Title
Is the All-NBA Selection Process Biased? Examining the Effect of Market Size on Media Member Votes for All-NBA Teams

Author
Tam, Kevin

Publication Date
2022-07-13

Undergraduate

Is the All-NBA Selection Process Biased?

Examining the Effect of Market Size on Media Member Votes for All-NBA Teams

Kevin Tam

Department of Economics, University of California, Santa Barbara

Econ 196 Senior Honors Seminar

Advisor: Professor Ted Frech

March 14, 2022

Abstract: This paper examines the effect of market size on the votes for the National Basketball Association's All-NBA teams. While the All-NBA teams are intended to represent the league's best players, prior research suggests that media members vote based on factors other than performance. I hypothesize that a player on a large market team is more likely to receive votes than a player with comparable performance but on a small market team. Using All-NBA ballots from five consecutive seasons and Nielsen television market sizes, I employ a two-part model approach to determine the effects of market size, being born outside the United States, representing an East Coast team, and Team Win-Loss Percentage on All-NBA votes. I find that the effect of market size on All-NBA votes is unclear, but Team Win-Loss Percentage has a significant positive effect. Future studies could introduce other factors to the models such as age and race.

1. Introduction

The end of every regular season for the National Basketball Association is followed by the playoffs and selection of the three All-NBA teams. All-NBA is a prestigious individual award that distinguishes a professional basketball player as one of the best in the league for a particular season. Every player dreams of being selected and adding the award to their career resume, as just being chosen once can considerably improve a player's career prospects. If selected to the All-NBA teams, not only does a player receive increased name recognition on a league-wide level, but he also becomes eligible for higher future compensation from his current team. Therefore, the All-NBA selection process has financial consequences for players and their teams, and the results can impact the competitive landscape of the league.

While the All-NBA teams are intended to be composed of the league's best players, prior research suggests that player performance alone does not fully explain the selections made by media members. In other words, there is at least one unidentified factor influencing media members to vote for certain players over others. This study examines the question of whether market size has an effect on how media members select players to the National Basketball Association's All-NBA teams. Given two players with comparable regular season performance, I hypothesize that the player on the team with the larger market size is more likely to receive All-NBA votes than the player on the team with the smaller market size.

To address this hypothesis, I construct a dataset consisting of All-NBA vote results, Nielsen television market sizes by team, and player statistics from five consecutive NBA seasons (2014-2015 through 2018-2019). Using this dataset, I attempt to determine the effects of the following four factors on the distribution of All-NBA votes while controlling for player performance: market size, whether a player was born outside the United States, whether the player represents an East Coast team, and Team Win-Loss Percentage. Since the majority of players in the dataset did not receive any All-NBA votes, I employ a two-part model approach by running one set of regressions on the whole dataset and another set of regressions on a partial dataset consisting of only players who received at least one All-NBA vote. The first set of regressions consists of probit and logit models determining the differences between players that received at least one All-NBA vote versus players that did not receive any All-NBA votes. The second set consists of OLS models intended to examine the effect of market size and the other variables of interest conditional on the outcome that the player received at least one All-NBA vote. I control for player performance by using different combinations of PER and its twelve box score statistic components, and the explanatory power of the box score statistics is evaluated using Wald tests.

I find that the effect of market size on All-NBA votes is positive in some of the models but negative or insignificant in the others. I obtain similarly inconclusive results for both the Foreign and East indicator as the effects vary in magnitude, sign, and significance level across the models without a discernible pattern. Team Win-Loss Percentage is the only factor with a positive effect on All-NBA votes in every single model.

I am unable to conclude whether market size has an effect on a player's All-NBA votes. Despite this unsatisfying result, this study contributes to the established literature by providing evidence that Team Win-Loss Percentage may be the factor introducing bias to the All-NBA selection process. A future study could expand on the effect of Team Win-Loss Percentage or lead to different results by incorporating other factors into the models such as player age and race.

2. Literature Review

Background on the All-NBA Teams and the All-NBA Selection Process:

The All-NBA teams are intended to represent the best professional basketball players by position in the league for a season. During the playoffs, after the conclusion of the regular season, the NBA announces the three All-NBA teams (First, Second, and Third) that each consist of five players: two Guards, two Forwards, and one Center (Fernandes, 2017). Only fifteen players representing approximately the top 3.33% of the entire league receive the distinction each season, so All-NBA selections are considered very prestigious.

Players are selected to the All-NBA teams through a vote by a group of media members appointed by the league. For example, the All-NBA teams for the 2020-2021 season were voted on by a global panel of one hundred sportswriters and broadcasters instructed to “vote for the player at the position he plays regularly” (NBA, 2021). The votes are tabulated by Ernst & Young, an independent accounting firm, and the All-NBA teams are formed by converting votes into points: five points per First Team vote, three points per Second Team vote, and one point per Third Team vote (NBA, 2021). Based on the vote results, the First Team is comprised of the two Guards, two Forwards, and Center with the highest points in their respective positions; the two Guards, two Forwards, and Center with the next highest point totals by position are included in the Second Team, and the next five after that make up the Third Team (Fernandes, 2017).

All-NBA selections are important because the players selected become eligible for higher yearly compensation. The maximum amount of compensation a player can receive per season, known as the ‘Supermax’ contract extension, is defined in the current Collective Bargaining Agreement (CBA) between the league and players’ association as a percentage of the current salary cap (NBA & NBPA, 2017). If signed to a ‘Supermax’ contract extension, a player with less than six years of experience in the league can receive up to 25% of the salary cap per year, a player with seven to nine years of experience can receive up to 30%, and a player with ten or more years of experience can receive up to 35% (NBA & NBPA, 2017). Applying this definition to the current season (2021-2022), a player signed to a 'Supermax' before the season started will have a salary of around $28 to $40 million per year (Sports Reference, 2021). To become eligible for a 'Supermax,' a player must meet one of the following requirements: "(i) named to the All-NBA first, second, or third team or (ii) named defensive player of the year, in each case in the immediately prior season or in two of the three prior seasons" (Freedman, 2019). Since there can only be one Defensive Player of the Year (DPOY) each season, it is more likely for a player to be selected to All-NBA than to receive the DPOY award. Only the player's current team can offer the 'Supermax' to the player, as it represents a contract extension, so an All-NBA selection can be considered as both a reward for a player's individual accomplishments and a monetary advantage for teams to re-sign their star players.

The problem with requiring an All-NBA selection to be eligible for a 'Supermax' contract extension is that a player's potential future earnings are directly impacted by the votes of a few media members. Consider the case of Klay Thompson, a shooting guard who plays for the Golden State Warriors. At the end of the 2018-2019 season, Thompson narrowly missed being selected to the All-NBA Third Team (Bednall, 2019). If Thompson were selected to the Third Team, then he would have been eligible for a five-year, $221 million 'Supermax' extension from the Warriors that offseason; however, by being left off the All-NBA teams, the maximum that the Warriors could offer decreased by a total of $30 million to five years for $191 million (Bednall, 2019). While the Warriors were able to re-sign one of their star players for less money, consider the case of another NBA team that lost its star player due to the outcome of the media vote. After the 2016-2017 season, the Utah Jazz were unable to offer their star forward Gordon Hayward a $207 million “godfather offer” to stay with the team because he was not selected to any of the All-NBA teams (Fernandes, 2017). Because the Jazz could not offer Hayward more money than any other team, Hayward left in free agency that offseason to sign with the Boston Celtics, reuniting with his former college (Fernandes, 2017). The stories of Klay Thompson and Gordon Hayward show how the votes of a few media members impacts not only the futures of individual players but also their teams and the league as a whole.

