2017

Height, Weight, and Durability in Major League Baseball

Joshua Yeager
Claremont McKenna College

Claremont McKenna College

HEIGHT, WEIGHT AND DURABILITY IN MAJOR LEAGUE BASEBALL

SUBMITTED TO

PROFESSOR HEATHER ANTECOL

BY

JOSHUA YEAGER

FOR

SENIOR THESIS

SPRING 2017

APRIL 24th, 2017

Table of Contents

Abstract

Using data from the 2000-2016 Major League Baseball seasons, this paper looks at the determinants of durability amongst baseball athletes, durability is measured in games played for batters and innings pitched for pitchers, with a particular focus on height and weight. This paper finds evidence that lighter, shorter batters play significantly more games than taller, heavier batters. Additionally, amongst pitchers, there is only circumstantial evidence that height and weight are important determining factors of player durability. Finally, I find that starting pitchers increase the likelihood of becoming injured in the regular season by 10.3% and 21.8% if their hits per nine innings and home runs per nine inning totals, respectively increase by one.

Acknowledgements

I would like to thank all of the people who have helped me achieve this academic milestone. Gary Cohen helped me pull the data from baseballcube.com. Professor Antecol offered unparalleled guidance throughout this semester and always went above the call of duty in order to make my paper the best it could be. My parents and extended family, who have continually taught me the importance of education. My uncles, grandpa and dad, for passing down the importance of baseball. Additionally, I would like to thank the Claremont McKenna community and Professor Brennan for helping me refine my writing throughout my time here.

I. Introduction

In the United States, professional baseball is the second most popular sport in terms of revenue, bringing in 9.5 billion dollars each year (Maury Brown, 2016). Since the 1980s, baseball teams have become more competitive and balanced through draft pick compensations, revenue share agreements and laws like the Luxury tax which penalizes teams for spending over a certain threshold each year (Luis Delgado, 2015). Thus, baseball management has been forced to find new ways in order to stay ahead of the curve and beat their competition (Robb Karskamp et. al., 2016). In doing so, new practices regarding sabermetrics, or advanced baseball metrics, have been increasingly used (Baumer and Zimbalist, 2003). Advanced baseball metrics use mathematical formulas to determine the production capabilities of a player. Despite this, in some instances, many baseball clubs remain surprisingly old-fashioned in an era dominated by mathematical formulas and technological progress (Robb Karskamp, 2016).

Christopher Crawford (2017) and others argue that one of the largest misguided principles baseball organizations and scouts continue to use to determine the production capabilities of a player is their height and weight. Often, the height of a pitcher is used to determine his strength and how vulnerable he is to an increased workload, or innings pitched. Additionally, taller pitchers are believed to have increased control of the strike zone and may even be able to throw faster. Moreover, many position players in the league are expected to play more games and be more productive if they are taller, or heavier. Therefore, amongst all players, height and weight are two metrics used to perhaps unjustly determine the durability of athletes.

In previous literature, many baseball enthusiasts and researchers have tried to determine the effect of a baseball player’s height and weight on their productivity. Glen Greenberg (2010) finds that tall pitchers are equally as effective as short pitchers in terms of their strikeouts and earned run averages. Additionally, Hamburg (1999) finds no statistical significance towards height on the batting average of a player, while McLennan (2010) finds that players who are taller and heaver have increased OPS (variable definitions are defined in Table 8). Finally, Gassko (2006) finds that shorter players have higher OBP and fewer strikeouts, while heavier, taller players have higher home run totals. Despite all of the literature and research on baseball, to the best of my knowledge, no study has examined the effect of height and weight on a player’s durability.

The purpose of this paper is to determine if pitchers and batters in the Major League Baseball are more durable based on their weight and height as well as other productivity metrics. Traditionally, baseball players have predisposed stereotypes by position (Christopher Crawford, 2017). Pitchers are expected to be taller so their arms do not wear out over time, while batters are expected to be heavier and taller in order to be productive and last throughout the regular season. These stereotypes could potentially disadvantage shorter, lighter players who do not fit them and may also hurt the productivity of teams that overemphasize these metrics, since these stereotypes predominately use starting players who are less durable. It is important to analyze if coaches and scouts are correct in using specific types of players because they believe they are less likely to break down.

Specifically, I use Major League Baseball data for every player from seasons between 2000-2016. I investigate pitcher’s and batter’s height separately, as well as the interaction between these variables in order to determine the joint effects height and weight have on a player's durability. I argue that this study is important as it helps inform baseball clubs about how to maximize their potential and make the best decisions they can toward determining which players start in games. If a baseball player's height and weight do not influence their durability throughout the regular season, baseball clubs should not rely on these metrics to determine which athletes play in games. Alternatively, if height and weight are important metrics with respect to a player's durability, scouts should continue to use these measures to create the most productive team possible.

This paper finds support that short and light batters play more games than other players. However, I find only circumstantial evidence that weight and height influence durability of pitchers (starters and relievers). These results suggest that, amongst batters, height and weight discrimination against short and lighter batters should not exist since these players play more games relative to all other players. Moreover, it does not appear that height and weight should be used to determine which pitchers should play in games.

The remainder of the paper is organized as follows. Section II discusses the existing literature. The data is presented in Section III. Section IV presents the estimation strategy and results. The conclusions are presented in the final section.

II. Literature Review

There is a large amount of academic literature that uses baseball statistics to assess different questions surrounding players and management, as baseball data is concrete and readily available. For example, BaseballReference.com and a multitude of other sites have baseball statistics dating back to the 18^{th} century. A number of studies have used height and weight to determine productivity and strength of players in the

MLB. Glen Greenberg (2008) looks at the effectiveness of pitchers and their height. He uses pitcher data from 1990-2007 and finds that height does not determine pitchers' productive capabilities. In his research, he uses traditional baseball metrics like earned run average and strikeouts. However, he does not study if pitchers are more durable based on their height and weight profiles. Since the inception of the MLB, players have become taller and heavier. Specifically, Bill Petti (2014) states that in the 1870's the average player was 68.9 inches and 168.4 pounds, while in 2016 an average player was 73.7 inches and 190.6 pounds. This suggests that heavier and taller players make better athletes in baseball. Moreover, this alludes to the idea that taller and heavier athletes are more productive, but evidence on this assertion is mixed.

Hamburg and Mines (1999) find that height and weight are not statistically significant in determining the batting average of a hitter. However, weight is statistically significant in determining the number of home runs hit. On the other hand, McLennan (2010) finds that height and weight show little correlation to any traditional baseball metrics in the MLB including earned runs, earned runs+, home runs allowed and strike outs. Moreover, for hitters he finds that increased height and weight have a statistical correlation to slugging percentage, OPS, BB%, HR% and K%. Hamburg's findings suggest that height and weight are more important and relevant to a batter than a pitcher. Intuitively, this could be true since batters play every day, but pitchers throw once every few days. The higher number of games could prove more difficult and strenuous for heavier and taller batters than for heavier and taller pitchers.

