

Implementatie grootzakelijke contracten en ALM voor Zwitserleven

Stageverslag

Rianne Lurvink
v 1.1 30 juli 2007

vrije Universiteit *amsterdam*
Faculteit der Exacte Wetenschappen
Studierichting Bedrijfswiskunde en Informatica
De Boelelaan 1081a
1081 HV Amsterdam

Stagebedrijf:
Zwitserleven
Burgemeester Rijnderslaan 7
1185 MD Amstelveen

Voorwoord

Het afsluitende gedeelte van de studie Bedrijfswiskunde en Informatica is een stage van minimaal zes maanden. Het doel van de stage is om een probleem van een bedrijf te onderzoeken, te analyseren, en zo mogelijk op te lossen. Belangrijk is om inzicht te verkrijgen in het functioneren van een organisatie. Ik heb mijn afstudeerstage bij ZwitserLeven in Amstelveen gelopen. Ik wil ZwitserLeven dan ook bedanken voor de mogelijkheid die ze mij hebben geboden om mijn afstudeerstage bij hen uit te voeren.

In het bijzonder wil ik mijn verschillende begeleiders binnen ZwitserLeven bedanken. Allereerst Wendy Wes voor het begeleiden van het projectgedeelte en Steve Djadoenath voor het begeleiden van het ALM gedeelte van dit verslag. Daarnaast wil ik ook Jitske Meijering en Tonny Verbaken bedanken voor de nuttige feedback op het totale verslag. Graag zou ik ook Dennis Roubos en André Ran, mijn begeleiders van de Vrije Universiteit, willen bedanken voor het commentaar op dit verslag.

Rianne Lurvink
Amstelveen, juli 2007

Management Summary

Implementatie grootzakelijke contracten

Vertrouwelijk.

Asset Liability Management

ALM is de afkorting voor Asset Liability Management. Deelnemers van een pensioenverzekering bouwen ieder jaar pensioenrechten op die gefinancierd moeten worden. De financiering vindt plaats door premiebetaling en het rendement op de portefeuille. Op een bepaald moment moeten er uitkeringen worden gedaan. Deze uitkeringen worden ook wel de verplichtingen genoemd. Het mag niet voorkomen dat op enig moment niet aan de verplichtingen kan worden voldaan. De toezichthouder van het fonds (DNB) houdt dit scherp in de gaten. Er zijn echter een aantal onbekende factoren in het spel. Zo is het niet duidelijk hoe de activa zich ontwikkelen in de loop der tijd. Het rendement op de activa is afhankelijk van het gekozen investeringsbeleid. Aan de andere kant zijn de toekomstige verplichtingen ook niet bekend. Hoeveel mensen gaan er met pensioen? Wat zijn de verwachte salarisontwikkelingen? Wat zijn de toekomstige indexatieverplichtingen? Hoeveel mensen overlijden er? De totale verplichtingen zijn hier afhankelijk van. Wat is dan het optimale investeringsbeleid? Hoe hoog moet de premie zijn?

Dit vraagstuk wordt ALM genoemd. ALM studies worden vaak gedaan voor pensioenfondsen om de ontwikkeling van de verschillende onderdelen op de balans te modelleren bij verschillende macro-economische scenario's. Bij het ontwikkelen van deze scenario's moeten er dus modellen worden opgesteld om voorspellingen te kunnen doen voor bijvoorbeeld de aandelenkoersen, de inflatie en de obligatiekoersen. Hoe de waarde van de activa zich ontwikkelt is afhankelijk van de voorspellingen van de macro-economische variabelen. Ook voor grootzakelijke contracten waarbij de beleggingsportefeuille gesepareerd beheerd wordt, is dit een belangrijk vraagstuk. Tijdens de sales fase is het belangrijk dat aan potentiële klanten getoond kan laten worden wat de ontwikkeling van de dekkingsgraad is bij een gekozen beleggingsbeleid en welke beleidsruimte dit geeft voor het verlenen van indexatie.

In dit verslag zijn twee verschillende typen ALM modellen beschreven. Dit zijn stochastisch lineair programmering (SLP) modellen en simulatiemodellen. Bij een SLP probleem wordt een doelstellingsfunctie geoptimaliseerd om zo een optimaal beleid voor een pensioenfonds te bepalen. Een nadeel is dat een SLP model snel erg complex wordt. Bij een simulatiemodel zijn er twee losstaande zaken: de economische scenario generator en het optimalisatiemodel. Dit is dus verschillend van SLP modellen, waar zowel de scenario's en het optimalisatieprobleem in één model worden gestopt. SLP modellen zijn gericht op het optimaliseren van een beleid terwijl simulatiemodellen gericht zijn op het evalueren van een beleid, en door het doorrekenen van verschillende strategieën streven naar een optimaal beleid.

Dit onderzoek heeft zich gericht op het modelleren van macro-economische variabelen: de inflatie, de kort- en langlopende obligatie index en het rendement op de AEX index. Om de inflatie en de kort- en langlopende obligatie index te modelleren is er gebruik gemaakt van verschillende modellen: VAR modellen, ARIMA modellen en Holt Winters (Exponential Smoothing). Wanneer de performance van deze modellen wordt getest op een deel van de dataset, blijkt dat het VAR(1) model het beste model is voor de gekozen dataset en bij de gebruikte technieken. Er blijkt dat een VAR(1) model met daarin ook de AEX index opgenomen slechtere voorspellingen geeft voor de inflatie, kort- en langlopende obligatie index, dan een VAR(1) model met daarin alleen de inflatie, kort- en

langlopende obligatie index. Er dient verder onderzoek te worden gedaan naar het vinden van een bevredigend model waarin wel de correlatie tussen aandelen en de andere factoren is meegenomen.

Het optimale model voor het aandelenrendement blijkt een Regime Switching Lognormal model te zijn: RSLN(2). Er dient echter verder onderzoek te worden gedaan naar de voorspelling prestaties voor het aandelenrendement van het RSLN(2) model ten opzichte van het GARCH(1,1) model. Ook blijkt het VAR(1) model met daarin de AEX index opgenomen niet beter te voorspellen dan het GARCH(1,1) of RSLN(2) model. Hier moet echter uitgebreider onderzoek naar worden gedaan.

Met de gegenereerde scenario's is het dan mogelijk om, voor grootzakelijke contracten met een gesepareerd beleggingsdepot, de kans te bepalen dat de dekingsgraad onder een bepaald niveau komt, de indexatieruimte te bepalen en het effect van andere activa percentages op de dekingsgraad te bepalen. Een simpele opzet voor zo'n model is gegeven.

Inhoudsopgave

Voorwoord.....	3
Management Summary.....	5
Lijst met gebruikte afkortingen en definities.....	9
1. Introductie.....	13
1.1 Zwitserleven.....	13
1.2 Probleemstelling.....	14
1.2.1 Probleemstelling projectgedeelte.....	14
1.2.2 Probleemstelling ALM gedeelte.....	14
1.3 Pensioenen in het kort.....	14
1.3.1 Pensioenen algemeen.....	14
1.3.2 Pensioenfondsen.....	15
2. Implementatie van grootzakelijke contracten binnen Zwitserleven.....	17
3. ALM algemeen.....	19
3.1 ALM.....	19
3.1.1 Geschiedenis van ALM.....	20
3.1.2 Waarom ALM studies steeds belangrijker worden.....	20
3.2 Verschillende ALM modellen.....	22
3.2.1 SLP modellen.....	22
3.2.2 Simulatiemodellen.....	22
4. Stochastic Linear Programming.....	23
4.1 Optimalisatiemodel SLP.....	23
4.2 Het genereren van scenario's.....	24
4.2.1 De toekomstige waarde van de activa.....	24
4.2.2 De verplichtingen en toekomstige pensioenuitkeringen.....	25
4.3 De uitgangspunten van een SLP model.....	26
4.3.1 De beslissingsvariabelen.....	26
4.3.2 De restricties van het model.....	26
4.3.3 Doelstellingsfunctie.....	27
4.4 Voordelen, nadelen en conclusie.....	27
5. Simulatiemodellen.....	29
5.1 Optimalisatiemodel simulatiemodel.....	29
5.2 Economische scenario generator.....	30
5.2.1 Bekende economische scenario generatoren.....	30
5.2.2 Scenario's voor economische grootheden genereren.....	31
5.2.3 Scenario's voor de actuariële grootheden.....	31
5.3 ALM tools.....	32
6. Data analyse en modelkeuze voor economische scenario's.....	33
6.1 Gebruikte data.....	33
6.2 Definities.....	33
6.3 Stationaire data verkrijgen.....	35
6.3.1 Small trend methode.....	35
6.3.2 Differencing van een tijdserie.....	35
6.4.3 Het bepalen van de seizoensinvloeden.....	36
6.4 Modellen voor tijdseries.....	37
6.4.1 AR model.....	37
6.4.2 MA model.....	37
6.4.3 VAR model.....	38
6.4.4 AR(I)MA model.....	38
6.4.5 Exponential Smoothing.....	38
6.4.6 Economische scenario generator van Triple A.....	40
6.5 Data analyse.....	41
6.6 Modelleren van de inflatie, lang- en kortlopende obligatie-index.....	43
6.6.1 ARIMA model.....	43

6.6.2 VAR model	51
6.6.3 Holt Winters methode	59
6.7 Conclusie: welk model is het best geschikt om scenario's te genereren voor de inflatie, kort- en langlopende obligatie-index?	60
6.8 Het modelleren van het rendement op de AEX index	62
6.8.1 Aandelenrendementen beschrijven met een normale verdeling	62
6.8.2 Aandelenrendementen beschrijven met een lognormale verdeling	63
6.8.3 Aandelenrendementen beschrijven met een AR(1) model	64
6.8.4 Tijdsvariërende variantie	64
6.8.5 Aandelenrendementen beschrijven met een ARCH proces	65
6.8.6 Aandelenrendementen beschrijven met het Markov Regime Switching Lognormal model (RSLN)	66
6.9 Conclusie: welk model is optimaal om het aandelenrendement te voorspellen?	67
7. Opzet model om de dekingsgraad te bepalen bij gegenereerde economische scenario's	69
8. Verder onderzoek	71
Conclusie	73
Literatuurlijst	75

Lijst met gebruikte afkortingen en definities

Afkortingen

ZL	Zwitserleven
ALM	Asset Liability Management
AOW	Algemene Ouderdomswet
OP	Ouderdomspensioen
AOP	Arbeidsongeschiktheidspensioen
NP	Nabestaandenpensioen
FTK	Financieel Toetsingskader
DNB	De Nederlandsche Bank
SLAM	Swiss Life Asset Management
T&I	Technologie en Implementatie
M&S	Marketing en Sales
CC	Corporate clients
CCR	Corporate clients Rekentechniek
MVCX	Sales Corporate Clients Experts
MVCE	Sales Account Executive Corporate Clients
MVCC	Sales Corporate Clients
MMCC	Market Corporate Clients
BO	Back Office
PV	Private View
PE	proceseigenaar
PL	projectleider
TM	teammanager
RM	relatiemanager
PM	project manager
KPI's	Key Performance Indicators

Wiskundige afkortingen

VAR(p)	Vector Autoregressief Model van de orde p
PaR	Pension at Risk
AR(p)	Autoregressief model van de orde p
MA(q)	Moving Average model van de orde q
ARIMA(p, q)	Autoregressief (Integrated) Moving Average model
OLS	Ordinary Least Squares
GARCH	Generalized Autoregressief Conditioneel Heteroscedasticity process
RSLN	het Markov Regime Switching Lognormal model
RMSE	Root Mean Squared Error

Gebruikte definities

T	tijdsperiode
t	lengte subperiodes, $0 \leq t \leq T$
α	de minimum hoogte van de dekkinggraad
η_t	de vereiste kans dat de dekkinggraad groter is dan een bepaalde waarde
A^*_{t+1}	de waarde van de activa vlak voor er een extra storting wordt gedaan
L^*_{t+1}	de waarde van de verplichtingen vlak voor er een extra storting wordt gedaan

F_{t+1}^*	stochast die de dekkingsgraad aangeeft vlak voor een extra storting wordt gedaan, $F_{t+1}^* = A_{t+1}^* / L_{t+1}^*$
β_t	een bovengrens aan het verwachte tekort
(t,s)	de staat waarin het model zich bevindt op tijdstip t en in scenario s
p_s	de kans dat scenario s voorkomt
$O_k(t)$	opgebouwde rechten van deelnemer k op tijdstip t
d_k	het aantal jaren dat deelnemer k , rechten heeft opgebouwd
$S_k(t)$	het salaris van deelnemer k op tijdstip t
$PP(t)$	franchise op tijdstip t
$L_k(t)$	netto contante waarde van de verwachte pensioenuitkeringen
r	de rekenrente
$p_{l_{kt}}^{(j)}$	de kans dat een persoon met leeftijd l_{kt} nog steeds levend is over j jaar
ε_t	error term
c	een constante
A_i	de autocorrelatie matrix
φ_i	de parameters van het model
X_t	de vector voor k verschillende variabelen, $X_t = [X_{1t}, X_{2t}, \dots, X_{kt}]'$.
Σ_ε	de covariantie matrix tussen de error termen van de verschillende variabelen in de vector X
$\gamma_X(t',t)$	de autocovariantie functie, $\gamma_X(t',t) = \text{Cov}(X_{t'}, X_t) = E[(X_{t'} - E[X_{t'}])(X_t - E[X_t])]$
$\gamma_X(h)$	De autocovariantie functie van een stationair proces, $\gamma_X(h) = \gamma_X(h,0) = \text{Cov}(X_{t+h}, X_t)$. Voor een multivariaat proces geldt: $\gamma_X(h) = (\text{cov}(X_{t+h,i}, X_{t,j}))_{i,j=1,\dots,d} = E(X_{t+h} - EX_{t+h})(X_t - EX_t)^T$
$\rho_X(h)$	De autocorrelatie functie voor lag h van een stationair proces $\{X_t\}$, $\rho_X(h) = \text{Cor}(X_{t+h}, X_t) = \frac{\gamma_X(h)}{\gamma_X(0)}$. Voor een multivariaat proces geldt: $\rho_X(h) = (\rho(X_{t+h,i}, X_{t,j}))_{i,j=1,\dots,d} = \left(\frac{\gamma_X(h)_{i,j}}{\sqrt{\gamma_X(0)_{i,i} \gamma_X(0)_{j,j}}} \right)_{i,j=1,\dots,d}$
$r(t)$	de t -jarige yield op obligaties
β_0, β_1	de parameters van het Nelson Siegel model, β_0 wordt gezien als de langlopende rente, en $\beta_0 + \beta_1$ als de kortlopende rente
τ en β_2	vormparameters van het Nelson Siegel model
P	de matrix die ontstaat na een Cholesky factorisatie: $\Sigma_\varepsilon = PP'$
I	identiteitsmatrix
$X_t = m_t + s_t + Y_t$	Small trend methode om trend- en seizoenscomponenten te verwijderen
m_t	vaste trend component
s_t	vaste seizoenscomponent met bekende periode d
Y_t	de random error term die stationair verondersteld wordt

w_i zijn de gewichten die zijn toegekend aan de vorige observaties
 α de smoothing constante
 l_t constante op tijdstip t

1. Introductie

In dit hoofdstuk zal Zwitserleven worden geïntroduceerd. Daarna wordt de probleemstelling gegeven. De probleemstelling bestaat uit twee afzonderlijke delen. In de afsluitende paragraaf van dit hoofdstuk worden in het kort pensioenen besproken.

1.1 Zwitserleven

Zwitserleven (volledige naam Zwitserse Maatschappij van Levensverzekering en Lijfrente) is de Nederlandse vestiging van het internationale Swiss Life concern. Zwitserleven is sinds 1901 gevestigd in Nederland en is de grootste buitenlandse levensverzekeraar. Via tussenpersonen biedt Zwitserleven klanten oplossingen voor vermogensbouw en inkomstenbescherming. Hieronder vallen pensioenen, levensverzekeringen, spaarplannen, hypotheek en beleggingsfondsen. Het hoofdkantoor van Zwitserleven is gevestigd in Amstelveen. Tevens zijn er nog twee zogenoemde advieslounges, die zich in Scheveningen en Amsterdam bevinden. Zwitserleven heeft zo'n 795 personeelsleden, georganiseerd in de afdeling Algemeen, de divisies Customer Services, Financiën, Marketing en Sales en Technologie en Implementatie.

Wereldwijd heeft Swiss Life ongeveer 9.000 medewerkers. Swiss Life is opgericht op 28 september 1857 door Conrad Widmer in Zürich. Dit maakt Swiss Life de oudste levensverzekeringsmaatschappij van Zwitserland. Swiss Life heeft kantoren, vertegenwoordigingen en partners in zo'n 40 landen. Zwitserleven had in 2006 een premie-inkomen van 1,2 miljard euro. Dit is een stijging van 16% ten opzichte van 2005. De nettowinst nam in 2006 toe van 67 miljoen naar 155 miljoen euro [14].

Kijkend naar de totale verzekeringsbranche is Zwitserleven nummer 8 met een marktaandeel van 2,8%. Achmea vormt de absolute leider met een marktaandeel van 22%, ING volgt met 18,1%, Fortis met 11,7%, Delta Lloyd met 10%, Reaal/AXA en Aegon met beiden 9%, daarna volgt Allianz met 3,8%. Hierbij moet worden vermeld dat deze andere verzekeringsmaatschappijen ook schadeverzekeringen aanbieden, terwijl Zwitserleven alleen levensverzekeringen aanbiedt.

In de collectieve levensverzekeringsmarkt had Zwitserleven in 2005 een marktaandeel van 7%, dit is de 5^e positie.

1.2 Probleemstelling

1.2.1 Probleemstelling projectgedeelte

Vertrouwelijk.

1.2.2 Probleemstelling ALM gedeelte

Deelnemers van een pensioenverzekering bouwen ieder jaar pensioenrechten op die gefinancierd moeten worden. De financiering vindt plaats door premiebetaling en het rendement op de portefeuille. Op een bepaald moment moeten er uitkeringen worden gedaan. Deze uitkeringen worden ook wel de verplichtingen genoemd. Het mag niet voorkomen dat op enig moment niet aan de verplichtingen kan worden voldaan. Aan de andere kant moet een zo hoog mogelijk rendement op de activa worden verkregen. Wat is dan het optimale beleggingsbeleid? Hoe hoog moet de premie zijn? Met dit vraagstuk houdt ALM (Asset Liability Management) zich bezig.

Voor grootzakelijke contracten waarbij de beleggingsportefeuille gesepareerd beheerd wordt, is dit een belangrijk vraagstuk. Dit geldt dus ook voor de nieuwe afgesloten grootzakelijke contracten.

Er zijn softwarepakketten op de markt om dit soort ALM studies te verrichten. Deze werken vaak op basis van economische scenario's. Een economisch scenario is een toekomstige ontwikkeling van economische tijdreeksen zoals het aandelenrendement, rentestand, enzovoorts. Er worden veel van deze scenario's gegenereerd. Het gekozen beleid wordt dan toegepast op deze scenario's en zo wordt geëvalueerd hoe goed het gekozen beleid is. Het is dus heel belangrijk om een goede set scenario's te kunnen genereren. Zwitserleven is in gesprek met aanbieders van economische scenario generatoren. Hoe is het mogelijk om zelf economische scenario's te genereren? Er zijn een aantal subonderzoeksvragen opgesteld.

- Welke verschillende ALM modellen kunnen worden onderscheiden?
- Welke verschillende modellen kunnen worden gebruikt om economische scenario's te genereren?
- Wat zijn de verschillen in uitkomsten tussen deze modellen?
- Welk model kan als optimaal worden beschouwd?
- Hoe kunnen deze economische scenario's vervolgens gebruikt worden?

1.3 Pensioenen in het kort

Dit hoofdstuk geeft een korte inleiding in pensioenen. De verschillende pensioenvormen en de verschillende pensioenregelingen zullen worden besproken. In de tweede paragraaf zal het een en ander over pensioenfondsen worden verteld. De lezer met verstand van pensioenen zal dit hoofdstuk naar alle waarschijnlijkheid kunnen overslaan.

1.3.1 Pensioenen algemeen

Met pensioen wordt niet alleen de uitkering bedoeld die gedaan wordt na het bereiken van de 65 jarige leeftijd. Ook de uitkering die de weduwe, weduwnaar of partner van een overleden persoon ontvangt behoort hiertoe, net zoals de uitkering die de kinderen van een overleden persoon ontvangen. Bij velen is niet bekend dat pensioen ook de uitkering betreft die gedaan wordt wanneer een persoon arbeidsongeschikt wordt voordat hij de pensioengerechtigde leeftijd bereikt heeft.

De uitkering die ontvangen wordt na het bereiken van de pensioengerechtigde leeftijd (op dit moment 65 jaar), bestaat uit twee delen. Allereerst de AOW (Algemene ouderdomswet). Iedereen die vanaf 15 jarige leeftijd in Nederland woont heeft recht op AOW.

Ten tweede de pensioentoezegging door de werkgever. In een Pensioenovereenkomst worden de pensioentoezeggingen door de werkgever vastgelegd. In de Pensioenovereenkomst staat onder andere vermeld welke toetredingseisen er zijn, wat de hoogte van het pensioen is, het aantal pensioenvormen, enzovoorts. Het toezeggen van een pensioen is een steeds belangrijker onderdeel van de arbeidsvoorwaarden geworden.

Een mogelijke derde uitkering bij ouderdom is een vrijwillige pensioenvoorziening. Deze vorm is voornamelijk bedoeld voor personen die geen vast dienstverband hebben of personen waarvan het pensioen ontoereikend is.

Enkele pensioenvormen die bestaan zijn het OP (ouderdomspensioen), het prépensioen, het AOP (arbeidsongeschiktheidspensioen) en het NP (nabestaandenpensioen). Welke pensioenvormen er precies verzekerd zijn hangt af van de Pensioenovereenkomst.

Er zijn verschillende pensioenregelingen of combinaties van pensioenregelingen bekend. De twee bekendste pensioenregelingen zijn:

- *Defined Benefits* (uitkeringsovereenkomst)
Het verzorgende karakter staat centraal, het pensioen is afhankelijk van het salarisverloop en het aantal dienstjaren. Dit betreft een uitkeringsovereenkomst, de werkgever zegt een uitkering toe. De uitkering is bekend, de premie niet. Het risico ligt niet bij de deelnemer. Er zijn twee vormen te onderscheiden, te weten het middellood en het eindlood. Bij het middellood wordt ieder jaar een percentage over het verdiende salaris in het betreffende jaar gebruikt als pensioenopbouw. Bij eindlood is het uiteindelijke pensioen afhankelijk van het laatstverdiende salaris en het aantal dienstjaren. Wanneer het salaris van een werknemer stijgt, moet de werkgever over de achterstallige jaren pensioen inkopen. Doordat het uiteindelijke pensioen afhankelijk is van het salarisverloop en het aantal dienstjaren, kan dit leiden tot gigantische premieverhogingen. Het eindlood is niet per definitie beter dan het middellood, bijvoorbeeld wanneer het salaris afneemt door minder werken. Ook kan middellood beter zijn dan eindlood wanneer de opgebouwde rechten bij middellood geïndexeerd worden ter bescherming tegen inflatie en bij eindlood niet.
- *Defined Contribution* (premieovereenkomst)
Het belonende karakter staat centraal. Dit betreft een premieovereenkomst, de werkgever zegt een premie toe. De premie is dus bekend, maar de uitkering is onbekend. Het risico van onvoldoende rendement is geheel voor de deelnemer.

