Basketball Players' Long Wingspan Enhances Defensive Upside but Hinders Shooting Capabilities:

A Statistical Inquiry of the Effect of NBA Player Wingspans on Shooting Accuracy and Defensive Impact

Ali R. Kaazempur-Mofrad*
Department of Statistics, University of California, Los Angeles

This manuscript is a **preprint**.

Example citation:

Kaazempur-Mofrad, Ali R. "Basketball Players' Long Wingspan Enhances Defensive Upside but Hinders Shooting Capabilities: A Statistical Inquiry of the Effect of NBA Player Wingspans on Shooting Accuracy and Defensive Impact." SportRxiv.

Abstract

The objective of this study is to examine the relationship between an NBA player's wingspan and their ability to shoot the basketball and perform defensively. Using regression analysis methodologies, the results of this study suggest that NBA players with a longer wingspan are able to perform better defensively, but also perform worse in shooting the ball. Based on these findings, NBA teams who prioritize length when evaluating draft prospects should potentially reconsider their approach due to the inverse relationship between wingspan and shooting accuracy.

1. Introduction

In the world of professional sports, the ultimate goal of each franchise is to win a championship. In the National Basketball Association (NBA), the front office of each team has the important role of forming and developing a roster of players with the best chance of reaching the apex of the league, winning a championship. While there are many ways to form a

*Email: amofrad@ucla.edu

roster, such as through trades and free agency, one of the best opportunities of building for the future is the NBA Draft. Once per year, teams take turns making draft selections from a pool of incoming young players that they believe could develop into key players for the future of their respective franchises.

Prior to the annual NBA draft, teams go through the thorough process of scouting and hosting player workouts. Alongside the essential evaluation of on-court production at the college or international level, many prospects are also evaluated by their physical attributes. To provide teams with information on the physical traits of prospects, the NBA hosts an annual Draft Combine where prospects participate in various drills and are measured for their size with measurements ranging from height and weight to wingspan and hand size.

It is generally perceived that players with longer measurements are more of a premium as they are expected to have the physical tools to succeed, especially on the defensive end [1]. Due to this, some prospects who have standout measurements are viewed as draft selections with higher potential, even if they're a bit further behind in their development of skills such as shooting. The reasoning behind this outlook is that teams believe that the prospect, with the help of the training staff, can develop their skills over time and thus have a higher ceiling than a smaller player with a similar skillset.

To assess the validity of this draft selection approach, I will examine and evaluate the relationship between wingspan and the ability of players to shoot the basketball and perform defensively.

2. Methods

In this section, I will introduce the terminology, model, and model validation approaches designed to examine the potential relationship between a player's wingspan and their ability to shoot and defend opposing teams.

2.1. Terminology

Before introducing the statistical methods implemented in this study, some additional context and definitions are introduced below:

- Defensive Win Shares (DWS) - This is an advanced statistic in basketball that measures a player's impact by estimating how many wins the player contributes to their team as a result of their defensive contribution [2].
- Free Throw Percentage (FT%) - The free throw is a shot in basketball that is typically taken after a player is fouled in the act of play. During each free throw attempt, a player has 10 seconds to shoot the basketball from a stationary position of 15 ft from the court's baseline. This shot is taken without the contesting of an opposing defensive player. The FT% is calculated by dividing a player's free throws made by free throws attempted [3].

- Wingspan - A player's wingspan is measured by the distance between the tips of their fingers while extending both arms to the side and perpendicular to the body. The unit of measurement used is inches.

2.2. Model

In order to construct the optimal linear regression model for this study, I utilized several methodologies.

Prior to conducting model diagnostics and determining the structure of the model, I randomly divided the dataset into training and test sets in order to validate the model that I reached. The purpose of this process is to ensure that the final model is appropriate, not only for a sample of the data, but also the population sample. The model used in the study was fit onto the training set, which represents a randomly sampled 50% of the total data observations.