Gassko (2006) also discusses the correlation between height and the production of hitters using data spanning over seventy years. He finds unsurprisingly, taller and heavier hitters hit more doubles and homeruns than shorter and thinner batters, but also strike out more. Additionally, he finds that shorter players have more hits and a higher OBP, since their strike zone is smaller and more difficult for pitchers to locate. For heavier players, Gassko (2006) expects that an extra 9.5 pounds will lead to an expected increase by one home run. This may seem insignificant, but if a player weighs fifty pounds more than another player and stays in the league for fifteen years, he will garner 75 more home runs solely based on his increased size. Moreover, Gassko finds that an increase in player's height has little significance in strike out rates or walks. Additionally, he finds that larger players hit .7 more doubles and .25 more triples for each additional ten pounds. Concluding, Gassko gives an example of a 6'3", 216-pound player and a 5'9", 160-pound player where the only difference between the two players is their weight and height. He finds that the taller player is expected to hit five more doubles and six more home runs, while only having two fewer triples. While his results are statistically significant, Gassko (2006) finds that height and weight make up a small fraction of the overall production of a player.

Pitchers performance based on height and weight proves more difficult to gage than it is for batters. Gassko (2007) finds that short pitchers under 5'11" and heavy pitchers who are at least 21 pounds overweight seem to stay in the league longer than tall or light pitchers. Gassko believes this could be true because scouts have selection biases towards what a stereotypical pitcher looks like, so athletes that do not fit the norm have to be better than average. He also finds that tall pitchers above 6'5" have .48 more strikeouts per game than other pitchers. Logically, this makes sense as a taller pitcher throws the ball at a higher angle, which could lead to an increase in the speed of their pitches. Additionally, Gassko finds that light pitchers are 50% more likely than a pitcher of average weight to pitch in the bullpen, alluding to the fact that baseball organizations probably put lighter pitchers in a reliever role because they believe they are more likely to break down.

Alan Nathan (2009) discusses the correlation between swing speed in the MLB and an athlete's mass. He finds that if an athlete has about 50% of their weight as muscle, every 10% increase in muscle mass translates to roughly a 3.9% increase in bat speed. Therefore, Nathan argues that an increase in mass can lead to an increase in production for baseball athletes. However, Katherine Foley (2016) believes differently. Specifically, she argues that the Laws of Physics state it is easier to move something smaller than something bigger. Thus, larger athletes have more mass to move. These athletes may receive less momentum or velocity since Foley (2016) argues kinetic energy= (1/2)(mass)(volume). Taken together, it seems that an increase in mass has inconclusive results on bat speed.

The aforementioned existing literature suggests that height and weight impact batters more heavily than pitchers. With an increase in weight, hitters are expected to hit with significantly more power, while taller pitchers have marginal benefits with respect to their strikeout statistics. Despite strong evidence that height and weight impact a player's productivity, to the best of my knowledge, there is no literature that discusses the impact that height and weight have on the durability of a player. This is incredibly important, since scouts continue to use height and weight as metrics to determine the longevity and durability of players. Thus, shorter, lighter pitchers and taller, heavier batters are thought to be more susceptible to breaking down in a strenuous 162 game season. These biases remain unfair to athletes, since many scouts and front offices appear to neglect, or choose not to sign athletes that fit these profiles (Christopher Crawford, 2017). Therefore, it is imperative to determine if the stereotypes baseball front offices employ to sign tall, short, heavy or, light athletes is misguided, or if these players truly are less durable than their average sized counterparts.

III. Data

The data I use in my empirical analysis comes from baseballcube.com and includes information on every baseball player from 2000-2016. Batters and pitchers are separated in the data sets. This data comprises 18,297 entries for batters and 12,107 entries for pitchers. Moreover, the pitcher data is broken amongst two groups for relievers and starters. In the data there are 7,592 relievers and 4,515 starting pitchers. I confirm the sample of batters in my data by using the mean number of batters across each team in the league by season.[1] I also employ this method to ensure my sample of pitchers is reasonable.[2] While I find I have a balanced sample of batters, I have a larger sample of pitchers. One possible reason my data set includes more pitchers than the mean number of pitchers is because baseball clubs often put in younger, less experienced athletes to see if they are ready for the workload in the major league. Thus, this inflates the total number

---

[1] The MLB has a 40-man roster with 30 teams, making roughly 1,200 individual data points per season. Since I used 16 years of data this gives a rough estimate of 19,200 player data points.
[2] Traditionally, the MLB has a 12-man roster for their pitchers. With 30 teams and 16 years of data this would make roughly 5,760 player data points.

of pitchers in my data. In my data specifically, if information for athletes was missing from the baseballcube.com, I manually inputted these variable values from baseballreference.com.[3] I also confirmed that the data from baseballcube.com is consistent with the data from baseballreference.com to ensure the data provided is accurate. This data is ideal for my paper, as it includes height and weight, two difficult metrics to get from other baseball sites, as well as detailed information on durability and player's performance characteristics.

In order to measure durability of batters and pitchers, I look at the amount of games played for batters and innings pitched for pitchers. I look at the amount of games played for batters, as this metric allows me to determine how reliable and consistent they are. It is a better metric than at bats, since the total amount of at bats a player has is partially driven by the offensive production of the team they play on. Additionally, I use innings pitched for pitchers to determine their durability, since innings pitched allows me to determine how consistently they are playing. Looking at these metrics as durability metrics, I then use individual height, measured in inches and weight measured in pounds to assess if these are driving determinants of a player's durability. In addition to these two measures, I create a set of indicators for height and weight, where height is measured in inches and weight is measured in pounds. Specifically, short equals 1 if a pitcher is below 6'0", or if a batter is below 5'9", and zero otherwise. Similarly, tall equals 1 if a batter or pitcher is taller than 6'6", and zero otherwise. For weight, I create two indicator variables

[3] baseballreference.com is regarded as the most reputable database. It has more cumulative and specific results than the official MLB site.

as follows. Light equals 1 if the player weights less than 160 pounds, and zero otherwise. Similarly, heavy equals 1 if the player weighs more than 230 pounds, and zero otherwise. Very few batters, or pitchers are in the extreme categories for these height and weight specifications.

I create an age variable based on the player's birthdate and the season the athletes played in. For example, if a player was born in 1976 and played in Season 2001, age is calculated as 25. Additionally, I create an indicator variable for relievers that equals 1 if the player has at least two saves and equals zero if the player has fewer than two saves.

I also have a number of productivity measures for players. In particular, for batters these include, (At Bats, Runs, Hits, Doubles Hit, Triples Hit, Home Runs Hit, Runs Batted In, Stolen Bases, Caught Stolen, Bases on Balls, Intentional Bases on Balls, Strikeouts, Sacrifice Hits, Sacrifice Fly Balls, Hit by Pitch, Double Played Grounded into, Batting Average, Slugging Average, On Base Percentage and On Base Percentage + Slugging Percentages). For pitchers, these include, (Wins, Losses, Games Pitched, Complete Games, Shutouts, Games Finished, Saves, Hits Allowed, Runs Allowed, Earned Runs Allowed, Bases on Balls Allowed, Intentional Bases on Balls Allowed, Strikeouts, Wild Pitches, Balks, Hits per Nine Innings, Home Runs per Nine Innings, Walks per Nine Innings, Strikeouts per Nine Innings and WHIP=(BB+H/IP)). See data appendix in Table 8 for detailed variable definitions.