Het Nederlandse pensioenstelsel is een van de beste ter wereld. 90% van de beroepsbevolking heeft pensioentoezeggingen door de werkgever. Dit bijvoorbeeld in vergelijking met Griekenland, waar slechts 5% van de beroepsbevolking via zijn werkgever pensioenrechten opbouwt. Dit verschil kan worden verklaard door het feit dat het in Nederland verplicht is om toe te treden tot een pensioenregeling. Landen waar dit ook verplicht is zijn bijvoorbeeld Denemarken en Zweden, waar het percentage ook hoog ligt (respectievelijk 80% en 90%) [9].

1.3.2 Pensioenfondsen

Een pensioenfonds is een organisatie waarvan het doel is om deelnemers na pensionering van een uitkering te voorzien. In ruil daarvoor hebben de deelnemers jarenlang pensioenpremie voldaan. Er zijn een aantal typen pensioenfondsen. Hier noemen we bijvoorbeeld een bedrijfstakpensioenfonds en een ondernemingspensioenfonds. Een werkgever kan er voor kiezen om toe te treden tot een bedrijfstakpensioenfonds of ondernemingspensioenfonds, of hij kan ervoor kiezen om verzekeringsvoorzieningen te treffen.

Een pensioenfonds kent een aantal verschillende vormen. Allereerst kan een pensioenfonds ervoor kiezen om eigen beheer te voeren. Er moet dan voldaan worden aan de eisen van het FTK (Financieel Toetsingskader). Een pensioenfonds mag alleen zelfstandig een fonds vormen wanneer het volgens DNB (De Nederlandsche Bank) voldoende draagvlak heeft voor het zelfstandig dragen van het risico. Ook kan een pensioenfonds ervoor kiezen om een kapitaalcontract of een garantiecontract af te sluiten bij een verzekeraar. Een kapitaalcontract houdt in dat alle beleggingen aan het einde van de contractperiode worden overgenomen van de verzekeraar en de mogelijke tekorten moeten worden aangevuld door het pensioenfonds, hier moet ook worden voldaan aan de eisen van het FTK. Bij een garantiecontract komen eventuele tekorten voor rekening van de verzekeraar. Een bedrijf dat niet zelf een pensioenfonds opricht, kan rechtstreeks een garantiecontract afsluiten bij een verzekeraar. De laatste drie groepen (pensioenfonds met kapitaalcontract, pensioenfonds met garantiecontract en een relatief groot bedrijf met een rechtstreeks verzekerd garantiecontract) hebben bij een verzekeraar veelal een gesepareerd beleggingsdepot. Dit worden SA's (Separate Accounts) genoemd. Wanneer er in dit verslag gesproken wordt over pensioenfondsen worden hier automatisch ook SA's mee bedoeld.

2. Implementatie van grootzakelijke contracten binnen ZwitserLeven

Vertrouwelijk.

3. ALM algemeen

In dit hoofdstuk zal Asset Liability Management (ALM) worden geïntroduceerd. Allereerst zal er worden uitgelegd wat ALM inhoudt, en vervolgens wordt kort de geschiedenis van ALM besproken. Daarna zal het nut van ALM in de context van de nieuwe Pensioenwet worden geschetst. Dit hoofdstuk wordt afgesloten met een korte beschrijving van twee fundamenteel verschillende technieken voor ALM studies, die in verdere hoofdstukken zullen worden uitgediept.

3.1 ALM

Deelnemers van een pensioenfonds of deelnemers van een rechtstreekse verzekeringsregeling bouwen ieder jaar pensioenrechten op die gefinancierd moeten worden. De financiering vindt plaats door premiebetaling en het rendement op de portefeuille. Omdat het pensioenfonds op een bepaald moment uitkeringen moet gaan doen, en dus verplichtingen heeft (de liabilities of passiva van het fonds), moeten de activa van het fonds goed beheerd worden. Het mag niet gebeuren dat het fonds op een bepaald moment niet aan zijn verplichtingen kan voldoen. De toezichhouder van het fonds (DNB) houdt dit scherp in de gaten. Er zijn echter een aantal onbekende factoren in het spel. Zo is het niet duidelijk hoe de activa zich ontwikkelen in de loop der tijd. Het rendement op de activa is afhankelijk van het gekozen investeringsbeleid. Aan de andere kant zijn de toekomstige verplichtingen ook niet bekend. Hoeveel mensen gaan er met pensioen? Wat zijn de verwachte salarisontwikkelingen? Wat zijn de toekomstige indexatieverplichtingen? Hoeveel mensen overlijden er? De totale verplichtingen zijn hier afhankelijk van. Er kan wel iets over worden gezegd door middel van kansverdelingen en verwachte waarden.

ALM is een term die wordt gebruikt in de beleggingswereld. ALM kan worden omschreven als het verkrijgen van inzicht in de onderlinge afhankelijkheden in de ontwikkelingen van de beleggingen en verplichtingen in een organisatie. ALM studies worden vaak gedaan voor pensioenfondsen en SA's om de ontwikkeling van de verschillende onderdelen op de balans te modelleren bij verschillende macro-economische scenario's. Bij het ontwikkelen van deze scenario's moeten er dus modellen worden opgesteld om voorspellingen te kunnen doen voor bijvoorbeeld de aandelenkoersen, de inflatie en de obligatiekoersen. Hierbij geldt voor een pensioenfonds dat de activa kant van de balans bestaat uit de beleggingsportefeuille en de passiva kant uit de verplichtingen (dus de in de toekomst te betalen pensioenen) en een gedeelte eigen vermogen (het surplus genoemd). De vraag is dus hoe deze onderdelen zich ontwikkelen in de loop der tijd terwijl de toekomst onzeker is. Hoe de waarde van de activa zich ontwikkelen is afhankelijk van de voorspellingen van de macro-economische variabelen. Hoe de waarde van de passiva zich ontwikkelen is afhankelijk van overgangskansen. Hier wordt dan bijvoorbeeld de kans dat een actieve deelnemer arbeidsongeschikt raakt of overlijdt mee bedoeld [5]. Een actieve deelnemer is een deelnemer die nog pensioen opbouwt en premie betaalt. Een belangrijk gegeven is de dekkingsgraad van een pensioenfonds. De dekkingsgraad is gedefinieerd als de activa gedeeld door de passiva. De dekkingsgraad wordt gebruikt als graadmeter voor de solvabiliteit van een fonds. De overheid schrijft een solvabiliteitsmarge voor. Dit houdt in dat verzekeraars moeten beschikken over een aansprakelijk vermogen van ten minste een bepaalde omvang. Bij ieder scenario dat wordt doorgerekend kan de dekkingsgraad worden uitgezet tegen de tijd. Aan de hand van bijvoorbeeld een simulatie kan dan worden aangegeven wat de kans is dat de dekkingsgraad onder een bepaald niveau komt te liggen. Er bestaan wettelijke eisen voor de hoogte van de dekkingsgraad van een pensioenfonds. ALM houdt zich vooral bezig met wat optimale investeringsstrategieën zijn waarbij aan bepaalde financiële verplichtingen moet worden voldaan. Daarom kan

het belang van ALM niet worden onderschat [1]. Niet alleen het investeringsbeleid is van belang. Andere belangrijke beleidsvariabelen zijn de hoogte van de premies, de indexaties van de pensioenaanspraken en bij welke dekkingsgraad er extra moet worden bijgestort door de sponsor van het fonds. Indexatie houdt het corrigeren van de opgebouwde rechten in zodat ze niet minder waard worden vanwege inflatie [2]. Het investeringsbeleid is erg belangrijk. Zo geldt dat wanneer er tijdens het werkende leven van de deelnemers 1% meer aan opbrengsten wordt gegenereerd uit de activa, het pensioen ongeveer 30% hoger is (afhankelijk van de toezegging). [3]

3.1.1 Geschiedenis van ALM

De eerste ALM modellen werden ontwikkeld in de jaren '80. Tijdens deze periode ontstond er een scheve verhouding tussen de activa en passiva vanwege de volatiele rente. Dit leidde tot het beheersen van het zogenaamde asset liability risk. De vraag bij het beheersen van dit risico is in hoeverre de activa en passiva gelijke tred houden. In de praktijk is dit vaak niet het geval. Daarom kan een kleine percentuele wijziging in de activa of passiva leiden tot grote verschillen in het surplus.

De eerste modellen die werden gebruikt om dit risico te beheersen zijn deterministische modellen en duration matching technieken. Deze worden gebruikt om de beste portfolio te vinden. Bij deterministische modellen wordt de stroom van toekomstige pensioenuitkeringen van tevoren met zekerheid bekend verondersteld. Deze modellen beschouwen alleen obligaties als mogelijke investering. Tot de jaren '80 waren deze modellen bruikbaar. Na deze tijd werden veelal obligatie modellen gebruikt. Hierin wordt de toekomstige stroom van pensioenuitkeringen stochastisch verondersteld. Alternatieve portfolio's werden gevonden door duration matching technieken. Onder de duration van bijvoorbeeld een obligatie wordt het gewogen gemiddelde van de uitkeringen verstaan. Hierbij tellen jaren waarin een grotere kasstroom plaatsvindt zwaarder mee dan jaren waarin een kleinere kasstroom plaatsvindt. Bij duration matching technieken wordt met behulp van de duration de vooraf vastgestelde verplichtingen gekoppeld aan de inkomsten uit de obligatieportefeuille. Deze methode heeft echter een aantal nadelen. Er worden bijvoorbeeld problemen verwacht wanneer de interest onverwacht verandert. Vanwege de nadelen die kleven aan bovenstaande modellen, maken verzekeringsmaatschappijen tegenwoordig veelal gebruik van scenario analyse. In een scenario krijgen economische grootheden zoals de rentecurve, aandelenkoersen en onroerend goed koersen, een waarde. Er wordt een groot aantal scenario's gesimuleerd voor een periode van bijvoorbeeld twintig jaar [8], [9], [12]. Zwitserleven maakt echter nog geen gebruik van scenario analyse.

3.1.2 Waarom ALM studies steeds belangrijker worden

Pensioenverzekeraars en levensverzekeraars moeten voldoen aan een aantal eisen. Deze worden hieronder beschreven.

- Er moeten technische voorzieningen aanwezig zijn. Voor elke gesloten verzekeringsovereenkomst moet een voorziening worden gevormd. Bij de berekening van deze voorziening moet rekening worden gehouden met toekomstige premies en verplichtingen.
- Naast de technische voorzieningen moeten verzekeraars ook beschikken over een aansprakelijk vermogen van een bepaalde omvang. Dit wordt de solvabiliteitsmarge genoemd. Wanneer er een eventueel tekort zou zijn in de technische voorziening, kan deze hierdoor worden opgevangen. Ook moet een garantiefonds aanwezig zijn. Er is een bedrag vastgesteld dat minimaal in het garantiefonds aanwezig moet zijn. Tevens is bepaald in welke activavormen het garantiefonds belegd mag worden.

- De technische voorzieningen moeten volledig door activa zijn gedekt. De activa moeten voor een groot deel in dezelfde muntsoort kunnen worden geïnd als de muntsoort van de verplichtingen. Als de activa technische voorzieningen binnen de Europese Unie dekken, moeten deze activa ook binnen de Europese Unie belegd zijn. Dit zijn dus beleggingsvoorschriften. Hierin worden ook maxima aangegeven voor bepaalde categorieën activa wat betreft de technische voorzieningen [22].

Sinds 1 januari 2007 is de nieuwe Pensioenwet ingegaan. De Pensioenwet vervangt de oude Pensioen- en spaarfondsenwet en beschrijft de regelgeving van het Financieel Toetsingskader. Het doel van de Pensioenwet is het veiligstellen van de pensioenaanspraken van de deelnemers. De nieuwe wet zorgt voor herziening en modernisering van de Pensioenwetgeving. Een belangrijk aspect van de Pensioenwet is dat de deelnemers het recht hebben om te weten waar ze aan toe zijn met hun pensioen. Ze hebben recht op duidelijke en begrijpelijke voorlichting door hun verzekeraar. Dit wordt ook wel de zorgplicht genoemd.

De groep deelnemers met een premieovereenkomst (premie bekend, uitkering niet) kan worden opgedeeld in twee subgroepen. Er zijn deelnemers met beleggingsvrijheid en deelnemers zonder beleggingsvrijheid. Alhoewel de groep met beleggingsvrijheid zelf verantwoordelijk is, dient de verzekeraar aan de zorgplicht te voldoen. Dit houdt in dat er advies en informatie moet worden gegeven en controle moet plaatsvinden. Het beleggingsrisico moet scherp worden bewaakt, aangezien dit geld moet resulteren in een levenslang ouderdomspensioen. Voor iedere deelnemer moet aan de hand van zijn risico aversie grenzen voor de beleggingsvrijheid worden vastgelegd. Zwitserleven moet dan in het kader van de zorgplicht jaarlijks bekijken of de vastgelegde grenzen niet worden overschreden. Zwitserleven geeft dus slechts advies, of de deelnemer ook daadwerkelijk actie onderneemt na dit advies is zijn eigen verantwoordelijkheid.

Bij de deelnemers zonder beleggingsvrijheid ligt de verantwoordelijkheid geheel bij Zwitserleven. Hiertoe kan Zwitserleven beleggen in zogenaamde Life Cycle fondsen. Aan de hand van de leeftijd van de deelnemer kan dan de optimale beleggingsmix worden bepaald. Hoe dichter de pensioendatum nadert, hoe minder risicovol de beleggingen mogen zijn. Het onderbouwd adviseren van de deelnemers moet volgens de Pensioenwet voor 1 januari 2008 gerealiseerd worden.

Voor de collectieve regelingen die onder de categorie uitkeringsovereenkomsten vallen, moet ook een onderbouwde beleggingsmix kunnen worden bepaald. Dit is belangrijk vanwege de eisen die gesteld worden aan pensioenverzekeraars en levensverzekeraars, zoals hierboven beschreven. Bij het bepalen van deze strategische asset allocaties wordt gebruik gemaakt van ALM studies. Om deze te verrichten is een goed ALM model noodzakelijk. Vanwege de Pensioenwet zijn ALM studies dus noodzakelijk geworden.

Niet alleen de Pensioenwet heeft het belang van ALM doen toenemen. Ook het instorten van de aandelenmarkt rond 2002 heeft hiermee te maken. Aangezien pensioenfondsen een groot deel van de activa belegd hadden in aandelen, had dit een grote invloed op de solvabiliteit van de fondsen. Er moet dus goed worden nagedacht over het investeringsbeleid van een pensioenfonds. Ook de sterke daling van de rentecurve in de afgelopen jaren heeft het belang van ALM doen toenemen. Er geldt namelijk dat wanneer de rente omlaag gaat, de waarde van de activa minder snel stijgt dan de waarde van de verplichtingen. Hierdoor daalt de dekkingsgraad.

Dus vanwege de eisen die gesteld worden aan pensioenfondsen en levensverzekeraars wat betreft de solvabiliteit, de nieuwe Pensioenwet, schommelende aandelenkoersen en onzekere rente is het belang van ALM nog groter geworden. Met behulp van een ALM

pakket is het mogelijk om een goed beleid te kiezen waarbij de onzekere factoren meegenomen worden in het model en de solvabiliteit beheerst wordt.

3.2 Verschillende ALM modellen

Er kunnen twee verschillende modellen worden onderscheiden om ALM studies te verrichten [9]. Deze twee modellen zijn SLP (stochastic linear programming) en simulatie. Hieronder zal een korte beschrijving van ieder van deze modellen worden gegeven en enkele voor- en nadelen worden besproken.

3.2.1 SLP modellen

Stochastic Linear Programming modellen zijn gericht op optimaliseren van een beleid. SLP zoekt naar de beste oplossing. Het beleid wordt dus ieder jaar opnieuw vastgesteld. Bij een SLP model wordt er rekening gehouden met onzekerheid (hoe verlopen de macro-economische scenario's). Het is een dynamisch model (meerdere tijdsperiodes, het aanpassen van de beslissingen later in de tijd). De voorwaarden van het ALM probleem worden geformuleerd in lineaire vergelijkingen. Dit zijn de grootste voordelen van een SLP probleem. Echter, de oplossingstijd van een dergelijk model kan erg oplopen, aangezien ze erg complex zijn. Wanneer het minder complex gemaakt wordt, dan zal het model erg afwijken van de werkelijkheid. Waarschijnlijk gaat het optimaliseren van een groot model via een heuristiek. Aangezien deze modellen theoretisch erg interessant zijn, zal er dieper op in worden gegaan in hoofdstuk 4. Een nadeel van SLP modellen is het beperkt aantal scenario's dat doorgerekend kan worden, aangezien het model anders erg complex wordt.

3.2.2 Simulatiemodellen

Een tegenhanger van SLP modellen zijn simulatiemodellen. Simulatiemodellen zijn gericht op het evalueren van een beleid, terwijl SLP modellen gericht zijn op het optimaliseren.

De eerste simulatiemodellen zijn eind jaren tachtig ontwikkeld. Deze modellen zijn erg populair omdat het mogelijk is om veel scenario's te genereren. Het is dus mogelijk om de financiële positie van een pensioenfonds te berekenen bij veel verschillende omstandigheden. In de volgende hoofdstukken zal het genereren van scenario's uitgebreid aan de orde komen. Scenario's zijn mogelijke ontwikkelingen van macro-economische variabelen, zoals inflatie, rente en de aandelenkoersen. Het voordeel is dat veel verschillende scenario's kunnen worden doorgerekend. Een nadeel van simulatiemodellen is dat veel keuzes wat betreft het beleid vast worden verondersteld. Er moeten dus veel verschillende beleidkeuzes worden doorgerekend voordat het optimale beleid gevonden is.

Simulatiemodellen zullen in hoofdstuk 5 behandeld worden.

4. Stochastic Linear Programming

Bij een SLP probleem wordt een doelstellingsfunctie geoptimaliseerd om zo een optimaal beleid voor een pensioenfonds te bepalen. Uit welke beslissingen bestaat dit beleid? Wat is deze doelstellingsfunctie, en welke voorwaarden gelden er? In dit hoofdstuk zullen deze vragen worden behandeld.

4.1 Optimalisatiemodel SLP

Een SLP probleem is een optimalisatiemodel, dus wordt het model geformuleerd met een doelstellingsfunctie en een aantal voorwaarden. In het komende paragrafen zullen de uitgangspunten van een SLP model kort worden besproken: de beslissingsvariabelen worden geïntroduceerd, alsmede de doelstellingsfunctie en de restricties van het model. Aangezien het doel hier niet is om een volledig model te beschrijven, zullen alleen de uitgangspunten worden behandeld. Voor een compleet SLP model wordt de lezer verwezen naar een van de referenties, bijvoorbeeld [2] en [5].

Een ALM studie beschouwt een bepaalde tijdsperiode T . Deze tijdsperiode wordt vaak opgedeeld in een aantal subperiodes van lengte t . Deze subperiodes hebben vaak een lengte van 1 jaar. Ieder jaar moeten beslissingen worden genomen, terwijl er ook ieder jaar veranderingen in de status van het fonds plaatsvinden. Zo worden er pensioenuitkeringen gedaan, gaan er mensen met pensioen, overlijden er mensen, gaan er mensen uit dienst, enzovoorts. Aan het eind van het jaar worden de opbrengsten van de activa bekend. Al deze ontwikkelingen hebben hun weerslag op de omvang van het vermogen en de verplichtingen.

Als de waarden van de activa en passiva bekend zijn, kan de dekkingsgraad worden berekend. Zoals al eerder besproken, is dit een belangrijke graadmeter voor de solvabiliteit van het pensioenfonds. Wanneer al deze gegevens bekend zijn, wordt het tijd om vooruit te kijken. Welke beslissingen worden er gemaakt voor het komende jaar? Deze beslissingen betreffen bijvoorbeeld de hoogte van de premies, maar ook wat de samenstelling van de portefeuille is. Het is hierbij erg belangrijk dat de dekkingsgraad op voldoende niveau blijft. Wanneer bijvoorbeeld in het afgelopen jaar de dekkingsgraad erg hoog was, kan ervoor gekozen worden om te investeren in activa met een hoger verwacht rendement en dus een hogere risico. Wanneer de dekkingsgraad lager was kan ervoor gekozen worden om de premies te verhogen. Ook worden er besluiten genomen over de hoogte van de indexatie van de pensioenuitkeringen of pensioenrechten op het moment dat de inflatie bekend is. Wanneer al deze beslissingen genomen zijn, ligt de startsituatie voor het komende jaar vast. Deze beslissingen moeten ieder jaar opnieuw worden gemaakt. Het is dus een dynamisch probleem.

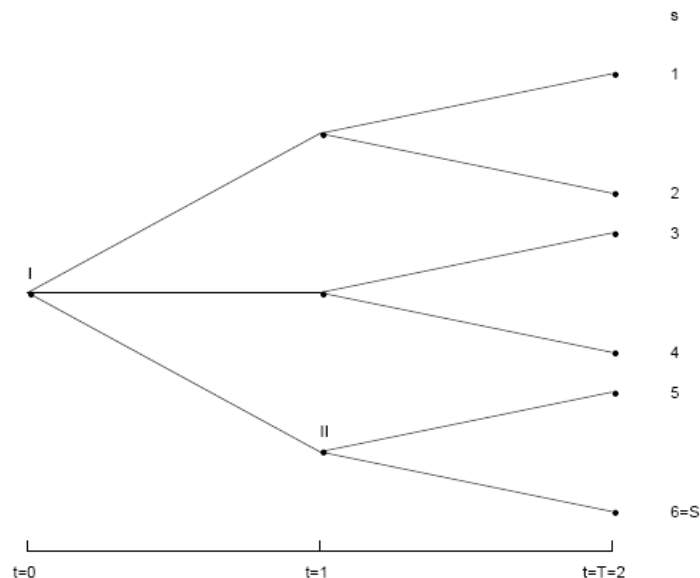
Het is duidelijk dat bij deze beslissingen rekening moet worden gehouden met alle belanghebbenden. De actieve deelnemers zullen niet te veel premie willen betalen. De gepensioneerden en de premievrije deelnemers zijn geïnteresseerd in het indexatiepercentage. De eventuele sponsor van het fonds betaalt niet alleen een deel van de premie, maar moet ook extra stortingen doen in het geval dat de dekkingsgraad onder een bepaald niveau komt. Het is dus duidelijk dat de sponsor het beleggingsbeleid scherp in de gaten houdt. Dan is er nog de toezichhouder (DNB) van het fonds. Er zijn een aantal regels waaraan pensioenfondsen moeten voldoen. De doelstellingsfunctie van een SLP model zal al deze belangen moeten behartigen.

Een optimalisatiemodel met beslissingsvariabelen waarbij er meerdere beslissingsmomenten en meerdere onzekerheden zijn wordt ook wel multistage stochastisch programming genoemd. Er dient een scenario tree opgesteld te worden om de

onzekere toekomst te modelleren. Scenario trees worden in de volgende paragraaf besproken.

4.2 Het genereren van scenario's

Er is een scenario tree nodig om de onzekere toekomst te modelleren. Een voorbeeld van een scenario tree wordt gegeven in figuur 11. In dit voorbeeld zijn er 6 verschillende scenario's. Iedere node in deze tree geeft een bepaalde staat aan. Een scenario s komt voor met kans p_s , waarbij $\sum p_s = 1$. Iedere staat heeft een set met variabelen die hun waarde krijgen gedurende het jaar. Dit zijn bijvoorbeeld de return op een bepaalde activa klasse, de waarde van de verplichtingen, de waarde van de uitkeringen, enzovoorts. Beslissingen kunnen alleen plaatsvinden op basis van informatie die stamt uit het verleden. Ieder scenario heeft een vector met beslissingvariabelen. Beslissingsvariabelen zijn dan bijvoorbeeld de te kopen hoeveelheid van een bepaalde activa klasse, de hoogte van de premie, enzovoorts. Ook bevat de tree voor iedere node de voorwaarden zoals later worden beschreven in dit hoofdstuk. Het optimale beleid is dynamisch.



