Factorbeleggen in het tijdperk van de machine

Beschrijving

De geschiedenis van factorbeleggen begint in een wereld met computers die krachtig genoeg waren om er iets mee te doen. Toen Eugene Fama en Kenneth French in 1992 hun driefactorenmodel publiceerden, was het inzicht dat kleine bedrijven en goedkope aandelen structureel beter presteerden dan de markt al langer bekend, maar het bewijs lag er nu in een robuuste academische vorm. Beleggers die kochten op basis van marktkapitalisatie, koers-boekwaardeverhouding en marktbÃ¨ta konden historisch een hoger rendement verwachten. Dat was de belofte. De praktijk bleek gecompliceerder.

Van academisch inzicht naar beleggingsproduct

De weg van het academische factormodel naar het verhandelbare beleggingsproduct duurde langer dan de theorie deed vermoeden. In de jaren negentig begonnen kwantitatieve fondsen zoals die van AQR en Dimensional de inzichten van Fama en French te operationaliseren, maar het grote publiek bleef factorbeleggen voornamelijk kennen als â??valueâ?• of â??small capâ?•, termen die intuÃ¯tief aanspraken zonder dat de onderliggende mechanica volledig begrepen werden. De definitieve democratisering van factorbeleggen vond pas plaats met de opkomst van smart beta-ETFâ??s in het eerste decennium van deze eeuw. Plotseling kon een particuliere belegger voor een fractie van de kosten van een actief beheerd fonds systematisch blootstelling kopen aan waarde, momentum, kwaliteit of lage volatiliteit. Factorbeleggen werd een product.

Met die producten kwamen ook de eerste problemen. Zodra een factor voldoende bekend is en voldoende kapitaal aantrekt, beginnen de rendementen te eroderen. De value-premie, die over tientallen jaren historisch robuust was gebleken, presteerde tussen 2010 en 2020 structureel lager dan de markt. Groeiaandelen, gedreven door de technologiesector, versloegen systematisch alles wat op basis van klassieke waarderingsmaatstaven goedkoop heette. Werkte de factor niet meer? Of was de meetperiode te kort? Het debat is tot op de dag van vandaag niet beslecht, maar het stelde een fundamentele vraag: zijn factorpremies economische werkelijkheid of statistische die weg gearbitreerd

worden zodra ze ontdekt zijn?

De grenzen van het klassieke model

Het klassieke factormodel, hoe verfijnd ook, rust op een reeks aannames die de werkelijkheid vereenvoudigt tot wat meetbaar is in historische koersdata. Dat is tegelijk de kracht en de zwakte. De kracht is reproduceerbaarheid: iedereen die dezelfde data en dezelfde methodologie gebruikt, komt op hetzelfde resultaat. De zwakte is starheid: de factoren worden gedefinieerd op basis van wat in het verleden gemeten is, niet op basis van wat economisch relevant is voor de toekomst.

Neem kwaliteit als factor. In het Fama-French vijffactorenmodel, dat in 2015 werd gepubliceerd als uitbreiding op het oorspronkelijke model, wordt kwaliteit geoperationaliseerd via winstgevendheid en investeringspatronen. Dat zijn boekhoudkundige maatstaven die per definitie achteruitkijken. Een bedrijf dat vandaag een hoge return on equity rapporteert, deed dat op basis van beslissingen die jaren geleden werden genomen. De vraag of dat bedrijf over vijf jaar nog steeds een hoge kwaliteit heeft, wordt niet beantwoord door historische boekhoudkundige data alleen. De markt verdisconteert informatie die ver buiten het bereik van een kwantitatief factormodel valt: de kwaliteit van het management, de veerkracht van het bedrijfsmodel, de technologische positie in een snel veranderende sector.

Momentum is een ander geval. Het is empirisch een van de robuuste factoren over activaklassen en geografieÃ«n heen, maar de theoretische fundering is altijd omstreden gebleven. Waarom zou iets dat in de afgelopen twaalf maanden gestegen is, de komende maanden blijven stijgen? Gedragseconomen wijzen op trage informatieverwerking en kuddegedrag. Risicodenkers wijzen op correlaties met economische cycli. De factor werkt, maar niemand is het er volledig over eens waarom. Dat maakt hem kwetsbaar: zodra de marktomstandigheden veranderen, kan momentum abrupt omslaan, zoals bleek in de marktcorrectie van 2020 en opnieuw in de draai van 2022.

De intrede van kunstmatige intelligentie

Kunstmatige intelligentie is niet de vervanger van factorbeleggen maar meer een katalysator. De meest directe toepassing is de uitbreiding van het data-universum. Klassiek factorbeleggen werkt met gestructureerde financiÃ«le data: koersen, balansen, winst-en-verliesrekeningen. Dat is een fractie van de informatie die beschikbaar is over beursgenoteerde bedrijven. Satelietbeelden van parkeerplaatsen bij winkelcentra, het sentiment in kwartaalverslagen gemeten via natuurlijke-taalverwerking, het patroon van mobiliteitsdata rond fabrieken, het verloop van vacatures op LinkedIn, dit zijn allemaal signalen die vroeger niet systematisch te verwerken waren maar nu, via machine learning, kunnen worden omgezet in kwantitatieve beleggingssignalen.

