

Simulatie als alternatief voor Black-Scholes in optiewaardering.

Praktische toepassingen van optieprijs modellen met alternatieve verdelingen

Salet, Ernst (290131)

Bachelor Scriptie Financiële Econometrie 2008
Bachelor Econometrie en Besliskunde
Erasmus Universiteit Rotterdam

Begeleider
Dr. L.F. Hoogerheide
*Econometrisch Instituut,
Erasmus Universiteit Rotterdam*

20 juni 2008

Abstract

In dit onderzoek wordt het onderzoek over optiewaardering uit het werkcollege financiële econometrie (geschreven in mei 2008 door M. Kok, E. Salet, G. Spijkers en M. Staarink) doorgezet waarbij de nadruk ligt op praktische toepassingen van het simulatiemodel. Waar dit model op basis van de scheve t-verdeling in het werkcollege de hemel wordt geprezen, volgt in dit onderzoek een kritische blik. Uit de resultaten van een “slim” optiemodel waarin de geïmpliceerde parameters over de uitoefenprijzen mogen variëren, blijkt dat het model op basis van de scheve t-verdeling zijn beperkingen heeft. De geïmpliceerde parameters blijken niet constant over uitoefenprijzen, wat voor het “juiste” model echter wel (*ongeveer*, vanwege inefficiënties op de optiemarkt) moet gelden. Hiernaast wordt een model gepresenteerd waarin delta-hedging kan worden toegepast. Zo kan naast de *statistische* vergelijking van het optiemodel werkcollege ook een *economische* interpretatie aan het resultaat worden gegeven. De meest veelbelovende praktische toepassing is een real-time versie van het simulatiemodel, waarbij alle relevante simulaties “van tevoren” in een algemene database worden opgeslagen. In dit onderzoek wordt de basis gelegd voor een optimaal gebruik van deze methode, die in staat is om de rekentijd van het optie-model op basis van simulatie sterk te reduceren.

Inhoudsopgave

1 Probleemstelling	3
1.1 Motivatie	4
1.2 Data	5
1.3 Methoden	5
2 “Slim schatten“	7
2.1 Stappenplan “Slim Schatten” o.b.v. gewichten	8
2.2 Resultaten Slim Schatten	10
2.3 Discussie resultaten Slim Schatten	12
3. Performance maatstaven	14
3.1 Diebold-Mariano (DM) toets	14
3.2 Delta-Hedging	14
3.3 Voorbeeld: Verschil Delta-Hedging en DM-toets	16
4 “Real-time” optiemodel	18
4.1 Methoden: “Real Time” optiemodel	18
4.2 Grid-Analyse	20
5 Conclusie & Discussie	23
6.Referenties	24
Appendix 1: Uitgebreide toelichting data	25
Appendix 2 Methoden: Monte-Carlo simulatie	26
Appendix 3 Gewichten analyse	28
Appendix 4 Slim schatten: Omringende Opties	30
Appendix 5 Slim schatten: Fijnere schatting γ	32
Appendix 6 Slim schatten voor andere dagen	33
Appendix 7 Pseudocode Delta-Hedging	34
Appendix 7 Pseudocode Delta-Hedging	35
Appendix 8 Grid-analyse uitgebreide resultaten	37
Appendix 9: Exponentieel df-grid	39
Appendix 10 Delta-Hedging resultaten: Verloop van Delta	41

1 Probleemstelling

In dit onderzoek staat het waarden van opties centraal. Deze scriptie vormt een vervolg op het werkcollege voor financiële econometrie “Simulatie als alternatief voor Black-Scholes”¹, waarin de onderliggende aannames van het optieprijsmodel van Black-Scholes (1973) worden bekritiseerd. Empirisch onderzoek in dit werkcollege wees uit dat de door de markt veronderstelde *risico-neutrale* verdeling van de rendementen van het onderliggende *asset* in de praktijk niet normaal verdeeld is, zoals het Black-Scholes model veronderstelt. De verdeling van de onderliggende rendementen blijkt in de praktijk negatief scheef, en heeft dikkere staarten ten opzichte van de normale verdeling. Op basis van de gebruikte data en de onderzochte verdelingen werd in het werkcollege geconcludeerd dat de modellen met een ‘split T-verdeling’ en ‘scheve T-verdeling’ (modellen die corrigeren voor zowel scheefheid als kurtosis) de beste optieprijzen afleveren. De optieprijzen die deze modellen afgeven komen statistisch gezien significant beter overeen met de beschouwde waargenomen optieprijzen dan bijvoorbeeld het Black-Scholes model en de onderzochte (stochastische) *jump*-modellen. De resultaten versterken het vermoeden van angst onder beleggers voor een crash, doordat de optimale verdelingen in vergelijking met de normale verdeling overwegend ‘negatief scheef’ zijn met dikkere staarten. Kanttekening bij deze resultaten is dat de impliciete parameters van de gebruikte verdelingen over alle strike prices van een specifieke uitoefendatum als constant worden verondersteld; deze aanname is verder niet getoetst. Daarnaast zijn de resultaten door de hoge rekentijd praktisch gezien helaas nauwelijks bruikbaar.

Het moet worden opgemerkt dat in het werkcollege een sterk vereenvoudigd model werd gebruikt voor het schatten van de geïmpliceerde volatiliteit. Per dag wordt de geïmpliceerde volatiliteit als constant verondersteld, aangezien dezelfde schatting van de geïmpliceerde volatiliteit wordt gebruikt voor alle bijbehorende uitoefenprijzen. De geïmpliceerde volatiliteit is in werkelijkheid echter niet constant, maar een convexe functie van de uitoefenprijzen (beter bekend als een *smile* of *skew*). Dit veroorzaakt een ruwe schatting van met name het Black-Scholes model, die de geïmpliceerde volatiliteit als enige impliciete parameter heeft. In dit onderzoek wordt een “slimmere” methode gepresenteerd, waarin de geïmpliceerde parameters afhangen van “omringende” opties qua uitoefenprijs. Zo worden per optie de bijbehorende parameters geschat, zodat het verloop van de parameters over verschillende uitoefenprijzen kan worden onderzocht.

Hutchinson, Lo & Poggio (1994) presenteren in hun artikel een praktische toepassing voor het vergelijken van niet-parametrische modellen en het Black-Scholes model door het gebruik van delta-hedging. Op vergelijkbare manier zal in dit onderzoek deze economisch relevante methode gebruikt worden om de verschillende verdelingen onderling te vergelijken. Het meest ambitieuze doel van dit onderzoek is echter een andere praktische toepassing van de resultaten: een *real-time* optie waarderings methode. Hier worden de simulatie resultaten van verschillende (slim gekozen) grids van parameter waarden opgeslagen en als database gebruikt. Zo zijn de simulaties “van tevoren” al uitgevoerd, zodat veel snellere *real-time* berekeningen voor optieprijzen mogelijk zijn.

¹ Geschreven in mei 2008 door M. Kok, E. Salet, G. Spijkers en M. Staarink,.

Kortom, in vergelijking met het werkcollege bevat dit verslag drie concrete toevoegingen:

- 1) Een fijnere schattingsmethode voor de geïmpliceerde volatiliteit en de andere geïmpliceerde parameters bij alternatieve verdelingen
- 2) Een numerieke methode voor delta-hedging van de opties
- 3) Een “real-time” optie waarderings methode op basis van “van tevoren” uitgevoerde simulaties.

Ten opzichte van het werkcollege moet dit onderzoek betere schattingen van de optiepreizen opleveren door een fijnere schatting van de geïmpliceerde parameters. Daarnaast biedt dit onderzoek praktische toepassingen, in de vorm van delta-hedging en een real-time optie prijs model.

1.1 Motivatie

In de data van het werkcollege, betreffende S&P500 futures opties van het jaar 2000 (zie ook Appendix 1), was vaak een typische *volatility skew* of een *volatility smile* waar te nemen. Dit houdt in dat de geïmpliceerde volatiliteit aflopend is (*skew*) of eerst aflopend is en daarna stijgend (*smile*) over toenemende uitoefenprijzen. Een mogelijke verklaring hiervoor is het verschijnsel ‘Crashphobia’, dat werd opgemerkt door Jackwerth en Rubenstein (1996). Beleggers houden volgens deze theorie rekening met de mogelijkheid van zeer negatieve ‘crash’ rendementen, waardoor de volatiliteit voor heel lage uitoefenprijzen vaak relatief hoog is. Toch werd in het werkcollege de geïmpliceerde volatiliteit geschat uit alle optiedata van 1 dag, waarbij aan elke uitoefenprijs hetzelfde gewicht wordt toegekend. Op deze manier wordt met name het Black-Scholes model (met volatiliteit als enige impliciete parameter) vrij ruw geschat. Dit is waarschijnlijk de belangrijkste reden voor de relatief grote afwijking van de geschatte Black-Scholes waarden met de werkelijke optiepreizen in het werkcollege. Vandaar dat in dit onderzoek een methode wordt gepresenteerd waarbij de geïmpliceerde volatiliteit “slimmer” wordt geschat, waarbij rekening wordt gehouden met het verloop van de geïmpliceerde volatiliteit. Ook aan de geïmpliceerde parameters van de andere onderzochte verdelingen wordt toegestaan te variëren over de uitoefenprijzen. Hopelijk heeft dit een betere benadering van werkelijke optiepreizen tot gevolg voor zowel het Black-Scholes model als de andere onderzochte verdelingen.

Het werkcollege heeft zich vooral gefocust op het prijzen van opties en het vergelijken van de “prestatie” van verschillende verdelingen d.m.v. de Diebold Mariano (DM) toets. Een nuttige toevoeging aan dit verslag is om naast het prijzen van de opties ook het *hedgen* van opties te beschouwen. Het (delta-) hedgen van opties biedt bovendien nieuwe inzichten in de verschillende verdelingen als economisch relevante toepassing. *Delta-hedging* kan naast de statistische vergelijking met de DM-toets, gebruikt worden als praktische evaluatietool bij de vergelijking van de prestaties van de verdelingen. Zoals de DM-toets aangeeft of het verschil tussen de methoden *statistisch* significant is, kan delta-hedging aangeven of het verschil *economisch* significant is.

In het werkcollege wordt veel inzicht verkregen in de door de markt veronderstelde risiconeutrale verdelingen binnen de optiemarkt, maar door lange rekentijd zijn de gebruikte modellen nauwelijks toepasbaar in de praktijk. In dit onderzoek wordt gezocht naar een oplossing van dit probleem, door een database op te stellen waarin alle relevante simulaties “van tevoren” al zijn uitgevoerd. Het beoogde eindresultaat is om op basis van interpolatie tussen de resultaten van de vooraf verkregen database een praktisch toepasbaar *real-time* optiemodel te maken, die het “closed form” optieprijs model van Black-Scholes significant verbetert, zowel statistisch gezien (via de DM-toets) als praktisch gezien (via delta-hedging).

1.2 Data

De dataset uit het werkcollege zal in dit onderzoek worden hergebruikt. Het betreft een dataset van call opties op futures van de S&P 500 index in het jaar 2000. Een uitgebreide toelichting van de dataset wordt gegeven in Appendix 1.

In dit verslag zullen alle modellen in eerste instantie worden getest aan de hand van de data van een enkele dag, zodat een goede onderlinge vergelijking mogelijk is. De geselecteerde dag is 8 december 2000, met looptijd 5 handelsdagen ($T=5$) en $F_0=\$1337,70$. Deze dag is relevant, aangezien er absoluut gezien veel is gehandeld (een relatief hoog totaal volume van 4028), waarvan 29 opties een volume hebben groter dan 0. Vervolgens zullen de resultaten ook voor andere dagen getest worden (uit verschillende periodes in het jaar en met verschillende looptijden), om een indicatie te krijgen van de robuustheid van de uitkomsten.

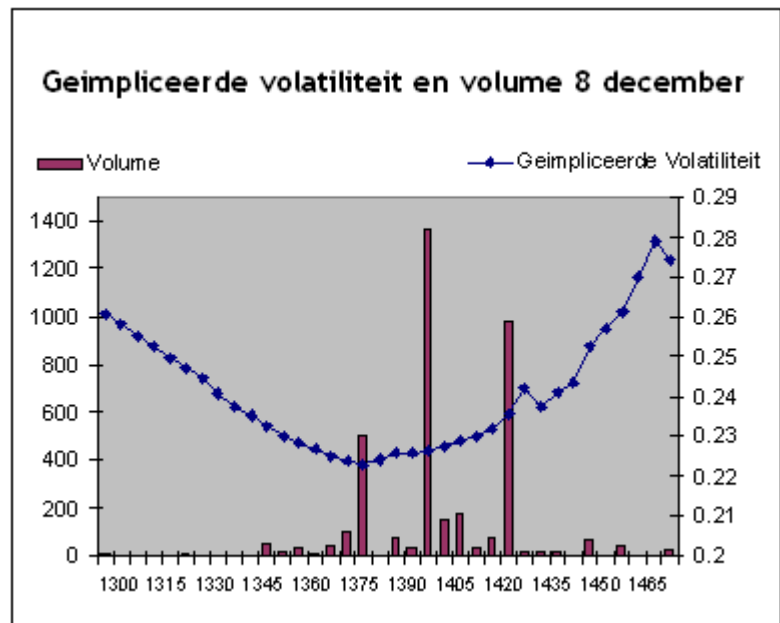
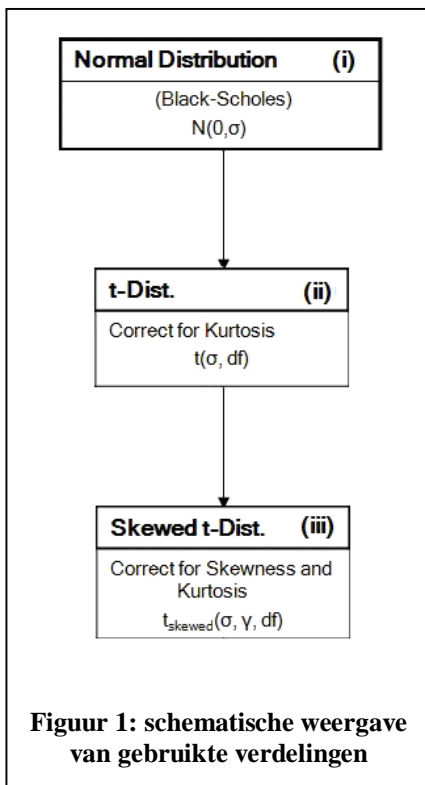
1.3 Methoden

In dit onderzoek zullen de 3 verdelingen uit figuur 1 onderzocht worden, waarbij de focus ligt op de Black-Scholes verdeling (verdeling i) en de Scheve t-verdeling (verdeling iii). De 3 verdelingen uit figuur 1 liggen op hetzelfde “pad” (d.w.z. dat de lage modellen generalisaties zijn van hoge modellen – zie het werkcollege verslag voor meer details) en zijn in dit onderzoek geselecteerd doordat verdeling (iii) in het werkcollege in veel gevallen het best presteert.

De fijnere schattingsmethode voor de geïmpliceerde volatiliteit komt tot stand door het slim kiezen van *gewichten* voor opties met verschillende strike prices van één uitoefendatum. Waar de geïmpliceerde volatiliteit over de verschillende strike prices van een dag in het werkcollege constant werd verondersteld, kan deze parameter dit in dit fijnere schattingsmodel variëren per strike price. Dat dit een grote verbetering kan zijn ten opzichte van het werkcollege, wordt geïllustreerd aan de hand van figuur 2. Dit specifieke voorbeeld toont aan dat de geïmpliceerde volatiliteit op 8 december de eigenschappen heeft van een *volatility smile*. Door aan omliggende opties qua uitoefenprijs een hoger gewicht toe te kennen dan aan optie op “afstand (qua uitoefenprijs) kan impliciet rekening worden gehouden met deze *volatility smile*. Hiernaast kan er ook voor gekozen worden om opties met een hoger handelsvolume zwaarder te wegen, omdat deze mogelijk efficiënter geprijsd zijn. In hoofdstuk 2 worden 2 toepassingen gepresenteerd van dit “slimme” schattingsmodel. In de eerste methode wordt alleen rekening gehouden met de 2 dichtst bijzijnde *omliggende opties* qua uitoefenprijs, in de tweede methode wordt zowel rekening gehouden met het volume als de ligging qua uitoefenprijs met behulp van *gewichten*.

In hoofdstuk 3 worden twee performance maatstaven toegelicht voor het evalueren van de optiewaarderingen. De gebruikte statistische maatstaf is de toetsgrootheid van de Diebold-Mariano toets. In deze toets wordt van de 2 onderzochte verdelingen het verschil in Mean Squared Error (MSE)² onderzocht door middel van een regressie van het verschil in kwadratische residuen op een constante. De t-waarde van de constante die uit de regressie volgt, fungeert hierbij als toetsgrootheid. Hiernaast wordt ook een economisch relevante maatstaf gepresenteerd waarbij gebruik wordt gemaakt van *delta-hedging*. De praktische maatstaf die in dit onderzoek wordt gehanteerd is de *tracking error* van een *delta-hedge* portfolio.

² MSE staat hiervoor het gemiddelde kwadratische verschil van de werkelijke optieprijs en de benadering van de geselecteerde verdeling.



Figuur 2 Geïmpliceerde volatiteit en volume uitgezet tegen Strike Price (K)

De methode voor het schatten van een dergelijke *economisch* relevante maatstaf is gebaseerd op het artikel van Hutchinson, Lo & Poggio (1994), waarin de prestatie van enkele niet-parametrische modellen wordt geëvalueerd op basis van delta-hedging. Hutchinson et al. (1994) presenteren een model waarin delta (de afgeleide van de optieprijs naar de prijs van het onderliggende asset) benaderd wordt, zodat een portefeuille kan worden samengesteld waarin een call-optie gehedged wordt. Deze portefeuille uit het verkopen van 1 call-optie (*short*) en het aankopen van delta maal de onderliggende future (*long*) en heeft een initiële waarde van 0. Door het toepassen van een “dynamic hedging” strategie op deze portfolio, zou de totale waarde van de portfolio aan het eind van de periode gelijk aan 0 moeten zijn. In de praktijk wijkt deze waarde echter vaak af van 0, mede door inefficiënties op de optiemarkt. De verwachte afwijking wordt de “tracking error” genoemd, en wordt in dit onderzoek gebruikt als economisch relevante maatstaf voor de prestatie van de gebruikte verdeling. Het idee hierachter is dat een beter model een betere benadering afgeeft van de werkelijke afgeleide van de optieprijs naar de prijs van het onderliggende asset³, wat gemiddeld genomen tot een lagere tracking error zal leiden.

Over het “real time” optiemodel (de methode gepresenteerd in hoofdstuk 4) is nog erg weinig literatuur verschenen. In dit hoofdstuk worden daarom eerst de gebruikte methoden uitgebreid toegelicht, waarna in H4.2 een zgn. *grid-analyse* wordt gedaan om het model te verscherpen.

Alle simulatie-experimenten die in dit onderzoek worden gedaan volgen de procedure van Monte-Carlo simulatie, analoog aan de methode gepresenteerd in het werkcollege. Dit model wordt uitgebreid toegelicht in Appendix 2 en is onder andere gebaseerd op het artikel van Boyle (1977), waarin optieprijsen worden benaderd met behulp van Monte-Carlo simulatie.

