

Regime-afhankelijke risk parity-allocatie met GARCH–HMM voor dynamisch portefeuillebeheer

Julius Sepmeijer

Begeleid door: Dhr. M. de Kok

Vak: Economie en Wiskunde D

Inleverdatum: 30 november 2025

Samenvatting

In dit profielwerkstuk wordt een beleggingsmodel ontwikkeld dat het risico zo klein mogelijk probeert te houden, vooral in perioden van marktstress. Het model combineert drie onderdelen: een EGARCH-model dat inschat hoe onrustig een asset wordt, een Hidden Markov Model dat bepaalt of de markt zich in een rustig of stressvol regime bevindt, en een risk-parityverdeling die het risico evenwichtig over de portefeuille spreidt. Het model is getest op data van 2007 tot 2025. Uit deze analyse blijkt dat de strategie grote dalingen tijdens crisisperioden goed opvangt en over de hele periode een gelijkmatige groei laat zien. Het model zorgt in de backtest voor een gemiddeld jaarlijks rendement van 5,2% en een volatiliteit van 8,7%. In zeer sterke markten blijft het rendement iets achter, maar het totale risico ligt duidelijk lager. Daardoor presteert de strategie als defensieve aanpak vergelijkbaar met – en in sommige gevallen beter dan – traditionele defensieve methoden. De resultaten laten zien dat een dynamische aanpak die meebeweegt met de volatiliteit een waardevol alternatief kan vormen voor beleggers die stabiliteit en rust belangrijk vinden.

Voorwoord

Het was een nazomerse middag in 2024 in Zadar, een stad aan de Kroatische kust. Ik zat met een vriend te praten over finance. Hij vertelde me over de zogenoemde One Billion Dollar Equation: de Black-Scholes-vergelijking. Dat was het moment waarop ik ontdekte hoe prachtig wiskunde en finance samen kunnen komen. Diezelfde dag heb ik urenlang doorgelezen over deze bijzondere vergelijking en daar is mijn passie voor kwantitatieve finance begonnen. Mijn proces begon met het lezen van het boek Rich Dad Poor Dad van Robert T. Kiyosaki en Sharon Lechter. Het was mijn eerste boek dat over geld ging, en het heeft mij veel waardevols geleerd. Een van de belangrijkste inzichten uit het boek is dat het rendement op investeringen in jezelf het hoogst is; geen andere investering is zo rendabel.

Dat advies heb ik opgevolgd en mijn doel was om dit actief toe te passen in mijn leven. Sindsdien ben ik me blijven verdiepen in finance. Boeken zoals The Psychology of Money en The 7 Habits of Highly Effective People hebben mij intuïtieve en fundamentele lessen over geld meegegeven. Geld is mooi; het is onderdeel van ons leven. Diegene die macht heeft over zijn emoties en zijn gedachten, heeft macht over geld. In deze periode heb ik me gericht op het leren van nieuwe vaardigheden: ik ben bijlessen gaan geven, heb veel boeken over geld gelezen en ben verkoper geworden bij een kledingzaak. Sinds kort ben ik ook lid van een politieke partij, en op 18 maart 2026 kan er op mij gestemd worden bij de gemeentelijke verkiezingen. Dit alles was niet mogelijk geweest als ik niet al die tijd in mezelf had geïnvesteerd.

Ik wilde mezelf uitdagen; ik besloot dat ik een kwantitatief model wilde programmeren dat voor mij investeert en het doel heeft om het risico te minimaliseren. Ik besloot om hier mijn PWS over te schrijven. Met mijn deelname aan Pre-University in Delft begon deze reis met het leren over wiskundig modelleren. Ik heb zowel leerlingen als professoren gevraagd hoe ik de leerweg van mijn PWS het beste kon voortzetten, en ik ben dankbaar voor alle adviezen die ik toen kreeg. Ik wil I. Goddijn, professor Wiskunde aan de universiteit, en L. Bos, student Econometrie in Rotterdam, extra bedanken voor de hulp die zij mij aan het begin van het profielwerkstuk hebben gegeven. Dankzij hen heb ik een goed plan kunnen maken voor de wiskunde die ik voor het model moest leren. Ook wil ik mijn begeleider, de heer M. de Kok, bedanken voor zijn ondersteuning en kritische blik; onze gesprekken waren inspirerend en verdiepend. Hij heeft mij geleerd hoe de wereld van finance in elkaar zit, en dat vind ik heel waardevol. Met veel enthousiasme en toewijding heb ik aan dit profielwerkstuk gewerkt. Ik hoop dat het resultaat recht doet aan de inzet die ik in dit traject heb gestoken en dat de lezer dezelfde fascinatie voor kwantitatieve finance ervaart als ik. Ik wens u veel leesplezier.

Voorburg, 30 november 2025
Julius Sepmeijer

Inhoudsopgave

1	Inleiding	5
1.1	Achtergrond en motivatie	5
1.2	Probleemstelling	6
1.3	Doelstellingen van het onderzoek	7
1.4	Onderzoeksvragen en hypotheses	7
1.5	Opbouw van het profielwerkstuk	8
2	Theoretisch kader	10
2.1	Inleiding tot defensieve beleggingsstrategieën	10
2.1.1	De effectieve markt hypothese	10
2.1.2	Adaptieve markt hypothese en behavioral finance	11
2.1.3	Definitie van defensieve strategieën	12
2.1.4	Doelstellingen (kapitaalbehoud, risicobeperking en stabiliteit)	12
2.1.5	Verschil tussen passieve, actieve en kwantitatieve defensieve strategieën	13
2.1.6	Defensieve strategieën: theoretisch overzicht	13
2.2	Fundamenten van portefeuilletheorie	14
2.2.1	Traditionele portfolio-optimalisatie (Markowitz, mean-variance)	14
2.2.2	De wet van grote getallen en diversificatie	16
2.2.3	Risico en volatiliteit in financiële markten	16
2.3	Reageren op veranderende volatiliteit	18
2.3.1	Dynamische volatiliteit voorspellen: ARCH/GARCH	18
2.3.2	Alternatief: EGRACH (1,1)	20
2.3.3	Schatting via maximale likelihood	21
2.3.4	Overzicht van risk parity-strategieën	23
2.4	Aanpassingen aan asset-allocatie na GARCH en risk parity: HMM's	24
2.4.1	Regime-switching modellen en Markov Chains	26
2.4.2	Verborgene regimes	27
2.4.3	HMM voor financiële tijdreeksen	27
2.5	Conclusie: antwoord op deelvraag 1 en 2	28
3	Methode	30
3.1	Onderzoeksopzet	30
3.2	Dataselectie (markten, activa, periode, frequentie)	30
3.3	Constructie van het portefeuillemodel (stappenplan)	32
4	Resultaten	35
4.1	Rendementen en de resultaten van de EGARCH-schattingen	35
4.2	Prestaties van het EH-RP-model	37
5	Discussie	39
5.1	Interpretatie van de resultaten	39
5.2	Vergelijking met andere strategieën: antwoord op deelvraag 3	39
5.3	Aanbevelingen voor vervolgonderzoek, validiteit en betrouwbaarheid	40
6	Conclusie en antwoord op de onderzoeksvraag	42
	Literatuurlijst	43

A Bijlage 1: Python-code van het model	45
B Bijlage 2: Logboek	51

1 Inleiding

1.1 Achtergrond en motivatie

Financiële markten laten plotselinge schommelingen zien met regimewisselingen. Wereldwijde aandelenmarkten kennen hoge en afwisselende risico's. Zo behaalde de MSCI World-index sinds eind jaren 70 een gemiddeld jaarrendement van circa 10,4% met een jaarlijkse volatiliteit van 14,8%, maar kende in crises maximale drawdowns van meer dan 50%. Zo was er in de periode van 2008-2009 een drawdown van 54,5%. Bovendien hebben globalisering en de toename van algoritmische trading ervoor gezorgd dat marktfases en trading steeds dynamischer en efficiënter werken. Dit heeft gezorgd voor een technische verandering in handelen. (MSCI, 2025)

Met de komst van kunstmatige intelligentie is handelen bovendien steeds complexer geworden. Politiek en economie raakt steeds meer met elkaar verweven. Grote techreuzen, zoals Nvidia en Palantir, zorgen voor een steeds grotere bijdrage aan de markten en zijn onmisbaar geworden. Zo heeft Nvidia een geschatte marktwaarde van 5 biljoen dollar; dat is bijvoorbeeld groter dan het BBP van heel Duitsland. Maar ook ING heeft als doel om duizend van de vijftienduizend banen in Nederland te schrappen en te automatiseren met kunstmatige intelligentie binnen één jaar. De vraag naar technische expertise wordt steeds groter. Waar vroeger fundamenteel investeren nog de beste strategie was, hebben grote firma's die technisch investeren nu het voortouw genomen. Dat was vroeger nog anders.

Zo heeft Renaissance Technologies, een fonds opgericht door de wiskundige Jim Simons, een gemiddeld jaarlijks rendement van 66% weten te realiseren. Rond de crisistijd in 2008 was er zelfs sprake van een enorme uitschieter van 152%. Dat kwam niet omdat ze meer wisten van business dan andere experts. Ze waren echter heel goed in één ding: wiskunde. Renaissance Technologies bestaat niet uit finance-experts, maar uit natuurkundigen en wiskundigen. In het boek *The man who solved the market* komt naar voren dat Jim Simons niet hield van business, maar wel van wiskunde en geld. De kwantitatieve analisten zijn goed in statistiek en stochastiek, en kunnen op die manier goed koersbewegingen analyseren en voorspellen, zonder naar macro-economische factoren te kijken. Ze keken puur naar de data. Dat was hun kracht.

Op die manier is de financiële industrie totaal gekanteld. Dat heeft ook gezorgd voor een verandering in markt-theorieën; theorieën die vroeger nog toepasbaar waren op de markt zijn nu onbetrouwbaar en ineffectief. Markowitz' idee dat de volatiliteit van markten constant is (Markowitz, 1952), gaat tegen resultaten van de markt in. Markten zijn dynamischer geworden en kunnen in bepaalde regimes van hoge volatiliteit of lage volatiliteit zitten. Ook zijn markten op sommige momenten efficiënt en op sommige momenten inefficiënt. Dit onderzoek gaat in op de wiskundige aspecten van koersbewegingen en houdt rekening met de ontwikkelingen omtrent de technische vooruitgang. Zo richt het onderzoek zich op het ontwikkelen van een nieuw kwantitatief model voor de asset-allocatie door gebruik te maken van een Hidden Markov Model en een geïntegreerd EGARCH-model. Over deze econometrische modellen volgt uiteraard een verdere toelichting in het profielwerkstuk. Het doel van het onderzoek is om een defensief model te ontwikkelen dat risico minimaliseert en kan concurreren met andere defensieve strategieën, zoals sparen of beleggen in obligaties.

1.2 Probleemstelling

Een verantwoording van het probleem is belangrijk voor het onderzoek, omdat daarmee de urgentie van het model wordt benadrukt. Het probleem is dan ook dat traditionele portefeuille-optimalisatiemethoden, zoals de klassieke Mean-Variance-benadering van Markowitz onvoldoende rekening houden met de dynamiek en onzekerheid in financiële markten. Rendementen zijn vaak niet normaal verdeeld en parameters zijn niet constant, maar stochastisch. Dit model zal in veel situaties werken, maar toch zijn er praktische complicaties. Zo is uit meerdere onderzoeken af te leiden dat het model in de praktijk tot tegenstrijdigheden heeft geleid (Cont, 2001). Zo komen extreme rendementen vaker voor dan het Mean-Variance model laat zien. Het model schiet dus te kort op veel vlakken. Er is een genuanceerdere benadering nodig. Er is behoefte aan een methode die deze onzekerheden meeneemt en rekening houdt met mogelijke regime wisselingen.

Concreet is het doel van een (pensioen)fonds om de *sharpe-ratio* zo hoog mogelijk te houden en het risico zo laag mogelijk. Twee dingen zijn daarbij van belang: het portfolio moet zo min mogelijk risico nemen (1) en het rendement moet zo hoog mogelijk zijn (2). Vaak betekent minder risico ook minder rendement. Een fonds moet de perfecte verhouding vinden tussen risico en rendement. Hedgefondsen vinden het belangrijker om het risico bijna helemaal uit te sluiten tegen een normaal rendement, dan om veel risico te nemen met een hoog rendement. Een kwantitatieve analist zal bij een eenmalige keuze eerder 1 euro met zekerheid accepteren dan 100 euro met slechts 10% kans op winst, omdat de verwachte waarde van de zekere optie hoger is en er geen risico aan verbonden is. Dat komt omdat analisten zoeken naar een bepaalde zekerheid tijdens handelen. Dat is lastig, want markten bewegen stochastisch en zijn daarom dynamisch. De sharpe-ratio moet zo hoog mogelijk zijn, door het rendement zo hoog mogelijk te houden tegen een lage variantie. Oftewel:

$$S = \frac{E[R_p] - R_f}{\sigma(R_p)},$$

waarbij $E[R_p]$ het verwachte rendement van het portfolio voorstelt en R_f het rendement zonder risico. Met de R_f wordt meestal de rente van een staatsobligatie bedoeld, omdat dat een rendement zonder risico is. Door ze van elkaar af te trekken, kom je op het totale deel van het rendement mét risico. Een portfolio kan dan wel 3% rendement hebben, maar dat stelt niks voor, omdat dat kapitaal ook in een staatsobligatie had kunnen zitten, waarbij de investeerder geen risico had hoeven te nemen. $\sigma(R_p)$ staat voor de variantie, of ook wel volatiliteit genoemd, van een asset of verzameling van assets binnen een portfolio. Onder dit begrip wordt de mate van schommelingen bedoeld binnen het portfolio P . Logisch is dat als een portfolio een hoge variantie heeft, dat dat minder gunstig is voor een hedgefonds. Dat komt omdat daarbij een grotere risico komt kijken en onzekerheid groter wordt. De kans op een (extreem) verlies is groter. Concluderend, er volgt uit de vergelijking dat een lagere variantie en hogere $E[R_p] - R_f$ voor een grotere sharpe-ratio zorgt. Een normale sharpe-ratio van bijvoorbeeld een pensioenfonds ligt rond de 0.4 (Ilmanen, 2011). Jim Simons' fonds heeft in sommige perioden een sharpe-ratio van 7 behaald (Zuckerman, 2019). Dat is uitzonderlijk, want daarbij moet dus het rendement extreem hoog zijn in combinatie met de lage variantie. Dat zijn twee factoren die moeilijk zijn te combineren. Het doel van dit onderzoek is niet perse om een hoog rendement te behalen. Het doel is om een

model te programmeren dat risico minimaliseert en kan concurreren met andere defensieve strategieën.

Met deze informatie kan een probleemstelling worden geformuleerd. Het is in de huidige markt steeds lastiger om de variantie van een verzameling activa te voorspellen, omdat technische innovatie heeft geleid tot een dynamische en complexe markt. Er is technische expertise nodig en de vraag naar nieuwe modellen die de markt begrijpen is enorm toegenomen. Door beroep te doen op bestaande modellen kan een nieuw wiskundig model worden opgesteld dat de huidige markt beter begrijpt. Dit soort modellen wiskundige modellen zijn onmisbaar. Oude modellen kunnen varianties niet goed voorspellen (Engle, 1982) en ontbreken daarin in hun diepgang.

1.3 Doelstellingen van het onderzoek

Op basis van de probleemstelling kan een doelstelling voor het onderzoek worden opgesteld. Dit onderzoek wil EGARCH-modellen combineren om de volatiliteit van een asset binnen het portfolio te voorspellen en vervolgens op basis van risk parity een portfolio samen te stellen. Het HMM-onderdeel zorgt daarbij voor een nuance in de samenstelling, omdat dit kern-onderdeel van het model zich bezighoudt met het indelen van marktregimes. Het EGARCH-onderdeel houdt zich bezig met de volatiliteit, het onderste deel van de breuk van de sharpe-ratio vergelijking. Het doel van het onderzoek is om een sharpe-ratio groter of gelijk aan 0.4 te behalen. Om gericht onderzoek te kunnen doen, zijn er voorwaarden opgesteld.

Er is sprake van een eerlijke en bias-vrije opzet (1). Deze voorwaarde gaat eigenlijk over dat er niet vals gespeeld mag worden. Er mogen alleen dingen gebruikt worden die op dat moment bekend waren. Er wordt niet in de toekomst gekeken bij bijvoorbeeld backtesting. Er wordt alleen oude data gebruikt om het model te trainen en daarna andere, nieuwe data om te kijken of het model het goed doet. Ook zijn er consistente en realistische model- en portefeuillekeuzes (2). Er wordt duidelijk aangegeven hoe het model werkt en het wordt niet steeds veranderd om mooiere resultaten te krijgen. Een aandeel mag niet alles zijn, er wordt niet oneindig veel geld geleend en er wordt niet elke seconde gehandeld. Dat zorgt voor grotere transactiekosten en maakt het model complexer. Handelen per dag of maand zorgt voor een beter overzicht. Er gelden duidelijke grenzen voor gewichten, leverage en turnover. Resultaten worden gerapporteerd en transactiekosten worden meegenomen in het totale plaatje. De resultaten worden vergeleken met andere strategieën.

Zolang aan deze voorwaarden wordt voldaan, kan er een betrouwbare en valide conclusie worden gedaan. De resultaten worden op die manier verwerkt en vergeleken met andere modellen. Bij een positief resultaat kan het model ook in de praktijk worden gebruikt om een bijdrage te leveren aan de technische ontwikkelingen.

1.4 Onderzoeksvragen en hypotheses

Nu de probleem- en doelstelling bekend zijn, kan om doelgericht onderzoek te doen een hoofdvraag worden geformuleerd. De hoofdvraag wordt op de volgende manier geformuleerd: *“In welke mate presteert een risico-minimaliserende beleggingsstrategie die zich aanpast aan veranderende marktvolatiliteit beter of slechter dan andere defensieve beleggingsstrategieën?”* Deze

hoofdvraag gaat in op drie elementen. Allereerst onderzoekt het of het model überhaupt praktisch gezien lucratief kan zijn. Ook kan op deze manier worden vergeleken met bestaande modellen en houdt het rekening met de drie voorwaarden, zoals geformuleerd in paragraaf 1.3. Het is belangrijk dat deze drie kernelementen terugkomen in het onderzoek.