Summary of Prior Research on the All-NBA Selection Process:

Prior research on the All-NBA selection process has focused on whether classification models can accurately predict the All-NBA teams. For example, Levine (2019) attempted to connect on-court performance with All-NBA voting patterns by comparing multiplication classification models on a controlled dataset. The purpose of this research was to find the best possible classification model for explaining voting patterns and to identify the player statistics that were the most accurate predictors of All-NBA selections. Levine (2019) found that a bagged classification tree run on the original unbalanced dataset predicted voting patterns the best, and assists per game was the most sensitive player statistic in determining All-NBA selections. However, the author admitted that various off-court inputs, including social media following, endorsement deals, and city data, were omitted from the analysis. I build on Levine’s work by examining the effect of city data, specifically television market size, on All-NBA selections.

Using a similar approach as Levine (2019), Na (2021) created several different classification models to determine which NBA statistics are the most impactful in the selection of All-NBA guards. According to Na (2021), the most recent criticism and feedback on the All-

NBA selection process has been directed towards the selection of All-NBA guards. Therefore, the purpose of this study was to provide transparency into the criteria that media members apply to players for only one position in the All-NBA teams. Na (2021) found that the best model correctly predicted only four of the six guards in the All-NBA team for the 2019-2020 season, providing evidence that there is inconsistency in the current selection process. In other words, there was a metric omitted from the model that had an effect on the selection of All-NBA guards. I extend the implications of Na’s study to the other two All-NBA positions of forwards and centers. By including data for television market size, national origin, and team location, I attempt to identify the specific factor supposed by Na that introduces bias to the All-NBA selection process.

The effect of player statistics on All-NBA selections has been covered in studies such as Randrianasolo (2021). This study utilized empirical analyses based on league criteria (like PER) to develop a methodology for how players should be selected to the All-NBA teams. Randrianasolo (2021) performed a series of analyses of variance to compare player efficiency ratings (PER) across the three All-NBA teams and against a group of players that had a Top 15 PER in the league but were not selected to an All-NBA team. The most important finding of this study was that players with a Top 15 PER but not selected were equivalent to Second Team selections and had a higher PER than Third Team selections. As a result, Randrianasolo (2021) concluded that the current way the NBA selects All-NBA teams is unfair and should be changed because some of the league’s best players are not selected. I incorporate the findings of this study into my research by controlling for PER and other player statistics. Therefore, I can be more confident that factors other than player statistics, such as market size, introduce bias in the All-NBA selection process.

Other studies done on NBA-related topics have used Nielsen television market ratings as a measure for market size. For example, Lee, Leonard, and Jeon (2009) examine pay and performance in the NBA across several different markets measured by Nielsen television rating. The authors found that players' compensation was negatively correlated with the number of competing teams in a given local sports market.

Studies investigating the differences in salaries between NBA players have identified race as a potential factor; however, previous studies into the All-NBA selection process have found no relationship between All-NBA votes and race. One notable study by Johnson and Minuci (2020) analyzed free agency contract signings from 2011 to 2017 and found that Black athletes are paid significantly less than athletes of other racial identities. The authors argued that consumer discrimination is the primary source for the NBA's racial wage gap and referenced the theory that consumer perceptions about an employee's race are an important determinant of labor market outcomes. Despite the extensive amount of research done on the effect of race on NBA salaries, I decided not to include the effect of race in my models because previous studies such as Levine (2019) found no evidence of race having an effect on All-NBA votes. Instead, I turn to another demographic factor that may be introducing bias: a player's country of origin.

3. Empirical Strategy

<u>The Approach:</u>

To examine whether market size affects how media members vote for the All-NBA teams, I run several regressions with All-NBA points as the dependent variable and Nielsen television market size as the primary independent variable. All-NBA points is a measure that combines the votes a player received in a season for each of the three All-NBA teams into a single number comparable across players. To convert All-NBA votes into All-NBA points, I follow the convention used by the NBA in tabulating the results of the media vote: five points per First Team vote, three points per Second Team vote, and one point per Third Team vote. Since the voting panel for the All-NBA teams consists of one hundred media members, a player can receive at most one hundred First Team votes or five hundred All-NBA points. For the 2014-2015 and 2015-2016 seasons, the voting panel consisted of 129 members. To preserve comparability across seasons, I scale the voting results from these two seasons to base one hundred.

All-NBA votes are concentrated among only a few players, resulting in most players not receiving any All-NBA votes. This poses a major challenge to the empirical work because the coefficients in any regression models run on the dataset would be skewed towards zero by the mass of observations with a zero outcome, potentially concealing the effect of market size or the other variables. To address this problem, I apply a two-part model approach by running two sets of regressions: one on the whole dataset and the other only including players that received All-NBA votes. The ‘Part 1’ regressions are linear probability models on a zero-one binary variable, and their purpose is to compare players that received All-NBA votes with players that did not receive any All-NBA votes. I run both probit and logit regressions for the ‘Part 1’ models because there is no reason provided by the dataset to justify picking one form over the other. The ‘Part 2’ regressions are conditional on a positive outcome, that the player received All-NBA votes, and their purpose is to identify whether market size or any of the other variables of interest has a causal effect on All-NBA votes.

In addition to market size, I include three other factors in my models and observe how All-NBA points changed while holding market size constant to test for other sources of bias in the All-NBA selection process. To explore the possibility of East Coast bias in professional basketball, I include an indicator that distinguishes players on teams based in the Eastern Time Zone from players on teams not based in the Eastern Time Zone. I include a Foreign indicator that distinguishes between players born in the United States from players born outside the United States to investigate whether media members have a bias against foreign-born players. Lastly, I include the player's Team Win-Loss Percentage to examine whether media members vote more for the players on teams with good records versus players with comparable season statistics but on teams with bad records. Team Win-Loss Percentage is calculated as the number of games won by the player's team during the season over eighty-two, the total number of games each team plays in a regular season.

Based on the work of Randrianasolo (2021) providing evidence of a relationship between the Player Efficiency Rating (PER) and All-NBA selections, I choose to use PER as the primary control for player performance in my models. According to John Hollinger (2005), Senior NBA columnist at The Athletic, "The PER sums up all of a player's positive accomplishments, subtracts the negative accomplishments, and returns a per-minute rating of a player's performance" (p. 6). The advanced metric is calculated from twelve box score statistics recorded for each player in every game: Field Goals made, Free Throws made, missed Field Goals, missed Free Throws, Three-point Field Goals made, Offensive Rebounds, Defensive Rebounds, Assists, Steals, Turnovers, Blocked Shots, and Personal Fouls. To test whether PER and its components add explanatory value to the All-NBA selection process, I utilize the following three sets of controls: only PER, PER and its twelve box score statistic components, and only the twelve box score statistics. Then, I perform Wald tests on the models controlling for PER and its box score statistic components and the models controlling for only the twelve box score statistics. A Wald test determines whether a set of variables adds explanatory power in a regression model. The null hypothesis of a Wald test is that the coefficients of a particular set of variables in a regression are equal to zero, and the null is rejected if the statistic generated from the test has a corresponding p-value close to zero. The Wald tests on the models controlling for PER and the twelve box score statistics test the explanatory power added by the twelve box score statistics in addition to PER. The Wald tests on the models controlling for only the twelve box score statistics test the explanatory power of the twelve box score statistics in the absence of PER. The additional explanatory power of PER is determined by observing the coefficient on PER with and without also controlling for the box score statistics. As a last step, I regress PER on the twelve box score statistics and create a correlation matrix to examine how the statistics are individually and as a whole correlated with PER.

The Models:

Testing the effects of market size and the other variables of interest, while controlling for player performance in the form of PER and box score statistics, requires nine regression models. Based on the two-part model approach and control blocks outlined above, ‘Part 1’ consists of six models, three probit regressions and three logit regressions, and ‘Part 2’ consists of three models. For the ‘Part 1’ regressions, the dependent variable, ‘receivedVotes’, is a binary variable with the value ‘1’ if the player received any All-NBA votes and ‘0’ otherwise. In contrast, the dependent variable for the ‘Part 2’ regressions, All-NBA’, is a continuous variable equal to the number of All-NBA points the player received. ‘Market’ is the variable representing the Nielsen television market size corresponding to the player’s team. ‘East’ is the indicator for East Coast Bias, and it takes the value of ‘1’ if the player’s team is based in the Eastern Time Zone or ‘0’ otherwise.

