Relatieve prestatiemeting bij AEX-fondsen

Roeland Haanen, Victor Maas en Sander van Triest

SAMENVATTING Door bestuurders te belonen op basis van relatieve prestaties in plaats van absolute prestaties kunnen bedrijven enkele belangrijke nadelen van prestatiebeloning ondervangen. Het is dan wel cruciaal om de prestaties van bestuurders te vergelijken met die van een relevante referentiegroep. In theorie moet die referentiegroep zo worden samengesteld dat zoveel mogelijk onbeheersbare factoren uit de prestaties van de bestuurders worden weggefilterd. Dit artikel beschrijft een exploratief onderzoek naar de samenstelling van de referentiegroepen van AEX-fondsen die gebruik maken van relatieve prestatiemeting. We vinden dat de samenstelling niet altijd optimaal is, maar er zijn ook geen duidelijke aanwijzingen dat ondernemingen bij de samenstelling strategisch te werk zijn gegaan.

1 Inleiding

Steeds meer Nederlandse ondernemingen kiezen ervoor hun bestuurders te belonen op basis van relatieve prestaties ten opzichte van een bepaalde referentiegroep in plaats van op absolute scores op prestatiemaatstaven. ABN AMRO was in 2000 één van de eerste ondernemingen in Nederland die relatieve prestatie-evaluatie (RPE) voor het topmanagement introduceerde:

'... qua Totaal Rendement voor Aandeelhouders (Total Return to Shareholders (TRS)) willen wij steeds tot de top 5 van onze peer group behoren. [...] Onze peer group vormt een goede afspiegeling van onze ambitie om het beter te doen dan superieure managementteams in onze sector. Hiermee leggen wij de meetlat een flink stuk hoger. In het verleden namen wij immers een tiende positie in binnen deze groep (gemeten naar TRS).' (Persbericht ABN Amro 17 augustus 2000)

De zogeheten 'peer group' is de referentiegroep waaraan de prestatie wordt afgemeten. De samenstelling van de referentiegroep is een keuze van de onderneming zelf. De theorie stelt dat een referentiegroep beter is samengesteld naarmate deze in sterkere mate overeenkomsten vertoont met de onderneming waarvan de prestaties moeten worden vastgesteld (Gibbons & Murphy, 1990). Ondernemingen kunnen bij de samenstelling van hun referentiegroep echter ook strategisch te werk gaan om zo hun relatieve prestaties voor het oog wat beter te doen uitkomen (Porac, *et al.*, 1999).
De Nederlandse corporate governance code ('Code Tabaksblat', zie Commissie corporate governance, 2003) stelt dat het 'best practice' is om de samenstelling van de referentiegroep openbaar te maken via het renumeratierapport van de Raad van Commissarissen (Best practice bepaling II.2.10). De navolging van deze best practice maakt het mogelijk de samenstelling van de referentiegroepen van beursgenoteerde Nederlandse ondernemingen nader te onderzoeken (zie Akkermans, *et al.*, 2006 voor de mate waarin de code wordt toegepast).
In dit artikel gaan wij na welke criteria AEX-fondsen gebruiken bij de samenstelling van hun referentiegroep, tot in hoeverre de daadwerkelijke samenstelling van hun referentiegroep inderdaad de genoemde criteria weerspiegelt en of er tekenen zijn van strategisch gedrag rondom de groepssamenstelling.

Drs. R.A.J. Haanen was ten tijde van het schrijven van dit artikel werkzaam als consultant bij Ernst & Young. Hij leverde zijn bijdrage op persoonlijke titel. Dit artikel is gebaseerd op zijn afstudeerscriptie.
Drs. V.S. Maas en dr. ir. S.P. van Triest zijn beiden als universitair docent management accounting verbonden aan de Universiteit van Amsterdam Business School.

Het artikel is als volgt opgebouwd. In paragraaf 2 bespreken we de theorie achter de relatieve prestatiemeting. Vervolgens kijken we in paragraaf 3 naar de criteria die de AEX-fondsen hebben gebruikt bij de samenstelling van hun referentiegroep. In paragraaf 4 beoordelen we de samenstelling van de referentiegroepen op vier criteria (sector, omvang, geografische spreiding en systematisch risico) en in paragraaf 5 onderzoeken we of er indicaties zijn van strategisch gedrag. Paragraaf 6 vat onze bevindingen samen in enkele concluderende opmerkingen.

2 Theorie

Met de scheiding tussen eigendom en bestuur van ondernemingen ontstaat het vraagstuk hoe ondernemingsbestuurders ertoe kunnen worden gebracht om te handelen in het belang van de eigenaars (aandeelhouders) in plaats van in hun eigenbelang. Een belangrijke oplossing voor dit vraagstuk wordt gevonden in prestatiebeloningscontracten. Door bestuurders te belonen voor gedrag dat de aandeelhouders ten goede komt, worden de belangen van beide partijen op één lijn gebracht. Met de keuze voor een prestatiebeloningssysteem ontstaat echter direct een nieuw probleem, namelijk dat van de keuze van een prestatiemaatstaf aan de hand waarvan de bestuurder zal worden beloond. De kern van dit probleem is dat het veelal onmogelijk zal zijn een maatstaf te vinden die de bijdrage van een bestuurder aan de waardegroei van de onderneming in een bepaalde periode perfect weergeeft. Ondernemingen zullen dus noodgedwongen moeten vertrouwen op prestatie-*indicatoren*: maatstaven die wel een indicatie geven van het presteren van een bestuurder maar die daarnaast ook beïnvloed worden door allerlei andere factoren. Bekende voorbeelden van dergelijke indicatoren zijn winst, rendement op eigen of op totaal vermogen en ontwikkeling van de aandelenkoers.

