Schaller, J., P. Fishback, & K. Marquardt (2020): "Local Economic Conditions and Fertility from the Great Depression through the Great Recession." *AEA Papers and Proceedings* **110**: pp. 236–240.

Schwartz, B., S. McDonald, & R. Kloner (2013): "Super Bowl Outcome's Association with Cardiovascular Death." *Clinical Research in Cardiology : the Official Journal of the German Cardiac Society* **102**.

Silver, M. (1965): "Births, Marriages, and Business Cycles in the United States." *Journal of Political Economy* **73**: pp. 237–237.

Sobotka, T., V. Skirbekk, & D. Philipov (2011): "Economic Recession and Fertility in the Developed World." *Population and Development Review* **37**: pp. 267–306.

Spong, C. (2013): "Defining "Term" Pregnancy Recommendations from the Defining "Term" Pregnancy Workgroup." *JAMA : Journal of the American Medical Association* **309**: pp. 1–2.

Sports Media Watch (2021a): "NBA Finals Ratings History (1988-Present)." https://www.sportsmediawatch.com/nba-finals-ratings-viewership-history/, Last accessed on 2021-7-25.

Sports Media Watch (2021b): "Super Bowl Ratings History (1967-present)." https://www.sportsmediawatch.com/super-bowl-ratings-historical-viewership-chart-cbs-nbc-fox-abc/, Last accessed on 2021-7-25.

SportsOddsHistory.com (2021): "Archived Futures Odds." https://www.sportsoddshistory.com/, Last accessed on 2021-6-28.

The Economist (2019): "As much as $26bn could be gambled on American sport in 2020." https://www.economist.com/graphic-detail/2019/12/29/as-much-as-26bn-could-be-gambled-on-american-sport-in-2020, Last accessed on 2021-7-25.

Tolchin, M. (1966): "Births Up 9 Months After the Blackout; Hospitals Report Rise in Births 9 Months After Power Failure." *The New York Times* https://www.nytimes.com/1966/08/10/archives/births-up-9-months-after-the-blackout-hospitals-report-rise-in.html, Last accessed on 2021-7-16.

Truong, Q.-T., Q.-N. Tran, W. Bakry, D. N. Nguyen, & S. Al-Mohamad (2021): "Football Sentiment and Stock Market Returns: Evidence from a Frontier Market." *Journal of Behavioral & Experimental Finance* **30**: p. 100472.

Udry, J. (1970): "The Effect of The Great Blackout of 1965 on Births in New York City." *Demography* **7**: pp. 325–7.

Vollset, S., E. Goren, C.-W. Yuan, J. Cao, A. Smith, T. Hsiao, C. Bisignano, G. Azhar, E. Castro, J. Chalek, A. Dolgert, T. Frank, K. Fukutaki, S. Hay, R. Lozano, A. Mokdad, V. Nandakumar, M. Pierce, M. Pletcher, & C. Murray (2020): "Fertility, Mortality, Migration, and Population Scenarios for 195 Countries and Territories from 2017 to 2100." *The Lancet* **396**.

Wooldridge, J. (2003): *Introductory Econometrics: A Modern Approach: 5th edition.* South-Western Cengage Learning.

Appendix A

Deciding Matches

	Date	*Winning team*	*Losing team*	*Score*
LIII	03.02.2019	**New England Patriots**	Los Angeles Rams	12 - 3
LII	04.02.2018	**Philadelphia Eagles**	New England Patriots	41 - 33
LI	05.02.2017	**New England Patriots**	Atlanta Falcons	34 - 28
L	07.02.2016	**Denver Broncos**	Carolina Panthers	24 - 10
XLIX	01.02.2015	**New England Patriots**	Seattle Seahawks	28 - 24
XLVIII	02.02.2014	**Seattle Seahawks**	Denver Broncos	43 - 8
XLVII	03.02.2013	**Baltimore Ravens**	San Francisco 49ers	34 - 31
XLVI	05.02.2012	**New York Giants**	New England Patriots	21 - 17
XLV	06.02.2011	**Green Bay Packers**	Pittsburgh Steelers	31 - 25
XLIV	07.02.2010	**New Orleans Saints**	Indianapolis Colts	31 - 17
XLIII	01.02.2009	**Pittsburgh Steelers**	Arizona Cardinals	27 - 23
XLII	03.02.2008	**New York Giants**	New England Patriots	17 - 14
XLI	04.02.2007	**Indianapolis Colts**	Chicago Bears	29 - 17
XL	05.02.2006	**Pittsburgh Steelers**	Seattle Seahawks	21 - 10
XXXIX	06.02.2005	**New England Patriots**	Philadelphia Eagles	24 - 21
XXXVIII	01.02.2004	**New England Patriots**	Carolina Panthers	32 - 29
XXXVII	26.01.2003	**Tampa Bay Buccaneers**	Oakland Raiders	48 - 21