Figuur 11: Een voorbeeld van een scenario tree zoals gegeven in [9].

Bij het nemen van de beslissingen moet de meest recente informatie met betrekking tot onzekere factoren worden meegenomen.

4.2.1 De toekomstige waarde van de activa

Om een optimaal investeringsbeleid te bepalen moeten de scenario's informatie bevatten over de returns op de verschillende activa klassen, maar ook over de toekomstige inflatie, enzovoorts. Deze scenario's worden in de literatuur veelal gegenereerd met behulp van VAR (Vector Autoregressief) modellen. Het idee achter deze modellen zal in hoofdstuk 6 verder behandeld worden. Er zijn ook alternatieven voor VAR modellen. Deze zullen ook in hoofdstuk 6 worden behandeld.

4.2.2 De verplichtingen en toekomstige pensioenuitkeringen

Om een goed model op te stellen is het noodzakelijk om een set scenario's te verkrijgen met toekomstige prijzen en de actuariële status van het fonds. Dit is in de vorige paragraaf besproken. Het doel van een pensioenfonds is bijvoorbeeld om deelnemers uitkeringen te doen van 70% van het laatstverdiende salaris. Hoeveel dit uiteindelijk werkelijk is, hangt af van het aantal jaren dat een deelnemer rechten heeft opgebouwd door premies te betalen en de hoogte van zijn salaris. De opgebouwde rechten worden dan bepaald door de volgende formule

$$O_k(t) = 1.75\%d_k(S_k(t) - PP(t)),$$

waarbij d_k het aantal jaren is dat deelnemer k , rechten heeft opgebouwd, $S_k(t)$ is het salaris van deelnemer k op tijdstip t en $PP(t)$ is de franchise op tijdstip t .

De waarde van de verplichtingen wordt dan bepaald door de netto contante waarde van de verwachte pensioenuitkeringen. Deze netto contante waarde kan als volgt worden weergegeven:

$$L_k(t) = \sum_{j=\max(1,65-l_{kt})}^{\infty} O_k(t) \left(\frac{1}{1+r} \right)^j p_{l_{kt}}^{(j)}.$$

Hierbij is r de rekenrente (die binnenkort niet langer vast is, maar afhankelijk is van de marktwaarde), en $p_{l_{kt}}^{(j)}$ de kans dat een persoon met leeftijd l_{kt} nog steeds levend is over j jaar.

Er moeten scenario's worden gegenereerd voor de onzekere ontwikkeling van de verplichtingen en de toekomstige pensioenuitkeringen. Belangrijk is dan om eerst een schatting te geven van de toekomstige status van de deelnemers, dus hoeveel deelnemers gaan er met pensioen, worden invalide, en zijn actief?

Het simuleren van de ontwikkeling van de verplichtingen en de toekomstige pensioenuitkeringen kan met behulp van een Push Pull Markov model [3]. Met dit model wordt voor iedere node in de scenario tree de status van de deelnemers bepaald op dat moment. Met behulp van overgangskansen wordt deze status bepaald. De overgangskansen zijn dan bijvoorbeeld de sterftekansen. Deze overgangskansen worden verkregen aan de hand van actuariële tabellen. Dit gedeelte wordt het Push Markov model genoemd. Het tweede deel wordt het Pull Markov model genoemd. Hier wordt gekeken naar de grootte van het deelnemersbestand in de toekomst. Met behulp van deze getallen kan dan een goede schatting worden gegeven van de toekomstige pensioenuitkeringen en verplichtingen. De verplichtingen worden berekend door de toekomstige uitkeringen te verdisconteren met behulp van de Nelson Siegel interest rate structuur (dat is aan de hand van de toekomstige marktwaarde). De Nelson Siegel interest rate zal later worden besproken.

4.3 De uitgangspunten van een SLP model

In de komende paragraaf zullen de uitgangspunten van een SLP model worden behandeld: de beslissingsvariabelen worden geïntroduceerd, alsmede de doelstellingsfunctie en de restricties van het model

4.3.1 De beslissingsvariabelen

Voor ieder jaar in de planningshorizon T dienen beslissingen te worden gemaakt. De beslissingsvariabelen in een SLP probleem zijn veelal het beleggingsbeleid, de hoogte van de premievoet en het indexatiepercentage. Deze beslissingen zijn afhankelijk van de meest recent beschikbaar gekomen informatie. Welke beslissingsvariabelen allemaal precies een rol spelen hangt af van het gekozen model.

4.3.2 De restricties van het model

Er kunnen verschillende soorten risico worden onderscheiden. Zo bestaat er short term risico. De overheid heeft een aantal harde voorwaarden opgesteld voor de risico's op de korte termijn. Allereerst moet er een voorwaarde worden opgesteld die ervoor zorgt dat de kans dat de dekkingsgraad groter of gelijk is aan een minimum waarde groot genoeg is. Deze voorwaarde kan worden weergegeven met behulp van een chance constraint. De chance constraint hiervoor ziet er als volgt uit:

$$P(F_{t+1}^* \geq \alpha \mid (t,s)) \geq \eta_t.$$

Hierbij is de stochast F_{t+1}^* de dekkingsgraad vlak voor een extra storting wordt gedaan, α de minimum hoogte van de dekkingsgraad (bijvoorbeeld $\alpha = 105\%$), η_t de vereiste kans (bijvoorbeeld $\eta_t = 0.95$) en (t,s) is de staat waarin het model zich bevindt op tijdstip t en in scenario s . Met deze vergelijking wordt dus de kans op onderdekking gemodelleerd. Het nadeel van een chance constraint is dat deze de omvang van het tekort niet meeneemt. Integrated chance constraints zijn echter voorwaarden waarbij de omvang van het tekort wel een rol speelt.

Integrated chance constraints kunnen als volgt worden weergegeven:

$$E[(A_{t+1}^* - \alpha L_{t+1})^- \mid (t,s)] \leq \beta_t.$$

Hierbij is A_{t+1}^* de waarde van de activa vlak voor er een extra storting wordt gedaan, L_{t+1} de waarde van de verplichtingen en β_t een bovengrens aan het verwachte tekort. Deze voorwaarde zorgt er dus voor dat het verwachte tekort kleiner is dan de bovengrens β_t . Er geldt dat $(a)^- := \max\{-a, 0\}$.

Verder bestaat er nog mid term en long term risico. Mid term houdt in dat bekeken over een aantal jaren de dekkingsgraad korte tijd lager mag zijn, maar als dit te vaak gebeurt of het tekort is te groot (deze waarden zijn vastgesteld door de overheid) dan moet er actie worden ondernomen. Long term risico houdt in dat de solvabiliteit van een fonds, bekeken over een groot aantal jaren van voldoende niveau moet zijn.

Verder gelden er binnen een SLP model nog een groot aantal voorwaarden. Deze worden hieronder opgesomd. Voorwaarden wat betreft

- De cash flow: er moet gelden dat de cash inflow gelijk is aan de cash outflow;
- De dekkingsgraad: wanneer deze onder een bepaald niveau komt, moeten er extra stortingen worden gedaan door de eventuele sponsor van het fonds;

- Boetes: er wordt een boete betaald voor grote fluctuaties in de premies van het ene jaar op het andere. Tevens moet een boete worden betaald voor een te lage dekkinggraad;
- De minimale en maximale hoogte van het indexatiepercentage;
- De sponsor: de sponsor moet worden terugbetaald op het moment dat de dekkinggraad erg hoog is;
- De fracties geïnvesteerd in de verschillende activa klassen;

4.3.3 Doelstellingsfunctie

Het doel van een SLP model is om de te betalen premies van de deelnemers en de sponsor te minimaliseren. Tevens is het doel om ongunstige situaties (zoals een lage dekkinggraad, sterk fluctuerende premie, geen volledige indexatie) te voorkomen. Daarom worden deze situaties in de doelstellingsfunctie bestraft met een boete.

Het doel is dus een optimaal beleid te vinden waarvan iedere belanghebbende optimaal profiteert.

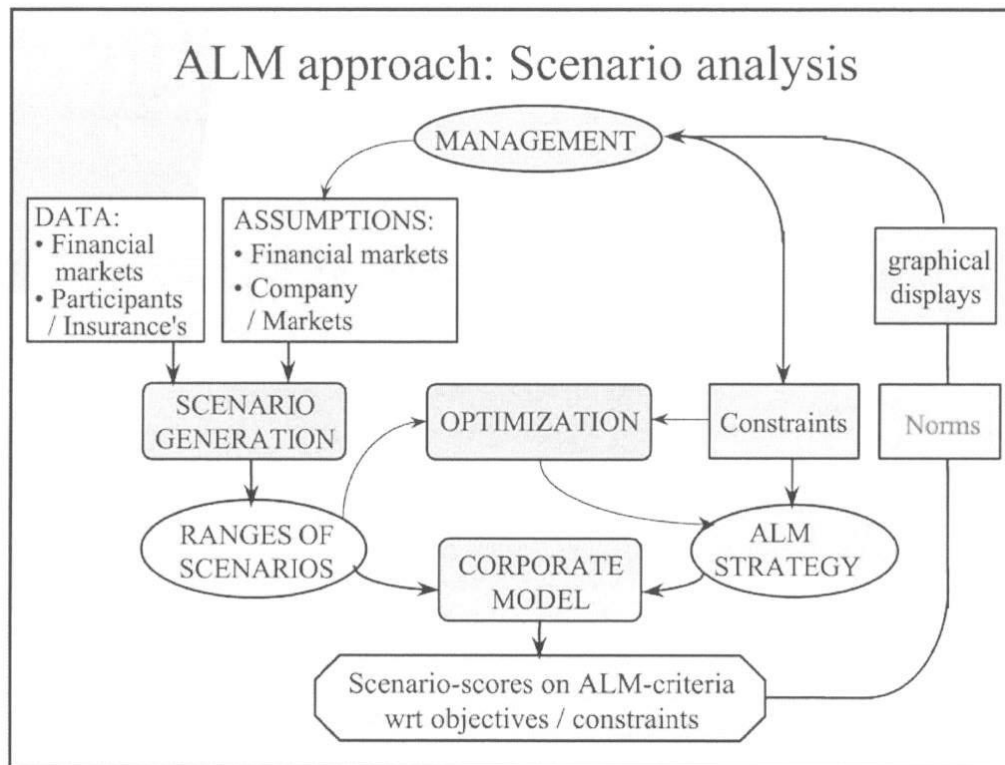
4.4 Voordelen, nadelen en conclusie

Het voordeel van een SLP model is dat er gezocht wordt naar een optimaal beleid. Het model is echter zo complex, dat het oplossen hiervan veelal gaat via een heuristiek. Het is dus nog maar de vraag of het optimale beleid wel gevonden wordt. Ook is het opstellen van een realistisch SLP model erg complex en tijdrovend. Nog een nadeel is de lange rekentijd. Er kunnen slechts een redelijk beperkt aantal scenario's worden gebruikt.

Vanwege deze nadelen kiest Zwitserleven er niet voor om een SLP model te implementeren om ALM studies mee te verrichten. Vanwege de praktische bruikbaarheid zal Zwitserleven waarschijnlijk kiezen voor een simulatiemodel [1], [9], [10]. In het volgende hoofdstuk zullen simulatiemodellen worden beschreven.

5. Simulatiemodellen

Simulatiemodellen zijn in Nederland vooral beschreven door Boender [3]. In figuur 12 worden de onderlinge relaties weergegeven binnen een simulatiemodel. Bij een simulatiemodel zijn er twee losstaande zaken: de economische scenario generator en het optimalisatiemodel. Dit is dus verschillend van SLP modellen, waar zowel de scenario's en het optimalisatieprobleem in één model worden gestopt.



Figuur 12: De onderlinge relaties in een simulatiemodel zoals beschreven in [3].

In dit hoofdstuk zal eerst iets verteld worden over een optimalisatiemodel. Tijdens dit onderzoek zal de nadruk echter komen te liggen op de economische scenario generator. Wat zijn de verschillende modellen om macro-economische scenario's te genereren? De tweede paragraaf van dit hoofdstuk zal dieper ingaan op economische scenario's.

5.1 Optimalisatiemodel simulatiemodel

In een simulatiemodel wordt gezocht naar een optimaal beleid. Bij een simulatiemodel is het idee dat er aan het begin van de planningshorizon een beleid wordt vastgesteld. De performance van het beleid wordt dan met behulp van een simulatie bepaald aan de hand van de gegenereerde scenario's. In deze simulatie wordt het beleid doorgerekend voor alle scenario's. De performance van een bepaald beleid wordt geëvalueerd aan de hand van een aantal maatstaven. Ook hier moet worden gezocht naar een beleid waarvan alle belanghebbenden profiteren.

Er wordt dus voor ieder beleid een simulatie uitgevoerd. Dan wordt er gekeken naar wat de consequenties zijn van dit beleid. Er vindt een 'learn and react' aanpak plaats: ALM strategieën worden geëvalueerd, geanalyseerd, en verbeterd in een iteratieve procedure. Dit wordt net zolang herhaald totdat een efficiënte en 'goede' strategie is gevonden. Deze aanpak wordt toegepast door middel van optimalisatie technieken. Net als bij de

SLP modellen bestaat het beleid uit het beleggingsbeleid, het premiebeleid en het indexatiebeleid. Het beleggingsbeleid betreft bijvoorbeeld grenzen aan de samenstelling van de portefeuille, maar ook doelstellingen wat betreft het rendement en het beheersen van het risico. Het premiebeleid geeft bijvoorbeeld de hoogte van de premies aan, en welk deel van de premie voor rekening van de werkgever komt. Het indexatiebeleid bepaalt bijvoorbeeld de hoogte van het indexatiepercentage, maar ook bij welke dekkingsgraad er wel of niet geïndexeerd wordt. Er wordt alleen geïndexeerd als de financiële positie van het fonds dit toelaat.

In iedere node (t,s) van de scenario tree wordt het pensioenfonds gekarakteriseerd door de samenstelling van de portefeuille, de dekkingsgraad en de premie. Allereerst wordt er het beleid gekozen. Wat betreft de samenstelling van de beleggingsportefeuille kan bijvoorbeeld worden gekozen voor buy and hold (niets veranderen aan de samenstelling) of rebalancing (dat wil zeggen dat de fracties in iedere klasse ieder jaar worden aangepast zodanig dat ze gelijk blijven). Wat betreft het indexatiebeleid kan worden besloten om niet te indexeren als de dekkingsgraad te laag is. Hetzelfde geldt voor de premie. Er kan worden gekozen voor extra premiebetaling of een 'premie holiday', afhankelijk van de dekkingsgraad. Een 'premie holiday' houdt in dat een pensioenfonds een werkgever de premie kwijtscheldt voor een bepaalde periode.

Voor het gekozen beleid wordt een simulatie uitgevoerd op de set scenario's. De performance van dit beleid kan worden bepaald aan de hand van een aantal maatstaven, zoals bijvoorbeeld de gemiddelde premie van het gesimuleerde beleid, de gemiddelde dekkingsgraad van het gesimuleerde beleid, de kans op onderdekking, enzovoorts. Tevens worden er vaak andere graadmeters gebruikt zoals PaR (Pension at Risk). Dit geeft het risico aan wanneer er over een langere tijd niet geïndexeerd wordt. Er moeten duizenden strategieën worden gegenereerd en geëvalueerd, zodat het beleid met de beste eigenschappen wordt gekozen.

5.2 Economische scenario generator

Tijdens dit onderzoek zal de nadruk liggen op de economische scenario generator. Wat zijn de verschillende modellen om macro-economische scenario's te genereren? De economische scenario's bevatten informatie over bijvoorbeeld de inflatie, interest rates, aandelenkoersen, maar ook bijvoorbeeld de ontwikkeling van het deelnemersbestand (of het deelnemersbestand ook behoort tot de scenario's hangt af van het gekozen model).

5.2.1 Bekende economische scenario generatoren

Er zijn een aantal bekende pakketten op de markt om macro-economische scenario's te genereren. Over het algemeen zijn dit erg robuuste pakketten, die gebaseerd zijn op verschillende onderliggende modellen. Een nadeel van deze modellen is dat ze vaak erg duur zijn. Het is vaak niet mogelijk om het model zelf aan te passen of de parameters te schatten. Omdat deze modellen vaak ontwikkeld zijn door grote consultancy bedrijven, kunnen alleen zij aanpassingen doen.

- **The Smith Model (TSM)**
- **Global CAP: Link**
- **Barrie and Hibbert**
- **The Wilkie model**

Een belangrijke eigenschap van een economische scenario generator is flexibiliteit. Het moet mogelijk zijn om extra asset klassen toe te voegen, de veronderstellingen achter de parameters te veranderen en deze veronderstellingen tijdsvariërend te maken. Ook erg

belangrijk is dat de gebruiker van het pakket weet welke technieken er gebruikt worden om de parameters te schatten [11].

5.2.2 Scenario's voor economische grootheden genereren

Er zijn verschillende modellen waarmee economische scenario's kunnen worden gegenereerd. Het genereren van economische scenario's gebeurt in een aantal stappen. Allereerst is historische data vereist. Op basis van deze data wordt dan een bepaald model geschat. Welk type model dit precies is, hangt af van de data. Dit wordt ook wel tijdserie analyse genoemd. Tijdserie analyse houdt zich bezig met de statistische analyse van een reeks afhankelijke random variabelen. Hier geldt vaak dat variabelen die dicht bij elkaar liggen afhankelijk zijn en variabelen die ver uiteen liggen onafhankelijk zijn. Het belangrijkste doel van tijdserie analyse is om toekomstige waarden te voorspellen met behulp van het model dat geschat is op basis van de geobserveerde waarden. Er wordt dus verondersteld dat historische waarden informatie bevatten over toekomstige waarden van de variabele. In dit geval is het doel om toekomstige waarden van verschillende macro-economische variabelen te kunnen voorspellen. Er wordt dan een reeks voorspellingen gegenereerd voor bijvoorbeeld een periode van vijf jaar, de normale duur van een contract bij Zwitserleven. Om een goede betrouwbaarheid te verkrijgen worden er dan een groot aantal van deze reeksen gegenereerd. Dit zijn dan de economische scenario's. Als er eenmaal een groot aantal economische scenario's zijn gegenereerd kan bijvoorbeeld de kans worden bepaald dat bij de huidige samenstelling van de portefeuille de dekkinggraad onder een bepaald niveau komt in de komende vijf jaar.

Er zijn verschillende typen modellen. Deze modellen proberen een tijdserie te 'begrijpen' zodat er nieuwe waarden van een variabele kunnen worden voorspeld. In hoofdstuk 6 zal een overzicht worden gegeven van de verschillende modellen om economische scenario's te genereren. De modellen die worden besproken zijn de volgende:

1. Autoregressief model (AR);
2. Moving average model (MA);
3. Vector Autoregressief model (VAR);
4. Autoregressief (Integrated) Moving average model (AR(I)MA);
5. Exponential Smoothing.

5.2.3 Scenario's voor de actuariële grootheden

Het genereren van de scenario's voor de actuariële grootheden gaat op dezelfde wijze als beschreven in paragraaf 4.2.

5.3 ALM tools

Vertrouwelijk.

6. Data analyse en modelkeuze voor economische scenario's

In dit hoofdstuk wordt allereerst de gebruikte data beschreven. Daarna worden de verschillende typen modellen besproken. Welk model is nu geschikt om economische scenario's te genereren? Deze vraag wordt in dit hoofdstuk beantwoord.

6.1 Gebruikte data

De data zijn verkregen uit DATASTREAM. Om het onderzoek in het begin beperkt te houden is ervoor gekozen om gebruik te maken van de aandelenkoersen en de inflatie, alsmede de rente op kort- en langlopende obligaties. Er wordt gedefinieerd:

- $X1$: Het Nederlandse inflatiepercentage;
- $X2$: De langlopende obligatie-index;
- $X3$: De kortlopende obligatie-index;
- $X4$: Het maandelijks rendement op de AEX index.

Voor de rente op de langlopende obligaties is gekozen voor de Lehman long index en voor de rente op de kortlopende obligaties is gekozen voor de Lehman treasury 1-3 jaar. Dit is vanwege het feit dat het Swiss Life Government Bond 10+ fonds, Lehman treasury 10+ als benchmark heeft. Dit wil zeggen dat het Swiss Life Government Bond 10+ fonds Lehman als vergelijkingsmaatstaf neemt. Aan de hand van de benchmark kunnen de prestaties van het fonds beoordeeld worden. Doet het fonds het beter dan de benchmark, dan doet het fonds het goed en omgekeerd. Het Swiss Life obligatiefonds gebruikt ook Lehman als benchmark. Daarom is het een logische keuze om Lehman data te gebruiken voor het opstellen van de scenario's. Het Swiss Life aandelenfonds heeft de AEX als benchmark. Daarom is historische data van de AEX totale return index gebruikt om het aandelenrendement te modelleren. De totale return index houdt rekening met het herinvesteren van dividend.

De maandelijks data voor $X1$, $X2$ en $X3$ loopt van februari 1979 t/m mei 2007. Dit zijn in totaal dus 340 datapunten. De maandelijks data voor het rendement op de AEX index loopt van januari 1983 t/m mei 2007. Dit zijn in totaal 294 datapunten. Om de data te analyseren is gebruik gemaakt van het softwarepakket R.

6.2 Definities

Voordat de verschillende modellen beschreven worden, moeten eerst een aantal definities worden gegeven.

Een tijdserie is niet stationair wanneer hij sterk stijgt of daalt en vrijwel nooit terugkeert op zijn startpunt. Een proces is stationair als de eigenschappen van een proces, zoals het gemiddelde en de variantie, niet variëren in de tijd [26]. Of een tijdserie stationair is kan worden getest met behulp van de Augmented Dickey-Fuller test.

Deze test ziet er als volgt uit:

$$\Delta X_t = \mu + \gamma X_{t-1} + \sum_{i=1}^m \delta_i \Delta X_{t-i} + \varepsilon_t.$$

Hierbij geldt dat $\Delta X_t = X_t - X_{t-1}$, γ en δ_i zijn de regressiecoëfficiënten.

De nulhypothese is dat de tijdserie niet stationair is:

$$H_0 : \gamma = 0,$$

en de alternatieve hypothese is dat de tijdreeks stationair is:

$$H_1 : \gamma < 0$$

[19].

De Autocorrelatie functie (ACF) geeft de correlatie tussen de variabele en de voorgaande waarden van zichzelf, dus bijvoorbeeld de correlatie tussen X_t en X_{t-2} . De Partiele Autocorrelatie functie (PACF) geeft de correlatie tussen de variabele en de voorgaande waarde van zichzelf, alleen dan rekening houdend met het feit dat als er correlatie is tussen X_t en X_{t-1} , en X_{t-1} en X_{t-2} zijn ook gecorreleerd, er ook correlatie gevonden wordt tussen X_t en X_{t-2} .