The primary shooting metric implemented in the study is the FT% because it is a consistent shot that is situationally equivalent with the lack of opposing defenders for every attempt. The primary defensive metric used in this study is the career average DWS.

In order to check whether any other predictor variables could be useful in the model, I conducted both backward and forward selection processes while ensuring that FT% and DWS are included. Using the Akaike's Information Criterion (AIC) to test other variables, the selection processes suggested implementing Three Point Percentage (3P%) as an additional variable. I decided, however, to reject this suggestion because this is a shot that is guarded and thus not equivalent for every player. Upon using the selection processes on the model, I concluded that no other variables shall be implemented in the model.

Upon choosing the variables for the study, the model diagnostics and test of assumptions were conducted. The standard model with the raw variables violates some assumptions (particularly with normality) and thus it was necessary to apply transformations on some variables[1].

The final model structure used in the study is:

$$Wingspan = \beta_0 + \beta_1 * FT\%^3 + \beta_2 * DWS^{1/3}$$

2.3. Model Validation

To validate the fitted model, I applied the same variable transformations on the test set. Evaluating the resulting model summaries between the training and test sets, the coefficients for both the $FT\%^3$ and $DWS^{1/3}$ variables of the training set are within range of the standard error of the corresponding coefficients of the test set. Additionally, the adjusted R^2 values of the corresponding models for each set are relatively similar. The variance inflation factor (VIF) of our predictors is approximately 1, which informs us that the model does not have any

[1]Please reference the Appendix for transformed model diagnostics

issues of variance with multicollinearity. Finally, the test set doesn't present any new model violations. Based on these observations, I concluded that the proposed model structure is validated.

3. Results

Prior to fitting the model, it is important to properly analyze the data sample. I began by conducting exploratory analysis (EDA) of the dataset and variables.

The goal of the study is to determine the effect of a player's wingspan measurement on their shooting and defensive capabilities. In order to obtain an understanding on the general trends and behavior of the data, I first investigate each variable (see Table 1 below).

Pos	Avg. Wingspan	Wingspan Var.	Avg. DWS	DWS Var.	Avg. FT%	FT% Var.
PG	78.18462	7.773978	0.9090551	0.5410080	0.7628133	0.0059244
SG	80.89552	3.589298	0.8638066	0.6208802	0.7564046	0.0087587
SF	83.13922	4.758931	1.1541577	0.7865430	0.7389160	0.0051216
PF	85.06410	3.979335	1.2693898	0.6757775	0.7101241	0.0035056
C	86.40385	6.213942	1.6500855	0.7143390	0.7335134	0.0039357

Table 1: Summary of the mean and variance of the variables

The variable summaries can also be illustrated with histograms (see Figure 1).