Op basis van deze hoofdvraag kunnen ook deelvragen worden geformuleerd om binnen het onderzoek een richting te vinden. Er is gekozen om de onderstaande selectie van deelvragen voor het onderzoek mee te nemen. Door de deelvragen uit te werken kan de hoofdvraag worden beantwoord.

1. *Welke theoretische principes liggen ten grondslag aan risico-minimaliserende en andere defensieve beleggingsstrategieën?*
2. *Hoe kan een strategie die zich aanpast aan marktvolatiliteit wiskundig worden gemodelleerd en welke aannames zijn daarvoor nodig?*
3. *Hoe verhouden het rendement, het risico en het risicogecorrigeerd rendement van de volatiliteitsafhankelijke strategie zich tot die van andere defensieve strategieën in verschillende marktregimes?*

Voor deze drie deelvragen is gekozen omdat ze logisch aansluiten bij de opbouw en het doel van het onderzoek. De eerste deelvraag richt zich op de theoretische basis: voordat de prestaties van verschillende defensieve strategieën kunnen worden vergeleken, moet duidelijk zijn welke concepten uit portefeuilletheorie, volatiliteitsdynamiek en risicominimaliserende beleggingsprincipes relevant zijn. De tweede deelvraag richt zich vervolgens op de modellering: om een volatiliteitsafhankelijke strategie te kunnen analyseren, moet worden vastgesteld hoe deze strategie wiskundig kan worden vormgegeven en welke aannames het model valide maken. De derde deelvraag maakt tenslotte de vertaalslag naar de praktijk door te onderzoeken hoe het rendement, het risico en het risicogecorrigeerde rendement van deze strategie zich verhouden tot die van andere defensieve strategieën in uiteenlopende marktregimes. Zo vormen de deelvragen samen een coherente keten van theorie naar modellering en uiteindelijk naar een empirische prestatieanalyse. De beantwoording hiervan leidt tot een betrouwbaar en valide onderzoeksresultaat.

De hypothese (H1) is dat een risico-minimaliserende beleggingsstrategie die zich dynamisch aanpast aan veranderende marktvolatiliteit een hoger risicogecorrigeerd rendement behaalt (bijvoorbeeld gemeten met de Sharpe-ratio) dan traditionele defensieve strategieën, zoals buy-and-hold, minimum volatility of een gelijkgewogen portefeuille, na aftrek van transactiekosten. De nulhypothese (H0) veronderstelt dat er geen significant verschil bestaat in risicogecorrigeerd rendement tussen de volatiliteitsafhankelijke strategie en deze benchmarkstrategieën. Deze nulhypothese is een standaarduitgangspunt en dient weerlegd te worden. In eerste instantie wordt aangenomen dat er geen effect of verschil is; pas wanneer de empirische resultaten voldoende statistisch bewijs leveren, kan geconcludeerd worden dat H0 niet blijkt te kloppen en dat de alternatieve hypothese H1 aannemelijker is.

1.5 Opbouw van het profielwerkstuk

Het profielwerkstuk focust zich op het minimaliseren van het risico tijdens investeren. Het opgestelde model zal uiteindelijk bestaan uit een geïntegreerd Hidden Markov Model en een EGARCH-onderdeel dat de volatiliteit voorspelt. Deze

voorspelde volatiliteit wordt gebruikt voor risk parity. De theorieën achter dit model worden in het Theoretisch kader uitgewerkt. Allereerst wordt de theorie over markten en informatie besproken, worden HMM's en EGARCH-modellen onderzocht om uiteindelijk in het hoofdstuk Methodologie een eigen model te ontwikkelen. De resultaten van het model worden besproken en vergeleken met andere modellen in het hoofdstuk Discussie. Tot slot wordt ingegaan op de hypothese en nulhypothese en wordt een conclusie gegeven. In deze conclusie wordt antwoord gegeven op de hoofdvraag.

2 Theoretisch kader

In dit hoofdstuk wordt het theoretisch kader geschetst dat nodig is om het verdere onderzoek te kunnen begrijpen en onderbouwen. Het hoofdstuk begint met een toelichting over hoe markten bewegen op basis van informatie. Later wordt ingegaan op de defensieve methoden en kijken we naar de klassieke portefeuille-theorie van Markowitz, waarin de relatie tussen rendement, risico en diversificatie centraal staat. Vervolgens komen risk parity-strategieën, volatiliteit en verschillende risicomaten aan bod, evenals de efficiënte markthypothese en de vraag in hoeverre markten voorspelbaar zijn. Ten slotte wordt ingegaan op tijdsvariërende volatiliteitsmodellen, zoals ARCH/GARCH/EGARCH, en op het concept van marktre regimes met behulp van Hidden Markov Modellen. In dit onderdeel van het profielwerkstuk wordt ingegaan op deelvragen 1 en 2. Het theoretisch kader eindigt met een conclusie waarin antwoord wordt gegeven op deze deelvragen.

2.1 Inleiding tot defensieve beleggingsstrategieën

2.1.1 De effectieve markt hypothese

De basis van asset-allocation ligt bij het verspreiden van het eigendom van assets. Dat kan worden gedaan als er sprake is van bruikbare informatie. Op basis van deze informatie kunnen beslissingen voor het portfoliomanagement worden genomen. We gaan dan uit van een markt waarin alle beschikbare informatie volledig de markt reflecteert. Zo'n markt noemt men efficiënt. Dergelijke markten kunnen op twee manieren worden geanalyseerd: fundamenteel en technisch. Wanneer beleggers fundamenteel beleggen kijken ze naar fundamentele factoren in de analyses, waaronder de intrinsieke waarde van een bedrijf. Technische analyses focussen zich vooral op trends en statistiek. Ons onderzoek focust zich alleen op de technische kant en laat de fundamentele kant volledig buiten beschouwing. Daarover volgt geen verdere informatie.

In het artikel *Efficient Capital Markets: A Review of Theory and Empirical work* (Fama, 1970) gaat Eugene Fama in op de efficiënte markt hypothese (hierna: EMH). Hij begint zijn artikel door onderscheid te maken tussen verschillende soorten efficiënte markten¹. Zo kan er een zwakke vorm aanwezig zijn waarbij alleen historische data de markt reflecteert. Ook bestaat er een variant, de semi-sterke variant, waarbij de markt ook wordt beïnvloed door publiekelijke informatie, bijvoorbeeld door het nieuws. Tot slot bestaat er een sterke variant waarbij monopolistische instituten alle informatie hebben die volledig de markt beïnvloed, waaronder ook privé informatie. Hierbij bestaan geen factoren die toegevoegde waarde bieden voor de markt. Zo kan een markt in verschillende gradaties van efficiëntie zitten. Wanneer informatie volledig wordt verwerkt in markt, kan dit voor complicaties zorgen tijdens het investeren. Dat betekent dat informatie direct verwerkt wordt in de markt. Wanneer bijvoorbeeld nieuwe informatie bekend is, wordt dat direct verwerkt en kunnen investeerders te laat reageren. Dat zorgt voor uitdagingen binnen het portfoliomanagement, omdat markten op die manier heel dynamisch en vaak onvoorspelbaar zijn. Het is moeilijk om consistent bovengemiddelde rendementen te behalen ten op zichte van de markt.

¹Dit onderscheid van efficiënte markten werd voor het eerst ter sprake gebracht door Harry Roberts.

Voor dit onderzoek en het te ontwikkelen EGARCH–HMM–risk-paritymodel is de EMH om meerdere redenen relevant. Als markten (minstens) in zwakke of semi-sterke vorm efficiënt zijn, is het lastig om structureel extra rendement te behalen door simpelweg publieke informatie of historische koersen te gebruiken. In plaats daarvan verschuift de focus van het voorspellen van rendementen naar het modelleren en beheren van risico. Het in dit profielwerkstuk voorgestelde model probeert dan ook niet om systematisch misprijzingen te verwerken, maar om de tijdsvariërende volatiliteit en regimeverschuivingen in de markt zo goed mogelijk te kwantificeren. EGARCH-modellen sluiten aan bij het idee dat rendementen op korte termijn moeilijk voorspelbaar zijn, terwijl volatiliteit wél een zekere structuur vertoont. Hidden Markov Models vullen dit aan door expliciet rekening te houden met verschillende marktregimes (bijvoorbeeld “rustig” versus “crisis”), zonder te veronderstellen dat deze regimes triviaal te voorspellen zijn op basis van publiek beschikbare informatie. Binnen de logica van de EMH is het dus zinvol om een model te bouwen dat zich richt op een efficiëntere risicotoedeling (via risk parity), in plaats van op het najagen van bovengemiddelde rendementen, omdat juist risicobeheersing en drawdownvermindering realistische en haalbare doelstellingen blijven in (semi-)efficiënte markten.

2.1.2 Adaptieve markt hypothese en behavioral finance

Toch zegt de efficiënte markthypothese niet alles. Markten fluctueren vaak in efficiëntie en zijn dus niet constant efficiënt of inefficiënt. Dit idee sluit aan bij de *Adaptive Markets Hypothesis* (Lo, 2004) (hierna: AMH). In dit kader wordt de markt gezien als een adaptief systeem waarin beleggers zich gedragen als economische “soorten” die zich voortdurend aanpassen aan veranderende omstandigheden. In plaats van volledig rationele beleggers, zoals de EMH veronderstelt, gaat de AMH uit van begrensde rationaliteit, leerprocessen en competitieve selectie. Strategieën die in één periode succesvol zijn, kunnen in een andere periode hun werking verliezen, omdat andere beleggers zich aanpassen of omdat marktomstandigheden veranderen. Emoties en gedragsfouten spelen daarmee een structurele rol in prijsvorming en marktdynamiek, waardoor efficiëntie tijdsafhankelijk wordt in plaats van absoluut. Dat resulteert onder andere in stochastische koersbewegingen.

Als markten zich adaptief gedragen en de mate van efficiëntie over de tijd fluctueert, dan is het aannemelijk dat ook volatiliteit, correlaties en risicopremies regime-afhankelijk zijn. Dit sluit direct aan bij het gebruik van een Hidden Markov Model, waarin verondersteld wordt dat de markt zich in verschillende, niet-direct observeerbare toestanden kan bevinden (bijvoorbeeld een “rustig” regime en een “stress”- of “crisis”-regime). In combinatie met EGARCH maakt dit het mogelijk om niet alleen tijdsvariërende volatiliteit te modelleren, maar deze ook expliciet te koppelen aan wisselende marktcondities en gedragsmatige factoren, zoals angst en hebzucht van beleggers. Vanuit het perspectief van de AMH is een statisch risk-paritymodel onvoldoende, omdat het impliciet uitgaat van relatief stabiele en rationele marktomstandigheden. Een EGARCH–HMM-gebaseerde risk-paritybenadering sluit beter aan bij een wereld waarin markten zich aanpassen, beleggers niet volledig rationeel zijn en efficiëntie dynamisch en stochastisch is. Het model gaat mee in de doelstelling van het onderzoek.

2.1.3 Definitie van defensieve strategieën

Een defensieve strategie is een strategie waarbij men risico wil minimaliseren en de assets kiest die stabiel zijn en zichzelf over de jaren heen hebben bewezen. Defensief beleggen is een langetermijnstrategie die niet gericht is op het behalen van snelle koerswinsten. Een van de bekendste vertegenwoordigers van deze denkwijze is Warren Buffett, die het belang van kapitaalbescherming vaak benadrukt. Zijn beroemde uitspraak – dat het vermijden van verlies de hoogste prioriteit heeft – illustreert de kern van defensief beleggen: voorzichtig omgaan met risico's en duurzaam waarde opbouwen (Buffett, 2008). Een defensieve strategie wordt gekenmerkt door de volgende eigenschappen (Ilmanen, 2011):

1. Allereerst is het belangrijk dat het risico wordt geminimaliseerd. Op deze manier wordt de variantie kleiner en dus de sharpe-ratio groter. De focus ligt op minder volatiliteit en drawdowns. Rendement is ook belangrijk, alleen een hoger rendement mag het risico niet zodanig beïnvloeden dat het risico te groot wordt.
2. Stabiliteit boven groei. Het is belangrijk om de focus te blijven houden op de volatiliteit verminderen. Het rendement is minder belangrijk. Risico en rendement zijn in veel gevallen evenredig. Als de ene variabele omlaag gaat, gaat de andere variabele dat ook.
3. Allocatie naar relatief veilige activa. Het portefeuille bestaat voornamelijk uit lage-volatiliteit activa, zoals staatsobligaties en spaarrentes. Op deze manier neemt het totale portefeuille minder risico. Dat sluit aan bij kenmerk 1.
4. Deze strategieën presteren relatief beter tijdens perioden van hoge volatiliteit en bear markten. Als er bijvoorbeeld een crisis is dan presteert deze strategie beter, omdat de kans op een extreme drawdown kleiner is dan bij een andere soort strategie.

De bedoeling is dat in dit profielwerkstuk een defensief model wordt ontwikkeld waarbij deze kenmerken worden overgenomen. Om te kijken of deze strategie het beter doet dan vergelijkbare strategieën moeten resultaten uit onder andere de backtesting worden vergeleken.

2.1.4 Doelstellingen (kapitaalbehoud, risicobeperking en stabiliteit)

Kapitaalbehoud vormt de kern van defensief beleggen. Het doel is om het ingelegde vermogen te beschermen tegen grote waardedalingen, vooral tijdens perioden van marktstress (Bogle, 1999). Deze benadering gaat ervan uit dat het vermijden van grote verliezen belangrijker is dan het najagen van hoge rendementen. Door neerwaarts risico te beperken, blijft het groeipotentieel van het vermogen op lange termijn intact. Het gaat om risico-minimalisatie; dat is waar de echte winst zit.

Defensieve strategieën streven naar een lager totaalrisico door blootstelling aan volatiele en cyclische marktfactoren te minimaliseren. Dit wordt bereikt door te investeren in posities met een relatief stabiel risicoprofiel of door risicomangementtechnieken toe te passen, zoals diversificatie, volatiliteitsgewogen allocatie of risicobudgettering. Het beperken van risico leidt tot voorspelbaardere prestaties en verkleint de kans op extreme negatieve rendementen.

Naast bescherming van vermogen richten defensieve strategieën zich op het realiseren van stabiele en consistente rendementen over de lange horizon. Door schommelingen te temperen en portefeuilles robuuster te maken voor marktdynamiek, ontstaat een rendementspatroon met minder pieken en dalen. Deze stabiliteit maakt de strategie aantrekkelijk voor beleggers die waarde hechten aan voorspelbare vermogensgroei en minder gevoelig willen zijn voor onzekerheden.

2.1.5 Verschil tussen passieve, actieve en kwantitatieve defensieve strategieën

Wanneer je naar defensief beleggen kijkt, merk je al snel dat mensen dit op heel verschillende manieren aanpakken. Sommigen kiezen voor een passieve defensieve strategie (Bogle, 2007), vooral omdat het rust en overzicht geeft. Je volgt een vaste, stabiele index of portefeuille, vaak opgebouwd uit minder volatiele aandelen. Het mooie hieraan is dat je niet elke dag hoeft in te grijpen; de strategie blijft op de achtergrond gewoon werken. Voor veel beleggers voelt dit prettig en veilig: je hoeft geen ingewikkelde beslissingen te nemen, je maakt weinig kosten en je weet precies waar je aan toe bent. Het nadeel is dat je bij onverwachte marktbevingen soms niet snel genoeg kunt reageren, simpelweg omdat de strategie ontworpen is om juist níet voortdurend te veranderen.

Andere beleggers kiezen bewust voor een actieve defensieve aanpak, juist omdat ze graag zelf aan het stuur staan. Zij proberen op basis van ervaring, marktgevoel of economische inzichten risico's te vermijden voordat ze zich echt manifesteren. In onrustige tijden stappen ze bijvoorbeeld eerder uit cyclische aandelen, of zoeken ze naar bedrijven met sterke balansen en stabiele kasstromen. Deze aanpak vraagt meer aandacht en kennis, maar geeft ook een gevoel van controle: je kunt inspelen op wat er in de markt gebeurt. Het brengt echter ook menselijkheid met zich mee — in de vorm van fouten, emoties en soms verkeerde inschattingen. Daardoor kan een actieve strategie zowel heel goed uitpakken als tegenvallen.

Tot slot zijn er beleggers die vertrouwen op kwantitatieve defensieve strategieën, vaak juist omdat ze de menselijke factor willen minimaliseren. In plaats van intuïtieve beslissingen gebruiken ze modellen die op basis van data inschatten hoe risicovol de markt op dat moment is. Zulke strategieën passen hun risico razendsnel aan wanneer volatiliteit stijgt of wanneer markten van regime veranderen. Het voelt bijna alsof je een nauwkeurige, objectieve assistent hebt die voortdurend voor je rekt en analyseert. Tegelijkertijd blijft er een menselijke kant: iemand moet die modellen bedenken, controleren en bijsturen. En als een model verkeerde aannames bevat, kan dat net zo goed misgaan als een menselijke fout. Toch spreken deze strategieën veel beleggers aan omdat ze een moderne, systematische manier bieden om risico's in een steeds dynamischere markt te beheersen.

2.1.6 Defensieve strategieën: theoretisch overzicht

Defensieve strategieën draaien om rust, stabiliteit en het beschermen van vermogen. Niet iedereen wil constant risico nemen of de markt proberen te verslaan; veel beleggers vinden het juist prettig om te weten dat hun geld veilig staat en geleidelijk kan groeien. Daarom kiezen sommigen voor staatsobligaties of spaarrentes, waar de kans op verlies klein is en de waardeschommelingen beperkt blijven. Anderen zoeken iets meer rendement, maar nog steeds met een gecontroleerd risicoprofiel, bijvoorbeeld via investment grade obligaties of defensieve aandelen.

