

**ONWEERLEGBAAR BEWIJS?
OVER HET BELANG EN DE WAARDE
VAN EMPIRISCH ONDERZOEK
VOOR FINANCIERINGS- EN BELEGGINGSVRAAGSTUKKEN**

Bibliographical Data

Library of Congress Classification (LCC)	5001-6182 : Business 4001-4280.7 : Finance Management, Business Finance, Corporation Finance HG 4001+ : Corporate Finance
Journal of Economic Literature (JEL)	M : Business Administration and Business Economics G 3 : Corporate Finance and Governance G 30 : Corporate Finance; General
European Business Schools Library Group (EBSLG)	85 A : Business General 220 A : Financial Management 220 B : Financial Management
Gemeenschappelijke Onderwerpsontsluiting (GOO)	
Classification GOO	85.00 : Bedrijfskunde, Organisatiekunde: algemeen 85.30 : Financieel management, financiering 85.30 : Financieel management, financiering
Keywords GOO	Bedrijfskunde / Bedrijfseconomie Financieel management, bedrijfsfinanciering, besliskunde bedrijfsfinanciering, redes (vorm)
Free keywords	data snooping, look-ahead bias, mutual funds, predictability, efficient market hypothesis

Erasmus Research Institute of Management (ERIM)
Erasmus University Rotterdam
Internet: <http://www.irim.eur.nl>

ERIM Inaugural Addresses Research in Management Series

Reference number ERIM: EIA-2002-09-F&A

ISBN 90 – 5892 – 026 – 7

© 2002, Marno Verbeek

All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission in writing from the author(s).

**ONWEERLEGBAAR BEWIJS?
OVER HET BELANG EN DE WAARDE
VAN EMPIRISCH ONDERZOEK
VOOR FINANCIERINGS- EN BELEGGINGSVRAAGSTUKKEN**

Oratie

Rede, in verkorte vorm uitgesproken bij de aanvaarding van het ambt van
hoogleraar Ondernemingsfinanciering aan de Faculteit der Bedrijfskunde en
de Faculteit der Economische Wetenschappen van de Erasmus Universiteit
Rotterdam op vrijdag 21 juni 2002

door

Prof. Dr. Marno Verbeek

Erasmus Universiteit Rotterdam
Rotterdam School of Management / Faculteit Bedrijfskunde
Rotterdam School of Economics / Faculteit der Economische Wetenschappen
P.O. Box 1738
3000 DR Rotterdam
M.Verbeek@fbk.eur.nl
<http://www.surf.to/marno.verbeek>

*Mijnheer de Rector Magnificus,
Geacht College van Decanen,
Waarde Collegae,
Dames en Heren,*

Inleiding

Het vakgebied financiering en belegging heeft zich de afgelopen twee decennia in snel tempo ontwikkeld. De manier en het niveau waarop momenteel in Nederland wetenschappelijk onderzoek wordt bedreven benadert steeds meer die van de Verenigde Staten, door velen beschouwd als toonaangevend. Met name in de beleggingstheorie en het risicomangement is het econometrische gehalte van het onderzoek zeer hoog en worden geavanceerde technieken niet geschuwd. Een snelle rondgang langs collegae die recentelijk een benoeming als hoogleraar op dit gebied hebben aanvaard, leert dat de meerderheid hiervan een gedegen achtergrond in de econometrie heeft. Ook ik pas in deze categorie.

Vandaag wil ik met u van gedachten wisselen over het belang en de waarde van empirisch onderzoek voor financierings- en beleggingsvraagstukken. Dit wordt geen lofrede op de econometrie, maar veeleer een kritische blik. Econometrische technieken kunnen vele antwoorden opleveren, maar het vinden van het meest geschikte antwoord bij een gegeven vraag is zelden triviaal. Zelfs resultaten gepubliceerd in internationale toptijdschriften worden nog vaak aangevallen op basis van de gebruikte methoden en de gehanteerde veronderstellingen.

Er zijn twee thema's die ik vandaag met u wil bespreken. Natuurlijk geven deze thema's geen volledig of representatief beeld van het vakgebied, maar dat is ook niet mijn bedoeling. Ze illustreren wel de belangrijke rol die empirisch onderzoek op bepaalde terreinen vervult, en vooral ook de mogelijke zwakke plekken en kwetsbare elementen daarbij. Als eerste wil ik aandacht geven aan de prestaties van beleggingsfondsen en aan de vraag in hoeverre historische rendementen van een fonds *iets* zeggen over de toekomst. Als tweede bespreek ik het gerelateerde thema van voorspelbaarheid van aandelenrendementen in het algemeen en de kritische rol van empirisch onderzoek daarbij. U zult merken dat er in deze twee thema's een onderliggend gemeenschappelijk probleem bestaat: de mogelijke misleiding van empirische resultaten doordat bepaalde selectieprocessen (bijvoorbeeld van de steekproef of van voorspellende factoren) niet worden waargenomen of door de onderzoeker worden verwaarloosd.

De prestaties van beleggingsfondsen

Een eerste onderzoeksthema dat ik met u aan de orde wil stellen, betreft het analyseren van de prestaties, of performance, van beleggingsfondsen, en de persistentie daarin. Vrijwel zonder uitzondering, en gelukkig in een steeds groter lettertype, voegen beleggingsfondsen de mededeling “In het verleden behaalde resultaten bieden geen garantie voor de toekomst” toe aan hun advertenties en andere publicaties. Hoewel iedereen er inmiddels wel van is doordrongen dat de waarde van een beleggingsfonds kan fluctueren, bestaat er op zijn minst de suggestie dat historische resultaten wel een *indicatie* zijn voor de toekomst.

Bij een empirische analyse van performance en persistentie spelen twee belangrijke problemen een rol. Ten eerste is er de vraag hoe je nu precies de prestaties van een beleggingsfonds moet meten. Of, iets anders geformuleerd, wat is een eerlijke manier om verschillende fondsen te vergelijken? Het tweede probleem is het feit dat prestaties noodzakelijkerwijs alleen achteraf kunnen worden bepaald, terwijl het cruciaal is bij beleggen dat vooraf beslissingen moeten worden genomen. Elke avond verlaat er wel iemand het casino met een grote winst van de roulettetafel; maar had u deze persoon ook van tevoren al uw geld toevertrouwd? Bij beleggingsfondsen komt hierbij het probleem dat relatief slechte fondsen vaak al in een vroeg stadium worden opgeheven of (zonder veel publiciteit) samengevoegd met andere fondsen. Indien deze selectieve uitval wordt verwaarloosd, kan onze inschatting van de prestaties van de fondsen sterk vertekend worden. Vergelijk dit met een situatie waarin het casino de verliezers langs de achterdeur laat vertrekken. Bij de voordeur zien we elke avond een verdacht groot aantal lachende gezichten!

Wanneer we de prestaties van een beleggingsfonds willen beoordelen kunnen we verschillende invalshoeken nemen: die van een individu die zich afvraagt of hij een deel van zijn portefeuille, of zelfs zijn totale (risicovol te beleggen) vermogen, bij dit beleggingsfonds zou moeten onderbrengen, of die van een beoordelaar die zich afvraagt in welke mate het beleggingsfonds in staat is (geweest) een beter resultaat te boeken dan vergelijkbare beleggingen. We zouden dit laatste eenvoudig kunnen vertalen met de vraag in hoeverre het fonds in staat is ‘de markt’ te verslaan. Heeft de fondsmanager superieure kennis of informatie die niet algemeen beschikbaar is, en, zo ja, weet hij deze om te zetten in een superieur beleggingsresultaat?

Een eerste mogelijkheid om een dergelijke benadering te implementeren is het kiezen van een ‘vergelijkbare’ belegging als referentiepunt, bijvoorbeeld een brede aandelenportefeuille, zoals de Standard and Poor 500 index. In plaats van het verschil in rendement te nemen tussen het beleggingsfonds en de referentie-index, kunnen we dit iets verfijnder doen. Gebruikmakend van historische rendementen van beide beleggingen, kunnen we met behulp van lineaire regressie bepalen welke portefeuille, bestaande uit de S&P 500 index en een risicovrije investering,

het meest overeenkomt met het te evalueren beleggingsfonds (d.w.z. de grootste correlatie heeft). Het gemiddelde verschil tussen het fondsrendement en dit portefeuillerendement, corresponderend met de constante term in de genoemde regressie, is dan een maat voor de performance van het beleggingsfonds. In de wetenschappelijke literatuur spreekt men meestal over ‘alpha’.