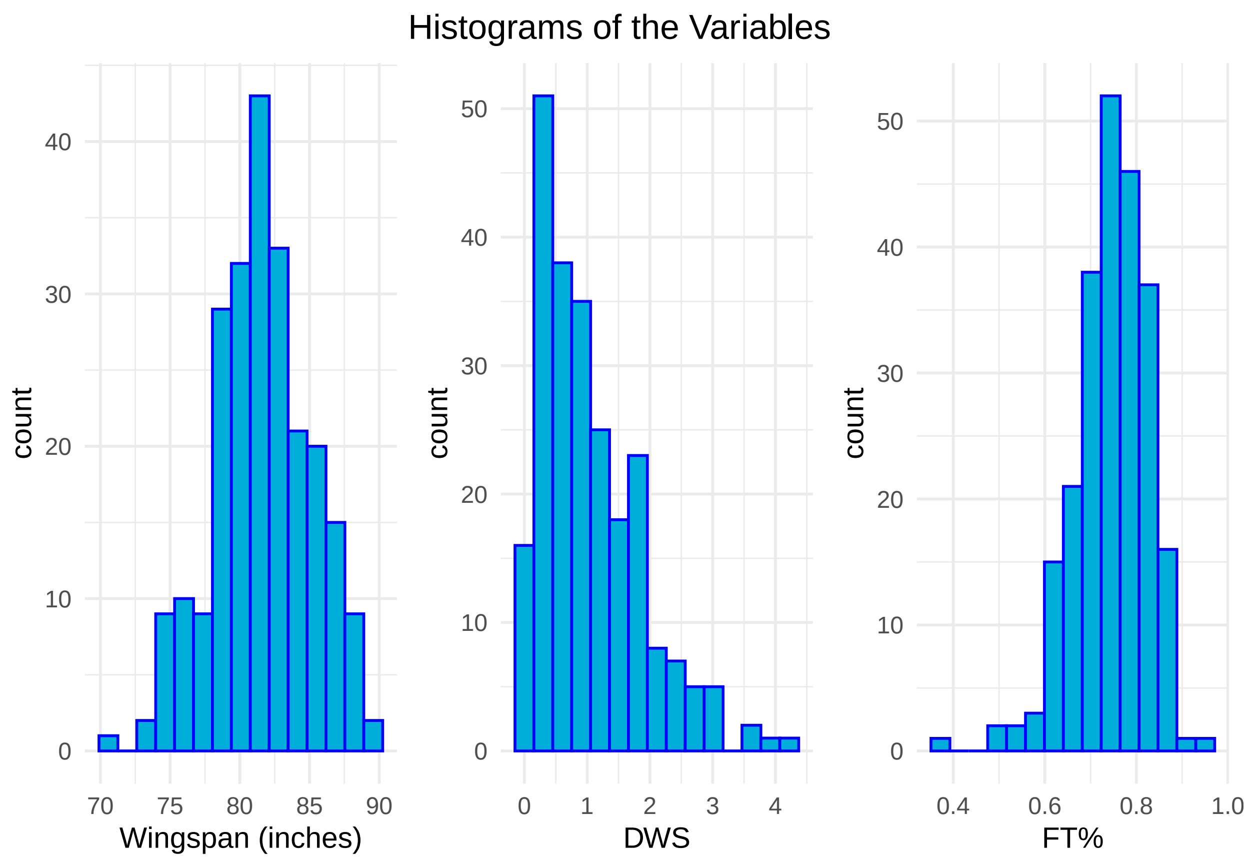

Figure 1: Histograms of Wingspan, Defensive Win Shares (DWS), and Free Throw Percentage (FT%) of NBA players

In order to visualize the relationship of the predictor variables (FT% and DWS) with the outcome variable, plots of each predictor against Wingspan are utilized. I first examine the scatterplot of FT% against Wingspan (see Figure 2). As evident in the figure, a slight decrease in FT% results in an increase of Wingspan, implying that there may be an inverse relationship between FT% and Wingspan.

