

# **Transferbedragen in het Nederlandse profvoetbal**

De transferwaarde van een voetballer statistisch gemodelleerd

BWI Werkstuk

Jarno Suijkerbuijk



# **Transferbedragen in het Nederlandse profvoetbal**

De transferwaarde van een voetballer statistisch gemodelleerd

BWI Werkstuk

Auteur: Jarno Suijkerbuijk  
Begeleider: Dennis Roubos

Vrije Universiteit Amsterdam  
Faculteit der Exacte Wetenschappen  
De Boelelaan 1081a  
1081 HV Amsterdam

Februari 2010



## Voorwoord

Een vast onderdeel van de master opleiding Business Mathematics and Informatics is het BWI werkstuk. Hierbij wordt zelfstandig onderzoek gedaan naar een praktische probleemstelling met wiskundige en informatica aspecten. In dit verslag is een probleemstelling uiteengezet en zijn de resultaten van het onderzoek beschreven.

Het doel van dit verslag is om de lezer te informeren over de factoren die het transferbedrag van een profvoetballer in Nederland bepalen en daarbij een vergelijking te geven met het Engelse voetbal. De lezer krijgt hierdoor een indruk welke factoren een profvoetballer in Nederland nou zo waardevol maakt en of er verschil is tussen de transfermarkten in Nederland en Engeland bij het bepalen van een transferbedrag.

Als laatste onderdeel van dit voorwoord wil ik graag van de gelegenheid gebruik maken om mijn begeleider Dennis Roubos te bedanken voor zijn hulp en ondersteuning tijdens het uitvoeren van mijn onderzoek en het schrijven van dit verslag. Ook wil ik mijn vader Jos Suijkerbuijk bedanken voor zijn voetbalarchief waar ik tijdens mijn onderzoek uitgebreid gebruik van heb mogen maken.

Jarno Suijkerbuijk  
Zoetermeer 2010



## Samenvatting

De vraag die aan de grondslag van dit werkstuk ligt is: hoe komen transferbedragen in het voetbal tot stand? Uit onderzoeken naar het Engelse voetbal is gebleken dat het transferbedrag van een speler bepaald wordt aan de hand van een aantal factoren. Omdat er alleen onderzoek gedaan is naar het Engelse voetbal is het interessant om te onderzoeken welke factoren in Nederland van belang zijn en om te vergelijken of ze overeenkomen met het Engelse voetbal. Hieruit volgt de probleemstelling van dit onderzoek:

*Welke factoren bepalen het transferbedrag van een profvoetballer in Nederland?*

In de voetbalwereld is er sprake van een transfer bij de overgang van een voetballer naar een andere club. Aan de ene kant is een transfer een onderhandelings-overeenkomst tussen twee clubs en aan de andere kant is het een overgang van een speler die van werkgever verandert. Aan de overeenkomst zit meestal ook een financiële vergoeding verbonden, het transferbedrag.

Carmichael en Thomas (1993) hebben als eerste onderzoek gedaan naar transferbedragen in het Engels profvoetbal. Zij gebruiken een onderhandelingstheorie waarmee ze uiteindelijk drie belangrijke categorieën van factoren onderscheiden die een rol spelen bij bepalen van de hoogte van een transferbedrag, dit zijn: spelers karakteristieken, karakteristieken van de verkopende club en karakteristieken van de kopende club. Reilly en Witt (1995) hebben het onderzoek van Carmichael en Thomas later uitgebreid door de factor huidskleur toe te voegen om racisme binnen het Engelse voetbal te onderzoeken. Gerrard en Howe (2000) hebben het onderzoek van Carmichael en Thomas uitgebreid door het onderzoek te doen voor het amateurvoetbal in Engeland. Hierbij introduceren ze een vierde categorie genaamd tijdseffecten.

De dataset voor het onderzoek bestaat uit 93 binnenlandse transfers van het Nederlandse profvoetbal afkomstig van het seizoen 2005/2006 tot en met het seizoen 2009/2010. Tijdens dit onderzoek zijn 52 factoren onderzocht, verdeeld volgens de vier categorieën die hierboven genoemd zijn. In bijlage A staan alle factoren in een overzicht. Van deze factoren is een meervoudig lineair regressie model opgesteld met het transferbedrag als afhankelijke variabele en de factoren als verklarende variabelen. Hierbij is gebruik gemaakt van het programma Microsoft Excel. Om het beste model te vinden is het volgende stappenplan gebruikt:

1. Met de Least Squares methode (LS) wordt een model gemaakt.
2. Om de fit van het model te bepalen wordt de determinatiecoëfficiënt berekend
3. Een F-toets voor het gehele model bepaald vervolgens of het model geldig is.
4. Aan de hand van een  $t$ -toets op de verklarende variabelen wordt bekeken welke variabelen (factoren) significant bijdragen aan het model.
5. Met de significante variabelen wordt vervolgens een nieuw model gemaakt, waarbij stappen 1 tot en met 4 herhaald worden.
6. Een F-toets bepaald of het model significant beter is dan zijn voorganger.
7. Stappen 1 tot en met 6 worden herhaald tot het beste model gevonden is.

Het uiteindelijke model bestaat uit 32 variabelen en heeft een fit van 71,65% op de gebaseerde data. Uit het model volgt dat het transferbedrag van een profvoetballer in

Nederland grotendeels bepaald wordt door 21 factoren, waaronder de positie van een speler op het veld, het gemiddeld aantal doelpunten per wedstrijd van een speler en in welke divisie de verkopende en kopende club spelen. De modellen uit de onderzoeken naar het Engelse voetbal bevatten een aantal dezelfde factoren, maar verschillen ook op een aantal factoren van het model voor het Nederlandse voetbal. Zo bevat het Engelse model onder andere ook de leeftijd van een speler en de positie van zowel de verkopende club als de kopende club op de ranglijst.

De resultaten van het onderzoek naar racisme zijn gelijk voor het Engelse en Nederlandse voetbal. In beide modellen spelen de racismefactoren geen significante rol. Er is dus geen bewijs gevonden voor racisme bij het bepalen van het transferbedrag voor een voetballer in Engeland en Nederland. Een interessante vervolgstudie op dit onderzoek is om hetzelfde onderzoek nogmaals te doen met een ander softwarepakket, omdat Excel moeite heeft om bij een meervoudig regressie model met veel variabelen tot een optimale oplossing te komen.



# Inhoudsopgave

Voorwoord .....	v
Samenvatting .....	vii
H1 Inleiding.....	1
H2 Achtergrond.....	3
2.1 Profvoetbal in Nederland.....	3
2.2 De transfermarkt.....	4
2.3 Eerder onderzoek naar transferbedragen.....	5
H3 Onderzoeksontwerp.....	7
3.1 De transferdata .....	7
3.2 De factoren .....	8
3.2.1 Spelers karakteristieken.....	8
3.2.2 Tijdseffecten.....	10
3.2.3 Karakteristieken van verkopende club .....	10
3.2.4 Karakteristieken van kopende club .....	11
3.3 Modellen en statistische methoden .....	12
3.3.1 Least Squares methode.....	13
3.3.2 De determinatiecoëfficiënt .....	13
3.3.3 De F-toets .....	14
3.3.4 De t-toets .....	15
H4 Resultaten .....	17
4.1 De onderzoeksresultaten .....	17
4.2 Vergelijking van resultaten .....	20
4.2.1 Resultaten van Carmichael en Thomas .....	20
4.2.2 Resultaten van Reilly en Witt.....	21
4.2.3 Resultaten van Dobson, Gerrard and Howe .....	21
H5 Conclusies en aanbevelingen.....	23
5.1 Conclusies .....	23
5.2 Aanbevelingen.....	24
Literatuurlijst.....	25
Bijlage A – Overzicht van factoren.....	27
Bijlage B – Het aangepaste model .....	31



# H1 Inleiding

Jaarlijks worden er in het profvoetbal enorme bedragen uitgegeven aan transferbedragen voor spelers. Zelfs in tijden wanneer er gesproken wordt over economische crisis legt een voetbalclub met gemak een paar miljoen euro op tafel voor een nieuwe speler. Zo kocht Ajax in aanvang op het seizoen 2009/2010 Demy de Zeeuw van AZ voor 8.000.000 miljoen euro en versterkte Feyenoord zich met Sekou Cissé van Roda JC voor het bedrag van 3.500.000 miljoen euro<sup>1</sup>. De vraag die aan de grondslag van dit werkstuk ligt is: hoe komen deze transferbedragen tot stand?

Wereldwijd is er nog niet veel studie gedaan naar de totstandkoming van transferbedragen voor voetballers. Alleen voor het Engelse voetbal zijn een aantal onderzoeken gedaan. Sporteconomen Carmichael en Thomas (1993) hebben de transfermarkt van het Engelse profvoetbal geanalyseerd volgens een onderhandelingstheorie waarbij ze gebruik maken van statistische modellen en methoden. Ze stellen dat een transferbedrag voor een voetballer tot stand komt door onderhandelingen tussen de huidige (oude) club van een speler en zijn aankomende (nieuwe) club. De hoogte van het bedrag is daarbij het resultaat van de onderhandelingen tussen beide clubs en verschilt per transfer van welke twee clubs erbij betrokken zijn en om welke speler het gaat. Per categorie (speler, verkopende club en kopende club) zijn een aantal factoren bepaald, waarvan een regressie model is gemaakt. Vervolgens is statistisch onderzocht welk van deze factoren een rol van betekenis spelen bij het bepalen van het transferbedrag van een speler.

Later hebben Reilly en Witt (1995) onderzocht of er sprake is van een racisme in het Engelse voetbal door het onderzoek van Carmichael and Thomas uit te breiden met de factor huidskleur. Ze hebben hierbij onderzocht of er voor spelers in het Engelse profvoetbal met een blanke huidskleur meer betaald wordt dan voor voetballers met een andere huidskleur. Ze vinden hiervoor uiteindelijk geen significant bewijs. Dobson, Gerrard en Howe (2000) hebben het onderzoek van Carmichael en Thomas herhaald voor de amateurtak van het Engelse voetbal. Hierbij zijn de factoren die het transferbedrag van een amateur voetballer bepalen vergeleken met de factoren van het profvoetbal. De conclusie van dit onderzoek is dat de factoren grotendeels overeenkomen.

