

Erasmus Universiteit Rotterdam

Erasmus School of Economics

Bachelor scriptie Economie en Bedrijfseconomie

Thuisvoordeel in de Nederlandse Eredivisie tijdens de COVID-19 pandemie 2018-2021

Een onderzoek naar het effect van de aanwezigheid van publiek tijdens voetbalwedstrijden in de Nederlandse Eredivisie gedurende de seizoenen 2018 t/m 2021.

Tom van Oeffel

15-2-2022

510352to

Begeleider: Jan van Ours

Tweede beoordelaar: Sam Hoey

Het geschrevene in deze scriptie is de opvatting van de auteur en niet noodzakelijk die van de begeleider, tweede beoordelaar, Erasmus School of Economics of Erasmus Universiteit Rotterdam.

Inhoud

Inleiding	2
Bestaande literatuur.....	3
Het thuisvoordeel	3
Afwezigheid van fans	4
Data en methodologie.....	6
Dataset.....	6
Variabelen.....	6
Afhankelijke variabele.....	6
Onafhankelijke variabele	7
Controle variabelen.....	7
Beschrijvende statistieken	9
Methodologie	10
Gevoeligheidsanalyse	10
Resultaten	11
Gevoeligheidsanalyse	12
Conclusie	13
Bibliografie	14

Inleiding

Met ongeveer 3,5 miljard fans is voetbal de grootste sport ter wereld. De sport heeft niveaus reikend van de kelderklasse tot de Champions League en er worden jaarlijks ongelooflijk veel wedstrijden gespeeld. Dit spektakel trekt dan ook een hoop toeschouwers. Bij verschillende topclubs kunnen tot wel honderdduizend mensen in een stadion een voetbalwedstrijd aanschouwen. Toeschouwers die je zullen prijzen als je kampioen wordt en beledigen wanneer je degradeert. Wetende dat er duizenden mensen aanwezig zijn die je nauwlettend in de gaten houden, kan de druk voor de spelers erg hoog doen oplopen. Dit is dan ook een factor die impact kan hebben op het spel van de spelers.

In december 2019 brak wereldwijd het coronavirus uit, ofwel COVID-19. In de paar maanden daarna raakten er miljoenen mensen besmet, waarvan er meer dan 1,4 miljoen overleden. Dit resulteerde in ernstige sociale beperkingen over de hele wereld. Mensen moesten binnen blijven, alle grote evenementen werden geannuleerd en publiek werd verboden bij sportwedstrijden. Op deze manier werden de grootste voetbalwedstrijden gespeeld zonder de aanwezigheid van publiek. In de Nederlandse Eredivisie werd in het seizoen 2019-2020 door de uitbraak van het coronavirus besloten dat de competitie niet werd uitgespeeld. Daarnaast werden er in het seizoen daarna maar drie speelrondes gespeeld onder de aanwezigheid van publiek. Deze drie speelrondes werden gespeeld met een maximale bezettingsgraad van ongeveer 25% (Transfermarkt). Normaliter worden er echter sporadisch wedstrijden zonder publiek gespeeld. Daarom creëert de uitbraak van COVID-19 een soort natuurlijk experiment. Dit schept een unieke mogelijkheid om te onderzoeken wat de impact van publiek is op voetbalwedstrijden.

In de voetballerij wordt het begrip “Thuisvoordeel” gezien als een gegeven. Thuisvoordeel betekent dat wanneer team x tegen team y speelt, de kans groter is dat x in het eigen stadion wint dan in het stadion van team y. Er is reeds veel onderzoek gedaan naar de aanwezigheid van thuisvoordeel in de grootste competities van Europa. Swartz & Barsky (1997) waren een van de eerste onderzoekers die lieten zien dat thuisvoordeel bestaat in de voetballerij. Echter bestaat er vandaag de dag nog wel veel discussie over de factoren die dit thuisvoordeel drijven. Aan de hand van de wedstrijden zonder publiek kan er onderzocht worden wat de impact van het publiek is op dit thuisvoordeel. Daarom staat in dit onderzoek de volgende onderzoeksvraag centraal:

Wat is het effect van de afwezigheid van voetbalfans tijdens wedstrijden, op het thuisvoordeel in Nederlandse Eredivisie in de jaren 2018-2019?