Source: *pro-football-reference.com*

Table A.1: The Super Bowls 2003 - 2019

Date	*Winning team*	*Losing team*	*Serie*	*Score*
28.10.2018	**Boston Red Sox**	Los Angeles Dodgers	4 - 1	5:1
01.11.2017	**Houston Astros**	Los Angeles Dodgers	4 - 3	5:1
02.11.2016	**Chicago Cubs**	Cleveland Indians	4 - 3	8:7*
01.11.2015	**Kansas City Royals**	New York Mets	4 - 1	7:2*
29.10.2014	**San Francisco Giants**	Kansas City Royals	4 - 3	3:2
30.10.2013	**Boston Red Sox**	St. Louis Cardinals	4 - 2	6:1
28.10.2012	**San Francisco Giants**	Detroit Tigers	4 - 0	4:3*
28.10.2011	**St. Louis Cardinals**	Texas Rangers	4 - 3	6:2
01.11.2010	**San Francisco Giants**	Texas Rangers	4 - 1	3:1
04.11.2009	**New York Yankees**	Philadelphia Phillies	4 - 2	7:3
29.10.2008	**Philadelphia Phillies**	Tampa Bay Rays	4 - 1	4:3
28.10.2007	**Boston Red Sox**	Colorado Rockies	4 - 0	4:3
27.10.2006	**St. Louis Cardinals**	Detroit Tigers	4 - 1	4:2
26.10.2005	**Chicago White Sox**	Houston Astros	4 - 0	1:0
27.10.2004	**Boston Red Sox**	St. Louis Cardinals	4 - 0	3:0
25.10.2003	**Florida Marlins**	New York Yankees	4 - 2	2:0
27.10.2002	**Anaheim Angels**	San Francisco Giants	4 - 3	4:1

Source: *baseball-reference.com*

Table A.2: The World Series 2002 - 2018

Date	*Winning team*	*Losing team*	*Serie*	*Score*
08.06.2018	**Golden State Warriors**	Cleveland Cavaliers	4 - 0	108:85
12.06.2017	**Golden State Warriors**	Cleveland Cavaliers	4 - 1	129:120
19.06.2016	**Cleveland Cavaliers**	Golden State Warriors	4 - 3	93:89
16.06.2015	**Golden State Warriors**	Cleveland Cavaliers	4 - 2	105:97
15.06.2014	**San Antonio Spurs**	Miami Heat	4 - 1	104:87
20.06.2013	**Miami Heat**	San Antonio Spurs	4 - 3	95:88
21.06.2012	**Miami Heat**	Oklahoma City Thunder	4 - 1	121:106
12.06.2011	**Dallas Mavericks**	Miami Heat	4 - 2	105:95
17.06.2010	**Los Angeles Lakers**	Boston Celtics	4 - 3	83:79
14.06.2009	**Los Angeles Lakers**	Orlando Magic	4 - 1	99:86
17.06.2008	**Boston Celtics**	Los Angeles Lakers	4 - 2	131:92
14.06.2007	**San Antonio Spurs**	Clevland Cavaliers	4 - 0	83:82
20.06.2006	**Miami Heat**	Dallas Mavericks	4 - 2	95:92
23.06.2005	**San Antonio Spurs**	Detroit Pistons	4 - 3	81:74
15.06.2004	**Detroit Pistons**	Los Angeles Lakers	4 - 1	100:87
15.06.2003	**San Antonio Spurs**	New Jersey Nets	4 - 2	88:77
12.06.2002	**Los Angeles Lakers**	New Jersey Nets	4 - 0	113:107

