



Wie gaan er winnen in Athene?

Auteur(s):

G.H. Kuper en E. Sterken

De auteurs zijn verbonden aan de faculteit der Economische Wetenschappen van de Rijksuniversiteit Groningen. e.sterken@eco.rug.nl

Verschenen in:

ESB, 89e jaargang, nr. 4439, pagina 375, 6 augustus 2004

Rubriek:

sport

Trefwoord(en):

Op basis van een econometrisch model werd twee jaar geleden in ESB voorspeld dat Nederland zeven medailles zou winnen op de Olympische Winterspelen in Salt Lake City. Het werden er acht. Dit artikel blikt terug op de voorspelling van toen en doet een voorspelling voor de Olympische Spelen in Athene.

De Olympische Zomerspelen van Athene beginnen op 13 augustus 2004. Het noc-nsf vaardigt ongeveer 215 atleten af in de hoop het succes van vier jaar geleden in Sydney op zijn minst te herhalen. Destijds zette de Nederlandse equipe een - in historisch perspectief - uitstekende prestatie neer met twaalf gouden medailles, negen zilveren en vier bronzen plakken. In dit artikel wordt met behulp van een modelvoorspelling van de medailleverdeling in Athene onderzocht of het Nederlandse succes naar verwachting gecontinueerd zal worden.

We presenteren de uitkomsten van een voorspelling aan de hand van een econometrisch model dat voor 118 landen de medaillewinst beschrijft. Ruim twee jaar geleden hebben we een vergelijkbaar artikel over de Winterspelen van Salt Lake City in 2002 geschreven (Kuper & Sterken, 2002). Met het toen gebruikte model hebben we gewaarschuwd voor een al te groot optimisme over de Nederlandse medaillewinstkansen. We voorspelden begin 2002 dat Nederland in Salt Lake City in totaal zeven medailles (twee goud, twee zilver en drie brons) zou behalen. In werkelijkheid werden het acht medailles (drie goud en vijf zilver). Omdat we een vergelijkbaar model voor de Zomerspelen opgezet hebben is het goed om, alvorens de voorspelling van Athene te presenteren, terug te komen op de nauwkeurigheid van de voorspelling voor Salt Lake City.

Terugblik

We hebben onze voorspelling voor Salt Lake City op twee verschillende manieren geanalyseerd. Ten eerste kan men een maat ontwikkelen voor de afwijkingen ten opzichte van de realisaties per medaillesoort of medailletotaal van de voorspellingen. De uitkomsten kunnen worden vergeleken met die van alternatieve voorspellingen. Ten tweede kan men zich richten op de voorspelde rangorde van landen in het medailleklassement.

In de eerste klasse beschouwt men de afwijking van de modelvoorspelling van het medailleklassement of -totaal tot die van een alternatieve voorspelling. Van belang is de definitie van een alternatieve voorspelling. Een eerste alternatieve voorspelling zou kunnen zijn dat alle landen eenzelfde aandeel in de medaillewinst zouden hebben. In dat geval kan men de zogenaamde determinatiecoëfficiënt (ofwel R^2) eenvoudig uitrekenen. De R^2 van de voorspelling van de totale medaillewinst per land is 0,84, maar voor de verdeling over goud, zilver en brons bedraagt de R^2 slechts 0,52. Dit bevestigt het beeld dat voor Nederland al schetsmatig geconstateerd is: de voorspelling van het medailletotaal is betrekkelijk nauwkeurig, maar de verdeling over de kleuren is moeilijker te voorspellen. Op basis van het verleden is het evenwel niet aannemelijk dat er een gelijke medailleverdeling per land resulteert. Sommige landen hebben immers een echte Olympische traditie. Daarom zou een volgende alternatieve voorspelling de uitslag van de vorige editie (in dit geval Nagano 1998) kunnen zijn. We kunnen dan de gesommeerde gekwadrateerde medailleafwijkingen van de voorspelling voor Nagano met de modelvoorspelling vergelijken. Ook nu ontstaat hetzelfde beeld: de alternatieve voorspelling van het medailletotaal is slechter dan onze modelvoorspelling (de voorspelfoutvariantie van de raming voor Nagano is 68 procent hoger dan de modelvoorspelling), maar de alternatieve voorspelling van de kleurverdeling is beter (4 procent lagere voorspelfoutvariantie). Een laatste alternatieve voorspelling is een concurrerende voorspelling. Een studie van Ali en Johnson (2002) geeft slechts de winst per land van het totaal aantal medailles en de gouden medailles. Onze voorspelfoutvariantie van het totaal aantal medailles is negen procent lager dan die van Ali en Johnson, terwijl de voorspelfoutvariantie van de gouden medailles voor beide modellen nagenoeg gelijk is.

