

Rapportcijfers voor voetballers: zin of onzin?*

Erik van der Aar Cor de Gooijer Yvonne Halfhide
 Ruud H. Koning Maarten Licht

13 augustus 1996

Samenvatting

In dit paper analyseren we de rapportcijfers die voetballers krijgen. Er zijn twee bronnen voor deze cijfers: *De Telegraaf* en de *Voetbal International*. Aan de hand van een simpel statistisch model zullen de cijfers worden geanalyseerd, waarin we onderscheid maken tussen wedstrijdspecifieke effecten en individu-specifieke effecten. We zullen nagaan in hoeverre de schattingen voor deze effecten zijn gecorreleerd tussen beide bronnen. Verder zullen we nagaan van welke factoren de wedstrijdspecifieke effecten afhangen.

1 Inleiding

Veel Nederlanders nemen zondagavond en maandagochtend de voetbaluitslagen door. Met collega's worden de verrichtingen van de favoriete teams doorgenomen: de helden van het weekend bejubeld en de scheidsrechters (indien nodig) beschimpt. Het aardige van voetballen is dat iedereen er een mening over heeft, niet noodzakelijkerwijze gebaseerd op kennis van zaken.

*Correspondentieadres: Ruud H. Koning, Vakgroep Econometrie, De Boelelaan 1105, 1081 HV Amsterdam, email rkoning@econ.vu.nl. De auteurs danken Bab Barends van *De Telegraaf* voor het beschikbaar stellen van de gegevens voor het seizoen 1994/95, O. Verdonk voor het gebruik van de twee jaargangen *Voetbal International*, Harry van Dalen, Bernard Liendenhovius en Geert Ridder voor commentaar. Erik van der Aar, Cor de Gooijer, Yvonne Halfhide en Maarten Licht studeren Econometrie aan de Vrije Universiteit, Amsterdam. Ruud Koning is als universitair docent verbonden aan de vakgroep Econometrie van de Vrije Universiteit en is lid van de Werkgroep Voetbal en Statistiek.

De voetballers worden overigens niet alleen beoordeeld door het publiek in de stadions en tijdens de gesprekken op maandag, ze krijgen ook rapportcijfers in een landelijk ochtendblad en een weekblad. In deze notitie analyseren we die rapportcijfers; met name gaan we na in hoeverre de beoordelingen tussen beiden bladen overeenstemmen. Die rapportcijfers dienen overigens niet alleen ter vermaak van de lezers: op basis van de rapportcijfers wordt een prestigieuze prijs als 'De voetballer van het jaar' uitgereikt en het gemiddelde rapportcijfer van 5.8 was voor Volendam één van de redenen om Danny den Ouden te ontslaan¹.

Zowel het dagblad *De Telegraaf* als het weekblad *Voetbal International*² nemen bij hun wedstrijdverslagen beoordelingen op van de voetballers. Deze beoordelingen hebben de vorm van rapportcijfers (gehele getallen tussen 1 en 10) en worden gegeven door speciaal daarvoor aangestelde deskundigen. Eerst zullen we nagaan in hoeverre de individuele beoordelingen zijn gecorreleerd en daarna zullen we aan de hand van een simpel regressiemodel die rapportcijfers analyseren. Een aantal vragen komen aan de orde:

- Bestaan er significante verschillen tussen de beoordelingen door de rapporteur van *De Telegraaf* en die van de *V.I.*?
- Wat is de correlatie tussen de beoordelingen door *De Telegraaf* en *V.I.*?
- Verschillen de beoordelingen tussen uit- en thuiswedstrijden?

De opzet van deze notitie is als volgt. In de volgende paragraaf behandelen we het datamateriaal. In paragraaf 3 bespreken we een statistisch model voor de rapportcijfers. Aan de hand van dat model worden de rapportcijfers in paragraaf 4 geanalyseerd. We eindigen in paragraaf 5 met enige conclusies.