Source: *basketball-reference.com*

Table A.3: The NBA Championships 2002 - 2018

Date	*Winning team*	*Losing team*	*Serie*	*Score*
07.06.2018	**Washington Capitals**	Vegas Golden Knights	4 - 1	4:3
11.06.2017	**Pittsburgh Penguins**	Nashville Predators	4 - 2	2:0
12.06.2016	**Pittsburgh Penguins**	San Jose Sharks	4 - 2	3:1
24.06.2015	**Chicago Blackhawks**	Tampa Bay Lightning	4 - 2	2:0
13.06.2014	**Los Angeles Kings**	New York Rangers	4 - 1	3:2*
24.06.2013	**Chicago Blackhawks**	Boston Bruins	4 - 2	3:2
11.06.2012	**Los Angeles Kings**	New Jersey Devils	4 - 2	6:1
15.06.2011	**Boston Bruins**	Vancouver Canucks	4 - 3	4:0
09.06.2010	**Chicago Blackhawks**	Philadelphia Flyers	4 - 2	4:3*
12.06.2009	**Pittsburgh Penguins**	Detroit Red Wings	4 - 3	2:1
04.06.2008	**Detroit Red Wings**	Pittsburgh Penguins	4 - 2	3:2
06.06.2007	**Anaheim Ducks**	Ottawa Senators	4 - 1	6:2
19.06.2006	**Carolina Hurricanes**	Edmonton Oilers	4 - 3	3:1
07.06.2004	**Tampa Bay Lightnings**	Calgary Flames	4 - 3	2:1
09.06.2003	**New Jersey Devils**	Anaheim Ducks	4 - 3	3:0
13.06.2002	**Detroit Red Wings**	Carolina Hurricanes	4 - 1	3:1

Source: *hockey-reference.com*

Table A.4: The Stanley Cup Finals 2002 - 2018

Appendix B

Home Stadiums

Team	Stadium	County	MSA
Baltimore Ravens	MT Bank Stadium	Baltimore County, MD	Baltimore-Columbia-Towson
Denver Broncos	Sports Authority Field at Mile High	Denver County, CO	Denver-Aurora-Lakewood
Green Bay Packers	Lambeau Field	Brown County, WI	Green Bay
Indianapolis Colts	RCA Dome	Marion County, IN	Indianapolis-Carmel-Anderson
New England Patriots	Gillette Stadium	Norfolk County, MA	Boston-Cambridge-Newton
New Orleans Saints	Louisiana Superdome	Orleans Parish, LA	New Orleans-Metairie
New York Giants	MetLife Stadium	Bergen County, NJ	New York-Newark-Jersey City
Philadelphia Eagles	Lincoln Financial Field	Philadelphia County, PA	Philadelphia-Camden-Wilmington
Pittsburgh Steelers	Heinz Field	Allegheny County, PA	Pittsburgh
Seattle Seahawks	MetLife Stadium	King County, WA	Seattle-Tacoma-Bellevue
Tampa Bay Buccaneers	Raymond James Stadium	Hillsborough County, FL	Tampa-St. Petersburg-Clearwater
Anaheim Angels	Edison International Field of Anaheim	Orange County, CA	Los Angeles-Long Beach-Anaheim
Boston Red Sox	Fenway Park	Suffolk County, MA	Boston-Cambridge-Newton
Chicago Cubs	Wrigley Field	Cook County, IL	Chicago-Naperville-Elgin
Chicago White Sox	U.S. Cellular Field	Cook County, IL	Chicago-Naperville-Elgin
Florida Marlins	Pro Player Stadium	Miami-Dade County, FL	Miami-Fort Lauderdale-Pompano Beach
Houston Astros	Minute Maid Park	Harris County, TX	Houston-The Woodlands-Sugar Land
Kansas City Royals	Kauffman Stadium	Jackson County, MO	Kansas City, MO-KS
New York Yankees	Yankee Stadium	Bronx County, NY	New York-Newark-Jersey City
Philadelphia Phillies	Citizens Bank Park	Philadelphia County, PA	Philadelphia-Camden-Wilmington
San Francisco Giants	AT&T Park	San Francisco County, CA	San Francisco-Oakland-Berkeley
St. Louis Cardinals	Busch Stadium	St. Louis County, MO	St. Louis
Boston Celtics	TD Banknorth Garden	Suffolk County, MA	Boston-Cambridge-Newton
Cleveland Cavaliers	Quicken Loans Arena	Cuyahoga County, OH	Cleveland-Elyria
Dallas Mavericks	American Airlines Center	Dallas County, TX	Dallas-Fort Worth-Arlington
Detroit Pistons	The Palace of Auburn Hills	Oakland County, MI	Detroit-Warren-Dearborn
Golden State Warriors	Oracle Arena	Alameda County, CA	San Francisco-Oakland-Berkeley
Los Angeles Lakers	Staples Center	Los Angeles County, CA	Los Angeles-Long Beach-Anaheim
Miami Heat	American Airlines Arena	Miami-Dade County, FL	Miami-Fort Lauderdale-Pompano Beach
San Antonio Spurs	AT&T Center	Bexar County, TX	San Antonio-New Braunfels
Anaheim Ducks	Honda Center	Orange County, CA	Los Angeles-Long Beach-Anaheim
Boston Bruins	TD Garden	Suffolk County, MA	Boston-Cambridge-Newton
Carolina Hurricanes	RBC Center	Wake County, NC	Raleigh-Cary
Chicago Blackhawks	United Center	Cook County, IL	Chicago-Naperville-Elgin
Detroit Red Wings	Joe Louis Arena	Wayne County, MI	Detroit-Warren-Dearborn
Los Angeles Kings	Staples Center	Los Angeles County, CA	Los Angeles-Long Beach-Anaheim
New Jersey Devils	Continental Airlines Arena	Bergen County, NJ	New York-Newark-Jersey City
Pittsburgh Penguins	PPG Paints Arena	Allegheny County, PA	Pittsburgh
Tampa Bay Lightning	St. Pete Times Forum	Hillsborough County, FL	Tampa-St. Petersburg-Clearwater
Washington Capitals	Capital One Arena	District of Columbia, DC	Washington-Arlington-Alexandria