Om de onderzoeksvraag te kunnen beantwoorden moet er eerst beter uitgelegd worden wat er in dit onderzoek mee bedoeld wordt met thuisvoordeel. In dit paper zullen we de impact van het thuispubliek op de volgende factoren meten: het aantal doelpunten wat de thuisploeg scoort, het aantal tegendoelpunten wat de thuisploeg incasseert en het aantal competitiepunten wat de thuisploeg verdient aan een wedstrijd. Hierbij zullen we rekening houden met andere factoren die hier invloed op kunnen hebben op het thuisvoordeel, zoals de kwaliteit van de teams. Om het effect van fans op de genoemde factoren van thuisvoordeel te schatten zullen er verschillende regressieanalyses worden uitgevoerd.

In het volgende hoofdstuk zal eerst de bestaande literatuur besproken worden waar dit paper aan bijdraagt. Vervolgens zullen de data en methodologie verder toegelicht worden. Daarna zullen de regressie analyses worden uitgevoerd en de resultaten hiervan worden besproken. Tot slot zal er in het laatste hoofdstuk een conclusie gevormd worden op de hoofdvraag.

Bestaande literatuur

Het thuisvoordeel in professioneel voetbal is een gegeven maar er bestaat nog wel veel discussie over de aard van dit voordeel. Er is reeds veel onderzoek gedaan naar de oorzaken van het thuisvoordeel in het hedendaagse voetbal. In dit hoofdstuk gaan we dieper in op de bestaande literatuur waar dit paper aan bijdraagt. Dit zal betrekking hebben op de factoren die het thuisvoordeel drijven en daarnaast op de impact van het uitblijven van voetbalfans door COVID-19.

Het thuisvoordeel

Dit paper draagt bij aan veel bestaande literatuur waarin verschillende redenen voor het thuisvoordeel worden onderzocht. Richard Pollard (1986) was een van de eersten die onderzoek deed naar het thuisvoordeel in voetbal. Pollard (1986) onderzocht variaties in het thuisvoordeel in de grootste vier Engelse voetbalcompetities. Hij concludeerde dat er in de hoogste Engelse competitie, sinds de oprichting in 1888 tot 1984, maar weinig variatie in het thuisvoordeel zit. Pollard (1986) bespreekt in zijn onderzoek meerdere factoren die invloed kunnen hebben op het thuisvoordeel. Hierbij concludeert hij dat de factoren, support van het publiek en vermoeidheid van de tegenstander (door het reizen), weinig bijdragen aan het

thuisvoordeel. De impact van andere factoren zoals de vertrouwdheid met de omgeving en scheidsrechter bias, zijn volgens Pollard (1986) nog niet duidelijk.

Swartz & Barsky (1997) bevestigen in hun onderzoek het bestaan van thuisvoordeel in verschillende sporten. Zij dragen echter bij aan de bestaande literatuur dat het thuisvoordeel groter is voor binnensporten (zoals ijshockey en basketbal) dan voor buitensporten (zoals honkbal en voetbal). Daarnaast beargumenteren ze dat het thuisvoordeel met name veroorzaakt wordt door een hogere aanvallende efficiëntie en niet door een defensief voordeel. Tot slot dragen Swartz & Barsky (1997) bij dat de steun die de thuisploeg krijgt van het publiek de grootste drijfveer is van het thuisvoordeel.

Naast factoren als grootte van het thuispubliek, vertrouwdheid met de omgeving en vermoeidheid van de tegenstander wordt in de bestaande literatuur ook scheidsrechter bias gezien als een van de drijfveren van thuisvoordeel. Boyko et al. (2007) onderzochten verschillende factoren die het thuisvoordeel kunnen beïnvloeden in de Engelse Premier League. Aan de hand van een OLS regressie onderzochten ze de impact van het publiek en de scheidsrechterbias op het thuisvoordeel, terwijl ze controleerden voor de kwaliteit van teams. Ze stellen dat er variaties bestonden in de beslissingen van de scheidsrechters met betrekking tot penalties en gele kaarten. Op basis daarvan concluderen ze dat scheidsrechterbias gedeeltelijk verantwoordelijk is voor thuisvoordeel.

Tot slot onderzochten Ponzo & Scoppa (2018) derbywedstrijden tussen teams die in hetzelfde stadion voetballen. Op basis van dit onderzoek konden ze bevestigen dat de vermoeidheid van de tegenstander geen impact heeft op het thuisvoordeel. Daar voegen ze echter nog aan toe dat ook vertrouwdheid met de omgeving geen impact heeft op het thuisvoordeel. Op basis hiervan worden in dit paper deze factoren buiten beschouwing worden gelaten. In dit paper wordt er enkel gefocust op de impact van het publiek op het thuisvoordeel. Eventuele andere factoren die impact kunnen hebben, zoals scheidsrechterbias, worden ook buiten beschouwing gelaten.