De economische literatuur (bijv. Holmström, 1979; Banker & Datar, 1989) stelt dat de geschiktheid van een maatstaf om te dienen als prestatie-indicator afhangt van de mate waarin deze maatstaf informatief is over het handelen van de bestuurder. De informatiewaarde (*informativeness*) van een maatstaf wordt daarbij in belangijke mate bepaald door de hoeveelheid ruis die deze bevat. Ruis in een maatstaf ontstaat doordat allerlei factoren buiten het handelen van de bestuurder van invloed zijn op deze maatstaf. De winst van een onderneming kan bijvoorbeeld worden beïnvloed door wetswijzigingen, macro-economische factoren en technologische ontwikkelingen buiten de invloedssfeer van het management (Merchant & Van der Stede, 2003).

De ruis in prestatiemaatstaven, veroorzaakt door voor bestuurders onbeheersbare factoren, zorgt ervoor dat de beloning van deze bestuurders op onvoorspelbare wijze kan gaan schommelen. Dit is om tenminste twee redenen problematisch. In de eerste plaats brengen de schommelingen risico met zich voor de bestuurders voor wat betreft hun periodieke inkomen. Omdat bestuurders over het algemeen risico-avers zijn, zullen ze voor dit risico gecompenseerd willen worden en bijvoorbeeld naast hun variabele beloning een hoger vast salaris eisen[1]. Per saldo zou de onderneming hierdoor zelfs slechter af kunnen zijn dan in een situatie waarin zij ervoor kiest om helemaal geen variabele beloning toe te kennen. Dit is de klassieke afweging tussen risico en *incentives* die het hart vormt van de economische agency theorie (zie bijvoorbeeld Indjejikian, 1999).

De tweede reden dat beloningsschommelingen veroorzaakt door ruis in prestatie-indicatoren problematisch zijn, is dat zij kunnen leiden tot gevoelens van oneerlijkheid. Een bestuurder die 'alles uit de kast' heeft gehaald om een doelstelling te halen maar zijn bonus misloopt door factoren buiten zijn invloedssfeer, zal zich gefrustreerd voelen en deze frustratie zal groter zijn naarmate dit vaker voorkomt en er een grotere discrepantie is tussen de 'werkelijke' en de gemeten prestaties (Hopwood, 1972; Cohen-Charash & Spector, 2001). Gefrustreerde bestuurders presteren slechter, voelen minder binding met de onderneming en gaan eerder op zoek naar een nieuwe functie (zie bijv. Aquino, *et al.*, 1997). Ruis in prestatiemaatstaven kan er ook voor zorgen dat bestuurders ondanks slecht presteren toch hoge beloningen ontvangen. Dit kan leiden tot onvrede bij werknemers, maar ook bij klanten en andere *stakeholders* van de onderneming (Fehr & Schmidt, 2003; Van de Wetering, *et al.*, 2003)

Eén mogelijkheid om ruis weg te filteren is het gebruik van relatieve prestatie-evaluatie (Gibbons & Murphy, 1990). Bij relatieve prestatie-evaluatie wordt niet gekeken naar de absolute scores op prestatie-indicatoren maar naar relatieve scores ten opzichte van een bepaalde referentiegroep, vaak aangeduid als de 'peer group'. De bonus van een bestuurder hangt er dus bijvoorbeeld vanaf of hij er in slaagt een winst per aandeel te laten zien die tot de 25% hoogste van zijn referentiegroep behoort. Het idee is dat de ondernemingen in de referentiegroep deels dezelfde externe invloeden zullen ervaren als de betreffende onderneming. Een onverwacht gunstig macro-economisch klimaat leidt bij relatieve prestatie-evaluatie dus bijvoorbeeld niet meer automatisch tot een hogere beloning voor bestuurders.

De keuze van een goede referentiegroep is uiteraard cruciaal bij de toepassing van relatieve prestatie-evaluatie. De ideale referentiegroep bestaat daarbij uit ondernemingen die identiek reageren op onbeheersbare factoren, of het nu gaat om een stijging van de olieprijs, de ontwikkeling van de inflatie of het mooie zomerweer. Vier factoren die iets zeggen over de mate waarin ondernemingen identieke of juist verschillende externe invloeden ervaren zijn de sector, de omvang, de geografische spreiding van de omzet en het systematisch risico (zie bijv. Gibbons & Murphy, 1990; Porac*, et al.*, 1999).

1 *Sector:* ondernemingen in dezelfde sector ervaren de invloed van dezelfde factoren in min of meer dezelfde mate. Zo is bijvoorbeeld de luchtvaartsector gevoelig voor een stijging van de olieprijs en worden de financiële prestaties van verzekeringsmaatschappijen in relatief sterke mate beïnvloed door (natuur)rampen.
2 *Omvang*: de omvang van ondernemingen wordt in vrijwel al het economisch en bedrijfskundig onderzoek meegenomen als control-variabele. De omvang van een onderneming is indicatief voor veel andere factoren, zoals bijvoorbeeld marktmacht, kwaliteit van het management, innovatiemogelijkheden, en in het algemeen de beschikbaarheid van middelen (in termen van geld, organisatievermogen en talent) om te reageren op invloeden van buitenaf.
3 *Geografische spreiding van omzet*: economische omstandigheden, marktontwikkelingen en politieke veranderingen kunnen sterk verschillen per regio.
4 *Systematisch risico:* de gevoeligheid voor marktontwikkelingen, zoals gemeten met het systematisch risico, is feitelijk een samenvatting van de inschatting die beleggers maken van de invloed van de markt op de onderneming.

