Human Capital, Negative Health Shocks and the Retirement Decisions of Peers: Evidence from the NFL

Joshua C. Martin*

April 2, 2023

This paper examines the effect of negative health shocks on the retirement decisions of peers. I do so by gathering detailed data on players from the National Football League (NFL) from 1980-2019. The introduction and advancement of research into chronic traumatic encephalopathy (CTE), coupled with increasing early-life mortality among NFL players, has significantly increased the salience of American football's long-term negative health effects. I exploit the quasi-random timing of early-life, CTE-related mortality of both former and current players on the retirement decisions of their teammates. I demonstrate that these players are significantly more likely to retire than comparable peers in the seasons following one of these deaths. This magnitude of this effect larger for players with fewer years of college attendance, but is uncorrelated with the amount of time shared as teammates.

---

*West Virginia University. Email: jcm0067@mix.wvu.edu. I am grateful for the tremendous help provided by two colleagues without whom this paper would not exist. Clay Collins provided the initial research idea and aided in identifying the relevant literature and important background information. My former student, Logan Tucker, analyzed the impact of the death of Junior Seau on the retirement decisions of his teammates as his undergraduate capstone project. His excellent work was the inspiration for this paper.

1 — Introduction

While both the personal and social consequences of retirement are relatively well documented, its causes are far less understood. I help to contribute to the literature on voluntary retirement decisions by examining the impact of negative health shocks on the retirement choices of peers. To do so, I gather and analyze detailed player-by-season data of professional American football players from 1980-2019 to investigate how increasing awareness of the link between work and long-term neurological damages has influenced labor supply.

First, I analyze the general determinants of retirement in this setting. I document four important variables: the expected probability of winning the Super Bowl, player productivity, age, and the number of games played with only age being positively correlated with the probability of retirement. However, this exercise explains relatively little of the variation in retirements.

Thus, next, I analyze the causal impact of negative health shocks on the retirement decisions of peers. I do so by relying on the quasi-random timing of chronic traumatic encephalopathy (CTE)-related deaths among both former and current players. Comparing the retirement decisions of former teammates of these recently-deceased athletes to their peers before and after the timing of these deaths, I find that this news shock increased voluntary retirements by 2-7 percentage points (10-30%).

In investigating the mechanisms of this result, I fail to reject the hypothesis that the length of time spent as teammates has an additive effect on the probability of retirement. However, I present evidence that this effect may be dwarfed by the retirement decisions of players without a college degree. Given the relatively young ages of these athletes, I provide evidence that human capital and occupation-specific negative health shocks have important interactions.

2 — Literature Review

2.1. American Football & Mortality Risk

While the brain-injury risks of participating in American football have been well-documented for many years, only recently has evidence arisen which suggests that, on average, these risks outweigh the benefits of income and elite physical fitness on health.

The early findings within the literature attempting to quantify the relationship between mortality risks and American football participation are mixed due to differences in selecting comparison units. Many papers written earlier in this literature rely upon either matching methods and/or multivariate regression analysis to "control" for differences between the observable characteristics of NFL players to the general population. For instance, Koning et al. (2014) find that playing in the NFL is associated with increased longevity, but that this advantage diminishes as athletes play in more games over the course of their careers. These findings were later confirmed as a part of a meta-analysis on the subject (Owora et al., 2018). Additionally, Lehman et al. (2016) find no evidence of elevated suicide risk among veteran, professional, American football players compared to the general population.

However, it would be incorrect to assume from these studies that there are positive health effects of American football participation. Professional American football players have much greater levels of physical fitness, significantly higher incomes, and face meaningfully stronger (indirect) financial incentives both to have better nutrition and to abstain from unhealthy behaviors over the course of their playing careers. These differences with the general population are plausibly large enough that the aforementioned statistical techniques may be either inadequate or underpowered (in terms of finding suitable comparison units) to answer such questions.

Venkataramani et al. (2018) sidestep some of these issues by using comparison units of "replacement players" who were temporarily hired to play during a league-wide player strike in 1987. They found higher, but statistically insignificant, mortality risks among (career) NFL players.[1] Relatedly, Nguyen et al. (2019) use a comparison unit of professional baseball players – individuals who are more likely to have similar income, fitness, and access to health care. They find that NFL players have significantly higher mortality rates from all causes (with neuro-degenerative diseases being particularly pronounced) when compared with their baseball-playing peers.

Further recent research has corroborated these findings. LeClair et al. (2022) use a comparison unit of high-school-level players and find that professionals are nearly 2.5 times more likely to develop CTE. Daneshvar et al. (2021) finds that these athletes had incidence and mortality rates of ALS that were 4 times greater than comparable (on observables) American males. Among NFL players, those who eventually developed ALS had careers that were 2.5 years longer. These studies support the idea that there is a dose-response relationship between American football participation and mortality risk.

2.2. Athlete's (Voluntary) Retirement Decisions

Within a traditional lifecycle framework in economics, agents maximize expected lifetime utility subject to their lifetime budget constraints. Worker's utility is a function of consumption (as determined by labor productivity) and leisure. As labor productivity diminishes over time, agents increasingly substitute time spent working for leisure until they retire (Gustman and Steinmeier, 1983).

In working environments such as American football where the amount of labor supplied by workers (playing time) is outside of their control, retirement decisions are discrete rather than continuous as measured by whether they work under contract with a team during that season. The surplus of players who are willing to play in the NFL relative to the number of positions available on a professional roster suggests

1. This identification is also likely important given the unobservable selection into playing football professionally. Allison et al. (2018) and White et al. (2021) document large racial and socio-economic background differences for NFL players which are unlikely to be represented in most datasets, thus the biasing results of more naive comparisons.

that the majority of retirement decisions are non-voluntary and productivity-related (Erpič et al., 2004; Fernandez et al., 2006). However, when players receive negative informational health shocks that are associated with their employment, this can influence both one's time horizon and their relative preference for the ratio of leisure to labor consumption. In both cases, we would expect to see increases in retirements.

Many important factors influence the voluntary retirement decisions of athletes. One such component is health. Worker health and physical ability are important determinants of retirement decisions as poor health or disability can either make it difficult to continue working or even require workers to retire earlier than initially planned. A large literature has demonstrated the impact of health shocks on workforce exits (Disney et al., 2006; Hagan et al., 2008; Jones et al., 2010) though the general decline of health with age explains only a small share of the labor supply response (French and Jones, 2017; Blundell et al., 2023).

Another related component is worker's finances (Park et al., 2013). The pervasiveness of workers' inability to save adequate amounts of income for retirement has led to a rise in both workplace and national pension plans. These suboptimal savings behaviors are particularly pronounced for those who are younger and have lower levels of financial literacy (Lusardi and Mitchelli, 2007). Given the combination of the extremely young average retirement age (and thus the expectation of substitution to other vocations), the unusually high levels of income, and the overrepresentation of athletes from lower socioeconomic areas, professional American football players are likely to be even less responsive to their financial conditions concerning retirement decisions than otherwise similar workers (Allison et al., 2018; White et al., 2021).

Brown and Laschever (2012) provide the most related research to this article. In their examination of teachers in Los Angeles, Brown and Laschever find strong evidence of peer effects on retirement decisions. The authors conclude that the retirement of a fellow teacher increases the likelihood of retirement by 1.5-2 percentage points.

3 — Data

The data used for this study come from `Pro-Football-Reference.com` (PFR). At the time of this writing, PFR provides the most comprehensive publicly-available for both detailed player-by-game and player-by-season data for professional and collegiate American football going back to the 1920s. Crucially, it also contains information on both active and retired players – including their year of death. I gather detailed playing data for every player who played at least one play or more during their NFL career from 1980 to 2020 in order to examine the determinants of retirement. When available and/or applicable, I also link this information to their date of death as well as the number of years they participated on a college football team.

