

Uniwersytet Warszawski
Wydział Nauk Ekonomicznych

Michał Wrzesiński
Nr albumu: 385197

**„Czynniki determinujące wartość rynkową
napastników występujących w
Premier League”**

Praca wykonana pod kierunkiem
dr inż. Janusza Gajdy
na zajęcia Ekonometrii

Warszawa, styczeń 2019

Streszczenie

Niniejsza praca obejmuje problematykę wyceny wartości rynkowej piłkarzy nożnych. Stworzono modele pośrednie analizujące wpływ poszczególnych analizowanych czynników, które w ostateczności doprowadziły do głównego modelu. Ostatecznie przeanalizowano parametry zmiennych wpływających na wartość rynkową zawodników w konfrontacji do hipotez, zawartych w literaturze, omawiającej badany problem. Analiza zawiera obserwacje piłkarzy nożnych, grających na pozycji napastnika w Premier League w sezonie 2018/2019. W celu przeprowadzenia badania posłużono się Klasycznym Modelem Regresji Liniowej oraz programem STATA, który został wykorzystany do uzyskania precyzyjnych obliczeń.

SPIS TREŚCI

WSTĘP.....	3
ROZDZIAŁ I. TEORIA ORAZ LITERATURA PROBLEMU.....	5
1.1. Fundamenty teoretyczne.....	5
1.2. Przegląd literatury.....	5
ROZDZIAŁ II. HIPOTEZY BADAWCZE.....	8
ROZDZIAŁ III. BAZA DANYCH.....	9
3.1. Etap tworzenia bazy danych.....	9
3.2. Podział zmiennych.....	9
ROZDZIAŁ IV. ESTYMACJA MODELU.....	13
4.1. Wstępna estymacja, poszukiwanie formy funkcyjnej modelu.....	13
4.2. Zmienne nieistotne.....	17
4.3. Współliniowość.....	18
4.4. Obserwacje nietypowe.....	19
4.5. Ostateczna estymacja, interpretacja parametrów	22
ROZDZIAŁ V. DIAGNOSTYKA MODELU.....	25
5.1. Test formy funkcyjnej.....	25
5.2. Testy sferyczności rozkładu reszt.....	25
5.3. Testy normalności rozkładu reszt.....	26
ROZDZIAŁ VI. WERYFIKACJA HIPOTEZ.....	28
6.1. Potwierdzenie hipotez badawczych.....	28
6.2. Zaprzeczenie hipotezom badawczym, próba interpretacji.....	28
ZAKOŃCZENIE.....	30
BIBLIOGRAFIA	31

WSTĘP

Z roku na rok pieniądze wydawane na piłkarzy rosną w zawrotnym tempie. Głównym powodem takiego stanu rzeczy jest przede wszystkim nieustająca popularność piłki nożnej. UEFA¹ w 2009 roku postanowiła spróbować zniwelować różnice pomiędzy najbogatszymi klubami Europy a resztą klubów, wprowadzając Finansowe Fair Play (FFP)², „zapewniające, że piłkarskie kluby nie będą wydawać więcej niż zarabiają, a także zabezpieczające je przed wpadnięciem w finansowe kłopoty, które mogłyby skutkować zagrożeniem ich długoterminowego przetrwania” (Goal 2018). Z perspektywy czasu można stwierdzić, że poczynione starania nie zniwelowały różnic, a dysproporcja dalej się pogłębiała. Należy w tym miejscu przytoczyć przykład ówczesnego sekretarza generalnego UEFA – Gianniiego Infantino, który zawarł układ z klubami: Paris Saint-Germain oraz Manchesterem City w 2014 r., w ramach którego „przymykał oko” na nieprawidłowości w stosowaniu się do Finansowego Fair Play (FFP) obu klubów (Buschmann & Wulzinger 2018). W konsekwencji tych wydarzeń, trudno uznać fakt wydawania coraz wyższych kwot na rynku transferowym za nadzwyczajny przypadek, a prędzej należy uznać to zjawisko za stałą tendencję. Pomijając narastającą dysproporcję w „gospodarce piłkarskiej” należy zwrócić także uwagę, że w ostatnich latach uwidacznia się trend w biciu światowych rekordów transferowych, a co za tym idzie także ligowych. Z oficjalnych statystyk wynika, że rekord transferowy Zinedine’a Zidane’a w wysokości 77.5 miliona euro trwał aż przez 8 lat. Gdy w 2009 roku Cristiano Ronaldo przeszedł do Realu Madryt za 94 miliony euro zdawało się, że jego rekord utrzyma się przez zdecydowanie dłuższy okres. Jednakże od tego czasu zmiany w liderowaniu nastąpiły już 3-krotnie, z czego ostatnie dwa rekordy padały w 2016 i 2017 roku. Tylko w okienku letnim obecnego sezonu piłkarskiego 2018/2019 nastąpiło około 30 transferów o wartości powyżej 31.5 miliona euro, co jeszcze w 1998 roku byłoby absolutnym rekordem transferowym (Transfermarkt 2018).

Na rynku światowym liderem wśród transferów oraz wartości zawodników jest oczywiście rynek europejski, gdzie zdecydowanie „na czoło” wysuwają się kluby TOP5³. Z kolei wśród nich niekwestionowanym przodownikiem wśród wartości rynkowych piłkarzy oraz kwot wydawanych na transfery jest analizowany w niniejszej pracy rynek Premier League⁴. Powodem takiego stanu rzeczy są między innymi pieniądze otrzymywane z praw telewizyjnych do transmisji meczów ligowych, które od lat są nieporównywalnie wyższe niż w pozostałych ligach TOP5. Wobec czego wydaje się być normalnym, iż to kluby angielskie wydają najwięcej gotówki na transfery. Podczas ostatniego kompletnego sezonu (2017/2018) wszystkie kluby ligi angielskiej wydały łącznie około 2.2 miliarda euro, co w porównaniu z wcześniejszym sezonem (2016/2017), jest wzrostem o ponad 0.5 miliarda euro (Transfermarkt 2018).

Dotychczas powstało wiele prac skupiających się wokół zagadnienia wartości rynkowych sportowców. Pierwsze badanie podejmujące ten temat należy datować na 1956 rok. Simon Rottenberg w swoim artykule pt. *The Baseball Players' Labor Market* zainicjował analizy charakterystyki rynku baseballistów, których zawód cieszy się dużą popularnością w Stanach Zjednoczonych. Aby przejść do tematu piłki nożnej należy przenieść się do dopiero 1993 r. - Carmichael i Thomas w *Bargaining in the Transfer Market: Theory and Evidence* podjęli się tematu wartości rynkowych piłkarzy nożnych grających w lidze angielskiej w sezonie 1990/1991. Od tego momentu powstawało coraz więcej prac skupiających się na wartościach rynkowych piłkarzy. Należy w tym miejscu przytoczyć nazwiska kolejnych autorów - między innymi: Dobson & Gerrard (1999), którzy jako jedni z wielu swoją pracą „podłożyli” fundamenty pod kolejne analizy, rozszerzające ich wyniki. Ważnym aspektem w kontekście badania rynku wartości zawodników jest również pole badań dotyczących wyceny

¹ UEFA - Unia Europejskich Związków Piłkarskich.

² Wprowadzenie Finansowego Fair Play (FFP) miało miejsce na początku sezonu 2011/2012. Z kolei zasady zostały zaakceptowane we wrześniu 2009 podczas Finansowego Panelu Kontrolnego UEFA.

³ TOP5 ligi – powszechnie przyjęte określenie pięciu najlepszych europejskich lig, w których skład wchodzi: angielska, niemiecka, hiszpańska, włoska i francuska.

⁴ Premier League – najwyższa, piłkarska klasa rozgrywkowa w Anglii

płac zawodników. Dotychczas powstało wiele prac podejmujących temat ekonomii supergwiazd – analizujących między innymi na przykładzie piłkarzy pensje supergwiazd.

Celem niniejszej pracy jest wyłożenie teorii funkcjonowania rynku piłkarskiego, analiza teoretyczna czynników wpływających na wartość rynkową zawodników oraz przeprowadzenie modelu empirycznego opierającego się o bazę zawodników z ligi angielskiej występujących na pozycji napastnika w sezonie 2018/2019. Następnie postarano się zinterpretować wyniki modelu oraz zweryfikować hipotezy badawcze.

Z uwagi na złożoność rozpatrywanego problemu postanowiono dokonać rozkładu niniejszej pracy według następującej struktury. W I rozdziale zaprezentowano fundamenty teoretyczne, a także przegląd dotychczasowych źródeł literackich, zajmujących się podobnym problemem. W II rozdziale przedstawiono oraz umotywowano zgodnie z teorią naukową hipotezy badawcze gotowe do zweryfikowania. Następnie w III rozdziale dokonano szczegółowego opisu uwzględnionych danych, a także scharakteryzowano ściśle wszystkie zachodzące ich transformacje. Dalej, w rozdziale IV zaprezentowano etapy estymacji modelu oraz interpretację końcowych parametrów. W rozdziale V przeprowadzono diagnostykę głównego modelu. Na koniec w rozdziale VI zweryfikowano hipotezy badawcze.

W niniejszej pracy użyto następującej metodologii. Model główny oraz modele pośrednie opracowane zostały z pomocą Klasycznego Modelu Regresji Liniowej. Z kolei parametry szacowano Metodą Najmniejszych Kwadratów.

Przygotowując referat, wykorzystano liczne artykuły naukowe, zawierające modele liniowe, a także inne ekonometryczne oraz predykcyjne modele. Ponadto odniesiono się do artykułów internetowych, publikacji naukowych oraz książek powiązanych z niniejszą tematyką. Wykorzystano dane empiryczne pochodzące z baz statystycznych.

ROZDZIAŁ I

Teoria oraz literatura problemu

1.1. Fundamenty teoretyczne

Wartość rynkowa zawodników jest niezwykle złożoną miarą. Głównym jej wyznacznikiem powinny być cechy piłkarza „skrojone” pod klub zainteresowany jego kupnem. W następstwie tego należy przyrzeć się możliwym charakterystykom klubów co przełoży się na wyłonienie pożądaných przez nie cech wśród zawodników:

- kluby światowej czołówki, pragnące pozyskiwać najlepszych zawodników, (głównie bez względu na wiek piłkarza), aby wzmocnić zespół na obecną chwilę;
- kluby ściągające młodych, perspektywicznych zawodników, których w przyszłości da się sprzedaż za wyższą kwotę;
- kluby bogate, lecz jeszcze nieuznawane za dobre, które decydują się na zatrudnienie popularnych piłkarzy, w celu wzmocnienia swojej marki;
- słabsze kluby, których nie stać na wielkie wydatki, poszukujące piłkarzy mających inne zalety niż aktualna forma – między innymi doświadczenie.