³ De benadering van de werkelijke afgeleide wordt weergegeven door delta, hierover meer in H 3.

2 “Slim schatten”

In het werkcollege worden de parameters σ , γ en df van de scheve t-verdeling voor alle opties met de zelfde afloopdatum over alle uitoefenprijzen gezamenlijk geschat. Dit resulteert in een verdeling met 3 constante parameters over de verschillende uitoefenprijzen. In de praktijk is het echter mogelijk dat de geïmpliceerde parameters verre van constant zijn, doordat niet aan de aannames van het model voldaan is. Herinner als voorbeeld de *volatility smile*, besproken in hoofdstuk 1. Het moet echter worden opgemerkt dat de geïmpliceerde volatiliteiten uit deze *smile* worden afgeleid uit de Black-Scholes formule. Dit model wordt in het werkcollege juist bekritiseerd doordat het normaal verdeelde rendementen aanneemt en uitsluitend σ gebruikt als impliciete parameter. De vraag rijst of deze *volatility smile* stand houdt als er parameters worden toegevoegd (zoals df en γ , zoals in figuur 1), of dat de *volatility smile* wordt “rechtgetrokken”. Als de aanname van de scheve t-verdeling correct is, dan moeten de geïmpliceerde parameters constant zijn over verschillende strike prices (of *ongeveer* constant door afrondingen en mogelijke inefficiënties op de optiemarkt).

Doel van dit “slimme” model is om een schatting van de *geïmpliceerde parameters* te verkrijgen, waarbij de parameters per optie apart geschat worden. Op deze manier wordt inzicht verkregen in het verloop van de verschillende parameters bij variërende uitoefenprijzen. Doelfunctie van dit model blijft het zo nauwkeurig mogelijk benaderen van de werkelijke optieprijs. Aangezien voor het waarderen van één optie een verdeling wordt gebruikt met 3 impliciete parameters (bij de scheve t-verdeling), worden ook de parameters van 2 omliggende opties (qua strike price) meegenomen in de schatting. Een algemene regel in de econometrie is namelijk dat het aantal waarnemingen minstens even groot moet zijn als het aantal parameters. Wordt deze regel niet gehonoreerd, dan zijn er vaak oneindig veel combinaties van parameter waarden mogelijk. De afzonderlijke parameters zijn in dat geval *niet geïdentificeerd*. Bij het schatten van σ van een optie met strike price (K) 1340 wordt daarom niet alleen gekeken naar de eigen geïmpliceerde volatiliteit, maar worden ook de geïmpliceerde volatiliteiten van de omliggende opties qua strike price meegenomen (in dit geval opties met K=1335 en K=1345). Op deze manier wordt de schatting van σ , γ en df uiteindelijk gebaseerd op 3 waarnemingen in plaats van 1. Aangezien elke schatting 2 omliggende waarnemingen nodig heeft, zijn er $n-2$ ⁴ schattingen mogelijk van de parameters; de laagste en hoogste uitoefenprijs vallen in die zin weg. Naast deze aanpak van *omliggende opties*, wordt in dit onderzoek ook een meer geavanceerde aanpak besproken op basis van *gewichten*. Hier wordt rekening gehouden met 2 specifieke kenmerken van de optie: het volume en de ligging qua strike price. Voor het schatten van 1 optie worden aan alle $(n-1)$ overige opties van die dag gewichten toegekend op basis van deze 2 kenmerken. In appendix 3 wordt een voorbeeld van een formule gepresenteerd die dergelijke gewichten aan de opties toekent. Deze gewichten worden vermenigvuldigd met bijbehorende impliciete parameters, zodat de schatting van σ , γ en df wordt gebaseerd op $n-1$ waarnemingen.

De praktische relevantie van deze modellen lijkt in eerste instantie niet erg groot, aangezien de rekentijd ten opzichte van het werkcollege alleen maar stijgt. Immers, waar in het werkcollege voor alle uitoefenprijzen slechts 1 gezamenlijke schatting nodig is, wordt hier voor elke uitoefenprijs afzonderlijk een schatting gemaakt. Wel levert dit zeer nauwkeurige schattingen op van de werkelijke optieprijs, doordat een schatting wordt gemaakt van de *impliciete parameters per afzonderlijke optie*. Hoofddoel van dit model is dan ook het bestuderen van het verloop van deze impliciete parameters. Zijn er patronen te ontdekken in

⁴ Waarbij n staat voor het aantal verschillende uitoefenprijzen van de optie waarin die dag gehandeld is.

het verloop van de optimale parameters, of is er sprake van een constante waarde plus *ruis*? Wat gebeurt er met de *volatility smile* als de parameters γ en df worden toegevoegd? Blijft de parameter γ (zoals in het werkcollege) structureel onder de waarde $\gamma = 1$ behorende bij de normale verdeling?

In hoofdstuk 2.1 wordt het model op basis van *gewichten* toegelicht aan de hand van een uitgebreid stappenplan. De bijbehorende resultaten worden in hoofdstuk 2.2 gepresenteerd. Ter vergelijking staan de resultaten van het model op basis van *omliggende opties* in Appendix 3. In hoofdstuk 2.3 worden alle resultaten besproken en wordt dieper ingegaan op de prestatie van het model en het verloop van de verschillende parameters.

2.1 Stappenplan “Slim Schatten” o.b.v. gewichten

De methode voor het fijnere schattingsmodel wordt beschreven aan de hand van een stappenplan en een schematisch overzicht, gegeven in figuur 3. In de figuur worden de parameters die gezamenlijk worden geschat (zoals in stap 3 en 4b), weergegeven binnen het grijs gearceerde gebied. Op basis van deze figuur kan worden gezien welke parameters worden gebruikt om tot de optimale schatting van de impliciete parameters te komen. Beginpunt van het stappenplan is het bepalen van de geïmpliceerde volatiliteiten uit de werkelijke optieprijsen. Per stap wordt het model uitgebreid door iteratief een extra parameter te schatten.

Stap 1: Geïmpliceerde volatiliteiten berekenen

Op basis van de optieprijsen kan de volatiliteit met behulp van de Newton-Raphson methode impliciet bepaald worden uit Black's formule. Deze formule wordt gegeven door (voor opties op futures geldt hierbij: $q=r$, de risicovrije interest):

$$c = S_0 e^{-qT} N(d_1) - K e^{-rT} N(d_2)$$

$$\text{waarbij } d_1 = \frac{\ln(S_0 / K) + (r - q + \sigma^2 / 2)T}{\sigma \sqrt{T}} \text{ en } d_2 = d_1 - \sigma \sqrt{T}$$

Stap 2: Optimale σ schatten (Black-Scholes)

σ wordt per optie afzonderlijk geschat door de geïmpliceerde volatiliteiten uit stap 1 te vermenigvuldigen met de $n-1$ gewichten op basis van volume en ligging qua strike price (zoals uitgewerkt in appendix 3)⁵. Deze schattingen kunnen worden verscherpt middels de optimalisatiefunctie *fminsearch* uit Matlab. Doelfunctie is hierbij het minimaliseren van de afwijking met de werkelijke optieprijs, gemeten door de kwadratische afwijking (SSR).

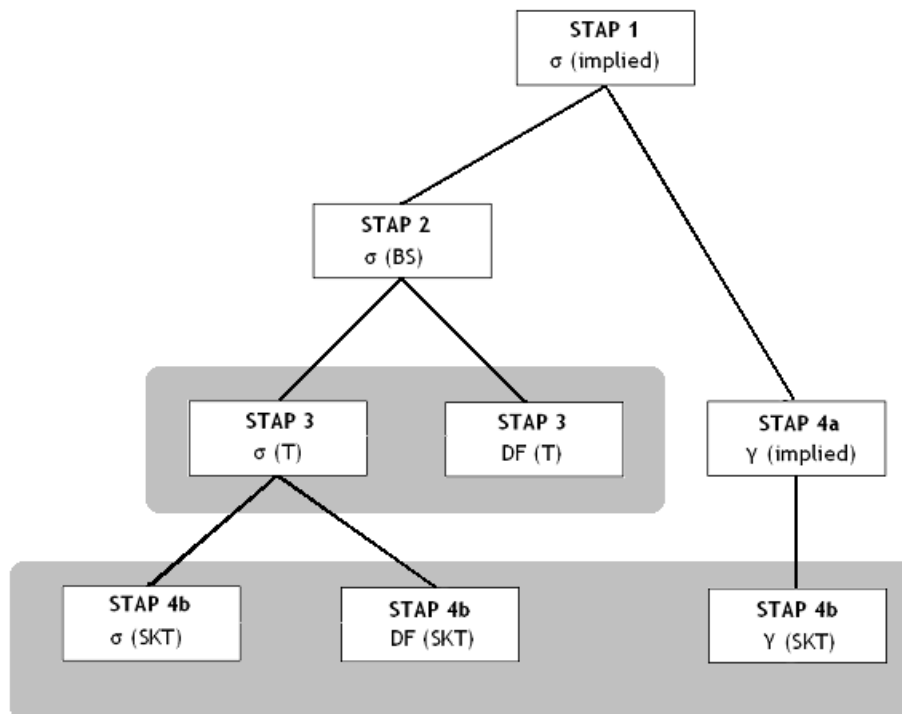
⁵ $n-1$ i.p.v. n omdat de prijs van een optie op basis van de andere $n-1$ opties met andere uitoefenprijzen willen schatten, omdat we de *out-of-sample performance* van het veronderstelde model willen bekijken.

Stap 3: Optimale σ en df schatten (T-verdeling)

Door het gebruik van een *grid* van *df*-waarden en de geschatte σ 's uit stap 2 kunnen de optimale impliciete parameters van de t-verdeling geschat worden. Gegeven de *df*-waarde uit het grid, kan de geïmpliceerde σ van de t-verdeling met behulp van *fminsearch* geoptimaliseerd worden. Dit levert een zeer goede benadering van de werkelijke optieprijs, maar de schatting kan nog worden verscherpt door toepassing van een gezamenlijke *fminsearch* methode, waarbij de geïmpliceerde σ 's en de optimale waarden uit het *df*-grid per optie gezamenlijk worden geoptimaliseerd.

Stap 4: Optimale σ , γ en df schatten (scheve T-verdeling)

- A) Voordat wordt overgegaan op het gezamenlijk schatten van de 3 parameters, moeten beginwaarden van γ worden bepaald. Dit kan door een grid te nemen over relevante γ -waarden (bijvoorbeeld tussen 0.5 en 1.5) en met hulp van de geïmpliceerde volatiliteiten uit stap 1 de SSR te minimaliseren.
- B) Op basis van de beginwaarden voor γ uit stap 4A wordt een schatting gemaakt van γ voor de scheve T-verdeling, analoog aan de procedure voor de σ 's in stap 2. Op basis van deze schatting van γ , de σ 's uit stap 3 en wederom een grid van *df*-waarden, kan met behulp van *fminsearch* de optieprijs benaderd worden. Ook deze benadering wordt nog verscherpt door toepassing van een gezamenlijke optimalisatie per optie van de geschatte σ , γ en *df*.

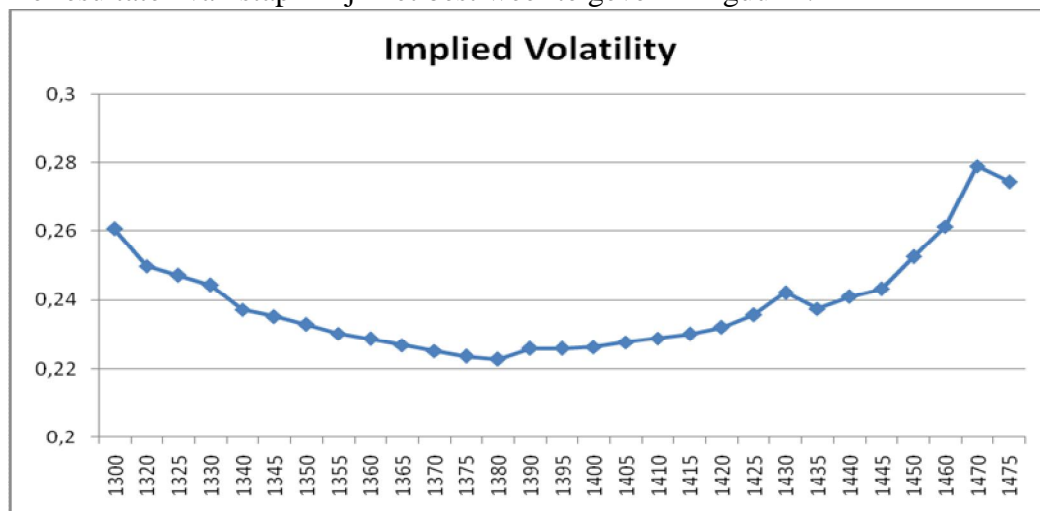


Figuur 3: Schematische weergave van het stappenplan.
In de 2 grijs gearceerde gebieden staan de parameters van de t-verdeling (stap 3) en de scheve t-verdeling (stap 4b), die ook gezamenlijk worden geoptimaliseerd.

2.2 Resultaten Slim Schatten

In eerste instantie worden de resultaten gepresenteerd van de opties van 8 december, looptijd 5 handelsdagen ($T=5$) en $F_0 = \$1337,70$. Van deze dag worden 29 opties (met een volume groter dan 0) onderzocht met uitoefenprijzen variërend van 1300 tot 1475.

De resultaten van stap 1 zijn het best weer te geven in figuur 4.



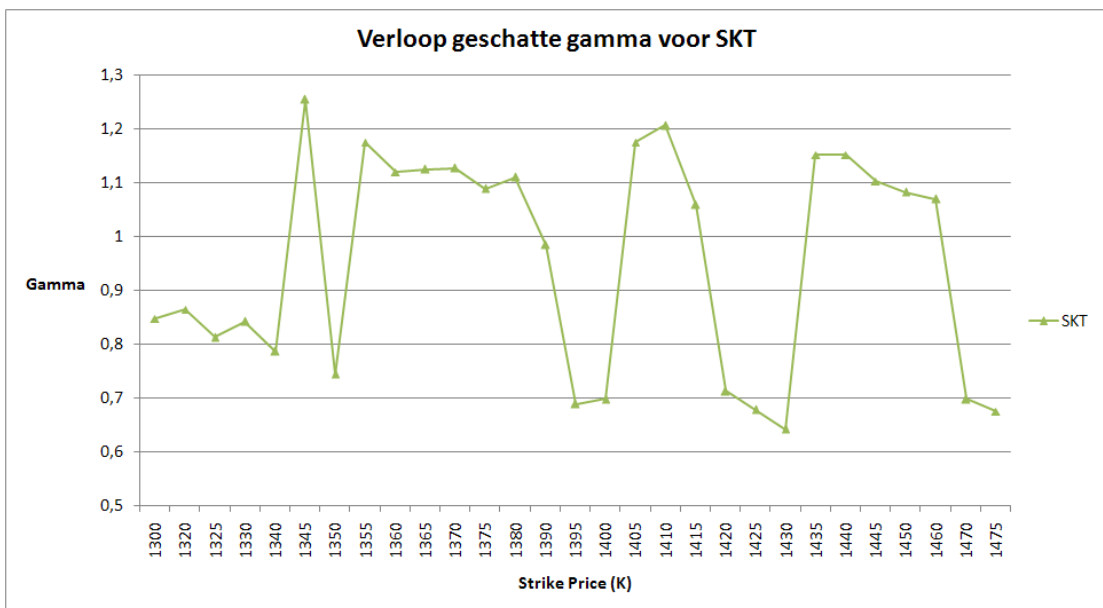
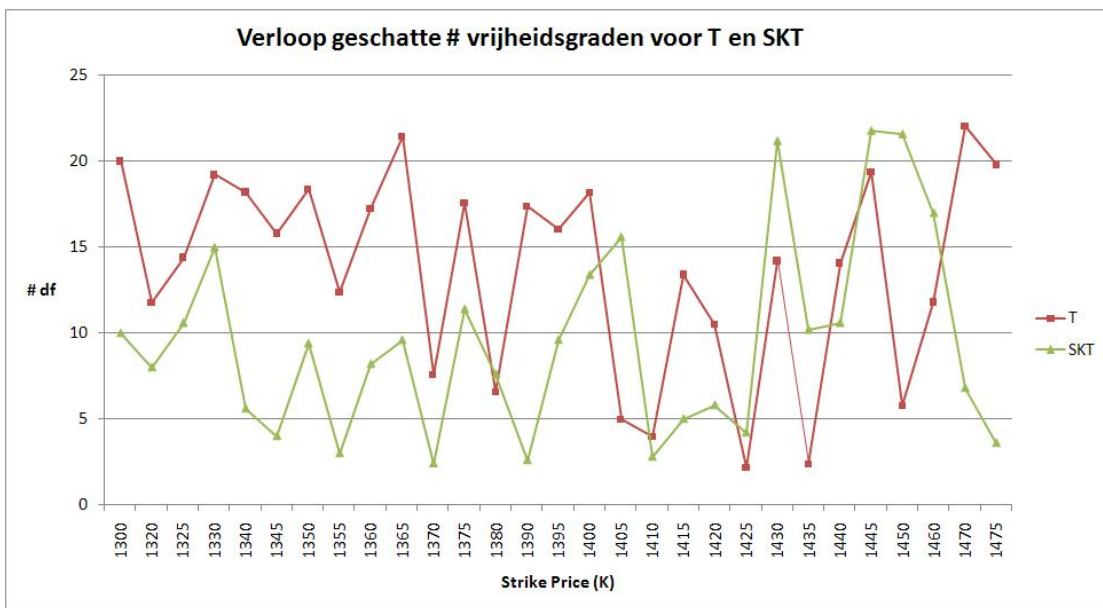
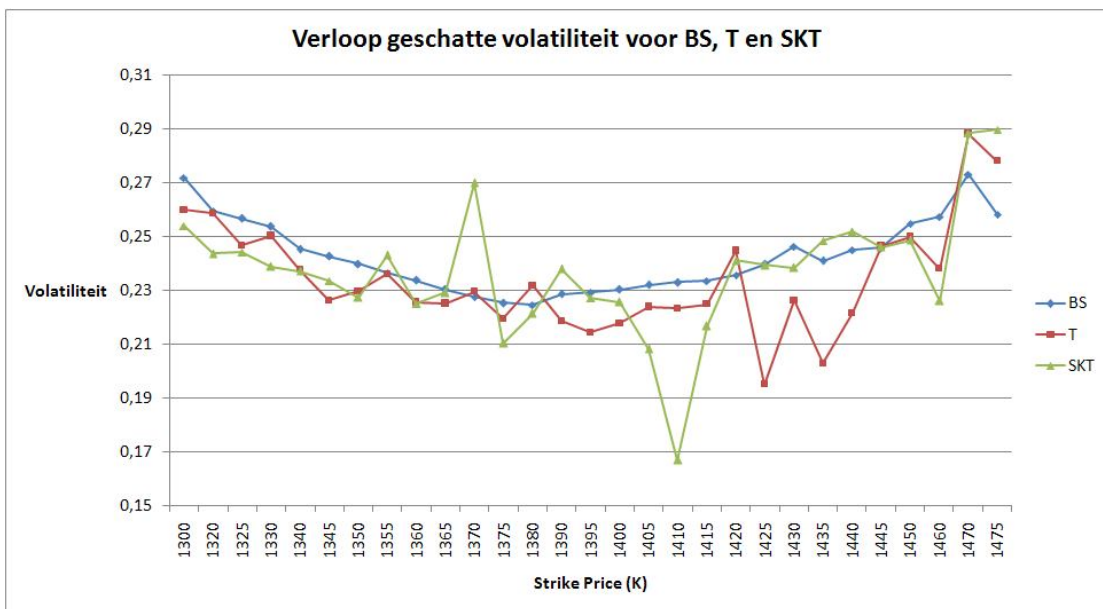
Figuur 4: Geïmpliceerde volatilititeit voor 8 december ($F_0=1337.70$, $T=5$ dagen)

De resultaten van stap 2 t/m 4 kunnen worden samengevat in tabel 1, met een korte beschrijving van de stappen en bijbehorende uitkomsten. In de tabel staat welke parameters worden geschat, welke numerieke methoden zijn gebruikt, wat de rekentijd van de methode is en de bijbehorende waarde van SSR (als som van de 29 afzonderlijke SSR-waarden). Bijbehorende optimale parameters uit stap 2 (BS, Black-Scholes), stap 3 (T, T-verdeling) en stap 4 (SKT, scheve T-verdeling) staan weergegeven in figuur 5, uitgezet tegen de uitoefenprijzen.