III.1 Summary Statistics

Table 1 presents summary statistics for batters, while Table 2 presents summary statistics for starters and relievers, respectively. There are several noteworthy patterns.

First, batters are approximately 6’1” ½ and 194.5 pounds, relievers are 6’2” ¼ and 196.2 pounds, and starters are 6’2” ¾ and 197.5 pounds. The Center for Disease Control and Prevention (2016) states that the male average for height in the United States is approximately 5’9’’½ with an average weight of 199.5. Therefore, baseball players are significantly taller than the average male, but roughly the same weight. This suggests that baseball players are in better shape than the national average since they are taller, but have similar weights as an average American. I look into this idea further by taking the average BMI of baseball players in my data set, where BMI= weight (kgs)/ ht^2 (meters). In my data from 2000 to 2016, the average MLB Batter has a BMI of 25.1, while relievers and starters both have a BMI of 24.8. However, the average American male has an average BMI of 29.0, which is overweight, and almost on the verge of becoming obese.[4] Thus, it seems height, weight, and BMI are important metrics in the screening process for the MLB.

As a first attempt to see if height and weight influence a player’s durability, I present average games played for batters and number of innings pitched for pitchers by height and weight in Table 3. It can be seen that the effect of height appears to be larger for batters than pitchers. Specifically, short batters play 72 games on average while tall batters play 27 games. On the other hand, for short and tall starters, the difference is 121 innings pitched versus 110 innings pitched. Additionally, the difference between short and tall relievers is 31 innings pitched versus 29 innings pitched. These summary

[4] Calculated from U.S. National Institute of Health

statistics demonstrate that short players seem to be more durable than tall players, potentially disproving common stereotypes that assume shorter batters are less durable.

Moreover, weight seems to have a smaller impact than height, irrespective of player position. It can be seen that the effect of weight appears to be larger for batters than pitchers. Specifically, light batters play 57 games on average while heavy batters play 49 games. On the other hand, the difference between light and heavy starters is 131 innings versus 106 innings. Moreover, amongst relievers the difference between light and heavy is 27 innings pitched versus 31 innings pitched. This provides potential evidence that lighter baseball players are more durable than heavy baseball players, as relievers pitch roughly the same amount of innings, irrespective of a relievers weight.

When height and weight are interacted to form a joint variable, players in these groups have greater differences in durability than players in individual height and weight categories. Short and light batters play 77 games, but tall and heavy batters only play 36 games. On the other hand, short and light starting pitchers pitch 131 innings, but tall and heavy starting pitchers only pitch 100 innings. Amongst relievers, short and light pitchers pitch 29 innings, but tall and heavy pitchers pitch 25 innings. Therefore, when both height and weight are looked at, it seems that all baseball players who are short and light are more durable than their heavier, taller counterparts.

Another important note is the difference in potential quality between starters and relievers displayed in Table 2. Interestingly, relievers seem to be less effective pitchers since they have higher efficiency numbers in categories like Whip, Walks per Nine Innings, Home Runs per Nine Innings, Strikeouts per Nine Innings and Hits per Nine Innings. Reliever's values for these categories are, (1.62, 4.57, 1.43, 7.21, 11.16) but batter's values for these categories are, (1.43, 3.33, 1.18, 6.48, 9.59) respectively. Therefore, relievers only strike out more batters than starters, but universally starters can be perceived as more effective by allowing fewer players on base and fewer home runs.

The remainder of the paper formally analyses these patterns between batters, starters and relievers to determine if height and weight are important determinants of a player's durability.

IV. Estimation Strategy and Results

IV.1 Determinants of Durability for Batters

In order to determine if height and weight are key determinants of a player's durability, I estimate a model of the following form:

$$Y_{pt} = \alpha + \gamma H_{pt} + ØW_{pt} + \beta PC_{pt} + U_b + Y_t + E_{pt} \qquad (1)$$

where Y is games played, H is the player's height in inches, W is the player's weight in pounds, PC is a vector of player characteristics (Age, Runs, Hits, Doubles, Triples, Home Runs, Runs Batted In, Stolen Bases, Caught Stolen, Walks, Intentional Walks, Strike Outs, Sacrifice Hits, Sacrifice Fly's, Hit by Pitch, Times Grounded into Double Plays, Batting Average, On Base Percentage, On-base Plus Slugging), U is a vector of team fixed effects, Y is a vector for year fixed effects and E is the error term.

Column 1 of Table 4 presents the results of equation (1) with height and weight as separate variables. Perhaps surprisingly, taller players play fewer games. For example, a 6'6" player would play roughly two games less than a 6'0" player, potentially due to injury. This proves contradictory to common beliefs since scouts often think taller batters should play more games because they are believed to be more productive. Additionally, heavier players also play fewer games. For example, if a player is roughly seventy pounds heavier than another player, they will play two fewer games. Therefore, both weight and height are important metrics to determine a batter's durability.

However, the relationship between weight and height and player durability may not be linear. In an attempt to address this, I re-estimate equation (1) by replacing weight with an indicator variable for light and heavy (the omitted category is average height). A player is considered light if they weigh less than 160 pounds and heavy if they weigh more than 230 pounds (an average weight lies between these two values). Moreover, I replace height with indicator variables for short (less than 5'9") and tall (greater than 6'6") and the omitted category is average height (between 5'9" and 6'6"). I refer to this as specification 2, and the results are presented in Column 2 of Table 4. While light batters play 2.6 more games than their average counterparts, heavier batters play .9 fewer games relative to their average counterparts. Similarly, short batters play 1 more game than average height batters, although the effect for short is less precisely estimated. Moreover, tall batters play 2.2 more games than average height batters. This evidence continues to go against the stereotype that shorter and lighter players are more likely to break down. Moreover, it seems that shorter, lighter players are actually more durable.

This leads to my third specification which allows for an interaction effect between height and weight. Specifically, I add interactions to short and light players and tall and heavy players. I am unable to add an interaction for tall and light and short and heavy as there are no players in these categories. Column 3 of Table 4 shows that a batter who is short and light plays an additional 3.4 games relative to their average height and weight counterpart, although the effect is less precisely estimated. Additionally, batters who were both tall and heavy played 2.66 fewer games than normal sized players. Across the board, these results continue to express that shorter, lighter players are more durable than heavier, taller batters and their average sized peers.

Before turning to the analysis of pitchers, I point out some interesting patterns with respect to player characteristics, particularly homeruns. Given the results for these measures are similar across specifications I focus on the results for specification 1 in column 1 of Table 4. Perhaps surprisingly, batters who earn more home runs play fewer games. It seems as if this batter is being penalized for adding runs to the scoreboard. However, it is quite likely that players who are more likely to hit home runs are also more susceptible to injury and breaking down. Thus, these home run hitters are most likely forced to sit during long stretches of the season due to injury. This explains how offensive power could be viewed as a detriment to durability within the MLB. Since power hitters are most likely heavier and taller in order to generate power, this means many power hitters are less durable. The patterns for the remaining player characteristics are in line with previous literature.