Niniejsza powierzchowna charakterystyka klubów pozwala na teoretyczne nakreślenie składowych wartości rynkowych zawodników. Najbardziej wymierną cechą są osiągnięcia liczbowe zawodnika – gole, asysty, czyste konta (w rozszerzonych badaniach, biorących pod uwagę pozycję bramkarza), a także statystyki opierające się na dokonaniach meczowych, takich jak: przebiegnięte kilometry, liczba sprintów, wślizgów, itp. Podobnie tyczy się to doświadczenia piłkarza w kontekście liczby meczów na danym poziomie rozgrywkowym. Związany z liczbą meczów jest bez wątpienia także wiek piłkarza. Piłkarze młodszy mają mniej rozegranych w karierze meczów, jednakże z reguły mają większe szanse na progres niż ci bardziej doświadczeni. Niezależnie od tego, wiek jest kluczowy w perspektywie występowania piłkarza w reprezentacjach młodzieżowych kraju, co przekłada się również w większości przypadków na brak możliwości występowania w dorosłej reprezentacji kraju (nie tylko ze względu na umiejętności). Następną składową jest bez wątpienia popularność piłkarza. Kiedyś mierzona przede wszystkim, że przez ekspozycję w mediach – telewizja, gazety. Obecnie – głównie poprzez budowanie swojej marki w Internecie.

1.2. Przegląd literatury

Literatura dotycząca omawianego problemu jest bardzo bogata. Modele wyceny wartości rynkowych sportowców, a także modele wyceny wartości ich zarobków (które są dosyć silnie związane z wyceną wartości sportowca) stanowią niezwykle interesujący punkt odniesienia co do przyszłych badań. Dotychczasowe prace dotyczą nie tylko piłki nożnej, ale także innych sportów, między innymi baseballu, futbolu amerykańskiego, czy też koszykówki.

Niniejszy przegląd literatury oparto w znaczącym stopniu o artykuły ściśle powiązane z tematem pracy – wartością rynkową piłkarzy nożnych, wykorzystujące w swoich badaniach Metodę Najmniejszych Kwadratów (MNK), tj. Majewski (2015), Wicker & Prinz, & Weimar & Deutscher & Upman (2013), He (2014). Obok tych autorów cenne informacje zostały zawarte również w pracach, wychodzących poza MNK: He & Cachucho & Knobbe (2015) – Lasso Regression, Tunaru & Clarrk & Viney (2005) - matematyczne losowe modele. Należy również zaznaczyć, iż ważny wpływ na dobór zmiennych miały prace, powiązane z ekonomią supergwiazd, tłumaczące zarobki sportowców: Lucifora & Simmons (2003), Lehmann & Schultze (2007), Adler (2006) oraz Franck & Nüesch (2010).

W poniższych akapitach zostały zaprezentowane wybrane aspekty prac podejmujących rozpatrywany problem według spójnych grup determinant zachowujących zgodny porządek z ich dekompozycją w kolejnych rozdziałach analizy.

Nawiązując ściśle do tematu niniejszej pracy można natrafić w badaniach na wiele odniesień co do parametrów mających pozornie bezdyskusyjny wpływ na wartość zawodnika, takich jak statystyki liczbowe piłkarza – mecze, gole, asysty (w większości wymienionych prac), ale też mniej oczywistych jak liczba żółtych oraz czerwonych kartek (Majewski 2015). Nawiązanie do goli oraz asyst przejawia się w pracach w wielu formach, między innymi jako transformacja w postaci klasyfikacji kanadyjskiej⁵ (Majewski 2015). Powyższe statystyki przedstawiane są w licznych odmianach, jako dane: z ostatniego sezonu; z kariery reprezentacyjnej (He 2014) oraz przefiltrowane – np. z najwyższej ligi rozgrywkowej podczas kariery zawodnika (Lucifora & Simmons 2003).

Kolejnym ważnym aspektem branym pod uwagę jest doświadczenie. Nie tylko wspomniana w powyższym akapicie liczba meczów charakteryzuje niniejszą cechę, lecz także wiek jest istotnym punktem w analizie wartości piłkarzy (w większości wymienionych prac). Istnieją także prace biorące pod uwagę nieliniowy charakter zmiennej wiek – uwzględniające transformację tej zmiennej (w postaci podniesienia do kwadratu) (Wicker & Prinz, & Weimar & Deutscher & Upman 2013) oraz bezpośrednio powiązane z wiekiem i umiejętnościami – występy w reprezentacji narodowej (Lucifora & Simmons 2003). Należy również odwołać się do badań biorących pod uwagę rozróżnienie na piłkarzy mogących występować w młodzieżowych reprezentacjach⁶ swojego kraju (Lucifora & Simmons 2003).

Ciekawym aspektem rozważań jest pochodzenie danego zawodnika. Badacze biorą pod uwagę nie wyłącznie kontynent, ale także konkretny kraj pochodzenia – np. wpływ narodowości niemieckiej na wartość zawodnika występującego w lidze niemieckiej (Wicker & Prinz, & Weimar & Deutscher & Upman 2013). Dodatkowo pojawia się podział kontynentalno-regionalny na: Wschodnią Europę, Unię Europejską oraz „resztę świata” (Lehmann & Schulze 2007).

Niemniej istotnym elementem analiz jest charakterystyka piłkarza. W literaturze nastąpiło dotychczas wiele różnych fragmentacji co do pozycji, na której dany zawodnik występuje. Najpopularniejszym jest standardowy podział na bramkarzy, obrońców, pomocników oraz napastników (Tunaru & Clark & Viney 2005; Lehmann & Schulze 2007), sporadycznie można napotkać na rozszerzony podział pozycji boiskowych - np. obrońców na środkowych oraz bocznych (Wicker 2013), a także okrojone próby w postaci analizy pojedynczej formacji – napastników (Majewski 2015). W dodatku niekiedy występuje rozróżnienie co do lepszej nogi zawodnika (He 2014). Nawiązując do charakterystyki piłkarza nie powinno zapominać się o parametrach fizycznych – czyli np. o wzroście, który być może nie jest tak istotny jak np. u koszykarzy (Margonis 2017), lecz także może odgrywać rolę w wycenie wartości piłkarza nożnego (Wicker 2013). Co więcej, dosyć ciekawym, a jednak niezbyt obszernie opisanym zjawiskiem jest wpływ kontuzji na wartość zawodnika (Tunaru & Clark & Viney 2005).

Kolejnym mocno zaznaczonym czynnikiem jest popularność zawodnika i klubu (głównie w literaturze odnoszącej się do wartości zarobków zawodników). Wszakże obie kwestie są od siebie mocno zależne. Niekiedy kluby zatrudniają piłkarza, który swoją własną marką podnosi ich popularność, niejako spłacając wartość swojego transferu. Przykładem takiej sytuacji jest chociażby transfer Cristiano Ronaldo z Realu Madryt do Juventusu Turyn. Umowa sponsorska Adidasa z Juventusem wzrosła z ponad 20 milionów funtów do około 46 milionów funtów rocznie, co powiązane jest w pewnej części z transferem Ronaldo (Shergold 2018). Ponadto Juventus sprzedał w 24 godziny od transferu 520 000 koszulek z nazwiskiem zawodnika, z czystym zyskiem dla klubu w postaci od 6 do 9 milionów dolarów (w zależności od umowy z Adidasem) (Dawson 2018). Popularny piłkarz może przynieść klubowi także nowych kibiców, podążających za swoim idolem, co przekłada się również na większe zyski klubu. W literaturze można znaleźć wpływ ekspozycji piłkarza (popularność) w mediach na jego zarobki (Adler 2006). Punktem odniesienia w badaniach jest także ilość corocznych publikacji w różnorodnych magazynach oraz gazetach (Franck & Nüesch 2010).

⁵ Klasyfikacja kanadyjska – klasyfikacja zawierająca statystykę sumującą gole i asysty dla danego piłkarza.

⁶ Z reguły najstarsza reprezentacja młodzieżowa ma limit wiekowy do lat 21, jednakże na igrzyskach jest wyjątek – najstarszy zawodnik może mieć 23 lata (dodatkowo może grać 3 piłkarzy bez limitu wiekowego).

Co więcej, powinno się zasygnalizować wpływ obiektywnych rankingów, bazujących na meczowych statystykach zawodników, prezentujących ich oceny – Carling Opta Index (Tunaru & Clark & Viney 2005) oraz WhoScored (He & Cachucho & Knobbe 2015). Analogicznie wpływ mają także pozycje w rankingach dotyczące: klubów – liga krajowa oraz reprezentacji – ranking FIFA (Majewski 2015), które mogą odgrywać niebagatelną rolę w wycenie ich wartości. Również dostrzegalna jest fragmentacja na najlepsze reprezentacje oraz te słabsze (Wicker & Prinz, & Weimar & Deutscher & Upman 2013).

Istnieje w literaturze również swego rodzaju podział na kluby większe i mniejsze. Między innymi w lidze włoskiej – Juventus oraz Milan a reszta klubów, ze względu na wystające ponad normę charakterystyki tych klubów (między innymi: przychody, liczba kibiców) (Lucifora & Simmons 2003).

ROZDZIAŁ II

Hipotezy badawcze

W silnym nawiązaniu do powyższej literatury postanowiono zweryfikować następujące hipotezy według podgrup, które zostały szerzej opisane w kolejnym rozdziale opracowania. Wskutek czego oczekuje się, iż dla:

- **osiągnięć indywidualnych** – wzrost liczby goli oraz asyst, będzie oznaczał wzrost wartości zawodnika;
- **agresji** – zmienne oznaczające liczbę żółtych oraz czerwonych kartek będą nieistotne;
- **doświadczenia** – wzrost liczby meczów będzie oznaczać spadek wartości rynkowej zawodnika. Oczekuje się także, że starszy piłkarz będzie mniej warty. Zmienna rozróżniająca piłkarzy młodzieżowych również nie będzie istotna ze względu na dość trudny do określenia jej punkt graniczny (w niniejszym badaniu – wiek 23 lat, Lucifora & Simmons – wiek 21 lat). Zaś zmienna określająca występ w reprezentacji seniorskiej będzie intensyfikować wartość rynkową piłkarza;
- **pochodzenia** – zmienne określające kontynent oraz czy piłkarz jest Anglikiem okażą się nieistotne;
- **charakterystyk boiskowych** – pozycja okaże się nieistotną zmienną, ze względu na badanie tylko jednej formacji - ataku. Zmienna określająca lepszą nogę zawodnika również okaże się nieistotna;
- **parametrów fizycznych** – wzrost zawodnika będzie zmienną nieistotną;
- **kontuzji** – zwiększenie liczby opuszczonych meczów w karierze będzie oznaczać spadek wartości zawodnika,
- **popularności** - wzrost liczby obserwujących w social mediach podwyższy wartość zawodnika;
- **popularności klubu** - przyrost liczby obserwujących klub w social mediach zintensyfikuje wartość zawodnika;
- **ocen** – wzrost punktów w rankingu przełoży się na wyższą wartość zawodnika;
- **siły reprezentacji** – zmienna określająca pozycję w rankingu państwa, z którego pochodzi zawodnik okaże się nieistotna;
- **siły klubu** – pozycja klubu, a także występowanie w silnym klubie będą nieistotne.