De implicaties zijn aanzienlijk. Factoren die vroeger maanden of kwartalen achterliepen op de economische werkelijkheid omdat ze afhankelijk waren van gerapporteerde financiÃ«le data, kunnen nu worden aangevuld met vrijwel realtime informatie. Een kwantitatief model dat de kwaliteitsfactor meet via boekhoudkundige winstgevendheid, kan worden gecombineerd met een taalmodel dat managementcommentaar analyseert op toon, consistentie en specifiekheid. De combinatie is rijker dan elk van de onderdelen afzonderlijk.

Alternatieve data en de erosie van de informatieasymmetrie

Toch schuilt hier een herhaling van het productiseringsdilemma. Alternatieve data die vandaag een competitief voordeel bieden, worden morgen breed beschikbaar en verliezen daarmee hun informatieve waarde. Satelietdata over grondstoffenvoorraden, ooit het exclusieve domein van gespecialisedderde hedgefondsen, worden inmiddels aangeboden als gestandaardiseerde datafeeds aan ieder institutioneel belegger die bereid is te betalen. Het signaal verdwijnt zodra het signaal breed bekend is. Kunstmatige intelligentie versnelt dat proces: de technologie om grote hoeveelheden ongestructureerde data te verwerken wordt democratischer, de verwerkingstijd korter en de toegangsdrempel lager.

Wat overblijft als informatieasymmetrie in brede zin verdampt, is de kwaliteit van het model zelf. Hier liggen de diepste vragen over de rol van kunstmatige intelligentie in factorbeleggen. Een traditioneel factormodel is transparant: de belegger weet welke variabelen worden gebruikt, hoe zij worden gewogen en waarom. Een neuraal netwerk dat honderden signalen verwerkt in een architectuur die voor mensen niet volledig doorgrondelijk is, biedt mogelijk betere voorspellingen maar schept een nieuwe vorm van onzekerheid. Is de outperformance van het model structureel en economisch verklaarbaar, of is het patroonherkenning in historische data die niet generaliseert naar nieuwe marktomstandigheden? Die vraag, die in de academische literatuur bekendstaat als het overfitting-probleem, is bij klassieke factormodellen reeds aanwezig maar wordt bij complexe machine-learningmodellen acuter.

De combinatie als toekomst

De meest veelbelovende richting is niet de vervanging van klassiek factorbeleggen door kunstmatige intelligentie, maar de integratie van beide. Factoren bieden structuur, theoretische fundering en interpreteerbare mechanica. Kunstmatige intelligentie biedt schaal, snelheid en het vermogen om niet-lineaire verbanden en alternatieve databronnen te verwerken. Een model dat de economische logica van een factor als winstgevendheid combineert met het vermogen van een taalmodel om de voorwaartse kwaliteit van een bedrijf te beoordelen op basis van tekst, audiofragmenten en gedragspatronen van insiders, is potentieel robuuster dan elk van de componenten afzonderlijk.

Dat vereist wel een herijking van de governancevraag. Wie is verantwoordelijk voor de beslissingen van een systeem dat deels door mensen en deels door machines wordt aangestuurd? Hoe wordt overfitting voorkomen in een model dat periodiek opnieuw wordt getraind op groeiende datasets? En hoe wordt de factor-zoo vermeden, de neiging om steeds meer variabelen toe te voegen totdat elke historische

periode in-sample perfect verklaard wordt maar out-of-sample geen voorspellende waarde meer heeft?

Wat blijft

Factorbeleggen in zijn oorspronkelijke vorm was een antwoord op een specifiek probleem: hoe kan een belegger systematisch en kostenefficiÃ«nt blootstelling verkrijgen aan rendementskarakteristieken die over lange perioden robuust zijn gebleken? Die vraag is niet veranderd. Wat veranderd is, zijn de instrumenten waarmee zij beantwoord kan worden en de complexiteit van het informatielandschap waarin factoren opereren. Kunstmatige intelligentie maakt het mogelijk factoren scherper te definiÃ«ren, sneller te actualiseren en te combineren met signalen die vroeger niet beschikbaar of niet verwerkbaar waren. Maar zij lost het fundamentele dilemma van factorbeleggen niet op: premies die voldoende zichtbaar zijn om systematisch te exploiteren, trekken kapitaal aan, en kapitaal dat een premie najaagt, erodeert haar op termijn.

De les van dertig jaar factorbeleggen is dat de meest duurzame premies verankerd zijn in economische logica die niet verdwijnt zodra ze bekend is. Risicocompensatie voor illiquiditeit of complexiteit, een gedragsmatig patroon dat diepgeworteld is in menselijke psychologie, een structureel informatieverschil dat niet volledig arbitrageerbaar is, dat zijn de fundamenten waarop een robuuste factor rust. Kunstmatige intelligentie kan helpen die fundamenten beter te identificeren en te exploiteren, maar zij kan ze niet creÃ«ren waar ze niet bestaan. De machine versterkt wat er is. Zij vervangt het economische denken niet.

Han Dieperink, Chief Investment Officer

Disclaimer

Abonneer op deze nieuwsbrief via

www.aureus.eu/nieuwsbrief