Er is dus onderzoek gedaan naar de factoren die een rol spelen bij het bepalen van het transferbedrag van een speler, waarbij het onderzoek geconcentreerd is op het Engelse voetbal. Het is interessant om te onderzoeken welke factoren in Nederland van belang zijn en om te vergelijken of ze overeenkomen met het Engelse voetbal. Hieruit volgt de probleemstelling voor dit onderzoek:

*Welke factoren bepalen het transferbedrag van een profvoetballer in Nederland?*

Omdat er over het amateurvoetbal in Nederland niet veel data beschikbaar is, wordt het onderzoek alleen gedaan binnen het profvoetbal. Om antwoord te vinden op de probleemstelling is eerst verder uitgezocht hoe de Engelse onderzoekers te werk zijn gegaan en welke methoden zij precies gebruikt hebben. Ook is het Nederlandse profvoetbal en zijn transfermarkt kort beschreven. Vervolgens zijn de te onderzoeken factoren bepaald en is er een dataset samengesteld. Met een regressie model en statistische testen is daarna onderzocht

---

<sup>1</sup> De bedragen zijn gebaseerd op de data van de website: [www.transfermarkt.de](http://www.transfermarkt.de)

welke factoren een rol spelen bij het bepalen van een transferbedrag van een profvoetballer in Nederland.

Het doel van dit verslag is om de lezer te informeren over de factoren die het transferbedrag van een profvoetballer in Nederland bepalen en daarbij een vergelijking te geven met het Engelse voetbal. De lezer krijgt hierdoor een indruk welke factoren een profvoetballer in Nederland nou zo waardevol maakt en of er verschil is tussen de transfermarkten in Nederland en Engeland bij het bepalen van een transferbedrag.

Het verdere verslag ziet er als volgt uit: In hoofdstuk 2 volgt verdere achtergrondstudie naar de gedane onderzoeken in Engeland en het Nederlandse profvoetbal. Hoofdstuk 3 beschrijft het onderzoeksontwerp waarin de factoren, de gebruikte data, het model en de statistische methoden worden beschreven. De resultaten zijn vervolgens in hoofdstuk 4 beschreven en tot slot volgen in hoofdstuk 5 de belangrijkste conclusies en aanbevelingen.

## H2 Achtergrond

In dit hoofdstuk wordt ingegaan op de achtergrond van het onderzoek naar de factoren die het transferbedrag van een profvoetballer in Nederland bepalen. In de eerste paragraaf wordt een kort historisch overzicht gegeven van profvoetbal in Nederland. In de tweede paragraaf wordt de bijbehorende transfermarkt beschreven. Tot slot worden in de laatste paragraaf de Engelse onderzoeken naar transferbedragen verder uitgewerkt.

### *2.1 Profvoetbal in Nederland*

De term profvoetbal of professioneel voetbal staat voor betaald voetbal en onderscheidt zich van het amateur voetbal doordat profvoetballers fulltime in dienst zijn bij hun club en daarbij salaris ontvangen. In het amateurvoetbal worden voetballers soms ook betaald, maar dan is er meer sprake van een kleine vergoeding (bijvoorbeeld een vergoeding van de reiskosten). Vanaf 1954 wordt er in Nederland voor het eerst profvoetbal gespeeld onder de vlag van de Koninklijke Nederlandse Voetbal Bond (KNVB). Er wordt dan in een competitiestructuur gespeeld met 80 clubs verdeeld over drie divisies: de eredivisie, de eerste divisie en de tweede divisie. De eredivisie is het hoogste niveau en de tweede divisie het laagste. In 1971 is het Nederlandse profvoetbal grondig gesaneerd, waardoor vele profclubs ophielden te bestaan, met elkaar fuseerden of verder gingen bij de amateurs. Hierbij is destijds ook de tweede divisie opgeheven.

Vandaag de dag (seizoen 2009/2010) bestaat de eredivisie uit 18 clubs, waarvan de nummer 18 aan het eind van het seizoen degradeert naar de eerste divisie. De eerste divisie, door sponsoring tegenwoordig de Jupiler League geheten, bevat 20 clubs. Aan het eind van het seizoen promoveert de nummer een naar de eredivisie. Een seizoen van de eerste divisie is, in tegenstelling tot de eredivisie, opgedeeld in vier perioden. De vier periodekampioenen aangevuld met de vier hoogst geëindigde clubs uit de eerste divisie spelen aan het eind van het seizoen volgens een play-offs systeem promotie/degradatie duels met de nummers 16 en 17 van de eredivisie om twee plaatsen in de eredivisie.

Vanaf het seizoen 1971/1972 is het niet mogelijk geweest te degraderen uit de eerste divisie. Vanaf het huidige seizoen 2009/2010 is dit wel weer mogelijk. In eerste instantie is het de bedoeling geweest dat er aan het eind van dit seizoen twee clubs naar de nieuw opgezette Topklasse, de nieuwe hoogste amateurklasse, zouden degraderen. Door het faillissement van de club HFC Haarlem degradeert er nu nog maar een club. Het aantal clubs in de eerste divisie wordt hierdoor teruggebracht naar 18. Vanaf het seizoen 2010/2011 degradeert aan het eind van elk seizoen de nummer 18 uit de eerste divisie naar de Topklasse en krijgt de beste club uit de Topklasse de mogelijkheid te promoveren. Ziet deze club van promotie af, dan krijgt de op een na beste club de mogelijkheid. Als ook die club weigert, dan degradeert er geen club uit de eerste divisie.

## ***2.2 De transfermarkt***

In de voetbalwereld is er sprake van een transfer bij de overgang van een voetballer naar een andere club. De transfermarkt is een verzamelnaam voor alle activiteiten van clubs omtrent transfers. Carmichael en Thomas (1993) beschrijven twee hoofddoelen van een transfer:

1. Een club haalt nieuwe spelers om zijn team te versterken en de prestaties van het team te verbeteren.
2. Een speler verruult van club om sportieve en/of financiële verbetering.

Aan de ene kant is een transfer dus een onderhandelingsovereenkomst tussen twee clubs en aan de andere kant is het een overgang van een speler die van werkgever verandert. Bij een transfer zijn dan ook drie partijen betrokken. De speler onderhandelt met zijn nieuwe club over zijn contract en arbeidsvoorwaarden en de twee clubs onderhandelen over een overeenkomst voor de aan- en verkoop van de speler. Aan deze overeenkomst zit meestal ook een financiële vergoeding verbonden, het transferbedrag (of de transfersom). Dit bedrag is een vergoeding die de nieuwe club van de speler betaalt aan de oude club ter compensatie van de vertrokken speler. De hoogte van de vergoeding wordt tijdens de onderhandelingen vastgesteld en hangt af van de betrokken clubs en om welke speler het gaat.

Bij een transfer is enkel sprake van een transfersom als een speler niet transfervrij is. Een speler is transfervrij als zijn contract aan het einde van het seizoen afloopt. Deze vrijheid bij een transfer aan het einde van de contractperiode is het gevolg van het Bosman-arrest uitgesproken in december 1995 door het Europese Hof van Justitie. Voor het arrest diende er bij een transfer van een speler met aflopend contract toch een afkoopsom betaald te worden. Het arrest is vernoemd naar de Belgische voetballer Jean-Marc Bosman die zijn vrijheid opeiste toen zijn huidige club Club Luik een hoge transfersom voor hem vroeg, terwijl zijn contract dat seizoen afliep.

Naast een definitieve overgang van een speler kunnen de clubs ook voor een tijdelijke overgang kiezen. Hierbij wordt een speler voor een periode van een aantal maanden (vaak een jaar of half jaar) gehuurd. Voor deze constructie wordt vaak gekozen als een speler bij een club weinig speeltijd krijgt, maar de club hem in de toekomst nog nodig denkt te hebben. Een tijdelijke overgang is voor een club ook handig als het last heeft van vele blessuregevallen. Bij een verhuur wordt een kleiner bedrag betaald dan bij een definitieve overname. Meestal gaat het hierbij om een (gedeeltelijke) overname van het salaris van de speler, aangezien de speler nog wel onder contract staat bij de verhurende club.

Tijdens het seizoen zijn er twee periodes waarin spelers van club mogen wisselen, een periode in de zomer voor aanvang van het nieuwe seizoen en een periode in de winter halverwege het seizoen. Deze periodes wisselen wel eens van data, maar over het algemeen loopt de transferperiode in de zomer van 1 juli tot en met 31 augustus en in de winter van 1 januari tot en met 1 februari. Op elke andere data is een transfer van een profvoetballer die onder contract staat niet mogelijk. Alle binnenlandse transfers in Nederland moeten worden geregistreerd bij de KNVB, die hierbij als controlerend orgaan optreedt. Bij Europese transfers is dit de UEFA (Union of European Football Associations). De regels voor een transfer die in deze paragraaf zijn beschreven gelden zowel voor een transfer in Nederland als in Europa.

### ***2.3 Eerder onderzoek naar transferbedragen***

Zoals in hoofdstuk 1 beschreven is, is er nog weinig onderzoek gedaan naar transferbedragen en hoe deze tot stand komen. Al het onderzoek wat is uitgevoerd richt zich op het voetbal in Engeland. Dit komt omdat het Engelse voetbal wereldwijd het meest populair en best bekeken is. Over het Engelse voetbal is daarom ook het meeste data bekend, waardoor het toegankelijk is voor onderzoek. Carmichael en Thomas (1993) doen als eerste onderzoek naar transferbedragen in het Engels profvoetbal. Zij gebruiken een onderhandelingstheorie waarmee ze uiteindelijk drie belangrijke categorieën van factoren onderscheiden die een rol spelen bij bepalen van de hoogte van een transferbedrag, deze zijn:

1. Spelers karakteristieken
2. Karakteristieken van de verkopende club
3. Karakteristieken van de kopende club

Onder elke categorie specificeren ze vervolgens een aantal factoren. Bij de spelers karakteristieken zijn dit de factoren leeftijd, het aantal gespeelde wedstrijden in het seizoen voorafgaand aan de transfer en het aantal goals in dat laatste seizoen. Voor de positie van een speler op het veld gelden de vier factoren keeper, verdediger, middenvelder en aanvaller. De laatste spelers karakteristiek is een factor die aangeeft of het transferbedrag van de speler is bepaald door het sporttribunaal. Wanneer de clubs bij een transfer in Engeland niet tot een overeenkomst kunnen komen over de hoogte van het transferbedrag is er de mogelijkheid om het bedrag te laten vaststellen door het sporttribunaal. In Nederland bestaat deze optie ook, maar is minder gebruikelijk en komt daardoor ook maar zelden voor. Deze factor is daarom verder niet meegenomen in het onderzoek van dit BWI werkstuk. Voor de karakteristieken van de verkopende en kopende club worden dezelfde factoren gespecificeerd, dit zijn: de begroting van de club, het gemiddeld aantal toeschouwers in het seizoen voorafgaand aan de transfer, de positie van de club aan het eind van de competitie voorafgaand aan de transfer en het doelsaldo van de club in dat laatste seizoen.