De autocovariantie kan worden gedefinieerd als de covariantie tussen twee waarden t en t' van de tijdreeks. Deze kan als volgt worden weergegeven:

$$\gamma_X(t', t) = \text{Cov}(X_{t'}, X_t) = E[(X_{t'} - E[X_{t'}])(X_t - E[X_t])].$$

De autocovariantie van een stationaire reeks wordt gegeven door:

$$\gamma_X(h) = \gamma_X(h, 0) = \text{Cov}(X_{t+h}, X_t).$$

De ACF voor lag h van een stationair proces X_t is dan:

$$\rho_X(h) = \text{Cor}(X_{t+h}, X_t) = \frac{\gamma_X(h)}{\gamma_X(0)}.$$

Voor een stationair multivariaat proces met d tijdseries $X_t = (X_{t,1}, \dots, X_{t,d})$ met $EX_t = (EX_{t,1}, \dots, EX_{t,d})$ de vector met de verwachting van iedere tijdserie, wordt de autocovariantie functie gedefinieerd als de matrix [29]:

$$\gamma_X(h) = (\text{cov}(X_{t+h,i}, X_{t,j}))_{i,j=1,\dots,d} = E(X_{t+h} - EX_{t+h})(X_t - EX_t)^T.$$

De autocorrelatie functie is gedefinieerd:

$$\rho_X(h) = (\rho(X_{t+h,i}, X_{t,j}))_{i,j=1,\dots,d} = \left(\frac{\gamma_X(h)_{i,j}}{\sqrt{\gamma_X(0)_{i,i} \gamma_X(0)_{j,j}}} \right)_{i,j=1,\dots,d}.$$

Een witte ruis proces is een stationair proces dat bestaat uit een reeks onafhankelijke variabelen $\dots, Z_{-1}, Z_0, Z_1, \dots$ ieder met een gemiddelde van 0 en variantie σ^2 . De autocovariantie functie van een witte ruis proces is gedefinieerd als:

$$\gamma_Z(h) = \sigma^2 \text{ als } h = 0,$$

$$\gamma_Z(h) = 0 \text{ als } h \neq 0.$$

De random variabelen zijn dus onafhankelijk.

6.3 Stationaire data verkrijgen

Wanneer blijkt dat de data niet stationair zijn, moeten de data zodanig worden bewerkt dat er een stationair proces overblijft. Twee oorzaken die ervoor zorgen dat data niet stationair zijn, zijn *trend*- en *seizoencomponenten*. Hierbij wordt een trend gedefinieerd als een regelmatige verhoging of daling in het gemiddelde van het proces. Seizoencomponenten komen voor wanneer het gemiddelde van het proces een cyclus vertoont. Bijvoorbeeld wanneer het gemiddelde in de zomer hoger ligt vergeleken met de winter. Voordat er gewerkt kan worden met een model waarvoor stationaire data benodigd is, moeten dus eerst deze trend en seizoencomponenten uit de data worden verwijderd. Om uit de niet-stationaire set een stationaire set te verkrijgen, kunnen verschillende methodes gebruikt worden. Hieronder worden twee methoden besproken, de small trend methode en de differencing methode [18].

6.3.1 Small trend methode

Om de trend en de seizoensinvloeden te elimineren uit de data kan gebruik worden gemaakt van de small trend methode. Bij deze methode wordt ervan uitgegaan dat de data realisaties zijn van het klassieke decompositie model dat er als volgt uitziet:

$$X_t = m_t + s_t + Y_t,$$

waarbij m_t een vaste trend component is, s_t een vaste seizoenscomponent met bekende periode d en Y_t de random error term die stationair verondersteld wordt. Zonder beperking der algemeenheid mag er worden verondersteld dat $EY_t = 0$.

De data dienen te worden gelabeld. Er geldt dan bijvoorbeeld dat $X_{j,k}$ de waarneming van variabele X is in de k -de periode van jaar j . Er geldt dan dus $X_{j,k} = X_{k+d(j-1)}$.

Als schatter voor de trend in het j de jaar wordt genomen:

$$\hat{m}_j = \frac{1}{d} \sum_{k=1}^d X_{j,k}, \quad j = 1, \dots, T.$$

Als schatter voor het seizoenscomponent in periode k wordt genomen:

$$\hat{s}_k = \frac{1}{T} \sum_{j=1}^T (X_{j,k} - \hat{m}_j), \quad k = 1, \dots, d.$$

Er geldt dat $s_{t+d} = s_t$ en $\sum_{j=1}^d s_j = 0$.

De random error term voor periode k in jaar j wordt dan gegeven door:

$$\hat{Y}_{j,k} = X_{j,k} - \hat{m}_j - \hat{s}_k \quad j = 1, \dots, T; \quad k = 1, \dots, d.$$

6.3.2 Differencing van een tijdserie

Een andere methode om de trend en de eventuele seizoencomponenten te verwijderen is de differencing methode van de orde k .

Hierbij wordt gedefinieerd:

$$\nabla_d X_t = X_t - X_{t-d}, \quad t > d,$$

waarbij d weer de periode van het seizoenscomponent aangeeft, dus bijvoorbeeld $d = 12$ voor jaarlijkse seizoenscomponenten en $d = 3$ voor seizoenscomponenten per kwartaal. Toegepast op $X_t = m_t + s_t + Y_t$ waarbij s_t periode d heeft, leidt dit tot het volgende model:

$$\begin{aligned} \nabla_d X_t &= X_t - X_{t-d} = \\ &= (m_t + s_t + Y_t) - (m_{t-d} + s_{t-d} + Y_{t-d}) = \\ &= (m_t - m_{t-d}) + (s_t - s_{t-d}) + (Y_t - Y_{t-d}) \end{aligned}$$

Aangezien $s_{t+d} = s_t$, geldt dat $(s_t - s_{t-d}) = 0$. Daarom leidt dit tot het volgende model:

$$\nabla_d X_t = m_t - m_{t-d} + Y_t - Y_{t-d}.$$

Nu is $\nabla_d X_t$ gedecomposeerd in een trend component $(m_t - m_{t-d})$ en een noise term $(Y_t - Y_{t-d})$. Het is duidelijk dat $Y_t - Y_{t-d}$ stationair is, aangezien Y_t dit ook is.

Allereerst dient ∇_d te worden toegepast op de dataset en daarna dient de trend te worden verwijderd, via onderstaande methode. Er volgt dan een stationair proces.

Een polynomiale trend van de orde k kan worden gereduceerd tot een constante door de operatie ∇ (waarbij $\nabla X_t = X_t - X_{t-1}$), k keer toe te passen. Dit wil in het algemeen zeggen dat het k keer toepassen van ∇ op een tijdserie X_t leidt tot een serie $\{\nabla^k X_t\}$ (welke een andere betekenis heeft dan ∇_d) die kan worden gemodelleerd als een realisatie van een stationair proces. De vereiste orde k is veelal klein, meestal 1 of 2.

6.4.3 Het bepalen van de seizoensinvloeden

Zijn er seizoensinvloeden te ontdekken in de data, en zo ja, wat is de periode? Deze vraag kan worden beantwoord met behulp van het periodogram van een tijdserie. Een cyclisch proces voor de tijdreeks $\{y(t); t = 1, \dots, N\}$ met n componenten wordt beschreven door:

$$y(t) = \sum_{k=1}^n \alpha_k \sin(\omega_k t + \theta_k) + \varepsilon(t), \quad t = 1, \dots, N.$$

Van de parameters θ_k wordt verondersteld dat ze onafhankelijk en uniform verdeeld zijn op $[0, 2\pi)$. Het proces $\varepsilon(t)$ is aanwezig aangezien het proces niet puur cyclisch is. Het is de bedoeling om de frequenties ω_k te vinden. Er geldt dan dat de frequentie van het cyclische proces kan worden bepaald door gebruik te maken van het spectrum, welke gedefinieerd is als:

$$S(z) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} R(k) z^{-k}.$$

Wanneer de covariantie $R(k)$ in het spectrum wordt vervangen door de steekproef covariantie:

$$\hat{R}(k) = \frac{1}{N} \sum_{t=k+1}^N y(t)y(t-k), \quad k = 0, \dots, N-1 \quad \text{met } \hat{R}(k) = \hat{R}(-k) \text{ voor } k < 0,$$

en met $z = e^{i\omega}$ leidt dit tot het periodogram. Uit het periodogram \hat{S}_N kan de periode van het cyclische proces worden afgelezen:

$$\hat{S}_N(e^{i\omega}) = \frac{1}{2\pi N} \left| \sum_{t=1}^n y(t)e^{-i\omega t} \right|^2$$

[36].

Wanneer het periodogram slechts één piek bij nul heeft, dan betekent dit dat de data nog een trend bezit. Deze zal eerst moeten worden verwijderd. Wanneer de trend is verwijderd kan opnieuw worden gekeken naar het periodogram. Wanneer het periodogram ergens een duidelijke piek heeft, is dit de periode van de seizoensinvloeden van een tijdserie.

6.4 Modellen voor tijdseries

In deze paragraaf zullen de verschillende modellen worden beschreven die geschikt zijn om de toekomstige waarden van het jaarlijks inflatiepercentage, de langlopende obligatie-index en de kortlopende obligatie-index te genereren, zoals opgesomd in hoofdstuk 5. Het aandelenrendement is een verhaal apart, en zal verderop in dit hoofdstuk beschreven worden. De laatste paragraaf van dit hoofdstuk beschrijft de manier waarop Triple A de economische scenario's genereert.

6.4.1 AR model

Een AR(p) model is een *Autoregressief* model van de orde p . Een autoregressief model heeft als verklarende variabelen alleen zijn eigen voorgaande waarden.

Dit model kan als volgt worden geformuleerd:

$$X_t = c + \sum_{i=1}^p \varphi_i X_{t-i} + Z_t.$$

Hierbij is c een constante, Z_t is de error term waarvan vaak wordt aangenomen dat deze i.i.d. normaal verdeeld zijn: $Z_t \sim N(0, \sigma^2)$. φ_i zijn de parameters van het model. X_t kan dus geheel worden beschreven door zijn eigen voorgaande waarden. Een AR model veronderstelt stationaire data.

6.4.2 MA model

MA staat voor *Moving Average*. Een MA(q) model ziet er als volgt uit:

$$X_t = c + \beta_0 Z_1 + \dots + \beta_q Z_{t-q}.$$

Hierbij is Z_t een witte ruis proces met variantie σ^2 . β_0, \dots, β_q zijn de parameters van het model. Een MA model veronderstelt stationaire data.

6.4.3 VAR model

Een VAR model is een multivariate uitbreiding van een AR model. VAR staat voor *Vector Autoregressief Model*. VAR modellen proberen dus meerdere tijdseries te ‘begrijpen’ zodat er nieuwe voorspellingen kunnen worden gedaan. De variabelen zijn onderling afhankelijk. Bij VAR modellen is het dus de bedoeling om onderling afhankelijke variabelen te modelleren. VAR modellen geven in de praktijk veel betere voorspellingen voor economische tijdreeksen dan de univariate modellen.

Een VAR(p) model kan dan als volgt worden weergegeven:

$$X_t = c + A_1 X_{t-1} + A_2 X_{t-2} + \dots + A_p X_{t-p} + Z_t.$$

Hierbij is A_i de autocorrelatie matrix (deze matrix geeft aan in hoeverre de te voorspellen variabele samenhangt met zijn ($t-i$)-de voorganger), X_t en ε_t zijn vectoren. X_t is de vector voor k verschillende variabelen (bijvoorbeeld aandelenkoersen, obligatiekoersen, enzovoorts). Er geldt dan $X_t = [X_{1t}, X_{2t}, \dots, X_{kt}]'$. ε_t is een zogenaamd multivariaat witte ruis proces. Hiervoor geldt dat $E[\varepsilon_t] = 0$ en $E[\varepsilon_t \varepsilon_{t'}'] = \Sigma_\varepsilon$ als $t = t'$ en $E[\varepsilon_t \varepsilon_{t'}'] = 0$ als $t \neq t'$. Hierbij is Σ_ε de covariantie matrix tussen de error termen van de verschillende variabelen in de vector X . Deze matrix geeft dus de onderlinge afhankelijkheid tussen de error termen weer. Een VAR model veronderstelt stationaire data. Het statistische softwarepakket R bevat echter een package (`vars`) dat rekening kan houden met een eventuele trend in de data [20].

6.4.4 AR(I)MA model

Een ARMA(p,q) (*Autoregressief Moving Average*) model is een combinatie van een AR en MA proces, met p AR termen en q MA termen. Het model ziet er dan als volgt uit:

$$X_t = c + \alpha_1 X_{t-1} + \dots + \alpha_p X_{t-p} + Z_t + \beta_1 Z_{t-1} + \dots + \beta_q Z_{t-q},$$

met X_t een stationair proces en Z_t een witte ruis proces met variantie σ^2 . Een ARMA model veronderstelt stationaire data.

Een model waarbij de data niet stationair verondersteld wordt, is het ARIMA(p,d,q) model. ARIMA staat voor *Autoregressief Integrated Moving Average*, en de variabele X_t in het ARMA model wordt vervangen door $\nabla^d X_t$ (de differencing methode, zoals beschreven in de vorige paragraaf) om zo een model te verkrijgen dat ook niet stationaire tijdseries kan beschrijven. De term *integrated* staat voor het feit dat het stationaire model dat is gefit op de gedifferenceerde data moet worden gesommeerd (integrated) om tot een model voor de niet stationaire data te komen. ARIMA modellen worden vaak gebruikt om economische tijdseries te beschrijven.

6.4.5 Exponential Smoothing

Bij de simpele vorm van exponential smoothing wordt de toekomstige waarde geheel voorspeld uit zijn voorgaande waarden. Er geldt:

$$X_t = w_0 X_{t-1} + w_1 X_{t-2} + \dots$$

w_i zijn de gewichten die zijn toegekend aan de vorige observaties van X_t . Vaak worden recente observaties zwaarder gewogen dan oudere observaties. Om dit te realiseren wordt vaak de volgende waarde voor de gewichten gekozen:

$$w_i = \alpha(1 - \alpha)^i.$$

Hierbij is α de smoothing constante en ligt tussen de 0 en 1. Voor deze vorm is gekozen omdat dan geldt dat de gewichten w_i tot 1 sommeren ($i = 0, \dots, \infty$). Dit wil zeggen dat w een kansverdeling is.

Een uitbreiding van exponential smoothing is de Holt Winters methode. De Holt Winters methode kan omgaan met data waarin een trend en/of seizoensinvloeden aanwezig is. Holt Winters gebruikt drie smoothing parameters: een voor de constante, een voor trend en een voor de seizoensinvloeden. Er zijn twee verschillende vormen te onderscheiden, namelijk:

1. De *multiplicatieve* vorm:

$$X_{t+1} = (l_t + m_t)s_{t-d}.$$

Waarbij l_t de level (constante) is op tijdstip t , m_t de trend op tijdstip t en s_{t-d} het seizoenscomponent met d de lengte van een periode en X_{t+1} de voorspelling over 1 tijdseenheid. l_t , m_t en s_{t-d} worden als volgt geüpdate.

$$l_t = \alpha_0 \frac{X_t}{s_{t-d}} + (1 - \alpha_0)(l_{t-1} + m_{t-1}).$$

$$m_t = \alpha_1(l_t - l_{t-1}) + (1 - \alpha_1)m_{t-1}.$$

$$s_t = \alpha_2 \frac{X_t}{l_t} + (1 - \alpha_2)s_{t-d}.$$

2. De *additieve* vorm:

$$X_{t+u} = l_t + um_t + s_{t-d+u}.$$

l_t , m_t en s_{t-d} worden als volgt geüpdate.

$$l_t = \alpha_0(X_t - s_{t-d}) + (1 - \alpha_0)(l_{t-1} + m_{t-1}).$$

$$m_t = \alpha_1(l_t - l_{t-1}) + (1 - \alpha_1)m_{t-1}.$$

$$s_t = \alpha_2(X_t - l_t) + (1 - \alpha_2)s_{t-d}.$$

[31].

R bevat de functie `HoltWinters` die automatisch de smoothing parameters berekent door het minimaliseren van de Mean Squared voorspellingsfout [31], [33].

Concluderend kan worden gesteld dat een aantal modellen geschikt zijn om met niet-stationaire data te werken. Dit zijn,

1. Het ARIMA model: dit model verkrijgt een stationaire dataset met behulp van de differencing methode;
2. Het VAR model: het pakket R bevat een package `vars` dat een VAR model kan schatten en hierbij rekening houdt met een eventuele trend in de data;

3. De Holt Winters methode: het pakket R bevat een functie `HoltWinters` die de smoothing parameters berekent.

Voor de rest van de beschreven modellen dient de data eerst stationair worden gemaakt, met een van de methodes zoals beschreven in paragraaf 6.3.

In het volgende hoofdstuk zal de data analyse worden beschreven. Is de data stationair, of bevat deze een trend of seizoensinvloeden?

6.4.6 Economische scenario generator van Triple A

De economische scenario generator van Triple A bestaat uit twee verschillende modules.

Allereerst is er de zogenaamde 'Real World module'. Deze methode baseert het schatten van de parameters van het model op historische gegevens en maakt gebruik van een VAR(1) model. Met behulp van het VAR model worden alle economische variabelen geschat, dus de aandelenrendementen, rentestanden, etc. Een tweede module in de generator van Triple A is de 'Market Implied module'.

De 'Real World module' maakt gebruik van historische data en de werkelijke yield curve. Hiermee kunnen veel verschillende variabelen worden gesimuleerd. Wanneer eenmaal de matrices A en Σ voor het VAR(1) model geschat zijn, worden de scenario's gesimuleerd door middel van een zogenaamde Cholesky factorisatie. Wat de Cholesky factorisatie precies inhoudt, wordt beschreven in het volgende paragraaf.

Het modelleren van de yield curve gaat met behulp van de Nelson Siegel vergelijking. Deze vergelijking ziet er als volgt uit:

$$r(t) = \beta_0 + (\beta_1 + \beta_2) \frac{(1 - e^{-t/\tau})}{t/\tau} - \beta_2 e^{-t/\tau}.$$

Hierbij is $r(t)$ de t -jarige yield op obligaties, τ en β_2 zijn vormparameters, β_0 wordt gezien als de langlopende rente, en $\beta_0 + \beta_1$ als de kortlopende rente [3], [11].

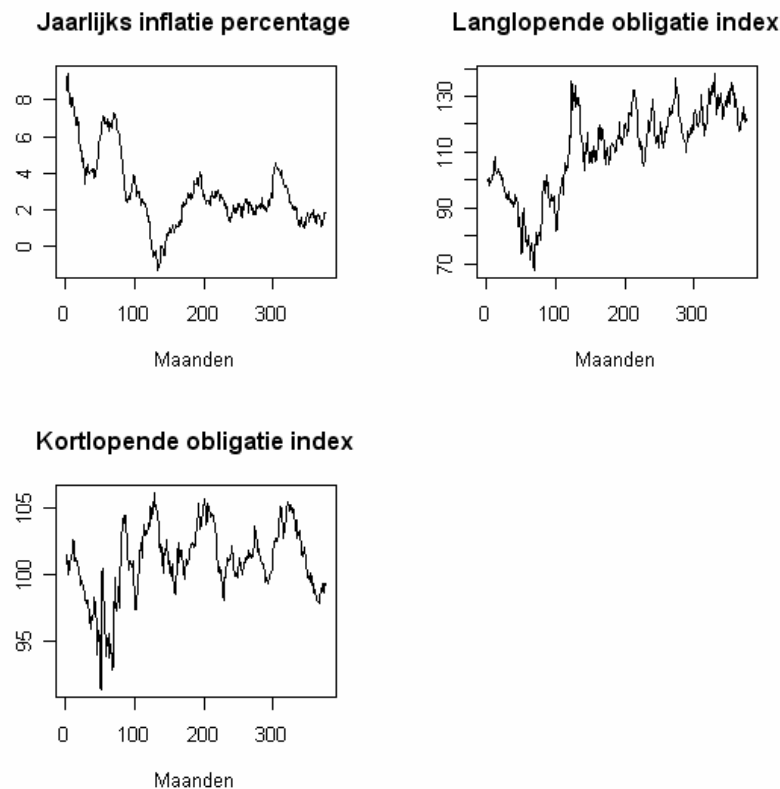
Er zijn verschillende methoden om de parameters van deze vergelijking te schatten. Een mogelijkheid is om de kort- en langlopende rente te genereren met behulp van een VAR model. De parameters τ en β_2 kunnen worden geschat uit de huidige yield curve en worden dan constant verondersteld tijdens de gehele tijdsperiode.

Een andere mogelijkheid is om alle vier parameters te schatten met behulp van een VAR model. Dit geeft natuurlijk betere uitkomsten, omdat het erg onrealistisch is om ervan uit te gaan dat τ en β_2 constant blijven. Een nadeel hiervan is dat er behalve historische data van de kort- en langlopende rente, ook historische data van de twee vormparameters τ en β_2 beschikbaar moet zijn.

Is het VAR(1) echter wel het meest optimale model om economische scenario's te genereren? In de volgende paragrafen zullen verschillende modellen worden vergeleken om dit te onderzoeken. Allereerst volgt een paragraaf over de data analyse.

6.5 Data analyse

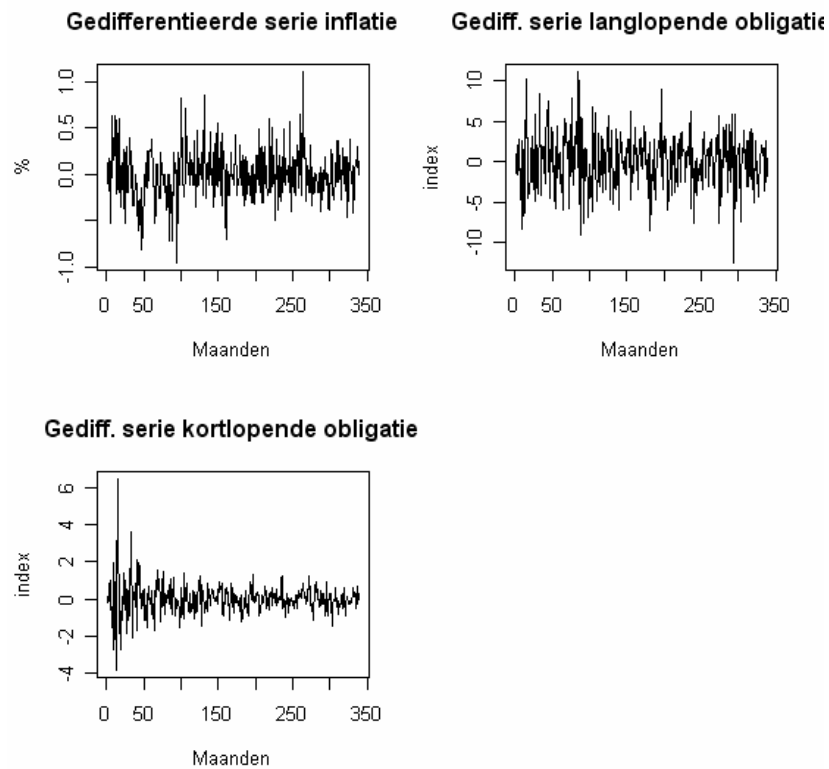
In de onderstaande figuur worden de historische waarden van de variabelen X_1 , X_2 en X_3 uitgezet tegen de tijd. X_4 wordt weergegeven in hoofdstuk 7.



Figuur 13: Historische waarden van X_1 , X_2 en X_3 .

In figuur 13 is de gekozen dataset weergegeven. Op het oog lijkt de dataset niet stationair. Om te testen of dit inderdaad het geval is, wordt de Augmented Dickey-Fuller test uitgevoerd. Hieruit blijkt dat de dataset inderdaad niet stationair is. De Holt Winters methode kan omgaan met zowel een trend als seizoensinvloeden. Het VAR model en het ARIMA model kunnen echter alleen een trend verwijderen uit de data. Het is daarom belangrijk om te weten of de data ook seizoensinvloeden bevat. Mochten de data inderdaad seizoensinvloeden bevatten, dan moet deze verwijderd worden uit de data.