'Foreign' is the indicator for bias against foreign-born players; the variable is set equal to '0' if the player was born in the United States and '1' if the player was born outside the United States. 'TeamWL%' is the variable for the player's team's win-loss percentage. The vector between the variable for Team Win-Loss Percentage and the error term (Ɛ) represents the particular control block used in the model and is defined below the equation when applicable. There are only six unique equations due to the equations for the probit and logit models being identical in appearance. The difference between the probit and logit models is the assumption made on the distribution of the dependent variable. Probit regressions are run on a cumulative distribution function assuming a standard normal distribution, and Logit regressions are run on a cumulative distribution function assuming a logistic distribution. The equation for each regression model is provided in Table 1 on the next page.

Table 1: The Model Equations

Part 1 - Model #1: Variables of Interest and PER
$receivedVotes = f(\beta_0 + \beta_1 Market + \beta_2 Foreign + \beta_3 East + \beta_4 TeamWL\% + B_5 PER + \mathcal{E})$
Part 1 - Model #2: Variables of Interest, PER, and its 12 Box Score Statistic Components
$receivedVotes = f(\beta_0 + \beta_1 Market + \beta_2 Foreign + \beta_3 East + \beta_4 TeamWL\% + B_5 X_5 + \ldots + B_{17} X_{17} + \mathcal{E})$ where $B_5 X_5 + B_{17} X_{17} = B_5 PER + B_6 FG + B_7 FT + B_8 MissedFG + B_9 MissedFT + B_{10} 3PM + B_{11} PF + B_{12} ORB + B_{13} DRB + B_{14} AST + B_{15} STL + B_{16} TOV + B_{17} BLK$
Part 1 - Model #3: Variables of Interest and the 12 Box Score Statistics
$receivedVotes = f(\beta_0 + \beta_1 Market + \beta_2 Foreign + \beta_3 East + \beta_4 TeamWL\% + B_5 X_5 + \ldots + B_{16} X_{16} + \mathcal{E})$ where $B_5 X_5 + B_{16} X_{16} = B_5 FG + B_6 FT + B_7 MissedFG + B_8 MissedFT + B_9 3PM + B_{10} PF + B_{11} ORB + B_{12} DRB + B_{13} AST + B_{14} STL + B_{15} TOV + B_{16} BLK$
Part 2 - Model #1: Variables of Interest and PER
$All\text{-}NBA = \beta_0 + \beta_1 Market + \beta_2 Foreign + \beta_3 East + \beta_4 TeamWL\% + B_5 PER + \mathcal{E}$
Part 2 - Model #2: Variables of Interest, PER, and its 12 Box Score Statistic Components
$All\text{-}NBA = \beta_0 + \beta_1 Market + \beta_2 Foreign + \beta_3 East + \beta_4 TeamWL\% + B_5 X_5 + \ldots + B_{17} X_{17} + \mathcal{E}$ where $B_5 X_5 + B_{17} X_{17} = B_5 PER + B_6 FG + B_7 FT + B_8 MissedFG + B_9 MissedFT + B_{10} 3PM + B_{11} PF + B_{12} ORB + B_{13} DRB + B_{14} AST + B_{15} STL + B_{16} TOV + B_{17} BLK$
Part 2 - Model #3: Variables of Interest and the 12 Box Score Statistics
$All\text{-}NBA = \beta_0 + \beta_1 Market + \beta_2 Foreign + \beta_3 East + \beta_4 TeamWL\% + B_5 X_5 + \ldots + B_{16} X_{16} + \mathcal{E}$ where $B_5 X_5 + B_{16} X_{16} = B_5 FG + B_6 FT + B_7 MissedFG + B_8 MissedFT + B_9 3PM + B_{10} PF + B_{11} ORB + B_{12} DRB + B_{13} AST + B_{14} STL + B_{15} TOV + B_{16} BLK$

<u>Expectations Before Running Regressions:</u>

From the regression models above, I expect to observe that a player's market size has a positive effect on the number of All-NBA votes they receive. A positive coefficient on market size, statistically significant at the 95% level, would provide evidence that my hypothesis is correct. If bias exists against foreign-born players, then I would expect foreign-born players to receive less All-NBA votes on average than American-born players with comparable season statistics. The coefficient on the Foreign indicator should be negative and significant at the 95% level. If bias exists toward players on East Coast teams, then I would expect players on East Coast teams to receive more All-NBA votes on average than players not on East Coast teams. The coefficient on the East indicator should be positive and significant at the 95% level. Lastly, I expect Team Win-Loss Percentage to have a positive effect on the number of All-NBA votes a player receives. In other words, players on teams with good records should receive more votes than players on teams with bad records. This expectation would be supported by a positive coefficient on Team Win-Loss Percentage significant at the 95% level.

4. Data

<u>Creating the Dataset:</u>

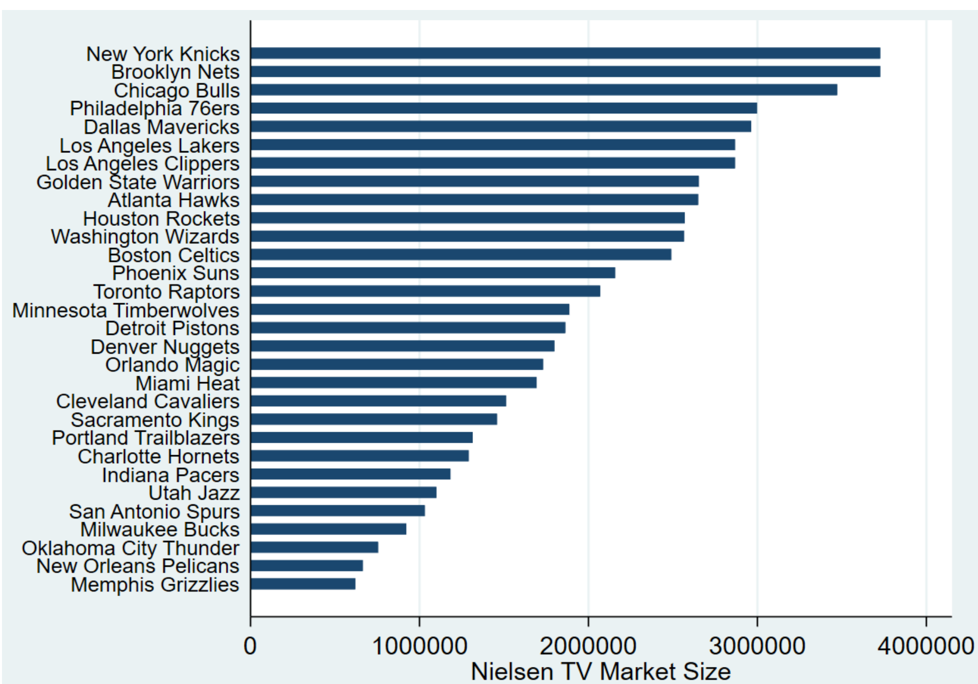
My dataset consists of NBA player data from five consecutive seasons (2014-2015 through 2018-2019). I did not include the two most recent NBA seasons (2019-2020 and 2020-2021) because both were shortened due to the coronavirus pandemic. For each season in the dataset, I select the top 175 players in minutes played, resulting in a total of 875 player seasons. The NBA consists of thirty teams of fifteen players, and in general the 'starters' for each team are the five players that play the most minutes. Accordingly, the top 175 players in minutes played should represent all the ‘starters’ plus the best ‘bench’ players in the league. While the same player could end up in the dataset for multiple years, double counting is not an issue because a player will have different on-court statistics and likely received a different number of All-NBA votes each season.