Een tweede manier om de voorspellingen te analyseren, is dat men zich afvraagt of de voorspelde rangorde van landen in het medailleklassement accuraat is. Daartoe hebben we de rangcorrelatiecoëfficiënt van Kendall tussen de voorspelde en gerealiseerde rangorde uitgerekend. Voor het totaal aantal medailles is de Kendall-correlatiecoëfficiënt 0,91, en voor goud, zilver en brons respectievelijk 0,79, 0,84 en 0,57. De vergelijkbare Kendall-rangcorrelatiecoëfficiënten voor de Ali-Johnson voorspellingen en de realisaties van de totaal gewonnen medaille en de gewonnen gouden plakken zijn iets lager, 0,90 en 0,74 respectievelijk. Uit deze analyse concluderen wij dat het gebruikte model vooral waarde toevoegt als het gaat om het voorspellen van de landvolgorde in het medailleklassement en het voorspellen van het totaal aantal medailles per land.

Model

In het artikel over de voorspelling van de medailleoogst in Salt Lake City is het gebruikte model besproken (Kuper & Sterken, 2002). We herhalen hier de belangrijkste elementen en bespreken de recent aangebrachte wijzigingen van het model voor de Zomerspelen uitgebreider. Centraal in het model staat een tweestapsprocedure. Eerst wordt de totale deelname per land voorspeld. Vervolgens wordt de medaillewinstkans conditioneel op de participatie geschat. Uiteraard is deelnemen noodzakelijk om medailles te winnen. De meeste deelnemende landen voeren een strak selectiebeleid, zodat een grotere ploeg veelal ook meer medaillekansen heeft. We hebben voor de Zomerspelen een nieuw model ontwikkeld (Kuper & Sterken, 2003). Dit is nodig omdat er veel meer landen mee doen aan de Zomerspelen (in Athene verwacht men 201 landen tegen 77 in Salt Lake City), de aard van de Zomerspelen anders is (de geografie van de deelnemende landen verschilt van die van de deelnemende landen aan de Winterspelen) en er veel meer nummers (sportonderdelen) zijn op de Zomerspelen (301 in Athene tegen 78 in Salt Lake City).

In totaal hebben ongeveer 120 landen tijdens de moderne Zomerspelen sinds 1896 ooit medailles gewonnen. Voor deze landengroep hebben we de parameters van het voorspelmodel vastgesteld op basis van de data van de laatste vier edities (Seoul, Barcelona, Atlanta en Sydney). Het model is voorts onderworpen aan een *out-of-sample* toets, waarin de medaillewinst van Sydney voorspeld is. De determinanten van de participatie zijn de bevolkingsomvang, het inkomen per hoofd van de bevolking, de breedte van de sportbeoefening en het thuisvoordeel. Meer inwoners levert meer talent op, een hoger gemiddeld inkomen maakt specialisatie in sport en investeringen in de sport eenvoudiger en het thuisland mag meer atleten afvaardigen (bij de moderne spelen komt dat neer op 100 tot 150 atleten extra). De breedte van de sportbeoefening per land is als volgt bepaald. Tijdens de Olympische Spelen beoefent men verschillende takken van sport (in Athene zijn dit er 39). Sinds 1896 is het bekend welk land aan welk evenement meegedaan heeft. Per land is de fractie van beoefende sporten sinds 1896 redelijk stabiel en dus voorspelbaar. Als een land in meer takken van sport actief is, is het waarschijnlijk dat de Olympische ploeg groter is. Medaillewinst (aandelen per land) wordt door de participatie (eveneens in aandelen per land) en het thuisvoordeel verklaard. Het thuisvoordeel werkt dubbel: men kan meer atleten sturen en men kan profiteren van een grotere supporterschare.

Innovaties van het model

Ten opzichte van het model voor de Winterspelen zijn er twee majeure innovaties doorgevoerd. Ten eerste is er gedetailleerd rekening gehouden met de rol van teamsporten. Ten tweede is er gebruik gemaakt van informatie binnen een Olympische periode om de voorspellingen te verbeteren.

In Athene staan veertien teamsportevenementen op het programma. Ruim twintig procent van de deelnemers is actief in deze evenementen, terwijl slechts minder dan vijf procent van de medailles verdiend kunnen worden in basketbal, handbal, hockey, honkbal, softbal, voetbal, volleybal en waterpolo. Omdat ons model medaillewinst laat afhangen van teamomvang zou een vertekening in de medaillekansen van landen met een grote vertegenwoordiging in teamsporten kunnen ontstaan. De kans is groot dat de deelname onderschat wordt, terwijl de medaillekansen overschat kunnen worden. Daarom zijn de participatie- en succesmodellen voor teamsporten en niet-teamsporten apart geschat.