Om te kijken of de data seizoensinvloeden bevatten, wordt gekeken naar het periodogram. Aangezien het periodogram van data dat nog een trend bevat een piek bij 0 heeft, moet eerst de eventuele trend verwijderd worden. Het blijkt dat alle drie de tijdseries een duidelijke trend bezitten. Wanneer de trend verwijderd wordt en de periodes worden bepaald, blijken de seizoensinvloeden zeer gering te zijn. Als deze worden verwijderd uit de tijdserie waar ook de lineaire trend uit is verwijderd, blijkt er geen stationair proces over te blijven. Een hogere orde trend fitten is ook geen optie, omdat een trend met een hoge orde de huidige data goed kan beschrijven, maar niets hoeft te zeggen over de voorspellingen voor de toekomst. Daarom valt de keuze op de differencing methode om een stationair proces te verkrijgen [29]. Wanneer een differencing methode van de orde 1 wordt toegepast, ziet de dataset er als volgt uit (zie figuur 14).



Figuur 14: De gedifferentieerde dataset.

Er wordt bekeken of deze series stationair zijn met behulp van de Augmented Dickey-Fuller test. De series, na differentiëren van de orde 1, blijken inderdaad stationair te zijn.

6.6 Modelleren van de inflatie, lang- en kortlopende obligatie-index

In deze paragraaf worden drie modellen vergeleken om zo tot een goed model te komen voor de inflatie, en de kort- en langlopende obligatie-index. De drie modellen die zullen worden vergeleken zijn,

1. ARIMA model (paragraaf 6.6.1);
2. VAR model (paragraaf 6.6.2);
3. Holt Winters (Exponential Smoothing) (paragraaf 6.6.3).

Allereerst zal een ARIMA model worden geschat. Dan wordt gekeken naar de voorspellingen die dit model geven. Daarna wordt een VAR model geschat, en de resultaten bekeken. In het laatste deel van deze paragraaf wordt een Holt Winters model op de data geschat. In de volgende paragraaf zullen deze modellen dan worden vergeleken, zodat een optimaal model kan worden gekozen.

6.6.1 ARIMA model

Box-Jenkins beschrijft een methode om ARIMA modellen te schatten. Allereerst dient een geschikt model te worden geformuleerd. Vervolgens moet het model worden gefit op de dataset. Daarna dient het model te worden geëvalueerd: is het inderdaad een goed model? Dit wordt een diagnostic check genoemd.

Om te bepalen welk ARIMA(p,k,q) model optimaal is voor de inflatie, en de kort- en langlopende obligatie-index, kan worden gekeken naar de autocorrelatie functie (ACF) en naar de partiele autocorrelatie functie (PACF).

Aangezien de werkelijke ACF onbekend is, moeten er schattingen voor worden gemaakt uit de data. Zo wordt als schatter voor de autocovariantie functie de steekproef autocovariantie functie $\hat{\gamma}_n(h)$ genomen:

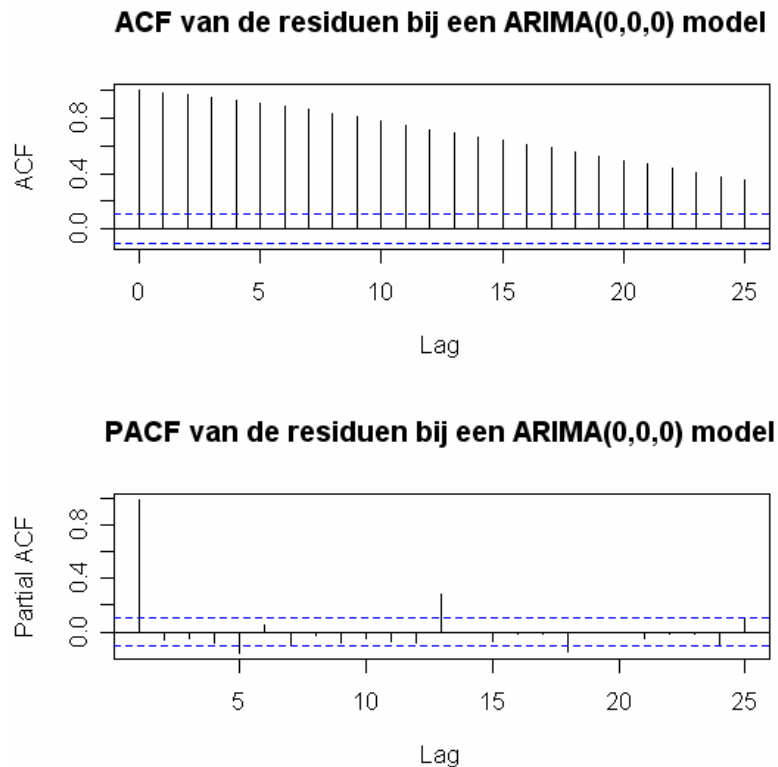
$$\hat{\gamma}_n(h) = \frac{1}{n} \sum_{t=1}^{n-h} (X_{t+h} - \bar{X}_n)(X_t - \bar{X}_n), \quad h \geq 0.$$

waarbij $\bar{X}_n = \frac{1}{n} \sum_{t=1}^n X_t$ de schatter voor de verwachting van het stationaire proces is.

Uit de steekproef autocovariantie functie kan dan de steekproef autocorrelatie functie worden berekend:

$$\hat{\rho}_n(h) = \hat{\gamma}_n(h) / \hat{\gamma}_n(0).$$

Een voorbeeld verduidelijkt dit. Figuur 15 geeft de ACF en PACF van de niet-gedifferentieerde tijdserie voor de inflatie (oftewel een ARIMA(0,0,0) model, met niet stationaire data).



Figuur 15: ACF en PACF van de niet- gedifferentieerde tijdserie van de inflatie.

De lineaire trend zoals te zien is in de plot van het ACF, is typisch voor een niet stationaire tijdserie. Alle lags hebben een significante correlatie met X_t , die waarschijnlijk kunnen worden verklaard door de correlatie van X_t en X_{t-1} . Dat dit inderdaad het geval is, is te zien in de PACF van de data. De grootste significante lag zit bij 1. Dit betekent dat de correlatie die te zien is in het ACF voor het grootste gedeelte verklaard wordt door de correlatie tussen X_t en X_{t-1} . De ACF en de PACF kunnen worden gebruikt om de orde van het ARIMA model te bepalen. Het is duidelijk dat de inflatie data eerst moeten worden gedifferentieerd voordat de orde kan worden afgelezen, aangezien deze duidelijk niet stationair is.

Aan de ACF van een stationaire tijdserie kan het aantal MA termen q worden afgelezen. Voor q wordt dan de grootste waarde waarvoor de ACF significant afwijkt van de nul gekozen. Aan de PACF van een stationaire tijdserie kan het aantal AR termen p worden afgelezen. Voor p wordt dan de grootste waarde waarvoor de PACF significant afwijkt van nul gekozen. Hierbij geven de stippellijnen in het ACF en PACF het 95% significantieniveau aan [18].

Een snellere en nauwkeurigere manier dan het aflezen van de ACF en PACF is om te kijken naar het Akaike's informatie criteria (AIC). Dit criteria is gebaseerd op het volgende principe. Het verhogen van de orde van het model kan een hogere log likelihood tot gevolg hebben, echter het model zal hier wel complexer van worden. Het aantal parameters moet zodanig worden gekozen dat het $AIC(k)$ wordt geminimaliseerd. Hierbij is k het aantal parameters en er geldt dat:

$$AIC(k) = -2\log \text{likelihood} + 2k .$$

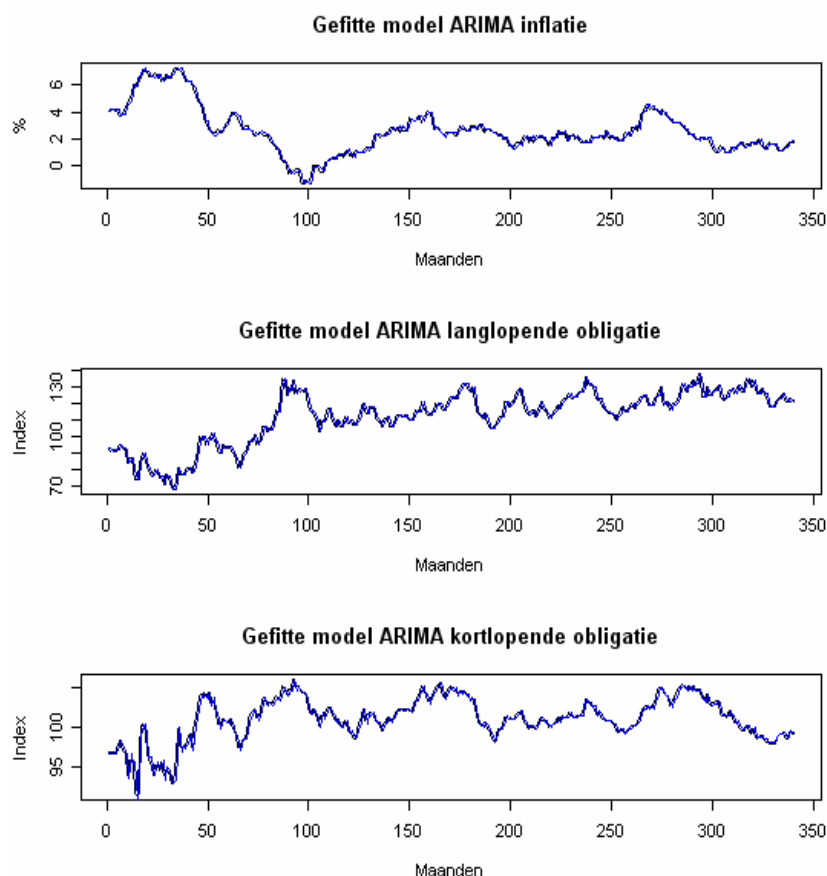
Het pakket R bezit een functie waarvoor voor $ARIMA(p,k,q)$ modellen kan worden bepaald wat de optimale p , k en q zijn. Deze functie, `auto.arima` (onderdeel van het forecast package), bepaalt aan de hand van het AIC de optimale order. In de onderstaande tabel worden de resultaten van deze functie weergegeven voor de inflatie, alsmede voor de kort- en langlopende obligatie-index.

Data	Model	AIC	AR ₁	AR ₂	MA ₁	MA ₂
Inflatie	ARIMA(2,1,2)	64,72	1,6708	-0,8412	-1,7383	0,9545
Langlopend	ARIMA(0,1,0)	1784,08	-	-	-	-
Kortlopend	ARIMA(2,1,2)	821,5	0,5495	0,3753	-0,3674	-0,6252

Tabel 2: De optimale modellen en de bijbehorende parameters.

Het blijkt dus voor alle drie de datasets dat differentiëren van de orde 1 voldoende blijkt te zijn. Opvallend is dat als optimale model voor de langlopende obligatie-index het $ARIMA(0,1,0)$ model wordt gekozen: dit komt gewoon neer op een random walk [30].

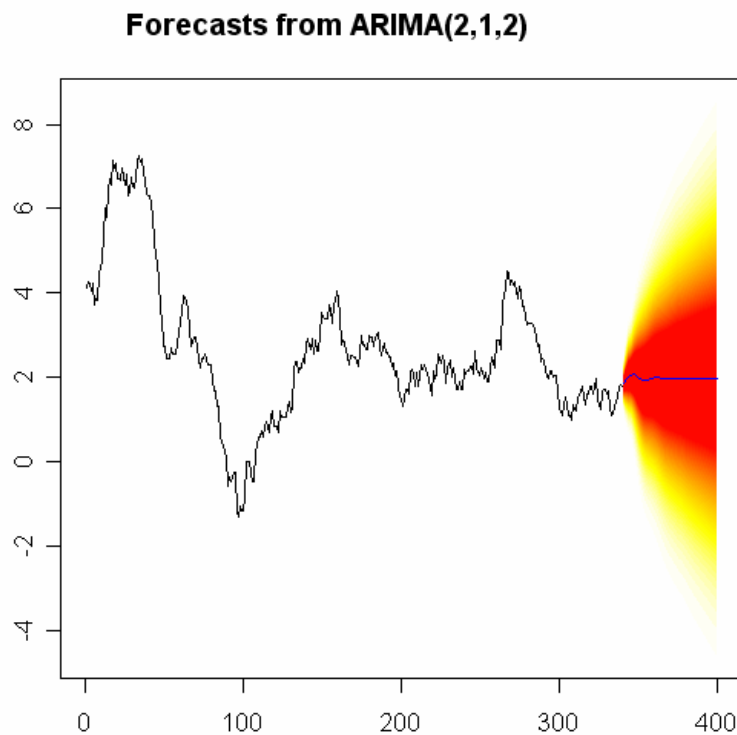
Figuur 16 geeft de fit van het model aan.



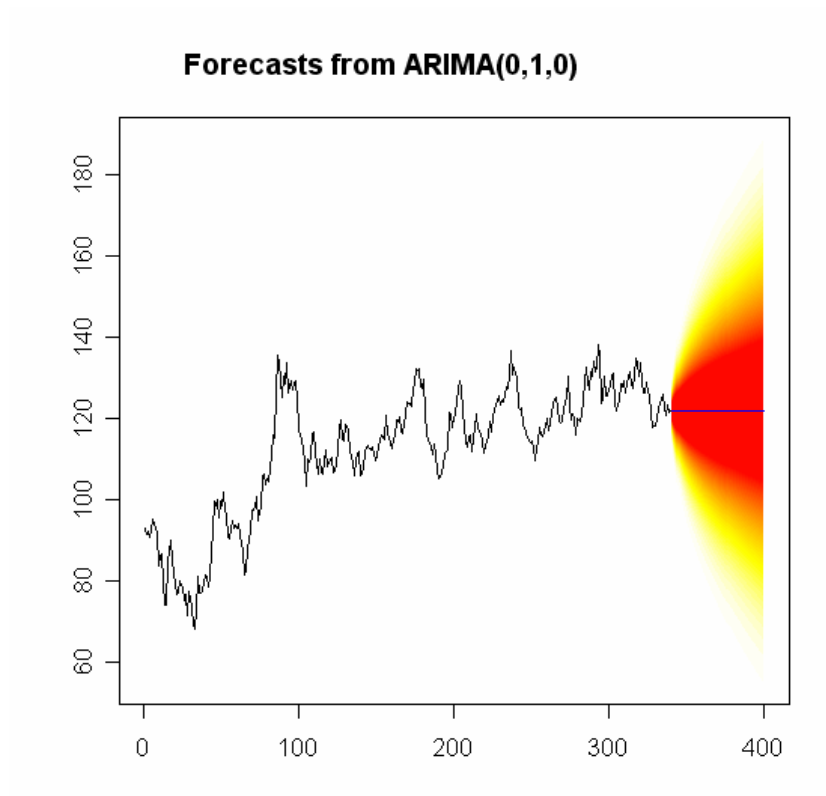
Figuur 16: De gefitte ARIMA modellen.

Voorspellen met het ARIMA model

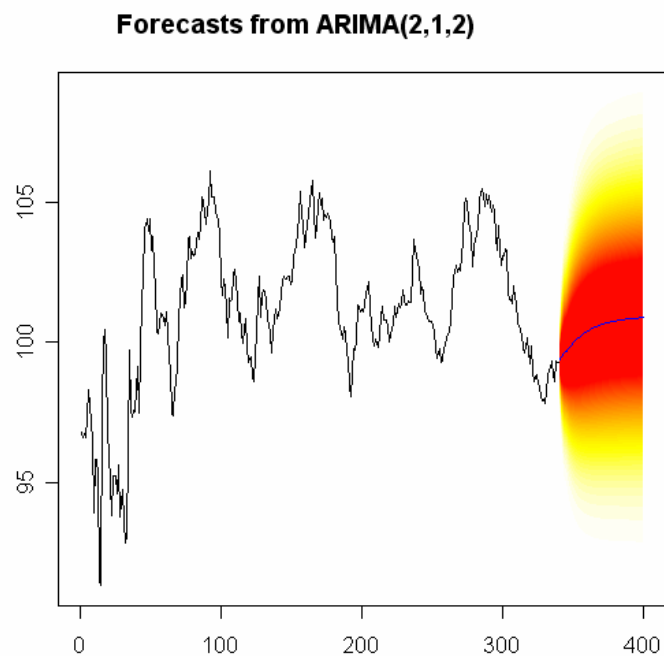
Nu de ARIMA modellen geschat zijn kunnen voorspellingen worden gedaan voor toekomstige waarden van de inflatie en de kort- en langlopende obligatie-index. De onderstaande figuren geven de voorspellingen voor een tijdperiode van 60 maanden, wat neerkomt op 5 jaar. Dit is de normale looptijd van een SA contract bij Zwitserleven.



Figuur 17: Voorspellingen voor de inflatie, 60 maanden vooruit.



Figuur 18: Voorspellingen voor de langlopende obligatie-index, 60 maanden vooruit.



Figuur 19: Voorspellingen voor de kortlopende obligatie-index, 60 maanden vooruit.

Zoals te verwachten is, is de voorspelling voor de langlopende obligatie-index gewoon de verwachting van de random walk (een rechte lijn, zoals te zien is in figuur 18).

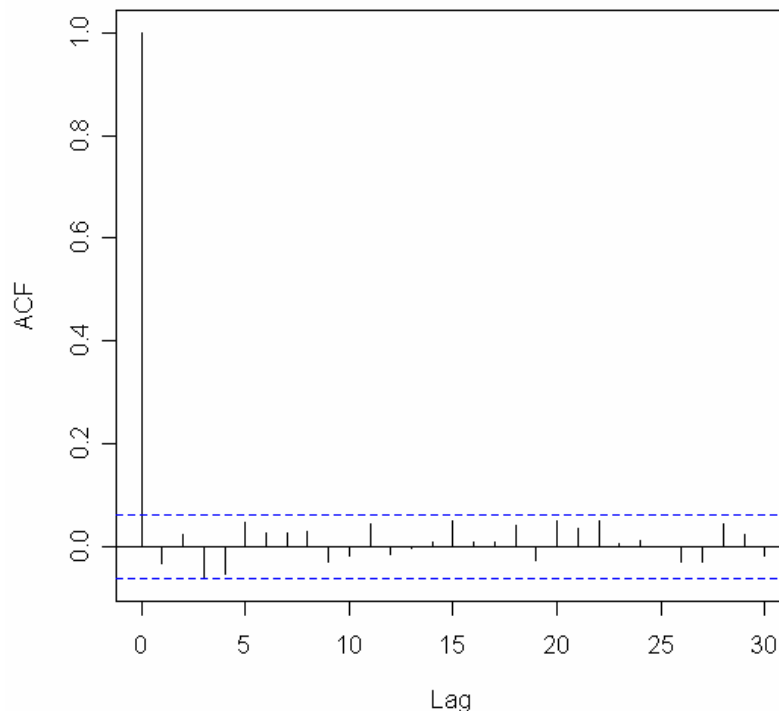
Diagnostic check ARIMA

Nu het model gekozen is en de parameters geschat zijn, moet worden gekeken of het model ook daadwerkelijk geschikt is. Hiervoor moet worden gekeken naar de residuen van het model. De statistische eigenschappen (bijvoorbeeld verwachting en variantie) van deze residuen moeten consistent zijn met een witte ruis proces. Om te bepalen of dit inderdaad het geval is, kunnen een aantal stappen worden ondernomen.

- 1) Er kan worden gekeken naar de ACF van de residuen. Als de residuen lijken op een witte ruis proces, dan moet de ACF van de residuen overeenkomsten vertonen met de ACF van een witte ruis proces.
- 2) Er zijn verschillende testen om te bekijken of de residuen lijken op white noise. Twee bekende testen zijn de Portemanteau test en de Ljung-Box test.

Ter vergelijking wordt in figuur 20 de ACF van een witte ruis proces weergegeven.

ACF van white noise



Figuur 20: De ACF van een witte ruis proces.

Zoals te zien is in figuur 20, geldt bij een witte ruis proces dat de observaties volkomen onafhankelijk zijn van elkaar.

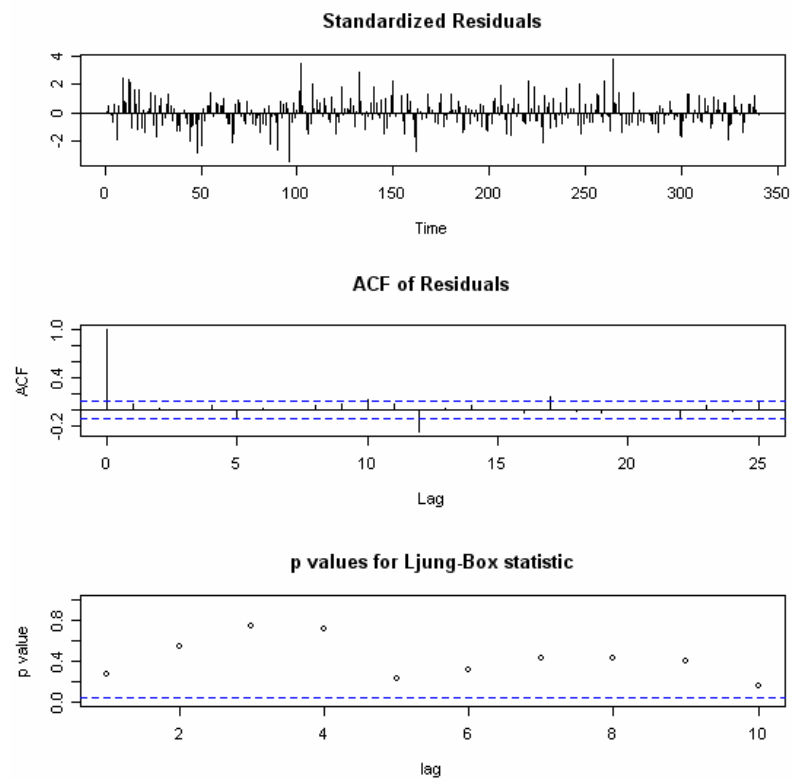
De onderstaande figuren geven de uitkomsten van de functie `tsdiag`. Deze functie kijkt naar de kenmerken van de residuen van het geschatte model. De residuen worden geplotted, de ACF en PACF van de residuen, en uiteindelijk wordt de Ljung-Box statistiek geplotted. De Ljung-Box test kan worden gebruikt om te testen of de residuen random zijn.