Uit tabel 1 kan worden afgelezen dat het eindresultaat van deze schattingsmethode een SSR is van $7.56E-15$, met een rekentijd van ca. 16,5 minuten (993.9s). Kanttekening hierbij is dat de rekentijd sterk afhankelijk is van de computer waarmee wordt gewerkt. Deze kolom is dus puur bedoeld voor een relatieve vergelijking van rekentijd van de genomen stappen.

Stap	Wat wordt geschat?	Welke numerieke methode wordt gebruikt?	Rekentijd	SSR
2	29 x σ 29 x σ	Gewichten uit Appendix 3 <i>fminsearch</i> (1 variabele: σ)	0.1 s 2.2 s	4.156 4.26E-05
3	100 x 29 x df 29 x σ en df gezamenlijk	100-staps df -grid en <i>fminsearch</i> (1 variabele: σ) <i>fminsearch</i> (2 variabelen: σ en df)	229.8 s 12.6 s	3.99E-05 2.92E-10
4a	100 x 29 x γ	100-staps γ -grid	4.9 s	0.010
4b	100 x 29 x df 29 x σ , γ en df gezamenlijk	100-staps df -grid en <i>fminsearch</i> (2 variabelen: σ en df) <i>fminsearch</i> (3 variabelen: σ , γ en df)	727.9 s 16.4 s	4.55E-13 7.56E-15
Eind			993.9 s	7.56E-15

Tabel 1 Samenvatting resultaten van stap 2 t/m 4 van het "Slim Schatten" model o.b.v. gewichten



Figuur 5 Van boven naar beneden: de geschatte parameters (σ , γ en df) van de verschillende verdelingen (BlackScholes, T en scheve T), uitgezet tegen de Strike Price (K) o.b.v. data van 8 december.

2.3 Discussie resultaten Slim Schatten

Ten koste van een relatief hoge simulatie tijd is het voor de data van 8 december 2000 gelukt om een zeer nauwkeurige benadering af te geven van de werkelijke optieprijs, met een totale SSR van $7,56E-15$. Dit is volgens verwachting, aangezien elke optieprijs afzonderlijk wordt bepaald uit een verdeling van 3 parameters. Elke “fout” die door één parameter wordt gemaakt, wordt door middel van de andere parameters (en de efficiënte *fminsearch* methode) gecompenseerd. De zeer lage SSR geeft dus eigenlijk aan dat het proces van het vinden van geïmpliceerde parameter waarden van de scheve t verdeling geslaagd is verlopen. Daarnaast geeft dit aan dat de gewichten vooral voor 2 omliggende uitoefenprijzen “belangrijk” te noemen zijn: bij de t-verdeling komt het er *ongeveer* op neer dat we telkens 2 parameters uit 2 vergelijkingen oplossen, waarbij de fout dus ongeveer 0 kan zijn. Vandaar lage SSR waarden als $3,99E-05$ en $2,92E-10$. Ook de resultaten van het model op basis van omliggende opties uit appendix 4 zijn qua SSR uitstekend, met een eindwaarde van $8,42E-14$. NB: we lossen hierbij telkens 3 parameters (voor het model van de scheve t-verdeling) op uit 3 vergelijkingen, waarbij de fout dus in wezen exact 0 kan zijn. De extreem lage waarde van $8,42E-14$ hangt hierbij waarschijnlijk af van afrondingsfouten en het convergentiecriteria van de gebruikte Matlab *fminsearch* functie.

Uit figuur 5 is het verloop van de geïmpliceerde parameters goed af te lezen. Uit de bovenste figuur blijkt dat de *volatility smile* van figuur 4 niet geheel wordt “rechtgetrokken” als de parameters γ en df worden toegevoegd. Zowel bij de t-verdeling als bij de scheve t-verdeling blijven de contouren van een *smile* zichtbaar. Wel valt het op dat de curve minder stabiel wordt naarmate er meer parameters worden toegevoegd. Dit is mede te verklaren uit het grillige verloop van de geïmpliceerde γ 's en df 's uit de overige 2 grafieken van figuur 5. De geïmpliceerde df 's vertonen ruis rondom een gemiddelde waarde van ongeveer 10 bij de scheve t-verdeling, en ongeveer 13 bij de t-verdeling. De geïmpliceerde γ 's van de scheve t-verdeling variëren rondom een gemiddelde van 0,95 en net zoals bij df is het lastig om een *patroon* te ontdekken. Zowel bij γ als bij df liggen deze gemiddelde waarden dus onder de waarden behorende bij de normale verdeling ($\gamma = 1$, $df > 20$ ⁶). Toch is de spreiding rondom deze waarden erg groot, zodat de geïmpliceerde parameters γ en df niet *voor elke uitoefenprijs* lager zijn dan de waarden behorende bij de normale verdeling (voor de data van 8 december 2000).

Het verloop van de parameters op basis van de methode met *omliggende opties*, zoals gepresenteerd in appendix 4, wijkt enigszins af van de patronen in figuur 5. Toch blijven de belangrijkste conclusies in stand: de *volatility smile* wordt niet “rechtgetrokken” bij het toevoegen van parameters, en de geïmpliceerde df 's en γ 's vertonen ruis rondom een hun gemiddelde waarden en liggen niet *voor elke uitoefenprijs* onder de waarden behorende bij de normale verdeling.

Zoals in hoofdstuk 1.2 is besproken zijn de 2 geselecteerde modellen (“slim schatten” op basis van *gewichten* en *omringende opties*) in eerste instantie alleen getest op de data van call opties op futures van de S&P 500 index van 8 december 2000. In Appendix 6 worden de resultaten van 2 andere dagen gepresenteerd met andere aflooptdata en aflooptijden, namelijk van de dagen 2 juni 2000 (T=10) en 15 maart 2000 (T=2). Wat betreft SSR-waarden ontlopen de 2 verschillende methoden (o.b.v. *gewichten* en *omringende opties*) elkaar wederom niet veel.

⁶ Met als doel een efficiënte gridkeuze in stap 3 en 4b van het model wordt ter vereenvoudiging aangenomen dat een scheve t-verdeling met $df > 20$ (bij benadering) normaal verdeeld is

Ook het verloop van de parameters is vergelijkbaar te noemen. Tussen de dagen zijn echter wel enkele verschillen te ontdekken. Waar bij de optimale geïmpliceerde σ 's van 8 december 2000 *volatility smile* te zien is, vormen de geïmpliceerde volatiliteiten op 2 juni en 15 maart een *volatility skew*. Bij toevoeging van andere parameters, wordt ook deze *skew* echter niet rechtgetrokken. De overige parameters lijken weer uit ruis te bestaan, met uitzondering van de geïmpliceerde γ 's van 15 maart. Hier is een soort “gamma skew” waar te nemen: de waarden voor γ lijken te dalen naarmate de uitoefenprijs stijgt. Het is interessant om te onderzoeken of dit verloop incidenteel is, of dat het eerder een fenomeen is voor kortlopende opties (zoals T=2). Deze hypothese wordt in dit onderzoek verder niet getoetst.

Uit tabel 1 kan worden opgemaakt dat de beginwaarden voor γ voor de scheve t-verdeling relatief oppervlakkig worden geschat. Waar men in stap 2 en stap 3 een extra *fminsearch* methode toepast om de schatting te verscherpen, bestaat stap 4a slechts uit het uittesten van een *grid* met verschillende γ -waarden. Het grote voordeel hiervan blijkt uit de lage rekentijd van ca. 5 seconden. De keuze voor deze oppervlakkige schatting voor γ wordt gerechtvaardigd in Appendix 5. Hier wordt aangetoond dat een fijnere schatting voor γ -waarden in stap 4a niet direct tot betere resultaten leidt. Een mogelijke verklaring hiervoor is dat de optimalisatie over df is een *complexere* aangelegenheid is dan de optimalisatie over γ voor het schatten van de parameters van de scheve t-verdeling⁷. In de Appendix (en in het werkcollege) wordt deze verklaring nader toegelicht.

⁷ In het werkcollege wordt df in de context van het optimalisatieproces een grotere “boosdoener” genoemd dan df . Een beknopte uitleg van de interpretatie volgt in Appendix 5.

3. Performance maatstaven

Een belangrijk onderdeel van dit onderzoek is het presenteren van een praktische toepassing van de verschillende gesimuleerde verdelingen. Waar het werkcollege zich vooral heeft beperkt tot het statistisch evalueren van de simulatieresultaten, richt dit onderzoek zich ook op de economische relevantie van alternatieve optieprijs modellen. In dit hoofdstuk worden de 2 gebruikte performance maatstaven nader toegelicht. De Diebold-Mariano toets wordt gebruikt om een *statistische* maatstaf te bepalen en wordt beknopt toegelicht in hoofdstuk 3.1. De *economische* interpretatie wordt uitgebreid toegelicht in hoofdstuk 3.2 en komt tot stand door het toepassen van Delta-Hedging. Voor dit onderzoek is deze procedure voor delta-hedging in Matlab geprogrammeerd. Voor de fijnproever is bijbehorende *pseudo-code* voor Matlab te vinden in Appendix 7.

3.1 Diebold-Mariano (DM) toets

Eenvoudig geformuleerd wordt bij deze performance maatstaf getoetst of de *errors* van de 2 geselecteerde modellen significant van elkaar afwijken. Deze *errors* worden in de DM-toets gedefinieerd als de gekwadrateerde fout tussen de gesimuleerde optieprijs en de waargenomen optieprijs voor de 2 geselecteerde modellen. De DM-toets bestaat nu uit een regressie van het verschil tussen deze *errors* op een constante:

$$(Fout_i^2 - Fout_j^2) = c + \varepsilon \quad (1)$$

De constante c kan worden gezien als de Mean Squared Error (MSE) van de toets, en wordt gebruikt als toetsgrootheid. Indien deze waarde significant afwijkt van 0⁸, kan worden geconcludeerd dat de modellen significant anders presteren. Aangezien in deze data rekening moet worden gehouden met serierecorrelatie en heteroskedasticiteit wordt gebruik gemaakt van Newey-West standaardfouten.

3.2 Delta-Hedging

Het schatten van de performance maatstaf ζ voor delta-hedging wordt in dit onderzoek voornamelijk gebaseerd op het raamwerk gepresenteerd in Hutchinson (1994). In dit artikel wordt $W(t)$ gedefinieerd als de totale waarde van de samengestelde 'hedge-portefeuille' op tijdstip t :

$$W(t) = W_F(t) + W_O(t) + W_C(t) \quad (2)$$

Hierbij staat $W_F(t)$ voor de dollar-waarde van de futures, $W_O(t)$ voor de dollar-waarde van obligaties en $W_C(t)$ de dollarwaarde van call-opties in de portefeuille op tijdstip t .

De initiële samenstelling van deze portefeuille ($t=0$) wordt als volgt vastgesteld:

$$W_F(0) = F(0) \Delta_{SKT}(0) \quad (3) \text{ met: } \Delta_{SKT}(t) \equiv \frac{\partial MC_{SKT}(t)}{\partial F}$$

$$W_C(0) = -C(0) \quad (4)$$

$$W_O(0) = -(W_F(0) + W_C(0)) \quad (5)$$

⁸ In dit onderzoek wordt een significantieniveau gebruikt van 10%, net zoals in het werkcollege.

Welke delta wordt gebruikt in formule (3) hangt af van de verdeling wordt geëvalueerd (in dit onderzoek BS of SKT). Als de hedge prestatie van BlackScholes wordt beschouwd, zal niet $\Delta_{SKT}(t)$ gebruikt worden, maar $\Delta_{BS}(t)$ (=dBS/dF). Hierbij staat BS voor de benadering van de optieprijs, welke via de BlackScholes formule analytisch te bepalen is. In formules (3)-(5) wordt het model echter op basis van de scheve t-verdeling geëvalueerd, met $MC_{SKT}(0)$ de Monte Carlo benadering van het optieprijs model met een scheve T-verdeling (SKT) op tijdstip $t=0$. $\Delta_{SKT}(t)$ kan worden gezien als afgeleide van de SKT-benadering (“scheve-T”) van de optieprijs ten opzichte van de futuresprijs (F). $\Delta_{SKT}(t)$ is analytisch niet te bepalen doordat de benadering van de optieprijs zelf door simulatie tot stand is gekomen. Bovendien veronderstelt de afgeleide functie een continue hedging strategie, met (theoretisch gezien) een oneindig kleine stapgrootte. Aangezien in dit onderzoek een discrete hedge-strategie wordt gebruikt, wordt $\Delta_{SKT}(t)$ benaderd op basis van een “discrete” stapgrootte. Een logisch alternatief voor de benadering van de afgeleide is:

$$\Delta_{SKT}(t) = [MC_{SKT}(t) (F(t) + stdevF) - MC_{SKT}(t) (F(t) - stdevF)] / (2 * stdevF) \quad (6)$$

$$\text{met: } stdevF = F(t) * \sigma / \sqrt{252} \quad (7)$$

waarbij $stdevF$ een benadering is voor de dagelijkse standaard deviatie van het futures proces. Hierbij fungeert $\sigma / \sqrt{252}$ als proxy van de standaarddeviatie van het dagelijkse rendement, waarbij σ de geïmpliceerde volatiliteit uit de geselecteerde verdeling voorstelt. Een combinatie van vergelijking (6) en (7) geeft als schatting voor delta:

$$\Delta_{SKT}(t) = [MC_{SKT}(t) (F(t) + F(t) * \sigma / \sqrt{252}) - MC_{SKT}(t) (F(t) - F(t) * \sigma / \sqrt{252})] / (2 * F(t) * \sigma / \sqrt{252}) \quad (8)$$

Op basis van deze delta wordt de portefeuille uit vergelijking (2) samengesteld, door het verkopen van 1 call-optie (short) en het aankopen van delta keer de onderliggende futures (long). Aangezien deze aankoop van futures volledig wordt gefinancierd door een combinatie van de opbrengsten van een risicovrije lening en de opbrengsten van de verkoop van 1 call-optie, is de initiële waarde van de portfolio gelijk aan 0:

$$W(0) = W_F(0) + W_O(0) + W_C(0) = 0 \quad (9)$$

Zolang de optie nog niet is verlopen, wordt de portefeuille dagelijks herzien om de volgende vergelijkingen te voldoen:

$$W_F(t) = F(t) \Delta_{SKT}(t) \quad (10)$$

$$W_O(t) = e^{r \tau / 252} W_O(t - \tau) - F(t) (\Delta_{SKT}(t) - \Delta_{SKT}(t - \tau)) \quad (11)$$

In (11) staat τ voor de stapgrootte voor het interval waarin de portfolio wordt herzien. Aangezien de portfolio in dit onderzoek dagelijks wordt herzien, wordt aan τ de waarde 1 toegekend. Vergelijking (11) geeft weer dat de schuld (obligaties) toeneemt door de rente over de schuld op de dag ervoor, en toe- of afneemt doordat de delta toe- of afgenomen is (zodat er aandelen gekocht of verkocht worden).

De *tracking error* van de portfolio op de afloopdatum van de optie ($t=T$) is nu simpelweg de waarde van de portfolio op tijdstip T; $W(T)$. Op basis van de *tracking error* kan de performance maatstaf ξ worden bepaald:

$$\xi \equiv e^{-rT} E[|W(T)|] \quad (12)$$

3.3 Voorbeeld: Verschil Delta-Hedging en DM-toets

Het grootste verschil tussen beide performance maatstaven is goed weer te geven door het evalueren van de resultaten van het “Slim Schatten” model op basis van *gewichten* (hoofdstuk 2.2). Het (aangepaste) Black-Scholes geeft bijvoorbeeld in slechts 2 seconden een nauwkeurige schatting af van de werkelijke optieprijs, met een SSR van $4.26 \text{ E-}05$. De scheve t-verdeling levert een nog scherpere optieprijs op, met een SSR van $7.56 \text{ E-}15$. Wel heeft de scheve t-verdeling maar liefst 994 seconden nodig voor het optimalisatieproces. De vraag rijst nu: welke optiewaardering is beter?

Statistisch gezien heeft deze vergelijking weinig waarde, aangezien de methode op basis van *gewichten* in staat is om een (bijna) perfecte fit van de optieprijs af te geven. Met andere woorden: het heeft weinig zin om een SSR van $4.26 \text{ E-}05$ te vergelijken met een SSR van $7.56 \text{ E-}15$ te vergelijken (aangezien beide resultaten dezelfde optieprijs afgeven). Economisch gezien heeft deze vergelijking echter meer diepgang: zijn deze optiewaarderingen geschikt om de **afgeleide** van de optieprijs ten opzichte van de futuresprijs te berekenen? Deze *afgeleide* is van belang voor het delta-hedgen van een optie, zoals eerder is besproken in dit hoofdstuk. Zowel door het BS-model als het SKT-model wordt de optieprijs precies teruggevonden, maar het is mogelijk dat de afgeleide van optieprijs naar futures prijs iets beter door het SKT-model wordt gegeven. Om deze hypothese te toetsen wordt in het onderstaande een Delta-Hedging experiment opgezet op grond van de “slim schatten” methode op basis van *gewichten*.

Nadeel van de procedure uit hoofdstuk 3.2 is dat er discreet “gehedged” wordt. Theoretisch gezien zou de portefeuille continu moeten worden herzien, doordat de afgeleide een oneindig kleine stapgrootte veronderstelt. Ter vereenvoudiging van het experiment wordt hier echter gekozen voor een stapgrootte van 1 dag, zodat slechts één update van de portfolio per dag mogelijk is. Op deze manier is het mogelijk dat een foutief model door een “toevalstreffer” beter presteert dan een model dat theoretisch gezien veel beter is. Om voor dit soort misleidende effecten te “filteren” is het belangrijk om het experiment voor verschillende data en looptijden te herhalen. Over meerdere experimenten middelen dit soort pech-en-geluk factoren, zodat een beter beeld ontstaat welk model *structureel* beter presteert.