High yield obligaties leveren vaak wat meer op, maar nemen tegelijkertijd meer risico met zich mee. Deze mix van strategieën laat zien dat defensief beleggen niet één vaste vorm heeft: het gaat vooral om het vinden van een balans tussen veiligheid en groei, passend bij de belegger zelf. Voor het onderzoek is het goed om vergelijkingsmateriaal te hebben. Op die manier kan de hoofdvraag beter beantwoord worden.

Tabel 1: Voorbeeld prestaties van defensieve strategieën

Strategie	Rendement	Sharpe	Vol.
Spaarrente	0-1.5%	0.40	0.2%
Staatsobligaties	2.8%	0.55	4%
Investment Grade	4.2%	0.60	6%
High Yield	6.5%	0.45	12%
Defensieve Aandelen	7.0%	0.65	10%

2.2 Fundamenten van portefeuilletheorie

2.2.1 Traditionele portfolio-optimalisatie (Markowitz, mean-variance)

De EMH en AMH beschrijven vooral hoe prijzen tot stand komen en in welke mate informatie en gedrag van beleggers al in koersen zijn verdisconteerd. Voor een belegger die vandaag een portefeuille moet samenstellen, is echter vooral de vraag relevant welke combinatie van assets past bij een gewenst risico- en rendementsprofiel. In de praktijk komt men dan al snel uit bij de klassieke mean-variance-benadering van Markowitz. Deze theorie gaat uit van verwachte rendementen, varianties en covarianties als input voor de portefeuillesamenstelling en levert een efficiënte grens op waaruit een belegger kan kiezen. Juist hier ontstaat een spanningsveld met de eerder besproken marktbeelden: als volatiliteiten en correlaties door regimes en gedrag sterk in de tijd veranderen, dan zijn de statische aannames van Markowitz problematisch. Dit vormt een directe aanleiding om verder te kijken dan het klassieke Markowitz-model en meer dynamische risicomodellen, zoals GARCH en HMM, te onderzoeken binnen de context van portefeuilleselectie.

In een van Markowitz' belangrijkste artikelen over de mean-variance-theorie (Markowitz, 1952) begon hij met het onderverdelen van verschillende fasen in het proces van het kiezen van een portfolio. Zo bestaat dit proces volgens Markowitz uit twee fasen. De eerste fase gaat volgens hem over het observeren van een asset en het vormen van een bepaald geloof over deze asset in de toekomst. De tweede fase begint bij dat geloof over de asset in de toekomst en leidt tot het daadwerkelijk kiezen van een portfolio. Zijn onderzoek richt zich voornamelijk op deze tweede fase: het opstellen van een portfolio. Het gaat hier dus over de asset-allocatie. Daarnaast worden bepaalde voorwaarden besproken voor het onderzoek. Zo wordt genoemd dat een belegger zou moeten proberen het verwachte rendement te maximaliseren, ongeacht risico of spreiding. Markowitz wijst deze regel echter vrijwel direct af, omdat dit in strijd is met zijn hypothese om beleggingsgedrag te verklaren. Onderzoek laat namelijk zien dat beleggers zich niet alleen laten leiden door het hoogste rendement, maar ook rekening houden met risico en spreiding. Ook wordt een andere voorwaarde besproken: een hoger rendement is wenselijk en een hogere variantie is onwenselijk. Hij duidt dus op een voorwaarde die niet uitsluitend gefocust is op hoog rendement. Deze voorwaarde is veel realistischer en sluit goed aan op de werkelijkheid. Bovendien gaat het hand in hand met de

doelstelling van het profielwerkstuk.

Ook gaat hij later nog een keer in op de eerste, verworpen voorwaarde. Zo zegt hij dat het maximaliseren van het verdiscoteerde verwachte rendement geen diversificatie oplevert. *Gedisconteerd* impliceert het verwachte rendement, omgerekend naar vandaag. Dus een asset kan dan wel 110\$ waard zijn in de toekomst, maar omgerekend naar vandaag slechts 100\$. Hij geeft aan dat een investeerder al zijn geld zal investeren in de asset met de hoogste gediscoteerde waarde. Wanneer er twee assets zijn met de hoogste gediscoteerde waarde, dan zal het vermogen verdeeld worden, en is elke verdeling goed. Hij onderbouwt dat door uit te gaan van een portfolio met N effecten. Hij stelt een verwacht rendement r_{it} voor in tijd t met een asset i ; ook bestaat er een disconteringsfactor d_{it} die het rendement terugreken naar een waarde voor vandaag; en tot slot bestaat een relatieve bijdrage X_i van de asset die de investeerder stopt in het portfolio. Omdat er geen short sales zijn toegestaan, geldt voor alle posities dat $X_i \geq 0$. Hierbij staat X_i voor het (relatieve) bedrag dat in asset i wordt geïnvesteerd. Het totale gediscoteerde verwachte rendement van de portefeuille kan dan als volgt worden geformuleerd:

$$R = \sum_{t=1}^{\infty} \sum_{i=1}^N d_{it} r_{it} X_i = \sum_{i=1}^N X_i \left(\sum_{t=1}^{\infty} d_{it} r_{it} \right).$$

In deze uitdrukking staat r_{it} voor het verwachte rendement van asset i in periode t , en d_{it} voor de bijbehorende disconteringsfactor die het rendement in periode t terugreken naar tijdstip 0. De binnenste som $\sum_{t=1}^{\infty} d_{it} r_{it}$ berekent dus de contante waarde van alle toekomstige verwachte rendementen van asset i . De buitenste som over i telt vervolgens de bijdragen van alle assets bij elkaar op, gewogen met hun portefeuillegewicht X_i . Om dit compacter te noteren, wordt vaak eerst het gediscoteerde verwachte rendement per individuele asset gedefinieerd als

$$R_i = \sum_{t=1}^{\infty} d_{it} r_{it}$$

Dit geeft ons weer nieuwe informatie. R_i is onafhankelijk van X_i . Ook is bekend dat de som van alle X_i gelijk is aan 1 (namelijk 100%) en dat R een gewogen gemiddelde is van R_i met $X_i \geq 0$. Om R te maximaliseren, stellen we $X_i = 1$ met i een maximum R_i . Als verschillende R_a , met $a = 1, \dots, K$ maxima zijn, dan zal elke allocatie in het portfolio R maximaliseren. Dat betekent dat als er meerdere assets de hoogste gediscoteerde waarde hebben, dan maakt het niet uit hoe de spreiding is. Dat impliceert dat een gespreid portfolio in geen situatie meer gewaardeerd zou moeten zijn dan een niet-gespreid portfolio. Deze situatie wordt op de volgende manier geformuleerd:

$$R_{\max} = \max_X \sum_{i=1}^N X_i R_i \quad \text{zodat} \quad \sum_{i=1}^N X_i = 1, \quad X_i \geq 0,$$

$$\Rightarrow R_{\max} = \max_{1 \leq i \leq N} R_i \quad \text{bij} \quad X_i = 1 \text{ voor een } i \text{ met } R_i = \max_j R_j.$$

Uit het voorgaande blijkt dat, als je alleen kijkt naar welk aandeel het hoogste verwachte rendement heeft, je in theorie al je geld in dat ene aandeel zou stoppen. In zo'n benadering wordt een gespreide portefeuille dus niet beloond: één asset

met het hoogste verwachte rendement wint altijd. Voor dit profielwerkstuk is dat belangrijk, omdat het laat zien dat we niet alleen naar rendement kunnen kijken. Het GARCH-model gaat juist over risico en schommelingen in de tijd (volatiliteit). Wil je GARCH echt gebruiken bij het kiezen van een portefeuille, dan moet risico dus een rol spelen in de beslissing. Pas als we naast rendement ook rekening houden met risico en spreiding, wordt een gespreide portefeuille aantrekkelijk en heeft het zin om een GARCH-model te gebruiken om de portefeuillegewichten te bepalen. Op die manier rijkt de sharpe-ratio zo hoog mogelijk.

2.2.2 De wet van grote getallen en diversificatie

Toch kiezen investeerders vaak voor diversificatie en daar is ook een grondige reden voor, aldus Markowitz. Zo bestaat er een strategie die zowel diversificatie als maximalisatie van verwacht rendement aanbeveelt. Volgens deze regel verdeelt een belegger zijn vermogen over alle assetklassen die het hoogste verwachte rendement opleveren. De wet van grote getallen zou dan ervoor zorgen dat het werkelijke rendement van de portefeuille dicht bij het verwachte rendement in de buurt ligt. De regel sluit aan bij de expected returns-variance (hierna: EV) regel, die uitgaat van een portefeuille die zowel maximaal rendement als minimale variantie biedt. Markowitz geeft aan dat deze EV-regel niet geaccepteerd kan worden. Hij gaat daarbij in op de covariantie tussen de assets en dat diversificatie niet alle variantie kan elimineren. Beleggers kunnen op die manier een trade-off maken, door bijvoorbeeld een deel van het risico te accepteren voor hoger rendement. Ook kunnen beleggers minder risico accepteren en een deel van het rendement opgeven. Op deze manier wordt er een nuance gebracht in het EV model van Markowitz. Het doel van het profielwerkstuk is om zo min mogelijk risico te zoeken tegen een rendement dat aanzienlijk groter is dan R_f , zodat de sharpe-ratio een aangename waarde bereikt ($S \geq 0.4$). Markowitz leert ons dat activa gekozen moeten worden op basis van verwacht rendement en variantie en dat daarin een goede verhouding gevonden moet worden. Toch mist er diepgang in het artikel van Markowitz op de gebieden van dynamica van markten. De volatiliteit is niet statisch, maar dynamisch (Mandelbrot, 1963).

2.2.3 Risico en volatiliteit in financiële markten

Markowitz begint ook met een introductie van enkele elementaire begrippen uit de statistiek. Zo heeft hij het over een variabele Y die afhangt van een kans. Hij gaat uit dat Y allerlei waarden $y_1, y_2, \dots, y_N$ kan aannemen. Hij stelt de kans van $Y = y_n$ gelijk aan p_n . Het verwachte rendement kan dan worden berekend met de volgende formule:

$$E(Y) = \sum_{i=1}^N p_i y_i$$

Dit is intuïtief te begrijpen. Neem bijvoorbeeld een eerlijke dobbelsteen met de uitkomsten 1, 2, 3, 4, 5, 6. De verwachtingswaarde $E(Y)$ van één worp is dan het gewogen gemiddelde van alle mogelijke uitkomsten, waarbij iedere uitkomst een kans van $\frac{1}{6}$ heeft. We krijgen dan:

$$E(Y) = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} = 3,5.$$

Dit betekent niet dat de dobbelsteen ooit daadwerkelijk 3,5 zal laten zien, maar dat 3,5 het gemiddelde resultaat is dat we op lange termijn verwachten als we de dobbelsteen heel vaak zouden gooien. De verwachtingswaarde is dus een maat voor het gemiddelde gedrag van een willekeurige uitkomst. Precies hetzelfde principe kan worden toegepast op financiële markten. In plaats van dobbelsteenuitkomsten kijken we dan naar mogelijke rendementen van een asset, met bijbehorende kansen. De verwachtingswaarde vertegenwoordigt dan het verwachte rendement van een belegging, als gewogen gemiddelde van alle mogelijke scenario's.

Naast de verwachtingswaarde is het echter ook belangrijk om te weten hoe sterk de uitkomsten rond dit gemiddelde schommelen. Twee beleggingen kunnen hetzelfde verwachte rendement hebben, maar toch heel verschillend risico laten zien. Hiervoor gebruiken we de variantie: een maat voor hoe ver de uitkomsten gemiddeld van de verwachtingswaarde af liggen. Er bestaat een standaardformule voor de variantie. Deze formule kan op de volgende manier worden geformuleerd:

$$V(Y) = \sum_{i=1}^N p_i (y_i - E(Y))^2$$

Hierbij weten we dat het totale rendement van de portefeuille afhangt van de gewichten a_i en de individuele rendementen R_i van de verschillende assets. Elk gewicht a_i geeft aan welk deel van het totale vermogen in asset i wordt belegd. Hoe groter het gewicht van een bepaalde asset, hoe groter de bijdrage van het rendement van die asset aan het totale rendement van de portefeuille. Het verwachte rendement van de portefeuille kan daarom worden berekend op basis van de afzonderlijke verwachte rendementen van de assets. In plaats van het rendement van de hele portefeuille in één keer te benaderen, kijken we eerst naar het verwachte rendement van iedere asset afzonderlijk en nemen vervolgens een gewogen gemiddelde. Dit levert de bekende relatie op:

$$E(R_p) = \sum_{i=1}^N a_i E(R_i),$$

waarbij $E(R_p)$ het verwachte rendement van de portefeuille is, $E(R_i)$ het verwachte rendement van asset i en $\sum_{i=1}^N a_i = 1$ (het totale belegde vermogen is 100%). Ter illustratie: als 60% van het vermogen wordt belegd in een asset met een verwacht rendement van 5% en 40% in een asset met een verwacht rendement van 3%, dan is het verwachte portefeuillerendement gelijk aan

$$E(R_p) = 0,6 \cdot 0,05 + 0,4 \cdot 0,03,$$

In een portefeuille met meerdere assets is het niet genoeg om alleen naar de afzonderlijke varianties (risico's) van de rendementen te kijken (Markowitz, 1952). Minstens zo belangrijk is de vraag hoe deze rendementen zich samen bewegen. Dit gezamenlijke gedrag wordt vastgelegd in de covariantie tussen twee assets. Als $\text{Cov}(R_i, R_j) > 0$ dan hebben de twee assets de neiging om in dezelfde richting te bewegen. Als $\text{Cov}(R_i, R_j) < 0$ dan bewegen ze gemiddeld genomen juist in tegenovergestelde richting. Bij een covariantie dichtbij nul is er weinig lineair verband tussen de twee rendementen. Wanneer we met N verschillende assets werken, is het onhandig om alle covarianties één voor één op te schrijven. In plaats daarvan verzamelen we alle varianties en covarianties in één object: de

covariantiematrix (Bodie, Kane, Marcus, 2014). De covariantiematrix Σ wordt dan gedefinieerd als

$$\Sigma = \text{Cov}(R) = \begin{pmatrix} \text{Cov}(R_1, R_1) & \text{Cov}(R_1, R_2) & \cdots & \text{Cov}(R_1, R_N) \\ \text{Cov}(R_2, R_1) & \text{Cov}(R_2, R_2) & \cdots & \text{Cov}(R_2, R_N) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(R_N, R_1) & \text{Cov}(R_N, R_2) & \cdots & \text{Cov}(R_N, R_N) \end{pmatrix}.$$

Hierbij staan de elementen op de diagonaal, $\text{Cov}(R_i, R_i)$, gelijk aan de variantie van asset i : $\text{Cov}(R_i, R_i) = \text{Var}(R_i)$. De elementen buiten de diagonaal geven de covariantie tussen twee verschillende assets i en j weer. In portefeuilletheorie speelt de covariantiematrix een centrale rol. Stel dat de gewichten van de portefeuille in een vector $\mathbf{a}$ staan:

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix}, \quad \text{met} \quad \sum_{i=1}^N a_i = 1.$$

We zijn nu geïnteresseerd in het risico van de hele portefeuille. Dit risico drukken we uit in de variantie van de portefeuille. De gewichten van de portefeuille staan in de vector $\mathbf{a}$ en alle varianties en covarianties staan in de matrix Σ . Met deze notatie kan de variantie van het portefeuillerendement compact worden geschreven als

$$\text{Var}(R_p) = \mathbf{a}^\top \Sigma \mathbf{a}.$$

Deze formule laat zien dat het portefeuillerisico niet alleen afhangt van de individuele risico's (varianties) van de assets, maar ook van de onderlinge samenhang (covarianties). Juist door assets te combineren die niet perfect met elkaar meebewegen (lage of negatieve covariantie), kan het totale risico van de portefeuille lager worden dan de simpele som van de individuele risico's. Dit is de kern van diversificatie.

In modellen zoals dat van Markowitz, en ook in meer geavanceerde modellen met EGARCH en HMM, vormt deze matrix de basis voor het berekenen en beheersen van het portefeuillerisico. De EV-regel van Markowitz gaat ook in op andere - vooral complexe - situaties die niet relevant zijn en niet aansluiten op de doelstelling van dit onderzoek. De statistische hulpmiddelen die Markowitz aan de lezer geeft, sluiten echter aan bij de doelstelling van het begrip investeren en diversificatie. Deze informatie wordt dan ook meegenomen in het maken van het model, waarbij het model wordt vergeleken met andere modellen.

2.3 Reageren op veranderende volatiliteit

2.3.1 Dynamische volatiliteit voorspellen: ARCH/GARCH

Het portfolio wordt dus verdeeld op basis van risk parity met de covariantie en variantie. Door deze volatiliteit te voorspellen - voor bijvoorbeeld de volgende maand - kan het risico voor het portefeuille van de volgende maand worden geminimaliseerd. Dat gaat mee in de doelstelling van het onderzoek. Een manier om deze volatiliteit te voorspellen is aan de hand van ARCH- en GARCH-modellen. Dit onderdeel van het profielwerkstuk focust zich op deze modellen.

In veel klassieke modellen wordt verondersteld dat het rendement van een asset een constante variantie heeft, zoals in artikelen van Markowitz. In de praktijk is dit echter niet realistisch: markten kennen perioden van relatieve rust, afgewisseld met perioden van stress en hoge volatiliteit ((Mandelbrot, 1963). Dit verschijnsel staat bekend als *volatility clustering*: grote koersschokken treden vaak in clusters op, net als rustige dagen. Om dit gedrag beter te kunnen beschrijven, zijn zogenoemde ARCH- en GARCH-modellen ontwikkeld. Deze modellen vormen het fundament van dit onderzoek.