The start date of 1980 was chosen in order to capture every player who could potentially have been impacted by the death of a former teammate who was formally diagnosed with CTE. The player with the most NFL experience following the first such of these publicly-confirmed cases began playing in 1980. The 2019-2020 season is chosen as the end date in order to avoid any confounding factors dealing with retirement that could have arisen from the COVID-19 pandemic.

Most crucially, the data include a variable titled "Approximate Value" (AV) which proxies for the productivity of each player for each season which is standardized to be easily comparable across both different positions and seasons. In the absence of a league-wide salary data (which isn't publicly-available until 2019), a measure which approximates the marginal product of labor for each player is highly valuable for attempting to answer this question.

The data used in this paper is at a player-by-season level. These include information on the player's player's age, the number of games played / started, their productivity, whether they made a pro-bowl or all-pro team, the position that they

Table 1 — Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Max
Age	70,723	26.546	3.335	20	47
Experience	70,723	3.507	3.295	0	25
Games Played	70,723	11.834	4.959	0	17
Games Started	70,723	5.947	6.340	0	16
Approx. Value	70,723	3.530	3.601	-6	26
Retire	70,723	0.201	0.401	0	1
p(Super Bowl)	70,723	0.033	0.032	0.001	0.209
Pro-Bowl	70,723	0.056	0.230	0	1
All-Pro	70,723	0.014	0.119	0	1
Years of College	41,791	3.100	1.030	1	6

most specialize in playing and their current team.[2] Players which change teams mid-season have the data from these observations aggregated so as to only include one observation.[3] The information for both the pre-season Super Bowl probabilities and the years of college completed were also accessed via PFR. The later is calculated by scraping the PFR web-links for each player's profile which contain links to similarly-detailed data for said player's college careers located on PFR's sister website SportsReference.com (`https://www.sports-reference.com/cfb`).

Table 1 provides the summary statistics. I observe 70,723 player-by-season observations in the data. This corresponds to 16,090 players over twenty seasons. The

2. Given the hyper-specialization of American football at the professional level, careful attention is paid to assigning players to a more general position so that accurate comparisons can be made in the analysis. Athletes are assigned to the most similar position of: Center (C), Guard (G), Tackle (T), Tight End (TE), Full Back (FB), Running Back (RB), Quarterback (QB), Wide Receiver (WR), Long Snapper (LS), Kicker (K), Punter (P), Defensive End (DE), Defensive Tackle (DT), Linebacker (LB), Cornerback (CB) or Safety (S). A deatiled breakdown of differences in average salary by position can be seen in Figure A1.

3. These individuals are assigned a *Team* variable value of "n-Team" where $n =$ the number of teams the athlete played for that season. This value is assigned after teammates have been identified so that players who were teammates for less than one season with an athlete who was later diagnosed with CTE are counted.

average player plays 3.5 years in the league. However, this is heavily right-skewed. Figure A2 examines a sub-sample of players in the NFL who played from 1980-2019 but retired by the start of the 2020 season. The figure shows that a majority of players do not remain in the league beyond their rookie year.

Table 1 shows that the average age of NFL players in the sample is 26.5 years old and has 3.5 years of experience in the league. The mean number of games played in per season is approximately 12 while the average number of games started is 6. The mean productivity of each player takes a value of 3 with 5.6% and 1.4% of players making the pro-bowl and all-pro teams respectively.[4] Approximately 20% of the sample retires in every season. The average pre-season probability of winning the Super Bowl is 3.3%.[5] The average player in the NFL has approximately 3 years of college experience as measured by the difference between the first and last year that statistics were recorded for their college careers.

4 — Identification

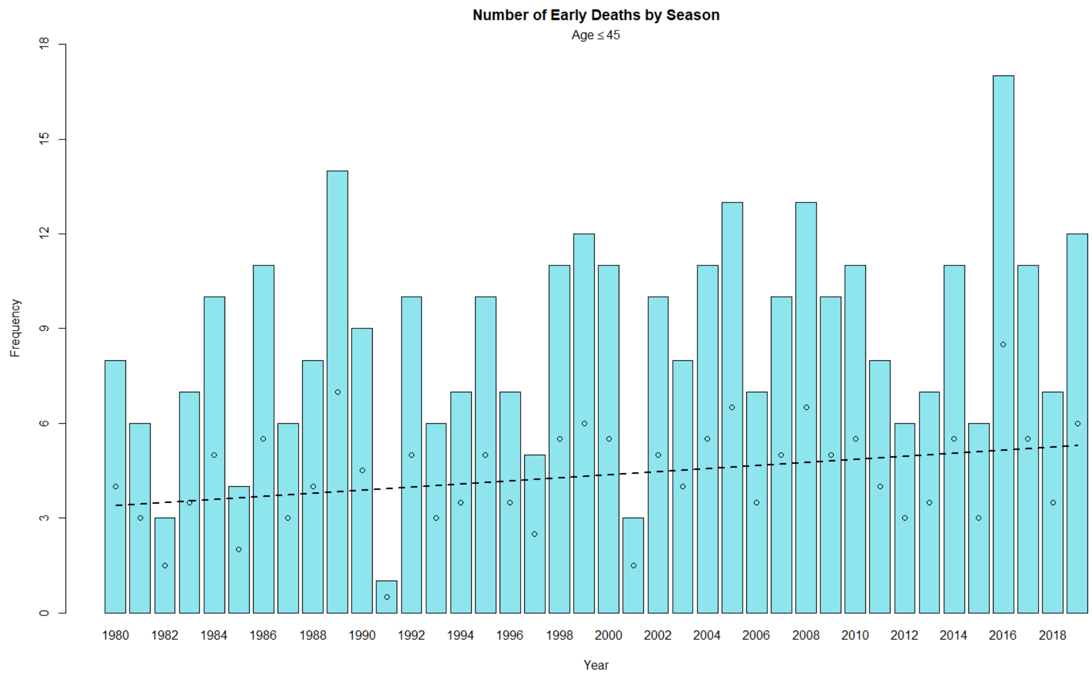
PFR contains data on both the date of death and age at the time of death for every current and former NFL player. I subset this data to include "early deaths" – those where current or former players die before the age of 45. Figure 1 demonstrates that these are on the rise by plotting out the number of these early deaths over time.

Developments in the identification of chronic traumatic encephalopathy (CTE) came to prominence within the context of growing numbers of early deaths of professional American football players. Most notably, one of the first confirmed and widely-notarized cases of CTE came from a post-mortem analysis of the brain of former Pittsburgh Steelers and Kansas City Chiefs player Mike Webster in 2002. Ta-

4. Figure A3 shows that this AV is strongly predictive of player's probability of earning these awards.

5. This variable is adjusted to account for book-keeper over-rounding. A time-series relationship of over-rounding by season can be seen in Figure A4.

Figure 1 — Frequency of Early Deaths of NFL Players



ble 2 presents the subset of players from Figure 1 who died at a young enough age to have been diagnosed with CTE while more than twenty of one's former teammates still played in the NFL. Table 2 shows information on the players, the number of teams they played for, the year they died, the number of former teammates that remained in the league 3 years before their death, and whether CTE was a prominently-credited factor that influenced their early death.

Given that most confirmed cases of CTE are only able to be confirmed following one's death, I rely upon the quasi-random timing of these incidents to estimate the causal impact of negative informational health shocks on one's labor supply at the extensive margin. Given that these information shocks presumably affect everyone, the estimates provided in the paper should be best understood as a dose-response effect whereby those with either greater prior familiarity with (and/or personal connections to) the subject of the news exhibit greater responses.