3 Samenstelling van de referentiegroepen

In dit onderzoek zijn alle ondernemingen uit de AEX-index per 1 februari 2006 opgenomen die in hun jaarverslag aangeven gebruik te maken van een referentiegroep voor het vaststellen van de prestatiegerelateerde langetermijnbeloning. AEX-fondsen zijn met ingang van het verslaggevingsjaar 2004 onderworpen aan de principes en best practices van de Commissie Corporate Governance (2003). Derhalve wordt er over het verslaggevingsjaar 2004 voor het eerst verplicht gerapporteerd over de samenstelling van de referentiegroepen. We onderzoeken deze samenstelling, waarbij we van de referentieondernemingen de gegevens per 31 december 2003 nemen.
Het onderzoek heeft plaatsgevonden in de vorm van een archiefdata-onderzoek. Voor het verzamelen van de benodigde gegevens voor dit onderzoek is gebruik gemaakt van Bloomberg en S&P Market Insight, alsmede van jaarverslagen, websites en remuneratierapporten. Alle financiële gegevens zijn omgerekend naar één munteenheid, US Dollars. Dit is nodig omdat in de referentiegroepen ook ondernemingen zitten die niet in euro's rapporteren (vooral ondernemingen uit de VS en Groot-Brittanië).[2]
Per 1 februari 2006 maken 23 fondsen deel uit van de AEX-index. Van deze fondsen rapporteren er 19 dat gebruik wordt gemaakt van RPE voor de beloning van de leden van hun Raad van Bestuur. De vier ondernemingen die geen gebruik maken van RPE zijn Hagemeyer, Reed Elsevier, SBM Offshore en Vedior.
In tabel 1 is aangegeven welke prestatiemaatstaven de 19 ondernemingen gebruiken, hoe groot hun referentiegroep is en welke criteria ten behoeve van de samenstelling worden vermeld. Een eerste opmerking bij deze tabel is dat Getronics zich niet houdt aan de best practice bepaling en geen informatie verschaft over de samenstelling van de referentiegroep. Vervolgens zien we dat de grootte van de referentiegroep aanzienlijk verschilt. Zo bestaat die van Shell uit 3 ondernemingen, terwijl Fortis 30 vergelijkbare ondernemingen onderscheidt. De gehanteerde prestatiemaatstaf op basis waarvan toekenning van de langetermijnbeloning geschiedt is in 16 van de 19 gevallen (84%) de ontwikkeling van het totale rendement voor aandeelhouders (Total Shareholder Return of TSR)[3]. Drie ondernemingen hanteren een andere prestatiemaatstaf: Fortis gebruikt 'share performance', ASML rekent af op 'Return On Invested Capital (ROIC)', en Getronics vermeldt slechts dat zij kijkt naar de winst per aandeel en dividend. De tabel laat voorts zien dat vijf ondernemingen niet aangeven welke samenstellingscriteria zij hebben gebruikt. De andere ondernemingen vermelden of omschrijven hun gebruikte criteria (in het geval van Shell bijvoorbeeld 'de andere grote oliemaatschappijen'). Als enige onderneming gebruikt Heineken een criterium dat afwijkt, namelijk 'dezelfde categorie aandeelhouders'.

Bij het bepalen van de definitieve totale beloning wordt door alle ondernemingen een ranglijst opgesteld aan de hand van de prestatiemaatstaf, waarbij de relatieve positie van de onderneming bepalend is voor de toekenning van de beloning. De periode die hierbij in beschouwing wordt genomen is bij 17 van de 19 ondernemingen een periode van drie jaar. ABN AMRO hanteert een periode van 4 jaar, Getronics vermeldt de duur van de periode die zij hierbij in beschouwing neemt niet.

Tabel 1 Algemene kenmerken van de referentiegroepen

Onderneming	Prestatiemaatstaf[b]	Aantal ondernemingen in referentiegroep	Criterium samentselling		
			sector	omvang	geografisch
ABN AMRO	TSR	20	x		
Aegon	TSR	11	x	x	x
Ahold	TSR	9	x		
Akzo Nobel	TSR	16			
ASML	ROIC	10			
Buhrmann	TSR	11	x	x	x
DSM	TSR	11	x		
Fortis	'share performance'	30	x	x	
Getronics	'winst per aandeel'	n.b.			
Heineken[a]	TSR	11			
ING	TSR	19	x		x
KPN	TSR	15	x		
Numico	TSR	18			
Philips	TSR	23	x		x
Shell	TSR	3	x	x	
TNT	TSR	4	x		
Unilever	TSR	20			
VNU	TSR	10	x		
Wolters Kluwer	TSR	15	x		
# ondernemingen			13	4	4

Noten
a Heineken hanteert als criterium 'dezelfde groep aandeelhouders'.
b ROIC (return on invested capital) wordt door ASML gedefinieerd als nettowinst na belastingen gedeeld door gemiddeld geïnvesteerd kapitaal, Fortis en Getronics geven geen verdere invulling van hun maatstaf.