From the website *Basketball-Reference*, I obtain the following information for each player season in my dataset: Team, Country of Birth, Team Record, Player Efficiency Rating, Field Goals Made, Free Throws Made, Missed Field Goals, Missed Free Throws, 3-Pointers Made, Offensive Rebounds, Defensive Rebounds, Assists, Steals, Turnovers, Blocked Shots, and Personal Fouls. I am confident in the accuracy of this data because *Basketball-Reference* has been cited in several Sports Economics papers written on NBA-related topics. Summary statistics for each of the twelve box score statistics and Player Efficiency Rating are provided in the Appendix.

I access both All-NBA voting results and All-Star teams by season through *NBA.com* and *NBA Communications*, official websites of the NBA. The All-NBA voting results for the 2014-2015 season were missing from both websites, so I obtain the ballots by accessing an archived image of *NBA.com* through the website *WayBack Machine - Internet Archive*. This website was created by The Internet Archive, a non-profit organization whose purpose is to maintain a digital library of Internet sites accessible for free to the general public (Internet Archive, 2021).

For my measure of market size, I obtain Nielsen television market sizes corresponding to every NBA team located in the United States (29 out of 30 teams) from *Sports Media Watch*. This website is a credible source because it has been cited in several Sports Economics research papers on market-related topics. Nielsen television market size is an estimate of the number of homes in a specific viewing area (Sports Media Watch, 2021). The only team without a Nielsen television market size was the Toronto Raptors, so I impute the market size for this team by assuming a linear relationship exists between Nielsen television market size and metro population. Using metro population data available through *Hoop Social*, I regress Nielsen television market size on metro population excluding Toronto. Then, I obtain an estimate for the Toronto market size by inputting its metro population into the model. I am confident in the accuracy of the metro population data obtained from *Hoop Social* because this website references the United States Census Bureau as the original author of the data. Figure 1 displays the market sizes, including the estimate for Toronto, of all thirty NBA teams. Since the New York and Los Angeles markets have two NBA teams each, I divide their market sizes by two to avoid double counting the number of households. The four teams with a market size affected by this choice were the New York Knicks, Brooklyn Nets, Los Angeles Lakers, and Los Angeles Clippers.

Figure 1: Nielsen Television Market Size by NBA Team



Summary Statistics:

Summary statistics for both the whole dataset, consisting of 875 player seasons, and the partial dataset, consisting of 176 player seasons, are presented in Tables 1 and 2 below. Outcome frequencies for each of the indicator variables are provided in parentheses next to the count of the outcome. Interestingly, by removing players that did not receive any All-NBA votes, the average Nielsen market size decreases from 2,025,539 to 1,948,980, the proportion of players born outside the U.S. decreases from 23.3% to 21.0%, and the proportion of players from Eastern Time Zone teams decrease and 37.8% to 32.4%; however, both the average Team Win-Loss Percentage and average PER increase from 0.509 to 0.602 and 16.03 to 22.26 respectively. Summary statistics for the twelve box score statistics are provided under the Appendix as Tables 10 and 11.

Table 2: Whole Dataset Summary Statistics (875 player seasons)

Variable Name	0	1
Received All-NBA Votes ('1' = Received at least 1 vote)	699 (0.80)	176 (0.20)
Foreign Indicator ('1' = Born outside the U.S.)	671 (0.77)	204 (0.23)
East Indicator ('1' = Eastern Time Zone team)	544 (0.62)	331 (0.38)

Variable Name	Mean	SD	Min.	Max.
Nielsen Television Market Size	2,025,539	881,779	620,000	3,726,500
Team Win-Loss Percentage	0.509	0.147	0.122	0.890
Player Efficiency Rating (PER)	16.03	4.63	5.7	31.5

Table 3: Partial Dataset Summary Statistics (176 player seasons)

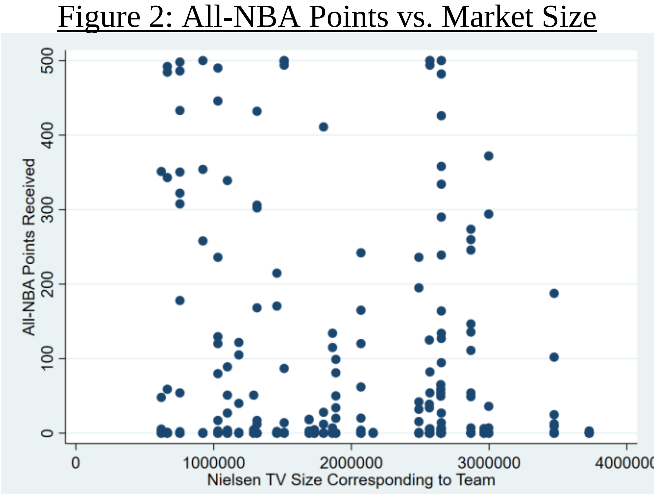
Variable Name	0	1
Foreign Indicator ('1' =Born outside the U.S.)	139 (0.79)	37 (0.21)
East Indicator ('1' = Eastern Time Zone team)	119 (0.68)	57 (0.32)

Variable Name	Mean	SD	Min.	Max.
All-NBA Points	127.80	166.33	0.78	500
Nielsen Television Market Size	1,948,980	856,900	620,000	3,726,500
Team Win-Loss Percentage	0.602	0.125	0.232	0.890
Player Efficiency Rating (PER)	22.26	3.84	11.8	31.5

Examining the Distribution of All-NBA Points versus the Continuous Variables:

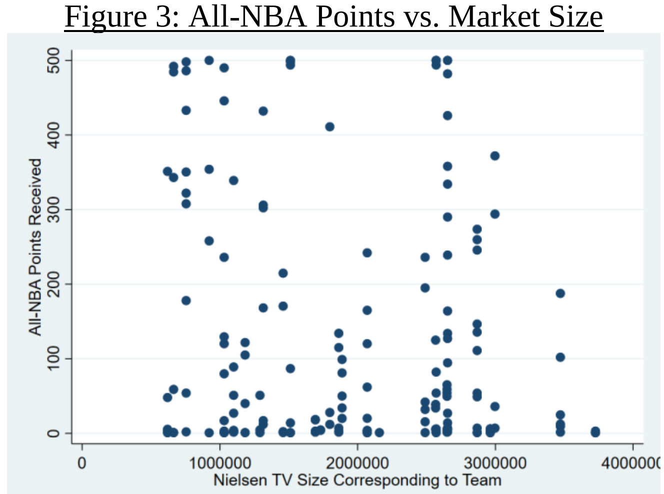
Figures 2 and 3 present the distributions of All-NBA points versus market size for both the whole and partial datasets. Even after removing player seasons with no All-NBA votes, there does not appear to be a clear positive relationship between All-NBA points and market size.

Figure 2: All-NBA Points vs. Market Size



Whole Dataset (875 player seasons)

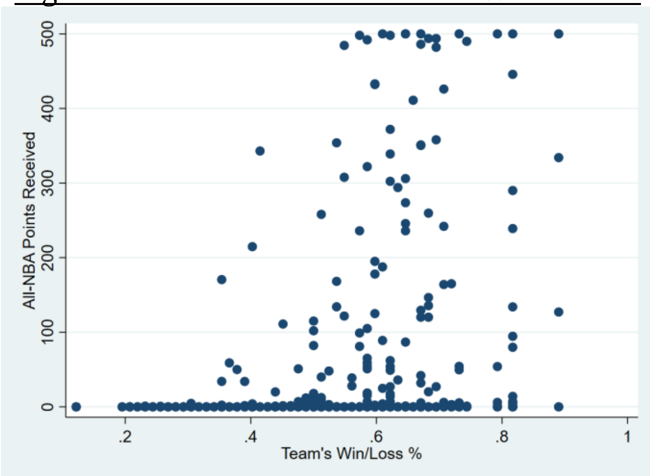
Figure 3: All-NBA Points vs. Market Size



Partial Dataset (176 player seasons)

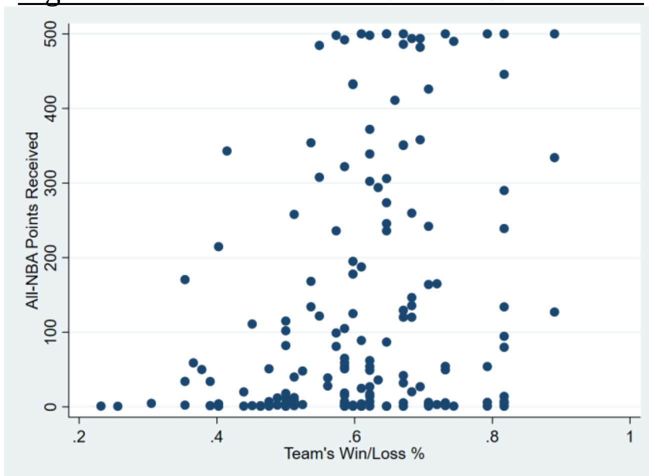
In contrast, a positive relationship can be observed in the distributions of All-NBA points versus Team Win-Loss Percentage for both the whole and partial datasets. The distributions are provided as Figures 4 and 5 below.