Table 2 — CTE Diagnoses Among Players with NFL-Active Former Teammates

Player	Teams	Career	Died	Teammates	Cause of Death
Mike Webster	2	1980-1990	2002	22	✓
Justin Strzelczyk	1	1980-1998	2004	55	✓
Chris Henry	1	2005-2009	2009	150	
Shane Dronett	3	1992-2001	2009	42	✓
Jovan Belcher	1	2009-2012	2012	127	✓
Junior Seau	3	1990-2009	2012	144	✓
Paul Oliver	1	2008-2011	2013	104	✓
Tyler Sash	1	2011-2012	2015	77	✓
Adrian Robinson	3	2012-2013	2015	145	✓
Aaron Hernandez	1	2010-2012	2017	48	✓
Daniel Te'o-Nesheim	2	2011-2013	2017	115	✓

Each individual listed represents a former NFL player who died at a young enough age to be diagnosed with CTE post-mortem and have former teammates still playing in the NFL. CTE was the prominently credited factor influencing the player's death in only one of the eleven incidents.

While the quasi-random nature of the identification strategy is helpful for (theoretically) balancing unobservable factors which could influence the outcome variable of interest, this analysis could still be biased if there are systemic differences between treated (former teammates who are still in the league) and control groups (everyone else). For instance, much research into retirement decisions focuses on access to health insurance.[6] Luckily (in terms of identification), all current players are compensated with full health insurance coverage as long as they are employed by an NFL team with more generous policies in place for those with longer playing careers. For instance, "vested" former players (those who played for three years or more) receive similar coverage as when they were playing for five years following their retirement. Former players can also begin receiving pensions at age 55 with the amount paid being determined by the number of seasons completed.

6. This is particularly pronounced in the United States where a sizable share of the population's health insurance is provided by their employer.

An equally concerning factor for identification is the issue of non-voluntary exits from the profession. Figure A2 shows a negative exponential relationship in the share of the total number of seasons played by each player. This relationship is likely driven by several factors including injury risk, competition from a growing supply of replacement athletes, and the near monopsony power that the NFL team owners enjoy (Rosen and Sanderson, 2001). Given these relatively high turnover rates, team owners suppress wages well below their market equilibrium amount in part due to the existence of rookie-maximum deals which stipulate pre-agreed upon salaries for players entering the league as determined by their draft order.[7] Given the restrictiveness of these rookie wages, players often enjoy massive increases in both their guaranteed and non-guaranteed wages due to the inter-team competition that arises at the end of the period when the team that drafted the player loses their exclusive rights to their talents. Most rookie deals end after the player's first four years in the league. Thus, the majority of voluntary exits from the profession occur if the player can reach the five-year mark of their career.

In attempting to address both the issues of health insurance coverage and non-voluntary exits, I rely upon proxies of the number of years of experience in the NFL as both non-rookie contracts and insurance benefits begin three to four seasons after one's career begins.

5 — Methods

I rely upon two broad categories of methods to analyze the retirement decisions of NFL players. The first method used is traditional multivariate regression analysis. This is used to measure the degree of the relationship between a group of independent variables with one dependent variable. Using this method, in conjunction with a wide array of fixed effects and estimation methods, calculates the predicted

7. The NFL uses a "draft" whereby teams select amateur athletes who often use their high school and college football teams to broadcast their skills.

average marginal effect of each variable on the expected likelihood of retiring.

However, this analysis can only reveal the correlates of retirement rather than what causes it. To answer whether negative informational health shocks lead players to retire earlier than they would have otherwise, I employ a stacked difference-in-differences estimator.[8] The primary benefit of using this stacking method is that it helps to avoid biased estimates that may arise when using difference-in-differences estimators when treatment events are staggered over time and when treatment effects differ as identified in Goodman-Bacon (2021). Baker et al. (2022) perform simulation analyses of various estimation techniques commonly used to correct these issues, including stacking. The results of their analysis indicate that stacking can accurately recover the true treatment path when used in event-study settings.

The empirical approach in this paper takes the following steps. First, I select a time window for the analysis of seven years. This provides four years of pre-treatment observations to test the parallel trends assumption in the event study analysis. This leaves three years remaining for the estimation of post-treatment effects. Second, I "slice" the data such that it only includes the relative event time window for the first group of treated players. Third, within this stack, a linear probability model is used to estimate how similar treated and control units are along observable characteristics.[9] The output is a continuous variable that helps in the construction of control groups which are more similar (along obsverable characteristics) to treated

8. The "stacking" process of involves organizing it based on event time rather than calendar time. To do this, a time window is chosen that balances the need for pre- and post-treatment periods with the availability of suitable control units. Then, a subset of the data is selected from this window, containing treated teammates and as many not-yet (or never-treated) control players as possible. This is done for each treatment event, and the resulting subsets are combined using a unique stack identifier. The result is that each treatment event is centered at the same relative time, allowing for unbiased difference-in-differences estimates using OLS with stack, unit, and time fixed effects.

9. Specifically, I regress whether a player was a teammate with someone who was diagnosed with CTE on player age, games played, games started, number of seasons played, AV, the expectation of winning the super-bowl that year, and whether they were named to the Pro Bowl of All-Pro teams. This is particularly useful as a non-arbitrary method for filtering out younger players from acting as control units if the remaining teammates happen to be older. Additionally, it allows for flexibility across different former teammate groups. For instance, Junior Seau's former teammates were primarily still playing with the New England Patriots who were perennial Super Bowl contenders.

cohorts. Fourth, the unique identifiers for each player that acted as a treated unit are stored so that they can be removed from acting as a control group later in the future. Fifth, the data are "stacked" into one dataset. This process is repeated for each treated cohort of players. The resulting dataset contains treatment timing for each cohort that is centered at the same relative treatment time.

6 — Results

6.1. Determinants (Correlates) of Retirement

Table 3 contains the results for the correlates of retirement by analyzing 70,000 player-by-season observations for every NFL player from 1980 to 2019. Each model contains common elements of both the binary dependent variable which identifies whether a player retired at the end of the season and *Year*, *Position*, and *Team* fixed effects. Models 1 and 2 are linear probability models that are estimated using Ordinary Least Squares (OLS). These models differ in that model 2 uses a *Player* fixed-effect.[10] Models 3 and 4 are the same in structure to model 1 but use different estimation techniques. Model three uses a probit model while model 4 relies on logit. The output of each of these regressions is reported as the average marginal effect for ease of interpretability.

Interpreting the coefficients of models 3 and 4 in conjunction with the patterns of statistical significance across each of the different specifications in Table 3 reveals several important variables which correlate with changes in the probability of retirement. Four variables of interest emerge. Interestingly, changes in the pre-season expected probability of winning the Super Bowl (as determined by betting markets) reflect the largest changes with a one percentage point (standard deviation) increase in the odds of winning the Super Bowl corresponding to a 34.5% (110%) reduction in the probability of retiring. Player productivity (*AV*) is estimated to have the second

10. *Age* is omitted from model 2 as it is co-linear with the *Player* fixed-effect.