Voor het experiment wordt voor alle afloopdata van het jaar 2000 (in maart, juni, september en december) de uitoefenprijs geselecteerd waarin het meest gehandeld is. Voor zowel maart, juni en december bleek dit $K=1450$ te zijn, voor september was dit $K=1500$. Van deze afloopdata worden 3 verschillende aflooptijden geëvalueerd: $T=10$, $T=20$ en $T=30$. Zo ontstaan in totaal 12 combinaties van afloopdatum/uitoefenprijs en aflooptijd. Door het toepassen van de methode zoals gepresenteerd in uit hoofdstuk 3.2 kunnen voor het BS-model en het SKT-model de tracking errors geschat worden. De resultaten van dit experiment zijn te vinden in tabel 2 waarin de tracking errors in matrixvorm worden weergegeven. Bijbehorende rekentijd (voor het SKT-model) staat weergegeven in tabel 3.

	MAART		JUNI		SEPTEMBER		DECEMBER	
	BS	SKT	BS	SKT	BS	SKT	BS	SKT
T=10	8,35	2,32	6,95	6,33	24,44	25,56	1,31	0,48
T=20	6,84	7	3,7	3,09	18,135	18,494	15,17	12,07
T=30	28,45	29,783	3,69	8,39	4,328	3,936	36,328	34,321

Tabel 2 De tracking errors van het experiment samengevat in matrixvorm.

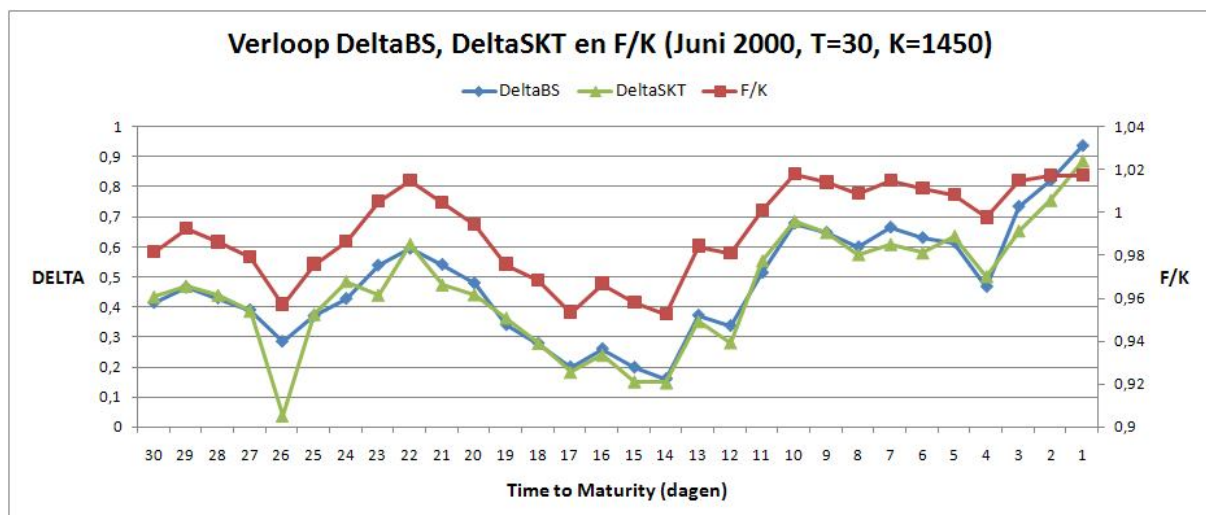
In de rijen staat de time-to-maturity (aflooptijd) van de optie, in de kolom de maand van de afloopdatum. Per kolom wordt onderscheid gemaakt tussen de uitkomst van het BS-model (linker subkolom) en het SKT-model (rechtsubkolom).

	MAART	JUNI	SEPTEMBER	DECEMBER
T=10	438	319	440	357
T=20	1497	1095	1297	1027
T=30	2983	2109	3045	2243

Tabel 3 De rekentijd van het SKT model voor delta hedging om tot de tracking errors uit tabel x te komen

Uit deze resultaten blijkt dat het SKT-model 7 van de 12 keer een betere tracking error afgeeft. Over deze 12 experimenten heeft de SKT een kleinere gemiddelde tracking error (van 12.66) ten opzichte van het BS-model (13.00). Het lijkt erop dat het SKT-model een net iets betere aanpak heeft dan BS. Het zal zeker voor kunnen komen dat SKT slechter presteert dan BS, maar gemiddeld genomen zijn de resultaten van het SKT-model dus licht in het voordeel. Groot nadeel van het SKT-model is echter de rekentijd, zoals weergegeven in tabel 3. In hoofdstuk 4 wordt een methode gepresenteerd waarin de rekentijd gereduceerd kan worden.

Om te zien of deze delta-hedging experimenten correct zijn verlopen, is het nuttig om het verloop van $\Delta_{SKT}(t)$ en $\Delta_{BS}(t)$ over de tijd te bekijken. Bij zeer hoge futures prijzen⁹ is de kans dat de optie uitgeoefend wordt ongeveer 1 en zal ook de delta richting 1 gaan. Bij zeer lage future prijzen is de kans juist klein dat de optie uitgeoefend zal worden: de delta zal dan richting 0 bewegen. Door naast het verloop van de 2 verschillende delta's ook de ratio F/K mee te nemen in de figuur, kan als het ware *gecheckt* worden of de berekening van delta op een correcte manier is gebeurd. In figuur 6 staat een dergelijke grafiek afgebeeld voor de uitkomsten van juni, met *delta* in de primaire verticale as (links), en de ratio F/K in de secundaire verticale as (rechts). Het valt op dat de delta's van BS en SKT een vergelijkbaar verloop hebben, en dat de delta richting 1 gaat wanneer de ratio F/K hoog blijft naarmate de afloopdatum in zicht komt. In appendix 10 staan de figuren van de overige afloopdata weergegeven. Ook deze figuren vertonen bovenstaande kenmerken, zodat het verloop van de delta's plausibel is te noemen. Een vergelijking van figuur 6 en tabel 2 kan ook verhelderend werken. Zo lijkt bijvoorbeeld de relatief slechte tracking error van SKT (ten opzichte van BS) voor $T=30$ van Juni (8,39 voor SKT ten opzichte van 3,69 voor BS) samen te hangen met de extreem lage Δ_{SKT} in de grafiek op $T=26$.



Figuur 6 Verloop van DeltaBS, DeltaSKT en F/K over de tijd van het Delta-Hedging experiment voor juni 2000. In de primaire verticale as staan de 2 delta's weergegeven, in de secundaire verticale as staat de ratio F/K .

⁹ Zodat de ratio future/strike (veel) groter is dan 1: $F/K \gg 1$

4 “Real-time” optiemodel

Uit de resultaten van hoofdstuk 2 en het werkcollege blijkt dat nauwkeurige benaderingen van optieprijsen gepaard gaan met relatief hoge rekentijden. Vooral het gezamenlijke optimalisatie proces neemt veel tijd in beslag, zodat de modellen nauwelijks toepasbaar zijn in de praktijk. Als oplossing voor dit probleem wordt in het onderstaande een model gepresenteerd waarin de resultaten van alle simulaties van tevoren in een database worden opgeslagen. In deze database worden op basis van slim gekozen *grids* alle relevante combinaties van de parameters opgeslagen. Met deze database kunnen de parameters in “real-time” met behulp van een (al dan niet) lineaire interpolatie functie geschat worden. Zo wordt het tijdrovende proces van simuleren en optimaliseren uit het werkcollege (2008) omzeild. Hoofddoel van dit onderzoek is het terugbrengen van de rekentijd met een zo nauwkeurig mogelijke schatting van de werkelijke optieprijs.

4.1 Methoden: “Real Time” optiemodel

Het vormen van een praktische en relevante database is cruciaal voor het slagen van het real-time optiemodel. Dit is in eerste instantie een lastige opgave, aangezien het gesimuleerde optieprijs model op basis van de scheve T-verdeling uit werkcollege (2008) maar liefst 7 argumenten telt:

$$\text{waarde call optie} = e^{-rT} * \text{mean}(\text{payoff})$$

$$\text{met: } \text{payoff} = \max(F_T - K, 0)$$

$$\text{en: } F_T = \text{functie}(F_0, \sigma, \gamma, df, T).$$

$$\text{zodat uiteindelijk: } \text{waarde call optie} = \text{functie}(r, K, F_0, \sigma, \gamma, df, T),$$

Aangezien deze argumenten binnen het optie-waarderings probleem onderling sterk afhankelijk zijn, moeten alle 7 argumenten worden meegenomen in het real-time optiemodel. Voor de database waarin de simulaties worden uitgevoerd is dit echter niet het geval. Het aantal argumenten voor de database kan worden teruggebracht door de dimensie van het probleem te reduceren. Dit is mogelijk door van tevoren alleen $\text{mean}(\text{payoff}/K)$ in de database op te slaan, in plaats van de totale functie $e^{-rT} * \text{mean}(\text{payoff})$. Als tijdens de optimalisatie methode van het real time-model (als het ware “achteraf”) met de factor $e^{-rT} * K$ wordt vermenigvuldigd, hangt de functie-waarde niet van r af. Dit resulteert in een functie met 5 argumenten:

$$\text{mean}(\text{payoff}/K) = \text{functie}(F_0/K, \sigma, \gamma, df, T).$$

Een database met 5 dimensies is echter nog steeds complex om mee te werken, en bovendien inefficiënt voor de interpolatie functie. Voor de vereenvoudiging van dit probleem, zal dit onderzoek zich beperken tot het uitvoeren van de analyse voor slechts 1 specifieke T. Op deze manier resteren er slechts 4 argumenten voor de database: F_0/K , σ , γ en df . Een overzicht van het resulterende tweeledige probleem wordt gegeven in tabel 2. Uit deze tabel is op te maken dat de database uit een 4-dimensionaal *array* bestaat met als argumenten *grids* van waarden voor F_0/K , σ , γ en df en de bijbehorende waarden van $\text{mean}(\text{payoff}/K)$. Het is van belang om deze *grids* slim te kiezen, om de rekentijd van de interpolatie functie in het “real-time” model

te minimaliseren en om de rekentijd voor genereren van de simulatie database ook nog enigszins te beperken. Een voorbeeld van grids voor de 4 variabelen is:

F_0/K	in [0.80, 1.20]:	40 subintervals van breedte 0.01
σ	in [0.10, 0.40]:	15 subintervals van breedte 0.02
df	in [2,15]:	26 subintervals van breedte 0.5
γ	in [0.40, 1.20]:	16 subintervals van breedte 0.05

De stap die nu resteert is het berekenen van $mean(payoff/K)$ voor alle gridpunten met behulp van Monte Carlo Simulatie.

Voor toepassing van het “real-time” model is meer optie-specifieke informatie nodig, zoals F_0, K, C, r en T . Een overzicht van de verschillende variabelen in de *database* en het *real time* model staat weergegeven in tabel 4. Ook is er een *grid* nodig van de impliciete parameters die geoptimaliseerd moeten worden: σ, γ en df . Met hulp van deze 7 argumenten en de database die van tevoren is opgesteld kan middels het volgende 5-stappenplan een schatting gemaakt worden van de optieprijsen:

Stap 1: creëren van fijner *grid* van waarden van σ, γ en df , zoals:

σ	in [0.10, 0.40]:	30 subintervals van breedte 0.01
df	in [2,15]:	52 subintervals van breedte 0.25
γ	in [0.40, 1.20]:	32 subintervals van breedte 0.025

Voor F_0/K kan de vector worden genomen van de waargenomen huidige futures prijs gedeeld door de strike prices (optie-specifieke informatie).

Stap 2: Door middel van interpolatie tussen de punten van het “ruwere” grid uit de database, kan voor alle gridpunten op het fijnere grid uit stap 1 de waarde $mean(payoff/K)$ berekend worden.

Stap 3: De gefitte optiewaarden kunnen nu berekend worden door de elke waarde $mean(payoff/K)$ met de factor $e^{-rT} * K$ te vermenigvuldigen.

Argumenten Database	Argumenten “Real Time” model
Grid van waarden voor: - F_0 / K - σ - df - γ	Specifieke informatie van optie: - F_0 - Vector van uitoefenprijzen (K) - Vector van optieprijsen (C) - r - T Grid van waarden voor: - σ - df - γ

Tabel 4 Overzicht van de “real-time” optiewaarderings methode obv simulatie. Links de argumenten die “van tevoren” in een *database* worden opgeslagen en rechts de argumenten die in *real-time* nodig zijn.

Stap 4: Bepaal voor elke (σ, γ, df) combinatie de SSR: de som van de gekwadrateerde verschillen tussen de gefitte optiewaarden en de marktprijzen van de opties voor de verschillende uitoefenprijzen.

Stap 5a: Bepaal de minimale SSR waarde van stap 4. De bijbehorende $(\sigma, \gamma$ en $df)$ combinatie is de optimale combinatie van σ, γ en df .

Stap 5b: De gevonden optimale combinatie van σ, γ en df kan nog gezamenlijk worden geoptimaliseerd met behulp van het *fminsearch* algoritme (toegepast op de SSR functie van de markt-optieprijzen en de - uit interpolatie tussen database punten afkomstige - gefitte optiewaarden behorende bij sigma, df, gamma). Nadeel van deze extra optimalisatie is dat het extra rekentijd kost. In het volgende hoofdstuk zal dieper in worden gegaan op deze afweging tussen rekentijd en nauwkeurigheid.

Met het hoofddoel van dit onderzoek in het achterhoofd, namelijk het vinden van een praktisch uitvoerbare optiewaarderingsmethode die in “real time” niet teveel rekentijd vergt (en beter presteert dan het Black-Scholes model), blijkt uit het stappenplan dat de keuze voor de grids van parameters een cruciale stap is. Kiest men de grids te ruim, dan worden er te veel “zinloze” combinaties van (σ, γ, df) geëvalueerd, wat ten koste gaat van de rekentijd van het model. In het volgende hoofdstuk wordt een *grid-analyse* gedaan, waar de effecten van het *verfijnen* en het *verruwen* van de grids worden onderzocht. Wat is het effect op de rekentijd en de SSR als je het aantal gridpunten in de database verdubbelt? Welke parameter heeft bij een verdubbeling van het aantal gridpunten in de database het grootste positieve effect? En welk grid kan je in het Real Time model juist *ruwer* nemen om tijdwinst te boeken?

4.2 Grid-Analyse

In het bovenstaande stappenplan zijn 2 voorbeelden gegeven van grids, zowel voor de **database**, als voor het **real-time** optiemodel. In deze analyse worden de intervallen van de parameters in eerste instantie vast gekozen. De geselecteerde intervallen zijn:

- [0.80, 1.20] voor F_0/K
- [0.10, 0.40] voor σ
- [2,15] voor df
- [0.40, 1.20] voor γ

In deze *grid-analyse* wordt onderzocht wat het effect is van de stapgrootte binnen deze intervallen. In het bovenstaande voorbeeld wordt voor *de grids* in de database een stapgrootte van 40 waarden van F_0/K gekozen, 15 waarden van σ , 26 voor df en 16 voor γ . Deze keuze van stapgroottes wordt voortaan aangeduid als *40-15-26-16* en dient in het onderstaande als “controlegroep” voor de database. Door telkens 1 grid van de database te verdubbelen (en daarmee ook het aantal *gridpoints*) en het effect te vergelijken met de controlegroep, kan het effect van een verfijning van de grids in de database onderzocht worden.

Ook binnen het real-time optiemodel zijn grids nodig voor σ, γ en df . Het bovenstaande voorbeeld verdubbelt het aantal *gridpoints* ten opzichte van de grids in de database: 30 stappen voor σ , 52 voor df en 32 voor γ . Deze keuze van de 3 stapgroottes wordt voortaan aangeduid als *30-52-32* en dient in het onderstaande als “controlegroep” voor het “real-time” model. N.B: de F_0/K waarden zijn hierbij gegeven door de waargenomen waarden, zodat in real time er slechts 3 gridkeuzes gemaakt worden. Er kan tijdwinst worden geboekt als een *verruwing* van het grid geen verstreckende gevolgen heeft voor de waarde van de SSR (en de optimale waarden van σ, γ en df). Door telkens het aantal gridpoints van 1 grid binnen het *real*

time model te halveren en het effect te vergelijken met de controlegroep, kan het effect van een *verruwing* van het grid van het real time model onderzocht worden.

In tabel 5 en 6 staan de resultaten van dergelijke onderzoeken, met in de kolommen de gridkeuze voor de database (met een fijner grid voor F_0/K , σ , df en γ) en in de rijen de gridkeuze voor het *real time* model (met een ruwer grid voor σ , df en γ). In tabel 5 staan de resultaten t/m stap 5a, in tabel 6 is ook de extra gezamenlijke optimalisatie van stap 5b uitgevoerd. In de tabel staan de SSR-waarden weergegeven, met tussen haakjes de benodigde rekentijd¹⁰. In Appendix 7 staan de resultaten uitgebreid weergegeven, inclusief de bijbehorende parameter-waarden.

STAP 5A		Controle 40-15-26-16	F_0/K 80-15-26-16	σ 40-30-26-16	df 40-15-52-16	γ 40-15-26-32
Controle	30-52-32	0.17 (29s)	0.24 (11s)	0.18 (22s)	0.17 (21s)	0.15 (21s)
σ	15-52-32	0.21 (11s)	0.25 (6s)	0.21 (12s)	0.33 (11s)	0.21 (12s)
df	30-26-32	0.17 (11s)	0.29 (6s)	0.18 (10s)	0.17 (10s)	0.15 (11s)
γ	30-52-16	0.21 (10s)	0.25 (6s)	0.21 (11s)	0.30 (11s)	0.21 (11s)

Tabel 5 Resultaten grid-analyse t/m stap 5a. In de kolommen staat een *verfijning* van de parameter-grids in de *database* weergegeven (resp. F_0/K , σ , df en γ). In de rijen staat een *verruwing* van de parameter-grids in het *real-time* model weergegeven (resp. σ , df en γ).

STAP 5B		Controle 40-15-26-16	F_0/K 80-15-26-16	σ 40-30-26-16	df 40-15-52-16	γ 40-15-26-32
Controle	30-52-32	0.13 (214s)	0.10 (23s)	0.13 (352s)	0.12 (463s)	0.12 (331s)
σ	15-52-32	0.13 (302s)	0.10 (25s)	0.13 (485s)	0.13 (680s)	0.12 (557s)
df	30-26-32	0.13 (119s)	0.10 (18s)	0.13 (343s)	0.12 (452s)	0.12 (311s)
γ	30-52-16	0.13 (301s)	0.10 (26s)	0.13 (488s)	0.13 (397s)	0.12 (559s)

Tabel 6 Resultaten grid-analyse t/m stap 5b. Interpretatie analoog aan tabel 5.