Source: The Internet

Table B.1: Home Stadiums Location

Appendix C

Metropolitan Statistical Areas

MSA	a	NFL	MLB	NBA	NHL	n
Baltimore-Columbia-Towson	*1*	1	0	0	0	1
Boston-Cambridge-Newton	*2*	1	1	1	1	4
Chicago-Naperville-Elgin	*3*	0	2	0	1	3
Cleveland-Elyria	*4*	0	0	1	0	1
Dallas-Fort Worth-Arlington	*5*	0	0	1	0	1
Denver-Aurora-Lakewood	*6*	1	0	0	0	1
Detroit-Warren-Dearborn	*7*	0	0	1	1	2
Green Bay	*8*	1	0	0	0	1
Houston-The Woodlands-Sugar Land	*9*	0	1	0	0	1
Indianapolis-Carmel-Anderson	*10*	1	0	0	0	1
Kansas City	*11*	0	1	0	0	1
Los Angeles County	*12*	0	0	1	1	2
Miami-Fort Lauderdale-Pompano Beach	*13*	0	0	1	1	2
New Orleans-Metairie	*14*	1	0	0	0	1
New York-Newark-Jersey City	*15*	1	1	0	1	3
Philadelphia-Camden-Wilmington	*16*	1	1	0	0	2
Pittsburgh	*17*	1	0	0	1	2
Orange County	*18*	0	1	0	1	2
Raleigh-Cary	*19*	0	0	0	1	1
San Antonio-New Braunfels	*20*	0	0	1	0	1
San Francisco-Oakland-Berkeley	*21*	0	1	1	0	2
Seattle-Tacoma-Bellevue	*22*	1	0	0	0	1
St. Louis	*23*	0	1	0	0	1
Tampa-St. Petersburg-Clearwater	*24*	1	0	0	1	2
Washington-Arlington-Alexandria	*25*	0	0	0	1	1
$\lvert MSA \rvert = 25$		11	10	8	11	40

Note: the Author's creation

Table C.1: Distribution of Winners by MSA

Appendix D

Composition of Control Variables

Variable	Source	Notes
Unemployment rate	US Bureau of Labor Statistics	collected monthly
House price index (HPI)	Federal Reserve Economic Data, FRED, St. Louis Fed	Jan 2000 = 100 , collected quarterly

