

Ranglijsten voor beleggingsfondsen: Sharpe versus Life

Wat is het 'ranglijst-risico' van de Sharpe- en de Life-index?

Recentelijk hebben Van Bussel e.a. een geactualiseerde versie van hun overzicht van de prestaties van beleggingsfondsen gepresenteerd¹. Van Bussel e.a. introduceren een methode om de resultaten van beleggingsfondsen te rangschikken, op basis van de z.g. Life-index. De Life-index is een originele methode om resultaten van beleggers te rangschikken, zij het dat de methode voor een belangrijk deel gebruik maakt van reeds bekende methoden en technieken. De rangschikking geschiedt op basis van twee eigenschappen die voor de belegger aantrekkelijk worden geacht: de mate waarin het fonds, voor risico gecorrigeerd, buitengewoon rendement heeft behaald, en de mate waarin het fonds gediversificeerd is.

Bij de introductie van de Life-index is deze gepresenteerd als een alternatief voor de Sharpe-index². De Sharpe-index wordt onder meer door ABN Amro in samenwerking met het weekblad *Elsevier* gebruikt voor het opstellen van een ranglijst van beleggingsfondsen. De Sharpe-index wordt berekend als het verschil tussen het rendement op een portefeuille en het rendement op een risicovrije belegging, gedeeld door de standaarddeviatie van het portefeuille-rendement. In deze bijdrage willen wij de betrouwbaarheid van beide indexen toetsen.

De zin van een ranglijst

Het opstellen van een ranglijst kan zinvol zijn om twee redenen. In de eerste plaats kan het nuttig zijn om de resultaten uit het verleden te rangschikken, bij voorbeeld als verantwoording over het in het verleden gevoerde beleid. In de tweede plaats kan een ranglijst belangrijk zijn voor het selecteren van beleggingsfondsen als toekomstige belegging. Daar-

bij is het wenselijk dat de ordening van de fondsen in de evaluatieperiode ook in de volgende periode kan worden waargenomen. Dit verschijnsel wordt de persistentie van performance genoemd. Van Bussel e.a. hebben de persistentie van de performance onderzocht door de ranglijsten van twee afzonderlijke, onderling niet overlappende perioden te vergelijken. Zij komen tot de conclusie dat de ranglijsten nauwelijks of geen onderlinge samenhang vertonen. Daarmee kan het nut van het opstellen van een ranglijst ten behoeve van de selectie van fondsen voor toekomstige beleggingen sterk betwijfeld worden. Het zegt wel iets over de beleggingsperformance in het verleden, maar hoeveel?

Voor de beantwoording van deze vraag zullen we nagaan wat het zogenaamde 'onderscheidend vermogen' is van de genoemde ranglijsten. Wanneer het onderscheidend vermogen beperkt is, kan aan de betrouwbaarheid van de opgestelde ranglijst getwijfeld worden.

Onderscheidend vermogen

Het onderscheidend vermogen van een index heeft betrekking op de mogelijkheid om de kwaliteiten van de beheerders van het fonds te kunnen meten, waarbij de invloed van toevallige factoren zoveel mogelijk uitgesloten wordt. Naarmate de kwaliteiten van beheerders minder van elkaar verschillen, zal het meetinstrumentarium om deze verschillen te kunnen meten fijnzinniger moeten zijn. Op grond van de hypothese van de efficiënte vermogensmarkt kan verwacht worden dat significante verschillen in de kwaliteiten van beheerders weinig zullen optreden. (De efficiënte-markthypothese veronderstelt dat alle beschikbare informatie in de koersen is verwerkt, en dat het dus

onmogelijk is om onder- of overgevalueerde fondsen te selecteren.) Desalniettemin zal het ook onder deze omstandigheden mogelijk zijn om een ranglijst op te stellen. Aan de betrouwbaarheid (in statistische zin) van deze ranglijst kan echter getwijfeld worden: immers de opgetreden verschillen zullen naar alle waarschijnlijkheid slechts aan toeval te wijten zijn.

Ten einde de betrouwbaarheid van een ranglijst in kaart te brengen, kan men in principe twee wegen bewandelen. Zo kan men voor ieder paar fondsen afzonderlijk toetsen of de scores significant van elkaar verschillen. Voor de Sharpe-index is een dergelijke toets beschikbaar³. Hierbij wordt verondersteld dat de rendementen van de beleggingsfondsen normaal verdeeld zijn. Voor de Life-index is echter geen formele toets om vast te stellen of de scores significant verschillen.