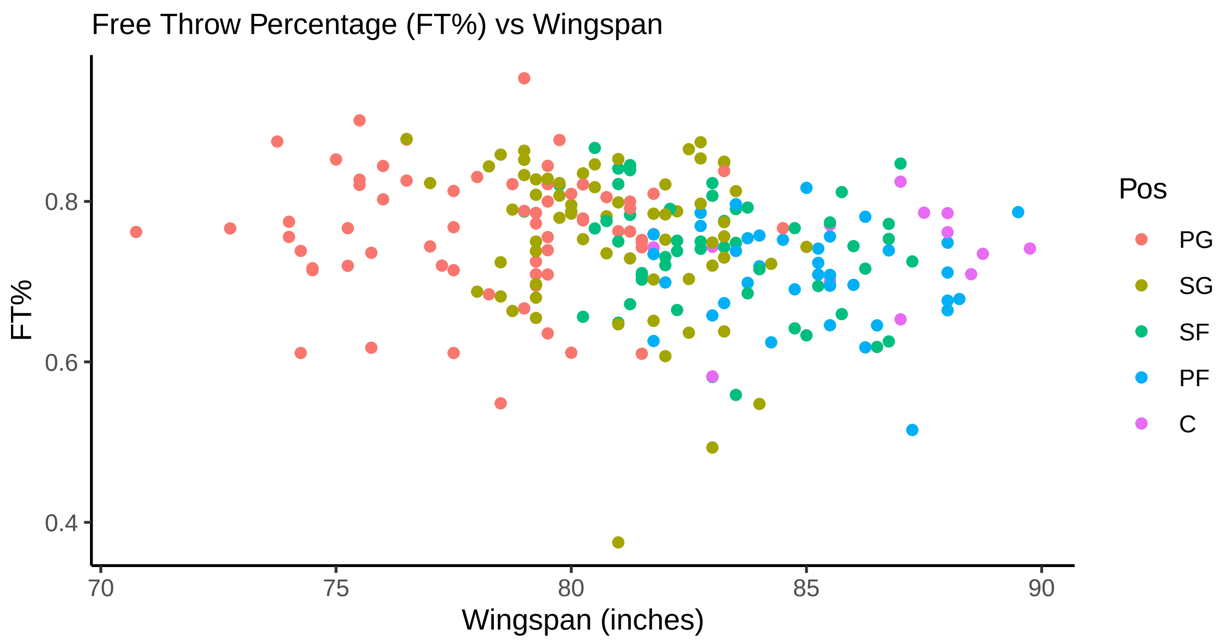

Figure 2: Scatterplot of Free Throw Percentage (FT%) vs. Wingspan for NBA players. The data points are color-coded to represent each player's specific position

In order to assess the relationship between DWS and Wingspan, I examine the scatterplot of the predictor against the Wingspan measurement (see Figure 3). As shown in the figure, an increase in Wingspan results in an increase in DWS, suggesting that there is positive relationship between the variables.

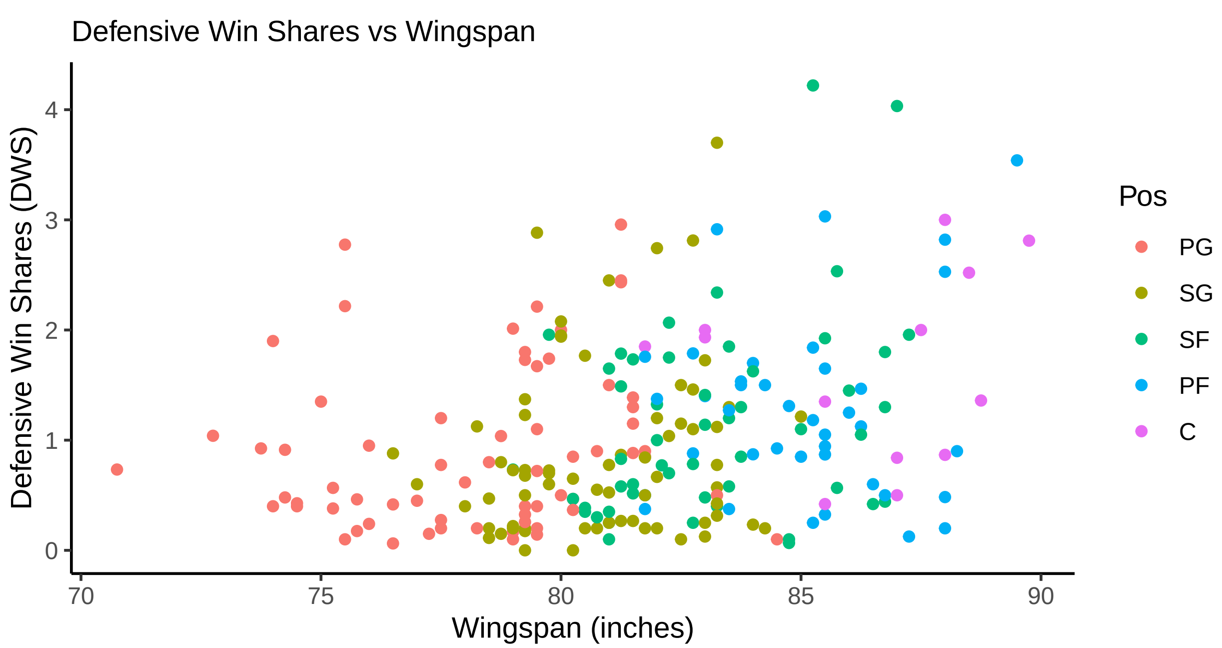

Figure 3: Scatterplot of Defensive Win Shares (DWS) vs. Wingspan for NBA players. The data points are color-coded to represent each player's specific position