ARCH staat voor Autoregressive (hangt af van vorige perioden), Conditional (op basis van het verleden) en Heteroskedasticity (variantie is niet constant). Oftewel: een model waarin voorwaardelijke variantie door de tijd heen verandert en afhangt van historische fouten. De ARCH-modellen zijn voor het eerst geïntroduceerd door Robert F. Engle in een artikel uit 1982 (Engle, 1982), oorspronkelijk om de variantie van inflatie in het VK te modelleren. Daarna zijn allerlei uitbreidingen ontwikkeld, waarvan GARCH (Bollerslev, 1986) de bekendste is. Bij het ARCH-model beschouwen we een dagelijkse reeks van rendementen r_t , die als volgt kunnen worden geformuleerd:

$$r_t = \mu + \varepsilon_t,$$

waarbij μ het (eventueel kleine) gemiddelde rendement weergeeft en ε_t de onverwachte schok op tijdstip t is. In plaats van een constante variantie aan te nemen, veronderstelt een GARCH-model dat de variantie van ε_t in de tijd varieert. Het gemiddelde van deze variantie is nul en het heeft een standaard uitwijking binnen σ_t^2 . De foutterm is dus stochastisch verdeeld en niet deterministisch:

$$\varepsilon_t \sim (0, \sigma_t^2),$$

Hierbij is σ_t^2 dus de voorwaardelijke variantie is, gegeven alle beschikbare informatie tot en met tijdstip $t - 1$. Er wordt dus alleen gekeken naar tijdstip $t - 1$ om een besluit mee te nemen. Het meest gebruikte model in de praktijk is het GARCH(1,1)-model. Dit GARCH-model specificeert de voorwaardelijke variantie met $t - 1$ als

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2,$$

met de volgende voorwaarden: $\omega > 0$, $\alpha \geq 0$, $\beta \geq 0$ en $\alpha + \beta < 1$. Hierbij is ω het langetermijngemiddelde niveau van de volatiliteit; α meet hoe sterk de volatiliteit reageert op nieuwe schokken in het rendement (de impact van ε_{t-1}^2); β meet hoe persistent de volatiliteit is: hoe groot is de invloed van de vorige volatiliteit σ_{t-1}^2 . Als $\alpha + \beta$ dicht bij 1 ligt, dan keert de volatiliteit slechts langzaam terug naar haar langetermijngemiddelde na een schok. Dit sluit goed aan bij empirische observaties op financiële markten, waar perioden van hoge volatiliteit vaak langer aanhouden (Bollerslev, 1986). Door de rest van de gegevens in de vergelijking te verwerken komt met op de volatiliteit op tijdstip t . Dat is dan een schatting van de volatiliteit. Door risk parity toe te passen komt met op een portefeuille waarvoor elke asset X_i een ERC geldt.

Om duidelijk te maken hoe een GARCH(1,1)-model in de praktijk werkt, nemen we een eenvoudig voorbeeld dat lijkt op de dagelijkse volatiliteit van een brede aandelenindex, zoals de S&P 500 of de MSCI World-index. Stel dat de parameters als volgt zijn:

$$\omega = 0,000002, \quad \alpha = 0,08, \quad \beta = 0,90.$$

Hieruit volgt dat

$$\alpha + \beta = 0,98,$$

wat betekent dat de volatiliteit zeer persistent is. Dit is een typische uitkomst in empirische studies van aandelenrendementen. Veronderstel bovendien dat de markt zich in een rustige periode bevindt, met een dagelijkse standaarddeviatie (volatiliteit) van ongeveer 1%. Dat betekent dat

$$\sigma_{t-1} = 1\% \Rightarrow \sigma_{t-1}^2 = 0,01^2 = 0,0001.$$

Stel dat de onverwachte schok in het rendement op dag $t-1$ klein was, bijvoorbeeld 0,5%, dus $\varepsilon_{t-1} = 0,5\% = 0,005$. De volatiliteit op tijdstip t volgt dan uit het GARCH(1,1)-model:

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 = 0,000002 + 0,000002 + 0,00009 = 0,000094.$$

De bijbehorende standaarddeviatie of volatiliteit is dan:

$$\sigma_t = \sqrt{0,000094} \approx 0,97\%.$$

Het antwoord verschilt dus niet extreem veel vergeleken met de dag ervoor. Dat is logisch, want de parameter α is klein en β is groot. De volatiliteit van tijdstip t hoort dus niet erg te verschillen ten op zichte van $t-1$. In het model kan een Python code worden opgesteld dat deze berekening voor elke maand doet en vervolgens op basis van risk parity w_i herverdeeld.

2.3.2 Alternatief: EGRACH (1,1)

Door een overzicht te krijgen van de volatiliteit binnen het portefeuille krijgt men een goed beeld van de situatie. Sommige activa hebben een groter risico en krijgen minder bijdrage binnen het portefeuille (met als doel om totale risico kleiner te maken). De volatiliteit per activum is niet constant maar dynamisch; de volatiliteit verandert. In financiële markten zien we vaak het *leverage effect*: negatieve rendementen resulteren vaak in een grotere variantie en positieve rendementen zorgen voor een lagere variantie (Black, 1976). GARCH-modellen kunnen dat asymmetrische gedrag niet goed vangen. GARCH reageert namelijk symmetrisch op schokken; positieve en negatieve schokken van dezelfde grootte hebben hetzelfde effect op de volatiliteit. Zoals bekend, kan het standaard GARCH (1,1) model op de volgende manier worden geformuleerd:

$$\sigma_t^2 = w + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2$$

Hieruit komt deze theorie van symmetrie duidelijk naar voren:

$$\sigma_t^2(+5\%) = \sigma_t^2(-5\%)$$

Een positieve schok heeft een even groot effect als een negatieve schok. Het enige wat duidelijk naar voren komt, is hoe groot de schok is; niet welke invloed een negatieve schok ten op zichte van positieve schok heeft. In financiële markten is het echter wel belangrijk om daar nog onderscheid tussen te maken.

Een EGARCH model neemt deze asymmetrie echter wel mee (Nelson, 1991). EGARCH staat voor exponentiële GARCH. Het is een variant van het GARCH-model. In plaats van σ_t^2 zelf modelleer je de logaritme van de variantie:

$$\log(\sigma_t^2).$$

Hierdoor kan de logaritme elke waarde aannemen, maar door het exponent zal σ_t^2 een positieve waarde aannemen. Er hoeven geen waarden als $\alpha \geq 0$ te worden afgedwongen. EGARCH houdt ook rekening met het feit dat negatieve schokken meer effect hebben op de volatiliteit dan positieve. Voor de genormaliseerde schok (Nelson, 1991) geldt:

$$z_t = \frac{\varepsilon_t}{\sigma_t}$$

Dat is de fout gedeeld door de huidige volatiliteit. Als $z_t \sim N(0, 1)$, heeft hij gemiddelde 0 en variantie 1. Een veelgebruikte vorm van EGARCH (1,1) is:

$$\log(\sigma_t^2) = w + \beta \log(\sigma_{t-1}^2) + \alpha(|z_{t-1}| - E|z_{t-1}|) + \gamma z_{t-1}.$$

Het logaritme heeft natuurlijk een bereik van $[0, \infty]$ en daardoor is σ_t^2 altijd positief, zonder beperkingen zoals $\alpha, \beta \geq 0$. De term $\alpha(|z_{t-1}| - E|z_{t-1}|)$ modelleert hoe de grootte van de schok de volatiliteit beïnvloedt. Als $|z_{t-1}|$ groter dan gemiddeld is, dan neemt $\log(\sigma_t^2)$ toe. Dan neemt de volatiliteit dus ook toe. $E|z_{t-1}|$ is een constante; over het algemeen is dat $\sqrt{2/\pi}$ bij een normaal verdeelde z_t . De formule kan in dit geval eenvoudiger worden opgeschreven:

$$\log(\sigma_t^2) = w + \beta \log(\sigma_{t-1}^2) + \alpha(|z_{t-1}| - \sqrt{2/\pi}) + \gamma z_{t-1}.$$

γz_{t-1} zorgt voor de asymmetrie in het model. Als $\gamma < 0$, dan hebben negatieve schokken een groter effect op de volatiliteit dan positieve schokken van dezelfde grootte; hierbij komt dus het leverage effect kijken. Positieve schokken hebben dan een kleiner effect op de volatiliteit dan negatieve. Tot slot komt ook nog de $\beta \log(\sigma_{t-1}^2)$ in de formule voor. β bepaalt hoeveel de volatiliteit afhangt van de vorige schommelingen. Hoe dichter $|\beta|$ bij 1 komt, hoe langzamer de volatiliteit terugkeert naar normaal niveau. Er gaat dan tijd overheen om weer terug te keren naar de oorspronkelijke volatiliteit.

Op deze manier zorgen de extra parameters voor een nuance in het normale GARCH(1,1)-model. Juist voor ons model is een geïntegreerd EGARCH-model beter dan een GARCH-model, omdat juist de nuance in de gevolgen van schokken zo belangrijk is. Door deze vergelijking met historische data in een Python code te verwerken kan de volatiliteit worden voorspeld. Schatting gebeurt meestal via maximale likelihood; software doet dit proces, maar het is zwaarder dan een simpele AR-schatting. Over de maximale likelihood volgt in de volgende paragraaf meer.

2.3.3 Schatting via maximale likelihood

Voor ons model kan een GARCH- of -EGARCH model worden gebruikt. In allebei de modellen zitten parameters die zo goed mogelijk geschat moeten worden. Deze schattingen worden gedaan via software (Python), maar toch is het belangrijk om de theorie achter deze berekeningen te begrijpen. Een MLE is een *Maximum Likelihood Estimator*. Een MLE is een parameterwaarde die de geobserveerde data het meest waarschijnlijk maakt, gegeven een statistisch model (Hamilton, 1994). Het schatten van een EGARCH-model wordt gedaan met (quasi-)maximum likelihood (Bollerslev & Wooldridge, 1992). In deze paragraaf wordt de theorie kort toegelicht. Een diepe uitleg over deze MLE bij EGARCH's is onnodig, want Python kan met een econometrische bibliotheek deze parameters schatten en al het

moeilijke werk doen. Toch volgt hieronder een intuïtieve uitleg over deze schatting, zodat het model goed begrepen kan worden. Want dan kunnen resultaten beter worden geïnterpreteerd.

In dit onderzoek gebruiken we een EGARCH(1,1)-model om de tijdsvariërende volatiliteit van rendementen te beschrijven. Laat $\{r_t\}_{t=1}^T$ een reeks (log-)rendementen zijn. We nemen aan dat elk rendement kan worden geschreven als een gemiddeld rendement plus een onverwachte schok:

$$r_t = \mu + \varepsilon_t,$$

waar μ een constante is en ε_t de foutterm voorstelt. Deze foutterm wordt verder opgesplitst in een grootte en een willekeurige factor:

$$\varepsilon_t = \sigma_t z_t.$$

Hier staat σ_t voor de (onbekende) volatiliteit op tijdstip t en z_t is een gestandaardiseerde willekeurige term met gemiddelde nul en variantie één. Intuïtief kun je σ_t zien als de “grootte van de schok” en z_t als de richting en relatieve sterkte. Het kenmerkende van een EGARCH-model is dat niet de variantie zelf, maar de logaritme van de variantie wordt gemodelleerd. Dit is in het vorige hoofdstuk al toegelicht. Daardoor blijft de variantie automatisch positief. In een EGARCH(1,1)-model schrijven we:

$$\ln(\sigma_t^2) = \omega + \beta \ln(\sigma_{t-1}^2) + \alpha(|z_{t-1}| - E|z_{t-1}|) + \gamma z_{t-1},$$

waar ω , β , α en γ parameters zijn die we willen schatten. Deze vergelijking zegt in woorden dat de huidige volatiliteit afhangt van (1) het niveau in de vorige periode, (2) de grootte van de vorige schok, en (3) het teken van de vorige schok. Op die manier kan het model bijvoorbeeld asymmetrisch reageren op slechte en goede nieuwsberichten (grote negatieve rendementen kunnen een ander effect hebben dan grote positieve).

Omdat z_t gestandaardiseerd is, wordt vaak aangenomen dat z_t voorwaardelijk normaal verdeeld is met gemiddelde nul en variantie één. Gegeven alle informatie tot en met tijdstip $t - 1$ (het verleden) is de verdeling van r_t dan normaal met gemiddelde μ en variantie σ_t^2 . In woorden: als we het verleden kennen, dan weten we hoe de verdeling van het volgende rendement eruitziet, en die verdeling hangt af van de parameters en de eerder waargenomen schokken.

De parameters van het EGARCH-model worden geschat met maximum likelihood (MLE). Dit betekent grofweg dat we die waarden van de parameters kiezen die het meest waarschijnlijk maken wat we in de data hebben waargenomen. Formeel wordt hiervoor een log-likelihoodfunctie opgebouwd op basis van de veronderstelde verdeling van de foutterm en de EGARCH-structuur voor de volatiliteit. Deze log-likelihood wordt vervolgens numeriek gemaximaliseerd (Bollerslev Wooldridge, 1992).

In de praktijk voeren we deze stappen uit met behulp van de `arch`-bibliotheek in Python. De onderzoeker specificeert het type model (EGARCH) en de orde van het model (hier 1,1), waarna de software automatisch:

- de reeks voorwaardelijke varianties σ_t^2 construeert op basis van de EGARCH-vergelijking;
- de bijbehorende waarschijnlijkheid van de geobserveerde data berekent (de log-likelihood);

- een numerisch optimalisatie-algoritme gebruikt om de parameterwaarden te vinden die deze waarschijnlijkheid maximaal maken.

De uitkomst van dit proces zijn schattingen van de modelparameters, samen met standaardfouten en diagnostische maten die we gebruiken om de geschiktheid van het model te beoordelen. Op deze manier kan het EGARCH model worden geïntegreerd en gebruikt voor een passend resultaat.

2.3.4 Overzicht van risk parity-strategieën

Nu de volatiliteit kan worden voorspeld, kan een verdeling binnen het portefeuille worden gemaakt. Hoeveel kapitaal stoppen we in elk activum? Daarvoor kan risk parity worden gebruikt. Elk activum bedraagt dan evenveel risico. Er zijn verschillende soorten risk parity-strategieën. In deze paragraaf worden alle strategieën besproken en wordt afgewogen welke strategie het beste bij dit onderzoek past.

Het doel van klassieke risk parity is een gelijke risico bijdrage per asset binnen het portfolio (Maillard, Roncalli, Teïletche, 2010). Hierbij wordt meestal alleen een long-only portefeuille opgebouwd over enkele brede asset classes. Vaak bestaat er leverage op defensieve assets, zoals obligaties, om een doelvolatiliteit te behalen. Deze vorm van risk parity staat bekend als equal risk contribution (hierna: ERC): elk onderdeel van het portefeuille levert een even groot deel risico. Voor een portefeuille X met n assets geldt dan: $RC_1 = RC_2 = \dots = RC_n$. Het risico is gelijk verdeeld over het portefeuille, zodat de totale variantie van het portefeuille wordt geminimaliseerd. Met de formule voor deze totale variantie is de logica achter de klassieke risk parity theorie logisch af te leiden. Deze formule is kan worden opgeschreven als:

$$\sigma_P^2 = \sum_{i=1}^n \sum_{j=1}^n w_i w_j \sigma_{ij}$$

Uit de formule is af te leiden dat als σ_{ij} groter is, dat de σ_P^2 ook groter wordt. Om dit effect te compenseren moeten de bijdragen kleiner zijn. Op deze manier kan de theorie ook wiskundig worden afgeleid. De risk parity theorie vormt een goede basis voor risico-minimalisatie.

Deze theorie kijkt dus naar de variantie en covariantie. Een manier is om alleen te kijken naar de variantie; deze manier wordt naïef genoemd (De Carvalho, Xiao, Zilberman, 2012), omdat het uitsluitend de variantie meet, maar niet de covariantie. De kernregel waarbij $RC_1 = RC_2 = \dots = RC_n$ wordt geprobeerd te bereiken, alleen het zal nooit aan deze voorwaarde voldoen. Daarvoor moet ook de covariantie meegerekend worden, omdat assets met elkaar samenhangen. Voor deze naïeve variant geldt:

$$w_i = \frac{\frac{1}{\sigma_i}}{\sum_{j=1}^n \frac{1}{\sigma_j}}$$

De inverse van alle varianties binnen het portefeuille wordt bij elkaar opgeteld en wordt in de noemer gezet van de breuk. Bovenaan in de teller komt de volatiliteit van asset X_i . Een extreme variant hiervan, waarbij grote volatiliteiten meer

worden afgekeurd, wordt genoteerd in de volgende vorm:

$$w_i = \frac{\frac{1}{\sigma_i^2}}{\sum_{j=1}^n \frac{1}{\sigma_j^2}}$$

Zoals duidelijk te zien is door de tweede macht, worden aan grote varianties nog minder bijdrage geleverd en aan kleine varianties juist meer. In onzekere tijden zou dit model aannemelijker zijn, omdat een lage totale variantie dan gewild is. Dan is de kans op extreme drawdown kleiner.

Een aannemelijker model is een waarbij ook de covariantie in rekening wordt genomen. De voorwaarde voor risk parity met $RC_1 = RC_2 = \dots = RC_n$ wordt dan vaker bereikt dan bij vorige varianten. Er wordt een volledige covariantiematrix Σ meegenomen. Deze covariantiematrix komt overeen met die van hoofdstuk 2.3. Voor de gewichten geldt: $w = (w_1, \dots, w_n)^T$ en voor de portefeuille variantie: $\sigma_P = \sqrt{w^T \Sigma w}$. Voor de marginale risicobijdrage van asset i geldt (Roncalli, 2013):

$$MCR_i = \frac{\partial \sigma_P}{\partial w_i} = \frac{(\Sigma w)_i}{\sigma_P}$$

Marginale risicobijdrage is het risico dat je loopt als je een gewicht van een asset een klein beetje verhoogd. De vergelijking is een partieel differentiaalvergelijking waarbij de verandering van σ_P tegen de verandering in w_i gezet wordt. Voor de risicobijdrage van asset i geldt het volgende:

$$RC_i = w_i \cdot MCR_i = w_i \frac{(\Sigma w)_i}{\sigma_P}$$

Voor deze bijdrage van asset i zou dus een Python code gerund moeten worden, dat dit per asset berekent. Eerst moet de MCR_i worden geschat en daarna kan deze input worden ingevuld in de bovenstaande formule. w_i kan worden berekent aangezien RC_i moet voldoen aan de volgende voorwaarde van ERC voor risk parity: $RC_1 = RC_2 = \dots = RC_n$.