Een alternatief is het gebruik van de z.g. 'bootstrap'-methode, waarmee eveneens betrouwbaarheidsintervallen kunnen worden opgesteld⁴. De bootstrap kan geïnterpreteerd worden als een simulatie, waarbij uit de gerealiseerde rentabiliteitsverdeling van de beleggingsfondsen steekproeven worden getrokken. Op basis van iedere steekproef kan vervolgens – na toepassing van de procedures om de verschillende indices te berekenen – een nieuwe ranglijst worden opgesteld. Per beleggingsfonds kan men nu een waarschijnlijkheidsverdeling bepalen, waarbij wordt aangegeven hoe groot de waarschijnlijkheid is dat het betreffende fonds op de ranglijst zal verplaatsen. In figuur 1 wordt zo'n waarschijnlijkheidsverdeling gepresenteerd voor het fonds dat op de zestiende plaats van de ranglijst staat. Zo blijkt de kans dat de ran-

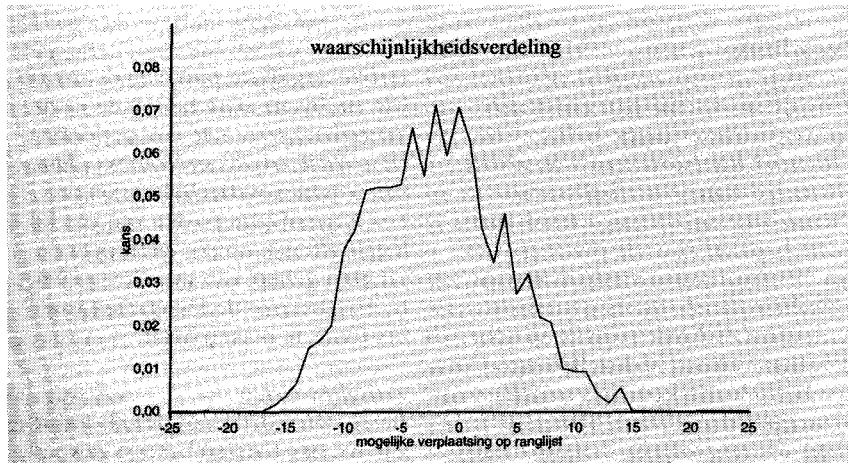
1. A. van Bussel, C.G. Koedijk, F. Nissen, L. Pijnenburg en P. Schotman, De prestaties van beleggingsfondsen 1989-94, *ESB*, 28 juni 1995, blz. 604-607.

2. A. van Bussel, J.-W. Goslings, C.G. Koedijk, P. Schotman en B. van Tuel, De Life-index voor beleggingsrendement, *ESB*, 13 juli 1994, blz. 643-645.

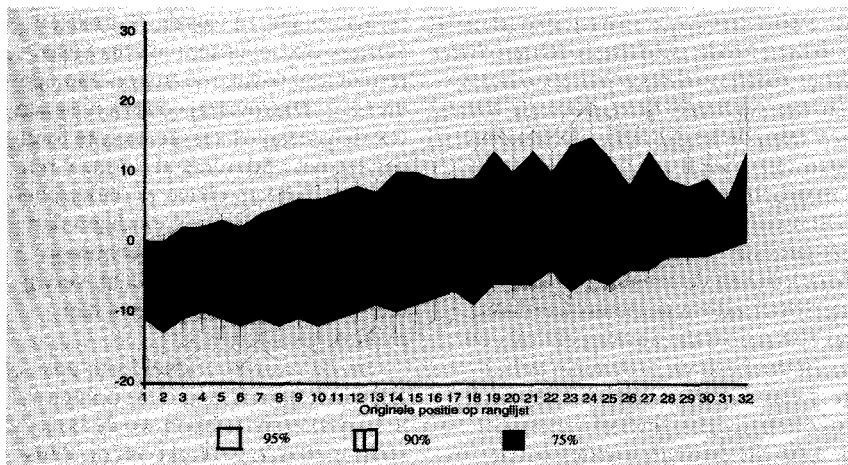
3. Zie Jobson en Korkie, Performance hypothesis testing with the Sharpe and Treynor measures, *Journal of Finance*, 1981, blz. 889-908.