IV.II Determinants of Durability amongst Starting Pitchers & Relievers

In order to determine if height and weight are key determinants of a pitcher's durability, I estimate a model of the following form:

$$Y_{ip} = \alpha + \gamma Hpt + \text{Ø}Wp_t + \beta PC_{pt} + U_b + Y_t + E_{pt} \quad (2)$$

where Y is number of innings pitched, PC is a vector of player characteristics (Wins, Losses, Games Played, Games Started, Complete Games, Games Finished, Shutouts, Saves, Hits Allowed, Homeruns Allowed, Runs Earned Runs Allowed, Walks Allowed, Wild Pitches Allowed and Strikeouts), U is a vector of team fixed effects, Y is a vector for year fixed effects and E is the error term.

Table 5 of Column 1 presents the results of equation (2) where I estimate the individual effects of height and weight on pitcher's durability. Perhaps surprisingly, both height and weight results were statistically insignificant and had small coefficients. Thus, these results have almost zero bearing on the durability of baseball starting pitchers. However, age was highly statistically significant with a coefficient of .1. Thus, if a starting pitcher's age increases by 1, he will play .1 more games throughout the regular season. These results are interesting since many players deal with injuries as they get older. However, these results imply that starting pitchers improve their ability during their time in the league, thus resulting in baseball organizations allowing them to pitch more games, which offsets any injuries or loss in durability they may experience later in their career.

However, the relationship between weight and height and starter durability may not be linear. In an attempt to address this, I re-estimate equation (2) by replacing weight with indicator variables for light and heavy (the omitted category is average weight). A player is considered light if they weigh less than 160 pounds and heavy if they weigh more than 230 pounds. Additionally, I replace height with indicator variables for short (less than 5'9") and tall (greater than 6'6"). I refer to this as specification 2, and the results are presented in column 2 of Table 5. While light starting pitchers play .05 more innings than their average size counterparts, heavier starting pitchers pitch .6 fewer innings than their average sized counterparts. Similarly, short starting pitchers pitch .5 fewer innings than their average sized counterparts, while taller starting pitchers pitch .1

more innings than their average sized counterparts. However, amongst short, tall, heavy and light batters, the effects were less precisely estimated. Therefore, in this second specification, there is inconclusive evidence to determine if height and weight are good determinants to the durability of a starting pitcher.

This leads to my third specification which allows for an interaction effect between height and weight. Specifically, I add interactions for short and light players and tall and heavy players. I am unable to add an interaction for tall, light and short, heavy as there are no players in these categories. Column 3 of Table 5 shows that a starting pitcher who is short and light plays 1.7 more games relative to their average weight and height counterparts, although the effect is less precisely estimated. Additionally, tall and heavy starting pitchers pitch 1 fewer innings relative to their average weight and height counterparts. These results offer mixed results that height and weight are good determinants of durability given the small sample size of these results.

The model I used for my relievers is displayed in Table 6. It is similar to the above model for starting pitchers, but without shutouts in the vector of player characteristics. Everything else is similar to the starting pitcher model above. Column 1 of Table 6 presents the results of equation (2). Perhaps unsurprisingly, height and weight have little effect on the durability of relievers, measured by innings pitched. Therefore, when all relievers are looked at, height and weight are poor determinants in measuring the durability of relievers.

However, the relationship between weight, height and reliever durability may not be linear. In an attempt to address this, I re-estimate equation (2) by replacing weight with indicator variables for light and heavy. Additionally, I replace height with indicator variables for short and tall. These variables are the same for starting pitchers. I refer to this as specification 2, and the results are presented in Colum 2 of Table 6. While light relievers pitch .3 more innings than their average counterparts, heavier relievers pitch .7 fewer innings than their average counterparts, although the effect for light relievers is less precisely estimated. Similarly, short players pitch .4 fewer innings than their average counterparts and tall relievers pitch only .05 more innings than their average counterparts, although the effect for tall relievers is less precisely estimated. Therefore, there seems to be inconclusive evidence that height and weight are good determinants for durability amongst these relievers.

This leads me to my third specification which allows for an interaction effect between height and weight. Specifically, I add interactions to short and light players and tall and heavy players. I am unable to add an interaction for tall and light and short and heavy, as there are not players in these categories. Column 3 of Table 6 shows that a reliever who is short and light pitches .03 fewer innings than their normal sized counterparts, and a reliever who is tall and heavy pitches the same amount as their normal sized counterparts. Throughout my specifications for relievers, there seems to be inconclusive evidence that size plays a strong contributing factor towards a reliever’s durability.

Before turning to the determinants of injury amongst starting pitchers, I point out some interesting patterns with respect to reliever’s production characteristics, specifically focusing on saves and home runs. It seems as if relievers are penalized for saving a game, thus winning the game for their team. With each ten saves, a reliever pitches one less inning, a very perplexing result. This may be because the dataset for relievers includes relievers who pitch in the middle of games who do not typically earn saves on their statistics sheet. As a result, this would artificially decrease the amount of innings pitched needed to garner a save. Moreover, home runs display counterintuitive results. Relievers who give up a home run pitch roughly .6 more innings in all three specifications. This seems as if relievers are incentivized to give up home runs because baseball organizations allow them to pitch more.

IV.III Determinants of Injury for Starting Pitchers

An alternative measure for durability is injury. This type of measure is only available for starting pitchers in my data. A starting pitcher is determined to be injured if they pitched less than 100 innings and not injured if they pitch more than 100 innings. I then estimate the effect of weight and height on the probability of becoming injured as follows:

$$Y_{pt} = \alpha + \gamma Hpt + ØWp_{t} + \beta PC_{pt} + U_{b} + Y_{t} + E_{pt} \qquad (3)$$

where Y is an indicator variable for injured player p in season t. All other variables are as previously defined in equation (2), with the exception that player characteristics controlled for are limited to (Wins, Losses, Games, Games Started, Complete Games, Games Finished, Shutouts, Saves, Runs Allowed, Earned Runs Allowed, Walks Allowed, Strike outs per nine innings and Home Runs per nine Innings). This model is a probit model and presents the marginal effects and marginal standard errors towards the likelihood that starting pitchers will become injured.

The results for equation (3) are presented in Table 7 Column 1. Interestingly, weight and height have little effect on the likelihood that a starting pitcher will become injured throughout the regular season. Each additional increase in pound and inch of a player has a .0001 and .0006 marginal effect towards the likelihood that a player will become injured. Therefore, individually, height and weight have little effect towards determining if a starting pitcher will become injured. This seems to contradict common stereotypes that shorter, lighter pitchers are less durable than heavier, taller pitchers.

However, the relationship between height and weight and the likelihood of a starting pitcher becoming injured may not be linear. In an attempt to address this, I re-estimate equation (3) by replacing weight with indicator variables for light and heavy. Additionally, I replace height with indicator variables for short and tall. The indicator variables thresholds are the same as Equation 1 and Equation 2. I refer to this as specification 2 and the results are presented in Column 2 of Table 7. Tall starting pitchers have a likelihood of becoming injured 8.6% of the time, but short pitchers actually become injured 1.9% less than their average sized counterparts. Additionally, light players have a likelihood of becoming injured 2.9% less than their average sized counterparts, while heavy starting pitchers have a likelihood of becoming injured 1.4% less than their average sized counterparts. Therefore, tall starting pitchers become injured significantly more often than all other pitchers.