ROZDZIAŁ III

Baza danych

3.1. Etap tworzenia bazy danych

Bazę danych, wykorzystaną w budowie modelu, skonstruowano za pomocą źródeł, gromadzących dane na temat charakterystyk i statystyk piłkarzy nożnych oraz klubów piłkarskich. Poza tego typu źródłami posłużono się informacjami z najpopularniejszych „social mediów”. Wszystkie poniższe dane są aktualne na dzień 20 grudnia 2018 r.

Głównym punktem wyjścia przy rozpoczęciu konstrukcji podstawy bazy danych było znalezienie informacji co do wiarygodnej strony internetowej, posiadającej aktualne wyceny wartości rynkowych piłkarzy nożnych. Na rynku takich portali zdecydowanym liderem pozostaje serwis – *transfermarkt.de*. Przede wszystkim wyceny wartości piłkarzy są kompetentne dla zawodników z najlepszych europejskich lig, co jest ważne w aspekcie budowania bazy danych opartej o obserwacje z ligi angielskiej. Wobec czego za zmienną objaśnianą przyjęto aktualne wartości rynkowe zawodników (w milionach euro) dla wszystkich wyselekcjonowanych piłkarzy nożnych, oznaczonych na portalu *transfermarkt.de* jako napastnicy (konkretnie: środkowi, cofnięci, lewi oraz prawi napastnicy), a następnie zebrano informacje dotyczące:

- liczby meczów, goli oraz asyst (w obecnym sezonie, w karierze w ligach z TOP5⁷);
- liczbę żółtych oraz czerwonych kartek (w karierze w ligach z TOP5);
- wieku;
- wzrostu;
- narodowości (w niektórych przypadkach występowały dwie narodowości, wobec czego brano pod uwagę tylko kraj, którego barwy dany zawodnik reprezentuje);
- wartości danej drużyny (w milionach euro);
- dokładnej pozycji, na której dany zawodnik najczęściej występuje (ŚN, CN, PN, LN⁸);
- liczby opuszczonych meczów z powodu kontuzji w trakcie trwania kariery (poszczególne okresy).

W kolejnym etapie procesu zbierania danych zweryfikowano obecnie najpopularniejsze źródła ocen oraz rankingów portali zajmujących się piłką nożną/ligą angielską. W konsekwencji wskazano na następujące strony:

- *whoscored.com* (oceny poszczególnych zawodników za obecny sezon Premier League w skali 0-10)⁹;
- *premierleague.com* (pozycja danego klubu w tabeli Premier League od 1 do 20 miejsca);
- *fifa.com/fifa-world-ranking* (pozycja danej reprezentacji w rankingu FIFA¹⁰ od 1 do 211 miejsca).

Na koniec użyto danych z najpopularniejszych „social mediów” zawodników:

⁷ Analogicznie do badania Lucifory oraz Simmonsa (Lucifora & Simmons 2003), którzy pod uwagę brali jedynie Serie A, lecz wtedy rynek był zdecydowanie bardziej hermetyczny, wobec czego przekładając to na dzisiejsze czasy oraz większy przepływ zawodników można wziąć pod uwagę jedynie statystyki z TOP5 (nie tylko Premier League).

⁸ Oznaczenia kolejno: środkowy, cofnięty, lewy oraz prawy napastnik

⁹ Niektórzy zawodnicy nie posiadali danych, najczęściej ze względu na brak (bądź małą liczbę) występów w obecnym sezonie

¹⁰ Ranking FIFA – ranking biorący pod uwagę drużyny zrzeszone w Międzynarodowej Federacji Piłki Nożnej

- *instagram.com* oraz *twitter.com* (liczba „obserwujących” danego zawodnika oraz jego klub);
- *facebook.com* (liczba „polubień” strony danego zawodnika oraz jego klubu).

W dalszej części pracy przyjęto ujednolicenie dla wszystkich trzech powyższych „social mediów” w postaci określenia – liczba obserwujących (również dla *facebook.com*).

3.2. Podział zmiennych

W nawiązaniu do przedstawionej literatury oraz wiedzy eksperckiej, dotyczącej specyfiki ligi angielskiej, dokonano dekompozycji zmiennych według następujących podgrup:

- Osiągnięcia indywidualne – *goals, assists, Agoals, Aassists*;
- Agresja – *yellow, red*;
- Doświadczenie – *matches, Amatches, age, age2, youth, represent*;
- Pochodzenie – *continent, english*;
- Charakterystyka boiskowa – *position, leg*;
- Parametry fizyczne – *height*;
- Kontuzje – *Injuries*;
- Popularność – *social*;
- Popularność klubu – *teamsocial*;
- Oceny – *ratingWho*;
- Siła reprezentacji – *positionFIFA*;
- Siła klubu – *positionPL, teamvalue*.

Powyższe zmienne zostały szczegółowo opisane w poniższej Tabeli 1. Ponadto w wyniku weryfikacji bazy danych odrzucono 12 ze 132 obserwacji ze względu na braki danych przy niektórych zmiennych (główny powód to brak występów w aktualnym sezonie). Aby przedstawić zmienne w powyższej formie przeprowadzono następujące transformacje:

- dla zmiennych: *goals, assists, yellow, red, matches*, zsumowano analogiczne statystyki uzyskane przez zawodnika w całej karierze w ligach TOP5;
- zmienną *age* zmodyfikowano w postaci podniesienia jej do kwadratu - otrzymano zmienną *age2*;
- dla zmiennej *youth* wyselekcjonowano piłkarzy mających maksimum 23 lata oraz tych, którzy mają powyżej;
- dla zmiennej *represent* wyselekcjonowano piłkarzy mających przynajmniej jeden występ w seniorskiej reprezentacji swojego kraju oraz tych którzy żadnego nie mają;
- dla zmiennej *continent* przyporządkowano narodowość zawodnika do odpowiedniego kontynentu;
- dla zmiennej *english* wyselekcjonowano piłkarzy angielskich;
- dla zmiennej *position* zespolono LN i PN w S (boczni napastnicy) oraz CN i ŚN w N (środkowi napastnicy);
- dla zmiennej *Injuries* zsumowano liczby opuszczonych meczów z powodu kontuzji;
- dla zmiennych *social* oraz *teamsocial*: zsumowano obserwujących z Facebooka, Twittera oraz Instagrama;
- dla zmiennej *strong* podjęto selekcję na kluby silne oraz słabe (na podstawie ostatnich 20 lat).

Tabela 1. Opis zmiennych użytych w modelu.

Podgrupa	Etykieta	Pełna nazwa zmiennej	Miara	Ilościowa/ Jakościowa	Dyskretna/ Ciągła
	<i>value</i>	Wartość zawodnika	Euro	ilościowa	ciągła
Osiągnięcia indywidualne	<i>goals</i>	Liczba goli w ligach TOP5 w karierze	Liczba goli	ilościowa	ciągła
	<i>assists</i>	Liczba asyst w ligach TOP5 w karierze	Liczba asyst	ilościowa	ciągła
	<i>Agoals</i>	Liczba goli w sezonie 2018/2019	Liczba goli	ilościowa	ciągła
	<i>Aassists</i>	Liczba asyst w sezonie 2018/2019	Liczba asyst	ilościowa	ciągła
Agresja	<i>yellow</i>	Żółte kartki w ligach TOP5	Liczba żółtych kartek	ilościowa	ciągła
	<i>red</i>	Czerwone kartki w ligach TOP5	Liczba czerwonych kartek	ilościowa	ciągła
Doświadczenie	<i>matches</i>	Liczba występów w karierze w ligach TOP5	Liczba meczów	ilościowa	ciągła
	<i>Amatches</i>	Liczba meczów w sezonie 2018/2019 dla klubu	Liczba meczów	ilościowa	ciągła
	<i>age</i>	Wiek	Lata	ilościowa	ciągła
	<i>age2</i>	Wiek do kwadratu	Lata ²	ilościowa	ciągła
	<i>youth</i>	Czy zawodnik ma możliwość gry w reprezentacji młodzieżowej	Zbiór liczb naturalnych {0,1}	jakościowa	dyskretna
	<i>represent</i>	Czy zawodnik wystąpił w seniorskiej reprezentacji	Zbiór liczb naturalnych {0,1}	jakościowa	dyskretna
Pochodzenie	<i>continent</i>	Kontynent, który reprezentuje zawodnik	Zbiór liczb naturalnych {0,1,2,3,4,5}	jakościowa	dyskretna
	<i>english</i>	Czy zawodnik jest Anglikiem	Zbiór liczb naturalnych {0,1}	jakościowa	dyskretna
Charakterystyka boiskowa	<i>position</i>	Pozycja na której występuje zawodnik	Zbiór liczb naturalnych {0,1}	jakościowa	dyskretna
	<i>leg</i>	Lepsza noga zawodnika	Zbiór liczb naturalnych {0,1,2}	jakościowa	dyskretna
Parametry fizyczne	<i>height</i>	Wzrost	Zbiór liczb naturalnych	ilościowa	ciągła
Kontuzje	<i>Injuries</i>	Liczba opuszczonych meczów w trakcie kariery	Zbiór liczb naturalnych	ilościowa	ciągła
Popularność	<i>social</i>	Suma liczby obserwujących zawodnika na Twitterze, Facebooku oraz Instagramie	Liczba obserwujących	ilościowa	ciągła
Popularność klubu	<i>teamsocial</i>	Suma liczby obserwujących klub zawodnika na Twitterze, Facebooku oraz Instagramie	Liczba obserwujących	ilościowa	ciągła
Oceny	<i>ratingWho</i>	Ocena zawodnika za występy w Premier League w sezonie 2018/2019	Zbiór liczb dodatnich rzeczywistych	ilościowa	ciągła
Siła reprezentacji	<i>positionFIFA</i>	Pozycja reprezentacji zawodnika w rankingu FIFA	Zbiór liczb naturalnych	ilościowa	ciągła
Siła klubu	<i>positionPL</i>	Pozycja klubu w Premier League w sezonie 2018/2019	Zbiór liczb naturalnych	ilościowa	ciągła
	<i>strong</i>	Czy klub jest zaliczany do grona potentatów Premier League	Zbiór liczb naturalnych {0,1}	jakościowa	dyskretna
	<i>teamvalue</i>	Wartość wszystkich zawodników występujących w klubie	Euro	ilościowa	ciągła

Źródło: Opracowanie własne na podstawie zebranych danych

Szczegółowy opis zmiennych dyskretnych zawartych w Tabeli 1.:

- *youth*: 1 – zawodnik maksymalnie 23-letni, 0 – zawodnik powyżej 23 lat;
- *represent*: 1 – zawodnik, który wystąpił w seniorskiej kadrze, 0 – zawodnik, który nigdy nie wystąpił w seniorskiej reprezentacji swojego kraju;
- *continent*: 0 – Europa, 1- Ameryka Południowa, 2 – Afryka, 3 – Ameryka Północna, 4 – Azja, 5 – Australia;
- *english*: 1 – Anglik, 0 – zawodnik nie jest Anglikiem;
- *position*: 1 – środkowy napastnik, 0 – boczny napastnik;
- *leg*: 0 – zawodnik ma dominującą prawą nogę, 1 – zawodnik ma dominującą lewą nogę, 2 – zawodnik obie nogi ma na tym samym poziomie;
- *strong*: 1 – zawodnik ze zbioru klubów {Manchester City, Manchester United, Liverpool FC, Chelsea London, Tottenham Hotspur, Arsenal London}, 0 – zawodnik z innego klubu..