Afwezigheid van fans

Het probleem met het onderzoeken van de impact van het publiek op het thuisvoordeel was dat er geen controlegroep bestond. Voetbalwedstrijden werden namelijk maar sporadisch zonder publiek gespeeld. In dit opzicht heeft COVID-19 een unieke mogelijkheid gebracht om

dit beter te onderzoeken. Doordat er vanuit veiligheidsoverwegingen geen publiek meer welkom was bij wedstrijden, werd er een soort natuurlijk experiment gecreëerd. Sinds de uitbraak van de corona pandemie is er dan ook al veel onderzoek gedaan naar de impact van het uitblijven van publiek op het thuisvoordeel. Er werden meerdere onderzoeken gedaan naar de verschillende factoren die het thuisvoordeel kunnen beïnvloeden.

Ten eerste, onderzochten Bryson et al. (2020) de impact van het uitblijven van publiek op de beslissingen van de scheidsrechters in 23 verschillende competities, gedurende het seizoen 2019-2020. Ze concluderen dat er geen verandering bestaat in de eindresultaten van de wedstrijden. Daarentegen vinden ze wel bewijs om te kunnen concluderen dat er zonder publiek minder gele kaarten werden uitgedeeld aan de tegenstander. Bryson et al (2020) stellen dat, doordat er nu helemaal geen publiek is wat de beslissingen van de scheidsrechter kan beïnvloeden, er een causaal verband bestaat tussen het spelen van wedstrijden zonder publiek en de daling in scheidsrechterbias. Daarnaast onderzochten Reade et al. (2020) of het ontbreken van de druk van het publiek impact had op het gedrag en de uitkomsten van wedstrijden. Dit paper sluit goed aan op het onderzoek van Reade et al. (2020). Ze gebruiken in hun onderzoek ELO ratings om te controleren voor de kwaliteit van teams. Ze onderzochten het uitblijven van fans in verschillende Italiaanse competities, UEFA-competities en de Franse Ligue 1. Op basis van een OLS regressie vonden ze significante resultaten dat het uitblijven van fans een negatieve impact heeft op de winkansen van de thuisploeg en de doelpunten van de thuisploeg, in alle competities. Daarnaast concluderen ze dat er een daling heeft plaatsgevonden in de scheidsrechterbias door het spelen van wedstrijden zonder publiek.

Tot slot, bevestigen Fisher & Haucap (2021), in hun onderzoek naar het thuisvoordeel gedurende de zogeheten "Ghost Games" in de Duitse Bundesliga, de bevindingen van Reade et al (2020). Fisher & Haucap (2021) onderzochten aan de hand van een OLS regressie, de impact van het publiek op de winstkans van het thuisteam. Ze hanteerden hierbij een dummy variabele voor publiek en controleerden voor een set van controle variabelen (waaronder bijvoorbeeld kwaliteit). Ze concluderen dat er door het uitblijven van het publiek er een daling van de winratio's van het thuisteam heeft plaatsgevonden. Daarnaast onderzochten ze ook de scheidsrechterbias bij deze wedstrijden, maar ze kunnen niet genoeg bewijs vinden om te stellen dat er wisselingen in de besluitvorming van de scheidsrechters hebben plaatsgevonden. Daarmee is de algehele conclusie dat het uitblijven van publiek een negatieve

impact heeft op het thuisvoordeel in de grote Europese voetbalcompetities. Dit paper draagt hieraan bij door deze impact te onderzoeken voor de Nederlandse Eredivisie.

Data en methodologie

In dit hoofdstuk wordt de dataset besproken. Vervolgens wordt er uitgelegd hoe de data gebruikt wordt om een antwoord te formuleren op de onderzoeksvraag. Tot slot zullen er beschrijvende statistieken worden getoond om een beter beeld te formuleren van de dataset.

Dataset

De dataset bestaat uit data die is verzameld over de Eredivisie seizoenen 18-21. Gedurende het seizoen 2020-2021 was er, omwille van het coronavirus, bij 31 speelrondes geen publiek welkom gedurende deze wedstrijden. Op basis van deze wedstrijden zonder publiek kunnen we schatten wat de impact hiervan was op het thuisvoordeel.

Het seizoen 2019-2020 is, omwille van de coronapandemie, echter niet volledig afgespeeld en stopgezet na speelronde 26. Een mogelijk probleem wat zich voor doet, is dat nog niet elk team in eigen huis te maken heeft gehad met dezelfde tegenstanders. Daarom kunnen sommige bestaande methodes om het thuisvoordeel te schatten, die gegevens over een geheel seizoen nodig hebben, niet gebruikt worden. Om dit probleem te verhelpen is de dataset per wedstrijd.