4 Analyse samenstellingcriteria

Sector

De sector waarin een onderneming actief is, is geen objectief gegeven. Er moet een keuze gemaakt worden voor een bepaalde sector, waarbij het vaak voorkomt dat de activiteiten die een onderneming uitvoert zo divers zijn dat ze niet in één sector te plaatsen zijn. Daarom gebruiken we standaard sectorindelingen. Sectorindelingen ('industry classifications') worden gehanteerd om ondernemingen te groeperen op basis van hun activiteiten. De meest gebruikte standaard is de Standard Industry Classification (SIC). SIC-codes kennen vier niveaus, waarbij de tweecijferige aanduiding in het algemeen als de relevante sector wordt genomen (Porac *et al*, 1999). Het meest nauwkeurige niveau is een aanduiding met vier cijfers, maar dit is vaak te specifiek om alle activiteiten van grote ondernemingen te omvatten.

Om te beoordelen of de resultaten robuust zijn passen we ook een andere indeling toe, de Global Industry Classification Standard (GICS), opgesteld door de financiële dienstverleners Morgan Stanley en Standard & Poor's (zie Bhojraj, *et al.*, 2003). Bij dit systeem bestaat de breedste classificatie ('sector') uit twee cijfers en de meest verfijnde classificatie (sub-industry) uit acht cijfers. Het GICS systeem classificeert ondernemingen op basis van hun primaire bedrijfsactiviteit (gemeten aan de hand van de omzet en winst), evenals aan de hand van de perceptie van de markt (zoals weerspiegeld in analistenrapporten). We kijken zowel naar de gangbare sectorindeling (de twee-cijferige SIC-code en de 'sector'-code volgens GICS) als naar de meest gedetailleerde indeling die de classificaties mogelijk maken. De resultaten staan in tabel 2, die het percentage ondernemingen in de referentiegroep weergeeft dat actief is in dezelfde sector als de onderneming waar het om gaat.

Tabel 2 laat duidelijk zien dat de ondernemingen bij de samenstelling van de referentiegroep rekening hebben gehouden met hun bedrijfstak, zoals zij zelf ook in grote meerderheid al aangaven. Opvallend is verder dat de mate van homogeniteit van de referentiegroep afhankelijk is van de gebruikte indeling, waarbij de matching volgens 'GICS sector'-indeling

Tabel 2 Vergelijking met referentiegroep op basis van sector

Onderneming	SIC		GICS	
	2 cijfers	4 cijfers	Sector	Sub-industry
ABN AMRO	75	75	100	50
Aegon	91	64	100	45
Ahold	78	67	100	56
Akzo Nobel	100	25	69	31
ASML	70	70	100	100
Buhrmann	36	0	55	55
DSM	100	0	100	55
Fortis	63	63	100	3
Heineken	73	45	91	45
ING	42	26	100	11
KPN	100	60	100	73
Numico	72	39	94	61
Philips	52	13	30	22
Shell	100	100	100	100
TNT	50	50	100	100
Unilever	40	20	95	25
VNU	30	20	80	60
Wolters Kluwer	87	0	100	100
Gemiddeld	**70**	**41**	**90**	**55**
Minimum	30	0	30	3
Maximum	100	100	100	100

Tabel 3 Vergelijking met referentiegroep op basis van omvang

Onderneming	Markt-waarde	Boek-waarde	Omzet	Gemiddelde rang
ABN AMRO	55	55	60	57
Aegon	55	45	55	52
Ahold	55	67	78	67
Akzo Nobel	81	75	81*	79
ASML	90*	90*	90*	90
Buhrmann	36	73*	82*	64
DSM	36	55	45	45
Fortis	67*	57*	53	59
Heineken	55	45	45	48
ING	74*	74*	95*	81
KPN	53	53	53	53
Numico	11*	6*	6*	8
Philips	64	70	70*	68
Shell	33	33	33	33
TNT	25	25	25	25
Unilever	75*	90*	90*	85
VNU	80*	90*	67	79
Wolters Kluwer	53	73	57	61
Gemiddeld	**55**	**60**	**60**	**59**

succesvoller is dan volgens het meer traditionele SIC systeem. Elf van de achttien ondernemingen hebben volgens de GICS indeling een referentiegroep die volledig in dezelfde sector actief is en gemiddeld is de overeenstemming 90%. Voor de tweecijferige SIC-indeling is er slechts bij vier ondernemingen sprake van een volstrekt homogene referentiegroep, en het gemiddelde is met 70% duidelijk lager. De resultaten voor de SIC-indeling komen overeen met die van Porac et al. (1999), die ook vinden dat gemiddeld 70% van de referentieondernemingen uit dezelfde SIC-sector komen. Niet verrassend is dat de subclassificaties het slechter doen, hoewel ook hier de GICS-indeling een grotere overeenstemming laat zien. Wel opvallend is echter dat voor vijf ondernemingen de 'GICS sub-industry'-indeling het beter doet dan de tweecijferige SIC-indeling. De betere prestaties van de GICS-indeling zijn waarschijnlijk het gevolg van de jongere leeftijd van het systeem, waardoor het beter aansluit op de huidige verhoudingen in het bedrijfsleven.