Van deze factoren zijn vervolgens meerdere lineair regressiemodellen gemaakt met het transferbedrag als afhankelijke variabele en de factoren als verklarende variabelen. Er is onderzocht welke combinaties van factoren het beste resultaat geeft. Hierbij is gebruik gemaakt van een dataset die bestaat uit binnenlandse transfers in Engeland uit het seizoen 1990/1991. Aan de hand van enkele statistische testen is bepaald welk model het best is. Deze statistische testen en de regressiemodellen zijn in het volgende hoofdstuk verder beschreven.

Reilly en Witt (1995) hebben het onderzoek van Carmichael en Thomas later uitgebreid door de factor huidskleur toe te voegen om racisme binnen het Engelse voetbal te onderzoeken. Zij gebruiken hiervoor de Engelse binnenlandse transferdata uit het seizoen 1991/1992. Dobson, Gerrard en Howe (2000) hebben het onderzoek van Carmichael en Thomas uitgebreid door het onderzoek te doen voor het amateurvoetbal in Engeland. Zij gebruiken hiervoor de Engelse binnenlandse transferdata van de amateurs van 1988 tot en met 1997. Vanwege deze langere tijdsperiode komen zij tot het inzicht dat het uitmaakt in welke seizoenen de transfer plaatsvindt. Dit is ook niet zo een vreemde conclusie aangezien elk seizoen de transferbedragen weer verder lijken te stijgen. Daarnaast nemen zij ook een factor mee die aangeeft of een transfer plaatsvond voor de start van een nieuw seizoen of midden in het seizoen. Deze extra factoren vormen samen de vierde categorie genaamd tijdseffecten.

Zowel Reilly en Witt als Dobson, Gerrard en Howe gebruiken bij hun onderzoeken dezelfde aanpak als Carmichael en Thomas. In het volgende hoofdstuk zijn de methoden van Carmichael en Thomas verder uitgewerkt en is beschreven hoe deze methoden in het onderzoek toegepast zijn op data van de Nederlandse transfermarkt.



## H3 Onderzoeksontwerp

In dit hoofdstuk wordt het onderzoeksontwerp beschreven. In de eerste paragraaf wordt beschreven welke transferdata er voor het onderzoek is gebruikt. In de tweede paragraaf worden de factoren die bij het onderzoek zijn gebruikt uiteengezet. Tot slot worden in de derde paragraaf de gebruikte modellen en statistische methoden uitgewerkt.

### *3.1 De transferdata*

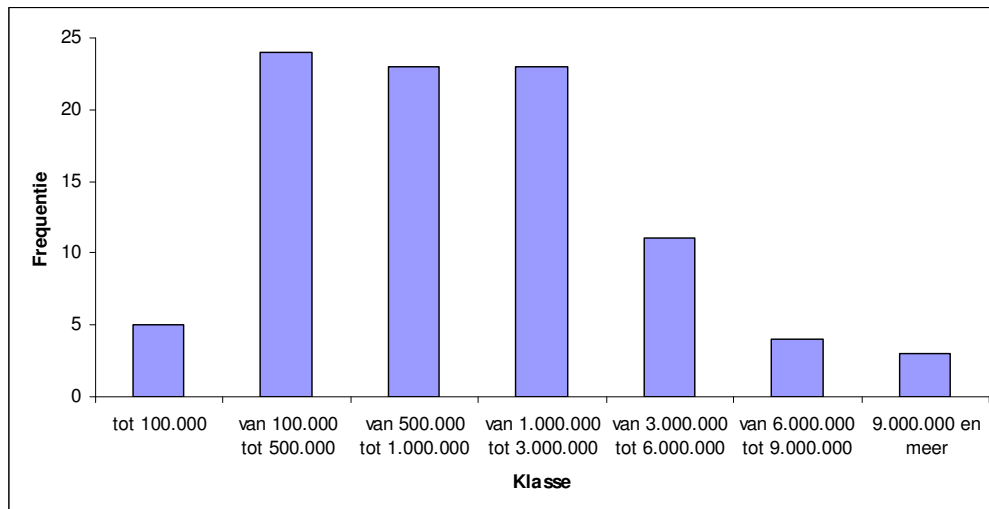
Voor het onderzoek is gebruikt gemaakt van de binnenlandse transferdata van het Nederlandse profvoetbal. In dit onderzoek gaat het om transfers van spelers die van een Nederlandse profclub naar een andere Nederlandse profclub gaan. De data is afkomstig van het huidige seizoen en de vier seizoenen ervoor, dus vanaf het seizoen 2005/2006 tot en met het seizoen 2009/2010. De data is gehaald van de websites en jaarboeken van de voetbaltijdschriften Voetbal International en Elf Voetbal. Daarnaast is gebruik gemaakt van het archief met oude voetbaltijdschriften van Jos Suijkerbuijk (de vader van de auteur). Als laatst is gebruik gemaakt van de website *www.transfermarkt.de*, waarop de transferbedragen terug te vinden zijn.

In de voetbalwereld en voornamelijk binnen Nederland is men niet zo scheutig met het naar buiten brengen van transferbedragen. Clubs zijn hier voorzichtig mee, omdat het hun onderhandelingspositie bij transfers in de toekomst kan beïnvloeden. Uiteraard worden de transferbedragen wel in de boekhouding opgenomen, maar de clubs zijn erg terughoudend met het naar buiten brengen van deze informatie. Dit hoeven zij ook niet te doen, alleen een beursgenoteerd bedrijf is verplicht om inzage in de boeken te geven. In Nederland is alleen de voetbalclub Ajax een beursgenoteerd bedrijf. Van de andere clubs in Nederland zijn alleen de bedragen bekend van transfers die in het nieuws zijn geweest. Vaak gaat het bij deze bedragen dan ook nog eens om een schatting van het echte bedrag. Het is voor buitenstaanders ook moeilijk om het precieze transferbedrag te schatten, aangezien er bij transfers vaak een aantal prestatie clausules worden afgesloten. Via deze clausules kan de vorige club van de speler achteraf extra geld ontvangen wanneer de speler of zijn nieuwe club binnen een voorafgestelde tijd aan bepaalde sportieve doeleinde voldoen (kampioen worden bijvoorbeeld). De transferbedragen in de dataset zijn daarom niet op de cent nauwkeurig, maar zijn wel consequent van dezelfde bron afkomstig waardoor de verhoudingen kloppen en de bedragen te gebruiken zijn voor wiskundig onderzoek.

In de uiteindelijke dataset zijn de transfers waarvan het transferbedrag ontbrak niet opgenomen. Ook de transfers waarvan de spelers een transfervrije status hadden en tijdelijke verhueringen van spelers zijn buiten de dataset gelaten. Van de ruim 2500 transfers van de Nederlandse profclubs in de gebruikte vijf seizoenen blijven uiteindelijk 93 transfers over die de dataset vormen voor dit onderzoek. Dat er maar zo weinig transfers overblijven heeft als hoofdredenen dat bijna de helft van de transfers spelers betreft die van/naar een buitenlandse club komen/vertrekken, er veel verhueringen van spelers tussen zitten en er maar weinig transferbedragen worden prijsgegeven door de clubs.

De duurste transfer in de dataset is de transfer van Miralem Sulejmani die voor aanvang van het seizoen 2008/2009 SC Heerenveen verruilde voor Ajax voor het bedrag van 16,25 miljoen euro. De goedkoopste transfer uit de dataset is die van Dominique van Dijk die voor aanvang

van het seizoen 2005/2006 RKC Waalwijk verruilde voor SC Cambuur (destijds nog Cambuur Leeuwarden geheten) voor het bedrag van 25 duizend euro. In Figuur 3.1 is een frequentietabel van de dataset opgenomen. Deze tabel laat het aantal transfer per prijsklasse van de dataset zien en daarmee de spreiding van de data. De meerderheid van de transferbedragen zitten tussen de 100 duizend euro en 6 miljoen euro met daarnaast een aantal uitschieters naar boven en naar beneden. Het gemiddelde transferbedrag is 1.8 miljoen euro.



Figuur 3.1: De frequentietabel van de dataset.

## 3.2 De factoren

De factoren die tijdens dit onderzoek zijn gebruikt zijn opgedeeld in de vier categorieën van Carmichael en Thomas (1993) en Dobson, Gerrard en Howe (2000): spelers karakteristieken, tijdseffecten, karakteristieken van de verkopende club en karakteristieken van de kopende club. Aan de hand van deze vier categorieën worden de factoren besproken in de volgende vier subparagrafen. Achter elke factor staat een afkorting tussen haakjes, hiermee wordt deze factor in het verdere onderzoek aangeduid. In bijlage A staan alle afkortingen van de factoren met hun uitleg in een overzicht. De factoren zijn in het vervolg van het onderzoek als verklarende variabelen in het regressie model gebruikt. Hierin zijn twee soorten variabelen te onderscheiden; continue variabelen die alle waarden kunnen aannemen en dummy variabelen die de waarden 0 en 1 kunnen aannemen. Bij de bespreking van de factoren geeft een sterretje (\*) achter de afkorting aan dat het om een dummy variabele gaat.

### 3.2.1 Spelers karakteristieken

De spelers karakteristieken bevatten factoren die de kenmerken en individuele prestaties van de spelers weergeven. Uit het onderzoek van Carmichael en Thomas (1993) zijn dit de factoren leeftijd (SLFT), het aantal gespeelde wedstrijden in het seizoen voorafgaand aan de transfer (SWEDCOM), het aantal goals in dat laatste seizoen (SDPCOM) en de positiefactoren keeper (SKPR\*), verdediger (SVERD\*), middenvelder (SMID\*) en aanvaller (SAANV\*). Daarnaast is de factor huidskleur (SBLANK\*), de racisme factor van het onderzoek van Reilly en Witt (1995), overgenomen. SBLANK heeft waarde 1 als de speler een blanke huidskleur heeft en waarde 0 als de speler een andere huidskleur heeft. Om het racisme aspect nog verder te onderzoeken is de factor nationaliteit (SNED\*) toegevoegd. Bij

racisme gaat het naast de huidskleur vaak ook om de afkomst en bovendien hebben niet alle buitenlanders een kleurtje. SNED heeft waarde 1 als de speler een Nederlandse nationaliteit heeft en waarde 0 als de speler een andere nationaliteit heeft.

Naast factoren uit bestaande onderzoeken zijn een aantal andere spelers karakteristieken in het onderzoek meegenomen die de individuele prestaties van een speler weergeven en waarvan het interessant is om te onderzoeken of ze ook een rol spelen bij het bepalen van het transferbedrag. Zo is naast SWEDCOM en SDPCOM ook onderzocht of het aantal wedstrijden en doelpunten in de beker in het seizoen voorafgaand aan de transfer (SWEDBEKER en SDPBEKER) en het totaal aan wedstrijden en doelpunten van de speler in zijn gehele professionele carrière (SWEDTOT en SDPTOT) een rol spelen. Ook de rol van het totaal aantal clubs waar een speler in zijn professionele carrière onder contract heeft gestaan (SCLUBS) is onderzocht. Een goede graadmeter van de aanvallende kwaliteit van een speler is het gemiddeld aantal doelpunten per wedstrijd (SD/W en SD/W TOT). Deze graadmeter is ook opgenomen in dit onderzoek. SD/W is daarbij het gemiddeld aantal doelpunten per wedstrijd in het seizoen voorafgaand aan de transfer en SD/W TOT het gemiddelde over de gehele professionele carrière van de speler voorafgaand aan de transfer.