Het gebruiken van de S&P 500 index is te verdedigen als het beleggingsfonds voornamelijk belegt in aandelen die in deze index zijn opgenomen. Indien het fonds eveneens in Europese aandelen of Amerikaanse obligaties zou beleggen, is de vergelijking met de S&P 500 benchmark natuurlijk minder opportuun. In dat geval zou je een andere referentie-index kunnen nemen. Echter, in veel gevallen is het niet evident welke dat zou moeten zijn. Veel fondsen kennen een combinatie van beleggingsstijlen of rapporteren een andere stijl dan zij in de praktijk brengen.¹ De meest eenvoudige manier om hiervoor een vergelijkingsportefeuille te bepalen is wederom de data te laten spreken: met behulp van een regressie waarbij het aantal verklarende factoren is uitgebreid wordt impliciet zo’n portefeuille bepaald. Met een dergelijke benadering kunnen we bijvoorbeeld voor een beleggingsfonds een benchmark construeren die bestaat uit een combinatie van de S&P 500 index, een Europese aandelenindex en een obligatie-index. Hiermee is het eveneens mogelijk om het rendement op te nemen van strategieën, waarvan wordt aangenomen dat zij een beter rendement opleveren dan een marktindex, bijvoorbeeld een strategie die een groter gewicht geeft aan kleine ondernemingen of aan ondernemingen die een lage marktwaaarde kennen in vergelijking met hun boekwaarde.² Het is duidelijk dat hoe meer additionele factoren we in een dergelijke benchmarkportefeuille toelaten, hoe moeilijker het wordt voor een beleggingsfonds om een dergelijke (‘passieve’) strategie te verslaan.

Het schatten van alpha, op basis van dergelijke regressies, als maat voor de performance van een beleggingsfonds is weliswaar eenvoudig met behulp van de kleinste kwadratenmethode (OLS) en gebruikmakend van tijdreeksgegevens, maar de onnauwkeurigheid van de schattingen is veelal groot. Dit komt doordat de meeste beleggingsfondsen een vrij korte historie hebben, terwijl de volatiliteit van

¹ Zie bijvoorbeeld Brown & Goetzmann (1997), DiBartolomeo & Witowski (1997), Kim, Shukla & Tomas (2000), en DeRoos, Nijman & ter Horst (2002).

² Het model met bovengenoemde drie factoren is tegenwoordig min of meer algemeen geaccepteerd en wordt het Fama-French (1993) 3-factor model genoemd. Carhart (1997) voegt hieraan een vierde factor toe, gebaseerd op een momentumstrategie. In een momentum strategie wordt relatief meer belegt in aandelen met een hoog rendement over de afgelopen (zes) maanden dan in die met een laag rendement. Men belegt dus meer in ‘winner stocks’ dan in ‘loser stocks’. Zie bijvoorbeeld Jegaseesh en Titman (1993) voor de Verenigde Staten, en Rouwenhorst (1998) en Swinkels, Nijman & Verbeek (2002) voor Europa.

de rendementen erg groot is.³ Dit probleem wordt nog verergerd als we de prestaties over subperioden met elkaar willen vergelijken. Bijvoorbeeld: is het zo dat van een beleggingsfonds dat in 2000 bij de beste 10% hoorde, verwacht kan worden dat het in 2001 ook relatief goed presteert? In dat geval hebben we dus twee perioden van slechts 12 maanden waarover we de prestaties willen bepalen. Als gevolg daarvan is de onnauwkeurigheid van de geschatte alphas vrij groot en kan een substantieel deel van de gemeten verschillen toegeschreven worden aan het toeval. Het is een interessante en belangrijke vraag hoe dit toevalseffect gereduceerd kan worden en de nauwkeurigheid van de geschatte alphas vergroot. Het antwoord is niet evident; een grotere steekproef biedt hier geen soelaas.

Twee recente onderzoeksprojecten waaraan ik samen met diverse collega's werk, sluiten aan bij deze problematiek. In samenwerking met onder andere Georges Bresson (Université de Paris II), wordt de mogelijkheid onderzocht informatie van andere beleggingsfondsen en uit eerdere perioden te gebruiken om onze inschatting van de performance van een gegeven beleggingsfonds te verbeteren. Dit kan met behulp van zogenaamde shrinkage schatters, die binnen de panel data literatuur voorgesteld zijn, of Bayesiaanse methoden.⁴ Bij deze laatste benadering is het mogelijk expliciet prior informatie te gebruiken, bijvoorbeeld over de kosten die een fonds in rekening brengt. De eerste resultaten van dit onderzoek suggereren dat dergelijke benaderingen beter in staat zijn beleggingsfondsen in de juiste rangorde te plaatsen dan de traditionele methoden. Een gerelateerd probleem wordt veroorzaakt door het feit dat beleggingsfondsen vaak actief hun beleggingsstijl aanpassen aan veranderende marktomstandigheden of veranderende inschattingen van de toekomst. Dat maakt een vaste benchmarkportefeuille vaak minder geschikt. Een oplossing is het expliciet modelleren van de tijdsafhankelijkheid in de benchmarkportefeuille. Dit is niet eenvoudig en vereist adequate econometrische technieken, waaronder schatten via het Kalman filter. Dit is een van de thema's waaraan ik samen met Laurens Swinkels (KUB en ABP) en Pieter-Jelle van de Sluis (ABP) momenteel werk.⁵ Op basis van dergelijk onderzoek kan onder meer nagegaan worden of beleggingsfondsen die van stijl veranderen dit op strategisch juiste momenten doen.

³ In econometrische termen: de steekproefperiode is kort en de variantie van de storingsterm is groot. Beide aspecten zorgen voor een grote variantie van de OLS schatter.

⁴ Recent hebben diverse auteurs het gebruik van Bayesiaanse alphas voorgesteld en verdedigd, zie bijvoorbeeld Baks, Metrick & Wachter (2001), Pástor & Stambaugh (2002a,b) en Busse & Irvine (2002).

⁵ Zie Swinkels, Van Der Sluis & Verbeek (2002).

Persistentie in performance: geven historische resultaten een indicatie voor de toekomst?

Graag maak ik nu de overstap naar het analyseren van de persistentie in de prestaties van beleggingsfondsen. Achteraf weten dat een beleggingsfonds goed gepresteerd heeft, is vooral interessant als dat ons ook iets zegt over de toekomst. Dat wil zeggen: geven historische resultaten een *indicatie* voor de toekomst?

Er zijn diverse manieren waarop de persistentie in performance geanalyseerd kan worden. Een gebruikelijke manier is de volgende. We delen de steekproefperiode op in twee subperioden: een 'ranking' periode en een evaluatieperiode. Voor beide perioden worden de prestaties geschat van de beschikbare beleggingsfondsen. Op basis van de resultaten van de eerste periode worden een indeling gemaakt in een aantal klassen. Uitgaande van een opdeling in 10 groepen, bevat deciel 1 de 10% slechtste fondsen en deciel 10 de 10% beste fondsen. Vervolgens bepalen we de gemiddelde performance over de tweede periode voor alle fondsen in een gegeven deciel. Indien er een positief verband bestaat, verwachten we dat de fondsen uit de betere decielen gemiddeld genomen ook betere prestaties vertonen in de tweede periode.

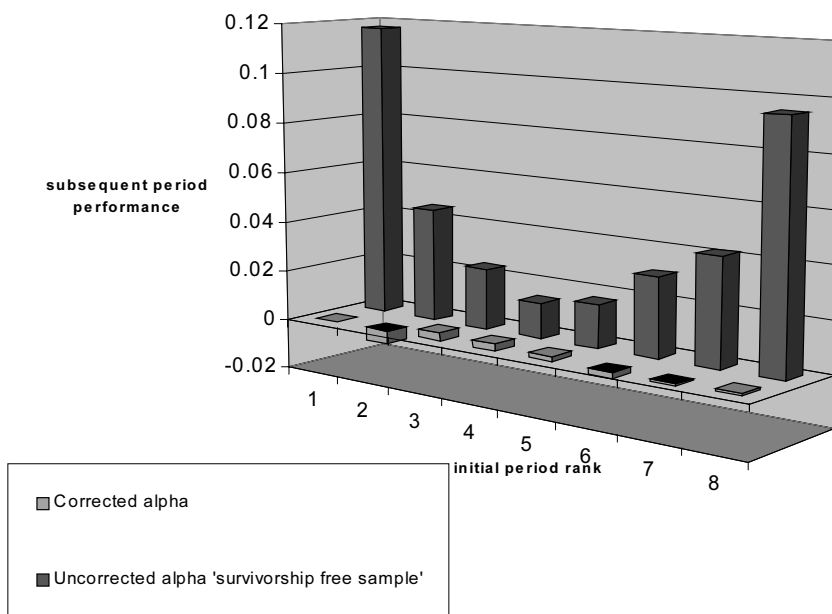
Zoals in de inleiding al is gesteld, is het bestaan van uitval een empirisch probleem bij het analyseren van performance en bij de persistentie van performance in het bijzonder. Voor Amerikaanse fondsen is er bijvoorbeeld tussen 1989 en 1994 sprake van een uitval tussen de 2% en 9% per jaar. Meestal zijn het de fondsen die relatief slecht gepresteerd hebben die worden opgeheven of worden opgenomen in een ander fonds van dezelfde maatschappij. Dit kan leiden tot serieuze vertekeningen.