Deze covariantie variant voor de risk parity voldoet aan de ERC voorwaarde en sluit aan bij het doeleinde van het profielwerkstuk. Het doel is namelijk om de totale variantie van een portefeuille zo laag mogelijk te houden en door de covariantie mee te nemen zal dit doel beter bereikt worden dan wanneer alleen de variantie wordt meegenomen. Maar ook de niet-covariantie vorm voldoet aan de voorwaarde om het risico zo laag mogelijk te houden en is veel simpeler dan de covariantie vorm. De berekeningen zijn dan veel simpeler en sneller door de computer uit te voeren.

2.4 Aanpassingen aan asset-allocatie na GARCH en risk parity: HMM's

Het fundament is bekend: de volatiliteit kan voorspeld worden met een GARCH model en risico kan geminimaliseerd worden op basis van risk parity. Elke asset X_i krijgt een bepaalde bijdrage w_i binnen het portefeuille. Toch is het belangrijk om ook het verwachte rendement in rekening te brengen, want ook het verwachte rendement is onderdeel van de sharpe-ratio. Om $S \geq 1$ te krijgen moet ook hier naar gekeken worden. In dit hoofdstuk wordt gekeken naar HMM's en wordt uitgepluisd hoe deze modellen ons model kan verbeteren.

Stochastische processen zijn het tegenovergestelde van deterministische processen (Ross, 1996). Deterministisch betekent dat een uitkomst vrijwel altijd gelijk is wanneer er sprake is van een bepaalde set beginvoorwaarden. Oftewel: de situatie is niet zomaar willekeurig, maar heeft altijd een bepaald resultaat bij een specifieke oorzaak. Bij stochastiek gaat het echter over hevige onzekerheid en willekeurige fluctuaties. Marktbewegingen fungeren op deze stochastische manier. Daar zijn verschillende redenen voor. Zo zijn data, renteaanpassingen, geopolitieke spanningen, nieuws en onverwachte gebeurtenissen onvoorspelbaar. Toch is het belangrijk om geen totale willekeur te claimen, omdat stochastiek dan geen zin heeft. Markten zijn deels onvoorspelbaar maar bewegen vaak ook in vaste patronen. Menselijke keuzes zijn vaak irrationeel door oncontroleerbare en onvoorspelbare emoties, wat zorgt voor willekeur in de markt. Dat is terug te lezen in de AMH uit 2.1.1. Sommige bewegingen in de economie zijn te voorspellen, maar het resultaat van deze gebeurtenissen is niet met volledige zekerheid te voorspellen.

Een markt is dus stochastisch vanwege de toevallige, onzekere variabelen die een bepaalde kansverdeling volgen. Zo is er een bepaalde kans dat een aandeel omhoog of omlaag beweegt. De toestand van een activum staat centraal, waarbij een bepaalde variabele de waarschijnlijkheid van verschillende uitkomsten weergeeft. De toestand van een variabele op tijdstip t kan worden voorgesteld aan de hand van de stochastische variabele X_t (Ross, 1996). Een stochastisch proces kan worden geformuleerd als een collectie:

$$\{X_t \mid t \in T\},$$

waarbij de elementen stochastische variabelen zijn en T een random variabele is. Binnen zo'n collectie zijn drie begrippen belangrijk:

- **Uitkomstenruimte Ω :** dit is de verzameling van alle mogelijke uitkomsten van een stochastisch proces. Als voorbeeld kan een dobbelsteen worden gebruikt. De uitkomstenruimte voor een zeszijdige dobbelsteen is $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- **σ -algebra $\mathcal{F}$:** dit is de verzameling van gebeurtenissen waarvoor we een kans kunnen berekenen. Een gebeurtenis is een deelverzameling van Ω . Voor de dobbelsteen kan een gebeurtenis bijvoorbeeld zijn: “de uitkomst is hoger dan 2”. Bij stochastische processen gaat men ervan uit dat de mogelijke uitkomsten (gebeurtenissen) gedefinieerd zijn door de σ -algebra; deze liggen dus theoretisch vast.
- **Kansfunctie P :** aan elke gebeurtenis in de σ -algebra wordt een kans toegekend. Zo kan aan de gebeurtenis “bij een dobbelsteen is de uitkomst groter dan 2” de kans $P = \frac{4}{6}$ worden toegekend. Met deze notatie is het ook eenvoudig te begrijpen dat $P(\Omega) = 1$, want alle mogelijke uitkomsten samen hebben kans 100%.

Zoals hierboven beschreven, kan een stochastisch proces worden aangeduid als een collectie $\{X_t\}$ in een kansruimte $(\Omega, \mathcal{F}, P)$ met random variabelen over een set T . Dit is een formeel kader waarin drie elementen voorkomen, waarin berekeningen kunnen worden gedaan die met kansberekening te maken hebben. Daarnaast komt nog de meetbaarheid van S kijken. Dit betekent dat er sprake is van een ruimte S waarin het proces plaatsvindt met een σ -algebra, waardoor we gebeurtenissen

in S kunnen meten. Dit is een benadering voor stochastische processen en deze theorie is toepasbaar op financiële markten.

2.4.1 Regime-switching modellen en Markov Chains

Al deze informatie kunnen we meenemen en toepassen op het volgende soort stochastische proces dat vaak wordt gebruikt bij het modelleren van financiële markten: het Markov-proces (Ross, 1996). Het Markov-proces is een specifiek stochastisch proces waarbij de toekomstige toestand alleen afhangt van de huidige toestand en niet van het verleden. Oftewel:

$$P(X_{t+1} = x_{t+1} \mid X_t = x_t, X_{t-1} = x_{t-1}, \dots, X_0 = x_0) = P(X_{t+1} = x_{t+1} \mid X_t = x_t)$$

Aan de hand van deze theorie zou de waarde van een aandeel vandaag alleen uitmaken voor de waarde van het aandeel morgen. Daarbij is het belangrijk om bepaalde overgangskansen te kunnen berekenen, bijvoorbeeld tussen de dagen. Dit wordt formeel gegeven door:

$$P(s, t; x, y) = P(X_t = y \mid X_s = x),$$

waarbij tijdstip t later is dan tijdstip s . In feite wordt berekend wat de kans is dat X_t gelijk is aan y op tijdstip t , wanneer X_s gelijk is aan x op tijdstip s . Als voorbeeld nemen we het aandeel X op tijdstip s , met een waarde van €100. Met de formule wordt berekend hoe groot de kans is dat aandeel X op tijdstip t bijvoorbeeld €110 waard is. Daarvoor is historische data nodig.

Dit proces kent veel verschillende toepassingen in de markt. De Markov-keten staat dan ook als basis voor modellering binnen financiële markten. Strikt genomen zijn marktbevingen natuurlijk niet volledig Markov, omdat de koers in werkelijkheid ook afhangt van het verleden. Toch worden Markov-processen vaak gebruikt, omdat ze veel eenvoudiger te analyseren en te simuleren zijn. Bovendien is de Markov-benadering meestal op korte tijdsintervallen redelijk accuraat, omdat dan alleen hoeft te worden gekeken naar de huidige prijs van het aandeel. Voor ons model en onze doelstelling is dit Markov-proces goed toepasbaar en wordt het meegenomen in de Methode.

Een Markov-proces gaat uit van een overgangsmatrix A . Hierbij wordt gekeken naar verschillende overgangen en hoe groot de kans is dat deze overgangen daadwerkelijk plaatsvinden. Bijvoorbeeld, een overgang van “stijgende koers” naar “dalende koers” met een kans van 80% kan worden weergegeven als $A_{12} = 0.8$. Een overgangsmatrix ziet er typisch uit als volgt:

$$A = \begin{pmatrix} P_{11} & P_{12} & \dots & P_{1n} \\ P_{21} & P_{22} & \dots & P_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ P_{n1} & P_{n2} & \dots & P_{nn} \end{pmatrix},$$

waarbij P_{ij} de kans is van de overgang van toestand i naar toestand j . Bij een Markov-keten spreken we van een *als* de verdeling over de toestanden op lange termijn niet meer verandert. In het begin hangt de verdeling nog sterk af van de starttoestand, maar na veel stappen “vergeet” de keten waar hij begonnen is. Onder milde voorwaarden bestaat er dan een stationaire verdeling π waarvoor geldt dat

$$\pi = \pi P,$$

waarbij P de overgangsmatrix is. Dit betekent dat als de keten eenmaal volgens π verdeeld is, één extra stap van de Markov-keten de verdeling niet meer verandert. In de tijd groeit de verdeling van toestanden geleidelijk naar deze stationaire verdeling toe: dat is wat we bedoelen met het bereiken van de stationaire fase. Voor ons model kan het handig zijn om zo'n up-to-date overgangsmatrix en verdeling π te hebben om kleine aanpassingen nog te kunnen doen aan onze allocatie waar nodig is.

2.4.2 Verborgen regimes

In veel tijdreeksmodellen, bijvoorbeeld in de econometrie en finance, gaat men uit van het idee dat er verschillende *regimes* of toestanden bestaan, zoals een rustige markt met lage volatiliteit en een turbulente markt met hoge volatiliteit (Hamilton, 1989). Deze regimes zijn vaak niet direct observeerbaar. Men spreekt dan van *verborgen regimes*. We nemen aan dat er een verborgen toestandsvariabele S_t bestaat (bijvoorbeeld $S_t \in \{1, 2\}$ voor een rustig en een turbulent regime), terwijl we alleen een geobserveerde variabele Y_t zien, zoals een rendement of een koers.

Het verborgen toestandsproces S_t wordt meestal gemodelleerd als een Markov-keten met bepaalde overgangskansen. Dit betekent dat de kans om van het ene regime naar het andere te gaan alleen afhangt van het huidige regime, en niet van de volledige voorgeschiedenis. Gegeven de toestand S_t heeft de geobserveerde variabele Y_t vervolgens een regime-afhankelijke verdeling. Zo kan in regime 1 een lager variantieniveau en eventueel een ander verwachtingswaarde gelden dan in regime 2. Op die manier kan het model verschillende vormen van gedrag in de data opvangen, bijvoorbeeld perioden met lage volatiliteit afgewisseld met perioden met hoge volatiliteit.

Omdat de regimes verborgen zijn, weten we niet precies in welk regime we ons op een bepaald moment bevinden. Wel kunnen we, op basis van de geobserveerde gegevens $(Y_1, Y_2, \dots, Y_t)$, voor elk tijdstip een kans toekennen aan elk regime. Na een periode met relatief kleine schommelingen zal de kans op het “rustige” regime bijvoorbeeld groot zijn, terwijl na een reeks grote uitslagen de kans op het “turbulente” regime toeneemt. In de praktijk gebruikt men speciale filter- en smoother-methoden (zoals het bekende Hamilton-filter) om deze regimekansen door de tijd heen te schatten. Verborgen-regimemodellen zijn nuttig omdat ze expliciet rekening houden met structurele omslagen in gedrag. In plaats van één enkel model met vaste parameters te gebruiken, gaat men uit van een mix van toestanden met elk hun eigen dynamiek. Dit is vooral relevant in financiële markten, waar men duidelijk onderscheid kan maken tussen normale perioden en crisisperioden. Op basis van de geschatte regimekansen kan men het risicobeheer aanpassen: in een regime met hoge geschatte volatiliteit kan men bijvoorbeeld de risicoblootstelling verlagen, terwijl in een kalm regime meer risico genomen kan worden. Zo bieden verborgen-regimemodellen een flexibele en realistische manier om de tijdsafhankelijke structuur van risico en rendement te modelleren.

2.4.3 HMM voor financiële tijdreeksen

Een Hidden Markov Model (HMM) is een model dat heel goed past bij financiële tijdreeksen waarin verschillende regimes een rol spelen, maar waarin die regimes niet direct zichtbaar zijn (Rabiner, 1989). Het idee lijkt sterk op dat van verborgen regimes: er is een niet-geobserveerd proces dat aangeeft in welk toestand of regime

de markt zich bevindt, en een geobserveerd proces dat we daadwerkelijk meten, zoals dagelijkse rendementen.

In een HMM gaan we uit van twee processen. Ten eerste is er de verborgen toestandsvariabele S_t , die aangeeft in welk regime de markt op tijdstip t zit. Dit kan bijvoorbeeld een rustig regime zijn met lage volatiliteit en een onrustig regime met hoge volatiliteit. Men schrijft dan vaak $S_t \in \{1, 2, \dots, K\}$, waarbij elk getal voor een ander regime staat. Het proces (S_t) volgt een Markov-keten: de kans op het volgende regime hangt alleen af van het huidige regime. Deze overgangskansen worden samengevat in een overgangsmatrix P , zoals omschreven in hoofdstuk 2.6.1. Deze matrix geeft bondig de kans op elke overgang weer.

Ten tweede is er de geobserveerde grootheid Y_t , bijvoorbeeld het rendement van een aandelenindex op dag t . De verdeling van Y_t hangt af van het regime S_t waarin we ons op dat moment bevinden. In een rustig regime kan het rendement bijvoorbeeld worden gemodelleerd met een lagere variantie, terwijl in een turbulente fase de variantie hoger is. Op die manier krijg je voor elk regime een eigen set van parameters, zoals een gemiddelde en een variantie.

In de praktijk observeer je alleen de Y_t 's en niet de S_t 's. Het centrale probleem bij een HMM is daarom om op basis van de data schattingen te maken voor de overgangskansen tussen de regimes én voor de kans dat de markt zich op een bepaald moment in een bepaald regime bevindt. Hiervoor bestaan speciale algoritmen. Een bekend voorbeeld is het zogenoemde voorwaarts-achterwaarts-algoritme, waarmee de kansverdeling van S_t gegeven de hele reeks waarnemingen $(Y_1, \dots, Y_T)$ kan worden berekend (Rabiner, 1989). Voor financiële tijdreeksen is een HMM aantrekkelijk omdat het expliciet rekening houdt met verschillende marktfases, zonder dat je vooraf precies hoeft te weten wanneer zo'n fase begint of eindigt. Denk aan perioden van lage volatiliteit die worden afgewisseld met crisisperioden waarin de volatiliteit sterk toeneemt. Met een HMM kun je op elk tijdstip een kans toekennen aan elk regime, bijvoorbeeld: "er is nu 80% kans op een rustig regime en 20% kans op een turbulent regime". Dit maakt het model geschikt voor toepassingen in risicobeheer, zoals het aanpassen van de risicoblootstelling wanneer de kans op een ongunstig regime oploopt. Op die manier biedt een HMM een flexibele manier om de structuur van financiële tijdreeksen te beschrijven, waarbij zowel het geobserveerde gedrag als de verborgen regimes een rol spelen.

Op deze manier kan een HMM gebruikt worden om in een rustig regime de σ_t variant te pakken voor risk parity in plaats van σ_t^2 voor onrustige perioden. Ook kunnen Markov-processen gebruikt worden met in de vorm van een overgangsmatrix P om de kans op stijgen of dalen te berekenen. Op basis van deze informatie kunnen kleine aanpassingen worden gedaan aan de asset-allocatie. Deze allocatie was voorheen al herverdeeld op basis van risk parity met een GARCH-model.

2.5 Conclusie: antwoord op deelvraag 1 en 2

In het theoretisch kader zijn veel dingen naar voren gekomen. Het is nu duidelijk hoe markten reageren op informatie en dat deze in regimes van efficiënte en inefficiënte schommelen. Ook hebben we gekeken naar Markowitz' theorie over verdelingen binnen een portefeuille. Het is nu bovendien duidelijk hoe de volatiliteit voorspelt kan worden aan de hand van ARCH, GARCH en EGARCH modellen en hoe Markov ketens kunnen berekenen hoe groot de kans op stijgen en dalen is. We kunnen nu antwoord geven op de eerste twee deelvragen. Deze deelvragen waren als volgt geformuleerd:

1. *Welke theoretische principes liggen ten grondslag aan risico-minimaliserende en andere defensieve beleggingsstrategieën?*
2. *Hoe kan een strategie die zich aanpast aan marktvolatiliteit wiskundig worden gemodelleerd en welke aannames zijn daarvoor nodig?*

Ten eerste kijken we naar vraag 1. Defensieve strategieën worden gekenmerkt door hun "saaiheid". Het zijn vaak portefeuilles die bestaan uit activa die weinig rendement hebben maar wel heel stabiel zijn. Dat is terug te zien in de gemiddelde volatiliteit van deze activa, die meestal tussen de 4% en 10% schommelt. Er zijn zelfs gevallen waarbij er bijna geen risico is. Dat is bijvoorbeeld zo bij spaarrentes. In theorie kiest een defensieve belegger dan ook voor een portefeuille met activa die stabiel zijn en een laag risico hebben. Ze leveren liever een beetje rendement in voor een lager risico. Juist het vermijden van extreme drawdowns, in plaats van het streven naar hoger rendement, dat is voor hun de manier waar de winst te behalen valt.