Uit de verschillen tussen tabel 5 en 6 is op te maken dat in alle gevallen een extra gezamenlijke optimalisatie betere SSR-waarden afgeeft. In de meeste gevallen (de controlegroep en de verfijnde grids in de *database* van σ , df en γ) heeft dit echter wel een drastische stijging in rekentijd tot gevolg. Het is zeer opvallend dat een fijnere keuze voor het grid voor F_0/K (80 stappen i.p.v. 40) zich aan deze aanzienlijke stijging in rekentijd onttrekt. Uit de relatief (extreem) lage rekentijd voor het fijnere grid van F_0/K in tabel 6 blijkt dat de interpolatiefunctie zeer efficiënt werkt bij een fijner grid voor F_0/K . Dit in tegenstelling tot een verfijning van σ , df of γ in de *database*: de daling van de SSR gaat gepaard met een zeer hoge rekentijd.

Verder is uit tabel 5 op te maken dat een *verruwing* van de grids in het *real time* model tijdwinst met zich meebrengt. Wordt echter een extra gezamenlijke optimalisatie uitgevoerd (stap 5B), dan hoeft een *verruwing* van het grid niet direct tot tijdwinst te leiden (zie tabel 6). Opvallend is dat een ruwer grid voor df in het *real time* model (26 stappen i.p.v. 52) vaak tot dezelfde SSR leidt, in combinatie met een lagere rekentijd. N.B. In wezen is dit ook telkens hetzelfde optimalisatie probleem, met alleen een mogelijk andere beginwaarde.

De interpolatiefunctie lijkt dus weinig moeite te hebben met een ruwer grid voor df , ook niet als een gezamenlijke optimalisatieprocedure wordt toegevoegd: de rekentijd neemt dan niet toe. Dit in tegenstelling tot een ruwer grid voor σ en γ : in tabel 6 schiet de rekentijd omhoog bij een *verruwing* van 1 van deze grids. σ en γ lijken nauw samen te hangen, aangezien de resultaten van *verruwing* van de grids vergelijkbaar zijn.

¹⁰ Zoals eerder vermeld kan de rekentijd alleen gebruikt worden voor een relatieve vergelijking.

Op basis van bovenstaande constatering kan geconcludeerd worden dat een fijner grid in de *database* voor F_0/K en een ruwer grid voor df in het *real time* model perspectief bieden voor winst in rekentijd en een verscherping van de SSR. Daarom is er vervolg onderzoek gedaan naar de gevolgen van een *nog* fijner grid voor F_0/K in de *database* in combinatie met het ruwe grid van het *real-time* model. In tabel 7 en 8 (resp. t/m stap 5a en stap 5b) staan de uitkomsten van 3 alternatieve grids in de *database* (met elk een verfijning van F_0/K), en 2 alternatieve grids in het *real time* model (de controlegroep en het ruwe grid voor df):

STAP 5A	80-15-26-16	120-15-26-16	160-15-26-16
Controle 30-52-32	0.24 (11s)	0.28 (11s)	0.29 (11s)
Df 30-26-32	0.29 (6s)	0.31 (6s)	0.32 (6s)

Tabel 7 Uitkomsten van vervolgonderzoek van de gridanalyse t/m stap 5a met een fijner grid voor F_0/K in de *database* (kolommen) en een ruwer grid voor df in het *real time* model (2e rij).

STAP 5B	80-15-26-16	120-15-26-16	160-15-26-16
Controle 30-52-32	0.10 (23s)	0.08 (31s)	0.09 (27s)
Df 30-26-32	0.10 (18s)	0.08 (21s)	0.09 (19s)

Tabel 8 Uitkomsten vervolgonderzoek t/m stap 5b. Interpretatie analoog aan tabel 7.

Uit een vergelijking van tabel 7 en 8 blijkt dat het zeker loont om een extra optimalisatie uit te voeren (stap 5b): de SSR daalt sterk ten opzichte van de SSR in stap 5a, terwijl de rekentijd in alle gevallen ver onder de minuut blijft. Bovendien is het in stap 5b geen probleem om het df -grid ruwer te kiezen, te zien aan de onveranderde SSR in de 2 rijen van tabel 8. Het is opvallend dat een verdere verfijning van het F_0/K -grid in de *database* niet altijd een lagere SSR en/of rekentijd met zich meebrengt. Van bovenstaande alternatieven presteert het grid met 120 stappen voor F_0/K het best.

In appendix 9 wordt aanvullend onderzoek gedaan naar het effect van een “exponentieel” df -grid in plaats van (standaard) lineair grid. De achterliggende gedachte van dit *exponentiële* grid is dat er zoveel mogelijk relevante waarden voor df geëvalueerd worden. De t -verdeling is namelijk gevoeliger voor een kleine stijging van df voor lage waarden van df dan voor hoge waarden. Immers, een stijging van 0.1 in df -waarde heeft meer impact op $df=3$ dan op $df=30$. Het is daarom interessant om te onderzoeken of een *exponentieel* verloop van stapgrootte binnen het grid van voor df tot betere resultaten leidt dan een *lineair* verloop. Het exponentiële verloop komt overeen met dat de waarden van $\log(df)$ equidistant uniform verdeeld zijn, wat in appendix 9 nader wordt toegelicht. In deze appendix worden ook de resultaten van het *exponentiële* grid gepresenteerd (voor uitgebreide resultaten wordt u wederom naar appendix 8 verwezen). Na de gezamenlijke optimalisatie van stap 5b zijn de resultaten van het *exponentiële* df -grid helaas niet substantieel beter dan de resultaten op basis van het standaard lineaire grid. De waarden van SSR zijn voor beide grids vergelijkbaar, terwijl de rekentijd van het *exponentiële* grid voor elk experiment hoger ligt. Het *exponentiële* grid presteert t/m stap 5a echter wel een stuk beter dan het lineaire grid. Ondanks een (iets) hogere rekentijd is de SSR-waarde veel lager dan bij het lineaire grid. De conclusie van appendix 9 is dat het *exponentiële* grid niet op voorhand moet worden afgeschreven door de hogere rekentijd na stap 5b. Indien blijkt dat de *fminsearch* methode niet voor iedere dataset even efficiënt werkt, zou de relatief goede prestatie na stap 5a van het *exponentiële* grid namelijk een groot voordeel kunnen opleveren.

5 Conclusie & Discussie

Een belangrijke conclusie uit dit onderzoek is dat de scheve t-verdeling (die in het werkcollege wordt gepresenteerd als “wondermiddel”) ook zijn beperkingen heeft. Uit hoofdstuk 2 blijkt dat lang niet altijd voldaan aan de aanname van constante impliciete parameters. Het meest sprekende voorbeeld hiervan is dat de *volatility smile* of *skew* van de onderzochte data niet wordt “rechtgetrokken” bij het toevoegen van extra parameters. Verder wordt aangetoond dat de parameters sigma en gamma niet voor alle uitoefenprijzen lager zijn dan de waarden behorende bij de normale verdeling, mede doordat ze relatief grote spreiding vertonen. Voor de data van 15 maart 2000 met aflooptijd 2 dagen ($T=2$) is zelfs een aflopend verloop van de gamma te zien (“gamma skew”). Vervolgonderzoek moet uitwijzen of dit een incident betreft of eerder een eigenschap is van kortlopende opties. Het “juiste” model heeft echter (*ongeveer*¹¹) constante impliciete parameters over uitoefenprijzen en looptijden (voor optiedata betreffende hetzelfde handelsmoment) omdat dit de eigenschappen van de onderliggende (risiconeutrale) verdeling van hetzelfde onderliggende proces zijn en niet van de opties.

De optieprijs wordt door zowel het BS-model en het SKT-model (scheve t-verdeling) in hoofdstuk 2 precies teruggevonden, maar de resultaten voor delta-hedging suggereren dat de *afgeleide* van optieprijs naar futures prijs door het SKT-model iets beter wordt gegeven. Vervolgonderzoek moet uitwijzen of de resultaten van het werkcollege (waar op basis van de statistische DM-toets gezien significant betere 'prestatie' Scheve-T tov BS) ook economisch gezien beter zijn. Nadeel van delta-hedging op basis van het SKT-model is echter de aanzienlijke rekentijd in vergelijking met delta-hedging op basis van het BS-model.

De gepresenteerde *real time methode* is een veelbelovende toepassing, aangezien dit model in staat is om de rekentijd van het optie-model sterk te reduceren. In hoofdstuk 4 is de basis gelegd voor een optimaal gebruik deze methode, waarin de keuze van grids voor de *database* en de beginwaarden voor de interpolatie cruciaal zijn. Voor de data van 8 december 2000 blijkt het bijvoorbeeld zeer efficiënt om in de database het grid van F_0/K fijn te kiezen, terwijl bij de keuze voor beginwaarden voor de interpolatie het grid voor *df*-waarden relatief ruim genomen kan worden¹². Vervolgonderzoek moet uitwijzen of deze constatering voor 8 december robuust te noemen zijn. Ook vervolgonderzoek naar een ander verloop van de grids in de database waarin meer *relevante* waarden voor de parameters in de database worden geëvalueerd, biedt perspectief. Een exponentieel verloop van het *df-grid* is getest, maar gaat gepaard met (iets) hogere rekentijden ten opzichte van een standaard lineair grid.

Het combineren van de real-time methode en delta-hedging is een zeer interessante toepassing voor vervolgonderzoek¹³. Het is bijvoorbeeld interessant om de resultaten van het werkcollege (een statistische verbetering van het *scheve-t* model t.o.v. het Black Scholes model) in het kader van de *real time* methode te herzien, en dit model toe te passen op de delta-hedging methode gepresenteerd in hoofdstuk 3. Zo kan naast de *statistische* vergelijking van het werkcollege ook een *economische* interpretatie aan het resultaat worden gegeven.

¹¹ Ongeveer, vanwege inefficiënties op de optiemarkt en bijvoorbeeld dat de optieprijs alleen veelvoud van \$0,10 kunnen zijn.

¹² Voor F_0/K bleek een fijn grid in de database van 120 stappen binnen bandbreedte [0.80, 1.20] zeer efficiënt, voor de beginwaarden van *df* voor de interpolatie bleek 26 stappen voldoende binnen de bandbreedte [2, 15].

¹³ Hier moet per dag waarop de portfolio wordt herzien een andere database worden gebruikt (elke database gebruikt 1 specifieke T , zie hoofdstuk 4.1). Uit tijdgebrek bleek dergelijk onderzoek voor dit verslag helaas niet haalbaar.

6.Referenties

Black, F., Scholes, M., 1973. *The pricing of options and corporate liabilities*. Journal of Political Economy 81: 637-659.

Boyle, P.P., 1977. *Options: a Monte Carlo approach*. Journal of Financial Economics 4, 32-38.

Hull, J.C., 2006. *Options, futures and other derivatives. 6editie*. Pearson Prentice Hal.

Hutchinson, J.M, Lo, A.W., Poggio, T., 1994. *A nonparametric approach to pricing and hedging derivative securities via learning networks*. Journal of Finance 49 (3), 851-89.

Jackwerth, J.C. and Rubinstein, 1996, M., *Recovering Probability Distributions from Option Prices*,. Journal of Finance, Volume 51, December 1996, 1611-1631

Appendix 1: Uitgebreide toelichting data

In het onderzoek wordt gebruik gemaakt van een dataset van call opties op futures van de S&P index in het jaar 2000. De data zijn afkomstig van de Chicago Mercantile Exchange en bevatten van elke handelsdag¹⁴ van het jaar 2000 de volgende variabelen:

- Dag, maand en jaar waarop de optie wordt aangeboden.
 - Maand en jaar waarop de optie afloopt.
 - Bijbehorende strike price (uitoefenprijs) per optie.
 - Settle prijs van de optie: ofwel de slotprijs van de optie op de bijbehorende dag.
 - Volume van de optie: ofwel het aantal keer dat in de optie gehandeld is op bijbehorende dag.
- De afloopdatum van de optie is de derde vrijdag van de maand waarin de optie afloopt, zoals beschreven in de regels van de Chicago Board Options Exchange¹⁵. Op basis van deze gegevens wordt per optie de maturity time bepaald (T), weergegeven in het aantal handelsdagen. Naast de gegevens van de opties wordt ook gebruik gemaakt van de gegevens van de onderliggende futures en de risicovrije rentevoet van alle handelsdagen in het jaar 2000.

Een specifiek voordeel van opties op de S&P 500 index is dat er relatief veel in gehandeld wordt. Het voordeel van futures opties is dat er geen dividend wordt uitbetaald, wat de analyse vereenvoudigt.

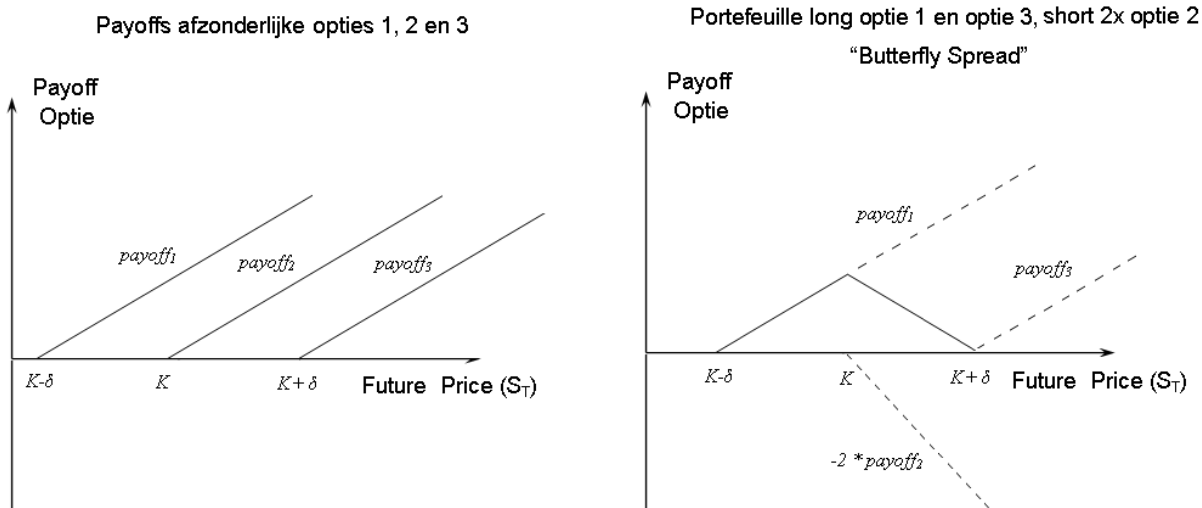
Een nadeel van markdata is dat er vaak niet synchroon wordt gehandeld. Dit kan bijvoorbeeld tot gevolg hebben dat $c_2 > 0.5(c_1 + c_3)$, waarbij c_1 , c_2 en c_3 prijzen zijn van Europese call opties met strike prices van resp. $K-\delta$, K en $K+\delta$ met dezelfde afloopdatum. Theoretisch gezien is dit onmogelijk, omdat dit arbitragemogelijkheden met zich meebrengt. Dit argument wordt in figuur 7 duidelijk gemaakt. In de linkerfiguur staat een grafische weergave van de payoffs van 3 afzonderlijke opties met prijzen c_1 , c_2 , c_3 en bovenstaande strike prices. In de rechterfiguur is een zogenaamde *butterfly spread* samengesteld zoals beschreven in Hull (2004). Deze spread callopties bestaat uit een portefeuille van 2 long opties met prijzen c_2 , c_3 en 2 short opties met prijs c_2 . Als je in de markt minder hoeft te betalen voor $c_1 + c_3$ dan voor $2 * c_2$ ¹⁶, dan brengt de *butterfly spread* arbitrage met zich mee. De payoff functie uit de rechter grafiek schuift naar boven, zodat structurele winsten worden behaald. De totale winst bestaat dan uit twee elementen: het verschil tussen de opbrengsten van de verkoop van 2 opties met prijs c_2 en de kosten van de aankoop van 2 opties met prijs c_1 en c_3 de payoff (af te lezen uit de rechter figuur)

In de data van dit onderzoek komt het voor dat $c_2 > 0.5(c_1 + c_3)$. Dit wordt mede veroorzaakt doordat we werken met slotprijzen van de onderliggende futures. Hierdoor zitten er af en toe vreemde waarnemingen in de dataset, zoals een optie settle-prijs van 0. Dit betekent uiteraard niet dat er tegen prijs 0 gehandeld is, maar uitsluitend dat er aan het eind van de handelsdag geen 'asks' geweest zijn. De dataset is daarom uitgebreid gefilterd op dit soort waarnemingen.

¹⁴ Handelsdagen zijn de dagen waarop de beurs geopend is. De beurs is dicht in het weekend en tijdens officiële feestdagen.

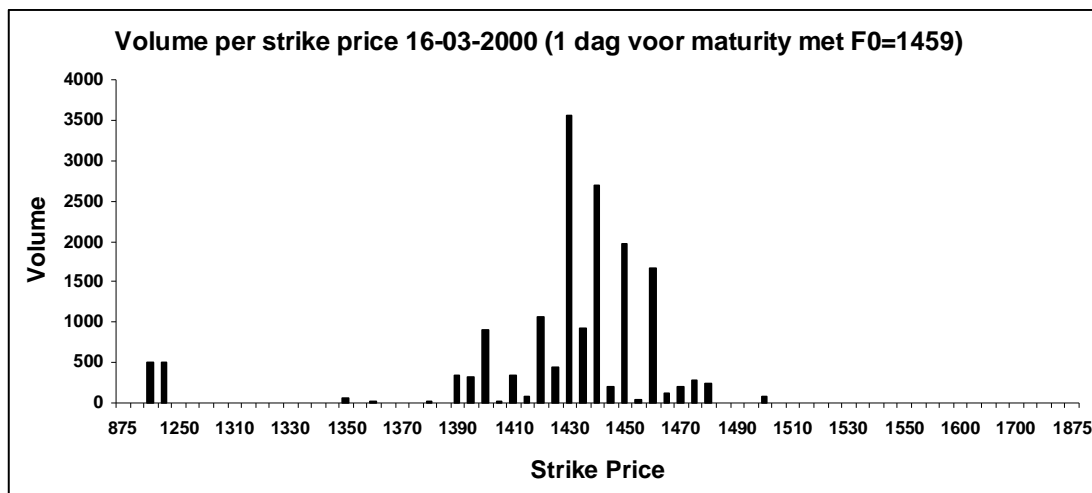
¹⁵ www.cboe.com

¹⁶ M.a.w. $c_2 > 0.5(c_1 + c_3)$



Figuur 7 Links: payoffs van 3 afzonderlijke callopties met strike prices $K-\delta$, K en $K+\delta$
 Rechts: payoffs van een "Butterfly spread" zoals vermeld in Hull (2004)

Het valt op dat dit verschijnsel het meest voorkomt bij opties waarin niet tot nauwelijks wordt gehandeld. Doordat de marktprijs van een optie tot stand komt door middel van vraag en aanbod, zijn opties efficiënter geprijsd naarmate er meer in wordt gehandeld. Vandaar dat marktprijzen van opties met een relatief laag volume minder betrouwbaar worden geacht dan opties met een hoog volume. Opties met een volume van 0 worden daarom in dit onderzoek buiten beschouwing gelaten. Zodoende wordt een groot gedeelte van de dataset verwijderd, aangezien het patroon uit figuur 8 in de praktijk relatief veel voor komt. Wat ook opvalt uit figuur 8 is dat het volume van relatief lage of juist hoge strike prices vaak erg laag is. Het zwaartepunt van het volume ligt normaliter rond de waarde van F_0 .