Laten we voor het gemak eens uitgaan van een situatie waarin geen enkel beleggingsfonds in staat is een buitengewoon rendement te behalen, d.w.z. een situatie waarin alpha voor elk fonds gelijk is aan nul. De fondsen verschillen wel in de hoeveelheid systematisch risico die ze nemen, risico gerelateerd aan de markt als geheel⁶, en de hoeveelheid fondsspecifiek risico (de variantie van de storingsterm). Fondsen met een hoog specifiek risico hebben een relatief grote kans om een schatting voor alpha te hebben die ver van nul ligt, zowel erboven als eronder. Met andere woorden, fondsen met een grote variantie in de eerste periode zullen met een grotere kans terechtkomen in een van de extreme decielen (met name 1 en 10). Fondsen met een lage variantie zullen relatief vaker in de middelste decielen te vinden zijn. Personen die aan de roulettetafel veel risico nemen, zullen aan het eind van de avond vaker grote bedragen gewonnen of verloren hebben, dan personen die steeds voorzichtig op zwart en rood spelen.

⁶ Om precies te zijn: risico waarvoor op de markt een risicopremie wordt betaald.

Vooralsnog is dit geen probleem, zolang we maar beseffen dat de relatieve positie van een beleggingsfonds in de eerste periode iets te maken heeft met het specifieke risico dat het fonds neemt. Laten we nu de tweede periode in ogenschouw nemen. Binnen elk deciel nemen we dan het gemiddelde ‘abnormale’ rendement, hetgeen correspondeert met het steekproefgemiddelde van de geschatte alphas. Dit gemiddelde is een maat voor de verwachte prestaties over de komende periode gegeven dat een fonds in de eerste periode bijvoorbeeld bij de beste 10% of slechtste 10% hoort.

Aangezien we verondersteld hebben dat de werkelijke alphas allemaal nul zijn, verwachten we geen systematisch patroon in de prestaties zoals gemeten over de evaluatieperiode. Zoiets dus als de voorste rij staafjes in Figuur 1.⁷



Figuur 1: Initiële ranking en daaropvolgende performance, met en zonder correcties (gesimuleerde gegevens).

Echter, stel nu dat fondsen die relatief slecht scoren met een grotere kans verdwijnen dan de relatief goede fondsen. Deze veronderstelling wordt door empirisch onderzoek ondersteund. Doordat de slechtst presterende fondsen

⁷ De fondsen in de ranking periode zijn hier opgedeeld in acht groepen (octielen). Voor meer details, zie ter Horst, Nijman en Verbeek (2001).

gemiddeld genomen een hoog risico kennen betekent dit dat fondsen die in de evaluatieperiode *niet* zijn verdwenen, gemiddeld genomen betere prestaties laten zien. Dat wil zeggen: een gemiddelde alpha hebben groter dan nul. Dit effect is het grootste voor fondsen die een groot risico kennen. Immers, hoe groter de variantie, hoe meer de staart van de verdeling wordt afgekapd door de uitval. Achteraf, als we de performance in de evaluatieperiode beschouwen - noodzakelijkerwijs alleen voor de fondsen die nog bestaan - betekent dit dat fondsen in de extreme decielen gemiddeld genomen betere prestaties kennen. We vinden een schijnverband dat uitsluitend te wijten is aan de selectieve uitval en niets te maken heeft met enig positieve verband tussen de prestaties van de fondsen in verschillende deelperioden. Immers, ons uitgangspunt was dat geen enkel fonds 'abnormale' rendementen wist te behalen, noch in de ranking periode, noch in de evaluatieperiode.⁸

Laat ik nog een keer de vergelijking met de roulettetafel maken. Indien spelers de tafel verlaten zodra ze een groot verlies hebben geleden, zullen spelers die veel risico nemen eerder verdwijnen dan voorzichtige spelers. Aan het eind van de avond verlaten de volhouders het casino door de voordeur. We zien, zoals eerder opgemerkt, opvallend veel winnaars. Daarnaast zien we dat de spelers die veel risico nemen gemiddeld met meest gewonnen hebben. Niet omdat ze de kansen van de roulettetafel in hun voordeel hebben kunnen ombuigen, maar omdat de personen die door de voordeur naar buiten komen gemiddeld genomen meer geluk gehad hebben. En geluk heeft bij voorzichtige spelers nu eenmaal minder effect op de winst dan bij spelers die veel risico nemen.

De precieze vorm van het schijnverband dat we vinden hangt af van het uitvalproces en hoe dit samenhangt met de fondsrendementen.⁹ Een mogelijk verband is weergegeven in de tweede rij staafjes in Figuur 1.¹⁰ We zien hierin bijvoorbeeld dat de 12,5% slechtste fondsen in de 3-jarige ranking periode, gemiddeld genomen een abnormaal rendement behalen in de evaluatieperiode van bijna 0,5% per jaar. Dit is in werkelijkheid niet zo, omdat een deel van deze groep beleggingsfondsen is verdwenen en die vinden we niet terug in het gemiddelde.

⁸ Natuurlijk, de meeste fondsen zullen positieve of negatieve schattingen hebben voor alpha. Onder onze veronderstellingen is dit uitsluitend een kwestie van geluk (c.q. pech).

⁹ Hendricks, Patel en Zeckhauser (1997) presenteren een J-vorm, terwijl ter Horst, Nijman en Verbeek (2001), op basis van een empirisch geschat uitvalproces, min of meer een U-vormig verband laten zien.

¹⁰ Deze figuur is gebaseerd op gesimuleerde gegevens, waarbij de data zijn gegenereerd op een wijze die dicht ligt bij de waargenomen verdeling van kenmerken van beleggingsfondsen.

Hoewel over de exacte terminologie wel wat verwarring bestaat, wordt een dergelijke vertekening veelal aangeduid met 'look ahead bias'. Bij het beantwoorden van de vraag of de prestaties in de eerste periode iets zeggen over die in de tweede periode hebben we immers stiekem al vooruit gekeken door alleen die fondsen te gebruiken die later blijken overleefd te hebben. Dat is informatie die we aan het eind van de eerste periode nog niet konden bezitten. Deze interpretatie maakt ook duidelijk dat het niet triviaal is deze vertekening te voorkomen doordat rendementen over fondsen die niet meer bestaan per definitie nooit beschikbaar zijn. De enige manier om rekening te houden met look-ahead bias en ervoor te corrigeren is een aanpassing van de schattingsmethode.¹¹

In de micro-econometrische literatuur zijn twee verschillende benaderingen te vinden voor het probleem van 'selection bias', waarvan look-ahead bias een speciaal geval is. Stel dat in abstracte termen onze interesse ligt in de verdeling van een of meer variabelen y gegeven een groep variabelen x . Bijvoorbeeld de verdeling van de rendementen van een beleggingsfonds in de evaluatieperiode gegeven de ranking uit de eerste periode en eventueel gegeven andere kenmerken van het fonds (bijv. beleggingsstijl). In beide benaderingen is het nodig het uitvalproces te modelleren en te identificeren als functie van waargenomen en niet-waargenomen variabelen (de storingsterm). De meest populaire benadering is gebaseerd op het werk van Nobelprijswinnaar James Heckman (o.a. Heckman, 1979). Hierin wordt verondersteld dat de waargenomen variabelen in het uitvalproces ofwel onafhankelijk zijn van y (voor gegeven x) ofwel deel uitmaken van x . In bovenstaand probleem is een dergelijke benadering niet geschikt omdat uitval voor een deel bepaald zal worden door de fondsrendementen die we niet in x willen opnemen. Een alternatief¹² veronderstelt dat de niet-waargenomen variabele in het selectieproces onafhankelijk is van y , maar laat alle mogelijke waargenomen variabelen toe. Deze benadering wordt gebruikt door ter Horst, Nijman en Verbeek (2001), waarbij uitval in een gegeven kwartaal wordt gemodelleerd als functie van zoveel mogelijk rendementen uit voorafgaande kwartalen.

Indien we een correctie willen doorvoeren voor de besproken vertekeningen hebben we dus informatie nodig over de manier waarop de fondsen uit de steekproef verdwijnen en hoe dit samenhangt met de al of niet waargenomen rendementen. We moeten met andere woorden ook aan de achterdeur van het casino gaan staan. Om te corrigeren voor look-ahead bias kunnen op basis van het geschatte uitvalproces (bijv. een probit model) eenvoudige gewichten bepaald worden. Vervolgens wordt in de analyse de gebruikelijke schattingsmethode

¹¹ Een alternatief is de zogenoemde 'follow-the-money' benadering, waarin het geld dat belegd was in de verdwenen fondsen gevolgd wordt tot daadwerkelijke liquidatie of opname in een ander fonds; zie Gruber (1996) en Wermers (2000).

