The English Tax: Discriminatory Practices in International Male Soccer Valuation

Ross Lojek

Economics 452: University of Massachusetts Amherst

Introduction

International soccer players are exchanged from club to club for money, rather than other stores of value such as other players or draft picks. Exchanging players in return for money creates a market for soccer players, based on a variety of player attributes. Due to variations in both supply and demand, subjectivity is pervasive in this market, and may lead to discrimination based on factors irrelevant to a players real value (Majewski, 2016). This study uses data openly available from transfermarkt (transfermarkt.us, 2022) and data.world (data.world, 2022) to compare the value of players from England and players not from England, while controlling for other factors such as goals scored and age. Although the scope of this project is relatively narrow, similar methods could be used to analyze more broad trends of racial or ethnic discrimination.
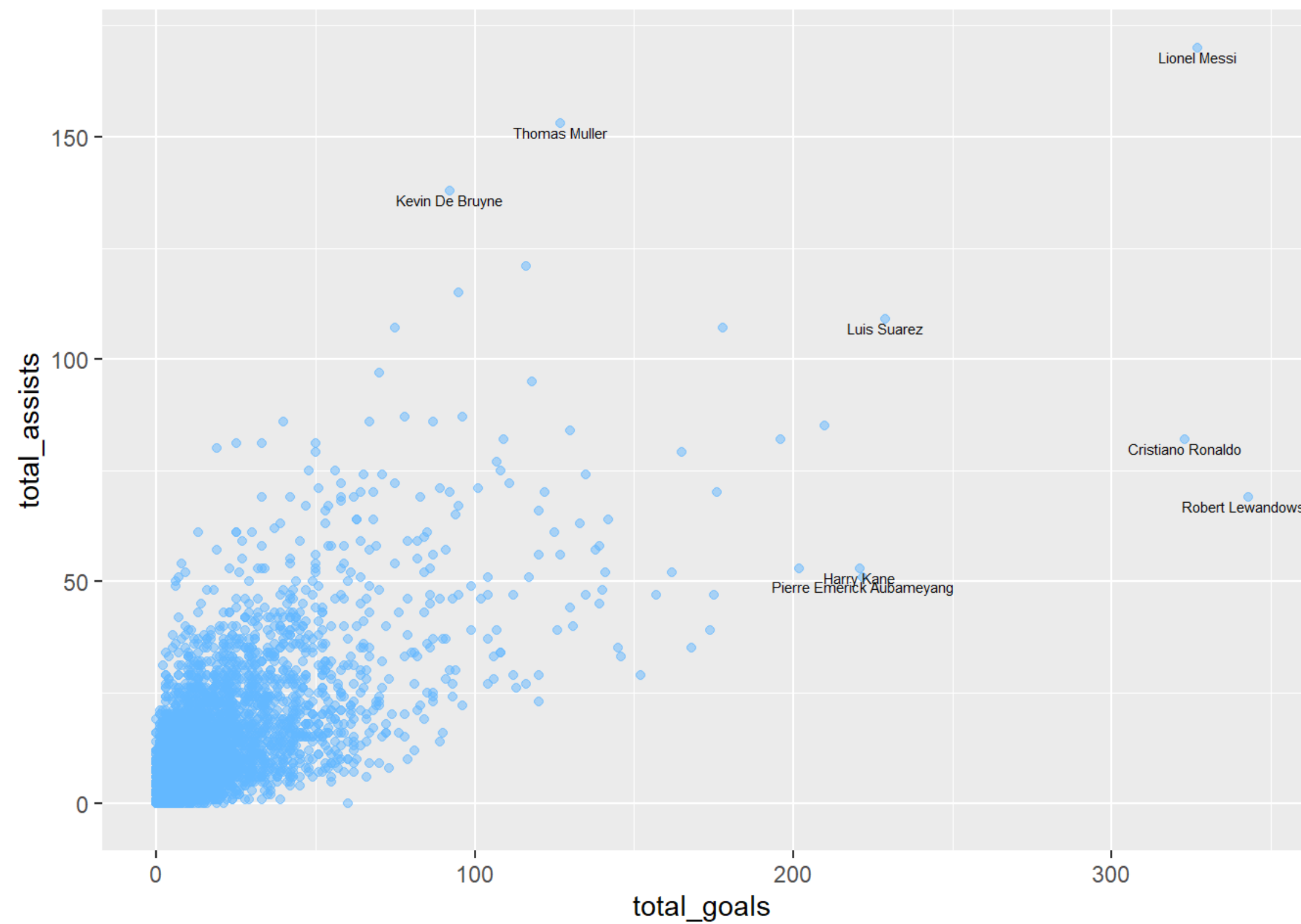