Figure 4: All-NBA Points vs. Team Win-Loss %



Whole Dataset (875 player seasons)

Figure 5: All-NBA Points vs. Team Win-Loss %



Partial Dataset (176 player seasons)

5. Results

I construct the nine regression models using the software package Stata. Standard errors are robust to account for heteroskedasticity introduced by the different distribution of each factor in the dataset. Coefficients statistically significant at the 95%, 99%, or 99.9% confidence levels are denoted using 1-star, 2-stars, and 3-stars, respectively. To make it easier to interpret the coefficients on Market size and Team Win-Loss Percentage, Market size is scaled to hundreds of thousands, and Team Win-Loss Percentage is converted from a decimal to a percent.

<u>Part 1 Regressions - Probit Models, Whole Dataset:</u>

Marginal effects for the three probit models are provided in Table 4 below, and the probit coefficients are provided as Table 8 in the Appendix. Market size is significant at the 95% level only in the model controlling for PER and the twelve box score statistics. Holding all factors in the model constant, an increase in market size of 100,000 households is associated with a 0.2% (0.002) increase in the average probability of a player receiving at least one All-NBA vote with a standard error less than 0.01. However, this effect should not be considered significant in practical terms because the increase in average probability is extremely small.

Similar to Market size, the Foreign indicator is significant at the 95% level in only one of the three models. For the model with PER as the only control, holding all factors constant, the average probability of receiving at least one All-NBA vote is 4.1% (-0.041) less for players born outside the United States than players born in the United States. While the effect of the indicator is not statistically significant in the other two models, the effect is negative in all three models which is consistent with my expectation of bias against foreign-born players.

Although the East indicator is not significant in any of the three models, Team Win-Loss Percentage is significant in all three models at the 99.9% level with the same marginal effect. Holding all factors constant, an increase of 1% in Team Win-Loss Percentage is associated with around a 0.5% (0.005) increase in the average probability of a player receiving at least one All-NBA vote with a standard error less than 0.01. While the magnitude of this effect is extremely small making it practically insignificant, the high confidence level indicates Team Win-Loss Percentage likely has a positive effect on All-NBA votes.

Table 4: Part 1 - Probit Marginal Effects, Whole Dataset

VARIABLES	(1) Only PER	(2) PER + Stats	(3) Only Stats
Nielsen TV Market Size (in hundred thousands)	0.002	0.002*	0.002
	(0.00)	(0.00)	(0.00)
Foreign Indicator = 1, Foreign	-0.041*	-0.014	-0.018
	(0.02)	(0.01)	(0.02)
East Indicator = 1, East Coast	-0.018	-0.023	-0.023
	(0.02)	(0.02)	(0.02)
Team Win-Loss Percentage (as %)	0.005***	0.005***	0.005***
	(0.00)	(0.00)	(0.00)
Player Efficiency Rating, PER	0.040***	0.017***	
	(0.01)	(0.00)	
Field Goals Made, FG		0.000	0.001**
		(0.00)	(0.00)
Free Throws Made, FT		0.000	0.001***
		(0.00)	(0.00)
Missed Field Goals		0.000	-0.000
		(0.00)	(0.00)
Missed Free Throws		0.000	-0.000
		(0.00)	(0.00)
3-Pointers Made, 3-PM		-0.000	0.000
		(0.00)	(0.00)
Personal Fouls, PF		-0.000	-0.001**
		(0.00)	(0.00)
Offensive Rebounds, ORB		-0.000	0.000
		(0.00)	(0.00)
Defensive Rebounds, DRB		0.000*	0.000**
		(0.00)	(0.00)
Assists, AST		0.000	0.000*
		(0.00)	(0.00)
Steals, STL		0.000	0.000
		(0.00)	(0.00)
Turnovers, TOV		0.000	-0.000
		(0.00)	(0.00)
Blocked Shots, BLK		0.001**	0.001***
		(0.00)	(0.00)
Observations	800	800	800

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05

The marginal effect of PER is significant at the 99.9% level both with and without the twelve box score statistics. When PER is the only control, an increase in PER of one unit is associated with an increase of 4.0% (0.040) in the average probability of a player receiving at least one All-NBA vote with a standard error of 0.01. With the addition of the twelve box score statistics, an increase in PER of one unit is associated with an increase of 1.7% (0.017) in the average probability of a player receiving at least on All-NBA vote with a standard error less than 0.01. The magnitude and significance of these marginal effects provide evidence that PER adds explanatory value to the models even with the inclusion of the twelve box score statistics.

Based on the results of the Wald tests, the twelve box score statistics have explanatory power in the models where they act as controls. For the model including PER and the twelve box score statistics, the Wald test returned a chi-squared of 57.89 with a p-value less than 0.0001. The extremely low p-value indicates that it is highly unlikely for the coefficients on all twelve box score statistics to be equal to zero. As a result, the box score statistics contribute explanatory value to the model also controlling for PER. For the model controlling with only the twelve box score statistics, the Wald test returned a chi-squared of 180.96 with a p-value less than 0.0001. The extremely low p-value indicates that it is highly unlikely for the coefficients on all twelve box score statistics to be equal to zero. Therefore, the twelve box score statistics as a control block provide explanatory value even in the absence of PER.

Part 1 Regressions - Logit Models, Whole Dataset:

Marginal effects for the three logit models are provided in Table 5 below, and the logit coefficients are provided in Table 9 under the Appendix. Overall, the marginal effects produced by the logit and probit models are very similar. For example, the marginal effect of Market size is the same across all six models. Holding all factors constant, an increase in market size of 100,000 households is associated with a 0.2% (0.002) increase in the average probability that a player received at least one All-NBA vote with a standard error less than 0.01. The magnitude of this effect is extremely small, so it should not be considered significant in practical terms. Market size is significant at the 95% level in all three logit models, compared with only two of three probit models, and the increase in confidence is likely attributable to a smaller standard error.

While the Foreign indicator is significant at the 95% level in both probit and logit models controlling for only PER, the magnitude of the effect is less in the logit model than the probit model. Holding all factors constant, a player born outside the United States is on average 3.2% less likely to receive at least one All-NBA vote than a player born in the United States with the same PER. The effect of the Foreign indicator is negative in the other two logit models which is consistent with my hypothesis that bias exists against foreign-born players.

For both the East indicator and Team Win-Loss Percentage, the results from the logit regressions are consistent with the results from the probit regressions. The East indicator is not significant in any of the three models, and Team Win-Loss Percentage is significant in all three at the 99.9% level. Holding all factors constant, an increase in Team Win-Loss Percentage of 1% is associated with a 0.4% in the average probability that the player received at least one All-NBA vote with a standard error less than 0.01. Since Team Win-Loss Percentage is significant at such a high confidence level in both the logit and probit models, it is likely that a relationship exists between Team Win-Loss Percentage and All-NBA votes.