De dataset bestaat uit 844 observaties over drie verschillende seizoenen. Per seizoen worden er 306 wedstrijden gespeeld in de Eredivisie. Zoals gesteld werd het seizoen 2019-2020 niet volledig gespeeld en daarom bevat de dataset 232 observaties over dit seizoen.

Variabelen

De modellen die geschat zullen worden om een antwoord te vormen op de deelvragen bestaan uit verschillende variabelen. De variabelen zijn onderverdeeld in afhankelijke-, onafhankelijke- en controlevariabelen. We zullen nu deze variabelen uitgebreider bespreken. De controlevariabelen zijn verkregen van clubelo.com (Clubelo) en alle andere variabelen zijn verkregen van Flashscore.nl (Flashscore).

Afhankelijke variabele

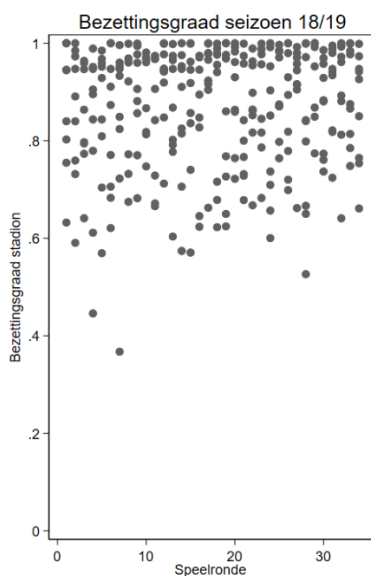
Zoals eerder gezegd zullen we drie verschillende afhankelijke variabelen gebruiken om het effect van publiek op het thuisvoordeel te schatten. Deze variabelen zijn het aantal

doelpunten wat de thuisploeg scoort, het aantal tegendoelpunten wat de thuisploeg incasseert en het aantal competitiepunten wat de thuisploeg verdient aan een wedstrijd. Er zullen drie verschillende regressies worden uitgevoerd met deze afhankelijke variabelen, waarna de resultaten zullen worden vergeleken om een antwoord te vormen op de onderzoeksvraag.

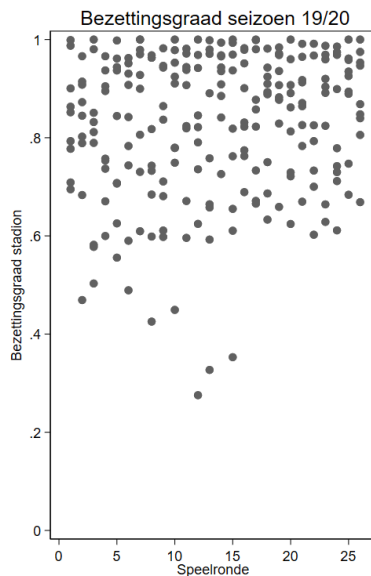
Onafhankelijke variabele

De onafhankelijke variabele in alle drie de modellen is de bezettingsgraad van het stadion. Deze variabele is berekend door het aantal toeschouwers per wedstrijd te delen door de capaciteit van het stadion. Zoals gezegd zijn in het seizoen 2020-2021, 31 van de 34 speelrondes gespeeld zonder de aanwezigheid van de fans. De onderstaande grafieken illustreren de hoeveelheid fans die aanwezig waren bij wedstrijden gedurende de drie verschillende seizoenen in de dataset.

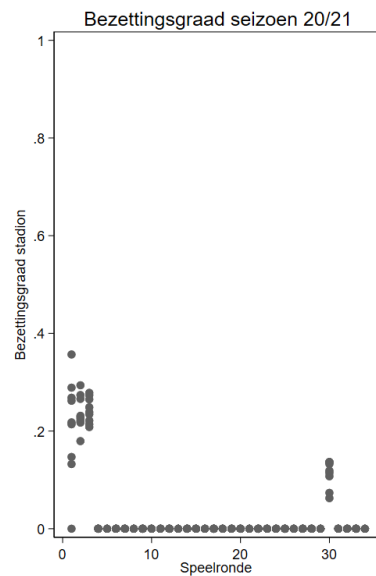
Figuur 1a: bezettingsgraad 18/19



Figuur 1b: bezettingsgraad 19/20



Figuur 1c: bezettingsgraad 20/21



Notities: de figuren 1 a t/m c geven, per seizoen, een scatterplot weer van de bezettingsgraad van het stadion per speelronde.