Figure 1: Each point represents a player, with goals scored on the horizontal axis and assists on the vertical axis. Points per game is derived from these data.

Data

The data collected were initially stored in three separate data frames: `players`, `player_valuations`, and `appearances`. These tables were joined using the common field `player_id` to create a final comparable data set. Below is a breakdown of columns included in the final dataset:

- `player_id`: unique value for each individual player
- `pretty_name`: name of player corresponding to `player_id`
- `country_of_citizenship`: individual player's country of citizenship
- `country_england`: dummy variable where value of 1 represents English citizenship
- `total_goals`: total goals for individual player since start date of data set
- `total_assists`: total assists for individual player since start date of data set
- `total_points`: sum of `total_goals` and `total_assists` for individual player since start date of data set

- `n`: total appearances by individual player since start of the data set
- `ppg`: points per game, or total goals plus total assists over appearances
- `market_value`: most recent trade value of individual player
- `age`: current age of player
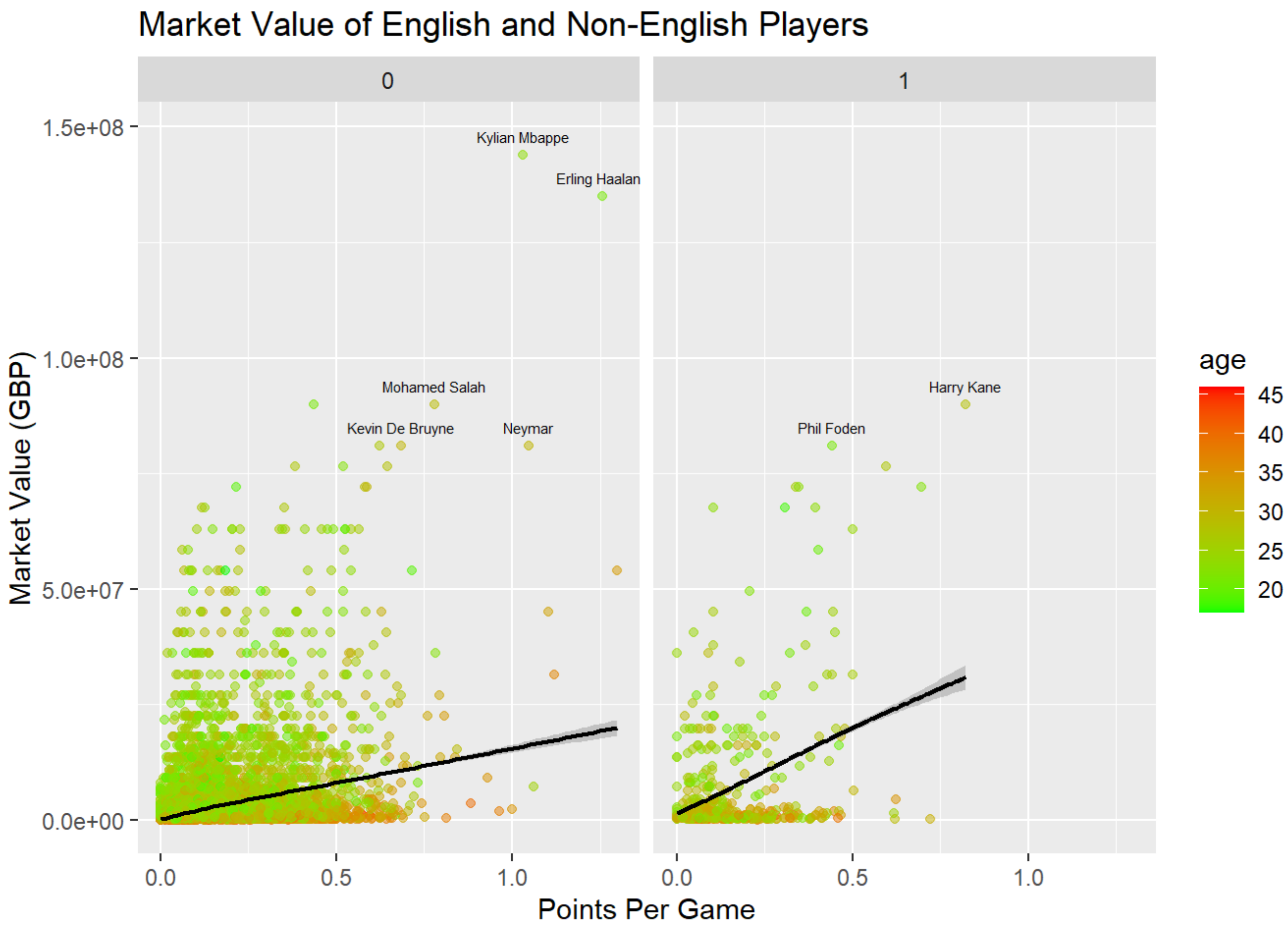- `position`: position divided into defender, midfield, and attack