ROZDZIAŁ IV

Estymacja modelu

4.1 Wstępna estymacja, poszukiwanie formy liniowej modelu

W początkowym etapie estymacji modelu przystąpiono do regresji wszystkich dostępnych zmiennych w nienaruszonej formie – wg Tabeli 1. Otrzymano następujące wyniki:

TABELA 2. Wstępna estymacja (Model 1).

Source	SS	df	MS	Number of obs	=	120
Model	9.2999e+16	30	3.1000e+15	F(30, 89)	=	15.83
Residual	1.7433e+16	89	1.9588e+14	Prob > F	=	0.0000
				R-squared	=	0.8421
				Adj R-squared	=	0.7889
Total	1.1043e+17	119	9.2801e+14	Root MSE	=	1.4e+07

value	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
goals	148572.2	99743.35	1.49	0.140	-49615.71	346760.1
assists	52701.82	223579.6	0.24	0.814	-391546.2	496949.8
Agoals	2557071	923901.1	2.77	0.007	721299	4392843
Aassists	2737195	990199.2	2.76	0.007	769690.3	4704700
yellow	94018.37	242286.7	0.39	0.699	-387400.2	575437
red	-2302211	2494289	-0.92	0.359	-7258309	2653888
matches	-28396.28	51537.45	-0.55	0.583	-130800.1	74007.54
Amatches	-343604.6	375299	-0.92	0.362	-1089316	402106.6
age	-235489.9	8718224	-0.03	0.979	-1.76e+07	1.71e+07
age2	-49264.05	156056.7	-0.32	0.753	-359345.4	260817.3
1.youth	-1.24e+07	6912172	-1.79	0.077	-2.61e+07	1359768
1.represent	3610970	4398251	0.82	0.414	-5128262	1.24e+07
continent						
1	-3710023	5060336	-0.73	0.465	-1.38e+07	6344756
2	4941879	6188563	0.80	0.427	-7354664	1.72e+07
3	-1.24e+07	9426238	-1.31	0.193	-3.11e+07	6359644
4	6671708	8765310	0.76	0.449	-1.07e+07	2.41e+07
5	1.09e+07	1.81e+07	0.60	0.549	-2.51e+07	4.69e+07
1.english	7558078	4012479	1.88	0.063	-414632.1	1.55e+07
1.position	-2969889	3871630	-0.77	0.445	-1.07e+07	4722958
leg						
1	-1591453	3474856	-0.46	0.648	-8495917	5313011
2	4995687	7430133	0.67	0.503	-9767830	1.98e+07
height	450759.5	252343.4	1.79	0.077	-50641.5	952160.5
Injuries	-139489.7	60713.53	-2.30	0.024	-260126.1	-18853.17
social	1241184	338344	3.67	0.000	568901.2	1913466
teamsocial	.070157	.0983662	0.71	0.478	-.1252947	.2656086
ratingWho	5781979	7679104	0.75	0.453	-9476236	2.10e+07
positionFIFA	-39792.96	105345.2	-0.38	0.707	-249111.7	169525.8
positionPL	60084.91	472489.4	0.13	0.899	-878741.5	998911.3
1.strong	-4038301	1.16e+07	-0.35	0.729	-2.71e+07	1.90e+07
teamvalue	.0197101	.0151767	1.30	0.197	-.0104458	.049866
_cons	-7.48e+07	1.35e+08	-0.56	0.580	-3.42e+08	1.93e+08

Źródło: Obliczenia własne za pomocą programu STATA

Na podstawie Tabeli 2. odrzucono hipotezę o łącznej nieistotności zmiennych ($p\text{-value} = 0.0000 < 0.05$). Następnie postanowiono przeprowadzić test RESET, aby stwierdzić czy model przyjmuje formę liniową.

TABELA 3. Test RESET dla Modelu 1.

Ramsey RESET test using powers of the fitted values of value

Ho: model has no omitted variables

$F(3, 86) = 32.32$

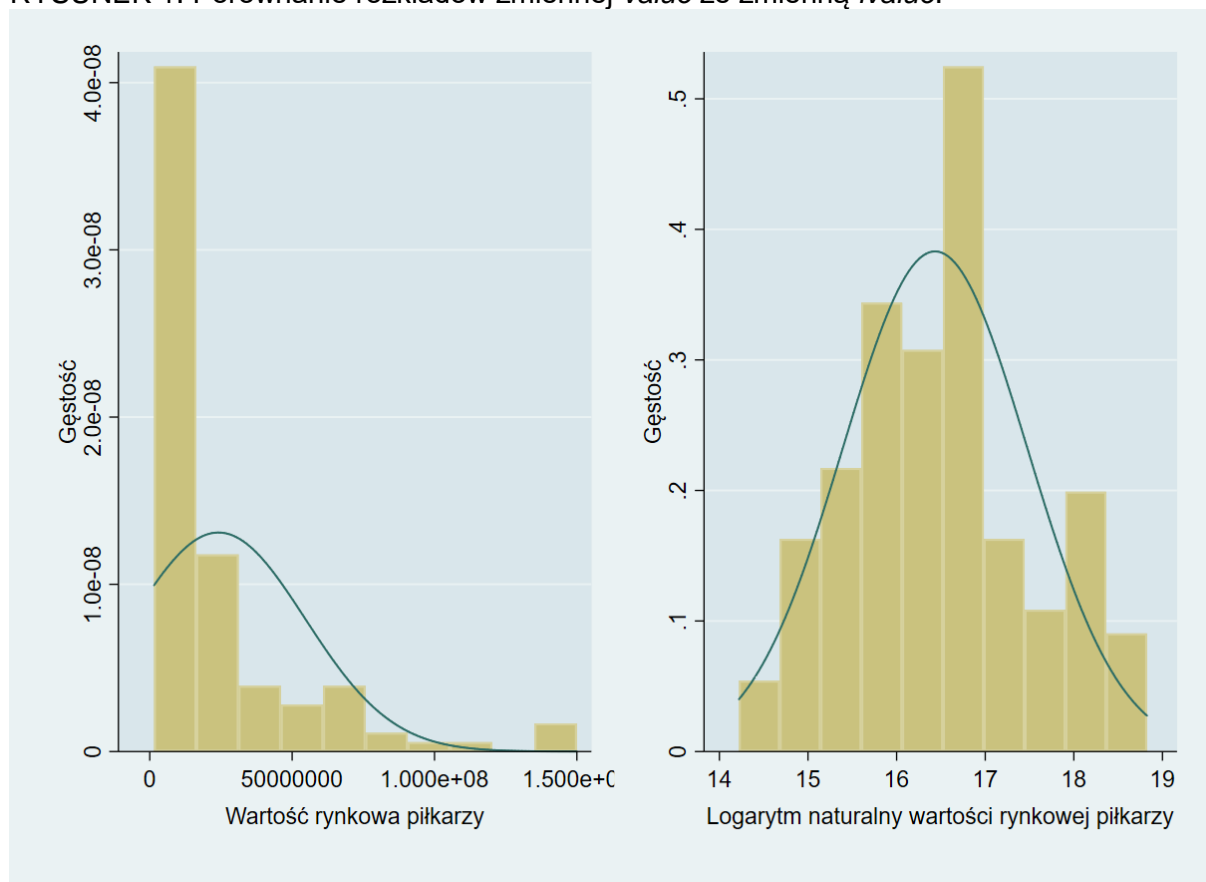
Prob > F = 0.0000

Źródło: Obliczenia własne za pomocą programu STATA

Test RESET jednoznacznie odrzuca hipotezę o formie liniowej modelu (0 mniejsze od poziomu istotności -0.05). W konsekwencji przystąpiono do poszukiwania takich zależności, które sprawiają, że model spełni niniejsze założenie KMRL.

W pierwszej kolejności rozpoczęto od analizy rozkładu zmiennej objaśnianej – value. Sprawdzono jak zmienna „zachowuje się” po zlogarytmowaniu jej.

RYSUNEK 1. Porównanie rozkładów zmiennej *value* ze zmienną *lvalue*.



Źródło: Wykresy wykonane za pomocą programu STATA

Z Rysunku 1. jednoznacznie wynika, że rozkład logarytmu naturalnego zmiennej *value* (*lvalue*) jest bardzo zbliżony do rozkładu normalnego w odróżnieniu od rozkładu zmiennej *value*. Na potwierdzenie wiarygodności powyższego wykresu przeprowadzono również test na normalność (Jarque-Bera), który jednoznacznie potwierdził niniejszą tezę ($p\text{-value}=0.2241>0.05$, wobec czego nie można odrzucić hipotezy o normalności rozkładu):

TABELA 4. Test Jarque-Bera na normalność zmiennej *lvalue*.