This leads to my third specification, which allows for an interaction effect between height and weight. Specifically, I add interactions for short, light players and tall, heavy players. I am unable to add an interaction for tall, light players and short, heavy players as there are no players in these categories. Colum 3 of Table 7 shows that a starting pitcher who is short and light is 4% more likely to become injured than his normal sized counterparts. Additionally, a starting pitcher who is tall and heavy is 9.8%

more likely to become injured than his normal sized counterparts. These results express that shorter, lighter starters are more durable than heavier, taller starters, but not normal sized players.

Additionally, I point out some interesting dynamics that significantly increase the likelihood that all starting pitchers will become injured. Traditionally, starting pitchers are expected to become more susceptible to injury the longer each inning lasts in duration and overall pitches. Pitchers become tired if the duration of their inning lasts an additional fifteen or twenty pitches, since they have little time to recover until the next inning. Therefore, according to Table 7, all starting pitchers are more likely to break down when their hits allowed per nine innings and home runs allowed per nine innings increases. When these production values increase, a pitcher's start lasts longer and more of the opponent's batters successfully reach base, which extends the duration of the inning. Hits per nine innings coefficient is . 10, while home runs per nine innings coefficient is .21. Thus, if a pitchers Hits per Nine and Home Runs per Nine increase by 1 from the mean, the pitcher is 10.3% and 21.8% more likely to become injured. Thus, a starting pitcher can increase his durability by becoming more efficient on the mound and allowing fewer base runners to score. Additionally, these results imply that more dominant or efficient starting pitchers are less likely to become injured since their hits per nine innings and home runs per nine inning totals are usually low.

If a starting pitcher records a save, this also highly increases the likelihood that he will become injured. During the regular season, a starting pitcher will likely pitch once every five games in order to ensure that he receives proper rest for his starts. However, occasionally in a high stakes game, or a game where many of the relievers have pitched recently, a general manager will put in one of his starting pitchers on the starting pitcher's rest day. However, if a general manager alters with one of his starters' rest schedules, my results show they are much more likely to become injured. If a starting pitcher receives a save by altering with his rest schedule, this increases his likelihood of becoming injured by 2.7%. Therefore, baseball organizations should universally try to minimize starting pitchers' role as a reliever in order to ensure their durability throughout the regular season.

IV.VI Potential Biased Data

Throughout my data and regressions, there are many shortcomings that can be amended to achieve more accurate results in the future. For example, in my batter data, there was no variable for injured players, which makes it difficult to determine if players played fewer games because they hurt themselves, or because their production fell and they were replaced with an alternative athlete. Although this is a problem in the data set, to my knowledge, there are no baseball sites or sources that have a database of MLB injuries alongside their height, weight and production characteristics.

The pitching data also has some biases that could skew the results and misrepresent my findings. Innings pitched might not be the best metric for determining durability, but was used because it was the best alternative available. In the future, to get precise results, it could be more accurate to use cumulative pitches instead of innings pitched. By looking at cumulative pitches instead of innings, this would adjust the data for pitchers that throw fewer pitches in an inning. Some pitchers have quick innings, while others have long innings that add additional strain. Additionally, I did not control for seasonal trends in innings pitched. For example, in 1980 there were fifty-six pitchers with over two hundred innings pitched, while in 2016 there were only fifteen. Although I did not use 1980 results in my data, this explains that pitchers have seen a reduction in their workload in the 21st century. By not adjusting for seasonality in pitching data, biasness could occur.

Another potential concern is how I created my indicator variables for both starters and relievers. If a pitcher had more than two saves, he was a reliever, but if he did not, he was a starter. This could have biased results if this was an inappropriate cut off. Some relievers could have zero saves if they pitched in the middle of games. Thus, they could have inadvertently been put in the starting pitcher data. However, there was no better way to control for this, as online data does not account for this metric. Another indicator variable that could have skewed results is my injured variable. Any starter who pitched less than one hundred innings was considered injured, as this roughly equated to pitchers throwing for half of the season. Although one hundred innings is an extremely conservative estimate in order to eliminate some biasness, I cannot say with certainty that all pitchers who pitched fewer than 100 innings were in fact injured. Many could have been resting or given few starts as they transitioned to a longer workload. Again, there was no way to account for these unforeseen difficulties within the data set. In order to have better and more accurate results in the future, it would be important to adjust for these potential biased characteristics.

V. Conclusion

Baseball organizations often discriminate against shorter, lighter batters because they are believed to be less durable. Holding the production of two players constant, taller or heavier players are often chosen instead of shorter, lighter players. Thus, shorter, lighter players have to outperform taller, larger players in order to counteract the discrimination in the league. This remains a central issue as baseball is one the United States' most popular sports and is a growing industry.

Looking at the literature of this topic, countless people have done research to determine if height and weight are determinants of the production of a baseball player. Glen Greenberg (2008) uses pitcher data from 1990-2007 and finds that a player's height has little impact on their production metrics like ERA and strikeouts. Moreover, Gassko (2006) finds that shorter players have more hits and a higher OBP, while heavier players have higher home run totals. Thus, height and weight are two determinants towards productivity in the MLB, but little literature discusses their importance towards durability. This paper adds to existing literature by determining if height and weight are determinants to a player's durability and not just a player's productivity, where durability is defined as games played for batters and innings pitched for pitchers.

The results of this study support the hypothesis that short and light batters and pitchers are not actually less durable, but only perceived to be less durable by scouts and baseball organizations. A baseball player's height and weight are overemphasized in choosing which athletes are signed to baseball organizations and given starting roles. Moreover, in many cases, shorter athletes are actually more durable than taller athletes. Amongst batters, a 6'6" player would play roughly two games less than a 6'0" player, potentially due to injury. In my second specification for batters, short and light players play 1.0 and 2.6 more games than normal players throughout the regular season, disproving the common notion that these players are less durable. Moreover, the tall and heavy batters play 2.2 and 1.9 fewer games than their normal sized peers. Additionally, in specification 3, there seems to be even more evidence that shorter batters are more durable, as these batters play 3.1 more games than their average sized peers. Across the board, shorter and lighter batters seem to be more durable than their normal sized and heavier, taller counterparts.

Amongst starting pitchers and relievers, there is only circumstantial evidence that height and weight have an effect on a player's durability within the MLB. Amongst starters, in the first specification, individual weight and height are poor determinants in explaining the durability of starters. Additionally, short and light players pitch only 1.7 more innings than normal sized pitchers. Alternatively, amongst relievers in all three specifications, the coefficients in innings between different height and weight groups is small and predominately negligible. Therefore, there is little evidence that height and weight are good determinants for the durability of MLB pitchers.