Table 5: Part 1 - Logit Marginal Effects, Whole Dataset

VARIABLES	(1) Only PER	(2) PER + Stats	(3) Only Stats
Nielsen TV Market Size (in hundred thousands)	0.002*	0.002*	0.002*
	(0.00)	(0.00)	(0.00)
Foreign Indicator = 1, Foreign	-0.032*	-0.010	-0.011
	(0.01)	(0.01)	(0.01)
East Indicator = 1, East Coast	-0.014	-0.017	-0.018
	(0.02)	(0.01)	(0.01)
Team Win-Loss Percentage (as %)	0.004***	0.004***	0.004***
	(0.00)	(0.00)	(0.00)
Player Efficiency Rating, PER	0.034***	0.013***	
	(0.00)	(0.00)	
Field Goals Made, FG		0.000	0.000*
		(0.00)	(0.00)
Free Throws Made, FT		0.000	0.001***
		(0.00)	(0.00)
Missed Field Goals		0.000	-0.000
		(0.00)	(0.00)
Missed Free Throws		0.000	-0.000
		(0.00)	(0.00)
3-Pointers Made, 3-PM		-0.000	-0.000
		(0.00)	(0.00)
Personal Fouls, PF		-0.000	-0.001**
		(0.00)	(0.00)
Offensive Rebounds, ORB		-0.000	0.000
		(0.00)	(0.00)
Defensive Rebounds, DRB		0.000*	0.000**
		(0.00)	(0.00)
Assists, AST		0.000	0.000*
		(0.00)	(0.00)
Steals, STL		0.000	0.000
		(0.00)	(0.00)
Turnovers, TOV		0.000	-0.000
		(0.00)	(0.00)
Blocked Shots, BLK		0.001**	0.001***
		(0.00)	(0.00)
Observations	800	800	800

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Although the magnitude of the effect is different between the probit and logit models, PER is significant at the 99.9% level in both models where it is a control. When PER is the only control, an increase in PER of 1 unit is associated with an increase of 3.4% (0.034) in the average probability that a player receives at least one All-NBA vote with a standard error of less than 0.01. With the addition of the twelve box score statistics, the associated increase in average probability is 1.3% (0.013) with a standard error also less than 0.01. Therefore, PER adds explanatory value to the logit models even with the inclusion of the twelve box score statistics.

The results of the Wald tests run on the logit regressions support the results of the Wald tests run on the probit regressions. For the model controlling with PER and the twelve box score statistics, the Wald test returned a chi-squared of 54.26 with a p-value less than 0.0001. For the model controlling for only the twelve box score statistics, the Wald test returned a chi-squared of 129.61 with a p-value also less than 0.001. The extremely low p-values suggest that it is unlikely for the coefficients on all the twelve box score statistics to be zero in either of the two models. Therefore, the box score statistics provide explanatory value to the models regardless of whether PER is included.

Part 2 Regressions - OLS Models, Partial Dataset:

The output of the OLS regressions run on the dataset consisting of only players that received at least one All-NBA vote is presented as Table 6 below. The models are run using 174 instead of 176 players because two player seasons in the partial dataset do not have a corresponding market size. This is possible if a player was traded during the season and consequently played for more than one team that season. Note that the mean number of All-NBA points, conditional on points being greater than zero, is 128, so the effects below are reasonable.

Compared with the probit and logit regressions, the effect of Market size on All-NBA points is notably not statistically significant at the 95% level in any of the three models. In fact, the coefficient on Market size is negative in all three models which contradicts my hypothesis that market size has a positive effect on All-NBA points. Instead, these results suggest that expected All-NBA points actually decrease as market size increases.

While the effect of the Foreign indicator is not significant in any of the models, the effect of the East indicator is significant at the 95% level in two of the three models. For the model controlling with PER and the twelve box score statistics, holding all other factors constant, players on East Coast teams have an average of 31.328 All-NBA points less than players not on East Coast teams with a standard error of 14.31. For the model controlling with only the twelve box score statistics, the difference increases to an average of 33.273 All-NBA points with a standard error of 15.12. Since the sign of these coefficients is negative, the results of the OLS regressions contradict my expectations by suggesting that East Coast bias could actually be against rather than in favor of players on East Coast teams.

Across all three models, the effect of Team Win-Loss Percentage on All-NBA points is significant at the 99.9% level. Holding all other factors constant, an increase in Team Win-Loss Percentage of 1% is associated with an increase of between 4.425 to 4.632 All-NBA points on average. Equivalently, an increase in Team Win-Loss Percentage of 10% is associated with an increase of about 45 All-NBA points on average. This increase could be the difference between a player making the All-NBA Third Team and being left off the All-NBA teams entirely. Therefore, the effect of Team Win-Loss Percentage on All-NBA points should be considered significant in practical terms.

Table 6: Part 2 - OLS Models, Partial Dataset

VARIABLES	(1) Only PER	(2) PER + Stats	(3) Only Stats
Nielsen TV Market Size (in hundred thousands)	-1.536	-1.213	-1.518
	(0.92)	(0.87)	(0.94)
Foreign Indicator	-21.023	6.273	5.998
	(19.69)	(18.40)	(19.22)
East Indicator	-18.432	-31.328*	-33.273*
	(17.06)	(14.31)	(15.12)
Team Win-Loss Percentage (as %)	4.456***	4.632***	4.425***
	(0.61)	(0.69)	(0.69)
Player Efficiency Rating, PER	29.717***	17.269***	
	(2.29)	(4.32)	
Field Goals Made, FG		0.183	0.587***
		(0.14)	(0.12)
Free Throws Made, FT		0.097	0.425***
		(0.12)	(0.09)
Missed Field Goals		-0.016	-0.352***
		(0.12)	(0.10)
Missed Free Throws		0.372*	0.066
		(0.17)	(0.18)
3-Pointers Made, 3-PM		0.179	0.317
		(0.15)	(0.17)
Personal Fouls, PF		-0.351	-0.788***
		(0.21)	(0.23)
Offensive Rebounds, ORB		-0.256	-0.034
		(0.16)	(0.17)
Defensive Rebounds, DRB		0.093	0.156*
		(0.08)	(0.08)
Assists, AST		0.016	0.113
		(0.08)	(0.09)
Steals, STL		0.177	0.498
		(0.27)	(0.27)
Turnovers, TOV		0.372	0.286
		(0.26)	(0.28)
Blocked Shots, BLK		0.595*	0.696**
		(0.25)	(0.24)
Constant	-760.437***	-750.278***	-491.336***
	(63.10)	(82.26)	(59.88)
Observations	174	174	174
R-squared	0.605	0.717	0.683

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Just like in the probit and logit regressions, the coefficient on PER in both OLS models supports the result that PER adds explanatory value even with the addition of the twelve box score statistics. When PER is the only control, holding all other factors constant, an increase in one unit of PER is associated with an increase in average All-NBA points of 29.717 with a standard error of 2.29. When controlling for the twelve box score statistics in addition to PER, holding all factors constant, an increase in one unit of PER is associated with an increase in average All-NBA points of 17.269 with a standard error of 4.32. Since the effect is significant at the 99.9% level in both models, PER adds explanatory value over the twelve box score statistics.

The results of the Wald tests run on the OLS regressions support the results of the other Wald tests providing evidence that the statistics add explanatory power to the models. The Wald test on the model controlling for PER and the twelve box score statistics returned a F-statistic of 6.69, and the Wald test on the model controlling for only the statistics returned a F-statistic of 26.89. Both statistics correspond to a p-value of less than 0.0001, indicating that it is highly unlikely for the coefficients on all the box score statistics to be equal to zero.

Analysis of Regression Controls - PER and the Twelve Box Score Statistics:

To examine the relationship between PER and its components, I perform an OLS regression of PER on the twelve box score statistics for both the whole and partial datasets. The output from each regression is presented as Table 7 below. The regression R-squared decreases from 0.881 for the whole dataset to 0.774 for the partial dataset which is expected because the distribution is being truncated. The coefficients in this regression are not meaningful because the box score statistics are assumed to be correlated with each other; however, it is important to recognize that the box score statistics are significant as a group in their effect on PER.