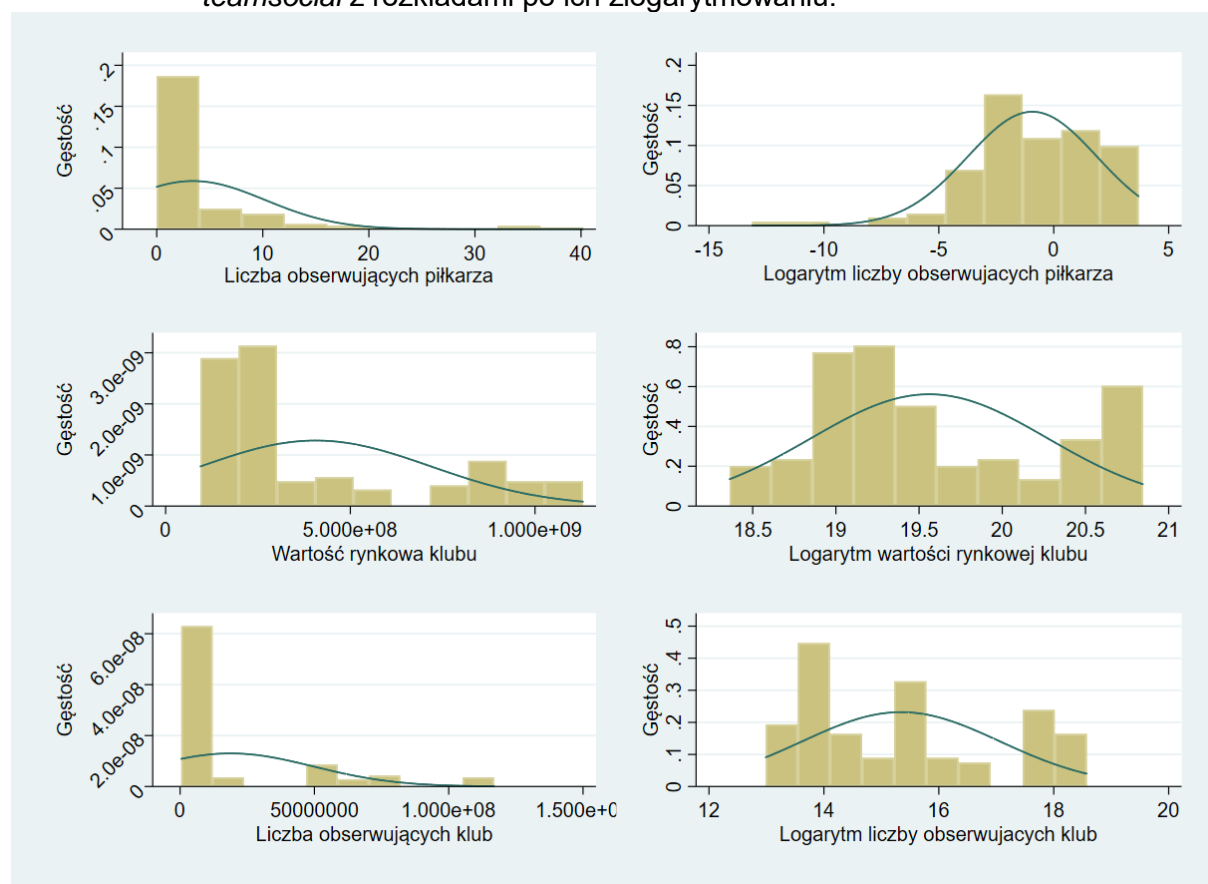
Skewness/Kurtosis tests for Normality

Variable	Obs	Pr (Skewness)	Pr (Kurtosis)	adj chi2 (2)	Prob>chi2
-----+-----					
lvalue	120	0.1433	0.3754	2.99	0.2241

Źródło: Obliczenia własne za pomocą programu STATA

Analogicznie postanowiono wykonać histogramy dla zmiennych *social*, *teamvalue* oraz *teamsocial*, które zgodnie z intuicją, po transformacji w logarytm naturalny powinny zachowywać się podobnie jak w przypadku zmiennej *value*.

RYSUNEK 2. Porównanie pierwotnych rozkładów zmiennych: *social*, *teamvalue* oraz *teamsocial* z rozkładami po ich zlogarytmowaniu.



Źródło: Wykresy wykonane za pomocą programu STATA

Histogramy pierwotnych zmiennych zamieszczone na powyższym Rysunku 2. (po lewej stronie) jednoznacznie wskazują na to, iż rozkłady zmiennych odbiegają w znaczącym

stopniu od rozkładu normalnego. Według przewidywań powyższe zmienne w wersji zlogarytmowanej (po prawej stronie rysunku) są dużo bardziej zbliżone do rozkładów normalnych, choć nie aż w tak dobrym stopniu jak przy zmiennej *value*. W konsekwencji w estymowanym modelu zamieniono te zmienne według powyższego wyjaśnienia.

Inne zmienne ilościowe pozostają w modelu bez transformacji ze względu na ich późniejszą potrzebę interpretacji w formie niezlogarytmowanej, a także ze względu na fakt istnienia zera przy niektórych obserwacjach (dla zmiennych: *goals*, *assists*, *Agoals*, *Aassists*, *yellow*, *red*, *matches*, *Amatches*, *Injuries*), co po zlogarytmowaniu usunęłoby takowe obserwacje z dalszych regresji.

TABELA 5. Estymacja ze zlogarytmowanymi zmiennymi (Model 2).

Source	SS	df	MS	Number of obs	=	120
Model	113.188185	30	3.77293951	F(30, 89)	=	21.18
Residual	15.8525989	89	.178119089	Prob > F	=	0.0000
				R-squared	=	0.8772
				Adj R-squared	=	0.8357
Total	129.040784	119	1.08437634	Root MSE	=	.42204

lvalue	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
goals	.0050913	.0027872	1.83	0.071	-.0004468	.0106294
assists	.0070226	.0062453	1.12	0.264	-.0053866	.0194318
Agoals	.0130839	.0276883	0.47	0.638	-.0419322	.0681
Aassists	.016255	.0300002	0.54	0.589	-.0433547	.0758647
yellow	.0087601	.0073722	1.19	0.238	-.0058882	.0234085
red	.100256	.0749183	1.34	0.184	-.0486051	.2491171
matches	-.0033886	.00143	-2.37	0.020	-.00623	-.0005472
Amatches	.0344126	.0114419	3.01	0.003	.0116779	.0571473
age	.505793	.2680768	1.89	0.062	-.02687	1.038456
age2	-.0110565	.004794	-2.31	0.023	-.020582	-.001531
1.youth	-.0647044	.2128476	-0.30	0.762	-.4876281	.3582193
1.represent	.2277839	.1344859	1.69	0.094	-.0394368	.4950046
continent						
1	-.1272571	.1499966	-0.85	0.398	-.4252971	.1707829
2	-.2047945	.1875506	-1.09	0.278	-.5774536	.1678645
3	-.382427	.2842711	-1.35	0.182	-.9472677	.1824136
4	.05645	.2646286	0.21	0.832	-.4693614	.5822614
5	.2852439	.5502132	0.52	0.605	-.808018	1.378506
1.english	.1085194	.1261604	0.86	0.392	-.1421587	.3591974
1.position	-.0578736	.1168224	-0.50	0.622	-.2899973	.17425
leg						
1	-.2086342	.1049435	-1.99	0.050	-.4171548	-.0001137
2	.1126721	.2238134	0.50	0.616	-.3320403	.5573845
height	.0097853	.0074821	1.31	0.194	-.0050814	.024652
Injuries	-.0021553	.0018518	-1.16	0.248	-.0058347	.0015242
lsocial	.1077111	.0226985	4.75	0.000	.0626097	.1528125
lteamsocial	.0258922	.0742205	0.35	0.728	-.1215822	.1733667
ratingWho	.4573005	.2320684	1.97	0.052	-.0038144	.9184155
positionFIFA	.0007874	.0032075	0.25	0.807	-.0055858	.0071606
positionPL	-.0143693	.0148953	-0.96	0.337	-.0439659	.0152273
1.strong	-.0559453	.218468	-0.26	0.798	-.4900365	.3781459
lteamvalue	.2295246	.2047217	1.12	0.265	-.1772531	.6363023
_cons	-.5914217	5.627981	-0.11	0.917	-11.7741	10.59126

Źródło: Obliczenia własne za pomocą programu STATA

R^2 dopasowane zdecydowanie się poprawiło, wobec czego przystąpiono do ponownego testu formy funkcyjnej modelu, którego wynik jest pokazany w Tabeli 6.

TABELA 6. Test RESET dla Modelu 2.

Ramsey RESET test using powers of the fitted values of lvalue

Ho: model has no omitted variables

F(3, 86) = 3.47

Prob > F = 0.0195

Źródło: Obliczenia własne za pomocą programu STATA

Test RESET w dalszym ciągu wykazuje, że model nie ma formy liniowej, jednakże wartość p-value wzrosła z 0 do 0.0195, co przy założonym poziomie istotności – 0.05, dalej odbiega od hipotezy zerowej, aczkolwiek po dalszych próbach poprawy modelu powinno zdecydowanie pozwolić na przyjęcie hipotezy zerowej o poprawności funkcyjnej modelu.

4.2. Poprawa wyników modelu, usunięcie zmiennych nieistotnych

Następnie, przeanalizowano wyniki regresji pod kątem istotności poszczególnych zmiennych oraz zachowania się R^2 dopasowanego w wyniku ich usuwania. Po wielu próbach doprowadzono do momentu, w którym usunięto następujące zmienne:

- ilościowe: *red*, *positionFIFA*, *positionPL*, *Agoals*, *Aassists*, *lteamsocial*;
- jakościowe: *strong*, *position*, *youth*, *continent*.

W wyniku powyższych “poprawek” w Tabeli 7. otrzymano następującą estymację modelu:

TABELA 7. Estymacja po usunięciu zmiennych nieistotnych (Model 3).

Source	SS	df	MS	Number of obs	=	120
Model	111.997746	16	6.99985911	F(16, 103)	=	42.30
Residual	17.0430386	103	.165466394	Prob > F	=	0.0000
				R-squared	=	0.8679
				Adj R-squared	=	0.8474
Total	129.040784	119	1.08437634	Root MSE	=	.40678

lvalue	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
goals	.0057768	.0021568	2.68	0.009	.0014992	.0100543
assists	.006529	.0055882	1.17	0.245	-.0045539	.0176118
yellow	.0103688	.006365	1.63	0.106	-.0022548	.0229923
matches	-.0036094	.0012596	-2.87	0.005	-.0061076	-.0011112
Amatches	.0358489	.0096873	3.70	0.000	.0166363	.0550614
age	.5263354	.1733704	3.04	0.003	.182496	.8701747
age2	-.0113293	.0032667	-3.47	0.001	-.0178081	-.0048505
1.represent	.250826	.1162246	2.16	0.033	.020322	.4813301
1.english	.1760424	.1044501	1.69	0.095	-.0311097	.3831944
leg						
1	-.1791588	.0946297	-1.89	0.061	-.3668344	.0085169
2	.2167853	.1852579	1.17	0.245	-.1506301	.5842006
height	.0108801	.0059632	1.82	0.071	-.0009466	.0227068
Injuries	-.0020988	.0016596	-1.26	0.209	-.0053903	.0011927
lsocial	.1022067	.0207277	4.93	0.000	.0610982	.1433153
ratingWho	.5875574	.1529292	3.84	0.000	.2842584	.8908565
lteamvalue	.3575125	.0864224	4.14	0.000	.1861141	.5289109
_cons	-4.252023	3.347357	-1.27	0.207	-10.89072	2.386669

Źródło: Obliczenia własne za pomocą programu STATA

TABELA 8. Test RESET dla Modelu 3.