4. Zie P. Diaconis en B. Efron, Computer-intensive methods in statistics, *Scientific American*, mei 1983.

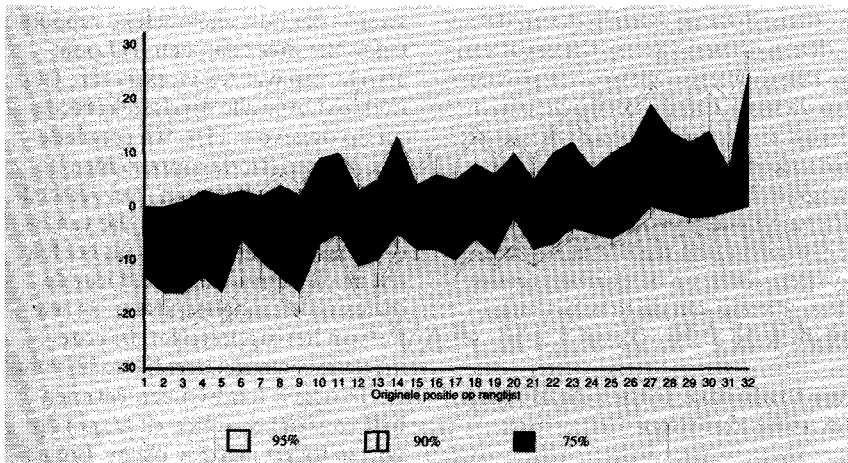
Figuur 1. Het Life ranglijst-risico voor een individueel beleggingsfonds



Figuur 2. Ranglijst-risico voor de Sharpe-index



Figuur 3. Ranglijst-risico voor de Life-index (alle beleggingsfondsen)



Tabel 1. Gemiddelde verplaatsing op de ranglijst

Betrouwbaarheidsniveau	AA Elsevier	Life
95%	22,9	22,9
90%	20,4	20,5
75%	15,4	15,1

king van het Transpacific Fund gelijk blijft ongeveer 8% bedraagt, terwijl 90% van de fluctuatie leidt tot een verplaatsing van ongeveer plus of min twaalf plaatsen op de ranglijst. Op grond van dit betrouwbaarheidsniveau kan dit beleggingsfonds bij een gelijkblijvende rendementsverdeling een notering krijgen tussen de vijfde en de 27ste plaats op de ranglijst. De verschillen tussen het ranglijst-risico blijkt voor beide indices in dezelfde orde van grootte te liggen.

Wanneer het onderscheidend vermogen van een ranglijst perfect is, dan zal voor ieder fonds in de simulatie de geobserveerde plaats met een kans van 100% gerealiseerd worden. Op basis van het fonds gepresenteerd in figuur 1 kan dus geconcludeerd worden dat het onderscheidend vermogen van de Life-index niet perfect is.

We hebben dit experiment uitgevoerd voor de Life-ranglijst van vorig jaar, en de gelijktijdig verschenen ranglijst van het weekblad *Elsevier*. Een probleem bij het vergelijken van deze ranglijsten is dat de Life-Index over zes jaar wordt berekend, terwijl de ranglijst van *Elsevier* over drie jaar wordt berekend. Ten einde tot vergelijkbare resultaten te komen hebben we de Life-index opnieuw berekend over een periode van eveneens drie jaar. Voorts omvat de ranglijst van *Elsevier* aanzienlijk meer fondsen, dus hebben we alleen die beleggingsfondsen opgenomen in het experiment die in beide ranglijsten voorkomen. De resultaten van het experiment moeten derhalve worden geïnterpreteerd als het antwoord op de vraag: welke verschillen in betrouwbaarheid bestaan er tussen een ranglijst op basis van de Life-index en een ranglijst op basis van de Sharpe-index?

In figuur 2 hebben we voor de Sharpe-index de resultaten voor alle beleggingsfondsen gepresenteerd. Deze afbeelding bevat een samenvatting van de waarschijnlijkheidsverdelingen voor alle fondsen, waarbij we van iedere individuele verdeling de punten hebben weergegeven die grenzen weergeven van de intervallen die respectievelijk 95%, 90% en 75% van de waarnemingen bevatten. Op de horizontale as staan de beleggingsfondsen geordend op basis van hun oorspronkelijke rangschikking in de ranglijst. Op de verticale as staan de mogelijke verplaatsingen

op basis van de door ons uitgevoerde simulaties voor verschillende niveaus van betrouwbaarheid.

In figuur 3 hebben we voor de Life-index de resultaten van de simulatie samengevat op een wijze gelijk aan die voor figuur 2. Ten einde de betrouwbaarheid van de ranglijsten te vergelijken, nemen we de oppervlakte van beide grafieken bij een gegeven betrouwbaarheidsniveau. Deze oppervlakte gedeeld door het aantal fondsen in de ranglijst geeft de gemiddelde verplaatsing van een fonds in de ranglijst weer.