Table 7: OLS - PER on Its 12 Box Score Statistic Components

VARIABLES	(1) Whole Dataset	(2) Partial Dataset
Field Goals Made, FG	0.027***	0.023***
	(0.00)	(0.00)
Free Throws Made, FT	0.022***	0.019***
	(0.00)	(0.00)
Missed Field Goals	-0.019***	-0.018***
	(0.00)	(0.00)
Missed Free Throws	-0.015***	-0.017***
	(0.00)	(0.00)
3-Pointers Made, 3-PM	0.008***	0.007**
	(0.00)	(0.00)
Personal Fouls, PF	-0.021***	-0.025***
	(0.00)	(0.00)
Offensive Rebounds, ORB	0.018***	0.014***
	(0.00)	(0.00)
Defensive Rebounds, DRB	0.002**	0.003
	(0.00)	(0.00)
Assists, AST	0.009***	0.005**
	(0.00)	(0.00)
Steals, STL	0.007**	0.019***
	(0.00)	(0.00)
Turnovers, TOV	-0.012***	-0.004
	(0.00)	(0.01)
Blocked Shots, BLK	0.014***	0.007
	(0.00)	(0.01)
Constant	9.858***	13.494***
	(0.27)	(0.91)
Observations	875	176
R-squared	0.881	0.774

Robust standard errors in parentheses

*** $p<0.001$, ** $p<0.01$, * $p<0.05$

After regressing PER on the twelve box score statistics, I construct a correlation matrix for both the whole and partial datasets to illustrate how the box score statistics are correlated with each other. For example, Field Goals are highly correlated with Missed Field Goals, Free Throws are highly correlated with Missed Free Throws, and Turnovers are highly correlated with Free Throws. Since several of the box score statistics are correlated with each other, it is likely that multicollinearity is introduced in the regression of PER on the twelve box score statistics. Therefore, it is unnecessary to examine the coefficients on each of the twelve box score statistics in these regressions. The correlation matrices for the whole and partial datasets are provided on the following page as Figures 6 and 7.

Figure 6: Correlation Matrix of PER and the Twelve Box Score Statistics - Whole Dataset

Variables	(PER)	(FG)	(FT)	(Missed FG)	(Missed FT)	(3-PM)	(PF)	(ORB)	(DRB)	(AST)	(STL)	(TOV)	(BLK)
PER	1.00												
FG	0.77	1.00											
FT	0.76	0.78	1.00										
Missed FG	0.48	0.86	0.72	1.00									
Missed FT	0.53	0.44	0.53	0.20	1.00								
3-PM	0.02	0.31	0.22	0.56	-0.30	1.00							
PF	0.17	0.21	0.18	0.05	0.42	-0.18	1.00						
ORB	0.45	0.21	0.15	-0.16	0.61	-0.55	0.50	1.00					
DRB	0.58	0.43	0.38	0.13	0.64	-0.27	0.51	0.75	1.00				
AST	0.41	0.45	0.49	0.53	0.16	0.26	0.02	-0.23	0.03	1.00			
STL	0.30	0.36	0.36	0.42	0.23	0.23	0.22	-0.05	0.13	0.56	1.00		
TOV	0.58	0.70	0.73	0.68	0.47	0.19	0.30	0.08	0.35	0.79	0.51	1.00	
BLK	0.42	0.22	0.19	-0.09	0.51	-0.39	0.46	0.67	0.68	-0.18	-0.03	0.10	1.00

Figure 7: Correlation Matrix of PER and the Twelve Box Score Statistics - Partial Dataset

Variables	(PER)	(FG)	(FT)	(Missed FG)	(Missed FT)	(3-PM)	(PF)	(ORB)	(DRB)	(AST)	(STL)	(TOV)	(BLK)
PER	1.00												
FG	0.61	1.00											
FT	0.63	0.64	1.00										
Missed FG	0.32	0.81	0.67	1.00									
Missed FT	0.20	0.09	0.23	-0.15	1.00								
3-PM	0.11	0.40	0.30	0.63	-0.39	1.00							
PF	0.05	0.10	0.07	-0.06	0.46	-0.17	1.00						
ORB	0.14	-0.08	-0.16	-0.44	0.60	-0.62	0.57	1.00					
DRB	0.29	0.11	0.06	-0.23	0.62	-0.42	0.61	0.76	1.00				
AST	0.29	0.37	0.41	0.53	-0.07	0.42	-0.02	-0.45	-0.21	1.00			
STL	0.28	0.32	0.35	0.45	0.08	0.37	0.16	-0.21	-0.10	0.59	1.00		
TOV	0.42	0.58	0.64	0.63	0.24	0.37	0.30	-0.16	0.13	0.78	0.54	1.00	
BLK	0.15	-0.10	-0.14	-0.45	0.43	-0.54	0.41	0.67	0.68	-0.47	-0.28	-0.21	1.00

6. Conclusion

This study investigates whether market size or a factor other than player performance explains how media members select players to the All-NBA teams. Using the results from the nine regression models, I analyze the effect of each potential factor on both the likelihood of a player receiving at least one All-NBA vote and the distribution of All-NBA points consisting of players that received at least one All-NBA vote.

I am unable to conclude whether market size has a positive effect on a player's All-NBA votes. The marginal effect of market size on the likelihood of a player receiving at least one All-NBA vote is the same in all probit and logit regressions, but the magnitude is very small and not significant in two of the three probit models. In addition, the effect of market size on All-NBA points for players that received at least one All-NBA vote is negative and not significant at the 95% level. This result contradicts my hypothesis as I would expect the effect to be positive.

I am also unable to conclude whether media members are biased against players born outside the United States or towards players on East Coast teams. While the marginal effect of the Foreign indicator on the likelihood of a player receiving at least one All-NBA vote is negative across all probit and logit models, the effect varies in magnitude and is not significant in the majority of the models. The negative sign is consistent with my expectation of bias against foreign-born players, but the variation in significance levels makes the results inconclusive. The East indicator is not significant in any of the probit or logit regressions, and the sign of the effect on All-NBA points is negative in two of the three models. I would recommend that further research consider East Coast bias occurs against players on East Coast teams as these results strongly refute my hypothesis of bias towards players on East Coast teams.

The most notable finding from my study is that Team Win-Loss Percentage has an effect on All-NBA votes. The marginal effect of Team Win-Loss Percentage on the likelihood of a player receiving at least one All-NBA vote is positive and significant at the 99.9% level across all probit and logit models. Also, the effect of Team Win-Loss Percentage on the distribution of All-NBA points is both positive and significant at the 99.9% level in all three OLS regressions. Holding all other factors constant, average All-NBA Votes received tends to increase with Team Win-Loss Percentage. A possible explanation for this result is that players on winning teams receive more media attention than players on losing teams, implying that media coverage is related with how well a team performs during a particular season.

This study contributes to the established literature on the All-NBA selection process by providing evidence that Team Win-Loss Percentage is the factor introducing bias to the All-NBA selection process. Although I would have liked to include more factors in my models, such as player age and race, I hope my results can lead towards a future study that uncovers how bias is introduced in the All-NBA selection process.