Omvang

We bekijken de referentiegroepen voor drie verschillende indicatoren van de omvang van de onderneming: de marktwaarde, de boekwaarde en de omzet. Voor het beoordelen van de mate van overeenstemming met de ondernemingen in de referentiegroep kiezen we twee benaderingen. In de eerste plaats bekijken we of de omvang van de onderneming statistisch verschilt van het gemiddelde van de referentiegroep[4]. Dat doen we middels een *t*-toets op de waardes van de referentiegroep, waarbij de omvang van de onderneming de toetswaarde is[5]. Daarnaast bekijken we de rang die de onderneming in haar referentiegroep inneemt, op een schaal van klein naar groot (de grootste onderneming heeft dus hoogste rang). Deze rang is gestandaardiseerd op een schaal van 0 tot 100, middels de volgende formule: score = 1 – ((rang – 1)/(groepsgrootte –1)). Op deze wijze wordt de relatieve positie duidelijk (een onderneming die rang 2 scoort in groepsgrootte 3 zit precies op de helft).

De resultaten staan in tabel 3. De scores in deze tabel geven de relatieve positie van de onderneming in de referentiegroep weer, op een schaal van 0 tot 100. Een * geeft aan dat de onderneming significant verschilt van het groepsgemiddelde op basis van een *t*-toets met een kans van 5%. De laatste kolom geeft de gemiddelde rang weer op basis van de drie maatstaven samen.

Tabel 3 laat uitschieters naar boven zien (ASML, Unilever) die duidelijk groter zijn dan de referentieondernemingen, en naar beneden (waar met name Numico er uit springt). Het algehele beeld is echter redelijk evenwichtig. De gemiddelde rang is ongeveer 60, wat betekent dat de ondernemingen iets groter zijn dan hun

gemiddelde referentieonderneming. Ook de *t*-toetsen zijn redelijk stabiel. De helft van de ondernemingen wijkt op geen enkele maatstaf statistisch significant af. De rang en de resultaten van de *t*-toetsen laten zien dat ondernemingen kiezen voor een referentiegroep die voor wat betreft de gemiddelde omvang redelijk overeenkomt met henzelf. Opvallend is wel dat twee van de vier ondernemingen die aangeven omvang doelbewust als selectiecriterium te hebben gebruikt (Buhrman en Fortis), voor wat betreft twee indicatoren, significant groter zijn dan hun gemiddelde referentieonderneming.

Geografische spreiding

Vier ondernemingen geven aan hun referentiegroep mede te hebben geselecteerd op basis van overeenkomst in de geografische spreiding van de activiteiten. Het beoordelen van de overeenkomst in geografische spreiding blijkt echter zeer lastig omdat er geen database beschikbaar is met objectieve gegevens over de geografische spreiding van ondernemingsactiviteiten. De enige beschikbare informatie betreft de rapportage daaromtrent door de ondernemingen zelf. Niet alle ondernemingen publiceren echter informatie over hun geografische spreiding, en waar er wel gegevens openbaar worden gemaakt gebeurt dit op vele verschillende manieren. Zowel het type informatie dat wordt verstrekt (soms omzet, soms winst) als de gebruikte indeling verschilt. De ene onderneming verdeelt haar omzet heel nauwkeurig over alle landen waarin zij actief is, andere ondernemingen groeperen deze omzet over continenten, of zelfs continenten overstijgende geografische gebieden ('Americas' als verzamelnaam voor alle landen in Noord- en Zuid-Amerika bijvoorbeeld). Ook komt het veelvuldig voor dat ondernemingen het grootste gedeelte van hun omzet nauwkeurig toerekenen aan enkele landen en/of regio's en de rest van hun omzet toerekenen aan 'de rest van de wereld' of 'overig'. Het gebrek aan duidelijke gegevens maakt het dus onmogelijk om op basis van een deugdelijke analyse een conclusie te trekken over de overeenkomst tussen ondernemingen en hun referentiegroep voor wat betreft de geografische spreiding van hun activiteiten.

Systematisch risico

Het systematisch risico van een onderneming wordt gemeten door middel van de beta uit het Capital Asset Pricing Model (CAPM, zie bijv. Brealey & Myers, 2002). De beta van een onderneming kan worden geschat door middel van een regressie-analyse van de rendementen van de onderneming en de rendementen van de markt als geheel. In het CAPM is de markt als geheel de gewogen portefeuille van alle activa, zowel verhandeld (zoals aandelen en obligaties) als niet verhandeld (zoals niet-beursgenoteerde ondernemingen en 'human capital'). In de Bloomberg-database die voor dit artikel is gebruikt worden beta's berekend op basis van locale indices. Zoals Koller *et al.* (2005) aangeven zijn de meeste landen oververtegenwoordigd in enkele sectoren, waardoor de lokale index sterk beïnvloed wordt door deze sectoren, en in sommige gevallen zelfs door slechts enkele bedrijven. Indien een lokale index gehanteerd wordt voor het schatten van een beta, wordt daardoor de gevoeligheid van die onderneming gemeten ten opzichte van deze sectoren of bedrijven, niet het systematisch risico van deze onderneming. Dit is een wezenlijk aandachtspunt, omdat alle referentiegroepen een internationale samenstelling kennen. Daarom berekenen we de beta voor iedere onderneming zelf, met als benadering voor de marktindex de MSCI World Index van Morgan Stanley, een gewogen index die bestaat uit de aandelen van grote ondernemingen uit 23 geïndustrialiseerde landen. De geschatte beta is de 'equity beta'. Dit is de relevante risicomaatstaf, omdat deze het risico meet waaraan aandeelhouders worden blootgesteld (Brealey & Myers, 2002).