The resulting coefficients for the model of the relationship between Wingspan and FT% and DWS are presented below (see Table 2).

Variable	β	Value
Intercept	β_0	80.32843
$FT\%^3$	β_1	-8.0914762
$DWS^{1/3}$	β_2	4.8102438

Table 2: The coefficient values for the fitted linear regression model

Inputting these coefficient values into the model structure gives the final linear regression model:

$$Wingspan = 80.328 - 8.091 * FT\%^3 + 4.81 * DWS^{1/3}$$

4. Discussion

This study was aimed at understanding the relationship between an NBA player's wingspan and ability to shoot the basketball and perform defensively. This information would enable scouts and team front offices to make a better judgement on the impact of the wingspan measurement at the annual NBA Draft Combine.

4.1. Model Intepretation

The results of the present linear regression model were analyzed to determine how a certain amount of change in one predictor will effect the outcome. Recall that for the FT% predictor I applied a cubic transformation and used $FT\%^3$ as the predictor. The corresponding coefficient for this predictor can be interpreted as follows:

- Controlling for the DWS, a 1 unit increase in $FT\%^3$ will result in an 8.09 inch decrease in Wingspan.

This interpretation confirms that players with longer wingspans are typically less accurate at shooting the ball.

For the DWS predictor, I applied a cubic root transformation and used $DWS^{1/3}$ as the predictor. The coefficient for this predictor can be interpreted as follows:

- Controlling for the FT%, a 1 unit increase in $DWS^{1/3}$ will result in a 4.81 inch increase in wingspan.

This result suggests that players with longer wingspans usually have a greater impact on the defensive end.

The coefficient of determination (R^2) of the model suggests that 17.9% of the variation in a player's wingspan measurement is described by the model.

4.2. Limitations

Although thousands of player have played in the league since the NBA started in 1946, the available dataset provides Draft Combine measurements from 2009 to 2017 only. While this still enabled creating a model using over 200 data observations, many of the corresponding players in this study are still in the midst of their professional careers. This means that, particularly for the younger players on the list, the DWS and FT% data may not end up being an accurate reflection of their full career. Several players tend to gradually develop and improve their statistics as they gain more playing experience.

4.3. Implications and Conclusions

The results of this study indicate that, on average, a player who has longer arms will typically have a greater defensive impact, but also will sacrifice a bit of shooting accuracy. This has quite significant implications for NBA franchises who are looking to make the best selection at the NBA Draft. The common perception that players with longer wingspans are a better selection may not always be an optimal approach to scouting young prospects.

While the exact cause of this decrease in shooting accuracy is not known, one possible hypothesis could be that an increased arm length creates greater room for error in the mechanics of shooting a basketball. It could be potentially more difficult for players with longer wingspans to shoot more accurately because the mechanics require greater precision.

Although the results of this study suggest that longer wingspans typically sacrifice shooting accuracy for defense, there are always outlier cases where a player with a long wingspan can develop their shot enough to be efficient as a shooter while also benefiting from having premium length. Players who can overcome this hurdle are typically elite players, such as NBA superstar Kevin Durant (89" Wingspan) who holds a career average of 3.08 DWS and 88.2% FT% [4].

References

1. Bleacher Report. “NBA Basketball: What Qualifies as a ‘Freakish’ Wingspan in the League?” Bleacher Report, Bleacher Report, 3 Oct. 2017, https://bleacherreport.com/articles/864549-nba-basketball-what-qualifies-as-a-freakish-wingspan-in-the-league.