De tweede innovatie heeft betrekking op de rol van informatie over sportevenementen in de Olympiade. Gedurende de Olympische periode worden diverse kampioenschappen gehouden. Het ligt voor de hand dat de informatie, met name over het succes, die in deze evenementen opgeslagen ligt van invloed is op de Olympische medaillewinst. Voor de atletiek (een groot evenement met 24 nummers voor mannen en 22 voor vrouwen) hebben we vanaf 1983 onderzocht in hoeverre de medailleverdeling in jaar $t-1$ de medailleverdeling per land bij de Olympische Spelen in jaar t bepaalt. De uitslag van de Wereldkampioenschappen (wk) draagt significant bij aan de verklaring van de uitslag van de Zomerspelen en bevat meer informatie dan de uitslag van de vorige editie van de Spelen. Hoewel de recente uitslagen op wk's dus informatief zijn, is het goed om te beseffen dat de aard van competities tijdens wereldkampioenschappen verschilt van die op de Spelen. Meestal zijn er per nummer meer deelnemers op een wk dan op de Spelen en is de competitie sterker. Immers, op de Spelen mag een land maar een beperkt aantal atleten per nummer inzetten. Zo kan het voorkomen dat een land alle medailles wint op een wk, terwijl dit op de Spelen veelal niet mogelijk is.

Op basis van de uitkomst van het experiment in de atletiek zijn voor de edities van Sydney voor de destijds driehonderd gehouden nummers de medailleverdelingen van in 1998/1999 gehouden wk's (en soms vergelijkbare en vervangende evenementen) verzameld en is een naïeve raming van de medaillewinst per land opgesteld. Tevens is met behulp van het model een *ex ante* raming van de medailleverdeling gemaakt. Vervolgens is de definitieve uitslag van Sydney geregresseerd op de naïeve voorspelling, de modelvoorspelling en de uitslag van Atlanta. Per medaillekleur zijn aldus gewichten vastgesteld die in de voorspelronde voor Athene gebruikt kunnen worden.

Opvallend is dat voor alle medaillekleuren het gewicht van de naïeve wk-voorspelling 0,7 tot 0,8 bedraagt. Voor de gouden medaille is de vorige editie van de Spelen de belangrijkste aanvullende bron, terwijl voor zilver en brons de modelvoorspelling van belang is. De genoemde weging hebben we toegepast op onze definitieve raming van de uitslag in Athene. Daartoe is dus ook een naïeve raming van de medaillespiegel op basis van wk-uitslagen in 2003 voor de 301 in Athene af te werken nummers opgesteld.

Resultaten

De uiteindelijke raming van de top dertig van de medaillespiegel staat vermeld in [tabel 1](#) (in totaal winnen naar verwachting zo'n zeventig landen één of meer medailles). De ramingen zijn gebaseerd op de bovengenoemde weging van de modelvoorspelling, de uitslag van Sydney en de naïeve voorspelling op basis van wk-resultaten in 2003. We bespreken eerst de algemene indruk van de tabel en vervolgens de speciale rol van Griekenland en Nederland.

Tabel 1. Voorspelling van de medailleverdeling

		goud	zilver	brons
1	Verenigde Staten	34	37	26
2	Rusland	31	25	27
3	China	27	21	23

4	Duitsland	19	21	29
5	Australië	15	16	14
6	Frankrijk	15	13	9
7	Japan	10	6	10
8	Zuid-Korea	9	8	8
9	Groot-Brittannië	8		10 11
10	Oekraïne	8		6 10
11	Italië	7		10 12
12	Cuba	7		8 5
13	Griekenland (a)	7	5	1
14	Zweden	7	3	5
15	Hongarije	5	10	7
16	Roemenië	5	8	8
17	Bulgarije	5	8	4
18	Canada	5	6	9
19	Polen	5	5	7
20	Argentinië	5	1	3
21	Spanje	4	8	9
22	Wit-Rusland	4	7	10
23	Nederland	4	6	11
24	Brazilië	4	3	2
25	Iran	4	1	2
26	Ethiopië	3	1	2
27	Denemarken	3	1	2
28	Uzbekistan	3	1	2
29	Nieuw-Zeeland	3	1	1
30	Zwitserland	2	3	1

a Ongewogen voorspelling

De top drie van Athene is ongewijzigd ten opzichte van Sydney. De Verenigde Staten zijn de verwachte winnaar. Sinds het uiteenvallen van de Sovjet-Unie lijkt de hegemonie van de Amerikanen onaantastbaar. Rusland verliest ten opzichte van Sydney, maar kan de tweede plaats vasthouden. In 2000 wonnen de Russen nog 88 medailles. In Athene zijn dit er een aantal minder. China staat steviger op positie drie en wint veertien medailles meer dan in Sydney. Met de organisatie van de Spelen in Beijing in 2008 in het vooruitzicht, is er duidelijk sprake van de opbouw van atletenkapitaal. Duitsland komt op plaats vier en Australië verliest het thuisvoordeel en zakt terug naar de vijfde plaats. Italië, Nederland en Cuba vallen weg uit de top tien en moeten hun posities afstaan aan Japan, Korea en de Oekraïne.