¹² Zie bijvoorbeeld Fitzgerald, Gottschalk en Moffitt (1998).

toegepast, maar waarbij de waarnemingen van elk fonds een verschillend gewicht krijgen.¹³ Een dergelijke gewogen analyse zorgt er voor dat het schijnpatroon, zoals te zien in Figuur 1, verdwijnt. Na correctie verkrijgen we inderdaad het vrijwel vlakke patroon van de eerste rij staafjes in deze figuur. Het schijnverband is verdwenen.

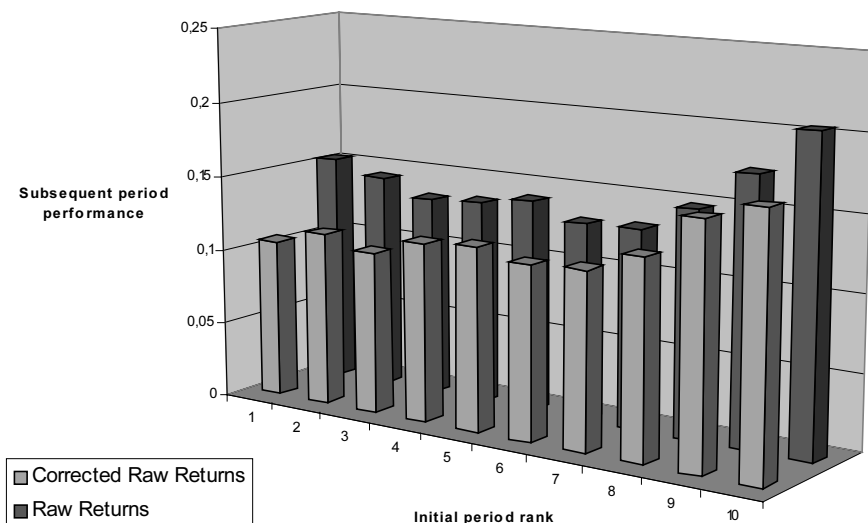
In samenwerking met Guillermo Bacquero en Jenke ter Horst (KUB) is bovenstaande methode, met correctie voor look-ahead bias, toegepast op Amerikaanse hedge fondsen. In zekere zin zijn hedge fondsen vergelijkbaar met beleggingsfondsen zonder al te veel beperkende regels. Dit betekent dat hedge fondsen vaak extreme posities innemen en dat alleen beleggers met voldoende vermogen en een reeds goed gespreide portefeuille een deel van hun geld hierin mogen investeren. Een moeilijk probleem is dan ook het bepalen van een geschikte benchmark voor hedge fondsen.

Tussen 1995 en 2000 ligt de uitval in databestanden van hedge fondsen tussen de 6% en 15% per jaar. Dit is substantieel hoger dan voor beleggingsfondsen. Bovendien wordt deze uitval voor een deel veroorzaakt door een vrijwillige selectie door de fondsen zelf. Fondsen die geen nieuwe investeerders meer willen aantrekken, kunnen besluiten hun gegevens niet meer openbaar te maken. Er is immers geen verplichting.

In Figuur 2 beschouwen we, als illustratie, de persistentie in de performance van 1797 Amerikaanse hedge fondsen, geschat over de periode 1994-2000. Het bepalen van geschikte referentieportefeuilles voor hedge fondsen staat nog volop ter discussie¹⁴; de resultaten hier zijn gebaseerd op de ruwe rendementen. Van links naar rechts zien we de decielen met de laagste tot en met de hoogste gemiddelde rendementen over het eerste jaar, waarbij de hoogte van de staafjes het gemiddelde rendement in het daaropvolgende jaar weergeeft. Zonder correctie voor look-ahead bias – de achterste rij staafjes – zien we wederom een J-vormig patroon: zowel de slechtere als de betere fondsen lijken in het daaropvolgende jaar goed te presteren. Het gemiddelde rendement wordt geschat op 15% voor de groep slechtste fondsen en op meer dan 20% voor de groep beste fondsen. Na correctie verdwijnt dit patroon grotendeels. De gemiddelde rendementen worden dan geschat op 10% respectievelijk 16%. Er lijkt nu nog wel sprake van positieve persistentie voor de betere fondsen, maar de slechtere fondsen behouden het daarop volgende jaar een relatief laag rendement.

¹³ In tegenstelling tot de meer gebruikelijke situatie waarin gewichten exogeen bepaald worden, zijn de gewichten in dit geval endogeen en hebben ze dus wel degelijk effect op de consistentie van de gebruikte schatters.

¹⁴ Zie bijvoorbeeld Brown en Goetzmann (2001) of Fung en Hsieh (2002).



Figuur 2: Persistentie in rendementen op hedge fondsen (jaarlijks).

Dit brengt mij aan het einde van het eerste deel van deze oratie. Een belangrijke les tot nu toe is dat empirische resultaten niet altijd aantonen wat ze lijken aan te tonen. Een zorgvuldige analyse is essentieel, niet alleen van de gebruikte methodes maar ook van de wijze waarop de gebruikte gegevensbestanden tot stand zijn gekomen. Deze les geldt veel algemener dan bij de analyse van beleggingsfondsen, zie bijvoorbeeld Brown, Goetzmann en Ross (1995).¹⁵ Hoewel onderzoekers zich vaak wel bewust zijn van de mogelijkheid van selectieproblemen, zijn er maar weinig die er op een adequate manier voor corrigeren.

¹⁵ Voor toepassingen op andere gebieden, zie bijvoorbeeld Nijman en Verbeek (1992), Verbeek en Nijman (1992) en Vella en Verbeek (1998, 1999). Inderdaad, deze vier verwijzingen vormen een bijzonder selectieve steekproef.

Voorspelbare voorspelbaarheid

Het feit dat maar weinig beleggingsfondsen in staat zijn de ‘markt’ te verslaan heeft alles te maken met de minimale voorspelbaarheid in aandelenrendementen. Tot in de jaren tachtig van de vorige eeuw was in de wetenschappelijk wereld het geloof in een efficiënte markt buitengewoon sterk. In zo’n efficiënte markt leidt elke bron van voorspelbaarheid tot extra handel zodanig dat de voorspelbaarheid zichzelf vrijwel meteen opheft.¹⁶ Populair gezegd: als de winsten voor het oprapen liggen dan had iemand anders die toch allang opgeraapt? Volgens de efficiënte markthypothese is het dus nagenoeg onmogelijk buitengewone winsten te behalen door gebruik te maken van informatie uit historische koersontwikkelingen of jaarverslagen. Talloze empirische onderzoeken bevestigen deze theorie. Natuurlijk staan hier grote groepen praktijkmensen tegenover die het tegendeel beweren en volhouden dat zij een methode of strategie hebben die wel werkt.

De laatste twee decennia verschijnen er ook in de wetenschappelijke literatuur meer artikelen die overtuigend lijken aan te tonen dat er wel sprake is van voorspelbaarheid.¹⁷ Weliswaar niet in bijzonder grote mate, maar voldoende om er, zelfs na transactiekosten, een hoger rendement mee te behalen (ook na correctie voor risico). Sinds enkele jaren verschijnt er echter weer onderzoek dat betoogt dat deze gerapporteerde voorspelbaarheden sterk overschat zijn of recentelijk zijn verdwenen.¹⁸ Een cruciaal probleem bij het onderzoeken van voorspelbaarheid is eveneens een vorm van selectie: van de vele duizenden patronen en verklarende factoren die men heeft onderzocht worden alleen de meest succesvolle gerapporteerd. En als je maar genoeg factoren en strategieën probeert, zijn er altijd wel die, puur toevallig, succesvol zijn geweest gedurende de periode van onderzoek. Dit probleem wordt data mining of data snooping genoemd.