Table 3 — Determinants of Retirement

Dependent Variable:	Retire			
Model:	(1)	(2)	(3)	(4)
	OLS	OLS	Probit	Logit
Variables				
Age	0.028***		0.028***	0.028***
	(0.000)		(0.002)	(0.000)
Games	-0.025***	-0.014***	-0.014***	-0.015***
	(0.000)	(0.000)	(0.000)	(0.000)
Games Started	-0.004***	-0.006***	-0.000	0.000
	(0.000)	(0.001)	(0.003)	(0.000)
Approx. Value	-0.015***	-0.008***	-0.036***	-0.041***
	(0.000)	(0.001)	(0.001)	(0.001)
p(Super Bowl)	-0.354***	-0.358***	-0.346***	-0.345***
	(0.045)	(0.052)	(0.045)	(0.046)
Pro Bowl	0.017**	-0.003	-0.018	-0.047**
	(0.008)	(0.008)	(0.013)	(0.016)
All Pro	0.065***	0.064***	-0.033	-0.084
	(0.013)	(0.013)	(0.040)	(0.057)
Fixed-effects				
Player		✓		
Year	✓	✓	✓	✓
Position	✓	✓	✓	✓
Team	✓	✓	✓	✓
Fit statistics				
Pseudo R^2	0.272	0.690	0.263	0.264

Heteroskedasticity-robust standard-errors in parentheses. n = 70,723
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*
Each of the coefficients and standard errors in models 3 and 4 are transformed the represent the average marginal effects.

largest effect with a one unit (standard deviation) increase corresponding to a 3.9% (13.9%) decrease in the probability of retiring. The third largest effect size is player age with a one year (standard deviation) increase corresponding to a 2.8% (9.3%) increase in the likelihood of retirement. Finally, an increase of one game (standard deviation) played in per season corresponds to a 1.5% (7.2%) reduction in the likelihood of exiting.[11]

Each of the other variables estimated in Table 3 is found to have inconsistent statistical significance, differing coefficient signs, and/or sufficiently small absolute effects such that they likely have trivial marginal explanatory power of variations of the dependent variable. Otherwise, it is noteworthy how few variables correspond to increases in the probability of retirement. This finding, coupled with the relatively low total explanatory power of the regressions as seen in the pseudo-R^2 variables, suggests that important omitted variables remain. For instance, even though both *Games* and *AV* should theoretically account for the effect of injuries on both the amount of time missed and changes in productivity, these will only be captured if the injury occurs within the season. Both the causes and severity of off-season injuries are unobservable with this data. Further, the length of players' contracts is also unobservable. This is true even amongst players on their first contracts given labor contract clauses such as the Franchise Tag whereby teams can effectively retain a player indefinitely by paying them maximum wages (as determined by the collective bargaining agreement) on yearly contracts.

Despite these limitations, the presence of voluntary retirement decisions amongst former players that occur neither as the result of severe off-season injuries nor the end of their contract implies an important role for other explanatory factors such as information shocks. I explore this possibility in the next section.

11. Table B1 replicates Table 3, but only using data for players with four or more years of experience playing in the NFL. Both the coefficients and the patterns of statistical significance are generally unchanged.

6.2. The Effect of a CTE Information Shock on Retirement

Table 4 shows the results of the analysis of the effect of the news of the post-mortem discovery of CTE among a former teammate on players' retirement decisions. Given the differential timing of treatment with such news, the data used in the analysis are constructed in "stacks" where each stack corresponds to one cohort of treated teammates and one cohort of control units who either never, or have not yet, had a similar incident occur while they were players. Thus, each model in the following tables includes *Stack-by-Player*, *Year*, *Position* and *Team* fixed effects for the estimation of a standard difference-in-difference model.

Table 4 — Difference-in-Difference (DiD) Estimates

Dependent Variable:		Retire	
Model:	(1)	(2)	(3)
Variables			
Teammate	0.013	0.022**	0.020*
	(0.010)	(0.010)	(0.010)
Fit statistics			
Observations	143,991	133,772	124,942
R^2	0.470	0.462	0.469
Control Group	100%	95%	90%

Clustered (Stack-by-Player) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*
Teammate = 1 if the player was teammates with someone who was diagnosed with CTE. The variable takes a value = 0 until the diagnosis is made. Each model includes *Stack-by-Player*, *Year*, *Position* and *Team* fixed effects.

Using a comparison group of every other player in the NFL within a seven-year window of the CTE discovery, Model 1 of Table 4 estimates that players who were teammates with the recently deceased person in question were 1.3% more likely to retire within three seasons of the event. However, this estimate is statistically indistinguishable from zero.

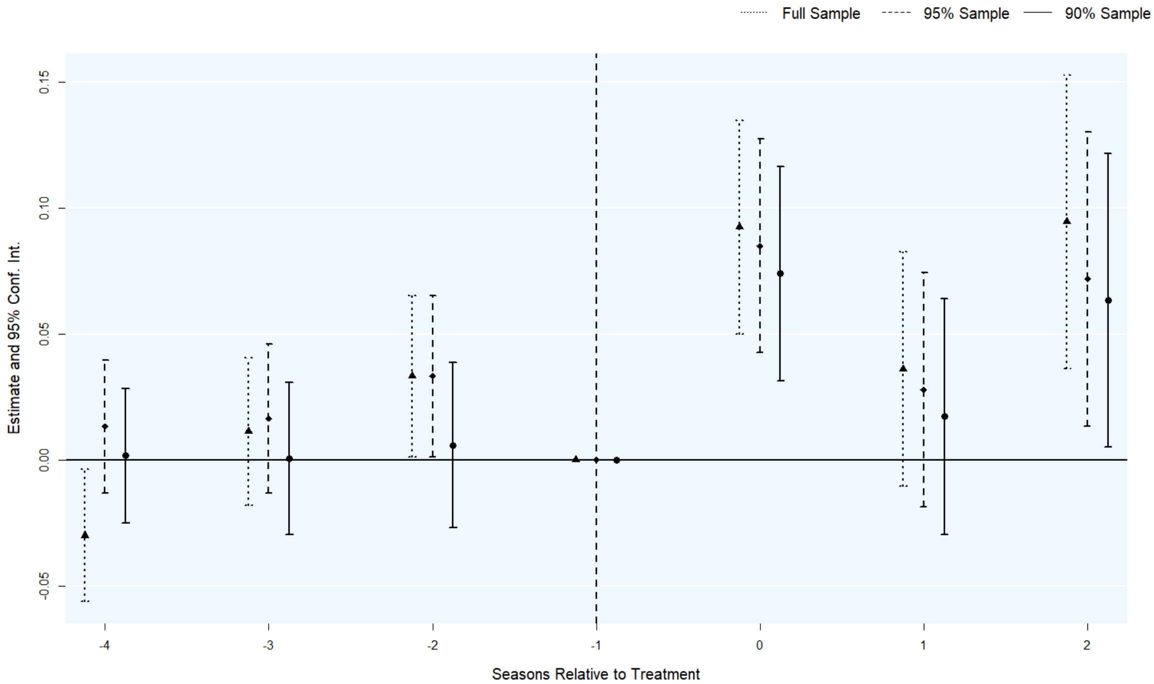
Given that a condition for receiving treatment is having been in the league long enough to have had a former teammate both die and receive a post-mortem CTE diagnosis, treated players likely differ significantly across observable characteristics – namely age. Thus, Models 2 and 3 of Table 4 drop individuals from the control group who are in the bottom fifth and tenth percentiles in terms of similarities to the treated group along observable characteristics.[12] These models find statistical significance in the average treatment effect – estimating a 2.0-2.2% increase in the likelihood of retiring within three seasons of the incident.