Indien men onze voorspelling vergelijkt met een voorspelling van de medailletotalen per land van Price Waterhouse Coopers (*Het Financieele Dagblad*, 13 juli 2004) valt op dat de pwc-voorspelling aanzienlijke dalingen van het aantal door de toplanden te winnen medailles laat zien (bijvoorbeeld voor de Verenigde Staten een reductie van 27 medailles ten opzichte van Sydney), terwijl wij een stabilisatie (en zelfs groei voor China) voorspellen.

Griekenland en Nederland

Er zijn twee speciale gevallen: de rol van het thuisland Griekenland en de positie van Nederland. Griekenland komt in de eigenlijke gewogen top dertig niet voor vanwege een slechte positie op de naïeve wk-voorspelling (1 goud, 2 zilver en 2 brons). Het thuisvoordeel komt in deze naïeve voorspelling evenwel niet aan bod. Derhalve is voor Griekenland gekozen voor de modelvoorspelling (7 goud, 5 zilver en 1 brons). Nederland zakt van de achtste positie op de ranglijst in Sydney naar de 23ste plaats nu. Dit wordt vooral veroorzaakt door het verlies van gouden posities. Met één gouden medaille meer klimt Nederland vijf plaatsen. Wat betreft de medailletotalen behaalt Nederland naar verwachting 21 medailles, hetgeen een dertiende positie oplevert. Voor Nederland is het interessant om de onderliggende deelvoorspellingen te beschouwen. Volgens de naïeve wk-schatting is de verwachte medailleoogst: twee goud, vijf zilver en twaalf brons. Het model voorspelt voor Nederland vijf goud, zes zilver en zeven brons. De uiteindelijke weging met de Sydney-uitslag (twaalf goud, negen zilver en vier brons) geeft de uiteindelijke positie op de ranglijst. Volgens de pwc-voorspelling is het thuisvoordeel veel sterker en zou Griekenland 29 medailles halen (tegen 13 medailles volgens onze voorspelling). Voor Nederland komt de pwc-voorspelling overeen met die in [tabel 1](#).

Besluit

De kosten van de Nederlandse Olympische ploeg zijn in de afgelopen vier jaar ongeveer 120 miljoen euro geweest. De kosten worden door sportbonden, het ministerie van vws en particuliere sponsorgelden gedragen. In dit artikel hebben we een inschatting gemaakt van het verwachte rendement van deze publiek-private investering. We schatten slechts een beperkt facet van het rendement (en letten bijvoorbeeld niet op ander succes tijdens bijvoorbeeld wereldkampioenschappen), maar doen dit wel onder een inschatting van de prestaties van concurrenten.

Ingegeven door het succes van de Olympische afvaardigingen naar Nagano 1998 en Sydney 2000 zijn de medaillewinstverwachtingen bij de Zomerspelen in Athene in Nederland hooggespannen. Onze statistische analyse levert een behoudende raming van de medailleoogst voor Nederland op. Een vergelijkbare behoudende raming voor de Winterspelen van Salt Lake City is achteraf goed gebleken. Hiermee komt wellicht de voornaamste waarde van ons model naar voren: met het model is van tevoren een normaal rendement en dus na afloop een zogenaamd 'abnormaal' rendement vast te stellen. Mocht Nederland toch dertig medailles winnen, dan heeft de ploeg boven verwachting goed gepresteerd. Hoe het ook zij, meedoen aan de Spelen is belangrijk, maar winnen lijkt, vanuit een economisch perspectief, steeds belangrijker te worden.

Literatuur

Johnson, D.K.N. en A. Ali (2002) A Tale of Two Seasons: Participation and Medal Counts at the Summer and Winter Olympic Games. Wellesley Working Papers 2002-2.

Kuper, G.H. en E. Sterken (2002) Wie gaat er winnen in Salt Lake City?, ESB,

8 februari 2002, blz. 110-111.

Kuper G.H. en E. Sterken (2003) Olympic Participation and Performance Since 1896. SOM Research Report 03C19.

Het Financieele Dagblad, Minder medailles Oranje, 13 juli 2004.