Ten tweede wordt antwoord gegeven op vraag 2. Markten bewegen. De mate van schommelingen noemt men ook wel de volatiliteit of variantie van een bepaald activum. Voor een belegger die risico wil minimaliseren is het gunstiger om (meer) geld te stoppen in activa met een lage volatiliteit. Dat brengt namelijk minder risico met zich mee. De kans op een extreme drawdown is aanzienlijk kleiner dan bij een activum met een hoge volatiliteit. Om deze reden wordt risk parity toegepast. Een activum met een hoge volatiliteit krijgt minder bijdrage binnen het portefeuille dan een activum met een lage volatiliteit. Zo wordt het risico met activum gelijk gesteld. Risk parity kan worden gedaan op verschillende manier. De meest waarschijnlijke variant voor dit profielwerkstuk is de naïeve variant, omdat berekeningen met deze variant sneller gaan. Deze variant is bovendien nog steeds uitstekend. Door met een EGARCH model de volatiliteit te voorspellen kan met deze volatiliteiten risk parity worden toegepast. Dat verlaagt het totale risico van ons model. We gaan daarbij uit van een markt waarbij er regimes zijn waartussen de volatiliteit kan schommelen. Door de volatiliteit telkens te voorspellen wordt rekening gehouden met deze dynamiek. Als nuance, om toch nog een beetje rekening te houden met het rendement, kunnen markov ketens een goede toevoeging zijn. Op basis van kansen op stijgen of dalen kunnen nog kleine aanpassingen worden gedaan aan de asset-allocatie. Tot slot zorgt een HMM voor extra risicovermindering, omdat dit model kan berekenen of een asset in een rustig regime zit of in een crisis regime. Op basis van die uitkomst kan het model extra defensief worden of minder defensief. Op deze manier speelt het model in op de stochastische dynamiek van de markt. Op deze manier is de variantie in theorie kleiner gemaakt en het rendement groter; het doel is dat daarbij de sharpe-ratio stijgt met $S \geq 0.4$

3 Methode

Dit is het deel van het onderzoek waar het om draait: het opstellen van het model. Het doel van het model is om het risico zo klein mogelijk te maken tegen een gunstig rendement. Het is daarbij niet heel erg als er rendement wordt ingeleverd voor een kleiner risico (zie paragrafen over defensief investeren). Het doel is dan ook dat de backtesting van het model op historische data leidt tot een plot met weinig variantie en een stijgende lijn. Een sharpe-ratio van $S \geq 0.4$ wordt daarbij als doel beschouwd. De optimale situatie zou zijn om een EGARCH-HMM-risk-parity model te maken dat dit risico binnen het portefeuille minimaliseert. Op die manier zal de probleemstelling worden behandeld en aangepakt.

3.1 Onderzoeksopzet

In dit hoofdstuk wordt behandeld hoe de hoofdvraag beantwoord wordt. De probleemstelling is dat markten dynamisch zijn, volatiliteit en regimes wisselen, en klassieke methoden houden daar weinig rekening mee. De onderzoeksvraag was als volgt geformuleerd: ***In welke mate presteert een risico-minimaliserende beleggingsstrategie die zich aanpast aan veranderende marktvolatiliteit beter of slechter dan andere defensieve beleggingsstrategieën?*** Het doel is om een dynamisch EGARCH-HMM-risk-parity (EH-RP-)model te bouwen en te kijken of het risico gewogen rendement beter is dan andere defensieve strategieën.

Onderzoeksaanpak: kwantitatieve backtesting Dit onderzoek is vooral praktisch van aard: het kijkt niet naar wat een model in theorie zou moeten doen, maar naar hoe het zich in het verleden daadwerkelijk had gedragen. Dat gebeurt via backtesting, waarbij het model wordt teruggezet in de tijd om te zien hoe het had gepresteerd in verschillende marktomstandigheden. Om een eerlijk beeld te krijgen, wordt de strategie vergeleken met andere defensieve strategieën. In tabel 1 in het theoretisch overzicht van defensieve strategieën zijn de prestaties van deze strategieën terug te vinden. Het doel is dat ons model kan concurreren met de andere strategieën.

3.2 Dataselectie (markten, activa, periode, frequentie)

Voor het onderzoek is het belangrijk om concreet te omschrijven welke data gebruikt wordt en hoe deze gebruikt wordt. Voor de data van de assets uit het onderzoek gelden de volgende regels:

1. De assets die voor het kwantitatieve onderzoek zijn meegenomen zijn gekozen op basis van verschillen in volatiliteit en weinig correlatie hebben met elkaar (doel: $Cov \leq 0$). Deze selectie zorgt voor een mix van risicovolle en defensieve beleggingscategorieën, zodat het model getest kan worden in zowel aandelen- als rente- en grondstoffenmarkten. Er is gekozen voor de volgende selectie van activa: GLD, IEF, SPY, HYG, QQQ en TLT.
2. De historische data wordt via Python uit de `yfinance`-library (Yahoo Finance) gehaald (Aroussi, 2023). Dit zorgt ervoor dat de data op een gestandaardiseerde manier wordt ingelezen, en dat er eenvoudig meerdere tijdreeksen tegelijk kunnen worden opgehaald. Bovendien maakt dit het gehele proces reproduceerbaar: andere onderzoekers kunnen in principe dezelfde code gebruiken om dezelfde datasets opnieuw te downloaden. Voor dit profielwerkstuk zijn datasets van onder andere Yahoo Finance, FRED datasets,

maar ook Robert Schiller (Shiller, 2023), opgenomen om een grotere, historische dataset te verkrijgen. Dat heeft geresulteerd in een data set die terug gaat tot rond 1970.

3. De historische data start minimaal vanaf 1970; indien het mogelijk is om deze set eerder te starten, wordt dat gedaan. Op deze manier wordt een zo lang mogelijke tijdshorizon gebruikt, zodat het model getest kan worden over meerdere marktregimes (bijvoorbeeld inflatoire perioden, crises en langdurige bull-markten). Wanneer een specifieke asset pas later beschikbaar is, wordt de analyse voor die asset uitgevoerd vanaf de vroegst beschikbare datum.
4. Er is sprake van dagelijkse data. Dagelijkse observaties maken het mogelijk om volatiliteit en regimewisselingen op relatief korte termijn te detecteren, wat belangrijk is voor GARCH- en HMM-modellen.
5. Bewerkingen worden gedaan door de prijzen naar (log-)rendementen om te zetten. Log-rendementen hebben wiskundige voordelen: ze zijn beter optelbaar over de tijd en sluiten aan bij veel theoretische modellen in de financiële econometrie.

Afhankelijke variabelen in het model In dit onderdeel wordt beschreven wat er precies wordt gemodelleerd en welke variabelen en modellen daarin een rol spelen. Er wordt onderscheid gemaakt tussen afhankelijke variabelen, risicovariabelen en de toegepaste modellen. De belangrijkste grootheden die in dit onderzoek worden gemodelleerd zijn de rendementen van de individuele assets en het totale portefeuillerendement. Het onderzoek maakt gebruik van een combinatie van modellen, die elk een specifiek onderdeel van het gedrag van de financiële tijdreeksen beschrijven:

- Voor elke afzonderlijke asset wordt een EGARCH(1,1)-model gespecificeerd op de tijdreeks van rendementen $r_{i,t}$. Dit model levert een tijdsvariërende inschatting van de volatiliteit $\hat{\sigma}_{i,t}$, waarbij onder meer rekening wordt gehouden met asymmetrische reacties op positieve en negatieve schokken (leverage effect). Zie voor een verdere toelichting het hoofdstuk Alternatief: EGARCH (1,1).
- Een Hidden Markov Model met twee regimes (laag- en hoogvolatiel) wordt toegepast op een gekozen referentiereeks (bijvoorbeeld de S&P 500-index of het portefeuillerendement). Het HMM levert door de tijd heen de kansverdeling over de regimes, zodat op elk tijdstip een inschatting gemaakt kan worden van de kans op een rustig versus een stressregime.
- De voortekens van de rendementen (UP voor positief, DOWN voor negatief) worden gebruikt om een eenvoudige Markov-keten te schatten. Dit resulteert in een overgangsmatrix met kansen zoals $P(\text{UP} \rightarrow \text{UP})$ en $P(\text{DOWN} \rightarrow \text{DOWN})$. Deze informatie wordt gebruikt als een extra, lichte richtingindicator naast het puur risicogedreven onderdeel.
- De basisgewichten in de portefeuille worden bepaald volgens een risk parity-logica, waarbij assets met een hogere geschatte volatiliteit een lager gewicht

krijgen. In de eenvoudigste vorm geldt dat de risk parity-gewichten evenredig zijn aan de inverse van de geschatte volatiliteit:

$$w_{i,t}^{\text{RP}} \propto \frac{1}{\hat{\sigma}_{i,t}}.$$

Na normalisatie zodat $\sum_{i=1}^N w_{i,t}^{\text{RP}} = 1$ ontstaat een portefeuille waarin de risicobijdragen meer in balans zijn. Vervolgens wordt een kleine correctie op deze basisgewichten aangebracht op basis van de UP/DOWN-kansen uit het Markov-model (bijvoorbeeld door $w_{i,t}^{\text{RP}}$ licht te verhogen als de UP-kans groter is dan de DOWN-kans). Daarna worden de gewichten opnieuw genormaliseerd. Op basis van de resultaten van het HMM, kan worden bepaald hoe defensief het model moet zijn.

3.3 Constructie van het portefeuillemodel (stappenplan)

In dit onderdeel wordt stap voor stap beschreven hoe het EH-RP-model in de tijd wordt toegepast. Het gaat om een iteratief proces, waarbij op elk tijdstip t alleen informatie wordt gebruikt die tot en met $t - 1$ bekend is. Dit voorkomt het gebruik van voorkennis (look-ahead bias) en sluit aan bij hoe een belegger in de praktijk beslissingen zou nemen. Het algoritme kan in woorden als volgt worden weergegeven:

1. Allereerst werd in de Python code de rendementen ingeladen, via yfinance en Robert Schillers datasets, en werden de logrendementen berekend per tijdseenheid. Voor dit onderzoek is dat per dag gedaan. Er geldt:

$$r_{t,i} = \ln \left(\frac{P_{t,i}}{P_{t-1,i}} \right)$$

Alleen historisch bekende data tot en met dag $t - 1$ werd gebruikt voor elke iteratieve stap. Voor elk nieuw beslismoment t werd alleen de informatie gebruikt die beschikbaar is tot en met dag $t - 1$. Dit omvatte de historische prijzen, rendementen en eerder geschatte volatiliteiten en regimes.

2. Voor elke asset i wordt op basis van de rendementen tot en met $t - 1$ een EGARCH(1,1)-model geschat of geüpdatet. Dit levert een inschatting van de voorwaardelijke volatiliteit $\hat{\sigma}_{i,t}$, die aangeeft hoe risicovol asset i wordt geacht op dag t . Er geldt:

$$\ln(\sigma_{t,i}^2) = \omega_i + \alpha_i (|\varepsilon_{t-1,i}| - E|\varepsilon_{t-1,i}|) + \gamma_i \varepsilon_{t-1,i} + \beta_i \ln(\sigma_{t-1,i}^2)$$

3. Het Hidden Markov Model werd toegepast op een gekozen referentiereeks om de kans te schatten dat de markt zich op dag t in een rustig dan wel stressregime bevindt. Dit resulteerde in een regime-indicator S_t . Voor deze S_t geldt:

$$S_{t,i} \in \{0, 1\} \quad \text{met} \quad P_i(a, b) = P(S_{t,i} = b \mid S_{t-1,i} = a)$$

4. Op basis van de historische UP/DOWN-sequenties van de rendementen van elke asset werd een overgangsmatrix geconstrueerd. Hieruit werden de kansen afgeleid dat een positieve (UP) of negatieve (DOWN) beweging zich

voortzet. Deze kansen werden gebruikt als een lichte richtinggevende factor (zie stap 6). Er geldt:

$$\mathbf{P} = \begin{pmatrix} P(\text{up} \rightarrow \text{up}) & P(\text{up} \rightarrow \text{down}) \\ P(\text{down} \rightarrow \text{up}) & P(\text{down} \rightarrow \text{down}) \end{pmatrix}$$

5. Uit de geschatte volatiliteiten $\hat{\sigma}_{i,t}$ werden de basisgewichten bepaald volgens de naïeve risk parity. In de eenvoudigste vorm geldt dat

$$w_{i,t}^{\text{RP}} \propto \frac{1}{\hat{\sigma}_{i,t}},$$

waarna de gewichten worden geschaald zodat $\sum_{i=1}^N w_{i,t}^{\text{RP}} = 1$. Assets met een hogere volatiliteit kregen zo automatisch een lager gewicht.

6. De basisgewichten $w_{i,t}^{\text{RP}}$ werden vervolgens licht aangepast op basis van de netto UP–DOWN-indicator per asset. Als de kans op een stijging (UP) duidelijk groter was dan de kans op een daling (DOWN), werd het gewicht van die asset iets verhoogd; in het omgekeerde geval werd het gewicht verlaagd. De aanpassing bleef bewust klein, zodat de risicocomponent (volatiliteit) dominant bleef.

$$\Delta_{t,i} = \lambda \left(\pi_{t,i}^{\text{up}} - \pi_{t,i}^{\text{down}} \right)$$

$$\hat{w}_{t,i} = w_{t,i}^{\text{RP}} (1 + \Delta_{t,i})$$

Lambda werd aangepast op basis van de resultaten die door het HMM waren verkregen. Als het HMM aangaf dat er sprake was van een crisisregime, dan werd $\lambda = 0.05$ gebruikt, en als er een rustig regime was, werd $\lambda = 0.1$ genomen, omdat er dan meer risico kon worden genomen.

7. Na de richtingstilt werden de aangepaste gewichten opnieuw genormaliseerd, zodat geldt:

$$\sum_{i=1}^N \hat{w}_{t,i} = 1.$$

Hiermee ontstond de uiteindelijke gewichtsverdeling voor dag t .

8. Om de backtesting ook goed te laten verlopen is het belangrijk om transactiekosten mee te laten rekenen. Daarvoor werd berekend hoe vaak een verandering het in portefeuille plaatsvond. De verandering werd vermenigvuldigd met de kosten voor de verandering:

$$\text{Turnover}_{t,i} = \left| w_{t,i}^{\text{nieuw}} - w_{t,i}^{\text{oud}} \right|$$

$$\text{Totale turnover} = \sum_i \left| w_{t,i}^{\text{nieuw}} - w_{t,i}^{\text{oud}} \right|$$

$$\text{TC}_t = c \cdot \text{Totale turnover}$$

Voor transactiekosten wordt een vaste kostenfactor van $c = 0,1\%$ per verandering in portefeuillegewicht gehanteerd ($tc_{rate} = 0.001$). De dagelijkse turnover wordt berekend als de helft van de som van de absolute veranderingen in asset-gewichten.

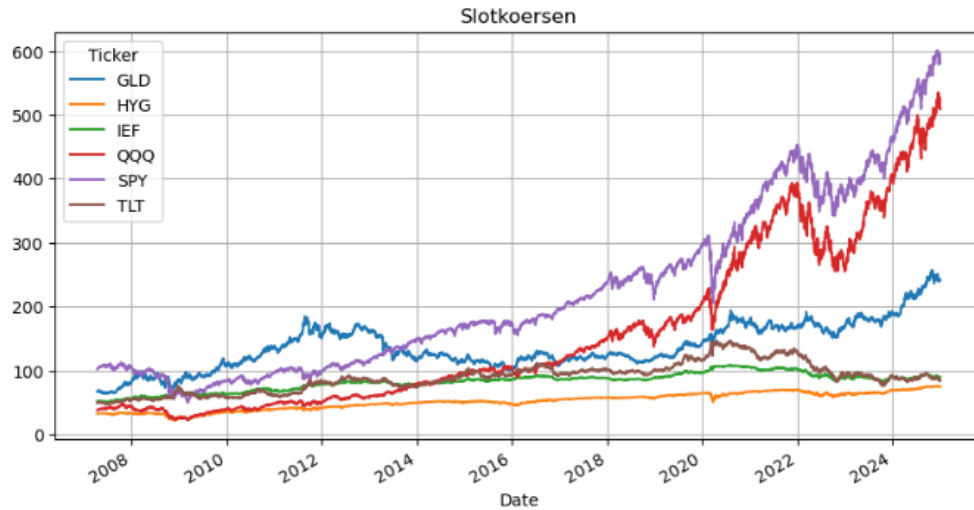
Voor de backtesting zijn ook andere dingen berekend, zoals de sharpe-ratio en het totale rendement (zie code in de bijlage); dat spreekt voor zich. De hele code kan aan het eind van het profielwerkstuk bij de bijlagen worden teruggevonden. Door dit stappenplan gedurende de hele onderzoeksperiode (2007-2025) herhaaldelijk toe te passen, ontstaat een tijdreeks van portefeuillerendementen die kan worden vergeleken met de andere defensieve strategieën. Hierdoor kan beoordeeld worden in hoeverre het GARCH-HMM-risk-parity model leidt tot een betere risico-gewogen performance. Deze resultaten worden in het volgende hoofdstuk uitgewerkt. Op die manier kan worden getest of het model ook daadwerkelijk werkt en kan een antwoord worden gegeven op de onderzoeksvraag.

4 Resultaten

In dit hoofdstuk worden de resultaten van de backtest van 2007 tot en met 2025 weergegeven. Er is bewust voor deze periode gekozen, omdat het verschillende risicovolle perioden bevat, maar ook perioden van rust. De backtest is uitgevoerd met Python in de Jupyter-omgeving. De code is in de bijlagen terug te vinden. Het doel van dit hoofdstuk is om alle resultaten zo bondig mogelijk weer te geven, zonder er verder op in te gaan. In het hoofdstuk Discussie worden de resultaten verder behandeld.