Figure 2 shows an event study analysis of the difference-in-differences model estimated with data that differ by the percentage of the control group included within the sample. The dynamic analysis of the full dataset fails to satisfy the parallel trends assumption for difference-in-difference models with two time periods that diverge significantly in the pre-treatment periods. The 95% dataset performs slightly better but still exhibits a significant pre-trend divergence one period before the timing of treatment. Though the size of the standard errors increases and the average treatment effect decreases, the pre-trend estimates from the model which uses the dataset which drops control group observations below the tenth percentile best satisfy the parallel trends assumption. Thus, the models using this dataset are my preferred specification for the estimation of treatment effects in each of the tables throughout the remainder of paper.

The point estimates showing differences in the probability of retiring between former teammates and other players are most pronounced in the season of the announcement. This is consistent with the idea of a news shock – particularly with the effect returning to being statistically indistinguishable from zero in the time period immediately afterwards. Interestingly, Figure 2 shows that the effect rebounds two

12. For more information, please see the Identification section of the paper.

Figure 2 — Event Study Analysis



seasons following the announcement. However, this is likely statistical noise as the the point estimate continues trending towards zero with increasing reductions in the sample size of the control group whereas the effect in the season of the announcement remains consistently in the 5-8% range.

6.3. Mechanisms & Robustness

Table 5 continues the analysis begun in Table 4, but controls for variables which were shown to be important correlates of retirement in Table 3.[13] Upon controlling for these variables, the magnitude of the treatment effect variable *Teammates* increases relative to those from Table 4.

Given the evidence that I provide that former teammates are more responsive to CTE news shocks, I test whether there is a differential impact on retirement de-

13. *Age* is excluded due to the presence of player and year fixed effects.

Table 5 — DiD Estimates w/ Controls

Dependent Variable:	Retire		
Model:	(1)	(2)	(3)
Variables			
Teammate	0.016*	0.025***	0.023**
	(0.009)	(0.010)	(0.010)
Games	-0.014***	-0.014***	-0.014***
	(0.0007)	(0.0007)	(0.0007)
Approx. Value	-0.014***	-0.014***	-0.014***
	(0.0009)	(0.0009)	(0.0009)
p(Super Bowl)	-0.274***	-0.273***	-0.291***
	(0.083)	(0.083)	(0.083)
Fit statistics			
Observations	143,991	133,772	124,942
R^2	0.502	0.493	0.499
Control Group	100%	95%	90%

Clustered (Stack-by-Player) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*
Teammate = 1 if the player was teammates with someone who was diagnosed with CTE. The variable takes a value = 0 until the diagnosis is made. Each model includes *Stack-by-Player*, *Year*, *Position* and *Team* fixed effects.

cisions that varies by the length of time spent as a teammate. This is done to test a "familiarity" hypothesis whereby the strength of a player's bond with their former teammate increases with time spent together (as measured by the number of seasons played on one team together). To do this, I create a *n-Year* binary variable to place former teammates into bins. Bins of 1 year, 2-4 year, and 5+ years are selected.[14]

Table 6 — DiD Estimates: Teammate Length

Dependent Variable:	Retire		
Model:	(1)	(2)	(3)
Variables			
Teammate × 1 Year	0.092***	0.102***	0.100***
	(0.018)	(0.018)	(0.018)
Teammate × 2-4 Years	-0.026*	-0.017	-0.019
	(0.014)	(0.014)	(0.014)
Fit statistics			
Observations	143,991	133,772	124,942
R^2	0.470	0.462	0.469
Control Group	100%	95%	90%

Clustered (Stack-by-Player) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*
Teammate = 1 if the player was teammates with someone who was diagnosed with CTE. The variable takes a value = 0 until the diagnosis is made. The *Year(s)* variables represent the number of years that the player was on the same team as the player who was later diagnosed with CTE. Each model includes *Stack-by-Player*, *Year*, *Position* and *Team* fixed effects.

Table 6 tests for the differential impact of increased former teammate familiarity. The *5+ Year* bin is omitted as a reference group. I find no effect of a differential

14. There are approximately half as many teammates in the 2-4 year bin as in the 1-year bin, and half as many in the 2-4 bin as in the 5+ year bin.

impact varying by teammate length as evidenced by the coefficients of the 2–4-year former teammates being significantly smaller than those in the 1-year bin. If anything, relative to those who were teammates for five or more seasons, former one-season teammates drive the result with two-to-four-year teammates having a slightly lower chance of retiring afterward.

Table 7 attempts to clarify the puzzling findings of Table 6. With linked data to players' college football statistics, I create a binary variable indicating whether the player was likely to have completed enough college credits to have graduated given the number of seasons they played. This is likely an important consideration as players without a degree are likely to face much less forgiving labor market prospects upon retiring than those with one given the relatively stagnant wages of those with only a high-school education (Taber, 2001; Carneiro and Lee, 2011; Castex and Kogan Dechter, 2014).

I create a variable *College Grad* which takes a value equal to one if the player played four or more seasons in college and a zero otherwise. Unfortunately, I am unable to link college attendance for every player in the sample – particularly for those who entered the league before 2000. However, given that CTE treatment timing does not begin until 2002, the issue of missingness is better for former teammates (31% missing observations = 320) than the entire sample (40% missing observations = 28,752).

Table 7 interacts the *Teammate* treatment variable with an indicator proxying for college degree completion. The presence of null results across each model provides evidence that the retirement result is being driven by teammates with fewer years of college experience.

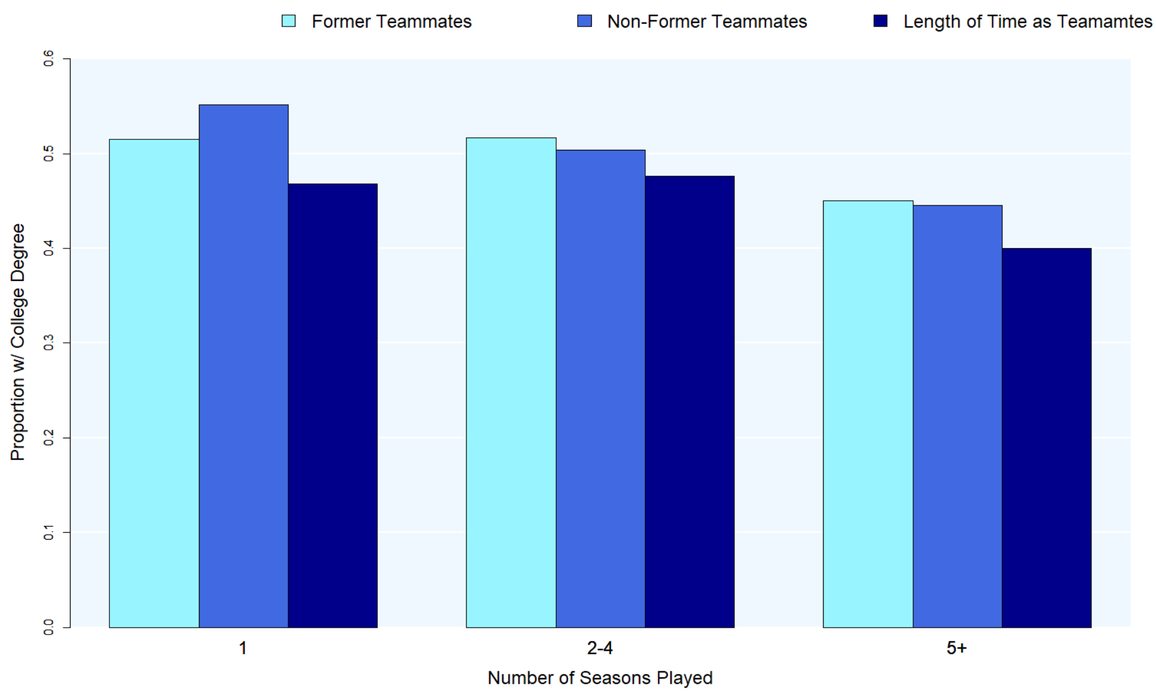
Figure 3 explores the findings of Table 7 by plotting the proportion of athletes with college degrees per number of seasons played in the league. The entire sample of players for which there is college data are split into two groups: players who were both former teammates with someone who died with CTE and were still in

Table 7 — DiD Estimates: College Graduates

Dependent Variable:		Retire	
Model:	(1)	(2)	(3)
Variables			
Teammate $\times$ College Grad	-0.011	0.0002	-0.0008
	(0.015)	(0.015)	(0.015)
Fit statistics			
Observations	143,671	133,452	124,622
R^2	0.470	0.462	0.469
Control Group	100%	95%	90%

Clustered (Stack-by-Player) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*
Teammate = 1 if the player was teammates with someone who was diagnosed with CTE. The variable takes a value = 0 until the diagnosis is made. *College Grad* = 1 if the player completed 4 or more years of college. The variable takes a value = 0 otherwise. Each model includes *Stack-by-Player*, *Year*, *Position* and *Team* fixed effects.