Hier geldt:

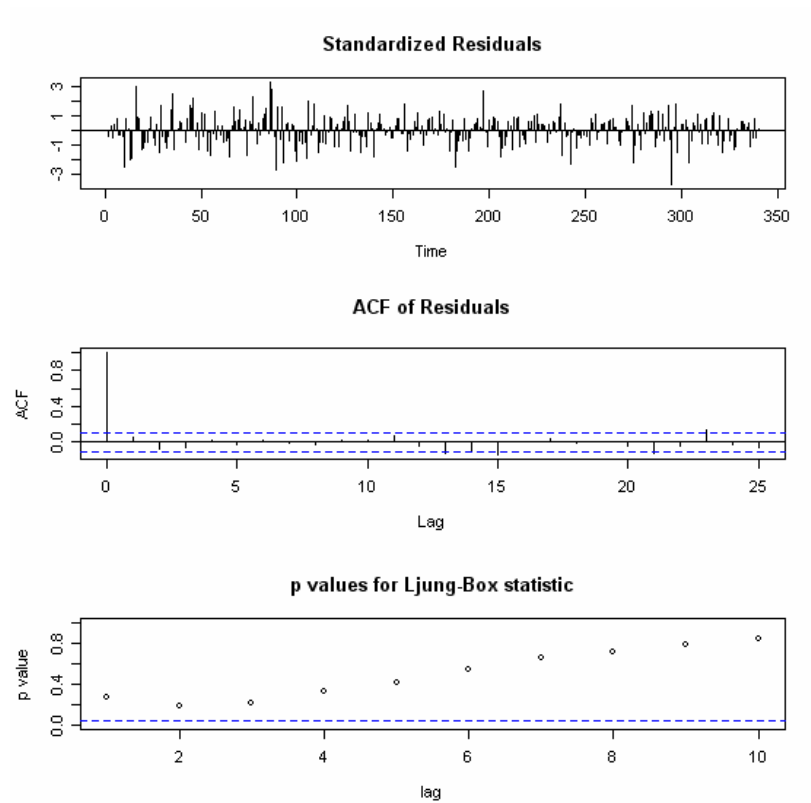
H_0 : *Onafhankelijk verdeelde residuen,*

H_1 : *De residuen zijn niet onafhankelijk verdeeld.*

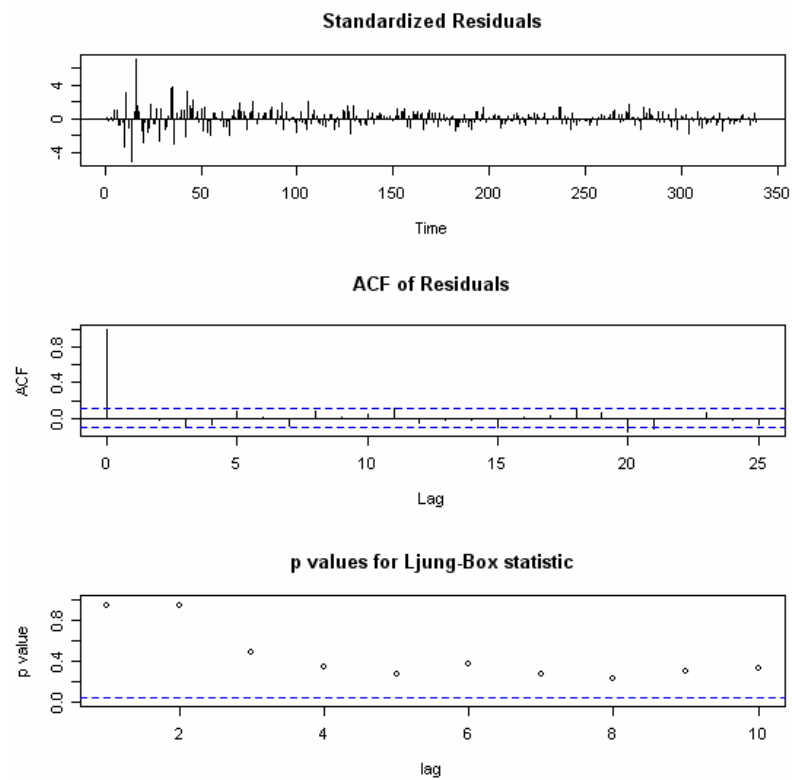
Deze test is gebaseerd op de ACF plot. De nulhypothese wordt verworpen bij een p waarde kleiner dan 5%.



Figuur 21: Kenmerken residuen ARIMA model inflatie.



Figuur 22: Kenmerken residuen ARIMA model kortlopende obligatie-index. .



Figuur 23: Kenmerken residuen ARIMA model langlopende obligatie-index.

Wanneer de figuren 21 t/m 23 worden bekeken blijkt dat de residuen van de geschatte ARIMA modellen dezelfde statistische eigenschappen hebben als witte ruis. Wanneer de ACF van de residuen van het geschatte model wordt vergeleken met de ACF van witte ruis, lijken ze overeen te komen. De residuen hebben geen onderlinge correlatie.

Wanneer naar de p waarde van de Ljung-Box test wordt gekeken, blijkt dat bij geen van de modellen de nulhypothese (de residuen zijn onafhankelijk verdeeld) wordt verworpen.

Aan de hand van de ACF en de Ljung-Box test, lijken ARIMA modellen dus goed geschikt te zijn om de economische tijdseries te modelleren.

6.6.2 VAR model

Om te bepalen welke orde optimaal is voor het VAR model, wordt gebruik gemaakt van de functie `VARselect`. Het VAR package kan omgaan met data met daarin een trend. Deze trend wordt opgenomen in het model.

De reden dat het aandelenrendement niet wordt meegenomen, is dat wanneer een VAR model wordt geschat met daarin ook het aandelenrendement, er een model overblijft waarvan geen enkele regressievariabele significant is. Het rendement op de AEX index is namelijk erg volatiel. Voor het aandelenrendement zal dus een andere aanpak gekozen moeten worden. Dit volgt verderop in dit verslag.

De functie `VARselect` maakt gebruik van een aantal informatie criteria om de optimale orde te bepalen. De informatie criteria die beschouwd worden, zijn de volgende:

- Het Akaike's informatie criteria:

$$AIC(n) = \ln \det(\tilde{\Sigma}_\varepsilon(n)) + \frac{2}{T} nK^2.$$

Dit kan worden beschouwd als de multivariate versie van het criteria zoals beschreven in 7.6.1.

- Het Hannan and Quinn informatie criteria:

$$HQ(n) = \ln \det(\tilde{\Sigma}_\varepsilon(n)) + \frac{2 \ln(\ln(T))}{T} nK^2.$$

- Het Schwarz criteria:

$$SC(n) = \ln \det(\tilde{\Sigma}_\varepsilon(n)) + \frac{\ln(T)}{T} nK^2.$$

- De Forecast Prediction Error:

$$FPE(n) = \left(\frac{T + n^*}{T - n^*} \right)^K \det(\tilde{\Sigma}_\varepsilon(n)).$$

Hierbij is n de te beschouwen lag grootte, T het aantal datapunten (hier geldt dus $T = 340$), K het aantal variabelen in het model (hier geldt dus $K = 3$) en $\tilde{\Sigma}_\varepsilon(n) = T^{-1} \sum_{t=1}^T \hat{\varepsilon}_t \hat{\varepsilon}_t'$.

De waarde p waarvoor geldt dat die een gekozen criterium minimaliseert, is de optimale lag grootte. Deze criteria zijn multivariate uitbreidingen van de standaard informatie

criteria, die gebruik maken van de log likelihood functie (zoals is beschreven in paragraaf 6.6.1).

Wanneer deze functie wordt toegepast op de dataset, blijkt dat twee van de vier criteria 1 als optimale orde geven en twee criteria beschouwen orde 3 als optimaal. Aangezien de verschillen tussen orde 1 en orde 3 minimaal zijn, wordt een VAR(1) model gekozen met een trend en een constante.

Vervolgens kunnen de parameters A en Σ_ε van het model worden geschat met behulp van het *vars* package. Dit gebeurt met behulp van Ordinary Least Squares (OLS). In de Triple A scenario generator worden de parameters geschat met behulp van de Yule-Walker schattingstechniek. Wat deze technieken precies inhouden wordt onder andere beschreven door Van der Vaart in [29].

Volgens Triple A ([11]) geeft de Yule Walker schattingstechniek betere voorspellingen dan OLS, vanwege het feit dat de Yule-Walker techniek rekening houdt met de multi-periode dynamiek van het VAR model, terwijl OLS alleen kijkt naar de voorspelfouten een periode verder. Asymptotisch is het verschil tussen de Yule-Walker en least squares schatters te verwaarlozen volgens [29]. Ze bezitten dezelfde normale limiet verdeling. Er dient verder onderzoek te worden gedaan naar de vraag of de Yule-Walker schattingstechniek betere voorspellingen geeft dan OLS.

Restricted VAR model

Een restricted VAR model is een VAR model waarbij de regressievariabelen die niet significant zijn verwijderd worden uit het model. Stel dat er een regressie wordt uitgevoerd die er als volgt uitziet:

$$X_t = c + \alpha X_{t-1} + \varepsilon_t.$$

De regressievariabele X_{t-1} is niet significant, dat wil zeggen dat geldt voor de nulhypothese dat:

$$H_0 : \alpha = 0.$$

De alternatieve hypothese is dat de regressievariabele significant is, dus:

$$H_1 : \alpha \neq 0.$$

De t -waarde kan worden bepaald:

$$\hat{t} = \frac{\hat{\alpha}}{se(\hat{\alpha})}.$$

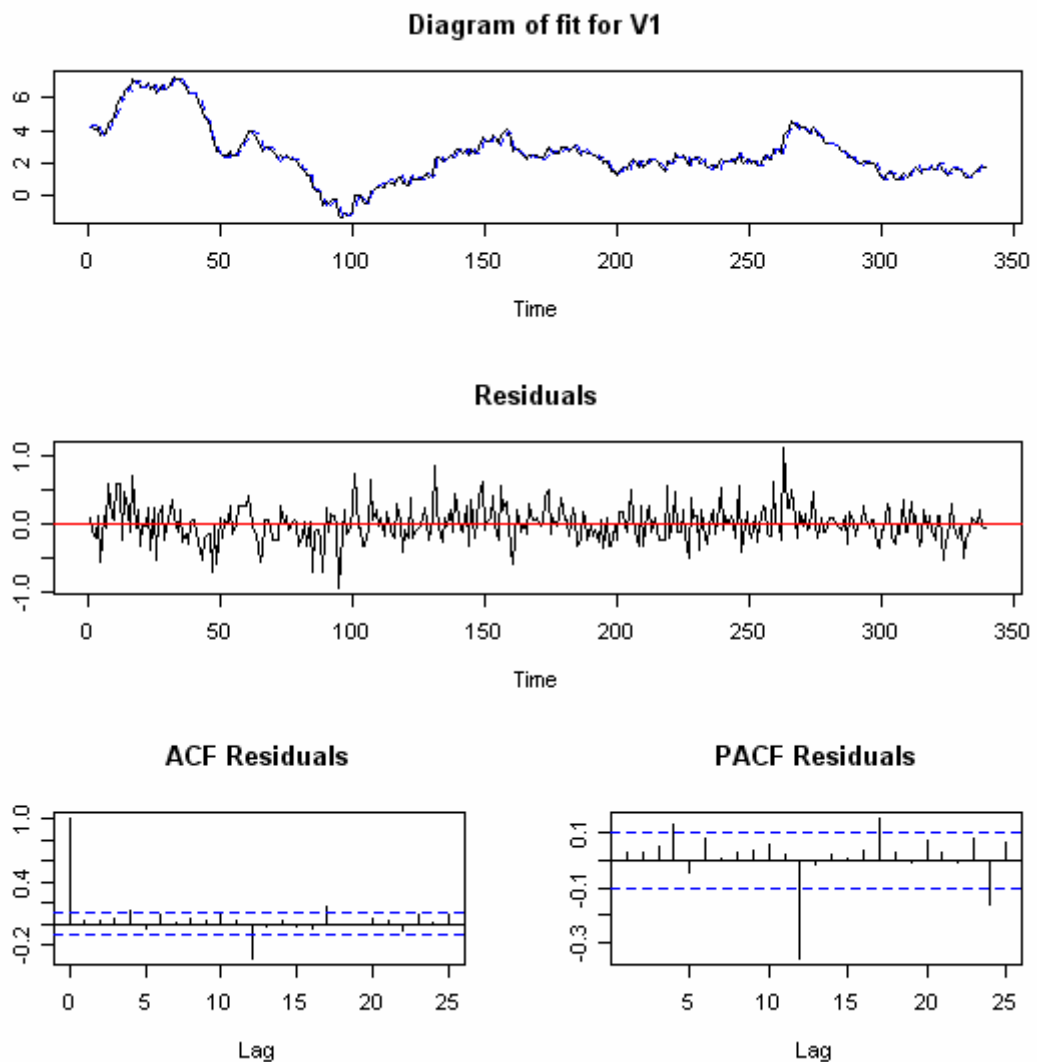
Waarbij $se(\hat{\alpha})$ de standaard error van de geschatte regressievariabele $\hat{\alpha}$ is, die bepaald wordt tijdens de regressie. De nulhypothese (geen significantie) wordt verworpen wanneer geldt dat $-1.66 < \hat{t} < 1.66$ (waarden verkregen uit de Student t -tabel met 334 vrijheidsgraden en $p = 5\%$).

Wanneer een restricted VAR model wordt opgesteld voor de gekozen dataset, blijkt dat een aantal regressievariabelen niet significant zijn. Deze worden in het model op nul gesteld. In de onderstaande tabel worden de significante en niet-significante variabelen aangegeven. Hierbij is de term 1 als hij significant is en 0 als hij niet significant is.

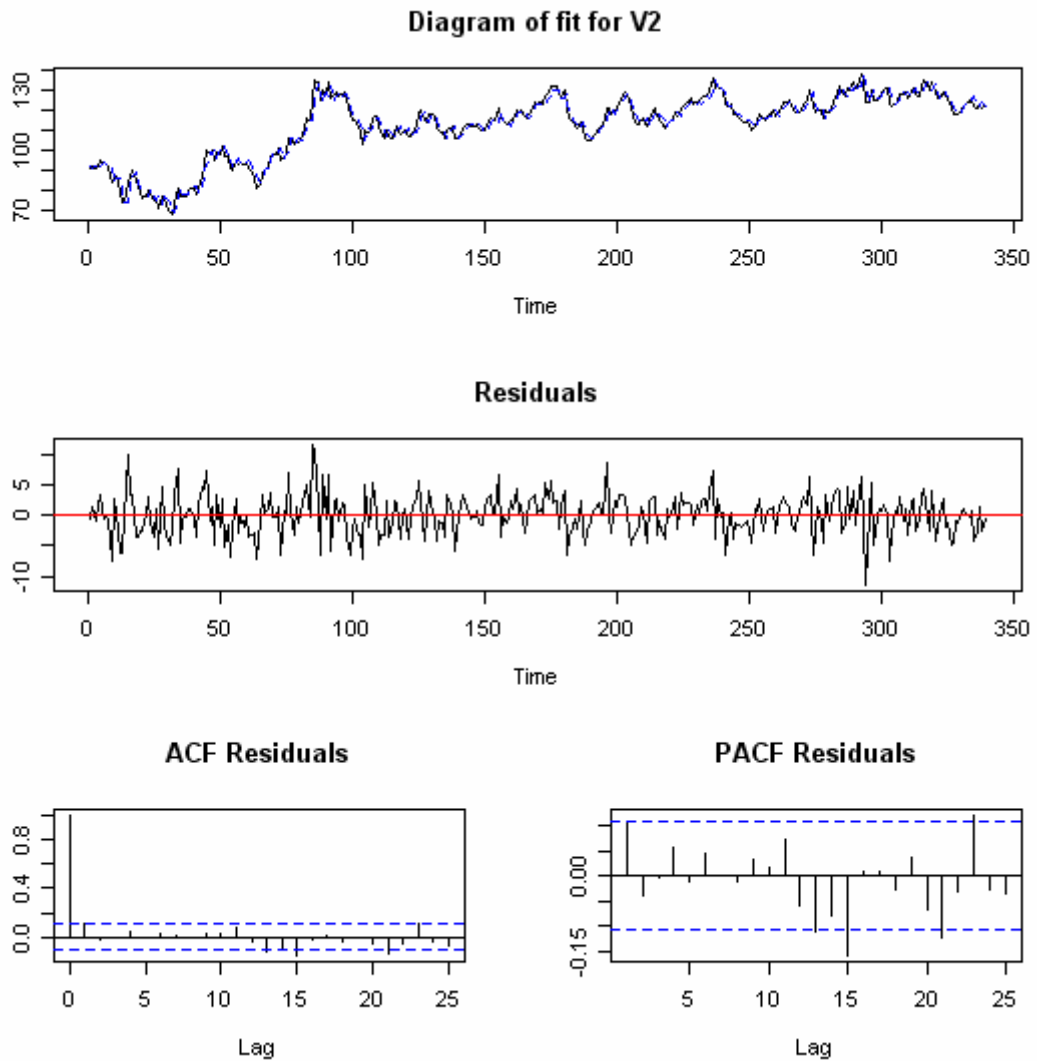
	Constante	Trend	$X1_{t-1}$	$X2_{t-1}$	$X3_{t-1}$
$X1_t$	1	0	1	0	1
$X2_t$	0	1	1	1	1
$X3_t$	0	1	0	1	1

Tabel 3: De significante en niet significante regressievariabelen.

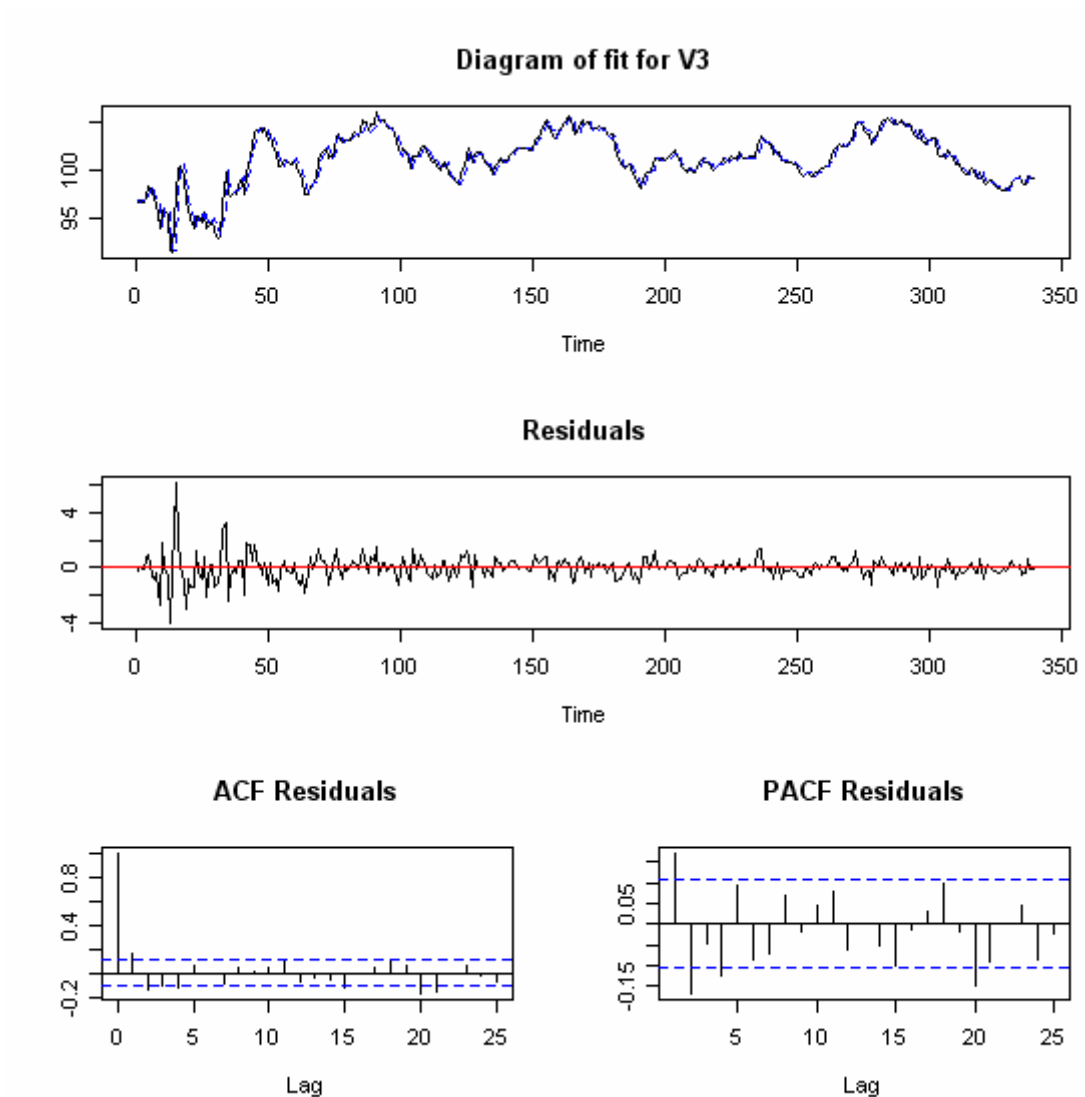
De onderstaande figuren geven de eigenschappen van het gefitte model weer. In de grafieken wordt X1 aangegeven als V1, X2 als V2 en X3 als V3.



Figuur 24: Gefitte model voor de inflatie.



Figuur 25: Gefitte model voor de langlopende obligatie-index.



Figuur 26: Gefitte model voor de kortlopende obligatie-index.

Het restricted VAR model heeft dan de volgende parameters.

	Constante	Trend	$X1_{t-1}$	$X2_{t-1}$	$X3_{t-1}$
$X1_t$	2.374	0	0.97	0	-0.023
$X2_t$	0	0.013	-0.473	0.859	0.148
$X3_t$	0	0.002	0	-0.017	1.016

Tabel 4: De coëfficiënten van het geschatte model.

Het var package bevat een functie waarmee Σ_ε kan worden bepaald. De covariantie matrix van de error termen ziet er als volgt uit.

$$\Sigma_\varepsilon = \begin{pmatrix} 0.071 & -0.120 & -0.014 \\ -0.120 & 10.744 & 2.015 \\ -0.014 & 2.015 & 0.681 \end{pmatrix}$$

Het bepalen van de covariantie matrix gebeurt door middel van een Cholesky factorisatie. Bij een Cholesky factorisatie vindt een decompositie van de matrix Σ_ε plaats, hierbij geldt dat $\Sigma_\varepsilon = PP'$. Hierbij geeft P' de getransponeerde van matrix P aan. Een mogelijkheid om de error termen met de vereiste eigenschappen te genereren is met behulp van een Cholesky factorisatie. Er moet gelden dat Σ_ε een positief definitie matrix is. Een matrix is positief definitief wanneer geldt dat $x'Ax > 0$ voor alle x (waarbij $x \neq 0$).

Wanneer matrix P gevonden is, is het makkelijk om error termen met de vereiste eigenschappen te genereren. Dit gebeurt dan door eerst witte ruis met de volgende eigenschappen te genereren:

$$E[u_t] = 0$$

$$E[u_t u_s'] = I \quad \text{als } s = t.$$

$$E[u_t u_s'] = 0 \quad \text{als } s \neq t$$

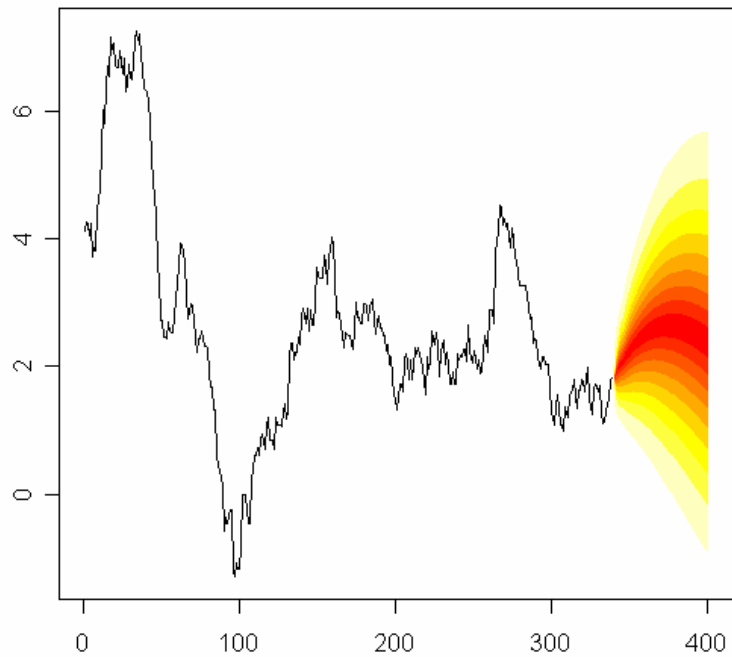
Dit witte ruis proces is niet onderling gecorreleerd, het is standaard normaal verdeeld. De error termen met de gewenste eigenschappen kunnen worden gegenereerd door eerst een vector met u_t termen te genereren en deze vervolgens te vermenigvuldigen met P .