References

2017 NBA-NBPA Collective Bargaining Agreement (Effective as of 7/1/2017). (2017, January 19). Retrieved October 17, 2021, from https://nbpa.com/cba

Bednalll, J. (2019, May 24). Millions gained and lost as All-NBA Teams announced. Retrieved November 13, 2021, from https://www.news.com.au/sport/american-sports/nba/millions-gained-and-lost-as-allnba-teams-announced/news-story/97dde85a236f79d63771591292905b13

Burns, M. (2021, July 2). *NBA Team Market Size Rankings 2021*. HoopSocial. https://hoop-social.com/nba-team-market-size-rankings/

Fernandes, A. (2017, May 20). All-NBA Teams, the who, what, when, why and how. Retrieved November 13, 2021, from www.sportskeeda.com/basketball/all-nba-teams-the-who-what-when-why-and-how

Freedman, D. (2019, April). CBA Breakdown - Compensation. Retrieved October 17, 2021, from https://cbabreakdown.com/compensation

Hollinger, J. (2005). *Pro basketball forecast, 2005-06*. Potomac Books. 2005

Johnson, C., & Minuci, E. (2020). Wage discrimination in the NBA: Evidence using free agent signings. *Southern Economic Journal*, *87*(2), 517-539. Retrieved from https://www.proquest.com/scholarly-journals/wage-discrimination-nba-evidence-using-free-agent/docview/2500885538/se-2?accountid=14522

Lee, Sang H., Leonard, John., & Jeon, Hyungjoon (2009). Sports market capacity, frequency of movements, and age in the NBA players salary determination. *Indian Journal of Economics & Business*, *8*(1), 127-134.

Levine, G. R. (2019). *All-NBA Team Voting Patterns: Using Classification Models To Identify How And Why Players Are Nominated* (Doctoral dissertation, Ohio University).

Na, A. (2021). *Transparency Is Needed in Voting for All-NBA Guards*.

NBA Media Ventures, LLC. (2021, June 15). *Nuggets' Nikola Jokić, Bucks' Giannis Antetokounmpo and Warriors' Stephen Curry lead 2020-21 All-NBA First team*. NBA Communications. https://pr.nba.com/2020-21-all-nba-team

NBA Media Ventures, LLC. (2021). *Year-by-year All-NBA Teams*. NBA History. https://www.nba.com/news/history-all-nba-teams

NBA Media Ventures, LLC (2015, May 21). *Warriors' Curry and Cavaliers' James unanimous picks for 2014-15 All-NBA First Team*. WayBackMachine. https://web.archive.org/web/20150524011324/http://www.nba.com/2015/news/05/21/2014-15-all-nba-first-team-official-release/index.html?ls=iref:nbahpts

Randrianasolo A.A. (2021) Hey Now, You are an All-Star…But are you All-NBA?. In: Triple Double. SpringerBriefs in Statistics. Springer, Cham. https://doi.org/10.1007/978-3-030-79032-5_4

Sports Media Watch (2021). *Major pro sports teams ranked by market size*. https://www.sportsmediawatch.com/nba-market-size-nfl-mlb-nhl-nielsen-ratings/

Sports Reference LLC. (2000-2021). *Basketball Statistics and History*. Basketball-Reference.com. https://www.basketball-reference.com/

The Internet Archive (n.d.). *About the Internet Archive*. https://archive.org/about/

7. Appendix

Table 8: Part 1 - Probit Coefficients, Whole Dataset

VARIABLES	(1) Only PER	(2) PER + Stats	(3) Only Stats
Nielsen TV Market Size (in hundred thousands)	0.018	0.023*	0.019
	(0.01)	(0.01)	(0.01)
Foreign Indicator	-0.428*	-0.184	-0.203
	(0.19)	(0.22)	(0.19)
East Indicator	-0.166	-0.288	-0.243
	(0.16)	(0.19)	(0.18)
Team Win-Loss Percentage (as %)	0.045***	0.055***	0.056***
	(0.01)	(0.01)	(0.01)
Player Efficiency Rating, PER	0.356***	0.206***	
	(0.03)	(0.05)	
Field Goals Made, FG		0.000	0.005**
		(0.00)	(0.00)
Free Throws Made, FT		0.002	0.007***
		(0.00)	(0.00)
Missed Field Goals		0.003	-0.001
		(0.00)	(0.00)
Missed Free Throws		0.001	-0.002
		(0.00)	(0.00)
3-Pointers Made, 3-PM		-0.001	0.000
		(0.00)	(0.00)
Personal Fouls, PF		-0.006	-0.009***
		(0.00)	(0.00)
Offensive Rebounds, ORB		-0.002	0.001
		(0.00)	(0.00)
Defensive Rebounds, DRB		0.003**	0.004***
		(0.00)	(0.00)
Assists, AST		0.001	0.003*
		(0.00)	(0.00)
Steals, STL		0.003	0.004
		(0.00)	(0.00)
Turnovers, TOV		0.002	-0.000
		(0.00)	(0.00)
Blocked Shots, BLK		0.009**	0.011***
		(0.00)	(0.00)
Constant	-9.829***	-10.860***	-8.460***
	(0.74)	(1.12)	(0.85)
Observations	800	800	800

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table 9: Part 1 - Logit Coefficients, Whole Dataset

VARIABLES	(1) Only PER	(2) PER + Stats	(3) Only Stats
Nielsen TV Market Size (in hundred thousands)	0.034*	0.046*	0.039
	(0.02)	(0.02)	(0.02)
Foreign Indicator	-0.707*	-0.304	-0.290
	(0.35)	(0.38)	(0.34)
East Indicator	-0.277	-0.490	-0.437
	(0.29)	(0.34)	(0.34)
Team Win-Loss Percentage (as %)	0.081***	0.100***	0.102***
	(0.01)	(0.02)	(0.02)
Player Efficiency Rating, PER	0.640***	0.368***	
	(0.05)	(0.09)	
Field Goals Made, FG		0.001	0.010**
		(0.00)	(0.00)
Free Throws Made, FT		0.005	0.014***
		(0.00)	(0.00)
Missed Field Goals		0.005	-0.002
		(0.00)	(0.00)
Missed Free Throws		0.001	-0.005
		(0.01)	(0.01)
3-Pointers Made, 3-PM		-0.002	-0.000
		(0.00)	(0.00)
Personal Fouls, PF		-0.010	-0.018**
		(0.01)	(0.01)
Offensive Rebounds, ORB		-0.004	0.003
		(0.01)	(0.01)
Defensive Rebounds, DRB		0.006*	0.007**
		(0.00)	(0.00)
Assists, AST		0.002	0.005*
		(0.00)	(0.00)
Steals, STL		0.004	0.007
		(0.01)	(0.01)
Turnovers, TOV		0.004	-0.001
		(0.01)	(0.01)
Blocked Shots, BLK		0.017**	0.021***
		(0.01)	(0.01)
Constant	-17.712***	-19.909***	-15.648***
	(1.44)	(2.12)	(1.73)
Observations	800	800	800

Robust standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05

Table 10: Controls - Box Score Statistics - Whole Dataset (n = 875)

Controls	Mean	SD	Min.	Max.
Field Goals Made (FG)	373.41	139.07	124	857
Free Throws Made (FT)	173.83	113.23	16	754
Missed Field Goals (missed_FG)	437.77	161.39	123	1,117
Missed Free Throws (missed_FT)	50.27	39.19	4	378
3-Pointers Made (threes_made)	91.98	64.50	0	402
Offensive Rebounds (ORB)	87.61	70.78	7	437
Defensive Rebounds (DRB)	301.90	140.27	78	848
Assists (AST)	221.61	149.29	29	907
Steals (STL)	70.67	30.10	11	177
Turnovers (TOV)	126.67	58.76	24	464
Blocked Shots (BLK)	41.83	37.42	0	269
Personal Fouls (PF)	163.70	40.92	70	292

Table 11: Controls - Box Score Statistics - Partial Dataset (n = 176)

Controls	Mean	SD	Min.	Max.
Field Goals Made (FG)	543.02	132.94	190	857
Free Throws Made (FT)	313.74	137.15	22	754
Missed Field Goals (missed_FG)	584.87	186.06	147	1,117
Missed Free Throws (missed_FT)	84.02	57.66	12	378
3-Pointers Made (threes_made)	109.28	88.87	0	402
Offensive Rebounds (ORB)	121.10	90.35	15	437
Defensive Rebounds (DRB)	432.66	171.14	131	848
Assists (AST)	337.58	197.22	30	907
Steals (STL)	89.49	36.41	25	177
Turnovers (TOV)	185.53	71.95	52	464
Blocked Shots (BLK)	66.78	51.18	8	269
Personal Fouls (PF)	175.88	45.42	78	292