Figure 3 — Degree Completion and Career Longevity

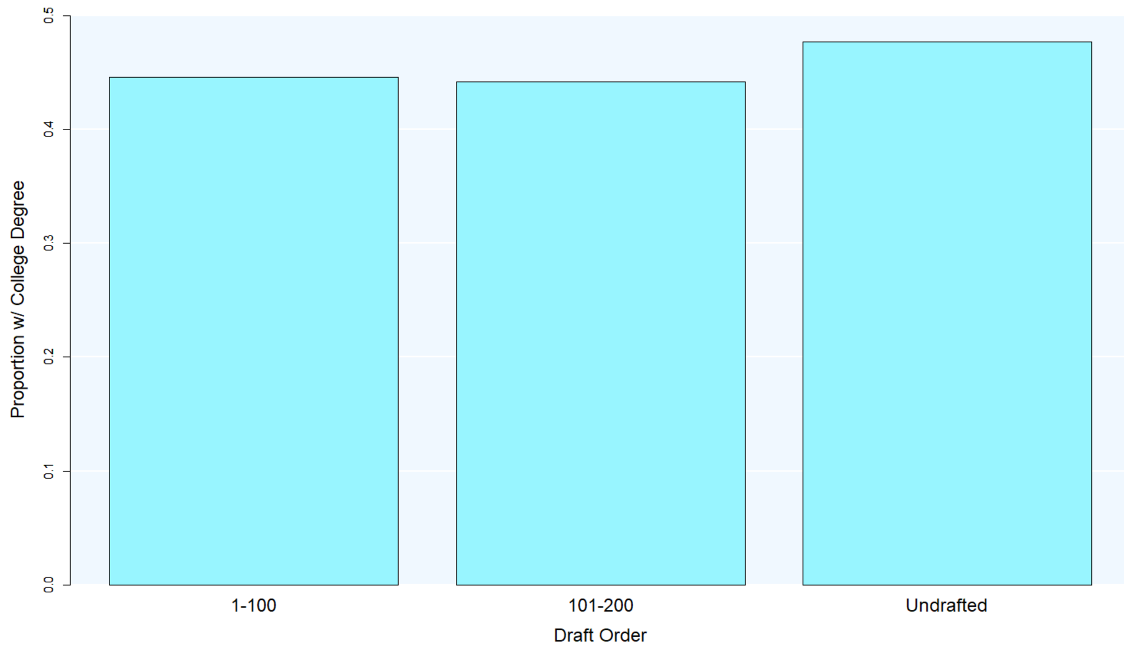


the league at least 4 years before their death and everyone else. Thus, the sample average can be roughly approximated by taking the mid-point of these two averages.

When looking at the first two columns in each group, there is a clear negative relationship between degree completion and the number of seasons played in the NFL. This is consistent with a rational model of human capital accumulation wherein players with the largest expected likelihood of having longer professional football careers trade off postponing their degree completion to begin their career as a professional football player earlier. I test this hypothesis in Figure 4. The figure shows that undrafted players (where a roster spot is not guaranteed, contracts are significantly shorter and pay is much less) are 3.5 percentage points more likely to have stayed for at least their fourth year of college than were players selected in the draft.

Of interest to the findings in Tables 6 and 7, the right-most columns in Figure 3 show the share of former teammates with at least four years of college attendance

Figure 4 — Degree Completion & Draft Status



by the length of time spent as a teammate. The figure shows lower levels of degree completion than would otherwise be expected relative to peers with similar career longevities. While there is no reason to suspect there is a causal relationship between the probability of graduating from college and being in the league at the time of the death of a former teammate who was later diagnosed with CTE, it does provide suggestive evidence for the hypothesis that lower levels of education likely dominate the effect of increasing familiarity with a former teammate.

7 — Conclusion

This paper focuses on understanding the causes of voluntary retirement decisions among professional American football players. I do so by first analyzing the general determinants of retirement. Upon discovering that changes in observable characteristics such as the expected probability of winning the Super Bowl, player productivity, and age can only explain one-quarter of retirement decisions, I turn my attention to the impact of negative informational health shocks. Using a difference-in-differences method that relies upon the quasi-random timing of chronic traumatic encephalopathy (CTE)-related deaths among both former and current players to show that this news shock increased voluntary retirements by 2-7 percentage points (10-30%).

Upon examining the mechanisms of this result, interesting differences emerge between players with and without a college degree. Traditional theory suggests that, all else equal, players with degrees will have higher opportunity costs (as determined by the higher wages they could expect to command in a labor market outside of professional football), and thus, should be more likely to retire.

However, if players use the income of their teammates as a benchmark for lifetime earnings goals rather than maximize expected lifetime earnings, it will take athletes without a degree more years of work upon retiring from professional athletics to maintain their earnings-based social status. Thus, as workers update their priors that continuing their career as a professional athlete could harm the number of productive years of work they expect to have once their first career is over, players without a degree would become more likely to retire. Thus, the relationship between education and negative informational health shocks is likely determined by the utility function of the workers in question.

References

Allison, Rachel, Adriene Davis, and Raymond Barranco. 2018. "A comparison of hometown socioeconomics and demographics for black and white elite football players in the US." *International Review for the Sociology of Sport* 53 (5): 615–629.

Baker, Andrew C, David F Larcker, and Charles CY Wang. 2022. "How much should we trust staggered difference-in-differences estimates?" *Journal of Financial Economics* 144 (2): 370–395.

Blundell, Richard, Jack Britton, Monica Costa Dias, and Eric French. 2023. "The impact of health on labor supply near retirement." *Journal of Human Resources* 58 (1): 282–334.

Brown, Kristine M, and Ron A Laschever. 2012. "When they're sixty-four: Peer effects and the timing of retirement." *American Economic Journal: Applied Economics* 4 (3): 90–115.

Carneiro, Pedro, and Sokbae Lee. 2011. "Trends in quality-adjusted skill premia in the United States, 1960–2000." *American Economic Review* 101 (6): 2309–2349.

Castex, Gonzalo, and Evgenia Kogan Dechter. 2014. "The changing roles of education and ability in wage determination." *Journal of Labor Economics* 32 (4): 685–710.

Daneshvar, Daniel H, Jesse Mez, Michael L Alosco et al. 2021. "Incidence of and mortality from Amyotrophic Lateral Sclerosis in National Football League athletes." *JAMA Network Open* 4 (12): e2138801–e2138801.

Disney, Richard, Carl Emmerson, and Matthew Wakefield. 2006. "Ill health and retirement in Britain: A panel data-based analysis." *Journal of Health Economics* 25 (4): 621–649.