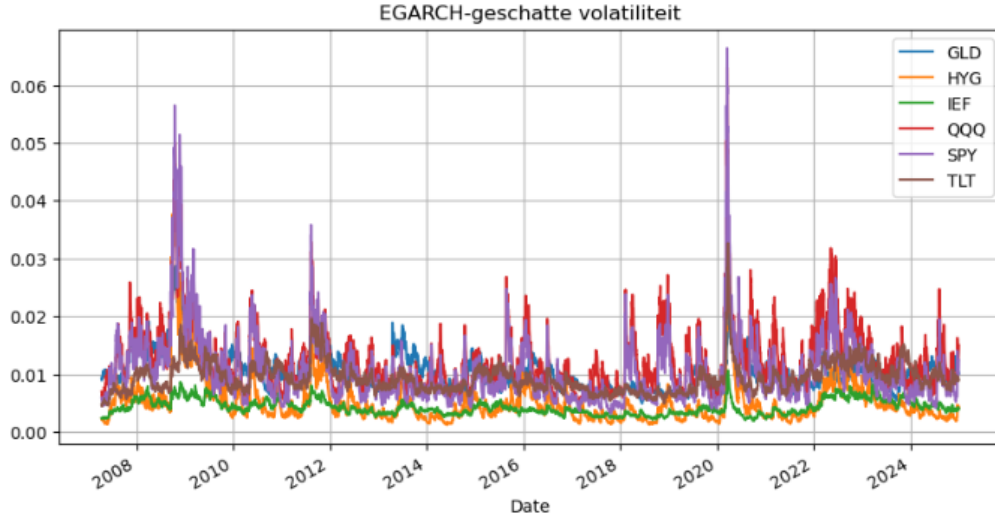
4.1 Rendementen en de resultaten van de EGARCH-schattingen

De rendementen van de set bestaande uit GLD, QQQ, TLT, IEF, SPY en HYG kunnen grafisch worden weergegeven met behulp van een tijdreeksplot. In figuur 1 is deze grafiek weergegeven, waardoor de dynamiek en volatiliteit van de verschillende assets duidelijk zichtbaar worden. Dit is voor dit onderzoek een heel



Figuur 1: Tijdreeks van de rendementen van GLD, QQQ, TLT, IEF, SPY en HYG

relevante plot. Alle theorie is zichtbaar. We zien dat er verschillende soorten activa bestaan. Activa met een kleinere variantie resulteren in een lager rendement. In de groei van elke asset is ook het risico te zien. Waar Nasdaq 100 en S&P 500 nog extreme drawdowns hebben (zie bijvoorbeeld 2022), hebben activa met minder risico (zoals obligaties) nauwelijks verlies. Het HMM kan op basis van deze rendementen berekenen of het portefeuille zich bevindt in een crisis regime of rustig regime en was een indicator voor λ . Daarnaast is met behulp van het EGARCH(1,1)-model de volatiliteit per t voorspeld. Dit is iteratief voor elke dag gedaan, en het resultaat is te zien in figuur 2.



Figuur 2: Resultaten EGARCH backtest 2007-2025

De onderstaande overgangsmatrices laten zien hoe vaak rendementen elkaar in dezelfde richting opvolgen. Zo zie je meteen welke assets meestal blijven stijgen na een stijging, en welke juist blijven dalen na een daling. Dit geeft een eenvoudig beeld van het momentum in de markt en helpt om de portefeuille iets mee te sturen met de richting van de trend. Deze data is nodig voor de tilt die plaatsvindt na de risk parity. De matrices per asset zijn als volgt weer te geven:

GLD

$$\begin{pmatrix} 0.45798526 & 0.54201474 \\ 0.49175212 & 0.50824788 \end{pmatrix}$$

HYG

$$\begin{pmatrix} 0.47923962 & 0.52076038 \\ 0.45677929 & 0.54322071 \end{pmatrix}$$

IEF

$$\begin{pmatrix} 0.46347363 & 0.53652637 \\ 0.50203528 & 0.49796472 \end{pmatrix}$$

QQQ

$$\begin{pmatrix} 0.44397906 & 0.55602094 \\ 0.44847973 & 0.55152027 \end{pmatrix}$$

SPY

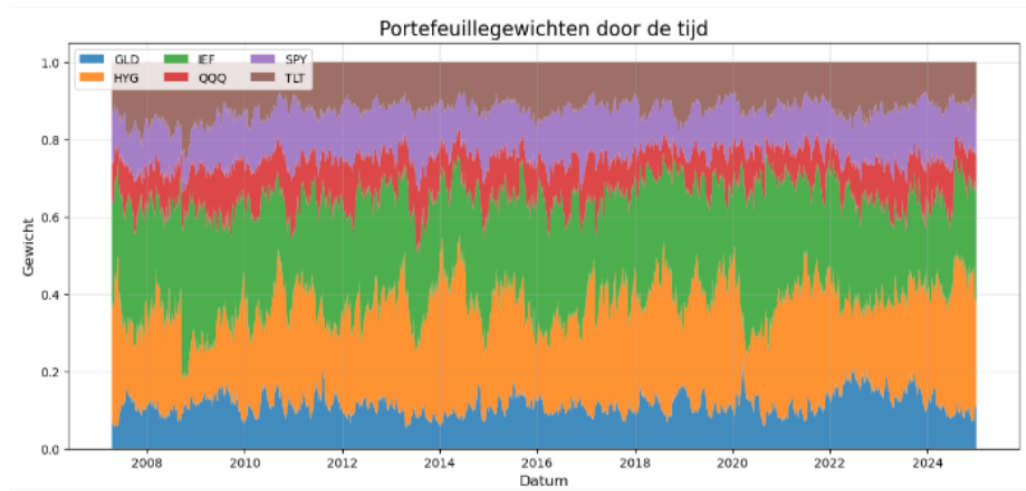
$$\begin{pmatrix} 0.44126819 & 0.55873181 \\ 0.45666950 & 0.54333050 \end{pmatrix}$$

TLT

$$\begin{pmatrix} 0.44985394 & 0.55014606 \\ 0.50854317 & 0.49145683 \end{pmatrix}$$

Op basis van deze resultaten (EGARCH-schattingen, de UP-DOWN-overgangsmatrices

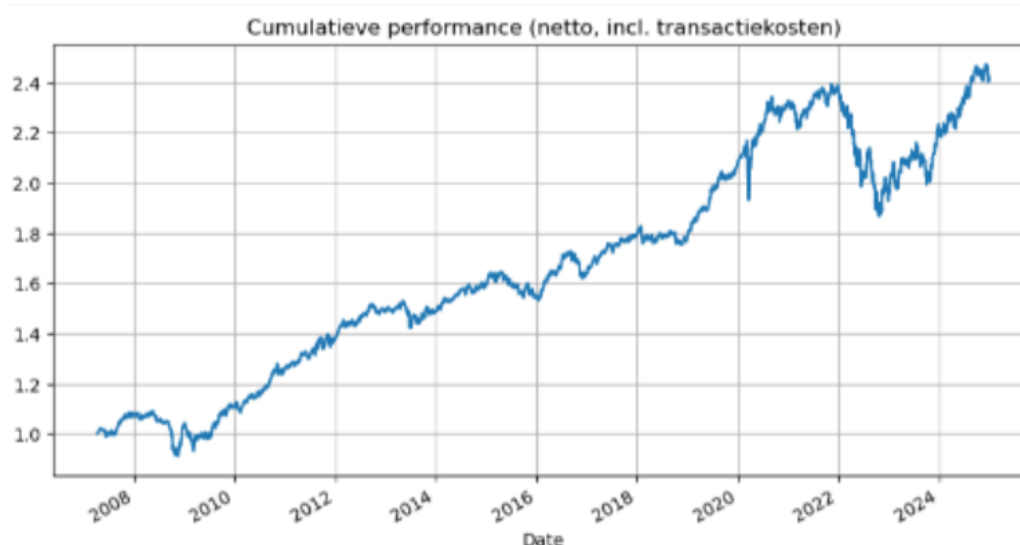
en de HMM-regime-indicator) is per dag een gewichtsverdeling gemaakt over de verschillende activa volgens naïeve risk parity. Hierbij wordt voor elke dag alleen gebruikgemaakt van informatie die tot en met de vorige dag beschikbaar was, zodat er geen sprake is van look-ahead bias. De GARCH-volatiliteiten bepalen de basisverdeling, terwijl de richtinggevende informatie uit de UP–DOWN-matrix dient als een kleine tilt op deze gewichten. Door dit proces iteratief toe te passen ontstaat een dynamisch portefeuilleepad waarbij het risicobudget continu wordt aangepast aan de veranderende marktomstandigheden. Deze dagelijkse gewichtsverdelingen kunnen overzichtelijk worden weergegeven in een tijdreeksplot, waarin duidelijk zichtbaar wordt hoe de allocatie door de tijd heen verschuift. De backtest is hieronder in figuur 3 in de tijdreeksplot grafisch terug te zien.



Figuur 3: Portefeuillegewichten backtest 2007-2025

4.2 Prestaties van het EH-RP-model

In deze paragraaf worden de resultaten van de cumulatieve performance van het EH-RP-model gegeven. De resultaten zijn in een tijdreeksplot verwerkt. Zie figuur 4. De cumulatieve performance-grafiek laat zien dat het EH-RP-model over de periode 2007–2025 een gelijkmatige opwaartse trend vertoont, ondanks meerdere perioden van aanzienlijke marktstress. Tijdens de financiële turbulentie van 2008–2009, de coronacrisis rond 2020 en de volatiliteit in 2022 blijft de terugval beperkt, wat erop wijst dat het model de risicoblootstelling effectief heeft verlaagd in deze onrustige fases. In rustige marktperioden weet het model daarentegen mee te profiteren van stijgende koersen door iets meer risico toe te laten. Hierdoor ontstaat een relatief stabiele rendementscurve.



Figuur 4: Cumulatieve performance backtest EH-RP-model 2007-2025

De prestatiestatistieken ondersteunen dit beeld (zie de tabel hieronder). Het jaarlijkse gemiddelde netto rendement van circa 5,2% in combinatie met een jaarlijkse volatiliteit van 8,66% wijst op een defensief maar efficiënt profiel. De Sharpe-ratio van 0,48 bevestigt dat het model in staat is om per eenheid risico een redelijk consistent rendement te genereren. Hoewel dit geen uitzonderlijk hoge Sharpe-ratio

Tabel 2: Prestatiestatistieken van het EH-RP-model (2007–2025)

Metriek	Waarde
Jaarlijks gemiddeld netto rendement	5,19%
Jaarlijkse netto volatiliteit	8,660%
Sharpe-ratio (netto)	0,4823

is, sluit het wel goed aan bij de doelstelling van het model: risicominimalisatie met behoud van voldoende opwaarts potentieel. De resultaten suggereren dat het EH-RP-model vooral waardevol is in omstandigheden waarin markten schokkerig, onzeker of wisselend van karakter zijn.

5 Discussie

5.1 Interpretatie van de resultaten

Als we kijken naar hoe het EH–RP-model uiteindelijk heeft gepresteerd, valt meteen op dat de groei van de portefeuille vrij gelijkmatig verloopt. De lijn stijgt rustig door de jaren heen, zelfs terwijl er verschillende perioden zijn waarin de markt juist erg onrustig was. Dat laat zien dat het model goed doet waar het voor bedoeld is: het beperken van risico, vooral op momenten dat de markt “in brand” lijkt te staan. Tijdens de coronacrash in 2020 en de onrustige periode rond 2022 schakelt het model bijvoorbeeld duidelijk terug in risicoblootstelling. Daardoor vallen de dalingen in de portefeuille minder heftig uit dan bij een simpele buy-and-holdstrategie.

Uit de cijfers komt dat beeld duidelijk naar voren. Een gemiddeld netto jaarrendement van ongeveer 5% is niet spectaculair hoog, maar in combinatie met een relatief lage volatiliteit van 8,66% laat het wél zien dat de strategie stabiel en beheerst werkt. De sharpe-ratio van 0,48 bevestigt dat het model aardig presteert in verhouding tot het genomen risico. Dit past ook precies bij het doel van de strategie: defensief beleggen, niet proberen de markt te verslaan met hoge risico's.

Er zijn ook een paar dingen die opvallen als je kritisch naar het model kijkt. Zo reageert het niet altijd even snel op plotselinge schokken. Dat komt omdat het model alleen naar het verleden kijkt: zowel EGARCH als het HMM hebben eerst meerdere dagen of weken aan data nodig voordat ze echt herkennen dat de markt van “rust” naar “stress” is gewicht. Daardoor komt de eerste klap bij een crisis soms alsnog binnen. Het model doet dus niet aan voorspellen, maar vooral aan zo snel mogelijk aanpassen — en dat lukt meestal behoorlijk goed, maar niet perfect.

Een andere beperking is dat een defensieve strategie minder profiteert in perioden waarin de markt juist sterk stijgt. Omdat risk parity automatisch minder gewicht geeft aan volatiele assets, zit de portefeuille nooit maximaal in aandelen. Dat betekent dat de portefeuille rustiger beweegt dan de markt, maar ook dat de rendementen niet zo hard omhoog schieten tijdens bull-markten. Dat is een bewuste keuze, maar wel eentje die gevolgen heeft voor het uiteindelijke rendement.

Alles bij elkaar laat het model zien dat het goed werkt voor beleggers die vooral stabiliteit willen en niet wakker willen liggen van grote pieken en dalen. Het vangt risico goed op tijdens stressvolle perioden en levert ondertussen een redelijk, betrouwbaar rendement op. Tegelijkertijd moet je accepteren dat je in hele sterke markten wat achterblijft en dat het model soms net iets te laat reageert op omslagpunten. Maar als defensieve strategie doet het precies waar het voor ontworpen is.

5.2 Vergelijking met andere strategieën: antwoord op deelvraag 3

Op basis van deze informatie kan antwoord worden gegeven op deelvraag 3: *Hoe verhouden het rendement, het risico en het risicogecorrigeerd rendement van de volatiliteitsafhankelijke strategie zich tot die van andere defensieve strategieën in verschillende marktregimes?* In het theoretisch kader kwam naar voren wat defensieve strategieën zijn en werden er een aantal voorbeelden genoemd. In de paragraaf, Defensieve strategieën: een theoretisch overzicht, werd een tabel getoond met een aantal defensieve strategieën. Aan deze tabel kunnen ook de resultaten

van het EH-RP-model worden gegeven. Dan krijgt men de volgende tabel: Deze

Tabel 3: Voorbeeld prestaties van defensieve strategieën, inc. EH-RP-model

Strategie	Rendement	Sharpe	Vol.
Spaarrente	0-1.5%	0.40	0.2%
Staatsobligaties	2.8%	0.3-0.5	6-10%
Investment Grade	4.2%	0.60	6%
High Yield	6.5%	0.45	12%
Defensieve Aandelen	7.0%	0.65	10%
EH-RP-model	5.19%	0,48	8,66%

tabel geeft weer hoe het model presteert ten op zichte van de andere modellen. De resultaten laten zien dat het model wel degelijk kan concurreren met andere defensieve strategieën. Het rendement is in sommige gevallen zelfs hoger dan andere strategieën, bijvoorbeeld bij spaarrente en staatsobligaties. Het risico is dan wel iets groter, maar nog steeds aanzienlijk klein. De resultaten geven aan dat het model werkt en dat het een goed alternatief kan zijn voor defensieve beleggers. Daarbij wordt antwoord gegeven op deelvraag 3. De hypothese (H1) klopt en de nulhypothese (H0) kan worden verworpen.

5.3 Aanbevelingen voor vervolgonderzoek, validiteit en betrouwbaarheid

Toch is het model nog niet compleet. Het doel bij wiskundig modelleren is om het model continu te verbeteren op basis van nieuwe inzichten. Ook voor dit model is dat belangrijk. Het model kan op verschillende manier verbeteren. Een paar manieren worden hieronder opgesomd:

1. In dit model is alleen variantie toegevoegd, maar niet covariantie. Een nieuwe versie van het model zou covariantie kunnen meenemen, zodat de totale variatie van het portefeuille door de tijd heen nog lager wordt. Op die manier wordt de sharpe-ratio groter.
2. Ook kunnen er andere volatiliteitsmodellen gebruikt worden. Denk aan GJR-GARCH, HAR-volatiliteit of Realized Volatility op basis van intraday data om te testen of deze een beter risicosignaal geven dan EGARCH. Ook daarvoor zijn backtests nodig om te controleren of het daadwerkelijk beter is.
3. Simulaties uitvoeren onder extreme scenario's. Dat is voor dit onderzoek niet gedaan, maar dat is enorm interessant om te onderzoeken. Bijvoorbeeld met stresstesten of Monte Carlo-simulaties, om te onderzoeken hoe robuust het model blijft als zich onverwachte, extreme gebeurtenissen voordoen die niet in de historische data voorkomen.
4. Een van de belangrijkste dingen om nog toe te voegen is kunstmatige intelligentie. Deze applicatie kan zorgen voor een nog betere aanpak van de probleemstelling en resulteert per definitie in betere resultaten. Veel fondsen integreren steeds met KI en ook voor dit model zou het een goede toevoeging kunnen zijn.
5. λ is nu nog binair: het is of 0.05 (crisis) of 0.1 (rust). Een toevoeging zou kunnen zijn om deze parameter dynamischer te maken op basis van de resultaten van het HMM. Er zou zelfs een keuze gemaakt kunnen worden om op

basis van de resultaten van het HMM om in sommige gevallen (bijvoorbeeld bij crisis) volledig over te gaan op goud of cash. Het doel is daarbij om het geld nog extra te beschermen tegen extreme drawdowns.

Kwantitatieve finance is enorm veelzijdig en er zijn enorm veel andere aanbevelingen nog voor vervolgonderzoek. Het is van groot belang om bij elke nieuwe mogelijke aanpassing goed te blijven kijken naar de doelstelling van het onderzoek. Het probleem is dat de markt dynamisch is en moeilijk te modelleren. De finance wereld is enorm competitief en er verschijnen elke dag van dit soort nieuwe technische modellen op de markt. Het is belangrijk om unieke modellen te modelleren die snel, handig en dynamisch zijn. De betrouwbaarheid en validiteit van het model zijn momenteel voldoende gewaarborgd, want er is een backtest gedaan van 2007 tot en met 2025. De resultaten zijn op die manier betrouwbaar en de resultaten zijn valide (laag risico). Toch kunnen de betrouwbaarheid en validiteit verder worden versterkt door de hierboven genoemde nuanceringen toe te voegen. Het toepassen van bijvoorbeeld Monte-Carlo-simulaties om fictieve scenario's door te rekenen, verhoogt de betrouwbaarheid doordat de robuustheid van het model onder uiteenlopende omstandigheden expliciet wordt getest.