2 Rapportcijfers voor voetballers

De data die in dit rapport worden gebruikt, zijn afkomstig uit *De Telegraaf* en *V.I.* Zij publiceren na elke wedstrijd in de PTT-Telecompetitie de wedstrijdgegevens. Deze wedstrijdgegevens omvatten ondermeer de opstellingen van de teams en de aan de spelers toegekende rapportcijfers. De procedure,

¹ *V.I.* 15 november 1995.

² Hierna zullen we dit afkorten tot *V.I.*

aan de hand waarvan de cijfers worden vastgesteld, verschilt tussen beide bladen. Deze verschillen zullen later in deze paragraaf worden besproken.

De analyse beperkt zich tot drie teams: in alfabetische volgorde zijn dat Ajax, Feyenoord en Heerenveen. De eerste twee representeren de top van het Nederlandse clubvoetbal en Heerenveen is een goede middenmoter. Voor deze ploegen zijn gegevens verzameld voor de seizoenen 1993/94 en 1994/95. Op het moment dat de data werden verzameld (april 1995) was het seizoen 1994/95 nog niet helemaal afgelopen, zodat van dat seizoen de gegevens van de laatste vijf wedstrijdronden ontbreken. In het seizoen 1993/94 werd Ajax kampioen, eindigde Feyenoord als tweede en werd Heerenveen dertiende. In het seizoen 1994/95 prolongeerde Ajax het kampioenschap, werd Feyenoord vierde en Heerenveen negende. In beide seizoenen won Feyenoord de Amstel Cup.

De volgende gegevens zijn voor elke wedstrijd verzameld: de rapportcijfers van de spelers, de uitslag, de scheidsrechter, de locatie van de wedstrijd (uit of thuis) en het aantal toeschouwers. *V.I.* geeft de scheidsrechter ook een rapportcijfer. Van de gegevens uit *De Telegraaf* is ook bekend wie de rapporteur was.

De rapportcijfers van *De Telegraaf* worden de eerste werkdag na een competitiewedstrijd gepubliceerd in *De Telegraaf* en *Het Nieuws van de Dag*. De rapporteurs zijn mensen die vroeger actief zijn geweest in het professionele voetbal: ze zijn allen oud-international of oud-trainer. Aan de hand van acht criteria kennen zij een cijfer toe. Deze criteria worden expliciet vermeld op de lijst die de rapporteurs inleveren. De criteria zijn: algehele presentatie (contact met supporters, tenue, gedrag), werklust, conditie, techniek, tactiek, dienstbaarheid aan team, sportiviteit en creativiteit. Het cijfer dat wordt gepubliceerd is het gemiddelde van de scores op bovengenoemde punten. Er zijn wel restricties verbonden aan het cijfer: een speler die een rode kaart krijgt kan niet een voldoende krijgen voor de betreffende wedstrijd en een speler moet minimaal 45 minuten hebben meegespeeld om een cijfer te kunnen krijgen.

De rapportcijfers van *V.I.* worden gegeven door rapporteurs ter plaatse. De rapporteurs zijn verbonden aan *V.I.* of sportjournalist van een regionaal dagblad. Het is derhalve mogelijk dat het merendeel van de thuiswedstrijden van een bepaalde club door dezelfde rapporteur worden beoordeeld. De rapporteurs geven een globaal cijfer aan de spelers en beoordelen ze niet expliciet op bepaalde criteria. Een speler moet minimaal 20 minuten hebben gespeeld om voor een cijfer in aanmerking te komen. Het is dus mogelijk dat

bepaalde spelers wel een waardering in *V.I.* hebben en niet in *De Telegraaf*.