Uit tabel 1 blijkt dat de verschillen in betrouwbaarheid tussen de beide ranglijsten nauwelijks van elkaar verschillen. Derhalve lijkt de conclusie gerechtvaardigd dat het onderscheidend vermogen van de Life-ranglijst niet groter is dan dat van de Sharpe-ranglijst.

Andere opmerkingen

Naast de hierboven beschreven problemen omtrent het onderscheidend vermogen van een ranglijst, spelen een aantal andere factoren een rol bij het beoordelen van de bruikbaarheid van de index. Dit wordt veroorzaakt doordat bij de berekening van de indices een aantal subjectieve stappen genomen. Zo wordt aangenomen dat een individuele belegger gewoonlijk slechts in één fonds belegt. Daardoor zijn zowel de Life- als de Sharpe-index ongeschikt voor beleggers die van plan zijn om in meerdere fondsen te beleggen. Deze beleggers zullen vooral kijken naar de mate waarin beleggingsfondsen buitengewoon rendement kunnen behalen. Diversificatie kunnen deze beleggers zelf in hun portefeuille aanbrengen door te beleggen in meerdere beleggingsfondsen.

Andere subjectieve elementen die er toe zouden kunnen leiden dat de belegger tot andere afwegingen komt dan de afweging die gemaakt is bij de opstelling van de ranglijsten. Zo kan de beleggingshorizon een rol spelen bij de risico-perceptie van de belegger. Beleggers met een lange beleggingshorizon zouden in principe bereid kunnen zijn om een groter risico te lopen. Ook de fiscale status van de belegger blijkt in de praktijk een belangrijke rol te spelen bij de keuze voor een beleggingsfonds. Beleggers die dividendbelasting moeten betalen zullen de voorkeur hebben

voor beleggingsfondsen die weinig dividend uitkeren.

Conclusie

Ranglijsten van beleggingsfondsen op basis van de Sharpe-index of de Life-index kenmerken zich door een vrij sterke mate van 'ranglijst-risico'. De omvang van dit risico wordt met name bepaald door de precizie van het meetinstrument en de mate van efficiëntie van de vermogensmarkt. Aangezien de beide ranglijsten voor dezelfde fondsen in dezelfde periode zijn bestudeerd, kunnen verschillen in ranglijst-risico worden toegerekend aan de precizie van het meetinstrument. Wanneer de rangschikking op basis van de Sharpe-index vergeleken wordt met de rangschikking op basis van de Life-index, dan blijkt dat beide ranglijsten een grote (hoge) mate van ranglijst-risico bevatten. De mogelijke fluctuaties die kunnen optreden liggen bij een betrouwbaarheidsniveau van 75% in de orde van grootte van twaalf plaatsen.

Een voordeel van de Life-index is het gebruik van rapportcijfers. Hierdoor kunnen de kwaliteiten van beleggingsfondsen gemeten worden op een schaal die aanspreekt bij een groot deel van de potentiële gebruikers van de index, namelijk de particuliere belegger. Nadeel van zowel de Life- als de Sharpe-index is dat een aantal subjectieve keuzes, die eigenlijk door de gebruiker van de ranglijst genomen zouden moeten worden, door de opstellers van de index genomen zijn. In het geval van de Life-index zou dit er voor pleiten om de mate van diversificatie afzonderlijk te vermelden, en deze niet te gebruiken bij het opstellen van de ranglijst. De belegger kan dan zelf de afweging maken voor het aantal beleggingsfondsen dat hij/zij zou willen gebruiken en de daarbij vereiste diversificatie van de te selecteren fondsen.

Age Bruinsma, Jeroen de Munnik en Auke Plantinga

Age Bruinsma is werkzaam bij Institutional Management Services, Jeroen de Munnik is werkzaam bij Spaarbeleg, en Auke Plantinga is verbonden aan de Economische Faculteit Groningen. De auteurs danken de heer T.K. Dijkstra voor enkele nuttige suggesties inzake de gevolgde methodologie.

Het belangrijkste punt van de reactie van Bruinsma et al. betreft de nauwkeurigheid waarmee een performance-analyse uitgevoerd kan worden. Het zou bij voorbeeld kunnen dat een hoge of lage positie van een beleggingsfonds op een performance-ranglijst geheel aan toeval te wijten is, en niets met de kwaliteiten van het fonds te maken heeft.