Bijna net zo belangrijk als het maken van doelpunten voor een speler is het geven van assists (de laatste voorzet bij een doelpunt). Van assists is alleen niet zoveel data bekend als van doelpunten, daarom is alleen het aantal assists meegenomen dat een speler maakt in het seizoen voorafgaand aan de transfer (SASS). Als laatste zijn ook nog toegevoegd: het aantal gele en rode kaarten die de speler in het seizoen voorafgaand aan de transfer heeft ontvangen van de scheidsrechter (SGEEL en SROOD), het aantal interlands dat de speler in zijn gehele professionele carrière voor het hoogste nationale team heeft gespeeld (SINT) en hoe vaak een speler door zijn oude club is uitgeleend in het verleden (SLEEN). In Tabel 3.1 en 3.2 is een overzicht gegeven van de zojuist beschreven spelerskarakteristieken. Tabel 3.1 geeft een overzicht van de continue factoren met de gemiddelden, minima en maxima van de dataset. Tabel 3.2 geeft een overzicht van de dummy factoren met het aantal procenten 1'en en 0'en van de dataset. In Tabel 3.2 is te zien dat de dataset ongeveer evenveel verdedigers als aanvallers bevat, dat er iets meer middenvelders bijzitten en maar een enkele keeper.

Factor	Gemiddelde	Maximum	Minimum	Factor	1	0
SLFT	23,28	34,00	18,00	SKPR	2,15%	97,85%
SCLUBS	2,19	5,00	1,00	SVERD	29,03%	70,97%
SWEDTOT	117,54	419,00	2,00	SMID	37,63%	62,37%
SWEDCOM	24,12	37,00	0,00	SAANV	31,18%	68,82%
SWEDBEKER	2,01	6,00	0,00	SBLANK	76,34%	23,66%
SDPTOT	21,18	102,00	0,00	SNED	63,44%	36,56%
SDPCOM	5,20	24,00	0,00			
SASS	2,22	10,00	0,00			
SDPBEKER	0,69	4,00	0,00			
SINT	3,68	56,00	0,00			
SD/W TOT	0,18	1,00	0,00			
SD/W	0,20	1,13	0,00			
SGEEL	2,32	8,00	0,00			
SROOD	0,14	1,00	0,00			
SLEEN	0,11	2,00	0,00			

Tabellen 3.1 en 3.2: Een overzicht van de spelers karakteristieken.

### 3.2.2 Tijdseffecten

De tijdseffecten bevatten factoren die aangeven wanneer de transfers plaats hebben gevonden. Voor elke seizoen is een factor aangemaakt die de waarde 1 aanneemt als een transfer dat seizoen plaats heeft gevonden (SEI05/06\*,..., SEI09/10\*). Ook zijn er twee factoren opgenomen die aangeven of een transfer in de zomer voorafgaand aan het nieuwe seizoen plaatsvond (ZOMER\*) of in de winter halverwege het seizoen (WINTER\*). Bij factoren uit dit onderzoek waarbij gekeken wordt naar het seizoen voorafgaand aan de transfer, geldt bij een wintertransfer dat er gekeken wordt naar het halve seizoen voorafgaand aan de transfer. Tabel 3.3 geeft een overzicht van de tijdseffecten van de dataset.

Factor	1	0
SEI05/06	16,13%	83,87%
SEI06/07	23,66%	76,34%
SEI07/08	25,81%	74,19%
SEI08/09	23,66%	76,34%
SEI09/10	10,75%	89,25%
ZOMER	80,65%	19,35%
WINTER	19,35%	80,65%

Tabel 3.3: Een overzicht van de tijdseffecten.

In Tabel 3.3 is goed te zien dat de meeste transfers plaats vinden in de zomer. De zomerstop is langer dan winterstop wat een speler meer tijd geeft om aan zijn nieuwe omgeving te wennen. In de winter worden vaak alleen transfers uit nood gedaan bij personele problemen (zoals blessures) of matige prestaties van het team.

### 3.2.3 Karakteristieken van verkopende club

De karakteristieken van de verkopende club bevatten factoren die de kenmerken en team prestaties van de club die de speler verkoopt weergeven. Uit het onderzoek van Carmichael en Thomas (1993) zijn dit de volgende factoren: de begroting van de club (VCBEG), het gemiddeld aantal toeschouwers in het seizoen voorafgaand aan de transfer (VCGS), de positie van de club aan het eind van de competitie voorafgaand aan de transfer (VCPOS) en het doelsaldo van de club in dat laatste seizoen (VCDS).

VCBEG en VCGS zijn kenmerken die de grootte van de club aangeven. Om meer inzicht in de grootte van de club te geven is nog een extra factor toegevoegd die de stadioncapaciteit van de club weergeeft (VCCAP). VCCAP hangt samen met VCGS, waarbij VCCAP weergeeft hoeveel supporters er in totaal in het stadion kunnen en VCGS hoeveel supporters er gemiddeld ook daadwerkelijk naar het stadion komen. Om aan te geven in welke divisie een club speelt zijn twee factoren toegevoegd (VCED\* en VCJL\*). Als VCED de waarde 1 heeft dan speelt de club in het huidige seizoen van de transfer in de eredivisie en als VCJL de waarde 1 heeft dan speelt de club in het huidige seizoen van de transfer in de Jupiler League.

VCPOS en VCDS zijn factoren die de team prestaties van de club in het seizoen voorafgaand aan de transfer weergeven. In dit onderzoek zijn de team prestaties van de club verder gespecificeerd door middel van een aantal dummy factoren. Zo zijn er factoren gecreëerd

voor als de club het seizoen voorafgaand aan de transfer de competitie heeft gewonnen (VCWCOM\*), de KNVB beker heeft gewonnen (VCWBEEKER\*), gepromoveerd is naar een hogere divisie (VCPROM\*), gedegradeerd is naar een lagere divisie (VCDEG\*) en of de club Europees voetbal heeft weten af te dwingen (VCEUR\*). De prestaties van clubs in Europees verband is in dit onderzoek niet meegenomen. Het gaat in dit onderzoek om binnenlandse transfers en er wordt daarom ook alleen gekeken naar de prestaties binnen Nederland. Het afdwingen van Europees voetbal wordt wel meegenomen, omdat dit wel een binnenlandse prestatie is. Europees voetbal kan door een club worden afgedwongen door de KNVB beker te winnen of door aan het einde van het seizoen op de eerste tot en met de vierde plaats in de eredivisie te eindigen. Een laatste Europese ticket is weggelegd voor de club die de play-offs wint waaraan de vier clubs meedoen die aan het einde van het seizoen in de eredivisie op de plaatsen vijf tot en met acht zijn geëindigd. De tabellen 3.4 en 3.5 geven een overzicht van de karakteristieken van de verkopende club van de dataset.

Factor	1	0	Factor	Gemiddelde	Maximum	Minimum
VCWCOM	3,23%	96,77%	VCPOS	12,57	38,00	1,00
VCDEG	7,53%	92,47%	VCDS	3,43	71,00	-52,00
VCWCOM	6,45%	93,55%	VCCAP	18276,13	51628,00	2500,00
VCWBEEKER	4,30%	95,70%	VCGS	14820,92	49595,00	1471,00
VCEUR	37,63%	62,37%	VCBEG	16935483,87	68000000,00	1500000,00
VCED	75,27%	24,73%				
VCJL	24,73%	75,27%				

Tabellen 3.4 en 3.5: Een overzicht van de karakteristieken van de verkopende club.

In Tabel 3.4 is te zien dat 75% van de transfers is gedaan door clubs uit de eredivisie en 25% door clubs uit de Tweede Divisie. Dit verschil komt niet omdat clubs uit de Jupiler League minder spelers verkopen dan club uit de eredivisie. Bij verkoop van spelers uit de Jupiler League gaat het meestal om kleinere bedragen dan bij verkoop van spelers uit de eredivisie, waardoor het minder interessant is voor de media om er aandacht aan te besteden en er daardoor ook minder informatie over bekend is.

### 3.2.4 Karakteristieken van kopende club

De karakteristieken van de kopende club bevatten factoren die de kenmerken en team prestaties van de club die de speler koopt weergeven. Dit zijn dezelfde factoren als voor de karakteristieken van de verkopende club. Om aan te geven dat het om de karakteristieken van de kopende club gaat, zijn in de afkortingen van de factoren de letters VC door KC vervangen. De tabellen 3.6 en 3.7 geven een overzicht van de karakteristieken van de kopende club van de dataset.

Factor	1	0	Factor	Gemiddelde	Maximum	Minimum
KCPROM	4,30%	95,70%	KCPOS	7,37	34,00	1,00
KCDEG	2,15%	97,85%	KCDS	17,91	52,00	-32,00
KCWCOM	11,83%	88,17%	KCCAP	26949,86	51628,00	4662,00
KCWBEKER	9,68%	90,32%	KCGS	22542,94	49595,00	1462,00
KCEUR	60,22%	39,78%	KCBEG	29221505,38	68000000,00	2200000,00
KCED	93,55%	6,45%				
KCJL	6,45%	93,55%				

Tabellen 3.6 en 3.7: Een overzicht van de karakteristieken van de kopende club.

In Tabel 3.6 is ook een veel hoger percentage eredivisie clubs dan clubs uit de Jupiler League waarneembaar. Dit komt voornamelijk omdat clubs uit de Jupiler League minder geld te besteden hebben en daardoor vooral transfervrije spelers halen en spelers huren (data die niet is opgenomen in de dataset voor dit onderzoek).

### 3.3 Modellen en statistische methoden

Na het bepalen van de factoren is een meervoudig lineair regressie model opgesteld. Regressie is een statistische techniek voor het analyseren van data waar mogelijke sprake is van samenhang tussen de data. Deze samenhang wordt regressie genoemd en houdt in dat een variabele afhangt van een of meerdere andere variabelen. Deze variabelen worden respectievelijk de afhankelijke variabele en de verklarende variabelen genoemd. Er is sprake van enkelvoudige regressie als de afhankelijke variabele van een verklarende variabele afhangt en er is sprake van meervoudige regressie als de afhankelijke variabele van meerdere verklarende variabelen afhangt. Bij dit onderzoek is getest met verschillende meervoudige regressiemodellen (lineair, exponentieel, etc.) en het lineaire model geeft de beste fit bij deze dataset. Hoe de fit van een model wordt bepaald, komt later aan bod.