Controle variabelen

Een belangrijke factor die een rol speelt bij de prestaties van voetbalteams is de kwaliteit van het team. Slechtere teams winnen vanzelfsprekend minder vaak en eindigen daarom ook lager op de ranglijst. We kunnen dus concluderen dat de kwaliteit van een team impact heeft op de verschillende afhankelijke variabelen die we zullen gebruiken in onze regressieanalyse.

Daarnaast heeft ook de kwaliteit van de tegenstander impact op het aantal goals wat de thuisploeg scoort in een gegeven wedstrijd.

Om te kunnen controleren voor de impact van de kwaliteit van beide teams op de verschillende afhankelijke variabelen, moet de kwaliteit onderdeel zijn van het model. Als proxy voor kwaliteit zullen de ELO ratings van teams voorafgaand aan een wedstrijd gebruikt worden.

De ELO rating is een numerieke waarde die de kwaliteit van een voetbal team op elk gegeven moment weergeeft, hierbij geldt hoe hoger hoe beter. Deze waarde wordt berekend aan de hand van de historische en huidige resultaten van een team. Hierbij wegen de huidige resultaten zwaarder dan de historische resultaten.

De ELO rating is direct verbonden met de winkansen van de twee teams die tegen elkaar spelen. De ELO vergelijking is als volgt (Clubelo):

$$E = 1 / (10^{(-dr/400)} + 1)$$

Hierbij is “dr” het verschil in de ELO ratings van de twee teams. Vervolgens wordt aan de hand van E de verandering van de ELO rating na een wedstrijd berekend aan de hand van de volgende formule:

$$\Delta ELO = (R - E) * 20$$

Hierbij is R het eindresultaat van een wedstrijd. R neemt de waarde 1 aan bij winst, 0,5 bij een gelijkspel en 0 bij verlies. Aan de hand van deze formules worden de ELO rating voor de verschillende teams in een competitie berekend en gemuteerd na afloop van een wedstrijd. Daarnaast kunnen factoren zoals het doelpunten-verschil en thuisvoordeel ook meewegen in de verandering van de ELO ratings. De waarde van deze ratings zijn verkregen via clubelo.com (Clubelo) en zijn per wedstrijd voor beide teams genoteerd.

De ELO ratings van de teams nemen grote waardes van tussen de 1000 en 2000 aan. Een mogelijk probleem hierbij is dat de resultaten van de regressieanalyse hele kleine waardes weergeven. Daardoor is de impact van de ELO ratings minder makkelijk is af te lezen. Om dit probleem te verhelpen zijn alle ELO ratings in de dataset gedeeld door 100.

Beschrijvende statistieken

De onderstaande tabel geeft de beschrijvende statistieken van de verschillende variabelen in de dataset, per wedstrijd weer.

Tabel 1: Beschrijvende statistieken

Variabele	Gemiddelde	Verschil ten opzichte van tegenstander
<i>Doelpunten thuisteam</i>	1,8	0,4
<i>Doelpunten tegenstander</i>	1,4	-0,4
<i>Punten thuisteam</i>	1,6	0,5
<i>Punten tegenstander</i>	1,1	-0,5
<i>ELO rating thuisteam</i>	1449.8	0,1
<i>ELO rating tegenstander</i>	1449.7	-0,1
<i>Win ratio thuisteam</i>	47,4%	16,9%
<i>Win ratio tegenstander</i>	30,5%	-16,9%

Notities: de bovenstaande tabel geeft de beschrijvende statistieken voor de verschillende variabelen in het regressiemodel weer voor het thuisteam en hun tegenstander. De linker kolom geeft de gemiddelde waarde weer en de rechter kolom het verschil ten opzichte van de tegenstander.

Tabel 1 laat zien dat er verschillen zitten in de gemiddelde prestaties van uit- en thuisteam. De thuisteam scoren gemiddeld 0,4 doelpunt en verdienen 0,5 competitiepunt meer per wedstrijd, ten opzichte van hun tegenstanders. Dit terwijl de gemiddelde kwaliteit van de teams en hun tegenstanders ongeveer gelijk is. Daarmee lijken deze statistieken erop te wijzen dat er een voordeel bestaat voor de thuisspelende ploeg.

De onderstaande tabel 2 laat zien dat er een duidelijk verschil bestaat tussen de prestaties van teams tussen wedstrijden met en zonder publiek. De tabel laat zien dat de thuisteam aanzienlijk vaker winnen wanneer er wel publiek aanwezig is ten opzichte van wedstrijden zonder publiek. Op basis van de onderstaande gegevens lijken de fans dus een impact te hebben op het eindresultaat van een wedstrijd.