Height and weight are interesting metrics in baseball, as they unfortunately sometimes help or inhibit the success of a player. For one, increased weight leads to higher production categories for home runs and OPS (Hamburg, 1999). Additionally, an increase in height can lead to an increase by .48 per nine innings for pitchers (Gassko). Therefore, both height and weight are found to help the production capabilities of an athlete. Baseball organizations and clubs constantly have to make choices between performance and durability on their rosters, as shorter, skinnier players are more durable, but heavier, taller players are often believed to be more productive, but are now found to potentially be less durable. In order for the MLB to counteract height and weight discrimination, when the production of two different sized players are equal, they should start the shorter and lighter player, as they typically will be more durable. If this is done, the MLB could become even more dominant in international competitions because their players will have relatively similar productive capabilities, but higher levels of durability.

Further studies could build on this paper by including more years of data before 2000. This could further the accuracy of future papers, and allow for more robust results in interactive variables like short, light players and tall, heavy players which both had relatively small sample sizes. Given all of these results amongst batters, starters and relievers, baseball clubs should try to recognize that height and weight are overplayed factors in determining which athletes play in games. Baseball organizations should ensure that shorter and lighter players have an equal chance in playing in the major league, compared to their normal sized and taller, heavier counterparts. This will ensure that baseball clubs are as economically efficient as possible with their resources and remain competitive amongst a sea of other baseball organizations.

References

Baumer, Benjamin, and Andrew S. Zimbalist. *The sabermetric revolution: assessing the growth of analytics in baseball*. Philadelphia: U of Pennsylvania Press, 2014.

"Body Measurements." *Centers for Disease Control and Prevention*. Centers for Disease Control and Prevention, 15 July 2016.

Brown, Maury. "MLB Sees Record Revenues For 2015." *Forbes*. Forbes Magazine, 04 Dec. 2015. Web.

Calculate Your Body Mass Index." *National Institutes of Health*. U.S. Department of Health and Human Services, n.d.

Crawford, Christopher. "Why Aren't There More Fat Baseball Players?" *VICE Sports*. Bleacher Report, 31 Jan. 2017.

Delgado, Luis. "Explaining the Luxury Tax in Major League Baseball." *SportingCharts.com*. N.p., n.d. Web. 16 Apr. 2017.

Foley, Katherine Ellen. "Are heavier baseball player's better hitters? Basically, no, says physics." *Quartz*. Quartz, 11 Oct. 2016. Web. 16 Apr. 2017.

Gassko, David. "Does Size Matter (Part 6)." *The Hardball Times*. N.p., 09 Nov. 2006.

Gassko, David. "Does Size Matter (Part 5)." Hardball Times, 08 July. 2007. Web.

Gassko, David. "Does Size Matter (Part 4)." Hardball Times, 14 Oct. 2008. Web.

Greenberg, Glen. "Does a Pitcher's Height Matter?" *Does a Pitcher's Height Matter? | Society for American Baseball Research*. N.p., n.d. Web.

Hamburg, Lydia, and Terence M. Mines. "Correlations For Weight, Height And Two Measures Of Batting Performance." *Perceptual and Motor Skills* 88.2 (1999):

Karkamp, Robb. "The Great Analytics Rankings." *ESPN*. ESPN Internet Ventures, n.d. Web. 18 Apr. 2017.

Lee, Christopher. "Baseball Player Height By Position." *Baseball Player Height By Position*. N.p., 24 Mar. 2010. Web. 03 Apr. 2017.

McLennan, Jim. "Baseball Players: Does Size Matter?" *AZ Snake Pit*. AZ Snake Pit, 05 July 2010. Web

"MLB Stats, Scores, History, & Records." *Baseball-Reference.com*. N.p., n.d. Web. 18 Apr. 2017.

Nathan, Alan. "Swing Speed vs. Bat and Batter Mass." *Sports Engineering* (2017): n. pag. Web.

Petti, Bill. "The Changing Size of MLB Players (1870-2010)." *Beyond the Box Score*. Beyond the Box Score, 19 Apr. 2011. Web. 16 Apr. 2017.

Appendix

Table 1: Batter Summary Statistics

Variable	Mean (Std. Dev)
Games Played	54.737 (49.724)
Height	73.403 (2.242)
Weight	194.575 (20.502)
Short	0.042 (0.200)
Normal Weight	0.929 (0.127)
Tall	0.029 (0.168)
Light	0.026 (0.158)
Normal Height	0.936 (0.085)
Heavy	0.038 (0.190)
Short & Light	0.008 (0.090)
Tall & Heavy	0.008 (0.089)
At Bats	154.411 (191.457)
Runs	20.583 (28.821)
Hits	40.292 (53.950)
Doubles	8.038 (11.273)
Triples	0.840 (1.742)
Home Runs	4.678 (8.194)

Runs Batted In	19.598 (28.485)
Caught Stolen	1.051 (2.170)
Walks	14.523 (21.629)
Intentional Walks	1.129 (2.949)
Strike Outs	31.391 (38.070)
Sacrifice Hits	1.421 (2.513)
Sacrifice Fly's	1.223 (2.066)
Hit By Pitch	1.561 (2.895)
Slugging	0.280 (0.220)
On Base Percentage	0.245 (0.755)
OPS	520.149 (363.680)
Observations	**18297**

Table 2: Pitcher Summary Statistics by Type of Pitcher

Variable	**Starters**	**Relievers**
	Mean (Std. Dev)	Mean (Std. Dev)
Innings Pitched	112.844 (66.496)	29.612 (24.840)
Height	74.710 (2.054)	74.264 (2.098)
Weight	197.572 (21.203)	196.211 (20.615)
Short	0.047 (0.212)	0.089 (0.285)
Normal Height	0.871	0.862

	(0.107)	(0.968)
Tall	0.082 (0.275)	0.049 (0.216)
Light	0.072 (0.258)	0.072 (0.259)
Normal Weight	0.872 (0.124)	0.888 (0.119)
Heavy	0.056 (0.231)	0.040 (0.196)
Short & Light	0.014 (0.115)	0.022 (0.148)
Tall & Heavy	0.020 (0.141)	0.009 (0.093)
Wins	6.653 (5.099)	1.481 (1.915)
Losses	6.619 (3.899)	1.500 (1.784)
Games	22.259 (11.627)	28.665 (24.835)
Games Started	18.042 (10.669)	0.146 (0.452)
Complete Games	0.583 (1.147)	0.001 (0.011)
Shutouts	0.237 (0.574)	0.001 (0.011)
Games Finished	1.312 (3.413)	9.741 (12.753)
Saves	0.139 (1.214)	2.668 (8.173)
Hits Allowed	115.435 (64.821)	28.452 (22.368)
Home Runs Allowed	13.582 (8.177)	3.195 (2.889)
Runs Allowed	58.940 (31.562)	14.550 (10.913)
Earned Runs Allowed	54.391 (29.104)	13.369 (10.043)
Walks Allowed	38.266 (22.390)	12.242 (10.090)