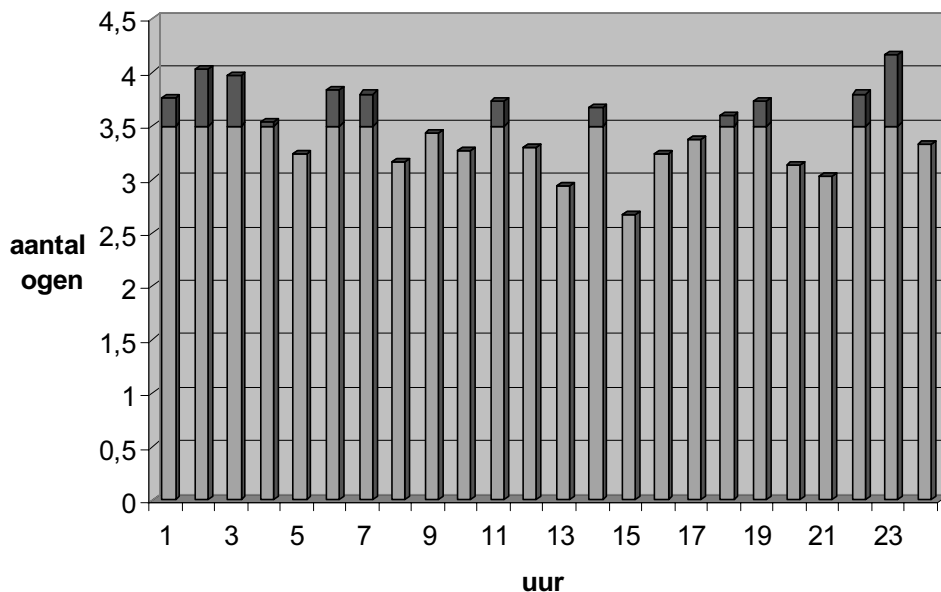
Om het probleem van data snooping te illustreren, wil ik met u het volgende experiment bespreken. Een maand lang werp ik elk uur van de dag met een dobbelsteen. In totaal is dat dus 30 dagen maal 24 uur is 720. Ik vraag mij af wat voor mij het meest succesvolle uur van de dag is om de dobbelsteen te gooien. Met andere woorden: op welk uur van de dag kan ik het hoogste aantal ogen

¹⁶ Om precies te zijn is voorspelbaarheid niet noodzakelijk in strijd met marktefficiëntie. Het is mogelijk dat er een rationele tijdsvariërende risicopremie bestaat. Kirby (1998) analyseert de vraag in welke mate rationele prijsvormingsmodellen consistent zijn met de voorspelbaarheid die empirisch wordt gevonden voor aandelenrendementen.

¹⁷ Zie bijvoorbeeld Campbell (1987), Breen, Glosten en Jagannathan (1989), en Pesaran en Timmermann (1995, 2000).

¹⁸ Zie bijvoorbeeld Bossaerts en Hillion (1999) en Ang en Bekaert (2001).

verwachten? Uitgaande van een zuivere dobbelsteen (we blijven wel eerlijk), is het duidelijk dat het uur van de dag waarop hij gegooid wordt volledig irrelevant is. De kans op zes ogen blijft 1 op 6 en het verwachte aantal ogen blijft 3,5.



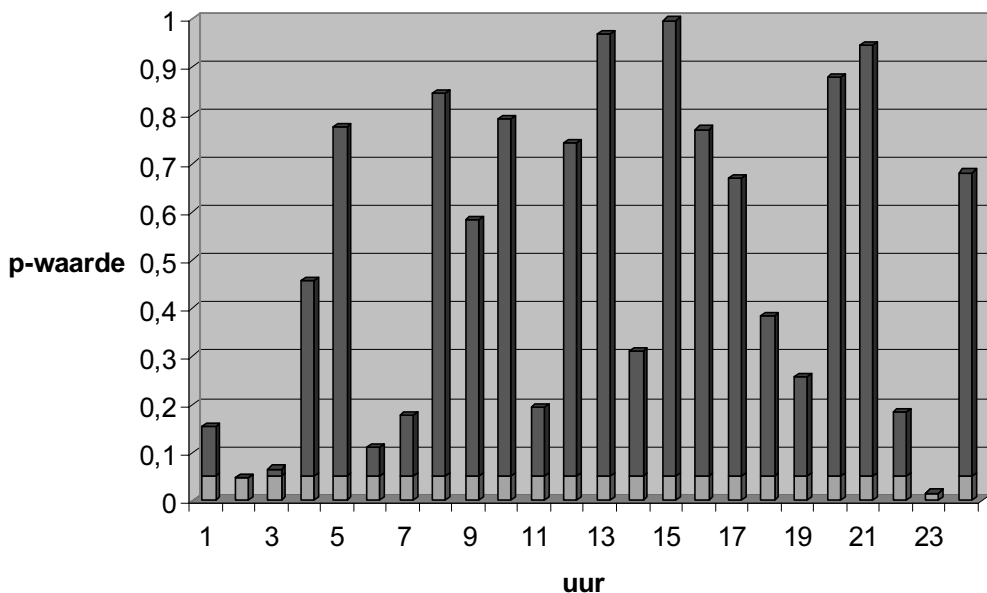
Figuur 3: Gemiddeld aantal geworpen ogen per uur van de dag (over 30 dagen)

In Figuur 3 wordt het gemiddelde aantal ogen gepresenteerd van de 30 worpen die ik in een bepaald uur heb gedaan.¹⁹ U ziet dat er behoorlijk wat variatie is in het gemiddelde. Het 15^e uur van de dag kent bijvoorbeeld een gemiddelde van slechts 2,67, terwijl het 23^e een gemiddelde kent van maar liefst 4,17. Natuurlijk weten we in dit geval dat het allemaal toeval is.

Stel nu eens dat ik voor elk uur zou toetsen of het gevonden gemiddelde gelijk is aan 3,5, de verwachte waarde uitgaande van een zuivere dobbelsteen. Bij zo'n toets kijken we in feite hoe onwaarschijnlijk het is het gerapporteerde gemiddelde of nog hoger te vinden, onder de veronderstelling dat de dobbelsteen in werkelijkheid zuiver is. Deze kansen, de zogenoemde p -waarden, staan weergegeven in Figuur 4. Het is gebruikelijk bij statistische toetsen om de nulhypothese, in dit geval zuiverheid van de dobbelsteen, als onwaarschijnlijk te

¹⁹ De worpen zijn gesimuleerd op de PC.

karacteriseren, en daarmee te verwerpen, als de p -waarde kleiner is dan 5%. We zien dat we op deze manier moeten verwerpen dat het te verwachte gemiddelde aantal ogen in uur 23 gelijk is aan 3,5. De gegevens lijken met andere woorden te duiden op een significant hoger gemiddelde in uur 23. Mijn conclusie is dan ook dat ik in dat magisch uur – op basis van jarenlange ervaring of bijzondere krachten – de wetten van de kansrekening in mijn voordeel weet om te buigen: ik zal graag met u dobbelen tussen 22 en 23 uur en ik verwacht met een grote kans te winnen.



Figuur 4: p -waarden voor de hypothese dat het verwacht aantal ogen gelijk is aan 3,5; per uur (eenzijdig)

In al zijn eenvoud legt dit voorbeeld de vinger op de zere plek: hoe kan het dat we de nulhypothese van een zuivere dobbelsteen verwerpen, terwijl we in dit geval weten dat de dobbelsteen echt zuiver is? Het antwoord is: dit is toeval. Of pech, zo u wilt.

Wanneer we één enkele statistische toets uitvoeren, is er een kleine kans dat we ten onrechte de nulhypothese verwerpen. De kans dat een van de toetsen ten onrechte verwerpt is 5%, en wellicht acceptabel. Maar in bovenstaand geval hebben we niet één toets uitgevoerd, maar 24. De kans dat één van de 24 toetsen ten onrechte verwerpt is veel groter dan 5%; in dit geval gelijk aan ongeveer

71%.²⁰ Als we nog meer toetsen zouden uitvoeren, bijvoorbeeld door de dagen op te splitsen in kleinere tijdseenheden, tendeert deze kans naar 100%. Anders gezegd: als we maar genoeg toetsen uitvoeren is er uiteindelijk altijd wel een die verwerpt! En daar schrijven we dan een mooi artikel over.

Het bovenstaande voorbeeld is vrij duidelijk. Als we de vertaling maken naar de vraag welke factoren aandelenrendementen kunnen voorspellen, wordt het allemaal wat ondoorzichtiger. Ten eerste is de nulhypothese – er is geen voorspelbaarheid – duidelijk aan twijfel onderhevig. Ten tweede is het aantal factoren dat onderzocht is en kan worden vele malen groter dan 24. Ten derde vertelt de gemiddelde onderzoeker niet welk zoekproces er aan de gerapporteerde resultaten vooraf is gegaan. Veelal worden alleen de statistisch interessante verbanden gerapporteerd. Ten vierde zullen de meeste toetsen die worden uitgevoerd niet onafhankelijk zijn van elkaar, waardoor het veel lastiger wordt de kans op een incorrecte verwerping te bepalen. Tenslotte maken vele onderzoekers gebruik van (gedeeltelijk) dezelfde gegevens, zodat resultaten van andere onderzoekers de eigen onderzoeksstrategie beïnvloeden.²¹

Het bovenstaande proces en de daarmee gepaard gaande problemen wordt veelal aangeduid als ‘data snooping’. Zoals de heks bij Hans en Grietje het niet kan waarderen dat er aan haar huisje geknabbeld wordt, zo moeten we als empirisch onderzoekers waakzaam zijn dat wijzelf, en anderen, niet te veel toegeven aan de verleiding te veel van de data te snoepen.

Het probleem van data mining of data snooping is weliswaar een bekend probleem in de econometrische literatuur²², maar een voor de hand liggende oplossing is er niet. Erger nog, de manier waarop hoofdredacteurs van wetenschappelijke tijdschriften te werk gaan lijkt het probleem eerder te vergroten dan te verkleinen. Conclusies dat er een significant verband is tussen, laten we zeggen, aandelenrendementen en het weer, zijn immers veel interessanter dan conclusies die vertellen dat zo’n verband niet aantoonbaar is.