In Formule (1) is een voorbeeld gegeven van een meervoudig lineair regressie model.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_m X_m = \beta_0 + \sum_{i=1}^m \beta_i X_i, \quad \text{voor } i = 1, \dots, m. \quad (1)$$

Hierbij is  $Y$  de afhankelijke variabele, zijn de  $X$ 'en de verklarende variabelen en zijn de bèta's de coëfficiënten van de verklarende variabelen. De  $m$  geeft het aantal variabelen aan. Het lineaire model geeft voor dit onderzoek het beste resultaat als de natuurlijke logaritme van de afhankelijke variabele  $Y$  wordt genomen. Het uiteindelijke model wat gebruikt is bij dit onderzoek is gegeven in Formule (2).

$$\ln(Y) = \beta_0 + \sum_{i=1}^m \beta_i X_i, \quad \text{voor } i = 1, \dots, m. \quad (2)$$

Voor het maken en berekenen van het model is gebruik gemaakt van het programma Microsoft Excel. Om het beste model te vinden is het volgende stappenplan gebruikt:

1. Met de Least Squares methode (LS) wordt een model gemaakt.
2. Om de fit van het model te bepalen wordt de determinatiecoëfficiënt berekend.
3. Een F-toets voor het gehele model bepaald vervolgens of het model geldig is.
4. Aan de hand van een  $t$ -toets op de verklarende variabelen wordt bekeken welke variabelen (factoren) significant bijdragen aan het model.
5. Met de significante variabelen wordt vervolgens een nieuw model gemaakt, waarbij stappen 1 tot en met 4 herhaald worden.
6. Een F-toets bepaald of het model significant beter is dan zijn voorganger.
7. Stappen 1 tot en met 6 worden herhaald tot het beste model gevonden is.

In de subparagrafen 3.3.1 tot en met 3.3.4 zijn de stappen verder uitgewerkt. In eerste instantie is het de bedoeling geweest om meteen van alle factoren een model te maken. Excel heeft echter moeite met het maken van een regressie model met LS wanneer het model veel

variabelen bevat (in totaal worden bij dit onderzoek 52 variabelen onderzocht). Door deze beperking is er gekozen om eerst per categorie een regressie model te maken om daarna de beste modellen samen te voegen. Op dit samengevoegde model is het hierboven genoemde stappenplan nogmaals toegepast om het uiteindelijke model te krijgen. In het hoofdstuk resultaten wordt het uiteindelijke model verder beschreven. Een interessante vervolgstudie op dit onderzoek is om het onderzoek nogmaals te doen maar dan met een ander softwarepakket (die wel goed kan rekenen met een regressiemodel van 52 variabelen) om te onderzoeken of dit dezelfde resultaten oplevert.

### 3.3.1 Least Squares methode

In stap 1 van het stappenplan wordt een model gemaakt met behulp van de LS methode. Bij LS, de kleinste kwadraten methode in het Nederlands, worden de coëfficiënten van de verklarende variabelen van het model bepaald door de som van de fouten te minimaliseren. Eerst wordt het verschil genomen tussen de echte  $Y$ -waarden ( $Y$ ) en de door het regressie model voorspelde  $Y$ -waarden ( $Y^c$ ). Dit verschil is de fout in de voorspelling. Een fout kan positief of negatief zijn en hangt per fout af of  $Y - Y^c$  wordt gedaan of  $Y^c - Y$ . Om dit verschil in fouten te corrigeren worden alle fouten gekwadraterd zodat ze allemaal positief zijn. Als laatst worden alle fouten bij elkaar opgeteld. Formule (2) geeft de LS methode in formulevorm weer.

$$\min \sum_{i=1}^n (Y_i - Y_i^c)^2, \quad \text{voor } i = 1, \dots, n. \quad (3)$$

Hierbij geeft  $n$  het aantal waarnemingen aan. Het minimaliseren van de som van de gemaakte fouten gebeurt met de oplosser in Excel. De oplosser past binnen een vooraf ingestelde precisie en tijd de coëfficiënten van het model aan, totdat een minimum van de som van de fouten gevonden is.

### 3.3.2 De determinatiecoëfficiënt

Om de fit van het zojuist met de LS methode gemaakte model te bepalen, wordt de determinatiecoëfficiënt van het model berekend. De determinatiecoëfficiënt is een maatstaf om vast te stellen in hoeverre de totale variantie van de variabele  $Y$  wordt verklaard door de  $X$  variabelen. De determinatiecoëfficiënt wordt aangegeven met  $R^2$  en wordt berekend volgens Formule (4).

$$R^2 = \frac{\sum (Y_i^c - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2}, \quad \text{voor } i = 1, \dots, n. \quad (4)$$

In deze formule is  $\bar{Y}$  de gemiddelde  $Y$ -waarde en  $n$  het aantal waarnemingen. Bij een correct model is de waarde van deze coëfficiënt tussen of gelijk aan 0 en 1. De determinatiecoëfficiënt geeft aan wat de fit van het model is ten opzichte van de echte data. Hoe dichter de waarde bij 1 komt hoe beter het model de data representeert en hoe dichter de waarde bij 0 komt des te slechter het model de data representeert.

### 3.3.3 De F-toets

In stappen 3 en 6 van het stappenplan wordt een F-toets gebruikt. De F-toets is een statistische toets om na te gaan of van twee modellen de varianties verschillen of overeenkomen. Als eerst wordt een hypothese opgesteld, daarna wordt de toetsingsgrootte ( $F^c$ ) berekend. De toetsingsgrootte volgt onder de nulhypothese een F-verdeling met een bepaald aantal vrijheidsgraden  $v_1$  en  $v_2$  op een vooraf gesteld significantie level  $\alpha$  ( $F_{v_1, v_2, 0.5 * \alpha}$ ). Afhankelijk van de gekozen hypothese wordt de nulhypothese verworpen als  $F^c$  groter of kleiner is dan  $F_{v_1, v_2, 0.5 * \alpha}$ .

In stap 3 wordt getoetst of het gemaakte model geldig is. Hierbij wordt getoetst of de variantie van het model overeenkomt met de variantie van de data. Hierbij wordt de volgende hypothese gebruikt:

$H_0$ : Geen van de verklarende variabelen heeft een lineaire relatie met de afhankelijke variabele  $Y$ .

$H_1$ : Minstens 1 verklarende variabele heeft een verband met  $Y$ .

De toetsingsgrootte  $F^c$  wordt berekend volgens Formule (5):

$$F^c = \frac{\sum (Y_i^c - \bar{Y})^2 / m}{\sum (Y_i - Y_i^c)^2 / (n - m - 1)}, \quad \text{voor } i = 1, \dots, n. \quad (5)$$

Hierbij is  $n$  het aantal waarnemingen en  $m$  het aantal variabelen. De toetsingsgrootte  $F^c$  volgt onder de nulhypothese een F-verdeling met  $m$  en  $n-m-1$  vrijheidsgraden ( $F_{m, n-m-1, 0.5 * \alpha}$ ). Als significantie level wordt  $\alpha=0.05$  gebruikt. Wanneer  $F^c > F_{m, n-m-1, 0.5 * \alpha}$  wordt de nulhypothese verworpen en is het model geldig. Wanneer  $F^c < F_{m, n-m-1, 0.5 * \alpha}$  wordt de nulhypothese niet verworpen en is er geen sprake van een geldig model waar statistisch verder mee gerekend kan worden. In het laatste geval kan het model meteen de prullenbak in en moet er gezocht worden naar ander model. In het eerste geval kan er met het model verder gerekend worden.

In stap 6 wordt getoetst of het aangepaste model significant beter is dan het model waarvan het is afgeleid. Dit is al deels zelf te concluderen door de determinatiecoëfficiënten van de twee modellen met elkaar te vergelijken. Een vergelijking van de determinatiecoëfficiënten geeft echter geen zekerheid, omdat modellen met meerdere variabelen over het algemeen een betere fit hebben. Voor statistische zekerheid wordt een F-toets toegepast. Hierbij wordt de volgende hypothese getoetst:

$H_0$ : De weggelaten variabelen dragen niet bij tot de verklaring van variabele  $Y$ , het huidige model is significant beter.

$H_1$ : De weggelaten variabelen dragen wel bij tot de verklaring van variabele  $Y$ , het oorspronkelijke model is significant beter.

De toetsingsgrootte  $F^c$  wordt berekend volgens formule (6)

$$F^c = \frac{(R_o^2 - R_h^2) / w}{(1 - R_o^2) / (n - m - 1)}. \quad (6)$$



Hierbij is  $R_o^2$  de determinatiecoëfficiënt van het oorspronkelijke model,  $R_h^2$  de determinatiecoëfficiënt van het huidige model,  $n$  het aantal waarnemingen,  $m$  het aantal variabelen van het oorspronkelijke model en  $w$  het aantal weggelaten variabelen in het huidige model. De toetsingsgrootte  $F^c$  volgt onder de nulhypothese een F-verdeling met  $w$  en  $n-m-1$  vrijheidsgraden ( $F_{w,n-m-1,0.5*\alpha}$ ). Als significantie level wordt  $\alpha=0.05$  gebruikt. Wanneer  $F^c > F_{w,n-m-1,0.5*\alpha}$  wordt de nulhypothese verworpen en is het oorspronkelijke model significant beter. Wanneer  $F^c < F_{w,n-m-1,0.5*\alpha}$  wordt de nulhypothese niet verworpen en is het huidige model significant beter. In het laatste geval is het beste model gevonden en in het eerste geval kan er worden onderzocht of ook dit model nog verder aangepast kan worden of dat dit het beste model is.

### 3.3.4 De t-toets

In stap 4 van het stappenplan wordt een t-toets gebruikt. De t-toets is een statistische toets die bij meervoudige regressie kan worden gebruikt om voor elke coëfficiënt  $\beta_i$  te toetsen of deze significant van 0 verschilt. Er wordt dus getoetst of een variabele significant bijdraagt aan het model. Ook bij de t-toets wordt eerst een hypothese opgesteld. Voor dit onderzoek is de hypothese:

$H_0: \beta_i = 0$ , variabele  $X_i$  levert geen significante bijdrage aan het model.

$H_1: \beta_i \neq 0$ , variabele  $X_i$  levert een significante bijdrage aan het model.