Ramsey RESET test using powers of the fitted values of lvalue

Ho: model has no omitted variables

$F(3, 100) = 3.21$

Prob > F = 0.0264

Źródło: Obliczenia własne za pomocą programu STATA

P-value testu RESET w dalszym ciągu wzrasta (z 0.0195 do 0.0264 w stosunku do poprzedniego wyliczenia) wobec czego należy uznać, że przyjęte wcześniej założenie, o tym iż model będzie się zbliżać do poprawnej formy jest prawidłowe.

4.3. Współliniowość

Z założenia zmienne objaśniające powinny być ze sobą nieskorelowane. W praktyce nie jest to wykonalne, gdyż rzeczywiste dane zawsze w pewnym stopniu są ze sobą skorelowane. Dlatego istotne jest, aby to zjawisko występowało w jak najmniejszym stopniu. Wobec czego posłużono się statystyką Variance Inflation Factor (VIF), opartą na statystyce R^2 , która jest ważnym elementem pomagającym w zidentyfikowaniu niepożądanych zmiennych.

Tabela 9. zawiera wszystkie zmienne użyte w regresji bez zmiennej *age2* ze względu na jej bezsporną współliniowość ze zmienną *age* (z powodu podniesienia jej do drugiej potęgi).

TABELA 9. Statystyka VIF dla Modelu 3.

Variable	VIF	1/VIF
goals	5.13	0.194857
assists	6.93	0.144353
yellow	3.34	0.299244
matches	10.43	0.095871
Amatches	2.09	0.478713
age	1.98	0.505807
1.represent	1.99	0.503371
1.english	1.54	0.649718
leg		
1	1.18	0.844914
2	1.36	0.732953
height	1.21	0.829104
Injuries	1.65	0.607263
lsocial	2.42	0.412588
ratingWho	2.99	0.334389
lteamvalue	2.65	0.377469
Mean VIF	3.13	

Źródło: Obliczenia własne za pomocą programu STATA

Zgodnie z teorią wyniki zmiennych poddanych statystyce VIF, przekraczające 10, należy uznać za „podejrzane” i rozważyć usunięcie ich z modelu (źródło). Dla zmiennej *matches* rezultat statystyki VIF jest najwyższy i wynosi 10.43, wobec czego zmienna ta minimalnie przekracza wartość „krytyczną” (10). Ze względu na jej ważne znaczenie w całości

4.4 Obserwacje odstające

W pierwszej kolejności wykonano wykres zależności między dźwignią a wystandaryzowanymi resztami podniesionymi do kwadratu, który przedstawia obserwacje, które odstają od większości. Z reguły najbardziej niepokojące są obserwacje znajdujące się w prawym górnym rogu tego typu wykresu.

Scatter plot showing the relationship between 'Dźwięnia' (Y-axis) and 'Kwadraty znormalizowanych reszt' (X-axis). The Y-axis ranges from 0 to 0.6, and the X-axis ranges from 0 to 0.1. A vertical red line is drawn at approximately x = 0.01, and a horizontal red line is drawn at approximately y = 0.15. Several data points are labeled with numbers. Three points are circled in red: 66, 54, and 110.

Analizując Rysunek 3. zdecydowano się rozważyć usunięcie obserwacji zaznaczonych w czerwonym kółkach, tj. nr: 54, 66 oraz 110. Obserwacja nr 66 cechuje się wysoką wartością dźwigni oraz jest nietypowa ze względu na zmienną objaśniającą. Z kolei obserwacja nr 110 cechuje się wysoką wartością wystandaryzowanych reszt podniesionych do kwadratu, nie powoduje jednak znacznych zmian w oszacowanych parametrach. Obserwacja nr 54 ma

zarówno stosunkowo wysoką dźwignię, jak i wartość wystandaryzowanych reszt podniesionych do kwadratów.

TABELA 10. Analizowane obserwacje, odnoszące się do Rysunku 3.

<i>nr</i>	<i>surname</i>	<i>value</i>	<i>age</i>	<i>Injuries</i>	<i>Amatches</i>	<i>Agoals</i>	<i>social</i>
54	Diangana	1500000	20	0	17	3	2
66	Defoe	2000000	36	64	7	0	1442000
110	Diakhaby	8000000	22	6	7	0	389

Źródło: Obliczenia własne za pomocą programu STATA

W konsekwencji usunięto obserwacje nr 54 oraz 110, które wydają się nieprawdopodobne ze względu na zmienną *social*, która w tym przypadku przyjmuje bardzo małe wartości. Obserwacja nr 66 zostaje w modelu, chociaż wydaje się być nietypowa, lecz nie powinna być błędna, gdyż piłkarze w wieku 36 lat często kontynuują kariery.

Następnie wyliczono studentyzowane reszty. W Tabeli 11. wyszczególniono obserwacje przekraczające wartość $> |2|$.

TABELA 11. Wysokie reszty studentyzowane wraz z charakterystykami danej obserwacji.

<i>nr</i>	<i>surname</i>	<i>value</i>	<i>age</i>	<i>Injuries</i>	<i>Amatches</i>	<i>Agoals</i>	<i>social</i>	<i>rstud</i>
6	Origi	10 000 000	23	26	4	2	1670000	-2.202
22	Moses	18 000 000	28	51	6	0	3030000	2.10177
48	Yarmolenko	20 000 000	29	41	10	2	300000	2.60631
59	Success	5 000 000	22	36	19	3	60000	-2.3395
66	Defoe	2 000 000	36	64	7	0	1442000	2.04339
96	Ward	1 750 000	28	29	10	1	80000	-2.3462
106	Barnes	8 000 000	29	43	22	3	29	2.08959

Źródło: Obliczenia własne za pomocą programu STATA

Niektóre z powyższych wydają się być błędne ze względu na nietypowe wartości dla zmiennej *social* oraz dla zmiennej *Injuries*. Pomimo tego, że część z tych obserwacji jest prawdopodobnie prawidłowa, zdecydowano się usunąć te przekraczające wartość $> |2|$, ze względu na potrzebę uzyskania wiarygodnych wyników modelu.

Po odrzuceniu obserwacji odstających kolejny raz przeprowadzono statystykę VIF, aby upewnić się, że nie spowoduje to działania powstania współliniowości.

TABELA 12. Statystyka VIF po odrzuceniu obserwacji odstających.

Variable	VIF	1/VIF
goals	5.63	0.177567
assists	7.02	0.142353
yellow	3.73	0.268066
matches	11.78	0.084855
Amatches	2.71	0.369208
age	2.13	0.468647
1.represent	1.85	0.541206
1.english	1.52	0.658961
leg		
1	1.19	0.840566
2	1.41	0.709267
height	1.22	0.820321
Injuries	1.67	0.600536
lsocial	2.95	0.338859
ratingWho	3.27	0.305632
lteamvalue	3.23	0.309456
Mean VIF	3.42	

Źródło: Obliczenia własne za pomocą programu STATA

VIF dla zmiennej *matches* wynosi 11.78. Pomimo tego, że wartość ta wzrosła porównując z poprzednim wyliczeniem, to jest to dalej niewiele ponad wartość niepokojącą – 10. Wobec tego zmienna pozostaje w modelu ze względu na to, że jest istotna.

4.5. Ostateczna estymacja, interpretacja parametrów

Po „oczyszczeniu” zbioru danych przystąpiono do ostatecznej estymacji modelu, w której swój udział miały wcześniejsze trzy pomocnicze estymacje. Statystyki R^2 modelu zmieniały się według następującej kolejności:

- Model 1: $R^2 = 84.21\%$, R^2 dopasowane = 78.89%;
- Model 2: $R^2 = 87.72\%$, R^2 dopasowane = 83.57%;
- Model 3: $R^2 = 86.79\%$, R^2 dopasowane = 84.74%;
- Model 4: $R^2 = 91.61\%$, R^2 dopasowane = 90.19%.

Ostateczny model wyjaśnia zmienną zależną w 90.19% (biorąc pod uwagę R^2 dopasowane) W Tabeli 13. została zaprezentowana całkowita estymacja niniejszego modelu.

TABELA 13. Ostateczna estymacja (Model 4).

Source	SS	df	MS	Number of obs	=	111
Model	104.174036	16	6.51087725	F(16, 94)	=	64.18
Residual	9.53674671	94	.101454752	Prob > F	=	0.0000
				R-squared	=	0.9161
				Adj R-squared	=	0.9019
Total	113.710783	110	1.03373439	Root MSE	=	.31852

lvalue	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
goals	.0059912	.0018177	3.30	0.001	.002382	.0096003
assists	.0074785	.004723	1.58	0.117	-.0018991	.0168562
yellow	.0100513	.0054031	1.86	0.066	-.0006768	.0207794
matches	-.004091	.0011094	-3.69	0.000	-.0062937	-.0018883
Amatches	.0361261	.0091201	3.96	0.000	.0180179	.0542343
age	.5265648	.1608685	3.27	0.001	.2071567	.8459729
age2	-.0116059	.0030018	-3.87	0.000	-.017566	-.0056458
1.represent	.2777856	.0935722	2.97	0.004	.0919958	.4635753
1.english	.2308481	.0846385	2.73	0.008	.0627963	.3988999
leg						
1	-.1920377	.0792584	-2.42	0.017	-.3494071	-.0346683
2	.3006288	.1486111	2.02	0.046	.005558	.5956996
height	.0122917	.0049612	2.48	0.015	.0024412	.0221422
Injuries	-.0021788	.0013269	-1.64	0.104	-.0048134	.0004559
lsocial	.1005035	.0217828	4.61	0.000	.0572533	.1437538
ratingWho	.7086653	.1292935	5.48	0.000	.4519501	.9653806
lteamvalue	.3347285	.0792847	4.22	0.000	.1773068	.4921502
_cons	-4.659886	3.238687	-1.44	0.154	-11.09037	1.770602

Źródło: Obliczenia własne za pomocą programu STATA

Model przyjmuje następującą formułę matematyczną: $lvalue = \beta_1 + \beta_2*goals + \beta_3*assists + \beta_4*yellow + \beta_5*matches + \beta_6*Amatches + \beta_7*age + \beta_8*age2 + \beta_9*1.represent + \beta_{10}*1.english + \beta_{11}*1.leg + \beta_{12}*2.leg + \beta_{13}*height + \beta_{14}*Injuries + \beta_{15}*lsocial + \beta_{16}*ratingWho + \beta_{17}*lteamvalue + \epsilon$.