2. “Defensive Win Shares - DWS Definition - Sporting Charts.” SportingCharts.com, https://www.sportingcharts.com/dictionary/nba/defensive-win-shares-dws.aspx.

3. “Basketball Free Throw Line.” Rookieroad.com, https://www.rookieroad.com/basketball/the-court/free-throw-line/.

4. “Kevin Durant Stats.” Basketball, https://www.basketball-reference.com/players/d/duranke01.html.

5. Goldstein, Omri. “NBA Players Stats since 1950.” Kaggle, 27 Apr. 2018, https://www.kaggle.com/drgilermo/nba-players-stats?select=player_data.csv.

6. “NBA Draft Combine Measurements - Dataset by Achou.” Data.world, 2 May 2018, https://data.world/achou/nba-draft-combine-measurements.

Appendix

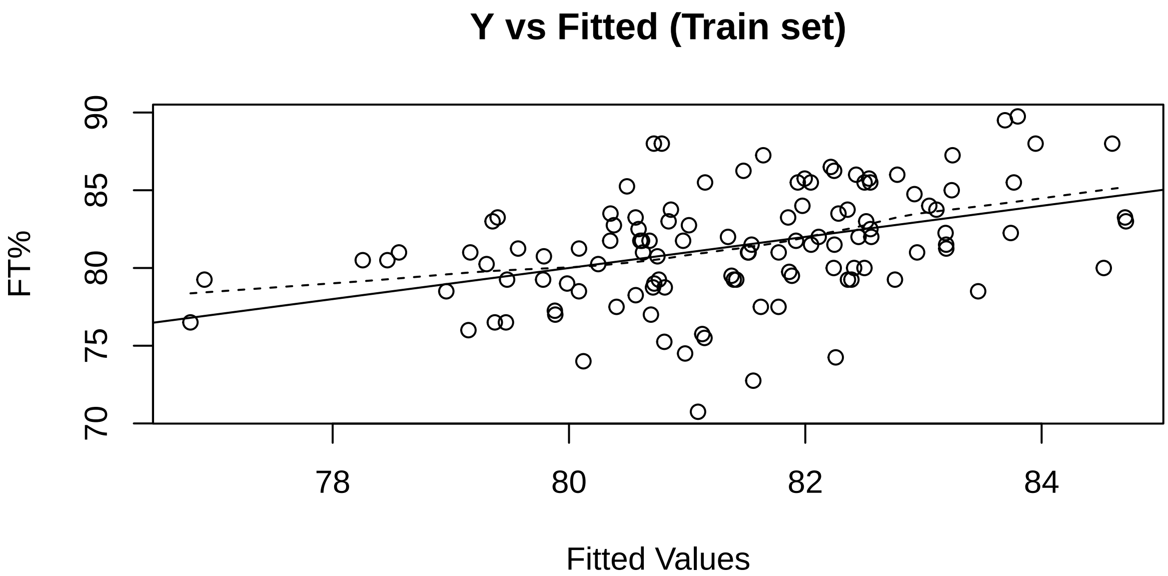

Figure 1: The plot suggests that the conditional mean response of the model with the applied transformations is a function of a linear combination of the predictors.