Erpič, S Cecić, Paul Wylleman, and Maja Zupančič. 2004. "The effect of athletic and non-athletic factors on the sports career termination process." *Psychology of Sport and Exercise* 5 (1): 45–59.

Fernandez, Anne, Yannick Stephan, and Evelyne Fouquereau. 2006. "Assessing reasons for sports career termination: Development of the Athletes' Retirement Decision Inventory (ARDI)." *Psychology of Sport and Exercise* 7 (4): 407–421.

French, Eric, and John Bailey Jones. 2017. "Health, health insurance, and retirement: A survey." *Annual Review of Economics* 9 383–409.

Goodman-Bacon, Andrew. 2021. "Difference-in-differences with variation in treatment timing." *Journal of Econometrics* 225 (2): 254–277.

Gustman, Alan L, and Thomas L Steinmeier. 1983. "A structural retirement model."Technical report, National Bureau of Economic Research.

Hagan, Ronald, Andrew M Jones, and Nigel Rice. 2008. "Health shocks and the hazard rate of early retirement in the ECHP." *Swiss Journal of Economics and Statistics* 144 323–335.

Jones, Andrew M, Nigel Rice, and Jennifer Roberts. 2010. "Sick of work or too sick to work? Evidence on self-reported health shocks and early retirement from the BHPS." *Economic Modelling* 27 (4): 866–880.

Koning, Ruud, Victor Matheson, Anil Nathan, and James Pantano. 2014. "The long-term game: An analysis of the life expectancy of National Football League players." *International Journal of Financial Studies* 2 (1): 168–178.

LeClair, Jessica, Jennifer Weuve, Matthew P Fox, Jesse Mez, Michael L Alosco, Chris Nowinski, Ann McKee, and Yorghos Tripodis. 2022. "Relationship Between Level of American Football Playing and Diagnosis of Chronic Traumatic Encephalopathy in a Selection Bias Analysis." *American Journal of Epidemiology* 191 (8): 1429–1443.

Lehman, Everett J, Misty J Hein, and Christine M Gersic. 2016. "Suicide mortality among retired National Football League players who played 5 or more seasons." *American Journal of Sports Medicine* 44 (10): 2486–2491.

Lusardi, Annamaria, and Olivia S Mitchelli. 2007. "Financial literacy and retirement preparedness: Evidence and implications for financial education: The problems are serious, and remedies are not simple." *Business Economics* 42 35–44.

Nguyen, Vy T, Ross D Zafonte, Jarvis T Chen et al. 2019. "Mortality among professional American-style football players and professional American baseball players." *JAMA Network Open* 2 (5): e194223–e194223.

Owora, Arthur H, Brittany L Kmush, Bhavneet Walia, and Shane Sanders. 2018. "A systematic review of etiological risk factors associated with early mortality among National Football League players." *Orthopaedic Journal of Sports Medicine* 6 (12): 2325967118813312.

Park, Sunghee, David Lavallee, and David Tod. 2013. "Athletes' career transition out of sport: A systematic review." *International Review of Sport and Exercise Psychology* 6 (1): 22–53.

Rosen, Sherwin, and Allen Sanderson. 2001. "Labour markets in professional sports." *The Economic Journal* 111 (469): 47–68.

Taber, Christopher R. 2001. "The rising college premium in the eighties: Return to college or return to unobserved ability?" *The Review of Economic Studies* 68 (3): 665–691.

Venkataramani, Atheendar S, Maheer Gandhavadi, and Anupam B Jena. 2018. "Association between playing American football in the National Football League and long-term mortality." *JAMA* 319 (8): 800–806.

White, Kristopher, Kathryn Wilson, Theresa A Walton-Fisette, Brian H Yim, and Michele K Donnelly. 2021. "Race and socioeconomic composition of the high schools of National Football League Players." *Sociology of Sport Journal* 38 (4): 376–385.