Uit het bovenstaande blijkt dat het niet moeilijk is een schijnbaar statistisch significant verband te vinden tussen een waargenomen grootheid en toekomstige aandelenrendementen.²³ Een dergelijk verband kan puur toeval zijn, en uitsluitend

²⁰ Te berekenen als $1-0.95^{24}$. De 24 toetsen zijn in dit geval onafhankelijk.

²¹ Dit probleem wordt ook wel aangeduid met ‘data sharing’; zie bijvoorbeeld Sullivan, Timmermann and White (2002).

²² Zie bijvoorbeeld Leamer (1978) en Lovell (1983).

²³ De volgende voorspellingsregel werkte bijzonder goed in de Verenigde Staten tussen 1900 en 2000: indien het jaar eindigt op 5 zijn de rendementen zeer hoog; indien het jaar eindigt op 0 of 7 zijn de rendementen zeer laag; zie Dimson, Marsh en Staunton (2002). De huwelijken van Elizabeth Taylor blijken ook jaren met hoge rendementen te voorspellen.

te wijten aan een uitgebreide zoektocht naar voorspellende factoren. Op dezelfde wijze is het niet moeilijk beleggingsstrategieën te construeren die – in een gegeven periode – aantrekkelijke resultaten laten zien. Indien het succes van zo'n strategie uitsluitend een historisch toeval betreft, dan is het evident dat we in de toekomst niets bijzonders van deze strategie hoeven te verwachten. Net zo min als ik kan verwachten van u te winnen als we in uur 23 gaan dobbelen.

Om een link met beleggingsfondsen te maken, wil ik even een hypothetisch beleggingsinstituut beschouwen. Stel dat dit instituut een aantal fondsen in de markt zet, zonder al te veel publiciteit. Deze fondsen beleggen volgens verschillende strategieën. Na een paar jaar zal zeker één ervan, puur toevallig, relatief hoge rendementen laten zien. Vervolgens begint men een uitgebreide publiciteitscampagne rondom dit fonds, terwijl de andere fondsen stilletjes worden afgevoerd. Puur toeval wordt op zo'n manier gemakkelijk verkocht als 'deskundigheid', en deskundigheid heeft het beleggingsinstituut natuurlijk volop in huis!²⁴

Wat zijn nu de lessen die we hieruit kunnen trekken en op welke wijze kan het probleem van data snooping aangepakt worden? Een belangrijke les is de bewustwording van het probleem. Dit geldt niet alleen als we zelf onderzoek doen, maar ook bij de beoordeling van de resultaten van anderen. We moeten ons niet laten misleiden door 'statistisch bewijs' en bijbehorende enthousiaste verhalen, als niet duidelijk is op welke manier dit bewijs verkregen is. Dit geldt a fortiori indien er geen overtuigend theoretisch argument voor het gevonden verband gegeven wordt.

Als we willen corrigeren voor het data snooping probleem wordt het al wat lastiger. Zoals we bij het corrigeren voor look-ahead bias informatie nodig hebben over de mensen die langs de achterdeur van het casino verdwijnen, zo hebben we in dit geval informatie nodig over het aantal alternatieven dat geprobeerd is en het karakter daarvan. Hoewel auteurs vaak wel iets laten doorschemeren over het zoekproces, is de informatie daarover zelden volledig. Soms kunnen we wel een universum van mogelijkheden construeren, bijvoorbeeld als we kijken naar kalendereffecten in aandelenrendementen, of een bepaalde klasse van (technische) handelsregels.²⁵ In het voorbeeld van de dobbelstenen betekent dit dat we het te gebruiken significantieniveau van onze toetsen zodanig aanpassen dat er in werkelijkheid nog steeds een kans van 5% is om een verkeerde conclusie te trekken, en niet 71%. Om dit te bereiken zouden we de nulhypothese van een zuivere dobbelsteen nog slechts verwerpen als de *p*-waarde kleiner is dan 0,21% in plaats van 5%. In complexere situaties kunnen methoden toegepast worden die

²⁴ Zie ook McQueen en Thorley (1999).

²⁵ Zie bijvoorbeeld Sullivan, Timmermann & White (1999, 2001). In het laatste artikel worden bijna 9500 mogelijke kalendereffecten geanalyseerd.

gebruikmaken van simulatie of zogenoemde bootstrap methoden.²⁶ Ook een Bayesiaanse benadering kan in bepaalde gevallen een adequate oplossing bieden.²⁷

Dergelijke methoden zijn vaak bewerkelijk en vereisen kennis over het aantal specificaties of strategieën dat impliciet of expliciet getoetst is. Er zijn andere mogelijkheden om het gevaar van data snooping te beperken. Bijvoorbeeld door een deel van de gegevens te gebruiken om toetsen uit te voeren of strategieën uit te proberen, en een ander deel om na te gaan wat er van de gevonden resultaten overblijft. Als we bijvoorbeeld over 1960-1990 zouden vinden dat aandelenrendementen in januari significant hoger zijn dan in andere maanden, kunnen we de periode 1990-2002 gebruiken om na te gaan of dit werkelijk zo is. Op een vergelijkbare manier zouden we ook gegevens van andere beurzen dan de Amerikaanse kunnen gebruiken om gevonden regulariteiten te toetsen.²⁸

Daarnaast is het aan te raden eerst de nul- en alternatieve hypothese te formuleren, op basis van intuïtie of economische theorie, en deze vervolgens te toetsen, in plaats van eerst een groot aantal toetsen uit te voeren en daarna de theorie te bedenken die bij de gevonden resultaten past. Niet altijd is dit advies praktisch uitvoerbaar, bijvoorbeeld omdat de gangbare theorie over voorspelbaarheid van aandelenrendementen stelt dat er geen voorspelbaarheid is.

Helaas is het een utopie te denken dat we vertekeningen als gevolg van data snooping volledig kunnen voorkomen. Data snooping gebeurt vaak subtiel of indirect, bijvoorbeeld door voort te bouwen op de resultaten van ander onderzoek. Bovendien hebben we als wetenschappers de natuurlijke neiging onze aandacht te richten op de afwijkingen van het normale, op de anomalieën en de ongewone correlaties. Die vormen immers de basis voor de vooruitgang van onze kennis.²⁹

Laat ik dit thema afsluiten met een voorbeeld, gebaseerd op recent onderzoek (Marquering en Verbeek, 2001). In dit voorbeeld wordt de Standard and Poor 500 index voorspeld voor de maanden januari 1966 tot en met augustus 2001. Dit voorspellen gebeurt op basis van eenvoudige lineaire modellen die geschat worden met gegevens die op het moment van voorspellen daadwerkelijk beschikbaar zijn. De verklarende variabelen in de modellen staan vast en worden door ons niet geselecteerd op basis van statistische criteria. Wel is het zo dat de gebruikte

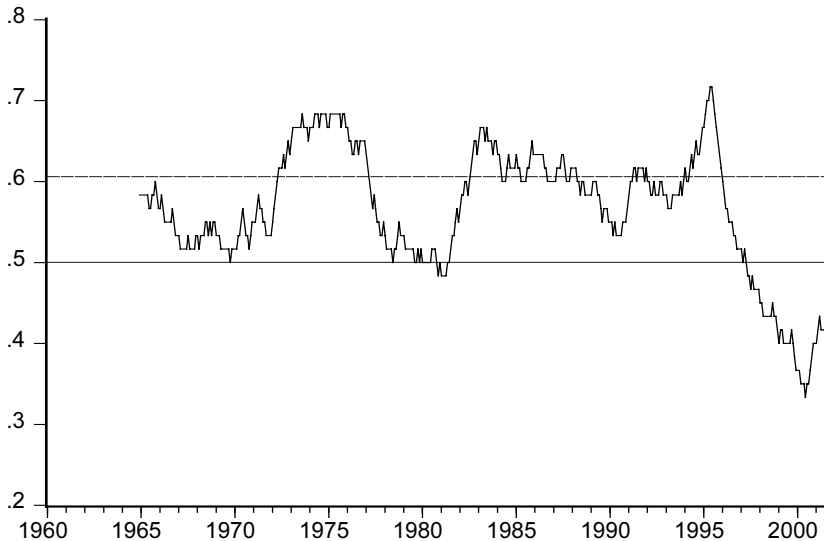
²⁶ Zie White (2000).

²⁷ Zie bijvoorbeeld Cremers (2002).

²⁸ Blijft natuurlijk de mogelijkheid dat door institutionele verschillen bepaalde effecten wel in Amerikaanse gegevens aanwezig zijn, maar niet op andere markten.