Tabel 2: Beschrijvende statistieken

Variabele	Gemiddelde met fans	Gemiddelde zonder fans	Verschil
<i>Doelpunten thuisteam</i>	1,9	1,6	0,3
<i>Doelpunten tegenstander</i>	1,3	1,4	-0,1
<i>Punten thuisteam</i>	1,7	1,5	0,2
<i>Punten tegenstander</i>	1,1	1,3	-0,2
<i>Win ratio thuisteam</i>	51,0%	39,9%	11,1%-punt
<i>Win ratio tegenstander</i>	28,8%	33,9%	5,1%-punt

Notities: de bovenstaande tabel geeft de gemiddelde waarden van de variabelen weer, met en zonder publiek. De eerste kolom geeft de ratio's met publiek weer, de tweede kolom de ratio's zonder publiek en de derde kolom laat het verschil tussen kolom 1 en 2 zien.

Methodologie

Om het effect van het uitblijven van fans in te schatten zouden we idealiter elke wedstrijd onder dezelfde omstandigheden moeten herhalen, mét en zonder fans, en de verschillen hiertussen vergelijken. Dit is echter niet realistisch, daarom wordt voor wedstrijd i de volgende OLS-regressie geschat om inzicht te geven in dit effect.

$$\text{Afhankelijke variabele } i = \alpha + \beta_1(\text{bezettingsgraad } i) + \beta_2(\text{ELO thuisteam } i) + \beta_3(\text{ELO tegenstander } i) + \varepsilon_i$$

Aan de hand van dit model worden er drie regressies geschat met elk hun eigen afhankelijke variabele. Elke regressie heeft de constante α_i en controle variabelen voor kwaliteit van de thuis- en uitploeg. Tot slot zal β_1 de impact van het publiek weergeven op het thuisvoordeel.

Gevoeligheidsanalyse

Om de regressieresultaten te controleren worden er twee gevoeligheidsanalyses uitgevoerd. Deze worden uitgevoerd door het regressiemodel een klein beetje aan te passen en te kijken of de resultaten hierop reageren.

Allereerst zal voor, wedstrijd i , de volgende regressie worden uitgevoerd:

$$\text{Afhankelijke variabele } i = \alpha + \beta_1(\text{fans } i) + \beta_2(\text{ELO thuisteam } i) + \beta_3(\text{ELO tegenstander } i) + \varepsilon_i$$

Hierbij is de nieuwe onafhankelijke variabele, fans, een dummy variabele die de waarde 1 geeft wanneer er fans aanwezig waren en 0 wanneer dit niet het geval was. Deze resultaten tonen of de impact van helemaal geen fans groter is dan variaties in de bezettingsgraad (zoals in het seizoen 2020-2021).

Ten tweede, wordt er een OLS-regressie uitvoeren met team fixed effects. Hierbij zullen we dezelfde OLS-regressie uitvoeren die we gebruiken om het effect van publiek op de wedstrijdresultaten te schatten, we zullen hier echter nog team fixed effects aan toevoegen. Deze regressie is, voor team i en wedstrijd t , als volgt:

$$\text{Afhankelijke variabele } it = \alpha_i + \gamma_t + \beta_1(\text{bezettingsgraad } it) + \beta_2(\text{ELO thuisteam } it) + \beta_3(\text{ELO tegenstander } xt) + \varepsilon_{it}$$

Door team fixed effects toe te voegen kunnen we zien of er vaste factoren zijn die impact hebben op het thuisvoordeel. Denk hierbij aan structurele verschillen tussen teams zoals de intensiteit van de fans of de afstand van het publiek tot het veld. Ten tweede, kan er uit deze resultaten afgelezen worden of de variabelen, die controleren voor de kwaliteit van de teams, worden weggelaten vanwege collineariteit. Dit zou betekenen dat kleine variaties in de kwaliteit van een team, zoals het hebben van een slechte dag, niet belangrijk zijn. Maar dat de impact van de variabele komt door verschillen tussen de teams.

Resultaten

In de onderstaande tabel staan de regressieresultaten voor de drie verschillende regressies, met elk hun eigen afhankelijke variabelen. Deze resultaten zijn per wedstrijd en zijn afgerond op drie decimalen.