Met een regressie-analyse is voor iedere doelonderneming en haar referentiegroep de beta geschat op basis van de maandelijkse rendementen ten opzichte van de MSCI World Index over een periode van 5 jaar. Daarbij is een ondergrens gehanteerd voor de R^2 van de regressie van 10%, zodat er sprake is

Tabel 4 Vergelijking met referentiegroep op basis van systematisch risico (beta)

Onderneming	Beta (rang)
ABN AMRO	76
Aegon	100*
Ahold	67
Akzo Nobel	75
ASML	50
Buhrmann	63
DSM	25
Fortis	39*
Heineken	n/b
ING	58
KPN	92*
Numico	n/b
Philips	79
Shell	100
TNT	n/b
Unilever	n/b
VNU	43
Wolters Kluwer	n/b
Gemiddeld	67

van een zekere minimale statistische verklaring. Deze methode leverde voor 13 van de 18 ondernemingen een beta op. De resultaten van een rangvergelijking van deze beta met de gemiddelde beta van de ondernemingen in de referentiegroepen staan in tabel 4. Van de ondernemingen waarvoor een beta kon worden bepaald blijken er slechts drie een beta te hebben die significant afwijkt van die van hun referentiegroep. Het systematisch risico van Fortis is lager dan gemiddeld terwijl het risico van Aegon en KPN hoger is dan gemiddeld. De gemiddelde rang voor alle ondernemingen samen is met 67 tamelijk hoog. Desalniettemin lijkt het gemiddeld risico van de meeste ondernemingen redelijk overeen te komen met dat van hun referentiegroep.

5 Samenstelling van referentiegroep versus relatieve prestatie

Uiteindelijk gaat het bij RPE om de relatieve prestatie van de onderneming ten opzichte van de referentiegroep. We vergelijken de prestaties van de AEX ondernemingen en hun referentiegroepen over de periode 1999-2003 en over de periode 2001-2003. Zoals vermeld in paragraaf 3 hanteren de meeste ondernemingen (17 van de 19) een periode van 3 jaar voor de beoordeling van de relatieve prestatie. De langere periode 1999-2003 wordt meegenomen om iets te zeggen over de stabiliteit van de relatieve prestatie (waarbij opgemerkt moet worden dat de keuze voor vijf jaar arbitrair is). Omdat de meeste ondernemingen zich richten op de Total Shareholder Return (TSR) nemen we deze indicator ook als vergelijkingsmaatstaf. De resultaten staan in tabel 5. Uit deze tabel blijkt duidelijk dat de AEX-ondernemingen veelal zeer matig scoren. Als we als bovengemiddelde prestatie een positie bij de beste helft nemen (dus rang 50 en hoger) zien we dat over de laatste vijf jaar gemeten slechts drie ondernemingen daarbij horen (DSM, Philips en VNU). Daarbij moet opgemerkt worden dat de lat vaak hoger wordt gelegd, zoals uit het voorbeeld van ABN AMRO blijkt: een top 5 positie in een groep van 20 betekent een rang van 80.[6] Over de periode 2001 – 2003 zijn de relatieve prestaties niet veel beter: slechts ABN AMRO, DSM, Philips en TNT zitten bij de beste helft van hun referentiegroep.

Tabel 5 Vergelijking prestaties (TSR) met referentiegroep

Onderneming	TSR 2001 – 2003	TSR 1999 – 2003
ABN AMRO	60	45
Aegon	9	0
Ahold	0	0
Akzo Nobel	25	33
ASML	0	25
Buhrmann	9	20
DSM	90	100
Fortis	36	18
Heineken	18	22
ING	37	42
KPN	43	0
Numico	0	6
Philips	60	81
Shell	0	0
TNT	50	0
Unilever	40	10
VNU	30	50
Wolters Kluwer	13	0
Gemiddeld	29	25

De vraag is nu of er sprake is van strategisch gedrag bij de samenstelling van de referentiegroep. Strategisch gedrag zal er in het algemeen op gericht zijn om een referentiegroep te vinden ten opzichte waarvan de relatieve prestaties van de onderneming bovengemiddeld zijn. Het ligt dan het meest voor de hand de referentiegroep te laten bestaan uit fondsen met historisch gezien relatief lage rendementen (Porac, *et al.*, 1999). Voor ondernemingen die voor het eerst RPE toepassen kan het echter ook aantrekkelijk zijn om een referentiegroep te kiezen ten opzichte waarvan men op basis van historische gegevens juist relatief slecht heeft gepresteerd. Het is immers makkelijker om procentueel veel te stijgen vanaf een relatief lager koersniveau. Uiteraard zal in beide gevallen het strategisch gedrag ten koste gaan van de representativiteit van de referentiegroep.
Wel is de vrijheid van ondernemingen om bepaalde ondernemingen al dan niet in hun referentiegroep op te nemen enigszins beperkt doordat analisten, beleggers en andere *stakeholders* goed bekend zijn met de spelers op de markten waarop de betreffende onderneming concurreert. Het zou bijvoorbeeld opmerkelijk zijn als ING niet tot de referentiegroep van ABN-AMRO zou behoren en andersom. Een onderneming die in een dergelijke situatie toch strategisch zou willen handelen, zou dat dan alleen kunnen doen door meer ondernemingen aan haar referentiegroep toe te voegen en zo de gemiddelde prestaties te beinvloeden. Concluderend kunnen we stellen dat ondernemingen hun relatieve prestaties voor het oog kunnen opvijzelen door hun referentiegroep groter te maken en dan in het bijzonder door minder representatieve maar relatief slecht (of juist goed) presterende ondernemingen toe te voegen.
Om te kijken of er inderdaad een verband is tussen de relatieve prestaties van ondernemingen en de