²⁹ Problemen van data snooping spelen ook bij andere vragen in de beleggingswereld een rol, bijvoorbeeld bij het toetsen van prijsvormingsmodellen voor financiële activa (zie Lo en MacKinlay, 1990).

variabelen gemotiveerd worden door andere studies waardoor indirecte data snooping, met name in het begin van de steekproefperiode, niet is uit te sluiten. Indien we de gehele steekproefperiode beschouwen blijkt dat we in 56,2% van de 500 maanden de richting van het rendement correct kunnen voorspellen. Dit is zeer beperkt, maar statistisch significant. In de tweede helft van de steekproef is de hit ratio 54,8%. Het lijkt er dus op dat de gevonden voorspelbaarheid afneemt in de loop der tijd, maar wel redelijk stand houdt. Echter, indien we kijken naar de voorspelkwaliteit in de loop van de tijd, zijn we dat de kwaliteit van de voorspellingen behoorlijk varieert en in het laatste decennium zeker te wensen over laat. In Figuur 5 zien we de ontwikkeling van de hitratio's, steeds over de voorafgaande vijf jaar.³⁰ Voor de vijfjaars periode tot 1995 vinden we een hitratio van bijna 70%, terwijl deze vijf jaar later is gedaald tot iets meer dan 30%.



Figuur 5: vijfjaarlijkse voortschrijdende gemiddelde hit ratio's voor rendement op S&P 500.

³⁰ De onderbroken lijn geeft de kritieke waarde aan. Een eenzijdige toets op het bestaan van positieve afhankelijkheid tussen voorspelde en gerealiseerde rendementen over de afgelopen vijf jaar verwerpt indien de hit ratio groter is dan 0,61 (bij een betrouwbaarheid van 95%).

Wat betekent dit? Is het toeval dat het de laatste vijf jaar zo slecht gaat? Of is het juist toeval dat het in eerdere jaren zo goed gaat? Of is de aandelenmarkt zo veranderlijk dat historische patronen zich zelden herhalen? Dit zijn moeilijke vragen, waarop met eenvoudige statistische analyses geen eenduidig antwoord op is te geven. Dit zijn desondanks vragen waarop talloze mensen in de beleggingswereld *een* antwoord formuleren, soms zelfs gestaafd door statistisch bewijs. Maar inmiddels hebt u wel begrepen dat statistisch bewijs vrijwel een *contradictio in terminis* is.

Voor ik ga afsluiten, nog even het volgende. In het bovenstaande heb ik regelmatig vergelijkingen gemaakt met de roulettetafel in het casino. Betekent dit dat ik er vanuit ga dat beleggen eigenlijk hetzelfde is als gokken? Niet helemaal. Op lange termijn word je als belegger beloond door een risicopremie. De hoogte van deze premie hangt af van de hoeveelheid risico die je bereid bent te nemen.³¹ Dit betekent dat je als belegger gemiddeld genomen een positief rendement kunt verwachten, hoewel het natuurlijk altijd slecht kan uitpakken. Afgezien van deze risicopremie, laten toekomstige rendementen zich amper voorspellen. We hebben dat in het bovenstaande gezien. Het onvoorspelbare karakter heeft beleggen gemeen met roulette spelen. Aan de roulettetafel verdwijnt gemiddeld 1/37^e deel van de inzet, zodat je daarbij op lange termijn gemiddeld genomen verliest.

Samenvatting en conclusies

Het afgelopen academische uur heb ik met u enkele concrete voorbeelden uit de financiële wereld besproken waarin empirisch onderzoek een belangrijke rol speelt. De selectie van voorbeelden was niet geheel willekeurig. In beide gevallen bleek de mogelijkheid te bestaan dat empirische resultaten iets lijken aan te tonen, terwijl dat in werkelijkheid niet zo is. Verschillende soorten van selectieprocessen speelden hierbij een rol.

Hoewel het probleem van misleidende resultaten in beide situaties gemeenschappelijk was, is de oplossing niet geheel identiek. Indien de steekproef die voor onderzoek gebruikt wordt onderhevig is geweest aan selectieprocessen, is het mogelijk hiervoor bij de analyse te corrigeren, mits we maar voldoende informatie hebben over het selectieproces. Bij het probleem van data snooping zou dit in principe ook kunnen, maar daar ontbreekt het ons in het algemeen aan voldoende informatie over het zoekproces, of is dit proces zodanig gecompliceerd en onoverzichtelijk dat dit zeer lastig wordt.

Mede gezien de overweldigende hoeveelheid gegevens die momenteel beschikbaar is, zal empirisch onderzoek op financierings- en beleggingsgebied een

³¹ Er bestaat een uitgebreide literatuur over de vraag welke bronnen van risico een risicopremie kennen. Zie bijvoorbeeld Campbell (2000) of Cochrane (2001).

belangrijke rol blijven spelen. De interactie van theoretische inzichten en empirische resultaten is de motor die ons begrip van de financiële markten doet groeien. Echter, empirische resultaten die op het eerste gezicht een statistisch overtuigend bewijs leveren, blijken dat bij nader onderzoek toch niet altijd te doen. Mocht u daarvan nog niet overtuigd zijn, dan nodig ik u graag uit met mij te dobbelen, vanavond tussen 22 en 23 uur.

Dankwoord

Dames en Heren,

Tot slot spreek ik graag enige woorden van dank uit. Het College van Bestuur van deze universiteit en de decanen van de faculteiten der Bedrijfskunde en Economische Wetenschappen ben ik zeer erkentelijk voor het in mij gestelde vertrouwen. Ik aanvaard mijn benoeming tot Hoogleraar Ondernemings-financiering met veel plezier.

Hooggeleerde Nijman, beste Theo,

Ik heb het geluk en genoegen gehad om in Tilburg als onderzoeker-in-opleiding onder jouw inspirerende leiding aan mijn proefschrift te mogen werken, in een tijd dat jouw agenda nog niet overheerst werd door bestuurlijke en andere verplichtingen. Gedurende die periode, en de jaren daarna, heb ik veel van je geleerd. Het valt niet in enkele zinnen samen te vatten op welke wijze jij van invloed bent geweest op mijn wetenschappelijke carrière maar ik ben er trots op dat jouw geest, ook bij de projecten waaraan we niet hebben samengewerkt, in mijn werk is terug te vinden.

Hooggeleerde Koedijk, beste Kees,

Enkele telefoontjes en een paar gesprekken met jou waren voldoende om mij te overtuigen de overstap naar Rotterdam te maken. Het is een bijzondere uitdaging om samen met jou te werken aan de verdere opbouw en uitbouw van een internationaal-vooraanstaande onderzoeksgroep in Financieel Management. Ik ben je bijzonder dankbaar voor het vertrouwen dat je in mij stelt.

Hooggeleerde van Dijk, beste Herman,

Een aantrekkelijk aspect van de leerstoel die ik vandaag met deze rede aanvaard, is dat deze gevestigd is aan twee faculteiten. Ik ben bijzonder vereerd dat ik mag bijdragen aan de uitstekende reputatie van het Econometrisch Instituut, en het is een groot genoegen de komende jaren met je mogen samen te werken. Een gezamenlijk aio-project is hiervan een eerste materialisatie.

Collega's van de vakgroep Financieel Management,

Onze vakgroep is een schip dat sinds enige tijd op de juiste koers zit. Naast de kapitein en de stuurman zijn jullie voor een belangrijk deel verantwoordelijk voor het vasthouden van deze koers. Jullie enthousiasme staat garant voor een stimulerende en opbouwende sfeer en ik hoop nog lang aan boord te kunnen blijven.

Lieve Timo, Thalia en Tamara,

Elke dag, wanneer ik thuiskom, zetten jullie me weer met beide benen op de grond: er zijn belangrijker zaken in het leven dan vakgroepbegrotingen, discussion papers en hooglerarenvergaderingen. Ik prijs mij gelukkig met jullie aanwezigheid in ons gezin en met het feit dat jullie geen idee hebben wat al deze moeilijke woorden betekenen.

Lieve Marcella,

Als geen ander weet jij me in mijn persoonlijke leven op de juiste koers te houden. Jouw liefde en onvoorwaardelijke steun zijn onvolprezen.

Ik heb gezegd.

Referenties

Ang, A. and G. Bekaert, 2001, Stock Return Predictability: Is it There?, Working paper, Columbia Business School (NBER Working Paper No. W8207).

Bacquero, G., J.R. ter Horst & M. Verbeek, 2002, Survival, Look Ahead Bias and the Performance of Hedge Funds, paper to be presented at the annual meeting of the European Financial Management Association, June 2002, London.

Baks, K., A. Metrick, & J. Wachter, 2001, Should Investors Avoid All Actively Managed Mutual Funds? A Study in Bayesian Performance Evaluation, *Journal of Finance*, 56, 45-85.

Bossaerts, P. & P. Hillion, 1999, Implementing Statistical Criteria to Select Return Forecasting Models: What Do We Learn?, *Review of Financial Studies*, 12, 405-428.

Breen, W., L.R. Glosten & R. Jagannathan, 1989, Economic Significance of Predictable Variations in Stock Index Returns, *Journal of Finance*, 44, 1177-1189.