3 Een regressiemodel

In deze paragraaf zullen we een eenvoudig statistisch model voor de rapportcijfers bespreken. We gaan uit van het volgende model voor de *waargenomen* rapportcijfers:

$$y_{it} = \mu_t + \alpha_i + \varepsilon_{it}, \quad i = 1, \dots, I_j, \quad t = 1, \dots, T. \quad (1)$$

In deze vergelijking is y_{it} het rapportcijfer dat speler i in wedstrijd t krijgt. μ_t is een wedstrijdspecifiek effect en α_i is een individu-specifiek effect. We veronderstellen dat de storingen ε_{it} onderling ongecorrleerd zijn en verwachting 0 hebben. I_j is het aantal spelers van club j . Aangezien in ons bestand geen speler voor twee verschillende clubs uitkomt, hoeven we geen extra teamsubscript toe te voegen aan y_{it} , μ_t en α_i .

De parameters van model (1) zouden eenvoudig te schatten zijn, als alle wedstrijden in dezelfde samenstelling zouden zijn gespeeld. In dat geval is het teamgemiddelde per wedstrijd een schatter voor het wedstrijdspecifieke effect μ_t . De individuele effecten zouden te schatten zijn door per speler het gemiddelde van de afwijkingen van de wedstrijdspecifieke effecten te nemen. Echter, teams variëren in hun samenstelling: bij Ajax speelden tussen 1993 en 1995 21 spelers zes wedstrijden of meer, bij Feyenoord 20 en bij Heerenveen 22³. Tevens raken spelers geblesseerd, uit vorm, stoppen met voetbal of worden getransfereerd. Bovendien wordt in een groot aantal wedstrijden meer dan 11 spelers ingezet: er wordt tijdens de wedstrijd gewisseld.

De parameters van vergelijking (1) kunnen worden geschat met behulp van kleinste kwadraten op basis van de volgende vergelijking:

$$y_{it} = \sum_{s=1}^T \mu_s I(s=t) + \sum_{j=1}^I \alpha_j I(j=i) + \varepsilon_{it}, \quad i \in \mathcal{I}_t, \quad t = 1, \dots, T. \quad (2)$$

In deze vergelijking is $I(\cdot)$ de indicatorfunctie die waarde 1 aanneemt als het argument waar is en 0 anders en is $\mathcal{I}_t$ de verzameling van spelers die hebben gespeeld in de t -e wedstrijd. Uiteraard zijn de parameters μ_s en α_i niet allen geïdentificeerd; het is niet mogelijk zowel een absoluut teamniveau te meten

³Alle resultaten zijn gebaseerd op de verzameling spelers die minimaal zes keer zijn beoordeeld.

als een absoluut individueel effect op basis van dezelfde data. We leggen de volgende identificerende restrictie op: $\sum_i \alpha_i = 0$ zodat het gemiddelde spelerseffect 0 is. Deze restrictie heeft als voordeel boven andere identificerende restricties dat de wedstrijdspecifieke effecten vergelijkbaar zijn tussen de drie teams.

4 Empirische resultaten

Een eerste ruwe indruk van de vergelijkbaarheid van de beoordelingen in *De Telegraaf* en *V.I.* wordt ontleend aan de correlaties tussen de beoordelingen van de spelers. De correlatie tussen de rapportcijfers van de individuele rapportcijfers in *De Telegraaf* en *V.I.* is voor Ajax 0.46, voor Feyenoord 0.49 en voor Heerenveen 0.48. Hoewel de correlatie niet buitengewoon hoog is te noemen, liggen ze alledrie in dezelfde orde van grootte.

De Telegraaf heeft de naam een 'Amsterdamse' krant te zijn en dit wordt door de gegevens niet tegengesproken. Met behulp van een tekentoets (zie onder meer Lehmann (1975)) hebben we getoetst of *De Telegraaf* significant vaker hogere waarderings geeft dan *V.I.*. Voor alledrie clubs geldt dat *De Telegraaf* vaker een hogere waardering geeft dan *V.I.*, maar alleen bij Ajax gebeurt dit significant vaker⁴.