Tabel 6 Spearman correlaties relatieve prestaties en kenmerken referentiegroep

	Grootte referentiegroep	Homogeniteit (% ondernemingen in zelfde sector)		Gemiddelde rang omvang
		SIC 2 cijfers	GICS sector	
rang TSR 01 – 03	.447*	–.096	.001	–.004
rang TSR 99 – 03	.350	–.272	–.417*	.374

omvang en de representativiteit van hun referentiegroep vergelijken we de rang van de TSR-prestatie van de onderneming met het aantal ondernemingen in de referentiegroep, het percentage referentie-ondernemingen in dezelfde sector en de gemiddelde omvang van de onderneming ten opzichte van de referentie-ondernemingen. De resultaten staan in tabel 6. Deze tabel geeft de verdelingsvrije Spearman correlaties weer tussen de relatieve prestatie op basis van 3- en 5-jaars TSR-rang en de kenmerken van de referentiegroep (correlaties met een * zijn significant met $p < .10$).
De positieve correlatie tussen het aantal ondernemingen in de referentiegroep en de 3-jaars TSR impliceert dat ondernemingen die meer referentie-ondernemingen toevoegen een betere relatieve prestatie laten zien. Dit zou kunnen wijzen op het toevoegen van ondernemingen in de referentiegroep om de relatieve positie te verbeteren. Als dat echter al gebeurt, dan niet doordat er buiten de sector wordt gezocht naar geschikte kandidaten. Gezien het ontbreken van een verband met de rang van de omvang wordt er evenmin naar 'kleine visjes' gespeurd. De negatieve correlatie tussen de homogeniteit volgens de 'GICS sector'-indeling en de 5-jaars TSR betekent dat de keuze voor een referentiegroep met meer ondernemingen buiten de eigen sector leidt tot een hogere relatieve prestatie. Dit komt overeen met de argumentatie van Porac et al. (1999). Aangezien de sectorindeling niet beïnvloedbaar is door de onderneming zelf is deze relatie te verklaren door aan te nemen dat ondernemingen buiten hun 'juiste' referentiegroep gaan zoeken naar ondernemingen die hen beter af doen zijn.

De aanwijzingen voor strategisch gedrag in het samenstellen van de referentiegroep zijn derhalve beperkt. Hoewel er correlaties gevonden worden die op strategisch gedrag kunnen wijzen, worden die niet voortdurend waargenomen: ze hangen af van de operationalisatie van de prestatie (3 of 5 jaar) en van de sectorhomogeniteit (SIC of GICS). Omdat de groepssamenstelling niet wijzigt (het gaat steeds om dezelfde referentiegroep per onderneming) is het moeilijk een eenduidig patroon in eventueel strategisch gedrag vast te stellen. Bovendien is de uiteindelijke relatieve prestatie van de meeste ondernemingen zeer matig, zoals tabel 5 laat zien. Als er echt gezocht zou zijn naar een relatief slecht presterende referentiegroep om zo de eigen prestatie beter te laten lijken (wat de theoretische verklaring is voor de twee significante correlaties) zijn de effecten van dit strategisch gedrag dus zeer beperkt.

6 Conclusie

Relatieve prestatie-evaluatie (RPE) wordt veel toegepast door AEX-fondsen bij de beloning van hun topbestuurders. RPE vereist een representatieve 'peer group': de groep ondernemingen ten opzichte waarvan de relatieve prestaties worden gemeten. In dit onderzoek hebben we gekeken in hoeverre de referentiegroepen van 18 AEX-fondsen representatief mogen worden geacht. De kwaliteit van de referentiegroepen blijkt redelijk. Ze bestaan voor het belangrijkste deel uit ondernemingen binnen dezelfde sector. Met name wanneer uitgegaan wordt van het GICS-classificatiesysteem is de homogeniteit groot. Ook voor wat betreft de omvang passen ondernemingen goed in hun referentiegroep. In hoeverre ondernemingen en hun referentieondernemingen overeenkomen voor wat betreft de geografische spreiding van hun activiteiten is niet goed vast te stellen, wat betekent dat dit criterium voor ondernemingen zelf ook moeilijk te gebruiken is. Voor wat betreft het systematisch risico tot slot, blijkt dat ondernemingen over het algemeen een iets hogere beta hebben dan hun referentiegroep, maar dit verschil is slechts in een enkel geval significant.
Ons onderzoek laat verder zien dat er weinig aanwijzingen zijn voor strategisch gedrag bij de samenstelling van de referentiegroepen. Ook moet worden opgemerkt dat voor zover er al sprake was van strategisch gedrag, dit weinig effect heeft gehad. De meeste onderzochte ondernemingen presteren namelijk slecht ten opzichte van hun referentiegroep. Toekomstig onderzoek zal moeten uitwijzen hoe ondernemingen reageren op langere periodes van relatief slechte prestaties. Aanpassing van de referentiegroep is een gemakkelijke manier om de eigen resultaten voor het oog te verbeteren. Ons onderzoek

laat daarbij zien dat zelfs wanneer ondernemingen de Code Tabaksblat volgen en de samenstelling van hun referentiegroep openbaar maken, het moeilijk is overtuigend vast te stellen dat zij daarbij strategisch te werk zijn gegaan. ■