Figuur 8 Volume per strike price 16-03-2000 (1 dag voor maturity met $F_0 = 1459$)

Appendix 2 Methoden: Monte-Carlo simulatie

Voor het prijzen van opties op futures wordt in dit onderzoek gebruik gemaakt van Monte Carlo simulatie, zoals onder andere beschreven in Boyle (1977). De optieprijs wordt berekend door het genereren van *pseudo*- random getallen voor de prijs van de onderliggende future. Een groot voordeel van Monte Carlo simulatie is dat er geen beperking is tot modellen waarvoor analytisch een optieprijs afgeleid wordt; men kan dus elke gewenste verdeling met Monte Carlo simulatie modelleren. Er wordt een x aantal tijdreeksen van rendementen gesimuleerd, waaruit de futureprijzen (F_T) op de uitoefendatum T berekend kunnen worden.

Nu hebben we x trekkingen \tilde{F}_T waarbij;

$$\tilde{F}_T = \exp\left(\ln(F_0) + \sum_{t=1}^T r_t\right) \quad (1)$$

Met r_t uit een rendementgenererend proces. De bijbehorende kansverdelingen die in dit onderzoek worden gebruikt zijn de:

Normale verdeling $r_t \sim N\left(0, \frac{\sigma^2}{252}\right),$ (2)

(Student's) t-verdeling $r_t \sim t\left(0, \frac{\sigma^2}{252} \frac{v-2}{v}, v\right),$ ¹⁷ (3)

Scheve t-verdeling $r_t \sim t\left(0, \frac{\sigma^2}{252} \frac{v-2}{v}, v, \gamma\right),$ (4)

Uiteindelijk moet \tilde{F}_T nog met een factor worden vermenigvuldigd om de drift van het proces op te nemen, zodat de uiteindelijke future prijs op de afloopdatum wordt:

$$F_T = \tilde{F}_T \left(\frac{F_0}{\text{gemiddelde}(\tilde{F}_T)} \right) \quad (5)$$

Waarbij $\text{gemiddelde}(\tilde{F}_T)$ het gemiddelde van de gesimuleerde trekkingen \tilde{F}_T is.

De achterliggende gedachte achter deze factor is dat onder de aanname van *risiconutraliteit* de gemiddelde futures prijs op de uitoefendatum gelijk moet zijn aan F_T . Dit heeft als oorzaak dat bij het aangaan van een future contract geen betaling plaatsvindt, maar pas op een gespecificeerd moment in de toekomst. Onder de *risico-neutrale* aanname is er dus geen verwachte betaling in de toekomst. De factor is een elegante manier om de verdelingen impliciet met een zodanige drift μ te specificeren dat de uiteindelijke gemiddelde gesimuleerde futures prijs F_0 is.

De gesimuleerde futures prijzen op uitoefendatum F_T worden dan gebruikt om de optieprijs \hat{c} behorende bij het simulatie-model te bepalen:

$$\hat{c} = e^{-rT} \times \text{gemiddelde}(\text{payoff}) \quad (6)$$

met $\text{payoff} = \max(F_T - \text{Strike}, 0)$ en met r de *risicovrije* interest, omdat de *risiconeutrale* verdeling van het onderliggende proces wordt verondersteld.

¹⁷ De variantie van een 'standaard' t-verdeling met v vrijheidsgraden is $v/(v-2)$ (zolang $v > 2$), zodat de variantie σ^2 met $v/(v-2)$ vermenigvuldigd moet worden. Door met $v/(v-2)$ te vermenigvuldigen hangt alleen de kurtosis van v af en niet de variantie en is σ^2 een directe interpretatie van de variantie.

Appendix 3 Gewichten analyse

In de formule die gewichten genereert voor stap 2 uit het stappenplan voor Slim Schatten (hoofdstuk 2.1) zijn 2 optie-eigenschappen van belang: de *ligging* qua strike price en het *volume*. De formule moet hogere waarden toekennen aan “omliggende” opties qua strike price en opties met een hoog volume. Voor een beeldvorming van deze eigenschappen, staan het volume en de strike price van alle opties van 8 december 2000 (T=5) samengevat in tabel 9.

De volgende kwadratische formule voldoet aan beide bovenstaande criteria:

$$gewicht_{ij}^* = \frac{1}{(K_{ij} - K_j)^2} * \log(1 + volume_i) \quad \text{met } i, j = 1, \dots \# \text{ Strike Prices} \quad (1)$$

Hierbij geldt de voorwaarde $i \neq j$. In het voorbeeld uit tabel 9 worden zodoende per optie i 28 gewichten toegekend aan de overige opties. De som van de gewichten per optie i is echter (meestal) niet gelijk aan 1. Uit praktisch oogpunt is het daarom handig om onderstaande formule te hanteren waarin de gewichten *genormaliseerd* worden:

$$gewicht_{ij} = \frac{gewicht_{ij}^*}{\sum_{i=1}^{28} gewicht_{ij}^*} \quad \text{met } i, j = 1, \dots \# \text{ Strike Prices} \quad (2)$$

Op basis van formule (2) kan een matrix worden gevormd met $i \times j$ gewichten. In tabel 10 staat het voorbeeld uit tabel 9 uitgewerkt.

Het is opvallend dat veel waarden in tabel 10 buiten de diagonaal de waarde 0 aannemen. Zodoende worden de schattingen in het model uit hoofdstuk 2.1 voornamelijk gebaseerd op *omliggende* opties. Toch speelt ook het volume een belangrijke rol. Bij de keuze tussen de 2 “dichtstbijzijnde” opties wordt namelijk de nadruk gelegd op de optie met het hoogste volume. Aangezien opties met een hoog volume mogelijk efficiënter zijn geprijsd (gevolg van vraag en aanbod in de optiemarkt, zie appendix 1), kan dit helpen in de zoektocht naar geïmpliceerde parameters uit hoofdstuk 2. In die zin is dit model *geavanceerder* te noemen dan het model dan het model gepresenteerd in Appendix 4, waarin aan de 2 omringende opties direct een gewicht van 1 wordt toegekend.

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
Vol.	5	2	5	3	1	1	48	16	37	6	39	99	504	77	35	1363	157	179	37	77	982	20	11	13	1	69	41	1	27
K	1300	1320	1325	1330	1340	1345	1350	1355	1360	1365	1370	1375	1380	1390	1395	1400	1405	1410	1415	1420	1425	1430	1435	1440	1445	1450	1460	1470	1475

Tabel 9 voorbeeld van strike prices met bijbehorend volume van opties op 8 december 2000

i\j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	
1	x	0,14	0,15	0,08	0,02	0,02	0,08	0,05	0,05	0,02	0,04	0,04	0,05	0,03	0,02	0,04	0,02	0,02	0,01	0,02	0,02	0,01	0,01	0,01	0,00	0,01	0,01	0,00	0,01	
2	0,04	x	0,63	0,12	0,02	0,01	0,04	0,02	0,02	0,01	0,01	0,01	0,02	0,01	0,01	0,01	0,01	0,01	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
3	0,02	0,33	x	0,41	0,02	0,01	0,05	0,02	0,02	0,01	0,01	0,01	0,02	0,01	0,01	0,01	0,01	0,01	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
4	0,02	0,08	0,55	x	0,05	0,02	0,07	0,03	0,03	0,01	0,02	0,02	0,02	0,01	0,01	0,01	0,01	0,01	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
5	0,01	0,02	0,06	0,10	x	0,20	0,28	0,09	0,06	0,02	0,03	0,03	0,03	0,01	0,01	0,01	0,01	0,01	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
6	0,00	0,01	0,02	0,02	0,10	x	0,57	0,10	0,06	0,02	0,02	0,02	0,02	0,01	0,01	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
7	0,00	0,01	0,01	0,01	0,03	0,12	x	0,47	0,15	0,04	0,04	0,03	0,03	0,01	0,01	0,01	0,01	0,01	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
8	0,00	0,00	0,01	0,01	0,01	0,02	0,40	x	0,37	0,05	0,04	0,03	0,03	0,01	0,01	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
9	0,00	0,00	0,00	0,00	0,01	0,01	0,12	0,34	x	0,23	0,11	0,06	0,05	0,01	0,01	0,01	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
10	0,00	0,00	0,00	0,00	0,00	0,00	0,04	0,06	0,32	x	0,33	0,10	0,06	0,02	0,01	0,01	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
11	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,03	0,08	0,18	x	0,43	0,15	0,03	0,01	0,02	0,01	0,01	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
12	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,01	0,03	0,04	0,29	x	0,49	0,04	0,02	0,02	0,01	0,01	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
13	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,01	0,03	0,02	0,10	0,52	x	0,12	0,04	0,05	0,02	0,02	0,01	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
14	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,01	0,01	0,01	0,02	0,05	0,16	x	0,38	0,19	0,06	0,03	0,02	0,01	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
15	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,02	0,04	0,28	x	0,46	0,08	0,04	0,01	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
16	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,01	0,03	0,08	0,27	x	0,39	0,10	0,03	0,02	0,02	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
17	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,02	0,03	0,05	0,44	x	0,31	0,05	0,03	0,03	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
18	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,01	0,02	0,03	0,13	0,36	x	0,26	0,08	0,05	0,01	0,01	0,01	0,00	0,00	0,00	0,00	0,00	
19	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,01	0,02	0,05	0,09	0,35	x	0,29	0,12	0,02	0,01	0,01	0,00	0,01	0,00	0,00	0,00	
20	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,01	0,01	0,03	0,04	0,09	0,25	x	0,46	0,05	0,02	0,01	0,00	0,01	0,00	0,00	0,00	
21	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,01	0,01	0,03	0,03	0,05	0,08	0,39	x	0,27	0,06	0,03	0,00	0,02	0,01	0,00	0,00	
22	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,01	0,02	0,02	0,02	0,02	0,03	0,08	0,53	x	0,19	0,05	0,01	0,02	0,01	0,00	0,00
23	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,01	0,01	0,02	0,01	0,02	0,02	0,05	0,18	0,31	x	0,27	0,02	0,05	0,02	0,00	0,01	
24	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,01	0,01	0,02	0,01	0,02	0,02	0,04	0,11	0,11	0,35	x	0,10	0,15	0,03	0,00	0,01	
25	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,01	0,01	0,01	0,02	0,05	0,04	0,06	0,28	x	0,44	0,04	0,00	0,01	
26	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,01	0,01	0,01	0,02	0,02	0,02	0,02	0,03	0,07	0,05	0,07	0,17	0,18	x	0,25	0,01	0,04	
27	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,01	0,01	0,01	0,02	0,02	0,02	0,02	0,03	0,05	0,03	0,04	0,06	0,03	0,41	x	0,07	0,14	
28	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,01	0,01	0,01	0,01	0,01	0,02	0,01	0,01	0,01	0,01	0,01	0,05	0,18	x	0,65
29	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,01	0,01	0,01	0,02	0,01	0,02	0,01	0,02	0,04	0,02	0,02	0,03	0,01	0,10	0,24	0,40	x	

Tabel 10 Voorbeeld van gewichten op basis van formule (2) voor 8 december 2000 (T=5)

Appendix 4 Slim schatten: Omringende Opties

In grote lijnen kan het stappenplan gepresenteerd in Hoofdstuk 2.1 hergebruikt worden voor het model op basis van omringende opties. Het grootste verschil hierbij is stap 2, waarin de geïmpliceerde volatiliteiten van het Black Scholes model geschat worden. Deze stap moet worden vervangen voor:

Stap 2: σ schatten (Black-Scholes)

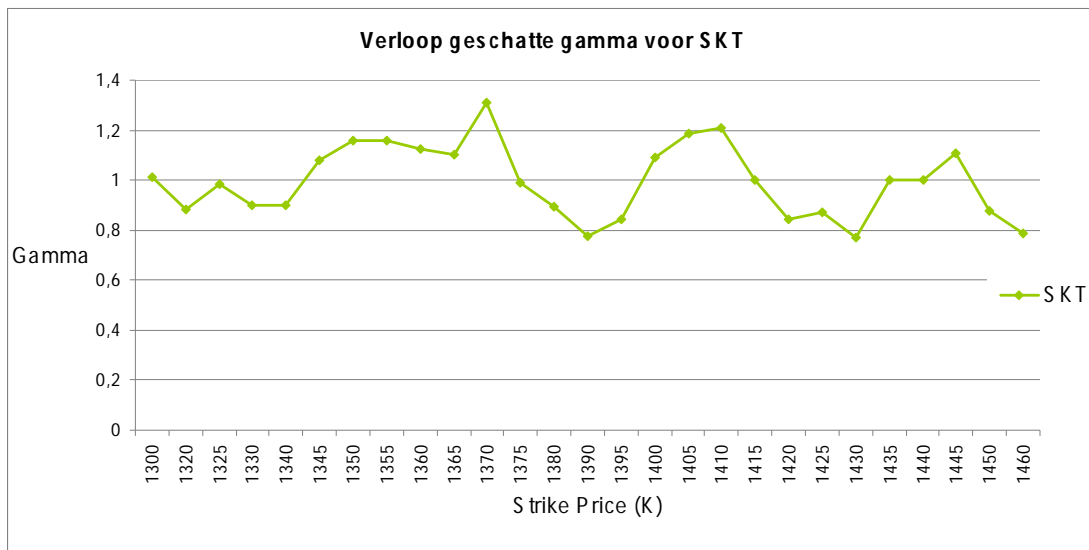
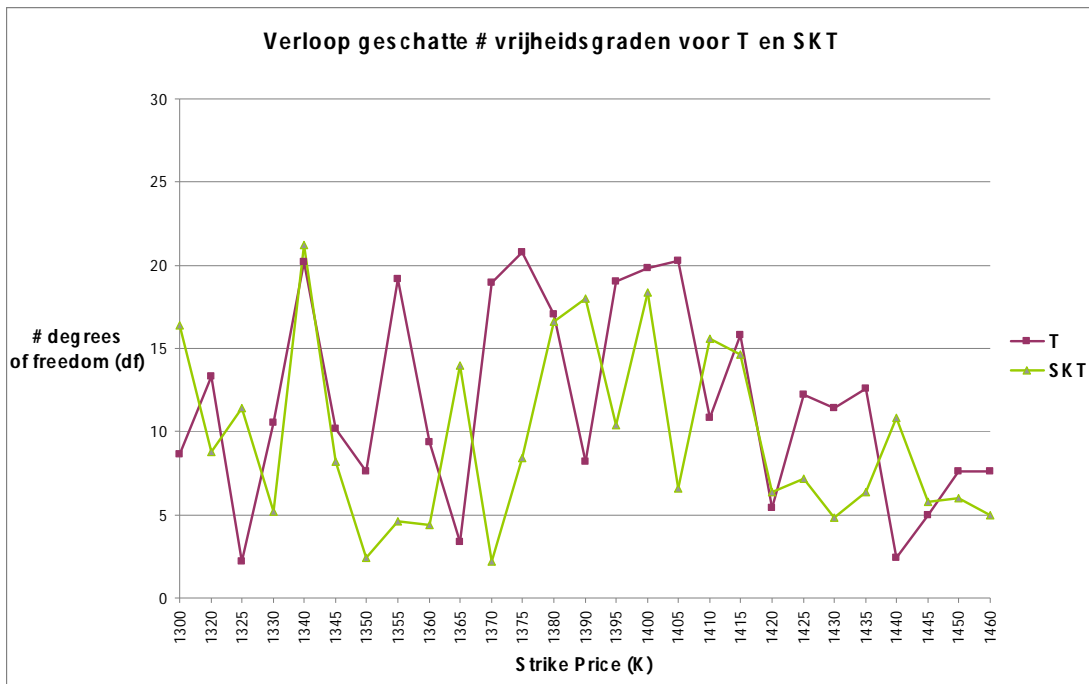
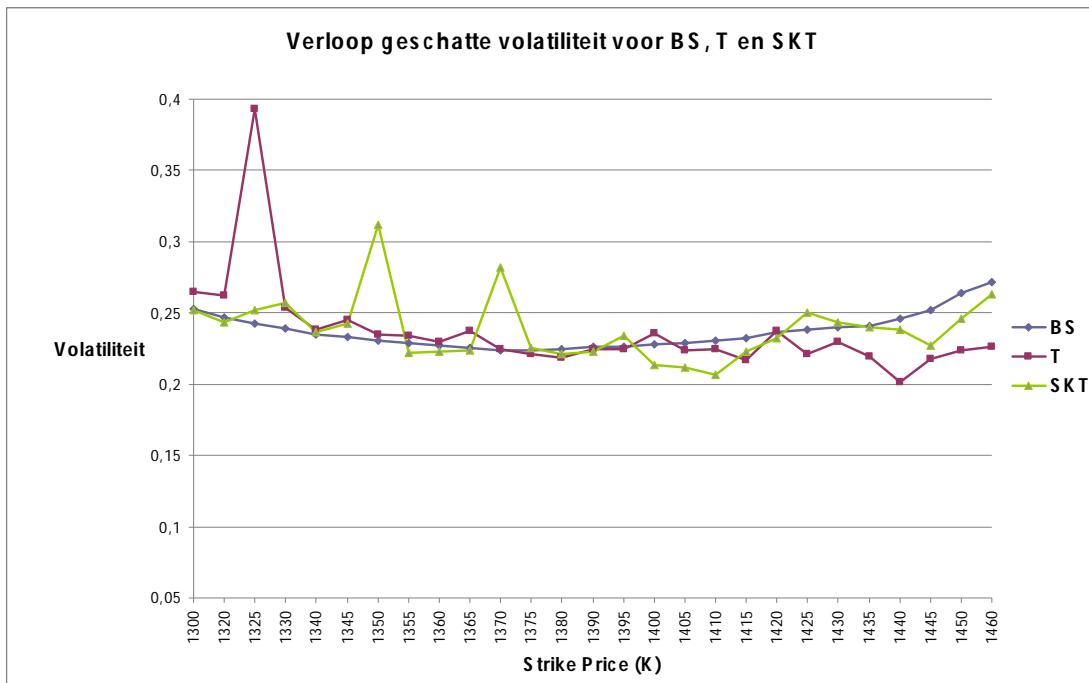
σ apart schatten door de eigen geïmpliceerde volatiliteit en de 2 omliggende (qua strike price) te middelen. Op deze manier worden 3 waarnemingen gebruikt om 1 parameter te schatten (σ).

Aangezien elke schatting 2 omliggende waarnemingen nodig heeft, kunnen de parameters van n-2 opties geschat worden. Op 8 december worden 29 opties geëvalueerd, zodat $29-2 = 27$ schattingen uitgevoerd kunnen worden; de laagste en hoogste uitoefenprijs vallen in die zin weg. De resultaten van het stappenplan kunnen worden weergegeven tabel 11 (analoog aan tabel 1). Hierin staat welke parameters worden geschat, welke numerieke methoden zijn gebruikt, wat de rekentijd van de methode is en de bijbehorende waarde van SSR (als som van de 27 afzonderlijke SSR-waarden). Bijbehorende optimale parameters uit stap 2 (BS, Black-Scholes), stap 3 (T, T-verdeling) en stap 4 (SKT, scheve T-verdeling) staan weergegeven in figuur 5, uitgezet tegen de uitoefenprijzen.