Brown S.J. & W.N. Goetzmann, 1997, Mutual Fund Styles, *Journal of Financial Economics*, 43, 373-399.

Brown S.J. & W.N. Goetzmann, 2001, Hedge Funds with Style, Yale ICF Working Paper No. 00-29.

Brown, S.J., W.N. Goetzmann & S.A. Ross, 1995, Survival, *Journal of Finance*, 50, 853-873.

Busse, J.A. and P.J. Irvine, 2002, Bayesian Alphas and Mutual Fund Persistence, working paper, Goizueta Business School, Emory University.

Campbell, J.Y., 1987, Stock Returns and the Term Structure, *Journal of Financial Economics*, 18, 373-399.

Campbell, J.Y., 2000, Asset Pricing at the Millennium, *Journal of Finance*, 55, 1515-1567.

Carhart, M.M., 1997, On Persistence in Mutual Fund Performance, *Journal of Finance*, 52, 57-82.

Cochrane, J.H., 2001, *Asset Pricing*, Princeton University Press.

Cremers, M., 2002, Stock Return Predictability: A Bayesian Model Selection Perspective, *Review of Financial Studies*, te verschijnen.

DeRoos, F.A., Th.E. Nijman & J.R. ter Horst, 2002, Evaluating Style Analysis, CEPR discussion paper 3181.

DiBartolomeo, D., & E. Witowski, 1997, Mutual Fund Misclassification: Evidence Based on Style Analysis, *Financial Analysts Journal*, 53, 32-43.

Dimson, E., P. Marsh & M. Staunton, 2002, *Triumph of the Optimists: 101 Years of Global Investment Returns*, Princeton University Press.

Fama, E.F. & K.R. French, 1993, Common Risk Factors in the Returns on Bonds and Stocks, *Journal of Financial Economics*, 33, 3-53.

Fitzgerald, J., P.Gottschalk & R. Moffitt, 1998, An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics, *Journal of Human Resources*, 33, 251-299.

Fung, W. & D.A. Hsieh, 2002, Hedge-Fund Benchmarks: Information Content and Measurement Error Biases, *Financial Analyst Journal*, 58, 22-34.

Gruber, M.J., 1996, Another Puzzle: The Growth in Actively Managed Mutual Funds, *Journal of Finance*, 51, 783-810.

Heckman, J.J., 1979, Sample Selection Bias as a Specification Error, *Econometrica*, 47, 153-161.

Hendricks, D., J. Patel & R. Zeckhauser, 1993, Hot Hands in Mutual Fund Performance: Short-Run Persistence of Relative Performance, 1974-1988, *Journal of Finance*, 48, 93-130.

Hendricks, D., J. Patel & R. Zeckhauser, 1997, The J-shape of Performance Persistence given Survivorship Bias, *Review of Economics and Statistics*, 79, 161-170.

Jegadeesh, N. & S. Titman, 1993, Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency, *Journal of Finance*, 48, 65-91.

Kim, M. R. Shukla and M. Tomas, 2000, Mutual Fund Objective Misclassification, *Journal of Economics and Business*, 52, 309-323.

Kirby, C., 1998, The Restrictions on Predictability Implied by Rational Asset Pricing Models, *Review of Financial Studies*, 11, 343-382.

Leamer, E., 1978, *Specification Searches*, John Wiley and Sons, New York.

Lo, A.W. and A.C. MacKinlay, 1990, Data-Snooping Biases in Tests of Financial Asset Pricing Models, *Review of Financial Studies*, 3, 431-468.

Lovell, M.C., 1983, Data Mining, *The Review of Economics and Statistics*, 65, 1-12.

Marquering, M. & M. Verbeek, 2001, The Economic Value of Predicting Stock Index Returns and Volatility, ERIM report 2001-75, Erasmus Universiteit Rotterdam.

McQueen, G. & S. Thorley, 1999, Mining Fool's Gold, Business Management Department, Brigham Young University.

Nijman, Th.E. & M. Verbeek, 1992, Nonresponse in Panel Data: The Impact on Estimates of a Life Cycle Consumption Function, *Journal of Applied Econometrics*, 7, 243-257.

Pástor, L. & R. Stambaugh, 2002a, Mutual Fund Performance and Seemingly Unrelated Assets, *Journal of Financial Economics*, 63, 315-349.

Pástor, L. & R. Stambaugh, 2002b, Investing in Equity Mutual Funds, *Journal of Financial Economics*, 63, 351-380.

Pesaran, M.H. & A.G. Timmermann, 1995, Predictability of Stock Returns: Robustness and Economic Significance, *Journal of Finance*, 50, 1201-1228.

Pesaran, M.H. & A.G. Timmermann, 2000, A Recursive Modelling Approach to Predicting UK Stock Returns, *Economic Journal*, 110, 159-191.

Rouwenhorst, G., 1998, International Momentum Strategies, *Journal of Finance*, 53, 267-284.

Sullivan, R., A. Timmermann & H. White, 1999, Data-Snooping, Technical Trading Rules, and the Bootstrap, *Journal of Finance*, 54, 1647-1692.

Sullivan, R., A. Timmermann & H. White, 2001, The Dangers of Data-Driven Inference: the Case of Calendar Effects in Stock Returns, *Journal of Econometrics*, 105, 249-286.

Sullivan, R., A. Timmermann & H. White, 2002, Scientific Progress with Data Sharing, *Journal of Forecasting*, te verschijnen.

Swinkels, L., Th.E. Nijman & M. Verbeek, 2002, Do Countries or Industries Explain Momentum in Europe?, paper to be presented at the annual meeting of the European Finance Association, 21-25 August 2002, Berlin.

Swinkels, L., P.J. Van Der Sluis and M. Verbeek, 2002, Mutual Fund Manager Ability when Styles are Changing, working paper, Erasmus Universiteit Rotterdam.

Ter Horst, J.R., Th.E. Nijman & M. Verbeek, 2001, Eliminating Look-Ahead Bias in Evaluating Persistence in Mutual Fund Performance, *Journal of Empirical Finance*, 8, 345-373.

Vella, F. & M. Verbeek, 1998, Whose Wages Do Unions Raise? A Dynamic Model of Unionism and Wage Rate Determination for Young Men, *Journal of Applied Econometrics*, 13, 163-183.

Vella, F. & M. Verbeek, 1999, Two-Step Estimation of Panel Data Models with Censored Endogenous Variables and Selection Bias *Journal of Econometrics*, Vol. 90, 239-263.

Verbeek, M. & Th.E. Nijman, 1992, Testing for Selectivity Bias in Panel Data Models, *International Economic Review*, 33, 681-703.

Wermers, R., 2000, Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transaction Costs, and Expenses, *Journal of Finance*, 55, 1655-1694.

White, H., 2000, A Reality Check for Data Snooping, *Econometrica*, 68, 1097-1126.

Erasmus Research Institute of Management
Inaugural Addresses Research in Management Series

www.erim.eur.nl

Triple inaugural address for the Rotating Chair for Research in Organisation and Management

Quality Management Research: Standing the Test of Time,

Prof. dr. B.G. Dale

Performance Related Pay - Another Management Fad?,

Prof.dr. R. Richardson

From Downsize to Enterprise: Management Buyouts and Restructuring Industry

Prof. dr. D.M. Wright

Reference number ERIM: EIA-2001-01-ORG

ISBN 90-5892-006-2

URL: <http://www.eur.nl/WebDOC/doc/iarm/erimia20010405124454.pdf>

Financial Regulation; Emerging from the Shadows

Prof. dr. Harald. A. Benink

Reference number ERIM: EIA-2001-02-ORG

ISBN 90-5892-007-0

URL: <http://www.eur.nl/WebDOC/doc/iarm/erimia20010628134057.pdf>

Opsporen van sneller en beter. Modelling through...

Prof. dr. Leo G. Kroon

Reference number ERIM: EIA-2001-03-LIS

ISBN 90-5892-010-0

East, West, Best: Cross cultural encounters and measures

Prof. dr. Slawomir Jan Magala

Reference number ERIM: EIA-2001-04-ORG

ISBN 90-5892-013-5

Leadership as a source of inspiration

Prof. dr. Deanne N. Den Hartog

Reference number ERIM: EIA-2001-05-ORG

ISBN 90-5892-015-1

Marketing Informatie en besluitvorming: een inter-organisatoneel perspectief

Prof. dr. ir. Gerrit H. van Bruggen

Reference number ERIM: EIA-2001-06-MKT

ISBN 90 -5892- 016 - X

**The residual: On monitoring and Benchmarking Firms, Industries and Economies
with respect to Productivity**

Prof. dr. Bert M. Balk

Reference number ERIM: EIA-2001-07-MKT

ISBN 90 –5892 - 018 – 6

“Nut en nog eens nut”

Over retoriek, mythes en rituelen in informatiesysteemonderzoek

Prof. dr. H.G. van Dissel

Reference number ERIM: EIA-2002-08-LIS

ISBN 90 –5892 - 018 – 6