Het is echter interessanter om naar de schattingsresultaten van model (2) te kijken. In dat model wordt onderscheid gemaakt tussen een teameffect en een individueel effect. Aangezien het aantal parameters van model (2) erg groot is presenteren we hier alleen enkele samenvattende en afgeleide grootheden⁵. In tabel 1 geven we de gemiddelden en standaarddeviaties van de geschatte wedstrijdspecifieke effecten $\hat{\mu}_s$. Beide bladen geven Ajax, gemiddeld genomen, de hoogste beoordeling, de verschillen tussen het gemiddelde wedstrijdspecifieke effect van Feyenoord en Heerenveen is marginaal. De spreiding van de wedstrijdspecifieke effecten is ongeveer even groot voor

⁴De p -waarden voor de nulhypothese dat er geen systematische verschillen zijn tegen het (eenzijdige) alternatief dat *De Telegraaf* hogere cijfers toe kent zijn respectievelijk 0 (Ajax), 0.058 (Feyenoord) en 0.20 (Heerenveen).

⁵Voor hen die het willen weten: de spelers met de hoogste geschatte individuele effecten zijn op basis van *De Telegraaf* gegevens Blind, Kluyvert, Litmanen (Ajax), De Goey, Larsson, Taument (Feyenoord) en Keur, L'Ami, De Visser (Heerenveen) en op basis van de *V.I.* gegevens Finidi, Petterson, Blind (Ajax), De Goey, Bosz, Van Bronckhorst (Feyenoord) Keur, L'Ami, Oosterveer (Heerenveen).

team	<i>De Telegraaf</i>				<i>V.I.</i>			
	m_μ	s_μ	R^2	s_α	m_μ	s_μ	R^2	s_α
Ajax	6.34	0.43	0.38	0.32	6.16	0.48	0.37	0.34
Feyenoord	5.96	0.42	0.35	0.22	5.87	0.43	0.37	0.39
Heerenveen	5.92	0.47	0.41	0.22	5.87	0.62	0.44	0.27

Tabel 1: Gemiddelde (m_μ), standaarddeviatie (s_μ) van de wedstrijdspecifieke effecten, R^2 van regressie (2) en standaarddeviatie individuele effecten (s_α)

alledrie de clubs en beide bladen en de R^2 voor beide bladen is bijna even hoog. De spreiding van de individuele effecten is bij Ajax het hoogst.

Het ligt voor de hand om de schattingen van de wedstrijdspecifieke- en individuele effecten te vergelijken tussen beide bladen. Met behulp van een Chow-toets (Chow (1960)) hebben we een aantal toetsen uitgevoerd. Ten eerste zijn we nagegaan of de individuele effecten gezamenlijk 0 zijn. Vervolgens hebben we getoetst of de wedstrijdspecifieke effecten en de individuele effecten gelijk zijn voor beide bladen. Ook hebben we getoetst op gelijkheid van de wedstrijdspecifieke effecten. Een laatste toets die we hebben uitgevoerd is die op gelijkheid van de individuele effecten. Alle nulhypotheseën moesten worden verworpen op elk redelijk betrouwbaarheidsniveau. We concluderen dan ook dat beide bladen significant verschillende waarderingen geven aan dezelfde wedstrijden en spelers.

Enig inzicht in de correlaties tussen de twee effecten uit beide bronnen kan worden ontleend aan tabel 2. We zien dat de individuele effecten sterker zijn gecorreleerd dan de wedstrijdspecifieke effecten. De rapporteurs zien een wedstrijd door een verschillende bril, maar ze zijn het er redelijk over eens wie de betere en mindere spelers zijn. De laatste twee kolommen van tabel 2 bevatten de autocorrelatiecoëfficiënt van de wedstrijdspecifieke effecten. Als de rapportcijfers iets als 'vorm' zouden meten, zou men verwachten dat deze positief gecorreleerd zijn: een team weet meestal goede vorm vast te houden om vervolgens enige tijd in een mindere vorm te raken. Echter, de autocorrelaties zijn soms zelf negatief en in het algemeen zijn ze laag. Voor Ajax en Feyenoord verschillen de autocorrelaties duidelijk tussen beide bladen en voor Heerenveen zijn ze 0.