Literatuur

Akkermans, D., H. Van Ees, N. Hermes, R. Hooghiemstra, G. Van der Laan, T. Postma, en A. Van Witteloostuijn, (2006), De code Tabaksblat: een jaar verder, in: *Maandblad voor Accountancy en Bedrijfseconomie,* vol. 80, no. 5, pp. 219-231.

Aquino, K., R.W. Griffeth, D.G. Allen, en P.W. Hom, (1997), Integrating justice constructs into the turnover process: a test of a referent cognitions model, in: *Academy of Management Journal,* vol. 40, no. 5, pp. 1208-1227.

Banker, R.D. en S.M. Datar, (1989), Sensitivity, precision, and linear aggregation of signals for performance evaluation, in: *Journal of Accounting Research*, vol. 27, no. 1, pp. 21-39.

Bhojraj, S., C. Lee en D.K. Oler, (2003), What's my line? A comparison of industry classification schemes for capital market research, in: *Journal of Accounting Research,* vol. 41, no. 5, pp. 745-774.

Brealey, R.A. en S.C. Myers, (2002), *Principles of Corporate Finance,* McGraw-Hill, New York.

Cohen-Charash, Y. en P.E. Spector, (2001), The role of justice in organizations: a meta-analysis, in: *Organizational Behavior and Human Decision Processes, vol. 86*, no. 2, pp. 278-321.

Fehr, E. en K.M. Schmidt, (2003), Theories of fairness and reciprocity - evidence and economic applications, in: M. Dewatripont a.o., *Advances in Economic Theory: Eighth World Congress of the Economic Society*, Cambridge University Press, Cambridge, UK.

Gibbons, R. en K.J. Murphy, (1990), Relative performance evaluation for chief executive officers, in: *Industrial and Labor Relations Review, vol. 43*, pp. 30-51.

Commissie corporate governance, (2003), *De Nederlandse corporate governance code: beginselen van deugdelijk ondernemingsbestuur en best practice bepalingen*, te downloaden via www.commissiecorporategovernance.nl

Holmström, B., (1979), Moral hazard and observability, in: *Bell Journal of Economics*, vol. 10, pp. 74-91.

Hopwood, A.G., (1972), An empirical study of the role of accounting data in performance evaluation, in: *Journal of Accounting Research,* vol. 10, pp. 156-182.

Indjejikian, R.J., (1999), Performance evaluation and compensation research: an agency perspective in: *Accounting Horizons*, vol. 13, no. 2, pp. 147-157.

Koller, T., M. Goedhart en D. Wessels, (2005), *Valuation: measuring and managing the value of companies*, John Wiley & Sons, Hoboken, NJ.

Merchant, K. A. en W.A. Van der Stede, (2003), *Management Control Systems: Performance Measurement, Evaluation and Incentives*, Pearson Education, Harlow, Essex, UK.

Porac, J.F., J.B. Wade, en T.G. Pollock, (1999), Industry categories and the politics of the comparable firm in CEO compensation, in: *Administrative Science Quarterly,* vol 44, pp. 112-144.

Wetering, H. van de, I. de Loo, en H. Spekreijse, (2003), Prestatiebeloning van topbestuurders: eclatante zelfverrijking?, in: *Maandblad voor Accountancy en Bedrijfseconomie,* vol. 77, no. 11, pp. 500-508.

Noten

1 Ter illustratie: geconfronteerd met de keuze tussen € 100 met zekerheid of 50% op € 200 kiezen de meeste mensen voor de eerste optie. Om hen voor de tweede optie te laten kiezen zal de verwachte waarde van deze optie moeten stijgen tot boven de € 100.

2 Voor het berekenen van het systematisch risico en aandelenrendementen worden gegevens over meerdere jaren gebruikt. In die gevallen wordt er met de wisselkoersen per 31 december van elk jaar gewerkt.

3 Onder TSR wordt verstaan de ontwikkeling van de aandelenkoers over een bepaalde periode, inclusief herbelegd dividend:

$$TSR = \frac{(P_t - P_{t-1}) + Div_t}{P_{t-1}}$$

4 Voor deze en andere statistische toetsen geldt dat voorzichtigheid geboden is bij het interpreteren van de significanties, vanwege de grote verschillen in omvang van de referentiegroepen. Hierdoor verschilt bijvoorbeeld de boekwaarde van Fortis (rang 57, groepsgrootte 30) significant van het gemiddelde, en die van Akzo Nobel (rang 75, groepsgrootte 16) niet.

5 We transformeren de waardes door de logaritmes te nemen; dit corrigeert voor de grote spreiding in omvangmaatstaven.

6 De referentiegroep van ABN AMRO bevat 20 ondernemingen. Met de doelonderneming zelf zijn dat 21 resultaten, waarbij de top vijf de nummers 17 tot en met 21 beslaat.