Table D.1: List of Control Variables

Appendix E

Complete Results

	Dependent variable:					
	ln(Births)					
comp	0.375**	0.230	0.002	0.002	0.002	−0.0004
	(0.153)	(0.155)	(0.007)	(0.007)	(0.007)	(0.012)
lag(Unemp, k = 9)		0.145	−0.002	−0.002	−0.002	−0.002
		(0.090)	(0.003)	(0.003)	(0.003)	(0.004)
lag(log(HPI), k = 9)		1.504***	0.148***	0.148***	0.148***	0.165***
		(0.531)	(0.042)	(0.042)	(0.042)	(0.060)
Constant	8.124***	−0.272				
	(0.176)	(2.904)				
Month dummy	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	No	No	Time	MSA	Time and MSA	Time and MSA
Observations	5,100	5,100	5,100	5,100	5,100	2,856
R^2	0.006	0.228	0.532	0.532	0.532	0.485
Adjusted R^2	0.0003	0.223	0.527	0.527	0.527	0.477
Note:						*p<0.1; **p<0.05; ***p<0.01

Table E.1: Regression Results - BT method

	Dependent variable:					
	ln(Births)					
nfl	−0.021	−0.167	−0.015	−0.015	−0.015	−0.005
	(0.210)	(0.146)	(0.009)	(0.009)	(0.009)	(0.024)
mlb	0.422**	0.139	0.017	0.017	0.017	0.024**
	(0.176)	(0.199)	(0.011)	(0.011)	(0.011)	(0.012)
nba	0.326	0.177	−0.007	−0.007	−0.007	−0.024**
	(0.266)	(0.225)	(0.013)	(0.013)	(0.013)	(0.012)
nhl	0.374	0.390*	−0.006	−0.006	−0.006	0.019**
	(0.241)	(0.223)	(0.010)	(0.010)	(0.010)	(0.009)
lag(Unemp, k = 9)		0.146	−0.002	−0.002	−0.002	−0.002
		(0.090)	(0.003)	(0.003)	(0.003)	(0.004)
lag(log(HPI), k = 9)		1.506***	0.148***	0.148***	0.148***	0.164***
		(0.531)	(0.042)	(0.042)	(0.042)	(0.060)
Constant	8.124***	−0.282				
	(0.176)	(2.905)				
Month dummy	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	No	No	Time	MSA	Time and MSA	Time and MSA
Observations	5,100	5,100	5,100	5,100	5,100	2,856
R^2	0.006	0.228	0.532	0.532	0.532	0.485
Adjusted R^2	-0.0003	0.223	0.527	0.527	0.527	0.477
Note:						*p<0.1; **p<0.05; ***p<0.01

Table E.2: Regression Results - Individual Leagues - BT method

	Dependent variable:					
	ln(Births)					
comp	0.380**	0.232	0.004	0.004	0.004	0.011
	(0.153)	(0.156)	(0.007)	(0.007)	(0.007)	(0.011)
Unemp		0.151	−0.001	−0.001	−0.001	−0.001
		(0.092)	(0.003)	(0.003)	(0.003)	(0.004)
log(HPI)		1.533***	0.160***	0.160***	0.160***	0.172***
		(0.529)	(0.041)	(0.041)	(0.041)	(0.062)
Constant	8.035***	−0.515				
	(0.175)	(2.892)				
Month dummy	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	No	No	Time	MSA	Time and MSA	Time and MSA
Observations	5,100	5,100	5,100	5,100	5,100	2,856
R^2	0.010	0.238	0.696	0.696	0.696	0.703
Adjusted R^2	0.005	0.233	0.693	0.693	0.693	0.699
Note:						*p<0.1; **p<0.05; ***p<0.01

Table E.3: Regression Results - GAT method

	Dependent variable:					
	ln(Births)					
nfl	−0.015	−0.164	−0.010	−0.010	−0.010	0.002
	(0.214)	(0.147)	(0.013)	(0.013)	(0.013)	(0.031)
mlb	0.427**	0.137	0.014	0.014	0.014	0.025
	(0.176)	(0.199)	(0.011)	(0.011)	(0.011)	(0.020)
nba	0.335	0.182	−0.001	−0.001	−0.001	−0.003
	(0.262)	(0.222)	(0.012)	(0.012)	(0.012)	(0.015)
nhl	0.376	0.393*	−0.003	−0.003	−0.003	0.025***
	(0.244)	(0.229)	(0.011)	(0.011)	(0.011)	(0.006)
Unemp		0.151	−0.001	−0.001	−0.001	−0.001
		(0.092)	(0.003)	(0.003)	(0.003)	(0.004)
log(HPI)		1.535***	0.160***	0.160***	0.160***	0.171***
		(0.529)	(0.041)	(0.041)	(0.041)	(0.062)
Constant	8.035***	−0.524				
	(0.175)	(2.893)				
Month dummy	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	No	No	Time	MSA	Time and MSA	Time and MSA
Observations	5,100	5,100	5,100	5,100	5,100	2,856
R^2	0.010	0.238	0.696	0.696	0.696	0.703
Adjusted R^2	0.004	0.233	0.692	0.692	0.692	0.698
Note:						*p<0.1; **p<0.05; ***p<0.01