Uit de tabel kan worden afgelezen dat het eindresultaat van deze schattingsmethode een SSR is van $8.42E-14$, met een rekentijd van ca. 15 minuten (880.6s). De resultaten uit de tabel zijn vergelijkbaar met de resultaten gepresenteerd in Hoofdstuk 2.2. De grafieken in figuur 9 wijken enigszins af van de grafieken in figuur 5, maar de belangrijkste conclusies blijven in stand: de *volatility smile* wordt niet "rechtgetrokken" bij het toevoegen van parameters, en de geïmpliceerde df 's en γ 's vertonen ruis rondom een hun gemiddelde waarden en liggen niet voor elke uitoefenprijs onder de waarden behorende bij de normale verdeling.

Stap	Wat wordt geschat?	Welke numerieke methode wordt gebruikt?	Rekentijd	SSR
2	27 x σ 27 x σ	- <i>fminsearch</i> (1 variabele)	0.1 s 1.9 s	5.911 4.70E-05
3	100 x 27 x df 27 x σ en df gezamenlijk	100-staps df -grid en <i>fminsearch</i> (1 variabele) <i>fminsearch</i> (2 variabelen)	217.3 s 10.3 s	4.54E-05 5.41E-11
4a	100 x 27 x γ	100-staps γ -grid	5.3 s	0.003
4b	100 x 27 x df 27 x σ , γ en df gezamenlijk	100 staps df -grid en <i>fminsearch</i> (2 variabelen) <i>fminsearch</i> (3 variabelen)	627.4 s 20.1 s	3.30E-13 8.42E-14
Eind			882.4 s	8.42E-14

Tabel 11 Samenvatting resultaten van stap 2 t/m 4 van het "Slim Schatten" model o.b.v. omringde opties



Figuur 6 Optimale parameter waarden van de procedure van *omliggende* opties (qua uitoefenprijs). Van boven naar beneden: de geschatte σ , df en γ van de verschillende verdelingen (BlackScholes, T en scheve T), uitgezet tegen de uitoefenprijs (K).

Appendix 5 Slim schatten: Fijnere schatting γ

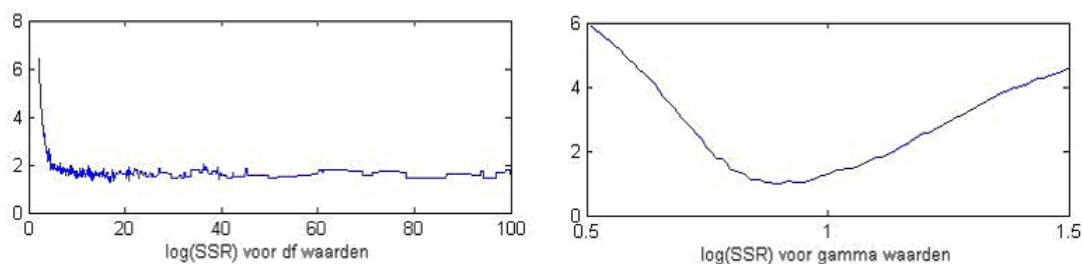
Het model uit Hoofdstuk 2.1 kan uitgebreid worden door voor elke waarde in het γ -grid een optimalisatieprocedure uit te voeren waarin σ wordt geoptimaliseerd (gegeven de *gridwaarde* voor γ). Doelfunctie van deze optimalisatieprocedure (in Matlab wordt de ingebouwde functie *fminsearch* gebruikt) is het vinden van een optieprijs die zo min mogelijk afwijkt van de werkelijke optieprijs (gemeten met behulp van SSR). De resultaten voor 8 december 2000 worden in onderstaande tabel samengevat:

Stap	Wat wordt geschat?	Welke numerieke methode wordt gebruikt?	Rekentijd	SSR
2	29 x σ	-	0.2 s	4.156
	29 x σ	<i>fminsearch</i> (1 variabele)	2.4 s	4.26E-05
3	100 x 29 x <i>df</i>	100-staps <i>df</i> -grid en <i>fminsearch</i> (1 variabele)	222.5 s	3.99E-05
	29 x σ en <i>df</i> gezamenlijk	<i>fminsearch</i> (2 variabelen)	11.4 s	2.92E-10
4a	100 x 29 x γ	100-staps γ -grid en <i>fminsearch</i> (1 variabele)	139.3 s	2.07E-08
4b	100 x 29 x <i>df</i>	100-staps <i>df</i> -grid en <i>fminsearch</i> (2 variabelen)	701.7 s	1.40E-12
	29 x σ , γ en <i>df</i> gezamenlijk	<i>fminsearch</i> (3 variabelen)	15.5 s	1.69E-13
Eind			1093.0 s	1.86E-13

Tabel 12 Samenvatting resultaten van stap 2 t/m 4 van het "Slim Schatten" model o.b.v. gewichten met fijnere schatting voor γ in stap 4a

Een vergelijking van tabel 1 uit hoofdstuk 2.2 en tabel 12 leert dat een fijnere schatting voor γ -waarden niet direct tot betere eind-resultaten leidt. De SSR van stap 4a is weliswaar verbeterd, maar het eindresultaat (1.86E-13) is er *niet* beter op geworden. Om deze reden is ervoor gekozen om de schatting van γ oppervlakkig te houden (zonder extra tijdrovende optimalisatieprocedure).

Hoe is het mogelijk dat een fijnere schatting van γ niet direct tot een beter eindresultaat leidt? Een mogelijke verklaring hiervoor is dat de *df* in de zoektocht naar een *globaal* minimum een grotere "boosdoener" is dan γ , zoals besproken in het werkcollege. Dit is te begrijpen aan de hand van figuur 7 waarin de logaritme van de SSR-waarde wordt gegeven in het optimalisatieproces voor een grid van waarden van *df* (links) en γ (rechts). Hieruit blijkt dat bij *df* de kans groter is dat je in een *lokaal* minimum terecht komt dan bij γ . De gezamenlijke *fminsearch* procedure in stap 4b is niet altijd in staat om de optimale parameters uit het (lokale) minimum van stap 4a te "trekken". Bovendien is de doelstelling van stap 4a om *beginwaarden* te creëren voor γ in het gezamenlijke optimalisatieproces van stap 4b. Uit figuur 7 blijkt dat een relatief oppervlakkige schatting voor γ op basis van een slim gekozen grid vaak al aardig in de richting komt van het optimum, zodat (voor een geschikte beginwaarde van γ) een nauwkeurige schatting overbodig is.



Figuur 7 Log SSR voor grid van *df* waarden (links) en een grid van γ waarden (rechts)

Appendix 6 Slim schatten voor andere dagen

Hieronder volgen de resultaten van 15 maart, met $T=2$, $F_0=\$1395,20$ en 25 uitoefenprijzen variërend van $\$1300,00$ tot $\$1480,00$. In onderstaande tabellen en figuren (tabel 13 en figuur 8) staan links de resultaten op basis van de *gewichten* procedure, en rechts de resultaten op basis van *omliggende opties*

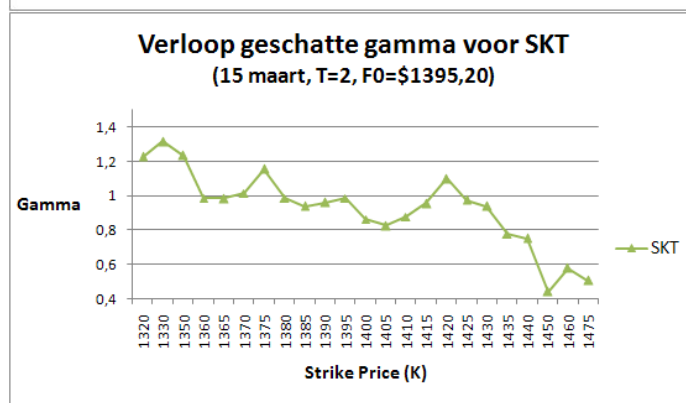
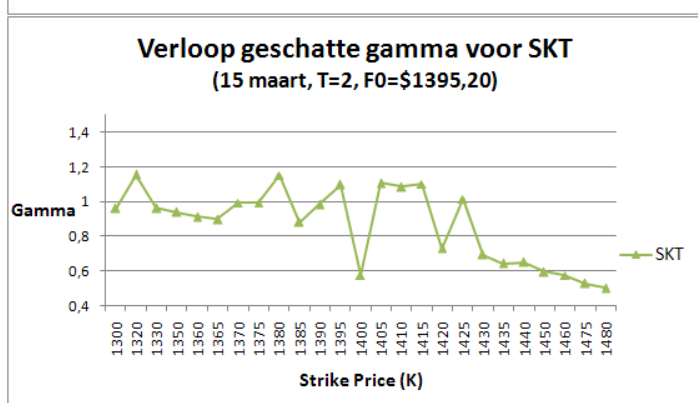
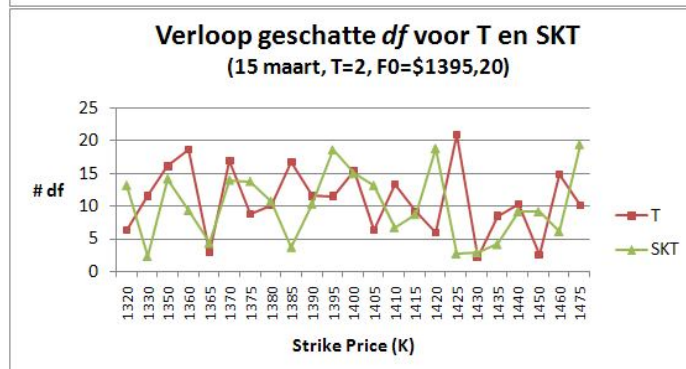
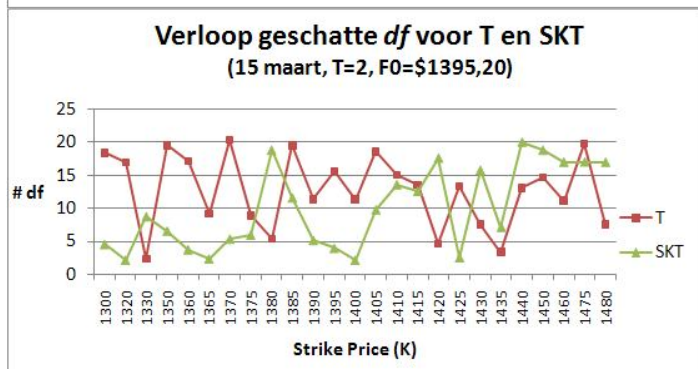
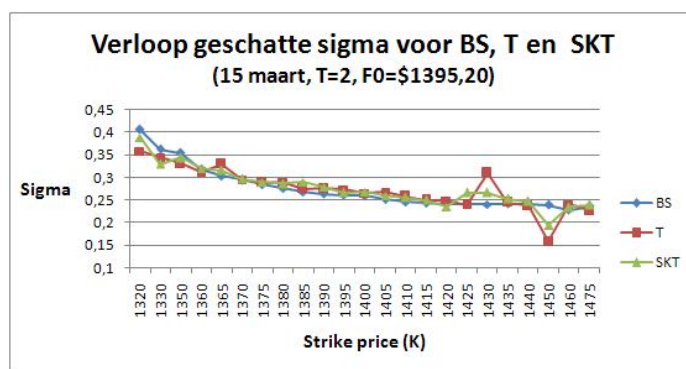
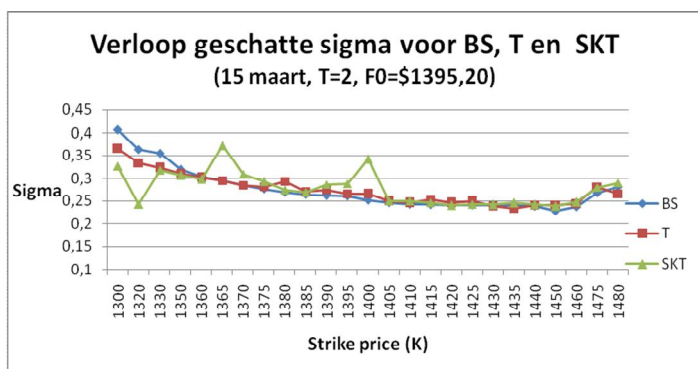
GEWICHTEN

Stap	Rekentijd	SSR
2	0.1 s	1.057
	0.7 s	1.32E-05
3	105.4 s	1.64E-07
	5.3 s	1.39E-13
4a	2.5 s	0.002
4b	321.3 s	8.14E-15
	11.9 s	1.32E-16
Eind	447.2 s	1.32E-16

OMLIGGENDE OPTIES

Stap	Rekentijd	SSR
2	0.1 s	1.640
	2.7 s	1.94E-05
3	89.0 s	2.63E-07
	5.1 s	9.56E-13
4a	2.2 s	0.004
4b	294.6 s	1.28E-14
	6.6 s	2.89E-16
Eind	400.5 s	2.89E-16

Tabel 13: SSR en rekentijd per stap in de procedure o.b.v. gewichten (links) en omliggende opties (rechts).



Figuur 1 Verloop van optimale parameters σ (bovenaan, voor de verdelingen BS, T en SKT), df (midden, voor T en SKT) en γ (onder, voor SKT). De 3 figuren links: optimale parameters na optimalisatie via de *gewichten* procedure. De 3 figuren rechts: de optimale parameters na door de procedure o.b.v. *omliggende opties*.

Hieronder volgen de resultaten van 2 juni, met $T=10$, $F_0=\$1476,00$ en 35 uitoefenprijzen variërend van $\$1350,00$ tot $\$1650,00$. In tabel 14 en figuur 9 staan links de resultaten op basis van de *gewichten* procedure, en rechts de resultaten op basis van *omliggende opties*

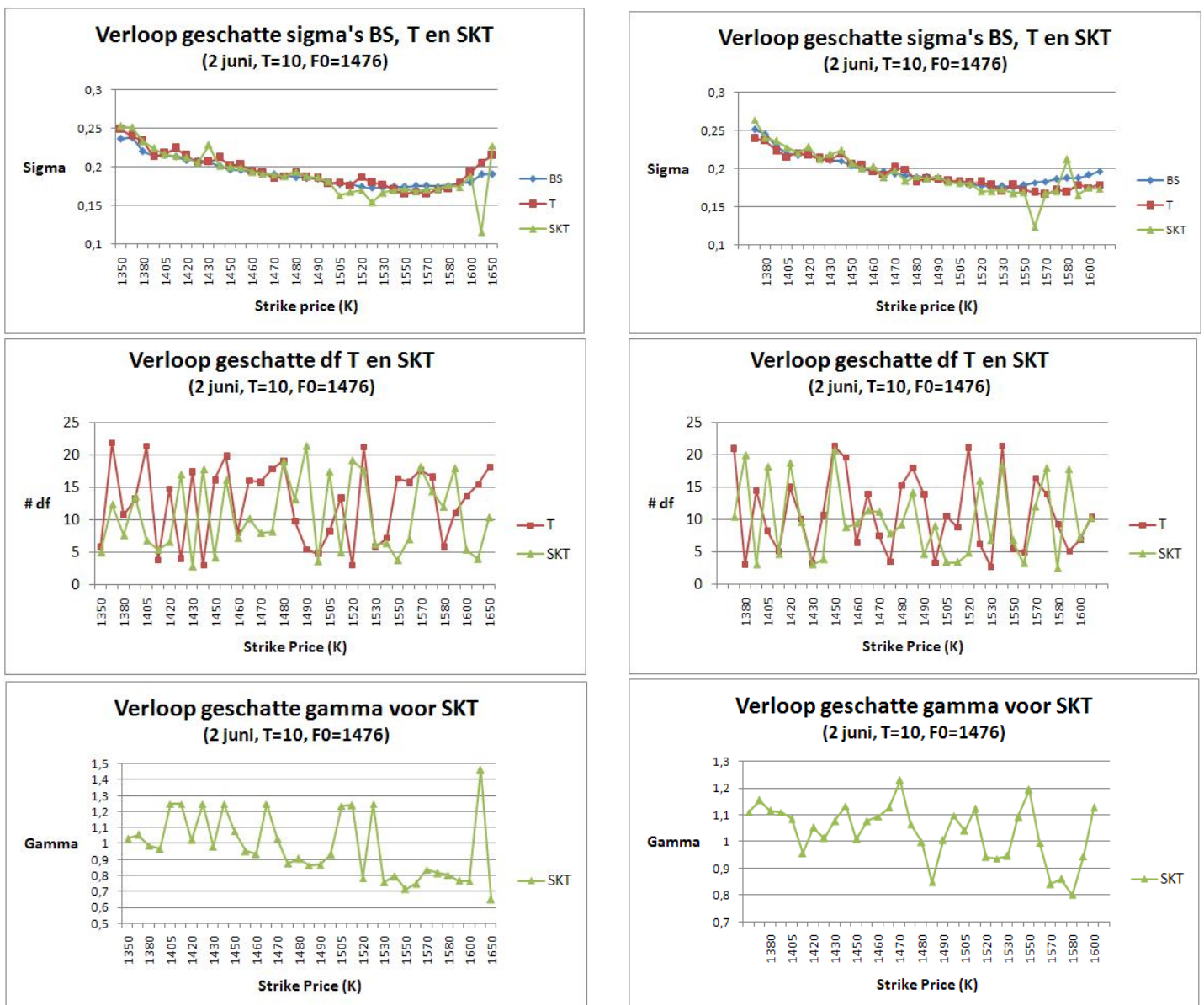
GEWICHTEN

Stap	Rekentijd	SSR
2	0.4 s	4.156
	5.5 s	1.24E-04
3	509.8 s	1.01E-05
	25.3 s	3.66E-11
4a	14.1 s	0.006
4b	1969.0 s	2.70E-11
	59.9 s	6.05E-12
Eind	2584.6 s	6.05E-12

OMLIGGENDE OPTIES

Stap	Rekentijd	SSR
2	0.2 s	4.589
	2.8 s	1.09E-04
3	490.8 s	1.64E-05
	24.1 s	2.87E-10
4a	12.8 s	0.020
4b	1453.3 s	1.73E-11
	16.1 s	1.83E-12
Eind	2000.1 s	1.83E-12

Tabel 14: SSR en rekestijd per stap in de procedure o.b.v. gewichten (links) en omliggende opties (rechts).



Figuur 9 Verloop van optimale parameters σ (bovenaan, voor de verdelingen BS, T en SKT), df (midden, voor T en SKT) en γ (onder, voor SKT). De 3 figuren links: optimale parameters na optimalisatie via de *gewichten* procedure. De 3 figuren rechts: de optimale parameters na door de procedure o.b.v. *omliggende opties*.

Appendix 7 Pseudocode Delta-Hedging

Hieronder wordt de *pseudocode* van de Delta-Hedging procedure uit Hoofdstuk 3.2 weergegeven. De code levert als output de tracking error van zowel het BS- als het SKT-model voor 1 specifieke optie (met afloopdatum T en strike price K_{target}).

De code gebruikt 3 andere functies:

- *IMPLIED.m* berekent de geïmpliceerde parameters van BS en SKT gegeven de optiedata en K_{target}
- *optieprijzen_BS* berekent de optieprijs van BS gegeven de volatiliteit, F_0 en T
- *optieprijzen_SKT* berekent de optieprijs gegeven de 3 SKT-parameters, F_0 en T

function DeltaHedging

Input: [data_input ; T; K_target]

Output: [TrackingErrorBS; TrackingErrorSKT]

% Het definiëren van de vectoren die later in het programma gevuld gaan worden.