Strikeout	83.820 (57.846)	25.807 (24.443)
Wild Pitches	3.543 (3.221)	1.437 (1.825)
Balks	0.383 (0.722)	0.112 (0.366)
Earned Runs Allowed	4.754 (1.586)	5.763 (7.214)
Whip	1.433 (0.262)	1.626 0.994
Walks per Nine Allowed	3.333 (1.274)	4.576 (4.325)
Home Runs Per Nine Allowed	1.183 (0.542)	1.432 (4.377)
Strikeouts Per Nine Allowed	6.483 (1.719)	7.211 (3.241)
Hits Per Nine Allowed	9.597 (1.793)	11.163 (16.048)
Observations (All Variables)	**4515**	**7592**

Table 3: Durability by Height and Weight

	Batters	Starters	Relievers
Variables	Mean (Std. Dev.)	Mean (Std. Dev.)	Mean (Std. Dev.)
Short	72.438 (50.388)	120.715 (66.276)	30.582 (26.459)
Normal Height	54.980 (49.808)	111.384 (65.926)	29.605 (24.662)
Tall	27.268 (34.388)	109.874 (69.967)	29.402 (25.642)
Light	57.255 (50.535)	131.436 (65.011)	26.991 (24.566)
Normal Weight	52.431 (46.876)	108.015 (63.801)	29.714 (25.938)
Heavy	49.742 48.634	106.136 67.484	31.375 (24.333)
Short & Light	76.804	131.737	29.317

	(55.123)	(66.231)	(26.521)
Tall & Heavy	36.177	100.505	24.909
	(43.586)	(65.994)	(22.326)

Note: Durability measured as Games Played for Batters and Innings Pitched for Pitchers

Table 4: Determinants of Durability for Batters

	Specification 1	**Specification 2**	**Specification 3**
G	Coefficients (Std. Err.)	Coefficients (Std. Err.)	Coefficients (Std. Err.)
Height	-0.297*** (0.067)		
Weight	(0.037) 0.007		
Short Players		1.027** (0.626)	1.469*** (0.676)
Tall Players		-2.292*** (0.733)	-2.897*** (0.840)
Heavy Players		-1.917*** (0.649)	-2.355*** (0.711)
Light Players		2.552*** (0.786)	3.416*** (0.922)
Short & Light Players			3.117 (1.743)
Tall & Heavy Players			(2.662) (1.723)
Age during Season	0.622*** (0.030)	0.616*** (0.029)	0.617*** (0.029)
Runs	-0.092*** (0.029)	-0.078*** (0.029)	-0.079*** (0.029)
Hits	0.601*** (0.016)	0.602*** (0.016)	0.603*** (0.016)
Doubles	-0.098*** (0.042)	-0.103*** (0.042)	-0.102*** (0.042)
Triples	0.350*** (0.109)	0.365*** (0.109)	0.369*** (0.109)
Home Runs	-0.700*** (0.060)	-0.728*** (0.060)	-0.726*** (0.060)
Runs Batted In	0.104*** (0.027)	0.098*** (0.027)	0.098*** (0.027)
Stolen Bases	-0.226***	-0.227***	-0.226***

	(0.036)	(0.036)	(0.036)
Caught Stolen	0.402***	0.441***	0.441***
	(0.103)	(0.103)	(0.103)
Walks	0.185***	0.182***	0.182***
	(0.017)	(0.017)	(0.017)
Intentional Walks	-0.111***	-0.120***	-0.119***
	(0.059)	(0.059)	(0.059)
Strike Outs	0.358***	0.356***	0.355***
	(0.008)	(0.008)	(0.008)
Sacrifice Hit	0.520***	0.539***	0.538***
	(0.054)	(0.054)	(0.054)
Sacrifice Fly	0.545***	0.559***	0.560***
	(0.104)	(0.105)	(0.105)
Hit by Pitch	0.388***	0.402***	0.399***
	(0.058)	(0.058)	(0.058)
Times Grounded into Double Plays	0.556***	0.556***	0.556***
	(0.052)	(0.052)	(0.052)
Batting Average	-8.285***	-8.061***	-8.021***
	(1.867)	(1.869)	(1.869)
Slugging Average	26.139***	24.958***	24.844***
	(6.032)	(6.038)	(6.037)
On Base Percentage	24.681***	24.021***	23.899***
	(5.603)	(5.608)	(5.607)
On-base plus Slugging	-0.025***	-0.024***	-0.024***
	(0.005)	(0.005)	(0.005)

Note:** is significant at the 10% level, *** is significant at the 5% level.

Table 5: Determinants of Durability amongst Starting Pitchers

	Specification 1	**Specification 2**	**Specification 3**
IP	Coefficients (Std. Err.)	Coefficients (Std. Err.)	Coefficients (Std. Err.)
Weight	0.001		
	(0.005)		
Height	0.019		
	(0.054)		
Short Players		0.444	-0.210***
		(0.449)	(0.520)
Tall Players		0.088	0.230

		(0.356)	(0.396)
Heavy Players		0.631 (0.418)	-1.156*** (0.502)
Light Players		0.047 (0.370)	0.187 (0.403)
Short & Light			1.735 (0.911)
Tall & Heavy			0.997 (1.033)
Age	0.094*** (0.023)	0.094*** (0.023)	0.095*** (0.023)
Wins	0.814*** (0.051)	0.811*** (0.052)	0.814*** (0.052)
Losses	0.012 (0.055)	0.009 (0.055)	0.008 (0.055)
Games	0.692*** (0.025)	0.694*** (0.025)	0.692*** (0.025)
Games Started	2.665*** (0.060)	2.666*** (0.060)	2.664*** (0.060)
Complete Games	3.047*** (0.137)	3.037*** (0.137)	3.043*** (0.137)
Games Finished	0.287*** (0.074)	0.286*** (0.074)	0.289*** (0.074)
Shutouts	0.200 (0.243)	0.193 (0.243)	0.207 (0.243)
Saves	-0.792*** (0.124)	-0.791*** (0.124)	-0.796*** (0.124)
Hits Allowed	0.420*** (0.010)	0.420*** (0.010)	0.420*** (0.010)
Home runs Allowed	0.268*** (0.026)	0.269*** (0.026)	0.266*** (0.026)
Runs Allowed	-0.189*** (0.032)	-0.191*** (0.032)	-0.192*** (0.032)
Earned Runs Allowed	-0.249*** (0.033)	-0.248*** (0.033)	-0.247*** (0.033)
Walks Allowed	0.101*** (0.010)	0.102*** (0.010)	0.102*** (0.010)
Wild Pitches Allowed	0.018 (0.038)	0.018 (0.038)	0.020 (0.038)
Strikeouts	0.159***	0.159***	0.159***

	(0.004)	(0.004)	(0.004)