#Methods The regression equation shown below was used to compare English and non-English player valuations:

$$MarketValue = \beta_0 + \beta_1(Age) + \beta_2(ln(PPG)) * \beta_3(CountryEngland)$$

Variables included in this regression model represent what should be the two most important factors in determining the value of players: productivity, and potential. The distribution of variables included in the model are similar for English and non-English players. Furthermore, the distribution of positions, is similar for English and non-English players, such that the same proportion of players are defenders, midfielders, and forwards. Differences in value cannot be a result of changes in the composition of the population English and non-English players.
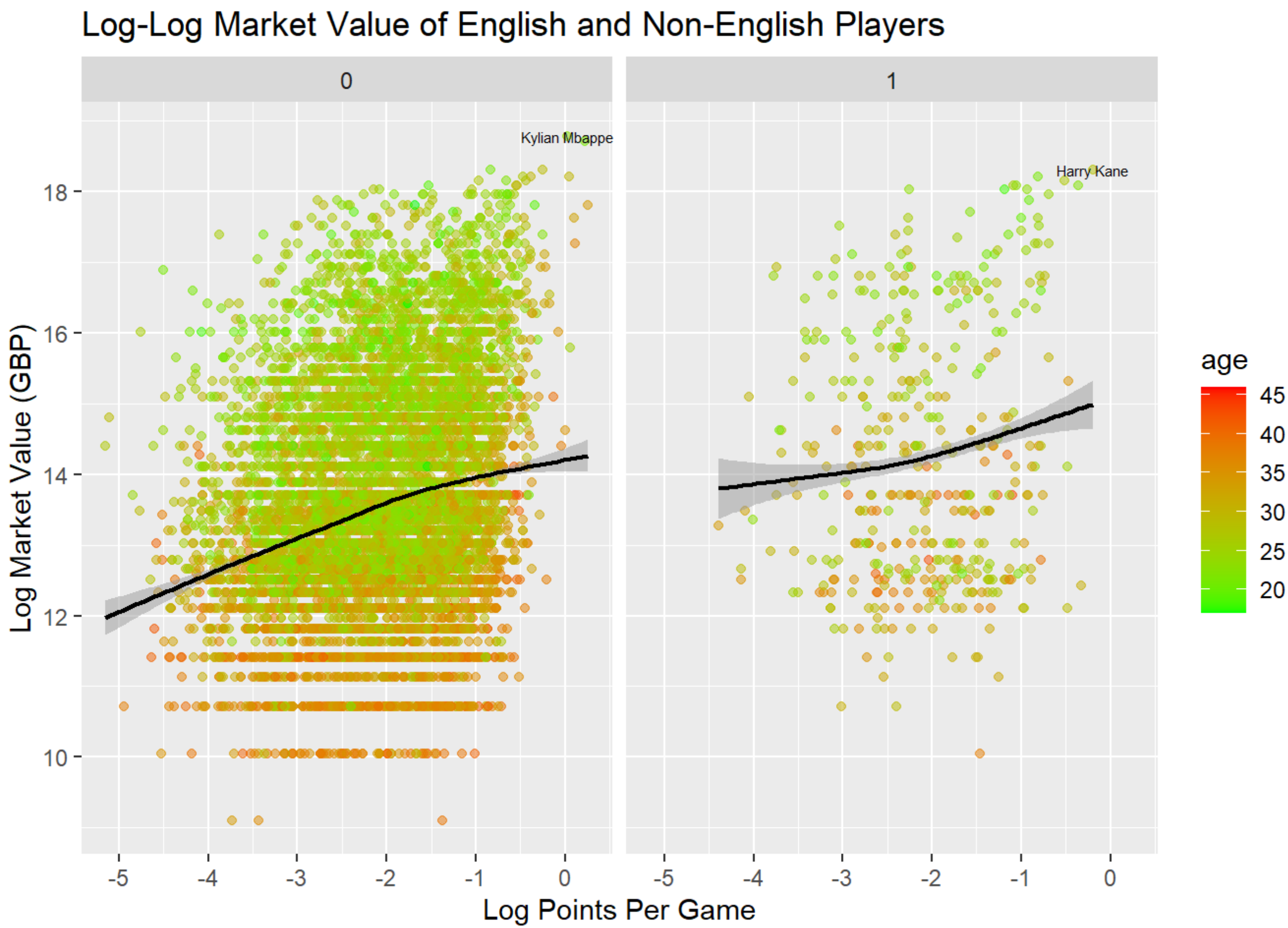
Results

term	estimate	std.error	statistic	p.value
(Intercept)	12414949.7	428777.50	28.954294	0.0000000
age	-407176.1	13954.05	-29.179775	0.0000000
ppg	15724697.8	547609.07	28.715189	0.0000000
country_england	941728.4	549779.20	1.712921	0.0867582
ppg:country_england	19914051.9	2600548.51	7.657635	0.0000000

Market Value of English and Non-English Players

Results show that English players are valued higher than non-English players, when controlling for age and points per game. This implies value in an area likely beyond that of productivity or potential, and from favorable bias toward English players. Overvaluation of English players indicates the potential for ethnic, racial, or other forms of discrimination.

term	estimate	std.error	statistic	p.value

term	estimate	std.error	statistic	p.value
(Intercept)	19.1728482	0.0859587	223.047283	0.0000000
age	-0.1612252	0.0026428	-61.005093	0.0000000
log_ppg	0.4217221	0.0165452	25.489062	0.0000000
country_england	0.3141390	0.1827714	1.718754	0.0856919
log_ppg:country_england	-0.1788412	0.0820050	-2.180857	0.0292185

Log-Log Market Value of English and Non-English Players

The second model, shown above, is a log-log transformation of the first model shown. This shows the effect of a percentage change in points per game on the percentage change of market value, which is the same concept as elasticity in economics. This model also supports the hypothesis that English players are overvalued.

References

https://data.world/dcereijo/player-scores. (2022).

Majewski, S. (2016). Identification of factors determining market value of the most valuable football players. *Central European Management Journal, 24 (3),* 91-104.

transfermarkt.us. (2022).