Tenslotte proberen we in tabel 3 de gevonden wedstrijdspecifieke effecten te verklaren uit andere waargenomen variabelen: het aantal gescorde doelpunten voor, het aantal doelpunten tegen en een dummy-variabele die de waarde 1 aanneemt als het een uitwedstrijd betreft. Uit de resultaten van

	$\rho(\mu_s^T, \mu_s^{VI})$	$\rho(\alpha_i^T, \alpha_i^{VI})$	$\rho(\mu_s^T, \mu_{s-1}^T)$	$\rho(\mu_s^{VI}, \mu_{s-1}^{VI})$
Ajax	0.40	0.61	-0.13	-0.025
Feyenoord	0.44	0.70	-0.12	0.25
Heerenveen	0.73	0.61	0.018	0.027

Tabel 2: Correlaties

variabele	Ajax		Feyenoord		Heerenveen	
	<i>De Telegraaf</i>	<i>V.I.</i>	<i>De Telegraaf</i>	<i>V.I.</i>	<i>De Telegraaf</i>	<i>V.I.</i>
intercept	6.12 (0.14)	5.97 (0.15)	5.79 (0.097)	5.75 (0.10)	5.82 (0.11)	5.67 (0.14)
goals voor	0.11 (0.032)	0.11 (0.035)	0.12 (0.030)	0.14 (0.031)	0.20 (0.043)	0.29 (0.058)
goals tegen	-0.13 (0.059)	-0.19 (0.063)	-0.13 (0.041)	-0.084 (0.040)	-0.12 (0.031)	-0.11 (0.041)
uit	0.060 (0.10)	0.062 (0.11)	0.13 (0.092)	-0.094 (0.095)	0.12 (0.094)	0.073 (0.13)
n	62	62	63	63	63	63
R^2	0.27	0.30	0.31	0.31	0.43	0.41

Tabel 3: Verklaring wedstrijdspecifieke effecten (standaardfouten tussen haakjes)

tabel 3 blijkt dat er geen sprake van is dat één van de onderzochte teams betere waarderingen krijgt bij een thuis- dan wel een uitwedstrijd. Doelpunten voor hebben een positief effect op het wedstrijdspecifieke effect en doelpunten tegen een negatief effect. Beide effecten zijn absoluut genomen ongeveer even groot voor elk team, tussen de teams verschillen ze. Een gescoord doelpunt voor Heerenveen resulteert in een toename die bijna twee keer zo groot is als de ‘beloning’ die Ajax voor een doelpunt krijgt.

5 Conclusies

In dit rapport hebben we de rapportcijfers voor voetballers van *De Telegraaf* en *V.I.* geanalyseerd. Aan de hand van een eenvoudig regressiecomponenten-model hebben we een aantal vragen proberen te beantwoorden. Een belangrijke vraag is natuurlijk of de rapporteur van *De Telegraaf* en *V.I.* dezelfde wedstrijd hebben gezien, met andere woorden, wat is de correlatie tussen beide beoordelingen?

Hoewel is gebleken dat de beoordelingen van *De Telegraaf* en *V.I.* zowel op

team- als op individueel niveau positief gecorreleerd zijn, is deze correlatie niet erg hoog te noemen. De deskundige rapporteurs zien bij benadering dezelfde wedstrijd en de rapporteur van *De Telegraaf* is geneigd de spelers van Ajax beter te waarderen dan zijn collega van *V.I.*. We hebben verder weinig bewijs gevonden voor het bestaan van de 'vorm' van een team: de wedstrijdspecifieke effecten vertonen nauwelijks enige autocorrelatie. Hoewel de correlatie tussen de beoordelingen van spelers positief is, is deze in absolute waarde gezien niet erg hoog. De waardering van voetballers is gelukkig geen exacte wetenschap, en dat is misschien maar goed ook...

Referenties

- Chow, G.C. (1960). Tests of equality between subsets of coefficients in two linear regressions. *Econometrica* 28, 591-605.
- Lehmann, E.L. (1975). *Nonparametrics: Statistical Methods Based On Ranks*. San Fransisco: Holden-Day.