Dus er geldt dat $u_t = P\varepsilon_t$.

Voorspellen met het VAR model

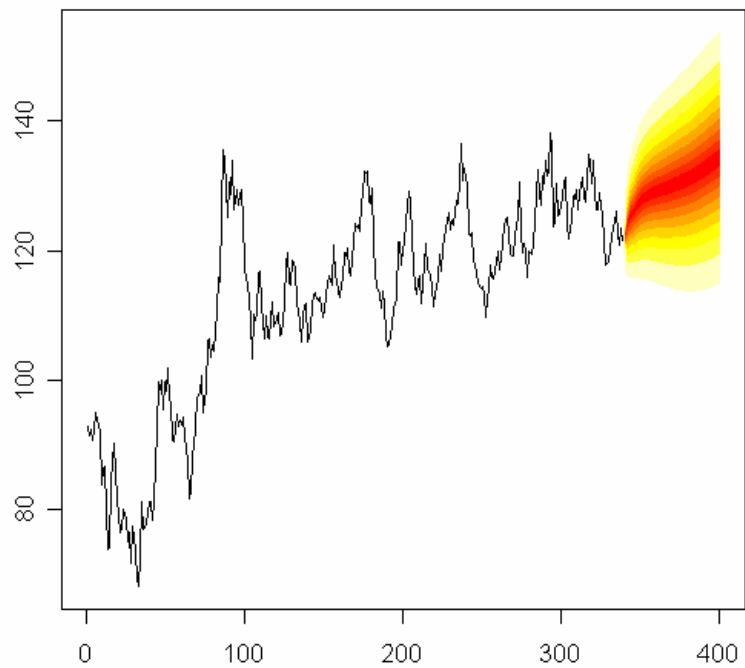
Met het geschatte model is het mogelijk om de toekomstige tijdreeks te voorspellen voor de aankomende 5 jaar. In deze paragraaf worden enkele figuren gegeven die zijn gegenereerd door het `vars` package. Dit package beschikt over een functie `predict` waarmee het mogelijk is om voorspellingen te doen voor de toekomstige tijdreeks op basis van het geschatte model. In figuur 27 wordt dit grafisch weergegeven. Iedere kleurschakeling in dit figuur geeft een kritieke waarde aan, met stappen van 0.1.

Fanchart for variable V1



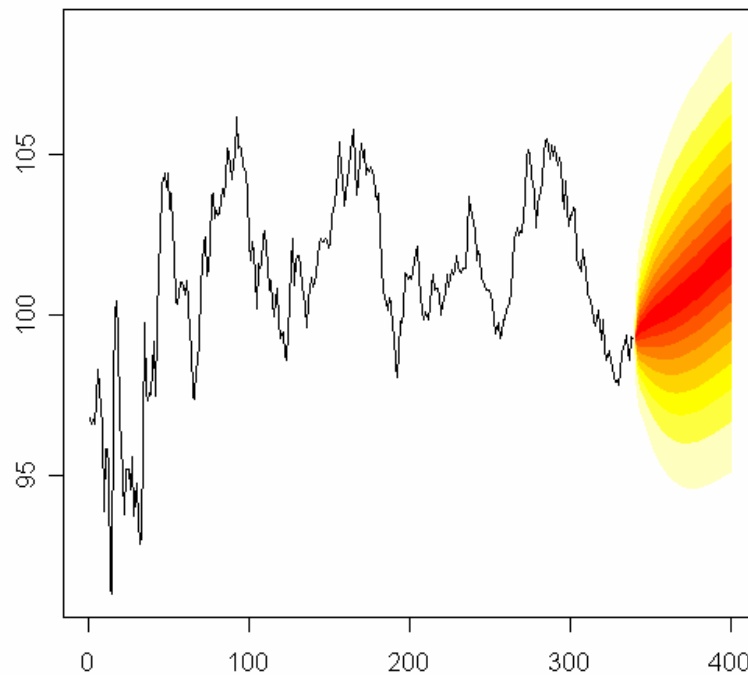
Figuur 27: Voorspellingen voor de inflatie, VAR model.

Fanchart for variable V2



Figuur 28: Voorspellingen voor de langlopende obligatie-index, VAR model.

Fanchart for variable V3



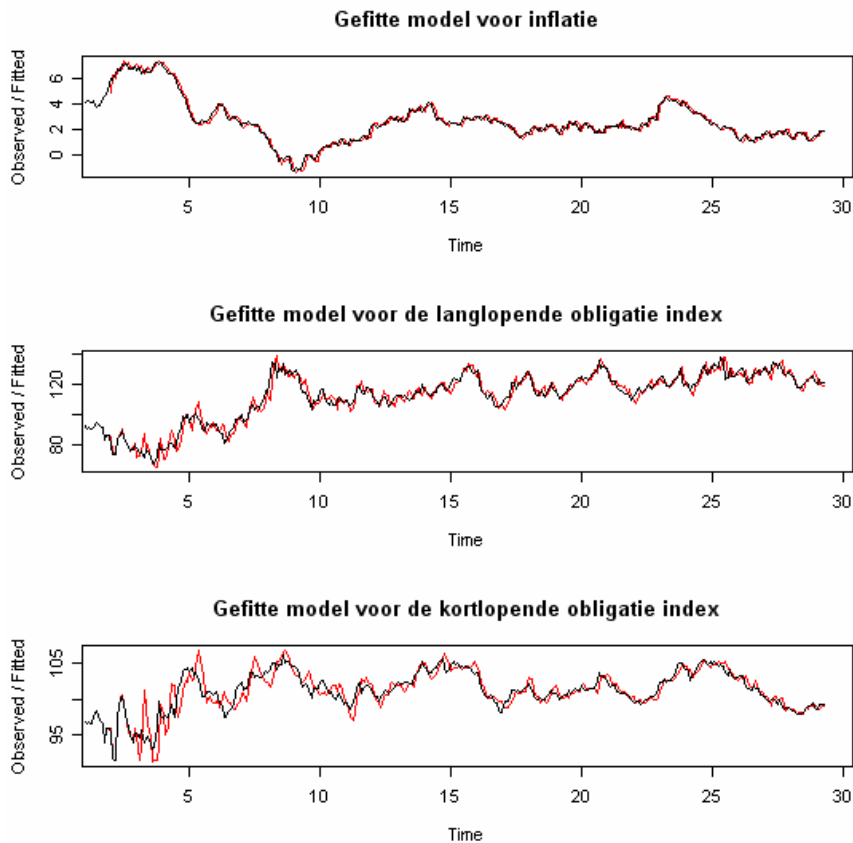
Figuur 29: Voorspellingen voor de kortlopende obligatie-index, VAR model.

Diagnostic check VAR

Om te kijken of het VAR model geschikt is, moet een diagnostic check worden uitgevoerd. Lijken de residuen inderdaad op een multivariaat witte ruis proces, zoals verondersteld? Dit principe is eerder beschreven in paragraaf 6.6.1. Een methode om dit te testen is met de multivariate Portmanteau test. Toegepast op het geschatte VAR model blijkt echter dat de residuen volgens deze test niet lijken op een multivariaat witte ruis proces. Volgens Francq en Raissi [34] blijkt echter dat deze test misleidende resultaten kan geven wanneer de witte ruis niet onafhankelijk is. In paragraaf 6.7 wordt onderzocht of het model toch goede voorspellingen geeft.

6.6.3 Holt Winters methode

Wanneer de Holt Winters methode wordt toegepast op de dataset, leidt dit tot de volgende gefitte modellen. De functie `HoltWinters` bepaalt automatisch de parameters van het model.



Figuur 30: Gefitte modellen en voorspelling met Holt Winters (exponential smoothing).

Met behulp van de geschatte parameters kunnen nu voorspellingen worden gedaan voor de komende 60 maanden.

In de volgende paragraaf zullen de drie beschreven modellen worden vergeleken. Welk model kan als optimaal worden beschouwd?

6.7 Conclusie: welk model is het best geschikt om scenario's te genereren voor de inflatie, kort- en langlopende obligatie-index?

Het is belangrijk om een geschikt model te kiezen om voorspellingen te doen voor minstens vijf jaar vooruit, aangezien dit de lengte van een SA contract is. De drie modellen zoals beschreven in de voorgaande secties zullen worden beschouwd, dus

1. ARIMA;
2. VAR;
3. Holt Winters (Exponential Smoothing).

Alleen kijken naar de fit van het model op de bekende data en op basis daarvan een model kiezen, is niet voldoende. De mate waarin een model de dataset fit hoeft niks te zeggen over hoe goed toekomstige waarden worden voorspeld.

Om te testen welk van de drie modellen de beste voorspellingen geeft, wordt de volgende methode gebruikt. De totale dataset is 340 maanden lang. Van deze dataset wordt 227 maanden (67%) gebruikt om een optimaal model op te schatten. Dit model doet een voorspelling voor de komende 60 maanden (5 jaar). Deze voorspelling wordt vergeleken met de werkelijke waarde van de tijdserie, en de RMSE (Root Mean Squared Error) wordt bepaald. Hierbij geldt dat:

$$RMSE = \sqrt{T^{-1} \sum_{t=1}^T e_{t+1}^2},$$

waarbij $e_{t+1} = x_{t+1} - \hat{x}_{t+1}$ met x_{t+1} de werkelijke waarde op $t + 1$ en \hat{x}_{t+1} de voorspelling op tijdstip $t + 1$.

Dit wordt herhaald, er worden modellen gefit voor 228, 229, ..., 280 maanden, wat leidt tot verschillende voorspellingen vijf jaar vooruit. Dan wordt de gemiddelde RMSE bepaald. Deze procedure wordt voor ieder van de drie modellen gevolgd. Het model met het laagste RMSE voor de variabele (inflatie, kort- en langlopende obligatie) kan dan worden aangewezen als beste model voor de dataset.

Hierbij geldt dat voor het ARIMA model telkens het optimale model wordt geschat met behulp van de functie `auto.arima` in R. Deze bepaalt de optimale orde van het ARIMA model. Voor het Holt Winters model wordt telkens het optimale model geschat met behulp van de functie `HoltWinters`. Voor het VAR model is een iets andere aanpak gekozen. Bij het schatten van het VAR model is telkens gebruik gemaakt van een VAR(1) model met daarin een trend en een constante. Wanneer blijkt dat het VAR(1) model minder goed presteert dan het ARIMA en Holt Winters model, dan kan alsnog voor iedere schattingsperiode een optimaal (met een hogere orde) VAR model worden geschat in plaats van een VAR(1) model. Wanneer echter zou blijken dat het VAR(1) model beter presteert dan het ARIMA en Holt Winters model, dan zou dit betekenen dat een optimaal hogere orde VAR model nog beter zou presteren dan een ARIMA en Holt Winters model. Voor de aankomende vijf jaar kan dan een optimaal VAR model worden geschat (met behulp van de functie `VARselect`), welke mogelijk van een hogere orde is. Samenvattend wil dit zeggen dat wanneer een niet-optimaal VAR(1) model al beter is dan een ARIMA en Holt Winters model, een optimaal VAR(p) model dit zeker ook is.

In tabel 5 staat de gemiddelde RMSE van de drie modellen weergegeven.

Variabele	Gem. RMSE ARIMA	Gem. RMSE HW	Gem. RMSE VAR(1)
Inflatie	1,52	2,17	1,04
Langlopende obligatie	8,78	9,23	6,90
Kortlopende obligatie	2,66	2,77	2,17

Tabel 5: Gemiddelde RMSE van de drie modellen.

Op dit moment zijn de parameters van het VAR(1) model geschat met behulp van OLS. In de literatuur wordt echter beschreven dat de Yule-Walker schattingstechniek betere voorspellingen geeft. Aangezien het VAR(1) model geschat met OLS al betere voorspellingen geeft dan ARIMA en Holt Winters, is het interessant om te onderzoeken of scenario's gegenereerd met Yule-Walker betere voorspellingen geven dan scenario's gegenereerd met OLS.

Conclusie:

Uit deze resultaten blijkt dat het VAR(1) model voor alle drie de variabelen de laagste RMSE heeft. Hieruit kan dus worden geconcludeerd dat een VAR model de beste resultaten geeft bij de gebruikte dataset en de gebruikte technieken.

NB: Waarom wordt niet gekozen voor een VAR model met daarin de AEX index?

Er wordt veelal aangenomen dat er een model gekozen moet worden waarin de correlatie tussen de aandelenindex en de obligatie-index meegenomen is. In de voorgaande secties is echter gebruik gemaakt van een VAR(1) model waarin alleen de inflatie, kort- en langlopende obligatie index meegenomen is. Waarom is dit het geval?

1. Wanneer de AEX index wordt meegenomen in het VAR(1) model, blijkt dat voor deze dataset het model slechtere voorspellingen geeft voor de inflatie en kort- en langlopende obligatie index, dan wanneer een VAR(1) model wordt geschat met daarin alleen de inflatie, kort- en langlopende obligatie index.

Voor de dataset met 294 punten (de AEX index loopt vanaf 1983), leidt dit tot de volgende waarden voor de RMSE.

	Gem. RMSE, VAR(1) met AEX index	Gem. RMSE, VAR(1) zonder AEX index
Inflatie	1,61	1,56
Langlopende obligatie index	3,48	2,53
Kortlopende obligatie index	8,72	5,67

Tabel 6: Gemiddelde RMSE voor VAR(1) met en zonder AEX index.

In de tabel is te zien dat de gemiddelde RMSE voor het VAR(1) model zonder de AEX index het laagst is.

2. Wanneer de AEX index wordt voorspeld met het VAR(1) model met vier variabelen, blijkt deze het aandelenrendement niet beter te voorspellen dan een

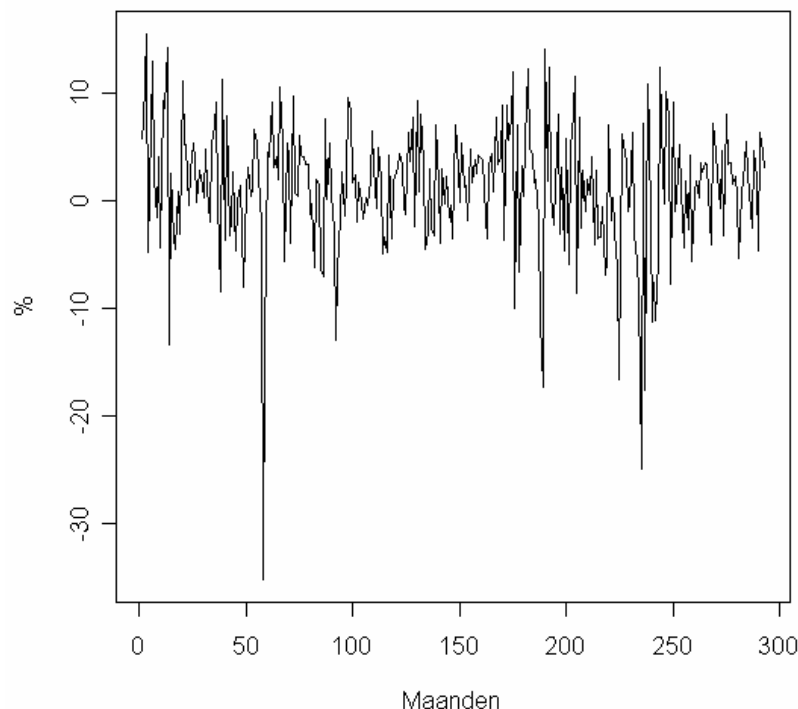
van de modellen die beschreven zullen worden in de volgende paragraaf (GARCH of RSLN). Hier is echter geen *uitgebreid* onderzoek naar gedaan.

Er dient dus verder onderzoek te worden gedaan (zoals genoemd in punt 2) naar de resultaten van de voorspellingen voor de AEX index van een VAR model ten opzichte van het GARCH of RSLN model, die zullen worden beschreven in de volgende paragraaf.

6.8 Het modelleren van het rendement op de AEX index

Het rendement op de AEX index vereist een andere aanpak dan de modellen die gebruikt zijn in de voorgaande paragrafen, zoals beschreven in de vorige paragraaf.

Logaritme van het maandelijkse rendement op de AEX



Figuur 31: De logaritme van het maandelijkse rendement op de AEX index.

In de literatuur zijn veel verschillende modellen beschreven om het rendement op de AEX index te modelleren. In deze paragraaf zullen de bekendste modellen worden beschreven, om vervolgens te kiezen voor één van deze modellen om de scenario's te genereren.

6.8.1 Aandelenrendementen beschrijven met een normale verdeling

Aandelenrendementen kunnen worden beschreven door een normale verdeling met de parameters μ en σ [12]. Er geldt dus voor het aandelenrendement r op tijdstip $t+1$ dat $r_{t+1} = N(\mu, \sigma^2)$, waarbij μ en σ verkregen zijn uit historische data. Er zijn een aantal redenen dat de normale verdeling geen goede voorspeller is voor het aandelenrendement. Ten eerste kan de normale verdeling alle waarden aannemen, dus ook alle negatieve waarden. Er bestaat dus een kans dat de prijs van een aandeel negatief wordt wanneer

de aandelenrendementen worden getrokken uit een normale verdeling. Dit is natuurlijk onmogelijk. Ook blijkt uit de data dat er meer grote negatieve rendementen voorkomen dan grote positieve rendementen. Dit duidt er ook op dat de rendementen niet normaal verdeeld zijn.

Tevens is de kans dat extreme prijsveranderingen voorkomen erg klein wanneer de normale verdeling als onderliggende verdeling wordt aangenomen. In de realiteit blijkt de kans op extreme prijsveranderingen een stuk groter. Ook houdt het model geen rekening met stochastische veranderingen in de volatiliteitparameter.

6.8.2 Aandelenrendementen beschrijven met een lognormale verdeling

Een andere methode om aandelenrendementen te beschrijven is met behulp van de geometric Brownian motion. Dit houdt in dat bekeken over een bepaalde tijdsperiode de rendementen lognormaal verdeeld zijn en dat de rendementen van de afzonderlijke tijdsperiodes onafhankelijk zijn. Als S_t de prijs van een aandeel is op tijdstip t , dan geldt:

$$\log \frac{S_t}{S_r} \sim N(\mu(t-r), \sigma^2(t-r)).$$

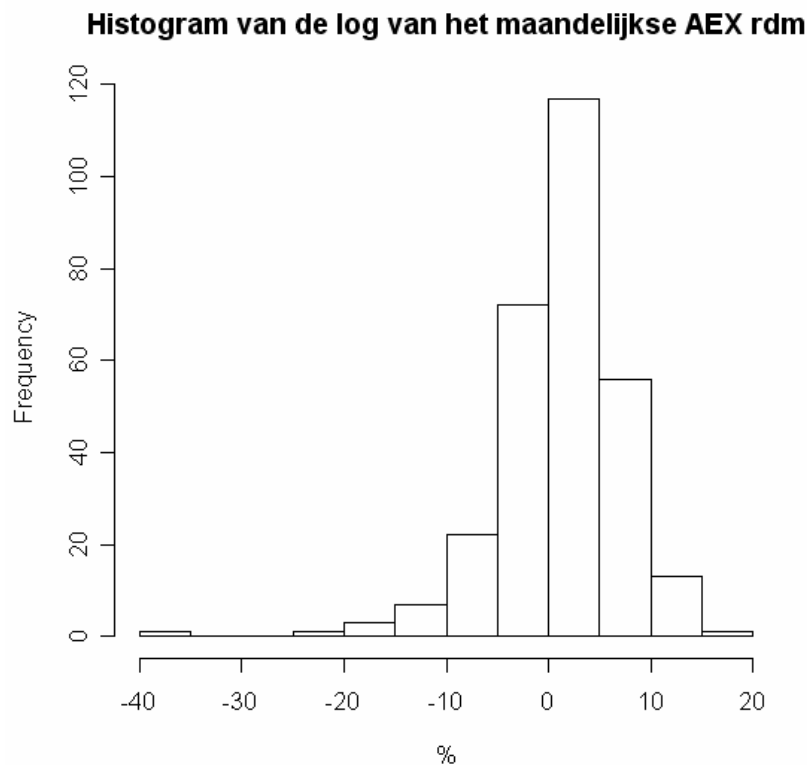
Waarbij μ en σ bekend worden verondersteld. Deze zijn bijvoorbeeld geschat uit een historische tijdreeks. Hier geldt dus dat het rendement $r_t = \frac{S_t}{S_r}$ lognormaal verdeeld is.

Voor $\log r_t = Y_t$ geldt dan het volgende:

$$Y_t = \mu + \sigma \varepsilon_t, \quad \varepsilon_t \text{ iid } \sim N(0,1).$$

Het blijkt dat dit model vooral voor kortere tijdperiodes goede voorspellingen geeft van het aandelenrendement, maar niet erg geschikt blijkt te zijn voor langere periodes. Tevens is de kans dat extreme prijsveranderingen voorkomen erg klein wanneer de lognormale verdeling als onderliggende verdeling wordt aangenomen. In de realiteit blijkt de kans op extreme prijsveranderingen een stuk groter. Ook houdt het model geen rekening met stochastische veranderingen in de volatiliteitparameter.

Figuur 32 geeft de histogram van de logaritme van het maandelijkse AEX rendement.



Figuur 32: Histogram van de logaritme van het maandelijkse AEX rendement.

6.8.3 Aandelenrendementen beschrijven met een AR(1) model

Het aandelenrendement kan worden beschreven met een AR(1) model zoals eerder beschreven in dit hoofdstuk. Hierbij geldt voor $\log r_t = Y_t$ het volgende:

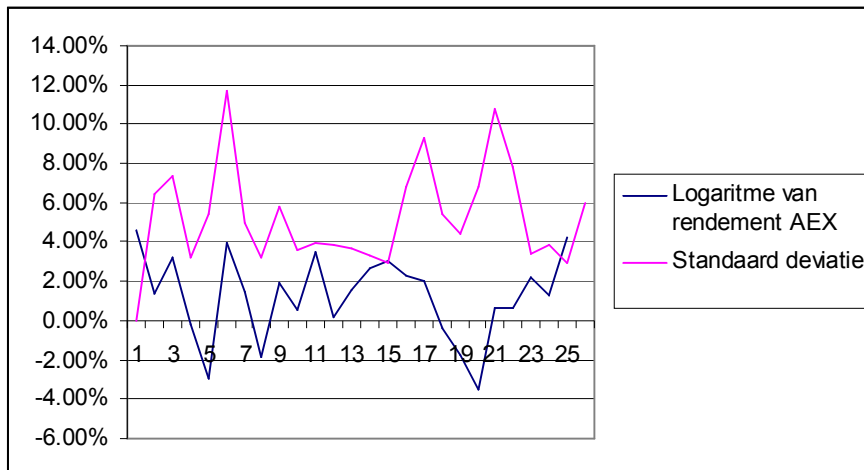
$$Y_t = \mu + \alpha(Y_{t-1} - \mu) + \sigma\varepsilon_t, \quad \varepsilon_t \text{ iid } \sim N(0,1).$$

Volgens Drijver [9] blijkt dat de autoregressieve term niet nodig is bij het beschrijven van de aandelenrendementen. Als de autoregressieve term niks toevoegt aan het model, dan komt dit model neer op de geometric Brownian motion zoals beschreven in de voorgaande paragraaf.

Wanneer binnen het softwarepakket R de optimale orde voor het AR model wordt bepaald met behulp van het Akaike Informatie Criteria blijkt dat deze 0 geeft als optimale orde. Er blijkt dus inderdaad dat de autoregressieve term geen toegevoegde waarde heeft en dus is het AR model niet geschikt om het aandelenrendement te genereren.