Table 6: Determinants of Durability for Relievers

	Specification 1	**Specification 2**	**Specification 3**
IP	**Coefficient (Std. Error)**	**Coefficient (Std. Error)**	**Coefficient (Std. Error)**
Weight	-0.009*** (0.003)		
Height	0.052*** (0.024)		
Short Players		-0.390*** (0.156)	-0.383*** (0.175)
Tall Players		0.050 (0.204)	0.050 (0.221)
Heavy Players		-0.662*** (0.225)	-0.661*** (0.250)
Light Players		0.323 (0.174)	0.332 (0.202)
Short & Light			(0.036) (0.389)
Tall & Heavy			0.001 (0.568)
Age	-0.036*** (0.011)	-0.038*** (0.011)	-0.038*** (0.011)
Wins	0.633*** (0.035)	0.636*** (0.035)	0.636*** (0.035)
Losses	-0.109*** (0.036)	-0.107*** (0.036)	-0.107*** (0.036)
Games	0.199*** (0.006)	0.199*** (0.006)	0.199*** (0.006)
Games Started	1.215*** (0.104)	1.210*** (0.104)	1.210*** (0.104)
Complete Games	1.406 (3.758)	1.544 (3.757)	1.545 (3.757)
Games Finished	0.114*** (0.013)	0.114*** (0.013)	0.114*** (0.013)
Saves	-0.105*** 0.016		

Hits Allowed	0.605*** (0.007)	-0.105*** (0.016)	-0.105*** (0.016)
Home runs Allowed	0.581*** (0.008)	0.605*** (0.007)	0.605*** (0.007)
Runs Allowed	-0.114*** (0.031)	0.286*** (0.026)	0.286*** (0.026)
Earned Runs Allowed	-0.250*** (0.032)	-0.115*** (0.031)	-0.115*** (0.031)
Walks Allowed	0.207*** (0.010)	-0.250*** (0.032)	-0.250*** (0.032)
Wild Pitches Allowed	-0.141*** (0.031)	0.208*** (0.010)	0.208*** (0.010)
Strikeouts	0.280*** (0.005)	-0.138*** (0.031)	-0.138*** (0.031)

Table 7: Determinants of Injury for Starting Pitchers

	Specification 1	**Specification 2**	**Specification 3**
Injured	**Coefficient (Std. Error)**	**Coefficient (Std. Error)**	**Coefficient (Std. Error)**
Weight	-0.001*** (0.000)		
Height	0.010 (0.010)		
Short		-0.019 (0.034)	-0.019 (0.039)
Tall		0.086*** (0.107)	0.089*** (0.114)
Light		-0.029 (0.031)	-0.029 (0.048)
Heavy		-0.014 (0.040)	-0.014 (0.056)
Short & Light			0.040 (0.215)
Tall & Heavy			0.098 (0.185)
Age	0.000 (0.000)	-0.003 (0.003)	-0.003 (0.003)

Wins	-0.029*** (0.010)	-0.029*** (0.007)	-0.029*** (0.007)
Losses	0.009*** (0.001)	0.011*** (0.006)	0.011*** (0.006)
Games	-0.019*** (0.00)	-0.020*** (0.004)	-0.020*** (0.004)
Games Started	-0.080*** (0.01)	-0.081*** (0.014)	-0.081*** (0.014)
Complete Games	0.060*** (0.020)	-0.060*** (0.024)	-0.059*** (0.024)
Games Finished	-0.020*** (0.010)	-0.015*** (0.006)	-0.015*** (0.006)
Shutouts	0.050 (0.040)	0.054 (0.042)	0.054 (0.042)
Saves	0.027*** (0.00)	0.026*** (0.009)	0.026*** (0.009)
Runs Allowed	-0.010*** (0.001)	-0.007** (0.004)	-0.007** (0.004)
Earned Runs Allowed	-0.010*** (0.001)	-0.009*** (0.004)	-0.009*** (0.004)
Walks Allowed	0.010*** (0.001)	0.007*** (0.002)	0.007*** (0.002)
Strike outs Per 9 innings	0.015*** (0.001)	0.060*** (0.008)	0.070*** (0.008)
Hits per 9 Innings	0.104*** (0.030)	0.106*** (0.025)	0.108*** (0.025)
Home Runs Per 9 innings	0.222*** (0.05)	0.222*** (0.052)	0.222*** (0.053)

Table 8: Variable Definitions for Batters and Pitchers

Batters	**Variable Name**
Height	Batter's Height in Inches
Weight	Batters Weight in pounds
Games Played	# of games a batter played in the 162 Game Season
Team Name	MLB Team the batter played for
Age	Batter’s Age during the season
At Bats	The Amount of Times a Batter Faces a Pitcher
Runs	A run is scored when a batter returns safely to home plate

Hits	When a batter successfully reaches first base
Doubles	When a batter successfully reaches second base
Triples	When a batter successfully reaches third base
Home Runs	When a batter circles all the bases and scores a run
Runs Batted In	When a batter’s actions leads a team member to reach home plate
Stolen Bases	When a batter advances to a base to which he is not entitled
Caught Stolen	When a batter fails an attempt to steal a base
Walks	If a batter receives four pitches that are called balls, he advances to first base without being called out
Intentional Walk	When a pitcher purposefully walks a batter so he cannot swing at the pitch
Strikeouts	When a batter accumulates three strikes during a time at bat, which ends his at bat
Sacrifice Fly Balls	When a batter sacrifices his at bat by hitting the ball into the outfield in order to advance a runner on base
Hit By Pitch	When a batter is hit by a pitch he is awarded first base
Double Plays Grounded Into	When a batter hits a ground ball that leads to two baserunners getting out
Batting Average	# of hits/# of At Bats
Slugging Average	# of Total Bases/ # of At Bats
On Base Percentage	(Hits + Walks + Hit by Pitch) / (At Bats + Walks + Hit by Pitch + Sacrifice Flies)
OPS	On Base Percentage + Slugging Percentage
Pitchers	**Variable Name**
Height	Pitcher's Height in Inches
Weight	Pitcher's Weight in pounds
Innings Pitched	# of Innings Pitched for Pitchers
Team Name	MLB Team the pitcher played for
Age	Age during Season
Wins	A pitcher receives a win when he is the pitcher who holds, or takes the lead
Losses	A pitcher is accredited with a loss when he gives up runs that lead to the other team taking the lead
Games Played	Total amount of games pitched in
Games Started	When a pitcher throws the first pitch for his team
Complete Games	When a pitcher throws the entire game regardless of length
Shutout	When a pitcher(s) holds the opposition to zero runs
Games Finished	When a pitcher pitches the last pitch of a game

Saves	When a pitcher finishes a game for the winning team
Hits Allowed	# of Hits a Pitcher allows
Runs Allowed	# of Runs a Pitcher allows
Earned Runs Allowed	# of Runs a Pitcher allows
Walks Allowed	# of Walks a pitcher allows
Intentional Bases on Balls	# of times a pitcher purposefully walks a batter
Strikeouts	# of Strikeouts a pitcher earns
Wild Pitches	When a pitch is unable to be caught by the catcher with reasonable effort, all baserunners advance
Balks	An illegal act by a pitcher that leads to runners advancing
Earned Runs per nine innings	9 x (Earned runs Allowed/Innings Pitched)
Hits per Nine Innings	9 x (Hits Allowed/Innings Pitched)
Home Runs per Nine Innings	9 x (Home Runs Allowed/Innings Pitched)
Walks per Nine Innings	9 x (Walks Allowed/Innings Pitched)
Strikeouts per Nine Innings	9 x (Strikeouts/Innings Pitched)
WHIP	(BB+H/IP)