Deze hypothesen worden getoetst met de zogenoemde  $t$ -ratio ( $t$ ) en wordt berekend volgens Formule (6).

$$t = \beta_i / s_{\beta_i}. \quad (6)$$

Hierbij is  $\beta_i$  de coëfficiënt van variabele  $X_i$  uit het model en  $s_{\beta_i}$  de standaarddeviatie van de richtingscoëfficiënt. Om de standaarddeviatie van de richtingscoëfficiënt te vinden is de covariatie matrix  $C$  berekend volgens Formule (7).

$$C = \frac{\sum_{i=1}^n (Y_i^c - Y_i)^2}{(n - m - 1)} * (X^T X)^{-1}, \quad \text{voor } i = 1, \dots, n. \quad (7)$$

In deze formule is  $Y_i$  waarneming  $i$ ,  $Y_i^c$  de geschatte waarneming  $i$ ,  $n$  het aantal waarnemingen en  $m$  het aantal variabelen in het model.  $(X^T X)^{-1}$  is de inverse matrix van  $X^T X$ , waarbij  $X$  de matrix van alle data uit de dataset is en  $X^T$  de getransponeerde matrix van  $X$  is. De variantie van de richtingscoëfficiënt voor  $\beta_i$  is te vinden op diagonaal van matrix  $C$  (op positie  $i,i$ ). De benodigde standaarddeviatie is de wortel van de variantie.

De uiteindelijk gevonden  $t$ -ratio wordt vergeleken met een waarde uit de tabel van de  $t$ -verdeling ( $t_{n-m-1,\alpha}$ ). Deze waarde hangt van het gekozen significantie level ( $\alpha$ ) en het aantal vrijheidsgraden. Het aantal vrijheidsgraden is gelijk aan  $n-m-1$ , waar  $n$  het aantal waarnemingen is en  $m$  het aantal variabelen in het model is. Als  $-t_{n-m-1,\alpha} > t > t_{n-m-1,\alpha}$  wordt de nulhypothese verworpen en levert variabele  $X_i$  een significante bijdrage aan het model. Als  $-t_n$

$t_{m-1,\alpha} < t < t_{n-m-1,\alpha}$  wordt de nulhypothese niet verworpen en levert variabele  $X_i$  geen significante bijdrage aan het model. Variabelen die geen significante bijdrage aan het model leveren worden in het volgende model buiten beschouwing gelaten.

## H4 Resultaten

In dit hoofdstuk worden de resultaten van het onderzoek beschreven. In de eerste paragraaf worden de resultaten uiteengezet. In de tweede paragraaf worden de resultaten van het onderzoek vergeleken met de resultaten uit de Engelse onderzoeken die in de achtergrondstudie van dit BWI werkstuk aan bod zijn gekomen.

### 4.1 De onderzoeksresultaten

Met het stappenplan uit het onderzoeksontwerp is voor elke categorie factoren het beste model bepaald. Deze modellen zijn samengevoegd tot een model, waarop wederom het stappenplan is toegepast. Het uiteindelijk beste model dat hieruit volgde is weergegeven in Figuur 4.1. Naast het model staat een tabel met de belangrijkste karakteristieken en toetswaarden van het model.

Variabele	Coëfficiënt	Significantie		
Constante	3,4311			
SCLUBS	-0,2204			
SASS	0,0537			
SDPBKER	-0,2280			
SKPR	1,5589	***		
SVERD	1,7947	***		
SMID	1,4311	***		
SAANV	1,6472	***		
SD/W TOT	2,2912	***		
SD/W	0,4252	*		
SROOD	0,2460			
SLEEN	-0,3511			
SBLANK	-0,1514			
SNED	-0,4881			
SEI05/06	1,2270	***		
SEI06/07	1,3304	***		
SEI07/08	1,4753	***		
SEI08/09	1,6986	***		
SEI09/10	1,7006	***		
ZOMER	2,2531	***		
WINTER	2,1786	***		
VCPRM	-1,0693	***		
VCDEG	0,6465	**		
VCWCOM	0,2910			
VCEUR	0,5797			
VCED	2,8775	***	$n$	93
VCJL	1,5540	***	$m$	32
KCPROM	-0,9769	***	$R^2$	0,7165
KCDEG	-1,6409	***	$F^c$	4,7378
KCWCOM	0,3413		$F(32;60; 0,025)$	1,7979
KCEUR	0,5077		$t(0,01;60)$	***
KCED	2,2283	***	$t(0,05;60)$	**
KCJL	2,2033	***	$t(0,10;60)$	*

Figuur 4.1: Het uiteindelijke model.

Het model is gebaseerd op 93 waarnemingen ( $n$ ), bestaat uit 32 variabelen ( $m$ ) en heeft een fit van 71,65% op de gebaseerde data ( $R^2$ ). In Figuur 4.1 geeft de eerste kolom de variabelen uit het model aan, in de tweede kolom staan de bijbehorende coëfficiënten en in de derde kolom is de uitkomst van de  $t$ -toets gegeven. Het aantal sterren komt hierbij overeen met het significantie niveau aangegeven door de drie onderste posities in de karakteristieke tabel. Hierbij is 60 het aantal vrijheidsgraden ( $n-m-1$ ) en zijn de significantie niveaus respectievelijk 1%, 5% en 10%. De variabelen met drie sterren hebben de grootste significante bijdrage aan het model en de variabelen zonder sterren hebben geen significante bijdrage aan het model. Ook zijn de resultaten van de F-toets naar de geldigheid van het model bijgevoegd. In dit geval is de berekende toetsingsgrootte  $F^c$  groter dan de toetswaarde  $F_{m,n-m-1,0.5*\alpha}$  met respectievelijk 32 en 60 vrijheidsgraden en een significantie level van 5%. De conclusie van de F-toets is daarom dat het model geldig is en er verder mee gewerkt mag worden.

In het uiteindelijk model hebben 21 van de 32 variabelen een significante bijdrage volgens de  $t$ -toets. Er is daarom ook een model gemaakt met daarin enkel de significante variabelen van het model uit Figuur 4.1. Dit tweede model is weergegeven in bijlage B. De determinatiecoëfficiënt van dit tweede model is 58.04%, een beduidend lagere fit dus. Aan de hand van een F-toets is berekend welk model significant beter is. Hierbij bleek de berekende toetsingsgrootte  $F^c$  groter dan de toetswaarde  $F_{w,n-m-1,0.5*\alpha}$ , wat betekent dat het eerste model uit Figuur 4.1 significant beter is. Dus ondanks dat het eerste model uit een aantal niet significante variabelen bestaat is deze beter dan het model wat alleen uit de significante variabelen bestaat. Dit betekent dat de niet significante variabelen samen wel van waarde zijn voor het model.

Uit het model blijkt uiteindelijk dat van de spelers karakteristieke de factoren SKPR, SVERD, SMID, SAANV, SD/W TOT en SD/W het meest significante van belang zijn en allemaal positief bijdragen aan het transferbedrag van een speler. Dit zijn de factoren die de positie van een speler op het veld en het gemiddeld aantal doelpunten per wedstrijd van een speler aangeven. De coëfficiënten van de positiefactoren schelen onderling niet veel van elkaar, maar zijn voor een verdediger en aanvaller het grootst. Uit het model blijkt dus dat het transferbedrag hoger is als een speler verdediger of aanvaller is. Uit de coëfficiënten blijkt tevens dat het totale doelpuntgemiddelde per wedstrijd over de gehele carrière van een speler een grotere bijdrage aan het transferbedrag levert dan het doelpuntgemiddelde per wedstrijd van het afgelopen seizoen.

De factoren SCLUBS, SASS, SDPBKER, SROOD, SLEEN, SBLANK en SNED spelen ook rol bij het bepalen van het transferbedrag, maar in mindere mate vanwege de significantie. Dit zijn onder andere de twee racisme factoren huidskleur en nationaliteit. Significante is er dus geen sprake van racisme bij het bepalen van transferbedragen binnen het Nederlandse voetbal. Hierbij valt aan de negatieve coëfficiënt op dat een blanke huidskleur en een Nederlandse nationaliteit een negatief effect hebben op het transferbedrag. Als er dus al gediscrimineerd wordt bij het betalen van transferbedragen, dan is dit ten opzichte van de blanke Nederlander. Ook kan dit worden uitgelegd dat voor spelers van niet Nederlandse nationaliteit met een niet blanke huidskleur een hoger transferbedrag moet worden betaald. Dit kan bijvoorbeeld te maken hebben met de kwaliteit van deze spelers.

Uit het model blijkt dat van de tijdseffecten alle factoren significant van belang zijn en dat zowel het seizoen zelf als het tijdstip binnen het seizoen waarin de transfer heeft plaatsgevonden van belang zijn bij het bepalen van het transferbedrag. Aan de oplopende

coëfficiënten van de seizoensfactoren blijkt dat de transferbedragen hoger worden naarmate de tijd vordert.

Van de karakteristieken van de verkopende clubs zijn de factoren VCPROM, VCDEG, VCED en VCJL significant van belang. Dit zijn de factoren die aangeven of een club het afgelopen seizoen gepromoveerd is, gedegradeerd is en in welke divisie het speelt in het huidige seizoen. Aan de coëfficiënt is te zien dat promotie van een verkopende club een negatieve invloed heeft op het transferbedrag van een speler en degradatie een positief effect heeft op het transferbedrag van een speler. Een vreemde conclusie, omdat het omgekeerde logischer lijkt. De verwachting is dat een speler die met zijn club promoveert meer waard wordt, vanwege de positieve sportieve teamprestatie. Zo is ook de verwachting dat de waarde van een speler die met zijn club degradeert minder wordt vanwege de negatieve sportieve teamprestatie. De dataset die gebruikt is voor dit onderzoek vertoont echter tegenovergestelde resultaten. Dit is mogelijk een toevallig resultaat van de gebruikte dataset. Om dit te controleren en verder te onderzoeken is vervolgstudie met een andere dataset nodig.

De divisiecoëfficiënten van de verkopende club laten zien dat het transferbedrag van een speler die onder contract staat bij een eredivisieclub bijna twee keer zo hoog is voor een verkopende club als het transferbedrag van een speler die onder contract staat bij een club uit de Jupiler League. Voor een speler die met zijn club in de hoogste divisie voetbalt, ontvangt de verkopende club dus een hoger transferbedrag dan voor een speler die met zijn club in de laagste divisie voetbalt. De factoren VCWCOM en VCEUR spelen ook rol bij het bepalen van het transferbedrag, maar in mindere mate vanwege de significantie.

Van de karakteristieken van de kopende clubs zijn de factoren KCAPROM, KCDEG, KCED en KCJL significant van belang. Dit zijn net als bij de verkopende clubs de factoren die aangeven of een club het afgelopen seizoen gepromoveerd is, gedegradeerd is en in welke divisie het speelt in het huidige seizoen. Aan de coëfficiënten is te zien dat zowel promotie als degradatie van een kopende club het transferbedrag negatief beïnvloedt. Net als bij de verkopende club is dit resultaat het verder onderzoeken waard. De divisiecoëfficiënten laten zien dat het transferbedrag van een speler die onder contract staat bij een eredivisieclub hoger is voor een kopende club als het transferbedrag van een speler die onder contract staat bij een club uit de Jupiler League. Een club uit de hoogste divisie betaalt dus meer voor een speler dan een club uit de laagste divisie. De factoren KCWCOM en KCEUR spelen ook rol bij het bepalen van het transferbedrag, maar in mindere mate vanwege de significantie.