W konsekwencji otrzymano następujące współczynniki dla zmiennych ostatecznego modelu:

TABELA 14. Parametry uzyskane w regresji modelu ostatecznego (Model 4).

zmienna	współczynnik
goals	0.00599**
assists	0.00748
yellow	0.0101
matches	-0.00409***
Amatches	0.0361***
age	0.527**
age2	-0.0116***
1.represent	0.278**
1.english	0.231**
1.leg	-0.192*
2.leg	0.301*
height	0.0123*
Injuries	-0.00218
Isocial	0.101***
ratingWho	0.709***
lteamvalue	0.335***
_cons	-4.66

Oznaczenia tabeli: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Źródło: Obliczenia własne za pomocą programu STATA

Następnie przystąpiono do interpretacji poszczególnych zmiennych (*ceteris paribus*):

- *goals* – jeśli zawodnik strzeli gola, to spowoduje tym samym wzrost swojej wartości o 0.599%;
- *assists* – jedna asysta więcej zintensyfikuje wartość zawodnika o 0.748% (zmienna nie jest istotna statystycznie);
- *yellow* – wzrost liczby żółtych kartek o jednostkę oznacza wzrost wartości zawodnika o 1.01% (zmienna nie jest istotna statystycznie);
- *matches* – jeden rozegrany mecz więcej powoduje spadek wartości o 0.409%;
- *Amatches* – jeden rozegrany mecz więcej (w obecnym sezonie) powiększy wartość o 3.61%;
- *age* – jeżeli zawodnik jest starszy o rok, to jego wartość spada o 8.27%;
- *represent* – reprezentowanie państwa na arenie międzynarodowej powoduje wzrost wartości zawodnika o 32%;
- *english* – posiadanie paszportu angielskiego intensyfikuje wartość piłkarza o 26%;
- *leg_1* – jeżeli zawodnik jest lewonożny jego wartość jest niższa o 17.5% w porównaniu gdyby był prawonożny;
- *leg_2* – jeżeli zawodnik potrafi grać jedną i drugą nogą na takim samym poziomie, jego wartość wzrasta o 35.1%, w porównaniu gdyby był tylko prawonożny;
- *height* – gdy zawodnik jest wyższy o centymetr jego wartość wzrasta o 1.23%;
- *Injuries* – jeden mecz więcej opuszczony z powodu kontuzji oznacza spadek wartości zawodnika o 0.218% (zmienna nie jest istotna statystycznie);
- *social* – jedno procentowy skok ilości obserwujących w social mediach powoduje wzrost wartości zawodnika o 0.101%;
- *ratingWho* – wzrost o jedną ocenę w rankingu WhoScored powoduje wzrost wartości zawodnika o 70.9%;

- *teamvalue* - większa wartość drużyny o 1% jest utożsamiana z wzrostem wartości zawodnika o 33.5%;
- *stała* – nieistotna statystycznie.

Zmienne: *red*, *rankingFIFA*, *positionPL*, *Agoals*, *Aassists*, *lteamsocial*, *strong*, *position*, *youth* zostały odrzucone we wcześniejszym etapie estymacji ze względu na ich nieistotność oraz „osłabianie” modelu.

ROZDZIAŁ V

DIAGNOSTYKA MODELU OSTATECZNEGO

W niniejszym rozdziale przeprowadzono testy, mające na celu weryfikację założeń KMRL. Kolejno sprawdzano założenia o:

- poprawności funkcyjnej modelu;
- homoskedastyczności składnika losowego;
- normalności reszt.

Ze względu na fakt, iż model zawiera jedynie dane przekrojowe, a tym samym nie obejmuje próby panelowej, nie wykonano testów na brak autokorelacji.

5.1. Test formy funkcyjnej

TABELA 15. Test RESET dla ostatecznego modelu.

Ramsey RESET test using powers of the fitted values of lvalue

Ho: model has no omitted variables

$F(3, 91) = 0.57$

Prob > F = 0.6345

Źródło: Obliczenia własne za pomocą programu STATA

Statystyka F testu RESET wynosi 0.57, a jego p-value 0.6345 (>0.05 , wobec czego nie można odrzucić hipotezy zerowej). Model spełnia założenie o poprawności funkcyjnej.

5.2. Testy sferyczności rozkładu reszt

W pierwszej kolejności przeprowadzono mniej restrykcyjny test na homoskedastyczność – test White’a.

TABELA 16. Test White’a dla ostatecznego modelu.

White's test for Ho: homoskedasticity

against Ha: unrestricted heteroskedasticity

$\chi^2(110) = 111.00$

Prob > χ^2 = 0.4554

Źródło: Obliczenia własne za pomocą programu STATA

Statystyka χ^2 testu White’a wynosi 111, a jego p-value 0.4554 (>0.05 , wobec czego nie można odrzucić hipotezy zerowej). Według tego testu model spełnia założenie o homoskedastyczności. Niemniej jednak przystąpiono również do bardziej restrykcyjnego testu – Breuscha–Pagana.

TABELA 17. Test Breuscha-Pagana dla ostatecznego modelu.

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity

Ho: Constant variance

Variables: fitted values of lvalue

chi2(1) = 3.03

Prob > chi2 = 0.0816

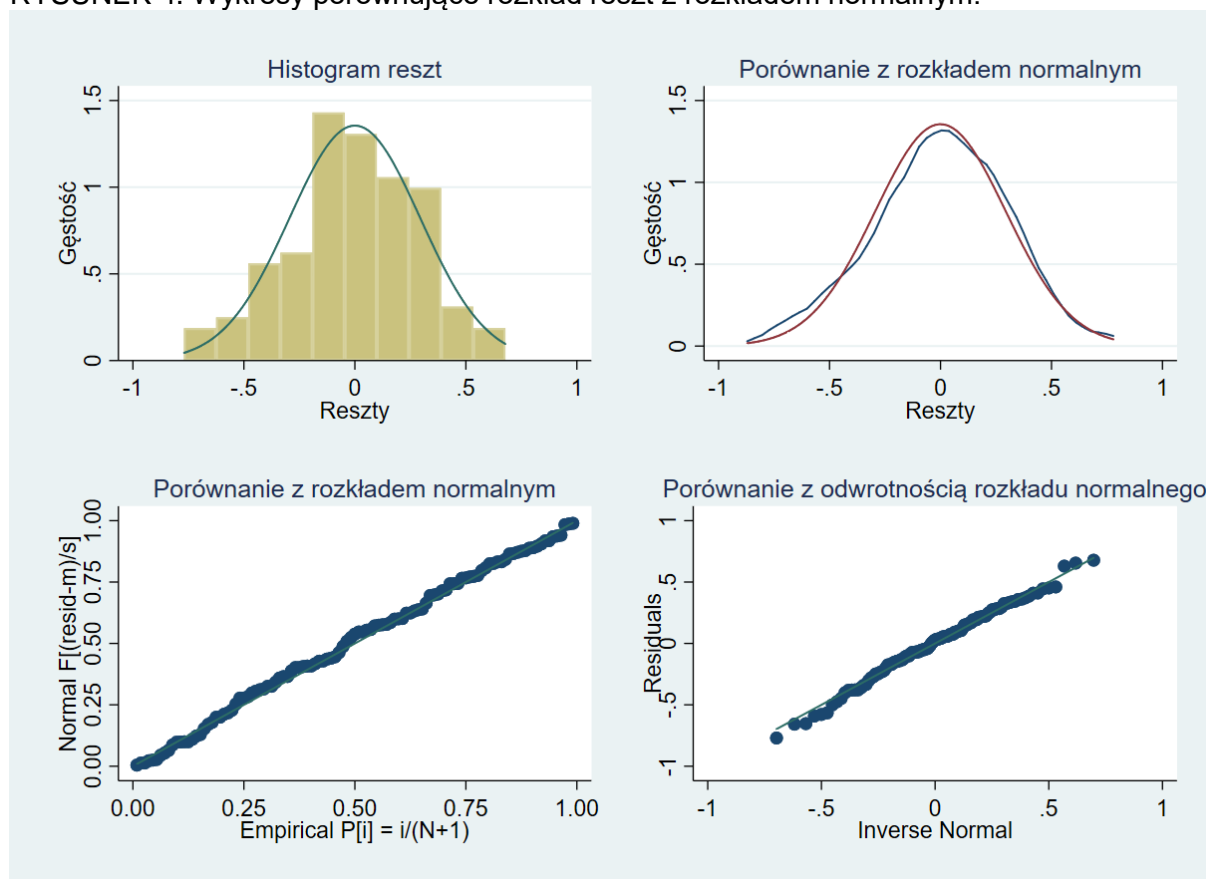
Źródło: Obliczenia własne za pomocą programu STATA

Statystyka Chi2 wyniosła 3.03, p-value 0.0816 (również powyżej poziomu istotności 0.05). Wobec czego kolejny test potwierdził homoskedastyczność składnika losowego.

5.3. Testy normalności rozkładu reszt

Na koniec przystąpiono do weryfikacji normalności reszt. Na Rysunku 4. można zauważyć cztery wykresy, badające powyższe założenie.

RYSUNEK 4. Wykresy porównujące rozkład reszt z rozkładem normalnym.



Źródło: Wykresy wykonane za pomocą programu STATA

Pomimo faktu, że każdy z wykresów potwierdza założenie o normalności rozkładu reszt przystąpiono do ostatniego testu – Jarque-Bera.

Tabela 18. Test Jarque-Bera dla ostatecznego modelu.

Skewness/Kurtosis tests for Normality

Variable	Obs	Pr (Skewness)	Pr (Kurtosis)	adj chi2 (2)	Prob>chi2
-----+-----					
reszty	111	0.3175	0.9529	1.02	0.6000

Źródło: Obliczenia własne za pomocą programu STATA

Test Jarque-Bera potwierdził założenie o normalności rozkładu reszt. P-value tego testu wyniosło 0.6, co jednoznacznie wskazuje na powyższy fakt (0.6 większe od przyjętego poziomu istotności – 0.05).

ROZDZIAŁ VI

Weryfikacja hipotez badawczych

W niniejszym rozdziale zweryfikowano hipotezy postawione na początku pracy. W pierwszej części zaprezentowano te hipotezy, które znalazły odzwierciedlenie w wynikach. Z kolei w drugiej przedstawiono hipotezy, które nie zostały spełnione, a także interpretację, z jakiego powodu taki wynik mógł mieć miejsce.