Tabel 3: Regressie resultaten

Variabelen	Aantal thuisdoelpunten (1)	Aantal tegendoelpunten (2)	Aantal competitiepunten (3)
<i>Bezettingsgraad</i>	0,465***	-0,039	0,311***
<i>ELO thuisteam</i>	0,357***	-0,243***	0,319***
<i>ELO tegenstander</i>	-0,279***	0,318***	-0,300***
<i>Constante</i>	0,439	0,667	1,194*

*Notities: in deze tabel worden de regressie resultaten weergegeven. De drie kolommen bevatten de regressieresultaten van een eenvoudige OLS regressie te zien met de drie verschillende afhankelijke variabelen. De sterretjes geven het significantieniveau bij de kritieke waarden aan: * is 0.1, ** is 0.05 en *** is 0.01.*

De resultaten in de kolommen 1 en 3 laten een significante positieve impact van de aanwezigheid van publiek op de prestaties van de thuisploeg zien. Er wordt bij een bezettingsgraad van 100% ongeveer 0,5 doelpunt meer gescoord per wedstrijd ten opzichte van wedstrijden in een leeg stadion. Dat betekent dat de gemiddelde Eredivisieclub ongeveer 7 doelpunten meer scoort in thuiswedstrijden wanneer het stadion het gehele seizoen vol heeft gezeten. Ten tweede wordt er bij een bezettingsgraad van 100% gemiddeld 0,3 competitiepunten per wedstrijd meer verdient. Dat betekent dat de gemiddelde Eredivisieclub ongeveer 5 competitiepunten ($0,3 \cdot 17$ thuiswedstrijden) meer verdient in thuiswedstrijden wanneer het stadion het gehele seizoen vol heeft gezeten ten opzichte van een geheel leeg stadion. Op basis van deze resultaten kunnen we concluderen dat, in de

seizoenen 2018-2021, een vol stadion een positieve impact had op de wedstrijdresultaten. Kortom, deze resultaten tonen aan dat er een thuisvoordeel bestond in de Eredivisie gedurende de desbetreffende seizoenen.

Daarnaast kunnen we op basis van de resultaten in kolom 2 concluderen dat het thuisvoordeel, dat gecreëerd wordt door de aanwezigheid van publiek, in een aanvallend voordeel zit. We kunnen namelijk geen significante impact van het publiek vinden op het aantal tegendoelpunten dat de thuisspelende partij incasseert. Daarmee zijn deze resultaten in lijn met die van Swartz & Barsky (1997), zij concludeerde al eerder dat het thuisvoordeel zat in hogere aanvallende efficiëntie.

Gevoeligheidsanalyse

De onderstaande tabel geeft de regressieresultaten weer voor de gevoeligheidsanalyse o.b.v. een eenvoudige OLS-regressie. De tabel geeft de waardes voor de drie verschillende regressies met hun eigen afhankelijke variabele weer. Hierbij is de afhankelijke variabele “Bezettingsgraad” echter vervuld voor een dummy variabele “Fans”. Deze variabele geeft de waarde 1 wanneer er fans aanwezig waren tijdens de wedstrijd en 0 wanneer dit niet zo was. De resultaten van de onderstaande tabel zijn per wedstrijd.

Tabel 4: resultaten gevoeligheidsanalyse

Variabelen	Aantal thuisdoelpunten (1)	Aantal tegendoelpunten (2)	Aantal competitiepunten (3)
<i>Fans</i>	0,331***	-0,090	0,270**
<i>ELO thuisteam</i>	0,357***	-0,244***	0,320***
<i>ELO tegenstander</i>	-0,272***	0,316***	-0,295***
<i>Constante</i>	0,372	0,384	1,093*

*Notities: in deze tabel worden de regressie resultaten van de gevoeligheidsanalyse weergegeven. De drie kolommen bevatten de regressieresultaten van een eenvoudige OLS-regressie met de drie verschillende afhankelijke variabelen. De sterretjes geven het significantieniveau bij de kritieke waarden aan: * is 0.1, ** is 0.05 en *** is 0.01.*

De bovenstaande regressieresultaten zijn in lijn met de resultaten uit tabel 3, het uitblijven van publiek heeft een negatieve impact op de prestaties van het thuisteam. Ook hier zijn de resultaten die wijzen op een hogere aanvallende efficiëntie statistisch significant. Daarnaast kunnen we geen bewijs vinden dat het publiek een negatieve impact heeft op het aantal tegendoelpunten wat geïncasseerd wordt. Wanneer we de resultaten van tabel 4 vergelijken

met die van tabel 3, kunnen we zien dat de impact van heel weinig fans ongeveer gelijk is aan helemaal geen fans. Op basis daarvan kunnen we concluderen dat ook variaties in de hoeveelheid publiek impact hebben op het thuisvoordeel.

De onderstaande tabel geeft de resultaten van de gevoeligheidsanalyse op basis van een OLS-regressie met team fixed effects weer.