A — Additional Figures

Figure A1 — Mean Salary by Position

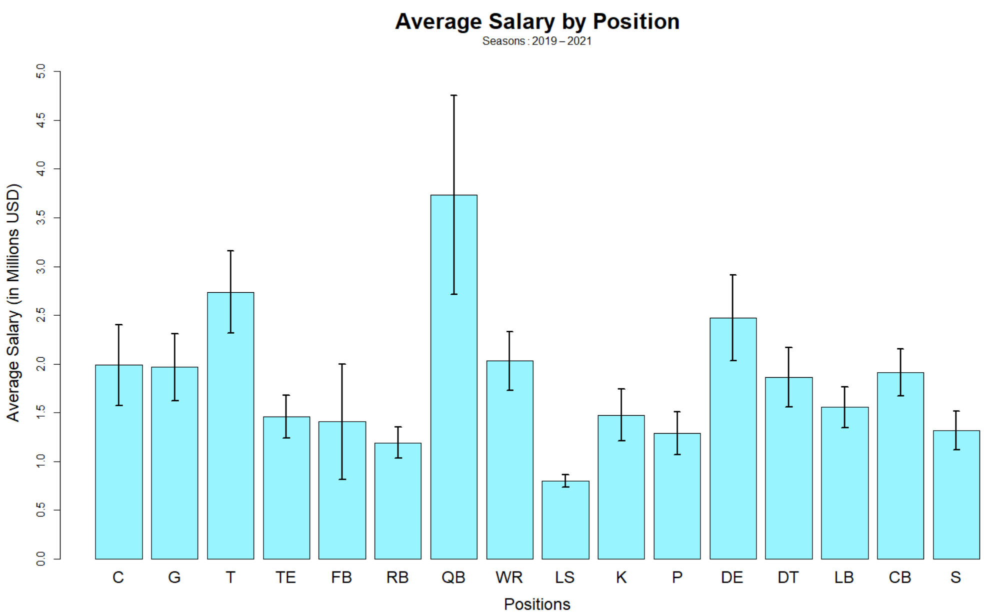


Figure A2 — Number of Seasons Played in NFL

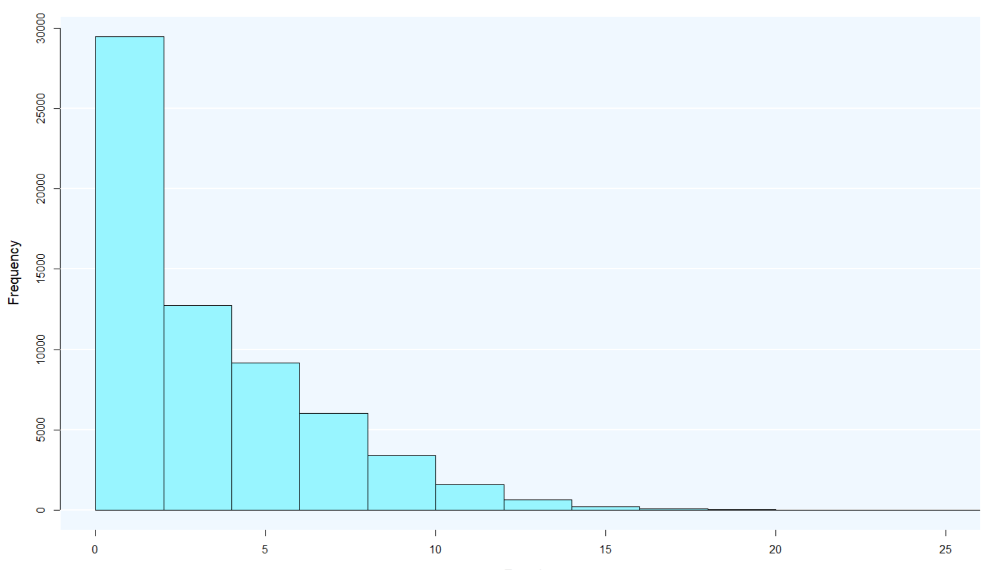


Figure A3 — Mean Productivity of by Award Type

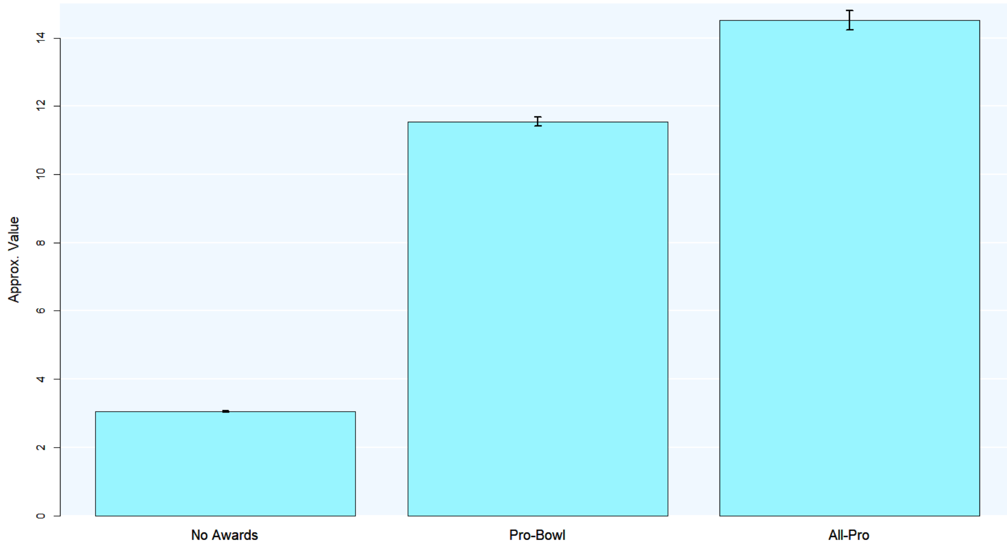


Figure A4 — Time-Series of Bookkeeper Over-Rounding

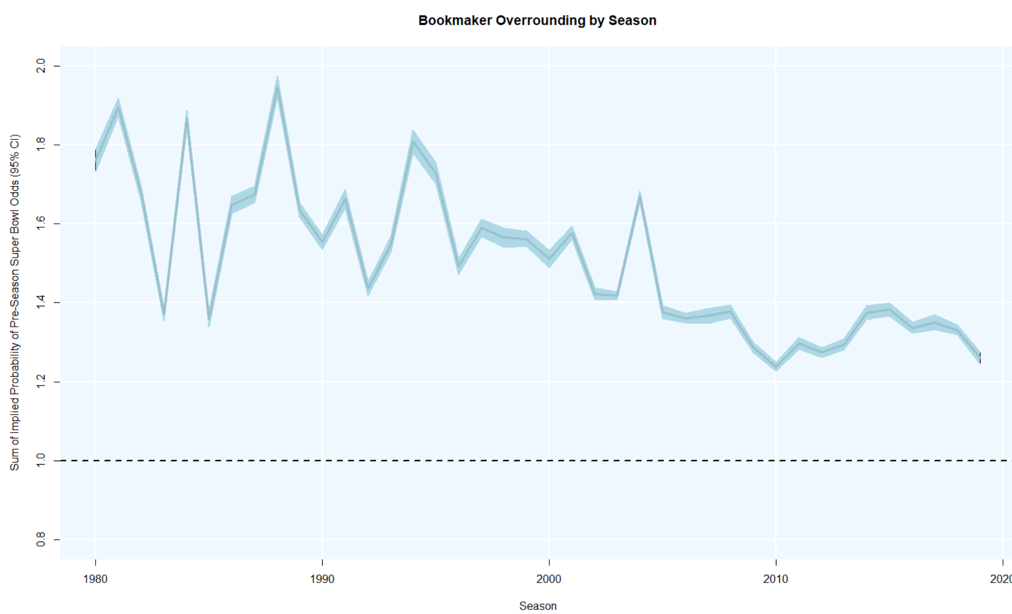
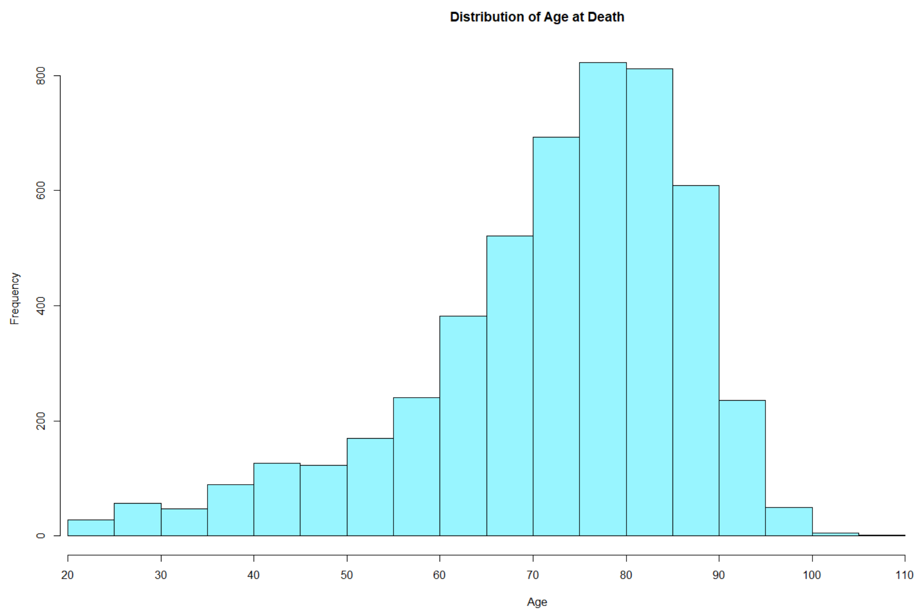


Figure A5 — Distribution of Age at Time of Death



B — Additional Tables

Table B1 — Determinants of Retirement ($>=$ 4 Years Experience)

Dependent Variable:	Retire			
Model:	(1)	(2)	(3)	(4)
	OLS	OLS	Probit	Logit
Variables				
Age	0.036***		0.031***	0.030***
	(0.000)		(0.000)	(0.000)
Games	-0.015***	-0.014***	-0.008***	-0.007***
	(0.000)	(0.000)	(0.000)	(0.000)
Games Started	-0.004***	-0.007***	-0.000	-0.000
	(0.000)	(0.001)	(0.000)	(0.000)
Approx. Value	-0.009***	-0.007***	-0.022***	-0.024***
	(0.001)	(0.001)	(0.001)	(0.001)
p(Super Bowl)	-0.234***	-0.277***	-0.195***	-0.194***
	(0.041)	(0.049)	(0.041)	(0.041)
Pro Bowl	-0.014**	-0.006	-0.023**	-0.037***
	(0.006)	(0.007)	(0.009)	(0.011)
All Pro	0.046***	0.055***	-0.013	-0.049
	(0.011)	(0.012)	(0.028)	(0.038)
Fixed-effects				
Player		✓		
Year	✓	✓	✓	✓
Position	✓	✓	✓	✓
Team	✓	✓	✓	✓
Fit statistics				
Pseudo R^2	0.428	0.651	0.321	0.320

Heteroskedasticity-robust standard-errors in parentheses. n = 56,549
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*
Each of the coefficients and standard errors in models 3 and 4 are transformed the represent the average marginal effects.