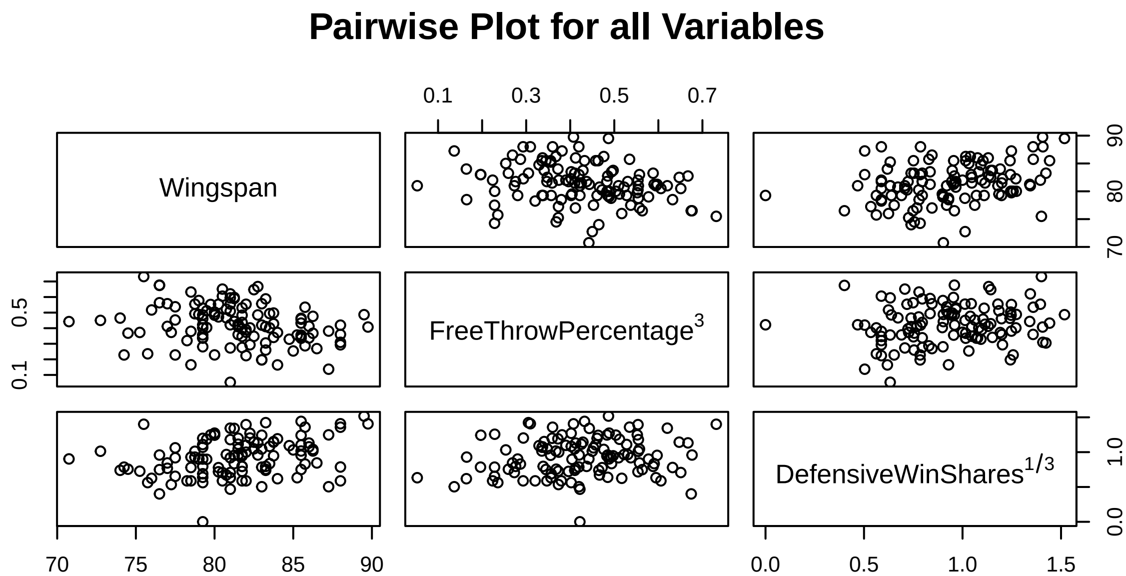

Figure 2: The conditional mean of each predictor in the transformed model has a linear function with the other predictor.

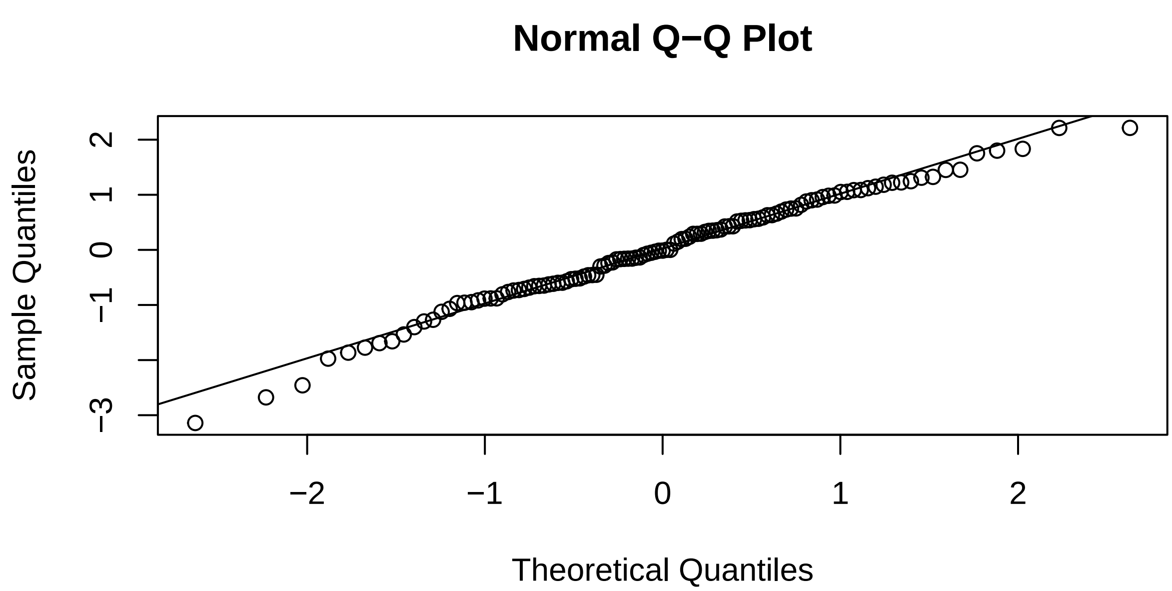

Figure 3: The model with the applied transformations satisfies the assumption of normality.