Ranglijst-risico

Bij de constructie van de Life-index zijn we ons hiervan bewust geweest, en hebben we hiermee rekening gehouden. Bij de door ons gehanteerde conversie van rapportcijfers scoren bij voorbeeld 15 van de 32 fondsen een zes op het criterium abnormaal rendement (t-waarde tussen min één en één). Daarmee geven wij juist aan dat de prestaties van de meeste fondsen op het onderdeel abnormaal rendement niet van elkaar te onderscheiden zijn. Veel fondsen eindigen dan ook op dit onderdeel ex-aequo op onze ranglijst. Bruinsma et al. verwijten ons hier een fout die wij niet maken.

Om de nauwkeurigheid van ons 'meetinstrument' nog eens opnieuw te controleren, hebben we de bootstrap-methode gebruikt om de statistische kwaliteit van de rapportcijfers te bepalen. Voor elk fonds hebben we onder andere uitgerekend wat de kans is dat hun uiteindelijke rapportcijfer met meer dan een half punt afwijkt van wat wij rapporteren. Gemiddeld over alle fondsen komt dit uit op ongeveer 15%. We concluderen daarom dat de rapportcijfers redelijk betrouwbaar zijn. Het verschil tussen onze conclusie en die van de discussianten komt voort uit het feit dat zij de betrouwbaarheid van de plaats op de ranglijst testen, en wij die van het rapportcijfer. In onze meest recente ranglijst halen bij voorbeeld negen fondsen een zevenenhalve en staan daarmee ex-aequo op de 12e tot en met 20e plaats. Bruinsma et al. concluderen nu een groot ranglijst-risico: een fonds dat op plaats twaalf staat had net zo goed op plaats twintig kunnen staan. Wij beweren echter alleen dat een fonds met een score van zevenenhalve echt anders is dan fondsen die boven de acht of onder de zeven scoren.

Juist omdat we schijn-nauwkeurigheid hebben willen uitsluiten, aggregeren we naar rapportcijfers. We hebben nooit beweerd dat we de prestaties in vijf cijfers achter de komma kunnen opschrijven, zodat we ieders plaats op de ranglijst exact kunnen bepalen. Daarmee wijken we bewust af van de Elsevier-lijst. Op die ranglijst uiten Bruinsma et al. dus ons inziens terecht kritiek. Dat is ook één van de redenen dat wij met een alternatief voor de Elsevier-lijst zijn gekomen.

Diversificatie

Een andere belangrijke motivatie om af te wijken van de Sharpe-index als maatstaf voor de prestaties van beleggingsfondsen is dat de Sharpe-index zich uitsluitend op het buitengewoon rendement richt. In tabel 3 van ons artikel van 28 juni laten we zien dat de mate van diversificatie een duidelijker en meer persistent kenmerk is. Hoeveel gewicht een indivi-

duële belegger aan dit kenmerk wil geven is natuurlijk een subjectieve zaak, zoals we overigens in onze bijdrage ook hebben aangegeven. Dat een individuele belegger ook zelf kan diversificeren is natuurlijk waar. Het voordeel van beleggingsfondsen is echter dat ze schaalvoordelen bieden en van de particuliere belegger geen al te actief beleggingsbeleid vereisen.

Technische bezwaren

Naast deze inhoudelijke punten, kleven er naar onze mening ook technische bezwaren aan de berekeningen van Bruinsma et al. Het eerste bezwaar betreft de beperking van de steekproefperiode tot drie jaar. In het algemeen geldt dat een meetinstrument minder nauwkeurig wordt naarmate er minder waarnemingen zijn. Door uit te gaan van slechts drie jaar wordt het ranglijst-risico overdreven.

Het tweede bezwaar betreft het feit dat geen rekening wordt gehouden met de afhankelijkheid van de

rendementen. In tabel 4 van ons artikel van 28 juni laten we zien dat de rendementen in zekere mate voorspelbaar zijn. Hierdoor kunnen ze niet zonder meer door elkaar gemengd worden, zoals in de bootstrap-methode gebeurt.

De door Bruinsma et al. opgeworpen kritiek is volgens ons dus niet van toepassing op de Life-index, maar hoogstens op de ook door hun onderzochte Elsevier-lijst.

Kees Koedijk
Francois Nissen
Peter Schotman
Arjan van Bussel

De auteurs zijn werkzaam bij het Limburg Institute of Financial Economics (Life), verbonden aan de Rijksuniversiteit Limburg.