Table E.4: Regression Results - Individual Leagues - GAT method

	Dependent variable:			
	ln(Births)			
	BT		GAT	
	(1)	(2)	(3)	(4)
Unexp	0.010	−0.001	0.013	−0.0003
	(0.010)	(0.028)	(0.010)	(0.034)
Unique	0.004	0.018	−0.00003	0.028
	(0.009)	(0.028)	(0.010)	(0.034)
Dram	−0.019***	−0.021	−0.011	−0.028
	(0.007)	(0.015)	(0.008)	(0.017)
lag(Unemp, k = 9)	−0.002	−0.002		
	(0.003)	(0.004)		
lag(log(HPI), k = 9)	0.148***	0.164***		
	(0.042)	(0.060)		
Unemp			−0.001	−0.001
			(0.003)	(0.004)
log(HPI)			0.160***	0.171***
			(0.041)	(0.062)
Month dummy	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes
Fixed effects	Time and MSA	Time and MSA	Time and MSA	Time and MSA
Observations	5,100	2,856	5,100	2,856
R^2	0.532	0.485	0.696	0.704
Adjusted R^2	0.527	0.477	0.692	0.699
Note:				*p<0.1; **p<0.05; ***p<0.01

Table E.5: Heterogeneity Effect - General

	Dependent variable:			
	ln(Births)			
	BT		GAT	
	(1)	(2)	(3)	(4)
Unexp_nfl	0.014	0.007	0.011	−0.008
	(0.013)	(0.022)	(0.017)	(0.029)
Unique_nfl	−0.005	0.018	−0.014	0.041
	(0.017)	(0.034)	(0.022)	(0.042)
Dram_nfl	−0.039***	−0.076***	−0.021	−0.110***
	(0.011)	(0.024)	(0.014)	(0.033)
Unexp_mlb	0.009	0.036***	0.012	0.064***
	(0.023)	(0.008)	(0.023)	(0.010)
Unique_mlb	0.013		0.012	
	(0.014)		(0.013)	
Dram_mlb	0.007	−0.015	0.0001	−0.052**
	(0.024)	(0.016)	(0.023)	(0.021)
Unexp_nba	−0.044*	−0.056***	−0.036***	−0.026
	(0.025)	(0.019)	(0.011)	(0.018)
Unique_nba	0.046	0.059*	0.043***	0.036
	(0.032)	(0.033)	(0.016)	(0.037)
Dram_nba	−0.017	−0.022	0.0003	−0.004
	(0.013)	(0.019)	(0.011)	(0.018)
Unexp_nhl	0.018	0.019**	0.029**	0.026***
	(0.013)	(0.009)	(0.015)	(0.006)
Unique_nhl	−0.003		−0.010	
	(0.012)		(0.014)	
Dram_nhl	−0.018		−0.013	
	(0.017)		(0.019)	
lag(Unemp, k = 9)	−0.002	−0.002		
	(0.003)	(0.004)		
lag(log(HPI), k = 9)	0.148***	0.164***		
	(0.042)	(0.060)		
Unemp			−0.001	−0.001
			(0.003)	(0.004)
log(HPI)			0.160***	0.171***
			(0.041)	(0.063)
Month dummy	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes
Fixed effects	Time and MSA	Time and MSA	Time and MSA	Time and MSA
Observations	5,100	2,856	5,100	2,856
R^2	0.533	0.486	0.696	0.704
Adjusted R^2	0.527	0.476	0.692	0.699
Note:				*p<0.1; **p<0.05; ***p<0.01

Table E.6: Heterogeneity Effect - Individual Leagues