Concluderend wordt het transferbedrag van een profvoetballer in Nederland grotendeels bepaald aan de hand van de positie van een speler op het veld, het gemiddeld aantal doelpunten per wedstrijd van een speler (zowel het gemiddelde over de speler zijn gehele carrière genomen als het gemiddelde van het afgelopen seizoen), in welk seizoen en op welk tijdstip binnen het seizoen de transfer heeft plaatsgevonden, of de verkopende en kopende club het afgelopen seizoen gepromoveerd zijn, of de verkopende en kopende club het afgelopen seizoen gedegradeerd zijn en in welke divisie de verkopende en kopende club spelen ten tijden van de transfer.

## **4.2 Vergelijking van resultaten**

In de volgende drie subparagrafen worden de resultaten van het onderzoek vergeleken met de resultaten van de drie Engelse onderzoeken die eerder aan bod zijn gekomen tijdens dit BWI werkstuk. Hierbij wordt het model van dit onderzoek aangeduid als het Nederlandse model. Als eerst worden steeds de resultaten van het model van het betreffende onderzoek uiteengezet, waarna vervolgens de vergelijking wordt gemaakt met de resultaten van het Nederlandse model.

### **4.2.1 Resultaten van Carmichael en Thomas**

Het definitieve model uit het onderzoek van Carmichael en Thomas (1993) is gebaseerd op 214 waarnemingen, bestaat uit 18 variabelen en heeft een fit van 81%. Van de 18 variabelen zijn er 13 significant. Dit zijn van de spelers karakteristieken de leeftijd van de speler, het aantal goals en gespeelde wedstrijden in het seizoen voorafgaand aan de transfer en de factor die aangeeft of het transferbedrag van de speler is bepaald door het sporttribunaal. Het model bevat geen spelers karakteristieken die niet van significant belang zijn. Het model bevat de volgende 7 karakteristieken voor zowel de verkopende als de kopende club:

1. De begroting van de club.
2. Het gemiddeld aantal toeschouwers van de club in het seizoen voorafgaand aan de transfer.
3. Het doelsaldo van de club aan het eind van het seizoen voorafgaand aan de transfer.
4. De positie van de club in de competitie aan het eind van het seizoen voorafgaand aan de transfer.
5. Divisie factor 1: De club speelt in de hoogste divisie.
6. Divisie factor 2: De club speelt in de tweede divisie.
7. Divisie factor 3: De club speelt in een van de twee laagste divisies.<sup>2</sup>

Van de karakteristieken van de verkopende club zijn de factoren 4, 5 en 6 van significant belang voor het bepalen van het transferbedrag. Van de karakteristieken van de kopende club zijn de factoren 2 tot en met 7 van significant belang voor het bepalen van het transferbedrag. In dit onderzoek zijn tijdseffecten niet onderzocht, omdat er van transferdata afkomstig uit een seizoen gebruik is gemaakt.

In vergelijking met het Nederlandse model heeft het model van Carmichael en Thomas (hierna het Engelse model genoemd) een 10% betere fit op de gebruikte data. Beide modellen concluderen dat het transferbedrag voor een speler wordt bepaald aan de hand van het aantal doelpunten en gespeelde wedstrijden van de speler. Het Engelse model bevat deze twee factoren alleen individueel en het Nederlandse model bevat deze twee factoren gecombineerd in het gemiddeld aantal doelpunten per wedstrijd. Ook volgt uit beide modellen dat het transferbedrag van een speler bepaald wordt door de divisiefactoren van zowel de verkopende als van de kopende club (op de laagste divisies voor de verkopende club in het Engelse model na).

---

<sup>2</sup> Het Engelse voetbal bestaat uit twee divisies meer dan het Nederlandse voetbal.

De twee modellen tonen naast overeenkomsten ook verschillen. Zo wordt volgens het Engelse model het transferbedrag van een speler ook nog bepaald aan de hand van de leeftijd van een speler, of het transferbedrag van de speler is vastgesteld door het sporttribunaal, de positie van zowel de verkopende club als de kopende club op de ranglijst, het gemiddeld aantal toeschouwer van de kopende club en het doelsaldo van de kopende club. Volgens het Nederlandse model wordt het transferbedrag van een speler ook nog bepaald aan de hand van de positie van een speler op het veld en of de kopende en verkopende club het afgelopen seizoen gepromoveerd of gedegradeerd zijn.

#### **4.2.2 Resultaten van Reilly en Witt**

Het definitieve model uit het onderzoek van Reilly en Witt (1995) is gebaseerd op 202 waarnemingen, bestaat uit 19 variabelen en heeft een fit van 74.70%. Van de 19 variabelen zijn er 13 significant. Dit model is een uitbreiding van het model uit de vorige paragraaf met de variabele huidskleur om racisme binnen het Engelse voetbal te onderzoeken. In het definitieve model is de variabele huidskleur niet van significant belang. Er wordt geconcludeerd dat de factor huidskleur geen rol speelt bij het bepalen van het transferbedrag voor een speler in het Engelse voetbal. De racisme factoren uit het Nederlandse model spelen ook geen significante rol van betekenis, waardoor ook in het Nederlandse voetbal racisme geen rol speelt bij het bepalen van het transferbedrag van een speler. De onderzoeken naar racisme vertonen dus gelijke resultaten.

#### **4.2.3 Resultaten van Dobson, Gerrard and Howe**

Het definitieve model uit het onderzoek van Dobson, Gerrard and Howe (2000) is gebaseerd op 114 waarnemingen, bestaat uit 21 variabelen en heeft een fit van 60%. Van de 21 variabelen zijn er 15 significant. Dit model (gericht op het Engelse amateurvoetbal) is een uitbreiding van het model van Carmichael en Thomas (gericht op het Engelse profvoetbal) met een variabele die het jaar aangeeft waarin de transfer plaats heeft gevonden en een variabele die aangeeft of een transfer plaats heeft gevonden voor de start van een nieuw seizoen of midden in het seizoen. De eerste variabele is een continue variabele in tegenstelling tot de dummy variabelen die in het Nederlandse model zijn gebruikt om het seizoen aan te geven waarin de transfer heeft plaatsgevonden. De resultaten van dit onderzoek naar het Engelse amateur voetbal komen grotendeels overeen met de resultaten naar het Engelse profvoetbal. De tijdseffecten zijn daarnaast (net als de Nederlandse tijdseffecten) van significant belang bij het bepalen van het transferbedrag van een speler.





## H5 Conclusies en aanbevelingen

In dit hoofdstukken worden de belangrijkste conclusies en aanbevelingen van het onderzoek op een rij gezet. Hierbij wordt antwoord gegeven op de probleemstelling van het onderzoek: *Welke factoren bepalen het transferbedrag van een profvoetballer in Nederland?* In de eerste paragraaf volgen de conclusies en in de tweede paragraaf de aanbevelingen.

### 5.1 Conclusies

Het onderzoek naar het transferbedrag van een profvoetballer in Nederland levert een meervoudig lineair regressie model op dat gebaseerd is op een dataset van 93 binnenlandse transfers. Het model bestaat uit 32 variabelen en heeft een fit van 71,65% op de gebaseerde data. Van de 32 variabelen hebben er 21 een significante bijdrage aan het model. Ondanks dat het model uit een aantal niet significante variabelen bestaat is deze beter dan wanneer het model uit alleen de significante variabelen zou bestaan. De niet significante variabelen zijn samen dus wel van waarde voor het model.

Uit het model volgt dat het transferbedrag van een profvoetballer in Nederland grotendeels bepaald wordt aan de hand van de positie van een speler op het veld, het gemiddeld aantal doelpunten per wedstrijd van een speler (zowel het gemiddelde over de speler zijn gehele carrière genomen als het gemiddelde van het afgelopen seizoen), in welk seizoen en op welk tijdstip binnen het seizoen de transfer heeft plaatsgevonden, of de verkopende en kopende club het afgelopen seizoen gepromoveerd zijn, of de verkopende en kopende club het afgelopen seizoen gedegradeerd zijn en in welke divisie de verkopende en kopende club spelen ten tijden van de transfer. Uit het model volgt ook dat de twee racismefactoren huidskleur en nationaliteit geen significante rol spelen bij het bepalen van het transferbedrag van een profvoetballer in Nederland.

De modellen uit de onderzoeken naar het Engelse voetbal bevatten een aantal dezelfde factoren, maar verschillen ook op een aantal factoren van het model voor het Nederlandse voetbal. In de modellen voor zowel het Engelse als Nederlandse voetbal wordt het transferbedrag bepaald aan de hand van het aantal doelpunten en gespeelde wedstrijden van de speler in het seizoen voorafgaande aan de transfer (in het Nederlandse model gecombineerd tot het gemiddeld aantal doelpunten per wedstrijd), in welke divisie zowel de verkopende als de kopende club spelen ten tijde van de transfer, in welk seizoen de transfer plaats heeft gevonden en wanneer in het seizoen de transfer plaats heeft gevonden. Verschillend is dat volgens het Engelse model het transferbedrag van een speler ook nog bepaald wordt aan de hand van de leeftijd van een speler, of het transferbedrag van de speler is vastgesteld door het sporttribunaal, de positie van zowel de verkopende club als de kopende club op de ranglijst, het gemiddeld aantal toeschouwer van de kopende club en het doelsaldo van de kopende club.

De resultaten van het onderzoek naar racisme zijn gelijk voor het Engelse en Nederlandse voetbal. In beide modellen spelen de racismefactoren geen significante rol. Er is dus geen bewijs gevonden voor racisme bij het bepalen van het transferbedrag voor een voetballer in Engeland en Nederland.

## ***5.2 Aanbevelingen***

Bij het uitvoeren van dit onderzoek is gebruikt gemaakt van Microsoft Excel. Daarbij is de functie oplosser van Excel gebruikt voor het optimaliseren van de gebruikte modellen met de Least Square methode. Tijdens het onderzoek is gebleken dat Excel niet de meest ideale software is voor meervoudige regressie modellen met veel variabelen. Hoe meer variabelen er worden gebruikt in het model des te meer moeite Excel heeft om tot een optimale oplossing te komen. Mocht er in de toekomst nog een vervolgstudie op dit onderzoek komen, dan is het interessant om het onderzoek nogmaals te doen met een ander softwarepakket en dan een vergelijking van de resultaten te maken.