6.1. Potwierdzenie hipotez badawczych

Zmienne które dla poniższych podgrup zostały spełnione według teorii:

- **osiągnięć indywidualnych** – wzrost liczby goli w całej karierze oznacza wzrost wartości zawodnika;
- **agresji** – zmienne oznaczające liczbę żółtych oraz czerwonych kartek okazały się nieistotne;
- **doświadczenia** – zarówno wzrost liczby meczów, jak i wzrost wieku oznacza spadek wartości rynkowej zawodnika. Zmienna rozróżniająca piłkarzy młodzieżowych nie jest istotna. Zaś zmienna określająca występ w reprezentacji seniorskiej zgodnie z założeniem intensyfikuje wartość rynkową piłkarza;
- **pochodzenia** – zmienna określająca kontynent okazała się nieistotna;
- **charakterystyk boiskowych** – pozycja okazała się nieistotną zmienną;
- **popularności** – zgodnie z hipotezą wzrost liczby obserwujących w social mediach podwyższa wartość zawodnika;
- **ocen** – wzrost punktów w rankingu przełożył się na wyższą wartość zawodnika;
- **siły reprezentacji** – zmienna określająca pozycję w rankingu państwa, z którego pochodzi zawodnik okazała się nieistotna;
- **siły klubu** – pozycja klubu, a także występowanie w silnym klubie okazały się nieistotne.

6.2. Zaprzeczenie hipotezom badawczym, próba interpretacji

Zmienne które dla poniższych podgrup nie zostały spełnione w nawiązaniu do teorii:

- **osiągnięć indywidualnych** – gole z obecnego sezonu, a także asysty zarówno z obecnego sezonu, jak i całej kariery okazały się nieistotne. W kwestii nieistotności asyst, prawdopodobnie wyszedł tenże wynik, ze względu na analizę jednej formacji – napastników, którzy przede wszystkim mają za zadanie strzelać gole. W przypadku badań rozszerzonych o formację pomocy – liczba asyst okazuje się istotna. Z kolei nieistotność zmiennych określających aktualny sezon można wytłumaczyć wymaganą regularnością piłkarza w strzelaniu bramek oraz notowaniu asyst. Obecny sezon nie jest wystarczający, aby w ogólności podnieść znacząco wartość takiego zawodnika;
- **pochodzenia** - posiadanie paszportu angielskiego okazało się istotne i podwyższa wartość piłkarza. Wynik ten „kłóci się” z rezultatem pracy - *No Pain, No Gain: Effort and Productivity in Professional Soccer*, gdzie paszport niemiecki nie wpływał na wartość piłkarza w lidze niemieckiej (Wicker & Prinz, & Weimar & Deutscher & Upman 2013). Przede wszystkim należy wziąć pod uwagę fakt, iż rynek angielski

różni się od niemieckiego. Wyniki niniejszej pracy mogą wskazywać, że angielskie kluby preferują Anglików nawet jeśli są w niewielkim stopniu gorsi od obcokrajowców;

- **charakterystyk boiskowych** – zmienna określająca lepszą nogę zawodnika okazała się istotna, gdyż wskazywała, że piłkarze z dominującą lewą nogą są mniej warci na rynku oraz piłkarze z obiema nogami na równym poziomie są więcej warci, w porównaniu do piłkarzy prawonożnych. Należy rozważyć, że wpływ na powyższy rezultat może mieć to, że większość piłkarzy jest prawonożna. Niemniej jednak jest to bardzo interesujące zjawisko i warte analizy na większej próbie – rozszerzonej o pozostałe formacje. Z kolei interpretacja piłkarzy obunożnych nie jest niespodziewana, gdyż taka zaleta jest pożądana przez kluby;
- **parametrów fizycznych** – wzrost zawodnika okazał się zmienną istotną oraz w konsekwencji podwyższa wartość rynkową zawodnika. W przytoczonym we wcześniejszej części badaniu *No Pain, No Gain: Effort and Productivity in Professional Soccer* wzrost nie jest istotny, prawdopodobnie ze względu na badanie wszystkich formacji piłkarskich (Wicker & Prinz, & Weimar & Deutscher & Upman 2013). W przypadku tylko napastników wzrost podnosi wartość ze względu na dodatkową zaletę – lepszą umiejętność gry głową. Również należy wziąć pod uwagę, że liga angielska jest dużo bardziej „fizyczna” od niemieckiej, a wzrost ułatwia odnalezienie się w takich warunkach;
- **kontuzji** – wbrew hipotezie liczba meczów opuszczonych z powodu kontuzji okazała się zmienną nieistotną. Jest to zaskakujący wynik, jednakże należy rozważyć fakt, że w dzisiejszych czasach dużo łatwiej dochodzi się do pełnej sprawności po kontuzji, co może powodować, że kluby potrafią zapłacić dużo pieniędzy, niezależnie od ryzyka poniesienia kolejnej kontuzji przez zawodnika;
- **popularności klubu** – przyrost liczby obserwujących klub w social mediach okazał się zmienną nieistotną. W obecnych czasach piłka nożna stała się w zdecydowanie większym stopniu sportem indywidualnym, co może przekładać się na to, iż tylko popularność zawodnika podwyższa jego wartość, a nie popularność klubu.

ZAKOŃCZENIE

Podsumowując, niniejsza praca potwierdziła, iż wycena wartości rynkowej piłkarza jest niezwykle złożoną miarą. Pomimo tego, iż zdawać by się mogło, że uwzględniono w modelu wszystkie zmienne, to jest z pewnością stwierdzenie na wyrost. Piłka nożna stała się w ostatnich latach tak złożonym zjawiskiem, że z pewnością kolejni badacze będą odkrywać i opisywać coraz to nowsze rezultaty oraz wnioski swoich złożonych analiz.

Niezależnie od faktu, że modele wyceny wartości rynkowych są dosyć powszechne, niewątpliwie udało się odkryć część ciekawych zależności. Między innymi wpływ lepszej nogi piłkarza na jego wycenę, co do tej pory w bardzo niewielu pracach było brane pod uwagę. Oprócz tego należy zwrócić uwagę na niezwykle interesującą zależność pomiędzy pochodzeniem piłkarza, a jego wartością. Powszechna opinia panująca w świecie piłki nożnej głosi, iż z każdym rokiem omawiana dyscyplina sportu jest coraz bardziej skomercjalizowana. W świetle takich głosów normalnym byłoby, iż paszport angielski nie powinien „dawać” korzyści, jednakże w wyniku niniejszej próby udało się zweryfikować ten pogląd, wobec czego na pozycji napastnika Anglik jest więcej warty od obcokrajowca w lidze angielskiej.

Niezaprzeczalnie, wartym uwagi aspektem w kontekście przyszłych badań jest umiejętne rozszerzenie analizy o więcej formacji, a także o więcej lig. Jednakże należy być ostrożnym, gdyż mniej popularne ligi są zaniedbane w kontekście wartości rynkowych piłkarzy, co w kontekście analizy, mogłoby przynieść wnioski, które byłyby niemal całkowicie niezgodne z prawdziwym stanem.

Perspektywa przyszłych badań jest również interesująca, gdyż niniejsza praca została sporządzona z użyciem Klasycznego Modelu Regresji Liniowej. Z całą pewnością warto byłoby skonfrontować powyższe wyniki oraz przyszłe badania z bardziej rozbudowanymi ekonometrycznymi, a w dalszej perspektywie predykcyjnymi modelami.

BIBLIOGRAFIA

Artykuły oraz Modele:

- Adler M. *Stardom and Talent*. Columbia University. New York 2006.
- Carmichael F., Thomas D. *Bargaining in the Transfer Market: Theory and Evidence*. Applied Economics., s. 1467-1476. 1993.
- Dobson S., Gerrard B. *The Determination of Player Transfer Fees in English Professional Soccer*. Journal of Sport Management. nr 13(4)., s. 259-279. 1999.
- Franck E., Nüesch S. *TALENT AND/OR POPULARITY: WHAT DOES IT TAKE TO BE A SUPERSTAR?* Institute for Strategy and Business Economics, University of Zürich. Zürich 2010.
- He M., Cachucho R., Knobbe A. *Football Player's Performance and Market Value*. LIACS, Leiden University, the Netherlands. Amsterdam 2015.
- He Y. *Predicting Market Value of Soccer Players Using Linear Modeling Techniques*. UC Berkeley Statistics. California 2014.
- Lehmann E., Schulze G. *What does it take to be a star? The role of performance and the media for German soccer players*. Discussion Paper Series, No. 1, University of Freiburg, Department of International Economic Policy (iep), Freiburg 2007.
- Lucifora C., Simmons R. *Superstar Effects in Sport*. Catholic University of Milan, Italy. Lancaster University, United Kingdom. Journal of Sports Economics 2003.
- Majewski S. *Identification of Factors Determining Market Value of the Most Valuable Football Players*. Uniwersytet Szczeciński. Wydział Nauk Ekonomicznych i Zarządzania. Szczecin 2015.
- Margonis A. *THE DETERMINANTS OF NBA PLAYER SALARIES: A REGRESSION ANALYSIS*. Athens University of Economics. Athens 2017.
- Rottenberg S. *The Baseball Players' Labor Market*. Journal of Political Economy Vol. 64, nr 3., s. 242-258. Chicago 1956.
- Tunaru R., Clark E., Viney H. *An Option Pricing Framework for Valuation of Football Players*. Business School. Middlesex University. London NW4 4BT. UK. Londyn 2005.
- Wicker, P., Prinz, J., Weimar, D., Deutscher, C., Upman, T. *No Pain, No Gain: Effort and Productivity in Professional Soccer*. International Journal of Sport Finance, 8(2), 124-139. 2013.
- Williams R. *Outliers*. University of Notre Dame. Notre Dame 2016.

Książki:

- Buschmann R., Wulzinger M. *Football Leaks: Die schmutzigen Geschäfte im Profifußball*. Spiegel 2018.

Strony internetowe:

Dawson A. *Juventus reportedly sold \$60 million worth of Ronaldo jerseys in 24 hours — almost half his transfer fee.* Business Insider 2018. On line. Dostęp: 28 grudnia 2018. <https://www.businessinsider.com/juventus-sold-60-million-in-cristiano-ronaldo-jerseys-in-24-hours-2018-7?IR=T>

Goal. *What is Financial Fair Play and how does it work? FFP rules explained.* Goal.com 2018. On line. Dostęp: 28 grudnia 2018. <https://www.goal.com/en/news/what-is-financial-fair-play-and-how-does-it-work-ffp-rules/1ihlynh8s59i319l6nxx1z6kg5>

Shergold A. *Juventus agree new deal with kit makers Adidas worth £46m-a-season over eight years from 2019 as promotional video features Paul Pogba celebration.* DailyMail 2018. On line. Dostęp: 28 grudnia 2018. <https://www.dailymail.co.uk/sport/football/article-6521097/Juventus-agree-new-deal-kit-makers-Adidas-worth-46m-season.html>

Transfermarkt. *Premier League Marktwerte.* Transfermarkt.de 2018. On line. Dostęp: 28 grudnia 2018. <https://www.transfermarkt.de/premier-league/marktwerte/wettbewerb/GB1>

Transfermarkt. *Transferrekorde.* Transfermarkt.de 2018. On line. Dostęp: 28 grudnia 2018. <https://www.transfermarkt.de/statistik/transferrekorde>