Tabel 5: Resultaten gevoeligheidsanalyse met fixed effects

Variabelen	Aantal thuisdoelpunten (1)	Aantal tegendoelpunten (2)	Aantal competitiepunten (3)
<i>Bezettingsgraad</i>	0,426***	-0,165	0,322***
<i>ELO thuisteam</i>	-0,253	0,012	-0,157
<i>ELO tegenstander</i>	0,316**	-0,098	0,410***
<i>Contante</i>	0,639	2,868	-2,386

*Notities: in deze tabel worden de regressie resultaten van de gevoeligheidsanalyse met team fixed effects weergegeven. De drie kolommen bevatten de regressieresultaten van een eenvoudige OLS-regressie met de drie verschillende afhankelijke variabelen. De sterretjes geven het significantieniveau bij de kritieke waarden aan: * is 0.1, ** is 0.05 en *** is 0.01.*

De coëfficiënten van de bezettingsgraad voor de regressies 1 en 3 uit tabel 5 liggen erg dicht bij de resultaten uit tabel 3. Op basis van deze resultaten kan er geconcludeerd worden dat er geen structurele verschillen bestaan tussen de teams in de dataset. Daarnaast worden de controlevariabele voor kwaliteit niet buiten de regressie gelaten. Dit betekent dat variaties in de kwaliteit van de teams impact hebben op de afhankelijke variabelen in het model.

Conclusie

In dit onderzoek is onderzocht of het uitblijven van publiek bij Eredivisie wedstrijden, in de seizoenen 2018 t/m 2021, impact heeft gehad op het thuisvoordeel. De resultaten van de uitgevoerde OLS-regressie laten zien dat er een positieve correlatie bestaat tussen de aanwezigheid van het publiek en de variabelen die het thuisvoordeel definiëren. De centrale onderzoeksvraag van dit onderzoek is:

Wat is het effect van de afwezigheid van voetbalfans tijdens wedstrijden op het thuisvoordeel in Nederlandse Eredivisie in de jaren 2018-2021?

Op basis van de resultaten is de conclusie, dat dat de afwezigheid van publiek een negatieve impact heeft op de prestaties van het thuisteam. Daarmee is het antwoord op de centrale

onderzoeksvraag dat de afwezigheid van het publiek een negatieve impact heeft op het thuisvoordeel in de Eredivisie in de jaren 2018 t/m 2021. Dit is echter geen causaal verband omdat in dit onderzoek factoren als scheidsrechterbias buiten beschouwing worden gelaten.

Voor vervolgonderzoek zou het ideaal zijn om meer te focussen op de factoren die veranderen bij het uitblijven van publiek. Er zou bijvoorbeeld meer gefocust kunnen worden op de scheidsrechterbias of de individuele prestaties van spelers.

Tot slot kan er worden geconcludeerd dat, wanneer de Nederlandse KNVB en overheid interesse hebben in het zo min mogelijk beïnvloeden van de Eredivisie, het zaak is om zo snel mogelijk weer publiek te verwelkomen bij voetbalwedstrijden.

Bibliografie

Boyko, R. H., Boyko, A. R., & Boyko, M. G. (2007). Referee bias contributes to home

advantage in English Premiership football. *Journal of Sports Sciences*, 25(11), 1185–1194.

Bryson, A., Dolton, P., Reade, J. J., Schreyer, D., & Singleton, C. (2020). Causal effects of an

absent crowd on performances and refereeing decisions during Covid-19. *Economics Letters*, 198.

Fischer, K., & Haucap, J. (2021). Does Crowd Support Drive the Home Advantage in

Professional Football? Evidence from German Ghost Games during the COVID-19 Pandemic. *Journal of Sports Economics*, 22(8), 982–1008.

Flashscore. (z.d.). Flashscore.nl: *Live voetbaluitslagen, tussenstanden, Eredivisie stand*.

<https://www.flashscore.nl/>.

Pollard, R. (1986). Home advantage in soccer: A retrospective analysis. *Journal of Sports*

Sciences, 4(3), 237–248.

Ponzo, M., & Scoppa, V. (2018). *Does the home advantage depend on crowd support? Evidence*

from same-stadium derbies. *Journal of Sports Economics*, 19(4), 562–582.

Reade, J.J., Schreyer, D., Singleton, C., (2020). Echoes: what happens when football is played behind closed doors? Economics Discussion Papers em-dp2020-14, Department of Economics, Reading University.

Schwartz, B., & Barsky, S. F. (1977). The Home Advantage. *Social Forces*, 55(3), 641–661.

Schiefler, L. (z.d.). *Football Club Elo Ratings*. Clubelo. <http://Clubelo.com>.

Transfermarkt. (z.d.). *Voetbaltransfers, transfergeruchten, marktwaarden & nieuws*.
<https://www.transfermarkt.nl/>.