Een andere mogelijkheid tot vervolgstudie op dit onderzoek volgt uit een tegenstelling die uit het definitieve model resulteert. Uit het model volgt dat promotie van een verkopende club een negatieve invloed heeft op het transferbedrag van een speler en degradatie een positief effect heeft op het transferbedrag van een speler. De verwachting is juist dat een speler die met zijn club promoveert meer waard wordt vanwege de positieve sportieve teamprestatie en dat een speler die met zijn club degradeert minder waard wordt vanwege de negatieve sportieve teamprestatie. Ook bij de kopende club volgt een dergelijke tegenstelling. Deze contradicties zijn mogelijk een toevallig resultaat van de gebruikte dataset. Om dit te controleren en verder te onderzoeken is vervolgstudie met een andere dataset nodig.

## Literatuurlijst

### Artikelen

Carmichael, F. en Thomas, D. (1993), Bargaining in the transfer market: theory and evidence. Gepubliceerd in: Applied Economics, volume 25, 12 december 1993, pagina's 1467 – 1476.

Reilly, B. en Witt, R. (1995), English league transfer prices: is there a racial dimension? Gepubliceerd in: Applied Economics Letters, volume 2, 7 juli 1995, pagina's 220 – 222.

Dobson, S. Gerrard, B. en Howe, S. (2000), The determination of transfer fees in English nonleague football. Gepubliceerd in: Applied Economics, volume 32, 9 July 2000, pagina's 1145 – 1152.

### Boek

Drs. A. Buijs (2000), Statistiek om mee verder te werken, tweede druk, Uitgever: Stenfert Kroese.



## Bijlage A – Overzicht van factoren

Deze bijlage bevat een overzicht van de factoren die gebruikt zijn tijdens het onderzoek. Hierbij wordt steeds eerst de gebruikte afkorting gegeven en vervolgens een korte uitleg van de factor. Als er een sterretje (\*) achter de afkorting staat, geldt de uitleg voor een zomertransfer. Voor een wintertransfer geldt in plaats van het hele seizoen voorafgaand aan de transfer, het halve seizoen voorafgaand aan de transfer.

SLFT	De leeftijd van de speler op de transferdatum.
SCLUBS	Het totaal aantal clubs uit de gehele professionele carrière van de speler voorafgaand aan de transfer.
SWEDTOT	Het totaal aantal gespeelde wedstrijden in de gehele professionele carrière van de speler op de transferdatum.
SWEDCOM*	Het aantal gespeelde competitiewedstrijden door de speler in het seizoen voorafgaand aan de transfer.
SWEDBEKER*	Het aantal gespeelde wedstrijden in de KNVB beker door de speler in het seizoen voorafgaand aan de transfer
SDPTOT	Het totaal aantal doelpunten in de gehele professionele carrière van de speler op de transferdatum.
SDPCOM*	Het aantal doelpunten van de speler in de competitie in het seizoen voorafgaand aan de transfer.
SASS*	Het aantal assist van de speler in de competitie in het seizoen voorafgaand aan de transfer.
SDPBEKER*	Het aantal doelpunten van de speler in de KNVB beker in het seizoen voorafgaand aan de transfer.
SINT	Het totaal aantal interlands gespeeld door speler voor het hoogste nationale team op de transferdatum.
SKPR	Een dummy factor voor de positie van de speler; 1=keeper; 0=anders.
SVERD	Een dummy factor voor de positie van de speler; 1=verdediger; 0=anders.
SMID	Een dummy factor voor de positie van de speler; 1=middenvelder; 0=anders.
SAANV	Een dummy factor voor de positie van de speler; 1=aanvaller; 0=anders.
SD/W TOT	Het doelpuntgemiddelde per wedstrijd van de speler in zijn gehele professionele carrière.
SD/W*	Het doelpuntgemiddelde per wedstrijd van de speler in het seizoen voorafgaand aan de transfer
SGEEL*	Het aantal gele kaarten die de speler in het seizoen voorafgaand aan de transfer heeft ontvangen.
SROOD*	Het aantal rode kaarten die de speler in het seizoen voorafgaand aan de transfer heeft ontvangen.
SLEEN	Het aantal keren dat de speler is uitgeleend door de verkopende club.
SBLANK	Een dummy factor voor de huidskleur van de speler; 1=speler is blank; 0=speler is niet blank.
SNED	Een dummy factor voor de nationaliteit van de speler; 1=speler heeft een Nederlandse nationaliteit; 0=speler heeft een andere nationaliteit.
SEI05/06,... ,SEI09/10	Een dummy factor die aangeeft in welk seizoen de transfer plaatsvond (een seizoen loopt van juni jaar 1 t/m mei jaar 2); 1=transfer vond plaats in het betreffende seizoen; 0=transfer vond plaats in een ander seizoen.

ZOMER	Een dummy factor die aangeeft in welke transferperiode de transfer heeft plaatsgevonden; 1=transfer is afgesloten in de zomerse transferperiode; 0=anders.
WINTER	Een dummy variabele die aangeeft in welke transferperiode de transfer heeft plaatsgevonden; 1=transfer is afgesloten in de winterse transferperiode; 0=anders.
VCPOS*	De positie op de ranglijst van de competitie van de verkopende club aan het eind van het seizoen voorafgaand aan de transfer. (Eredivisie: 1 t/m 18, Jupiler League: 19 t/m 38)
VCDS*	Het doelsaldo van de verkopende club in het seizoen voorafgaand aan de transfer (doelsaldo = doelpunten voor - doelpunten tegen)
VCPRM	Een dummy factor; 1=de verkopende club heeft in het seizoen voorafgaand aan de transfer promotie afgedwongen; 0=anders.
VCDEG	Een dummy factor; 1=de verkopende club is het seizoen voorafgaand aan de transfer gedegradeerd;0=anders.
VCWCOM	Een dummy factor; 1=de verkopende club is het seizoen voorafgaand aan de transfer kampioen geworden in de competitie; 0=anders.
VCWBEKER	Een dummy factor; 1=de verkopende club heeft het seizoen voorafgaand aan de transfer de KNVB beker gewonnen; 0=anders.
VCEUR	Een dummy factor; 1=de verkopende club heeft zich in het seizoen voorafgaand aan de transfer geplaatst voor Europees voetbal; 0=anders.
VCED	Een dummy factor; 1= de verkopende club speelt in het huidige seizoen van de transfer in de Eredivisie; 0=anders.
VCJL	Een dummy factor; 1= de verkopende club speelt in het huidige seizoen van de transfer in de Jupiler League; 0=anders.
VCCAP	De stadioncapaciteit van de verkopende club in het huidige seizoen van de transfer.
VCGS	Het gemiddeld aantal toeschouwers van de verkopende club in het seizoen voorafgaand aan de transfer.
VCBEG	De begroting van de verkopende club in het huidige seizoen van de transfer.
KCPOS*	De positie op de ranglijst van de competitie van de kopende club aan het eind van het seizoen voorafgaand aan de transfer. (Eredivisie: 1 t/m 18, Jupiler League: 19 t/m 38)
KCDS*	Het doelsaldo van de kopende club in het seizoen voorafgaand aan de transfer (doelsaldo = doelpunten voor - doelpunten tegen)
KCPROM	Een dummy factor; 1=de kopende club heeft in het seizoen voorafgaand aan de transfer promotie afgedwongen; 0=anders.
KCDEG	Een dummy factor; 1=de kopende club is het seizoen voorafgaand aan de transfer gedegradeerd;0=anders.
KCWCOM	Een dummy factor; 1=de kopende club is het seizoen voorafgaand aan de transfer kampioen geworden in de competitie; 0=anders.
KCWBEKER	Een dummy factor; 1=de kopende club heeft het seizoen voorafgaand aan de transfer de KNVB beker gewonnen; 0=anders.
KCEUR	Een dummy factor; 1=de kopende club heeft zich in het seizoen voorafgaand aan de transfer geplaatst voor Europees voetbal; 0=anders.
KCED	Een dummy factor; 1= de kopende club speelt in het huidige seizoen van de transfer in de Eredivisie; 0=anders.
KCJL	Een dummy factor; 1= de kopende club speelt in het huidige seizoen van de transfer in de Jupiler League; 0=anders.

KCCAP	De stadioncapaciteit van de kopende club in het huidige seizoen van de transfer.
KCGS	Het gemiddeld aantal toeschouwers van de kopende club in het seizoen voorafgaand aan de transfer.
KCBEG	De begroting van de kopende club in het huidige seizoen van de transfer.





## Bijlage B – Het aangepaste model

Het uiteindelijk model uit het resultaten hoofdstuk bevat een aantal niet significante variabelen. Om te onderzoeken of deze variabelen uit het model gelaten mogen worden is een tweede model gemaakt met enkel de significante variabelen. Uiteindelijk bleek dit tweede model niet significant beter dan het oorspronkelijke model. In Figuur B.1 is het aangepaste tweede model weergegeven met naast het model de belangrijkste karakteristieken en toetswaarden.

Variabele	Coëfficiënt	Significantie		
Constante	3,2217			
SKPR	1,5154	***		
SVERD	1,6700	***		
SMID	1,4521	***		
SAANV	1,5845	***		
SD/W				
TOT	2,7792	***		
SD/W	-0,0184			
SEI05/06	1,0001	***		
SEI06/07	1,4381	***		
SEI07/08	1,4623	***		
SEI08/09	1,5761	***		
SEI09/10	1,7455	***	<i>n</i>	93
ZOMER	2,0484	***	<i>m</i>	21
WINTER	2,1735	***	<i>w</i>	11
VC PROM	-1,6544	***	$R^2$	0,5804
VC DEG	0,7491	***	$F^{c, geldig}$	4,6765
VC ED	3,1413	***	$F(21;71;0,025)$	1,8893
VC JL	1,0805	**	$F^{c, verschil}$	2,6176
KC PROM	-0,9526	***	$F(11;60;0,025)$	2,2159
KC DEG	-2,3803	***	$t(0,01,71)$	***
KC ED	2,2947	***	$t(0,05,71)$	**
KC JL	1,9271	***	$t(0,10,71)$	*

Figuur B.1: Het aangepaste model.

Het model is gebaseerd op 93 waarnemingen ( $n$ ), bevat 21 variabelen ( $m$ ), 11 variabelen uit het eerste model zijn verwijderd ( $w$ ) en de fit van het model is 58.04%.  $F^{c, geldig}$  is de toetsingsgrootte van de F-toets naar de geldigheid van het model. Deze is groter dan de toets waarde  $F(21;71;0,025)$ , dus is het model geldig.  $F^{c, verschil}$  is de toetsingsgrootte van de F-toets die toetst welk model significant beter is. Deze is groter dan de toets waarde  $F(11;60;0,025)$ , dus is het eerste model significant beter.