6.8.4 Tijdsvariërende variantie

De bovenstaande modellen hebben allemaal als nadeel dat ze een vaste volatiliteit veronderstellen. In de realiteit is dit natuurlijk niet het geval, zoals onderstaande grafiek laat zien.



Figuur 33: De logaritme van het rendement op de AEX index en de standaard deviatie over de tijd.

Uit figuur 33 blijkt dat in tijden van een hoog of laag rendement, de volatiliteit erg hoog was. Het is dus niet juist om een model te kiezen waarin de volatiliteit constant wordt gehouden gedurende de gehele periode, omdat dit in de realiteit niet het geval blijkt te zijn. In tijden van economische onzekerheid is de volatiliteit hoog.

Dit is een typische eigenschap van economische data. Perioden met een hoge volatiliteit worden afgewisseld met een periode met een lage volatiliteit en andersom. Dit fenomeen wordt volatiliteit clustering genoemd.

Wanneer er een hoge volatiliteit is, dan blijkt dat de volatiliteit hoog blijft voor een bepaalde tijd. Het tegenovergestelde gebeurt ook: wanneer de volatiliteit laag is blijft deze een bepaalde tijd laag te blijven. De volatiliteit is dus conditioneel autoregressief. Wanneer een reeks stochasten verschillende varianties hebben wordt dit heteroskedasticiteit genoemd.

De modellen die nu besproken zullen worden houden allen wel rekening met een tijdsvariërende variantie. Het eerste model is het ARCH (Autoregressief Conditioneel heteroskedasticiteit) proces.

6.8.5 Aandelenrendementen beschrijven met een ARCH proces

In het algemeen ziet een ARCH(1) proces er als volgt uit:

$$Y_t = \mu + \sigma_t \varepsilon_t \quad \varepsilon_t \text{ iid} \sim N(0,1).$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 (Y_{t-1} - \mu)^2.$$

Een uitbreiding van dit model is het GARCH(1,1) model. Dit staat voor Generalized Autoregressief Conditioneel Heteroscedasticity proces [27]. Deze ziet er als volgt uit:

$$Y_t = \mu + \sigma_t \varepsilon_t, \quad \varepsilon_t \text{ iid} \sim N(0,1).$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 (Y_{t-1} - \mu)^2 + \beta \sigma_{t-1}^2.$$

Hierbij geeft α_1 aan in hoeverre de huidige schok doorwerkt in de volatiliteit van morgen, en $\alpha_1 + \beta$ is de mate waarin dit effect uitdooft over de tijd. ARCH modellen zijn ontwikkeld door Engle (1982) en uitgebreid tot GARCH door Bollerslev (1986).

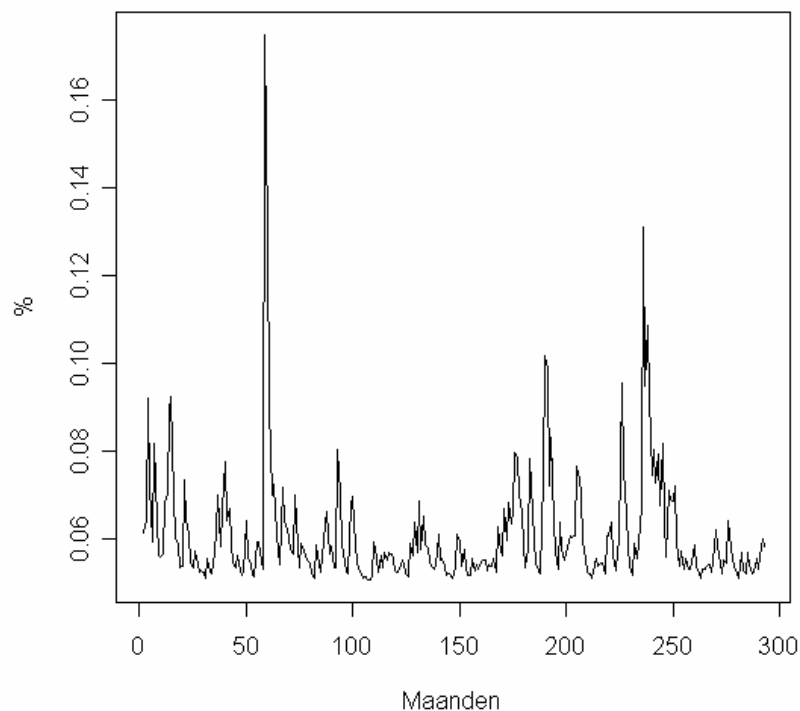
Wanneer het GARCH(1,1) model wordt geschat op de logaritme van het rendement op de AEX index, leidt dit tot de volgende parameters:

$$Y_t = 0.01123 + \sigma_t \varepsilon_t, \quad \varepsilon_t \text{ iid } \sim N(0,1)$$

$$\sigma_t^2 = 0.00162 + 0.22442 (Y_{t-1} - \mu)^2 + 0.35940 \sigma_{t-1}^2$$

De conditionele standaard deviaties van het geschatte model worden in de volgende grafiek weergegeven.

Conditionele standaard deviaties bij GARCH(1,1) proces



Figuur 34: De conditionele standaard deviatie bij het GARCH(1,1) proces.

Zoals in figuur 34 te zien is, geldt dat de conditionele standaard deviatie tijdens periodes met een laag rendement hoog is, zie ook figuur 31.

6.8.6 Aandelenrendementen beschrijven met het Markov Regime Switching Lognormal model (RSLN)

Het Markov Regime Switching lognormaal proces houdt rekening met de stochastische veranderingen in de volatiliteit van de rendementen. Het principe achter dit model is het volgende. Er wordt aangenomen dat de standaard deviatie K waarden (regimes) kan aannemen. Random wordt een van deze K waarden aangenomen. Ieder regime onderscheidt zich door verschillende parameters voor het model. De redenatie hierachter is dat de markt van tijd tot tijd wisselt tussen een stabiel regime met een lage volatiliteit en een minder stabiel regime met een hogere volatiliteit. Deze minder stabiele regimes treden bijvoorbeeld op in tijden van economische onzekerheid. Er wordt dus telkens gewisseld tussen K lognormale processen met ieder verschillende parameters.

Voor een RSLN(2) model geldt dat $K = 2$. Er zijn overgangskansen: de kans dat het proces naar regime 2 gaat als hij in 1 zit (p_{12}) en de kans dat het proces naar regime 1 gaat als hij in regime 2 zit (p_{21}). ρ_t geeft het regime aan in het interval $[t, t+1)$, $\rho_t = 1, 2$. De overgangskansen worden gegeven door:

$$p_{ij} = P[\rho_{t+1} = j \mid \rho_t = i], \quad i = 1, 2, j = 1, 2,$$

waarbij $\sum_j p_{ij} = 1$.

In totaal heeft dit model dus zes parameters, namelijk $\Theta = \{\mu_1, \mu_2, \sigma_1, \sigma_2, p_{12}, p_{21}\}$ [27].

6.9 Conclusie: welk model is optimaal om het aandelenrendement te voorspellen?

In deze paragraaf zullen drie verschillende modellen worden vergeleken, te weten:

- Geometric Brownian motion (log van de rendementen normaal verdeeld);
- Het GARCH(1,1) model;
- Het RSLN(2) model (hierbij geldt dus $K = 2$).

Het vergelijken van de modellen gebeurt aan de hand van de log likelihood functie en de likelihood ratio test. De log likelihood functie is als volgt gedefinieerd. Voor een vector onafhankelijke identiek verdeelde stochasten $X = (X_1, \dots, X_n)$ geldt dat de log likelihood functie er als volgt uitziet:

$$\theta \rightarrow \log L(\theta; x_1, \dots, x_n) = \sum_{i=1}^n \log p_{\theta}(x_i),$$

met p_{θ} de marginale dichtheid van een enkele X_i [27].

Er geldt dat een maximale waarde van de log likelihood geprefereerd wordt.

Het Society of Actuaries (SOA) heeft een Excel sheet waarin het RSLN(2) model geïmplementeerd is [35]. Voor de verkregen data worden de parameters voor het model geschat en ook wordt de log likelihood gegeven voor het geschatte model. Hoe de log likelihood voor het RSLN(2) model precies berekend wordt, staat uitgebreid beschreven in [28]. De log likelihood is tevens berekend voor het lognormale model en het GARCH(1,1) proces. De resultaten staan in tabel 7 weergegeven.

Model	Log likelihood	Aantal parameters
Geometric Brownian motion	406,21	2
GARCH(1,1) proces	415,19	4
RSLN(2) proces	418,16	6

Tabel 7: log likelihood van de verschillende modellen en het aantal parameters

Met de likelihood ratio test kan worden bepaald of het verschil in log likelihood tussen twee verschillende modellen ook significant is. Wanneer een model een hogere log likelihood heeft maar wel veel meer parameters, moet worden bekeken of het verschil in log likelihood wel opweegt tegen het extra aantal parameters. Dit kan worden bekeken met de likelihood ratio test. De likelihood ratio test vergelijkt verschillende modellen. Model 1 is een model met k_1 parameters en model 2 een model met k_2 parameters, waarbij geldt $k_2 > k_1$. Hierbij is l_1 de log likelihood van model 1 en l_2 is de log likelihood van model 2. De nulhypothese van de likelihood ratio test is:

H_0 : Geen significante verbetering in model 2,

H_1 : Wel significante verbetering in model 2.

De test statistiek is $2(l_2 - l_1)$. Onder de nulhypothese heeft de test statistiek een χ^2 -verdeling met $k_2 - k_1$ vrijheidsgraden [28].

Allereerst is de nulhypothese getest dat er geen significante verbetering optreedt door het GARCH(1,1) model te gebruiken ten opzichte van het lognormale model. De nulhypothese wordt verworpen, aangezien geldt dat $p = 0.000126$ bij 2 vrijheidsgraden.

Daarna is de nulhypothese getest dat er geen significante verbetering optreedt door het RSLN(2) te gebruiken ten opzichte van het GARCH(1,1). Bij een 5% significantieniveau wordt de nulhypothese verworpen, aangezien geldt dat $p = 0.05$ bij 2 vrijheidsgraden.

Conclusie:

Er blijkt dus dat het RSLN(2) model bij deze dataset een significante verbetering is ten opzichte van het GARCH(1,1) model. Op basis van dit model kunnen nu scenario's worden gegenereerd voor het aandelenrendement voor de komende 60 maanden.

Dat het model de huidige dataset fit wil niet zeggen dat hij ook de toekomst goed beschrijft. Er dient verder onderzoek te worden gedaan naar de prestaties van het RSLN model ten opzichte van het GARCH model. Wanneer een eerste vergelijking wordt gedaan, blijkt dat het GARCH model betere voorspellingen geeft. Dit dient echter uitgebreider te worden onderzocht.

7. Opzet model om de dekkingsgraad te bepalen bij gegeneerde economische scenario's

Nu eenmaal voor de inflatie en de kort- en langlopende obligatie index een model is geschat en de voorspellingen zijn gedaan voor de komende vijf jaar, en als eenmaal voorspellingen zijn gedaan voor het rendement op de AEX index, is het mogelijk om een model op te stellen om te bepalen wat de kans is dat de dekkingsgraad van een Separate Account onder een bepaald niveau komt. Onder alle gedane veronderstellingen en gebruikte technieken, blijkt uit dit onderzoek dat daarvoor het best het VAR(1) model kan worden gebruikt voor de inflatie en kort- en langlopende obligatie-index. Voor het rendement op de AEX index kan het RSLN(2) model worden gebruikt, er dient echter nog verder onderzoek te worden gedaan naar de voorspelling prestaties van het RSLN(2) model ten opzichte van het GARCH(1,1) model. Nu is alleen de fit op de huidige data onderzocht. In dit hoofdstuk wordt een eerste opzet gegeven van de manier waarop een model zou kunnen worden opgesteld om de kans te berekenen dat de dekkingsgraad van een SA onder een bepaald niveau komt. Dit is een simpele eerste opzet, en dient te worden uitgebreid.

Wanneer een model zou worden opgesteld om te bepalen wat de kans is dat dekkingsgraad van een SA onder een bepaald niveau komt, zijn er een aantal keuzes die moeten worden gemaakt:

- Welke *beleggingsstrategie* wordt doorgerekend?
 - In welke activa klassen wordt geïnvesteerd? En wat is het percentage in iedere activa klasse? Aangezien in dit onderzoek alleen rekening is gehouden met obligaties en aandelen, zou een keuze 80% obligaties en 20% aandelen zijn.
 - Wordt er jaarlijks gerebalanced? Wanneer er gekozen wordt voor een portefeuille samenstelling met 80% obligaties en 20% aandelen, is aan het einde van een bepaalde periode op de aandelen een bepaald rendement behaald en op de obligaties een bepaald rendement behaald. Het is nu zeer waarschijnlijk dat de verhouding in de portefeuille niet langer 80% obligaties en 20% aandelen is. Wanneer er op de aandelen een hoger rendement is behaald dan op de obligaties, zal dit percentage bijvoorbeeld 23% aandelen en 77% obligaties bedragen. Dit is vaak niet gewenst, omdat hierdoor de variantie van de totale portefeuille onbewust toeneemt. Bij rebalancing wordt dit percentage iedere periode aangepast zodat het oorspronkelijke percentage gehandhaafd wordt. In dit geval wordt het percentage aangepast tot 80% obligaties en 20% aandelen. Dit gebeurt, bij dit voorbeeld, door de verkoop van aandelen en de aankoop van obligaties.
 - Wordt er ingegrepen in de portefeuille mix wanneer de dekkingsgraad te laag wordt?
- Wordt *premiebetaling* meegenomen in het model?
- Zijn er *bijstortingsverplichtingen*? Dit wil zeggen dat er extra moet worden bijgestort wanneer de dekkingsgraad onder een bepaald niveau komt.
- Wat is het *indexatiebeleid*? Wordt er bijvoorbeeld niet geïndexeerd wanneer de dekkingsgraad onder een bepaald niveau komt?
- Hoe worden de *verplichtingen* gemodelleerd? In het simpele geval kan worden aangenomen dat de verplichtingen ieder jaar met 4% stijgen. De verplichtingen kunnen echter ook met de marktwaarde stijgen.

Als deze keuzes zijn gemaakt, is het gemakkelijk om het verloop van de dekkingsgraad te berekenen. Er worden dan 1000 voorspellingen gedaan voor de komende vijf jaar voor

het aandelenrendement en de obligatie-index. Aan de hand van de gemaakte keuzes (beleggingsstrategie, premiebetaling, et cetera) worden er dan 1000 verlopen van de dekkingsgraad berekend. Zo kan een betrouwbaarheidsinterval voor de dekkingsgraad worden bepaald. Het is duidelijk dat er verder onderzoek moet worden gedaan naar de betrouwbaarheid van zo'n model. Het hebben van een model dat de kans op een lage dekkingsgraad kan berekenen zou veel toegevoegde waarde hebben voor ZwitserLeven.

8. Verder onderzoek

In dit hoofdstuk zullen enkele aanbevelingen worden gedaan voor verder onderzoek.

- Hoe kan een gedetailleerder model worden opgesteld om de kans te bepalen dat de dekkinggraad onder een bepaald niveau komt? Een eerste opzet voor dit model is gegeven in hoofdstuk 7. Dit is echter een simpele opzet, en dient te worden uitgebreid.
- Nu de kort- en langlopende obligatie-index voorspeld zijn, is het ook mogelijk om de yield curve te modelleren. Dit kan met behulp van de Nelson Siegel vergelijking, zoals beschreven in hoofdstuk 6.
- Er dient verder onderzoek te worden gedaan naar de vraag of de Yule-Walker schattingstechniek betere voorspellingen geeft dan OLS.
- Wat zijn de prestaties bij het voorspellen van het rendement op de AEX index voor het GARCH model en het RSLN model?
- Er dient verder onderzoek te worden gedaan naar de voorspelling prestaties van het GARCH model ten opzichte van het RSLN model voor het rendement op de AEX index. Ook blijkt het VAR(1) model met daarin de AEX index het aandelenrendement niet beter te voorspellen dan het GARCH of RSLN model voor het aandelenrendement, maar hier moet uitgebreider onderzoek naar worden gedaan.
- Er moet verder onderzoek worden gedaan naar het vinden van een bevredigend model waarin wel de correlatie tussen obligaties en aandelen verwerkt zit. Dit zou wellicht een combinatie van de twee losstaande modellen, zoals nu geschat, kunnen zijn.

Conclusie

Welk model om economische scenario's te genereren kan als optimaal worden beschouwd, en hoe kunnen deze economische scenario's vervolgens gebruikt worden?

In dit verslag zijn twee verschillende typen ALM modellen beschreven. Dit zijn stochastische lineaire programmering (SLP) modellen en simulatiemodellen. Bij een SLP probleem wordt een doelstellingsfunctie geoptimaliseerd om zo een optimaal beleid voor een pensioenfonds te bepalen. Een nadeel is dat een SLP model snel erg complex wordt. Bij een simulatiemodel zijn er twee losstaande zaken: de economische scenario generator en het optimalisatiemodel. Dit is dus verschillend van SLP modellen, waar zowel de scenario's en het optimalisatieprobleem in één model worden gestopt. Simulatiemodellen zijn gericht op het evalueren van een beleid terwijl SLP modellen gericht zijn op het optimaliseren.

Dit onderzoek heeft zich gericht op het modelleren van macro-economische variabelen: de inflatie, de kort- en langlopende obligatie index en het rendement op de AEX index. Om de inflatie en de kort- en langlopende obligatie index te modelleren is er gebruik gemaakt van de verschillende modellen: VAR modellen, ARIMA modellen en Holt Winters (Exponential Smoothing). Wanneer de performance van deze modellen wordt getest op een deel van de dataset, blijkt dat het VAR(1) model het beste model is bij de gekozen dataset en bij de gebruikte technieken. Er blijkt dat een VAR(1) model met daarin ook de AEX index slechtere voorspellingen geeft voor de inflatie, kort- en langlopende obligatie index, dan een VAR(1) model met daarin alleen de inflatie, kort- en langlopende obligatie index. Het optimale model voor het aandelenrendement blijkt een Regime Switching Lognormal model te zijn: RSLN(2). Er dient echter verder onderzoek te worden gedaan naar de voorspelling prestaties van het RSLN model ten opzichte van het GARCH model. Ook blijkt het VAR(1) model met daarin de AEX index het aandelenrendement niet beter te voorspellen dan het GARCH of RSLN model. Hier moet echter uitgebreider onderzoek naar worden gedaan.

Met de gegenereerde scenario's is het nu mogelijk om voor grootzakelijke contracten met een gesepareerd beleggingsdepot de kans te bepalen dat de dekkingsgraad onder een bepaald niveau komt. Een simpele opzet voor zo'n model is gegeven.

Literatuurlijst

- [1] Kouwenberg, R. *Dynamic Asset Liability Management*. Tinbergen Institute Research Series, 2000
- [2] Drijver, S.J., Klein Haneveld, W.K. en van der Vlerk, M.H. *ALM model for pension funds: numerical results for a prototype model*. Research School SOM, 2002
- [3] Boender, G., Dert, C., Heemskerk, F. en Hoek, H. *A scenario approach of ALM*. Handbook of Asset and Liability Management, Volume 2, 2007
- [4] Steehouwer, H. *Macroeconomic Scenarios and Reality, A frequency domain approach for analyzing Historical Time Series and Generating Scenarios for the Future*. Optima, 2005
- [5] Kleynen, R. *Asset Liability Management binnen pensioenfondsen*. Katholieke Universiteit Brabant, 1996
- [6] *Inwerkmap Zwitserleven*, intern document Zwitserleven
- [7] <http://www.wikipedia.org>
- [8] http://www.riskglossary.com/link/asset_liability_management.htm
- [9] Drijver, S. *Asset Liability Management for Pension Funds using Multistage Mixed-integer Stochastic Programming*. Proefschrift Rijksuniversiteit Groningen, 2005
- [10] Haneveld, W.K., Streutker, M.H. en van der Vlerk, M.H. *An ALM model for pension funds using integrated chance constraints*. Research School SOM, 2005
- [11] *Economic Scenario Generators*, presentatie Triple A
- [12] Bodie, A., Kane, A. en Marcus, A. *Investments 6th*. Mc Graw Hill, 2005
- [13] Onna, M. van en Koning, A. *De kleine Prince 2: Gids voor projectmanagement*. Academic Service, 2006
- [14] *Jaarverslag 2006*, Zwitserleven
- [15] <http://www.zbc.nu>
- [16] Steehouwer, H. *Nelson and Siegel yieldcurve modelling and bank asset and liability management*. Ortec, 1998
- [17] *Wegwijzer proceseigenaarschap*, intern document Zwitserleven
- [18] Gunst, M.C.M de. *Statistical Models*. Dictaat Vrije Universiteit Amsterdam, Najaar 2004
- [19] Cuthbertson, K. en Nitzsche, D. *Quantative Fiancial Economics*. Wiley, 2004
- [20] Pfaff, B. *Using the vars package*. www.r-project.org, 2007
- [21] Lay, D.C. *Linear Algebra and its applications*. Addison-Wesley, 2003
- [22] *PVK Informatie, informatie over de wet toezicht verzekeringsbedrijf 1993*. versie januari 2002, Pensioen en Verzekeringskamer, www.dnb.nl
- [23] Veen, P., Alblas, G. en Geersing, J. *Mensen in organisaties: Een inleiding in de organisatiepsychologie*. Bohn Stafleu Van Loghum, 1991
- [24] Boer, H. *Inleiding Organisatiekunde*. Dictaat Universiteit Twente, 1988
- [25] Golub, G.H. en van Loan, C.F. *Matrix Computations 3th*. The Johns Hopkins University Press, 1996
- [26] Steehouwer, H. *Modellen voor de Termijnstructuur ten behoeve van het Asset and Liability Management van een Bank*. Master scriptie Econometrie, 1996
- [27] Hardy, M.R. *A Regime-Switching Model of Long-Term Stock Returns*. North American Actuarial Journal Volume 5 Part 2, pp. 41-53, 2001
- [28] Oosterhoff, J. en Vaart, A.W. van der. *Algemene statistiek*. Dictaat Vrije Universiteit Amsterdam
- [29] Vaart, A.W. van der. *Collegedictaat Time Series*. Vrije Universiteit, 1995-2002
- [30] *Forecasting Course outline*. Duke University, <http://www.duke.edu/~rnau/411home.htm>
- [31] Bergh, K. van den. *Predicting Arrivals in Call Centers*. BWI werkstuk, Vrije Universiteit. <http://www.few.vu.nl/stagebureau/werkstuk/werkstukken/werkstuk-bergh.pdf>
- [32] Fritzer, F., Moser, G. en Scharler, J. *Forecasting Austrian HICP and its components using VAR and ARIMA models*. Oesterreichische Nationalbank.

- [33] http://www.oenb.at/de/img/wp73_tcm14-6163.pdf
Zucchini, W. en Nenadic, O. *Time Series Analysis with R – Part I*.
- [34] http://www.staetk.wiso.uni-goettingen.de/veranstaltungen/zeitreihen/sommer03/ts_r_intro.pdf
Francq, C. en Raissi, H. *Multivariate Portmanteau test for autoregressive models with uncorrelated but nonindependent errors*. Journal of time series analysis, Vol. 28, No. 3, 2006
- [35] Society of Actuaries, Geïmplementeerd RSLN(2) model,
<http://www.soa.org/research/files/xls/RSLNL%20Excel%202003.xls>
- [36] Heij, C., Ran, A.C.M. en Schagen, F. van. *Mathematical Systems Theory: an introduction to Linear Systems, Identification and Control*