6 Conclusie en antwoord op de onderzoeksvraag

Voor het onderzoek was de hoofdvraag op de volgende manier geformuleerd: *In welke mate presteert een risico-minimaliserende beleggingsstrategie die zich aanpast aan veranderende marktvolatiliteit beter of slechter dan andere defensieve beleggingsstrategieën?* Op basis van de resultaten kan worden geconcludeerd dat de risico-minimaliserende strategie die gebruikmaakt van volatiliteitsschattingen en marktregimes over het algemeen stabiel en consistent presteert in sommige gevallen dan de andere defensieve strategieën (bijvoorbeeld spaarrente en obligaties) waarmee zij is vergeleken. De resultaten bij de backtest van 2007 tot en met 2025 waren namelijk: een gemiddeld jaarlijks rendement van 5,2% en een jaarlijkse volatiliteit van 8,7% (netto). Het model weet de grootste dalingen tijdens crisismomenten beter te beperken, waardoor het rendementspad gelijkmatiger verloopt. Hoewel het rendement niet altijd hoger is dan dat van simpelere strategieën in sterke bull-markten, levert het model wel een betere balans tussen risico en rendement. De sharpe-ratio en de relatief lage volatiliteit laten zien dat het model vooral effectief is voor beleggers die bescherming zoeken tegen grote uitschieters, zonder het opgeven van lange-termijngroei. In die zin presteert de strategie in dit onderzoek beter in termen van risico-gewogen rendement, wat aansluit bij het doel van defensief beleggen.

Literatuurlijst

MSCI. (2025). *MSCI World Index Factsheet*. Geraadpleegd van <https://www.msci.com/documents/10199/149ed7bc-316e-4b4c-8ea4-43fcb5bd6523>

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7 (1), 77–91.

Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1(2), 223–236.

Ilmanen, A. (2011). *Expected returns: An investor's guide to harvesting market rewards*. Wiley.

Zuckerman, G. (2019). *The man who solved the market: How Jim Simons launched the quant revolution*. Penguin Books.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4), 987–1007.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2), 383–417.

Lo, A. W. (2004). The adaptive markets hypothesis: Market efficiency from an evolutionary perspective. *Journal of Portfolio Management*, 30(5), 15–29.

Bogle, J. C. (1999). *Common sense on mutual funds: New imperatives for the intelligent investor*. Wiley.

Buffett, W. E. (2008). *Berkshire Hathaway Inc. Annual Letter to Shareholders*. Berkshire Hathaway.

Ilmanen, A. (2011). *Expected returns: An investor's guide to harvesting market rewards*. Wiley.

Bogle, J. C. (2007). *The little book of common sense investing: The only way to guarantee your fair share of stock market returns*. Wiley.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1), 77–91.

Mandelbrot, B. (1963). The variation of certain speculative prices. *The Journal of Business*, 36(4), 394–419.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1), 77–91.

Bodie, Z., Kane, A., & Marcus, A. J. (2014). *Investments* (10th ed.). McGraw-Hill.

Mandelbrot, B. (1963). The variation of certain speculative prices. *The Journal of Business*, 36(4), 394–419.

- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4), 987–1007.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327.
- Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59(2), 347–370.
- Hamilton, J. D. (1994). *Time series analysis*. Princeton University Press.
- Bollerslev, T., & Wooldridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews*, 11(2), 143–172.
- Maillard, S., Roncalli, T., & Teïletche, J. (2010). The properties of equally weighted risk contributions portfolios. *The Journal of Portfolio Management*, 36(4), 60–70.
- De Carvalho, R., Xiao, Y., & Zilberman, E. (2012). The case for naive risk parity. *Journal of Portfolio Management*, 38(4), 22–41.
- Roncalli, T. (2013). *Introduction to risk parity and budgeting*. Chapman and Hall/CRC.
- Ross, S. M. (1996). *Stochastic processes* (2nd ed.). Wiley.
- Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2), 357–384.
- Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2), 257–286.
- Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2), 257–286.
- Aroussi, R. (2023). *yfinance: Yahoo! Finance market data downloader*. GitHub repository. <https://github.com/ranaroussi/yfinance>
- Shiller, R. J. (2023). *Online data: Irrational exuberance*. Yale University. <http://www.econ.yale.edu/~shiller/data.htm>

A Bijlage 1: Python-code van het model

In deze bijlage is de code van het Python model verantwoord. Het model is opgesplitst in verschillende onderdelen. Elke cel is een onderdeel van het totale model. Bovenaan de cel wordt het doel van de cel benoemd.

```
# installeren van packages
!pip install pandas
!pip install numpy
!pip install matplotlib
!pip install arch
!pip install hmmlearn
!pip install yfinance
```

```
# installeren van imports
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

from arch import arch_model

from hmmlearn.hmm import GaussianHMM
import yfinance as yf

plt.rcParams['figure.figsize'] = (10, 5)
pd.options.display.float_format = "{:,.6f}".format

print("Imports klaar.")
```

```
# data ophalen van elke asset
tickers = [
    "GLD",
    "HYG",
    "IEF",
    "QQQ",
    "SPY",
    "TLT",
]

start_date = "1970-01-01"
end_date = "2025-01-01"

data = yf.download(tickers, start=start_date, end=end_date)

prices = data["Close"].dropna()

print(prices.head())

prices.plot(title="Slotkoersen")

plt.grid(True)
plt.show()
```

```
# Log-rendementen berekenen
returns = np.log(prices / prices.shift(1)).dropna()
```

```

returns.head()

returns.plot(title="Dagelijkse log-rendementen")
plt.grid(True)
plt.show()

```

```

# EGARCH voorspelt volatiliteit
def fit_egarch_vol_series(r, dist="normal"):
    am = arch_model(r * 100, vol="EGARCH", p=1, o=1, q=1, dist=
        dist)
    res = am.fit(update_freq=0, disp="off")

    sigma = res.conditional_volatility / 100.0
    sigma.name = r.name
    return sigma, res

vol_dict = {}
model_results = {}

for col in returns.columns:
    print(f"Schat EGARCH voor {col}...")
    sigma_col, res_col = fit_egarch_vol_series(returns[col])
    vol_dict[col] = sigma_col
    model_results[col] = res_col

sigma_df = pd.DataFrame(vol_dict).dropna()
sigma_df.head()

sigma_df.plot(title="EGARCH-geschatte volatiliteit")
plt.grid(True)
plt.show()

```

```

# overgangsmatrices
def compute_markov_up_down_matrix(r):
    r = r.dropna()
    sign = (r > 0).astype(int)

    s_prev = sign.shift(1).dropna()
    s_curr = sign.loc[s_prev.index]

    s_prev = s_prev.astype(int)
    s_curr = s_curr.astype(int)

    counts = np.zeros((2, 2), dtype=float)

    for i, j in zip(s_prev, s_curr):
        if i in (0, 1) and j in (0, 1):
            counts[i, j] += 1

    row_sums = counts.sum(axis=1, keepdims=True)
    row_sums[row_sums == 0] = 1
    P = counts / row_sums

    return P, sign

markov_matrices = {}

```

```

sign_dict = {}

for col in returns.columns:
    P, sign = compute_markov_up_down_matrix(returns[col])
    markov_matrices[col] = P
    sign_dict[col] = sign
    print(f"\nOverschakelingsmatrix voor {col}: \n{P}")

# volgende dag up/down-kansen
def compute_next_up_down_probs(sign_series, P):
    idx = sign_series.index
    pi_up = pd.Series(index=idx, dtype=float)
    pi_down = pd.Series(index=idx, dtype=float)

    for t in idx:
        state = sign_series.loc[t]
        if state not in (0,1):
            pi_up.loc[t] = np.nan
            pi_down.loc[t] = np.nan
            continue

        p_down_next = P[state, 0]
        p_up_next = P[state, 1]
        pi_up.loc[t] = p_up_next
        pi_down.loc[t] = p_down_next

    return pi_up, pi_down

up_probs = {}
down_probs = {}

for col in returns.columns:
    P = markov_matrices[col]
    sign_series = sign_dict[col]
    pi_up, pi_down = compute_next_up_down_probs(sign_series,
        P)
    up_probs[col] = pi_up
    down_probs[col] = pi_down

up_df = pd.DataFrame(up_probs).loc[sigma_df.index]
down_df = pd.DataFrame(down_probs).loc[sigma_df.index]

up_df.head()

# risk parity
def risk_parity_naive_weights(sigma_row):
    inv_sigma = 1.0 / sigma_row
    w = inv_sigma / inv_sigma.sum()
    return w

weights_rp = sigma_df.apply(risk_parity_naive_weights, axis
    =1)
weights_rp.head()

weights_rp.plot(title="Risk Parity gewichten (zonder tilt)")
plt.grid(True)

```

```
plt.show()
```

```
# tilt
lambda_dir = 0.05

def apply_directional_tilt(w_rp_row, up_row, down_row, lam=
    lambda_dir):
    tilt = lam * (up_row - down_row)
    w_tilted = w_rp_row * (1.0 + tilt)
    w_tilted = w_tilted.clip(lower=0.0)
    if w_tilted.sum() > 0:
        w_tilted = w_tilted / w_tilted.sum()
    else:
        w_tilted = w_rp_row
    return w_tilted

up_aligned = up_df.loc[sigma_df.index]
down_aligned = down_df.loc[sigma_df.index]

weights_tilted = pd.DataFrame(index=sigma_df.index, columns=
    sigma_df.columns, dtype=float)

for t in sigma_df.index:
    w_rp = weights_rp.loc[t]
    up_row = up_aligned.loc[t]
    down_row = down_aligned.loc[t]
    weights_tilted.loc[t] = apply_directional_tilt(w_rp,
        up_row, down_row, lambda_dir)

weights_tilted.head()

weights_tilted.plot(title="Gewichten met directionele tilt")
plt.grid(True)
plt.show()
```

```
# HMM
from hmmlearn.hmm import GaussianHMM

def rolling_hmm_no_lookahead(series, window=250):

    series = series.dropna()
    X = series.values.reshape(-1,1)
    regimes = []
    last_reg = 0

    for i in range(len(series)):
        start = max(0, i-window)
        X_window = X[start:i+1]

        # min data nodig
        if len(X_window) < 60:
            regimes.append(last_reg)
            continue

        try:
            hmm = GaussianHMM(
```



```

        n_components=2,
        covariance_type="diag",
        n_iter=30,
        init_params="",
        params="stmc"
        means, covars
    )

    # stabiele startwaarden
    hmm.startprob_ = np.array([0.5, 0.5])
    hmm.transmat_ = np.array([[0.90, 0.10],
                               [0.10, 0.90]])

    mean = np.mean(X_window)
    var = np.var(X_window) + 1e-6

    hmm.means_ = np.array([[mean], [mean*0.9]])
    hmm.covars_ = np.array([[var], [var]])

    hmm.fit(X_window)

    reg = hmm.predict(X_window)[-1]
    last_reg = reg
    regimes.append(reg)

except Exception:
    regimes.append(last_reg)

return pd.Series(regimes, index=series.index)

print("Rolling_HMM_klaar.")

```

```

# backtest 2007-2025
common_idx = weights_tilted.index.intersection(returns.index)
w = weights_tilted.loc[common_idx]
r = returns.loc[common_idx]

tc_rate = 0.0005 # 0.05% per verhandelde euro

dates = w.index

gross_ret_list = []
net_ret_list = []
turnover_list = []

for i, t in enumerate(dates):
    w_t = w.iloc[i]
    r_t = r.iloc[i]

    gross_ret_t = (w_t * r_t).sum()

    if i == 0:
        turnover_t = 0.0
    else:
        w_prev = w.iloc[i - 1]
        turnover_t = (w_t - w_prev).abs().sum() / 2.0

```

```

    tc_t = tc_rate * turnover_t
    net_ret_t = gross_ret_t - tc_t

    gross_ret_list.append(gross_ret_t)
    net_ret_list.append(net_ret_t)
    turnover_list.append(turnover_t)

gross_port_ret = pd.Series(gross_ret_list, index=dates, name=
    "gross_ret")
net_port_ret    = pd.Series(net_ret_list,    index=dates, name=
    "net_ret")
turnover        = pd.Series(turnover_list,   index=dates, name=
    "turnover")

gross_port_ret.head(), net_port_ret.head(), turnover.head()

cum_net = (1 + net_port_ret).cumprod()
cum_net.plot(title="Cumulatieve performance (netto, incl. transactiekosten)")
plt.grid(True)
plt.show()

```

```

#resultaten backtest
risk_free_rate = 0.04 # 4% per jaar
periods_per_year = 252

mean_ret_net = net_port_ret.mean() * periods_per_year
vol_ret_net  = net_port_ret.std() * np.sqrt(periods_per_year)
sharpe_net   = (mean_ret_net - risk_free_rate) / vol_ret_net

print(f"Jaarlijks gemiddeld NETTO rendement: {mean_ret_net:.4f}")
print(f"Jaarlijkse volatiliteit (netto): {vol_ret_net:.4f}")
print(f"Sharpe-ratio NETTO: {sharpe_net:.4f}")

```

B Bijlage 2: Logboek

Deze bijlage geeft een weergave van hoe het proces is verlopen. Een belangrijke sidenote is dat dit een indicatie is van hoe het profielwerkstuk is uitgevoerd. In feite heb ik de afgelopen jaren veel meer over finance geleerd, alleen die tijden zijn niet opgenomen.

Datum	Activiteit	Minuten
15-04-2024	Honours: <i>The Man Who Solved the Market</i> lezen	45
22-04-2024	Honours: notities over kwantitatieve traders	45
29-04-2024	Honours: <i>The Psychology of Money</i>	45
06-05-2024	Oriëntatie marktdynamiek & risicobeheer	60
13-05-2024	Honours: <i>Rich Dad Poor Dad</i>	45
20-05-2024	Honours: behavioural finance	45
27-05-2024	Honours: modellen van Renaissance Technologies	45
06-09-2024	Honours: volatiliteit & financiële markten	45
13-09-2024	Honours: basisstatistiek & variantie	45
20-09-2024	Honours: literatuur samenvatten	45
04-10-2024	Honours: portefeuilletheorie	45
11-10-2024	Brainstorm onderzoeksrichting	90
18-10-2024	Honours: correlatieanalyse markten	45
01-11-2024	Honours: risico-indicatoren	45
15-11-2024	Oriëntatie EGARCH & regime shifts	75
29-11-2024	Honours: literatuurnotities	45
13-12-2024	Onderzoeksopzet voorbereiden	45
14-12-2024	Technisch interview met L. Bos	60
14-12-2024	Feedback van L. Bos verwerken	45
10-01-2025	Onderzoeksopzet uitwerken	90
17-01-2025	Literatuur verzamelen	45
24-01-2025	Structuur onderzoeksvraag opstellen	45
31-01-2025	Fundament Markov-ketens bestuderen	45
07-02-2025	Methodesectie schrijven	45
14-02-2025	Dataverzameling plannen	60
21-02-2025	Statistische basis voorbereiden	45
28-02-2025	Dataset selecteren & ordenen	45
14-03-2025	HMM-basis programmeren (states & emissies)	180
18-03-2025	Gesprek De Kok – voortgang	30
06-05-2025	Gesprek De Kok – feedback modelopzet	30
06-07-2025	Black-Scholes theoretische basis	150
13-07-2025	Efficiënte Markt Hypothese (sterk/zwak/semi-sterk)	150
16-07-2025	Technisch interview met I. Goddijn	60
16-07-2025	Feedback van I. Goddijn verwerken	45
20-07-2025	Black-Scholes: optieprijs berekenen	120
27-07-2025	GARCH-variantie modelleren	150
03-08-2025	HMM: forward/backward algoritme programmeren	180
10-08-2025	EGARCH – asymmetrische volatiliteit	150
08-09-2025	Code verbeteren & HMM-states verfijnen	40
09-09-2025	Code organiseren & bronnen samenvatten	180
14-09-2025	Papers lezen over volatility regimes	30

23-09-2025	State-interpretatie verbeteren	120
27-09-2025	Gesprek De Kok – planning & structuur	30
11-10-2025	Feedback verwerken & testresultaten analyseren	60
11-10-2025	Inleiding herschrijven	90
12-10-2025	HMM-tabellen genereren & filtering toepassen	250
12-10-2025	Methodologie & bijwerken	100
22-10-2025	Diagrammen verbeteren + leren basis stochastiek	45
09-11-2025	HMM toepassen op volledige dataset	180
10-11-2025	Gesprek De Kok – inhoud	30
12-11-2025	Bronnenlijst opschonen	240
12-11-2025	Presentatie: kernboodschap opstellen	45
13-11-2025	Bronnen controleren	45
13-11-2025	Presentatie: storyline bouwen	45
13-11-2025	Literatuur uitbreiden	120
14-11-2025	Modellering uitschrijven (EGARCH/HMM/RP)	120
14-11-2025	Eerste slides maken	60
15-11-2025	Resultaten schrijven (EGARCH & regimes)	60
15-11-2025	Leren over monte-carlo	45
16-11-2025	Technische feedback verwerken	60
16-11-2025	Presentatie voorbereiden	45
16-11-2025	Resultaten uitbreiden (Risk Parity + regimes)	320
17-11-2025	Literatuur, discussie & conclusie afronden	330
17-11-2025	Presentatie: script & steekwoorden	45
17-11-2025	Modelverbeteringen implementeren	110
18-11-2025	Vorbereiding op de vragen	60
18-11-2025	Literatuurlijst doornemen	150
19-11-2025	Modelinconsistenties oplossen	210
19-11-2025	Hardop oefenen presentatie	45
20-11-2025	Bijlagen afronden (code & figuren)	60
20-11-2025	Slides finetunen	60
21-11-2025	PWS volledig doorlezen	180
21-11-2025	Twee oefenrondes presentatie	60
22-11-2025	Aanpassingen presentatie	30
22-11-2025	Feedback verwerken	45
23-11-2025	Presentatie afmaken	60
23-11-2025	Moeilijke stukken oefenen	60
24-11-2025	Commentaar verwerken	60
24-11-2025	Q&A simulatie	60
25-11-2025	Oefenen presentatie	300
25-11-2025	Gesprek De Kok – eindfeedback	30
25-11-2025	Presentatie oefenen	45
26-11-2025	Laatste oefening voor presentatie	240
30-11-2025	PWS perfectioneren (figuren, modellen, APA, overleaf)	400

Totale tijd: 13.345 minuten = 222,4 uur