DeltaSKT = zeros(T,1);

WaardeF_SKT = zeros(T,1);

WaardeO_SKT = zeros(T,1);

TotaleWaarde_SKT = zeros(T,1);

DeltaBS = zeros(T,1);

WaardeF_BS = zeros(T,1);

WaardeO_BS = zeros(T,1);

TotaleWaarde_BS = zeros(T,1);

t_vector = zeros(T,1);

F = zeros(T,1);

r = zeros(T,1);

% De stapgrootte (tau) is 1 dag

tau = 1;

% Definieren data met $t=T$ zodat initiële positie bepaald kan worden

[row,col]= find(data_input(:,13) == T);

data = data_input(row,:);

% Uit deze data kunnen de benodigde scalars en vectoren gedefinieerd worden:

[K_number,col] = find(data(:,5) == K_target);

c_vector = data(:,6);

Fnu_scalar = data(1,8);

F(1) = data(1,9);

r(1) = data(1,10);

t_vector(1) = T;

% o.g.v. data en K_{target} worden de optimale geïmpliceerde parameters geschat:

[implied_vol implied_sigma_gamma_df] = IMPLIED (data, K_target)

% op basis van deze geïmpliceerde parameters worden de optieprijzen SKT1 en SKT2 berekend,

% welke nodig zijn voor het schatten van DeltaSKT.

SKT1 = optieprijzen_SKT (implied_sigma_gamma_df, (F(1)+F(1)*implied_sigma_gamma_df(1)/sqrt(252)), T)

SKT2 = optieprijzen_SKT (implied_sigma_gamma_df, (F(1)-F(1)*implied_sigma_gamma_df(1)/sqrt(252)), T)

% Op basis van SKT1 en SKT2 wordt de initiële DeltaSKT gevormd. Deze komt op de eerste plek te

% staan van de vector DeltaSKT

DeltaSKT(1) = (SKT1-SKT2)/(2*F(1)*implied_sigma_gamma_df(1)/sqrt(252));

% DeltaBS wordt op analoge wijze bepaald, nu wordt gebruik gemaakt van implied_vol berekend uit % de IMPLIED functie.

BS1 = optieprijsen_BS (implied_vol, (F(1) + F(1)*implied_vol/sqrt(252)), T)

BS2 = optieprijsen_BS (implied_vol, (F(1) - F(1)*implied_vol/sqrt(252)), T)

DeltaBS(1) = (BS1-BS2)/(2*F(1)*implied_vol/sqrt(252));

% Op basis van de bovenstaande Delta's en de gedefinieerde vectoren en parameters wordt de % initiële positie bepaald van de delta-hedge portefeuille van BS en SKT

WaardeC = -c_vector(K_number);

WaardeF_SKT(1) = Fnul_scalar * DeltaSKT(1);

WaardeO_SKT(1) = -1*(WaardeC + WaardeF_SKT(1));

InitieleTotaleWaarde_SKT = WaardeC + WaardeO_SKT(1) + WaardeF_SKT(1); % gelijk aan 0

WaardeF_BS(1) = Fnul_scalar * DeltaBS(1);

WaardeO_BS(1) = -1*(WaardeC + WaardeF_BS(1));

InitieleTotaleWaarde_BS = WaardeC + WaardeO_BS(1) + WaardeF_BS(1); % gelijk aan 0

% Nu wordt een discrete hedge-strategie toegepast voor de overige T-1 dagen: per dag wordt eerst

% de data ge-update en vervolgens wordt de portefeuille dagelijks herzien:

for t = 2:T

TimetoMaturity = (T+1)-t;

% data update voor de nieuwe dag;

[row,col] = find(data_input(:,13) == TimetoMaturity);

data = data_input(row,:);

F(t) = data(1,8);

r(t) = data(1,9);

% nieuwe geïmpliceerde parameters en optieprijsen berekenen voor het scahtten van Delta [BlackScholes implied_vol SkewT implied_sigma_gamma_df] = IMPLIED (data, K_input);

SKT1 = optieprijsen_skewed_T_F0 (implied_sigma_gamma_df, ...
(F(t) + F(t)*implied_sigma_gamma_df(1)/sqrt(252)), TimetoMaturity);

SKT2 = optieprijsen_skewed_T_F0 (implied_sigma_gamma_df, ...
(F(t) - F(t)*implied_sigma_gamma_df(1)/sqrt(252)), TimetoMaturity);

DeltaSKT(t) = (SKT1-SKT2)/(2*F(t)*implied_sigma_gamma_df(1)/sqrt(252));

BS1 = optieprijsen_BS_F0 (implied_vol, (F(t) + F(t)*implied_vol/sqrt(252)), TimetoMaturity);

BS2 = optieprijsen_BS_F0 (implied_vol, (F(t) - F(t)*implied_vol/sqrt(252)), TimetoMaturity);

DeltaBS(t) = (BS1-BS2)/(2*F(t)*implied_vol/sqrt(252));

% Portefeuilles herzien op basis van de nieuwe Delta's en de update van de data:

WaardeF_SKT(t) = F(t) * DeltaSKT(t);

WaardeO_SKT(t) = exp(r(t)*tau/252)*WaardeO_SKT(t-tau) - (F(t)*(DeltaSKT(t) - DeltaSKT(t-tau)));

WaardeF_BS(t) = F(t) * DeltaBS(t);

WaardeO_BS(t) = exp(r(t)*tau/252)*WaardeO_BS(t-tau) - (F(t)*(DeltaBS(t) - DeltaBS(t-tau)));

% Totale waarden berekenen van de herziene portefeuilles:

TotaleWaarde_SKT(t) = WaardeC + WaardeO_SKT(t) + WaardeF_SKT(t);

TotaleWaarde_BS(t) = WaardeC + WaardeO_BS(t) + WaardeF_BS(t);

end

% Berkenen van de Tracking errors:

TrackingError_BS = exp(-r(T)*T/252) * abs(TotaleWaarde_BS(T))

TrackingError_SKT = exp(-r(T)*T/252) * abs(TotaleWaarde_SKT(T))

Appendix 8 Grid-analyse uitgebreide resultaten

In onderstaande tabellen staan de uitgebreide uitkomsten van de eerste *grid-analyse* uit Hoofdstuk 4.2, inclusief de bijbehorende impliciete parameters. In tabel 14 staat de controlegroep, eronder (tabel 14a t/m 14d) staan de uitkomsten van de 4 verfijnde *grids* in de database. In iedere tabel staan onder de kolommen 2 subkolommen: de linker beschrijft de uitkomst na stap 5a, de rechter beschrijft de uitkomst na stap 5b.

14) 40-15-26-16 (controlegroep) Rekestijd simulaties database: 243s.

Gridkeuze	Rekestijd		SSR		Sigma		DF		Gamma	
	5a	5b	5a	5b	5a	5b	5a	5b	5a	5b
30-52-32	29.39	213.60	0.172	0.130	0.230	0.232	3.500	3.541	0.725	0.737
15-52-32	10.90	301.53	0.211	0.130	0.240	0.232	3.250	3.541	0.750	0.737
30-26-32	11.16	119.15	0.172	0.130	0.230	0.232	3.500	3.541	0.725	0.737
30-52-16	10.45	300.54	0.211	0.130	0.240	0.232	3.250	3.541	0.750	0.737

14a) 80-15-26-16 (fijner grid voor F0/K) Rekestijd simulaties database: 425s.

Gridkeuze	Rekestijd		SSR		Sigma		DF		Gamma	
	5a	5b	5a	5b	5a	5b	5a	5b	5a	5b
30-52-32	11.38	22.52	0.241	0.095	0.230	0.234	3.750	3.587	0.725	0.740
15-52-32	6.10	24.56	0.250	0.095	0.240	0.234	3.250	3.587	0.750	0.740
30-26-32	6.05	18.13	0.286	0.095	0.230	0.234	4.000	3.587	0.725	0.740
30-52-16	6.10	25.56	0.250	0.095	0.230	0.234	3.250	3.587	0.750	0.740

14b) 40-30-26-16 (fijner grid voor sigma) Rekestijd simulaties database: 461s.

Gridkeuze	Rekestijd		SSR		Sigma		DF		Gamma	
	5a	5b	5a	5b	5a	5b	5a	5b	5a	5b
30-52-32	21.86	351.88	0.184	0.129	0.230	0.232	3.500	3.527	0.725	0.739
15-52-32	11.53	484.88	0.211	0.129	0.240	0.232	3.250	3.527	0.750	0.739
30-26-32	10.27	343.21	0.184	0.129	0.230	0.232	3.500	3.527	0.725	0.739
30-52-16	10.63	487.87	0.211	0.129	0.240	0.232	3.250	3.527	0.750	0.739

14c) 40-15-52-16 (fijner grid voor DF) Rekestijd simulaties database: 451s.

Gridkeuze	Rekestijd		SSR		Sigma		DF		Gamma	
	5a	5b	5a	5b	5a	5b	5a	5b	5a	5b
30-52-32	21.33	463.47	0.172	0.115	0.230	0.228	3.500	3.635	0.725	0.733
15-52-32	11.44	679.53	0.333	0.129	0.220	0.232	3.750	3.439	0.700	0.731
30-26-32	10.31	452.08	0.172	0.115	0.230	0.228	3.500	3.635	0.725	0.733
30-52-16	10.75	397.09	0.302	0.129	0.230	0.232	3.250	3.439	0.700	0.731

14d) 40-15-26-32 (fijner grid gamma) Rekestijd simulaties database: 420s.

Gridkeuze	Rekestijd		SSR		Sigma		DF		Gamma	
	5a	5b	5a	5b	5a	5b	5a	5b	5a	5b
30-52-32	21.09	331.26	0.151	0.122	0.230	0.232	3.500	3.550	0.725	0.733
15-52-32	12.30	557.13	0.211	0.122	0.240	0.232	3.250	3.550	0.750	0.733
30-26-32	11.10	310.81	0.151	0.122	0.230	0.232	3.500	3.550	0.725	0.733
30-52-16	11,19	558.84	0.211	0.122	0.240	0.232	3.250	3.550	0.750	0.733

De uitgebreide resultaten van het vervolgonderzoek met *nog* fijnere grids voor F_0/K staan in onderstaande 2 tabellen. Deze nieuwe grids voor de database zijn uitgetest op de controle-groep van het *real-time* model (uit de bovenstaande grid-analyse) en op een ruwer grid van *df*-waarden. De linker subkolom beschrijft wederom de uitkomst na stap 5a, de rechter beschrijft de uitkomst na stap 5b.

15a) 120-15-26-16 (fijner grid voor F0/K) Rekentijd simulaties database: 889s

Gridkeuze	Rekentijd		SSR		Sigma		DF		Gamma	
30-52-32	11.45	31.30	0.278	0.083	0.240	0.234	3.250	3.591	0.750	0.740
30-26-32	6.13	21.26	0.305	0.083	0.230	0.234	4.000	3.591	0.725	0.740

15b) 160-15-26-16 (nog finer grid F0/K) Rekentijd simulaties database: 891s.

Gridkeuze	Rekentijd		SSR		Sigma		DF		Gamma	
30-52-32	10.97	26.71	0.290	0.085	0.240	0.234	3.250	3.592	0.750	0.741
30-26-32	6.05	18.75	0.315	0.085	0.315	0.234	4.000	3.592	0.725	0.741

Het onderzoek naar het effect van een slim *df*-grid wordt in Appendix 9 uitgebreid besproken. Hieronder staan de uitgebreide resultaten, waarin in elke tabel (16a t/m 16c) het grid voor F_0/K verder verfijnd wordt. Ook hier is onderscheid gemaakt tussen stap 5a (linkerkolom) en stap 5b (rechterkolom).

16a) 40-15-slim26-16 (standard grid F0/K) Rekentijd simulaties database: 478s

Gridkeuze	Rekentijd		SSR		Sigma		DF		Gamma	
30-slim52-32	42.8	241.3	0.166	0.148	0.230	0.231	3.591	3.560	0.725	0.722
30-slim26-32	16.8	252.9	0.260	0.148	0.230	0.231	3.704	3.560	0.725	0.722

16b) 80-15-slim26-16 (fijner grid voor F0/K) Rekentijd simulaties database: 1180s

Gridkeuze	Rekentijd		SSR		Sigma		DF		Gamma	
30-slim52-32	42.4	73.0	0.117	0.089	0.230	0.233	3.704	3.432	0.725	0.716
30-slim26-32	13.6	35.3	0.117	0.089	0.230	0.233	3.704	3.432	0.725	0.716

16c) 120-15-slim26-16 (nog finer grid F0/K) Rekentijd simulaties database: 1388s

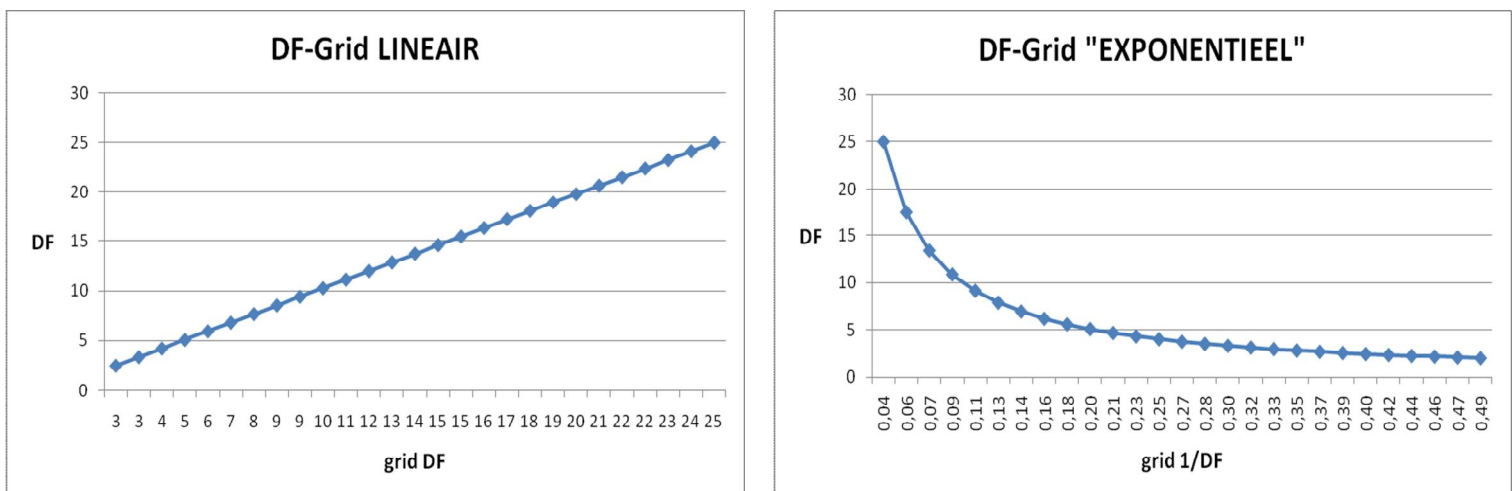
Gridkeuze	Rekentijd		SSR		Sigma		DF		Gamma	
30-slim52-32	30.3	50.9	0.115	0.081	0.230	0.233	3.704	3.447	0.725	0.718
30-slim26-32	11.9	35.0	0.115	0.081	0.230	0.233	3.704	3.447	0.725	0.718

Appendix 9: Exponentieel df-grid

In de *gridanalyse* van hoofdstuk 4.2 wordt het standaard grid vervangen voor een “exponentieel” grid, zodat meer relevante waarnemingen van df worden geëvalueerd. Dit idee is ontstaan uit het feit dat de *gevoeligheid* van de t-verdeling groter is bij hoge waarden van df dan bij lage waarden van df . Zodoende wordt onderzocht of het gebruik van een df -grid met een grotere stapgrootte bij hoge waarden van df dan bij lage waarden van df , een positieve invloed heeft op de resultaten van het *real-time* model.

In dit onderzoek wordt een “exponentieel” verloop van het df -grid getest. Met het oog op de efficiëntie van het ingebouwde interpolatie-algoritme van Matlab is het echter wel zaak om de stapgrootte van de grids evenredig te houden. Voor het lineaire grid is dit geen probleem: deze loopt simpelweg van df_{LB} (de ondergrens van df) tot df_{UB} (de bovengrens van df) met evenwijdige stapgrootte. De waarde in het df -grid is hier dus gelijk aan de waarde df zelf. Het is ook mogelijk om in het grid getransformeerde waarden van df te gebruiken, die bij gebruik in het *real-time* model eenvoudig teruggerekend kunnen worden. In dit onderzoek wordt een grid gebruikt dat loopt van $1/UB(df)$ tot $1/LB(df)$, met evenwijdige stapgrootte. Op deze manier is $\log(df)$ equidistant verspreid. Om de juist waarde van df te gebruiken voor de berekening van de optieprijs, moet df worden teruggerekend uit de waarde van het grid door de factor $1/df$ te gebruiken. Gevolg van dit getransformeerde grid is goed te zien in in figuur 10. Hier wordt een voorbeeld uitgewerkt met $df_{LB}=2.5$ en $df_{UB}=25$ met 26 stappen. Het is te zien dat het aantal gridpoints voor lage waarden van df “exponentieel” toeneemt en wordt voortaan aangeduid als *exponentieel grid*.

Het *exponentieel grid* is uitgebreid getest, en de belangrijkste resultaten worden weergegeven in tabel 17 en 18. Hierbij moet worden opgemerkt dat voor het df -grid in de database alleen een *exponentieel* df -grid is gebruikt met 26 stappen, aangezien de variant met het dubbele aantal stappen opmerkelijk veel rekentijd kostte. Vandaar dat is gekozen voor het testen van een drietal varianten van grids voor de *database* (waarbij steeds voor een verdere verfijning van het F_0/K -grid is gekozen) en 2 varianten van grids voor het *real time* model (het controle-grid en een verruwing van het df -grid). Aangezien de df -grids in de database bestaan uit de waarden $1/df$, wordt ook voor de grids in het real time model een *exponentieel* df -grid gebruikt.



Figuur 10 Voorbeeld van het lineaire grid (links) en het getransformeerde grid (rechts). De ondergrens van DF is 2.5, de bovengrens is 25 en de grids bestaan uit 26 stappen (gridpoints).

In combinatie met een fijn grid voor F_0/K (met 80 of 120 stappen) in de database en een ruw (*exponentieel*) *df*-grid in het real time model, presteert het *exponentiële* grid qua SSR beter dan het lineaire grid. Met behulp van het exponentiële grid kan bijvoorbeeld in slechts 12s. een SSR van 0.12 worden bewerkstelligd. Helaas kan deze trend niet worden doorgezet na de extra optimalisatie van stap 5b. Naarmate het grid voor F_0/K in de database fijner wordt genomen, lijkt de optimale SSR voor beide grids te convergeren naar dezelfde waarde (0.08). Helaas kost de optimalisatieprocedure voor het exponentiële grid in alle gevallen meer rekentijd dan de lineaire variant. Het lijkt erop dat de *fminsearch* dermate efficiënt werkt, dat de voorsprong van het exponentiële grid na stap 5a (volledig) teniet wordt gedaan. Voordat echte conclusies kunnen worden getrokken uit deze onderlinge vergelijking van de grids, zou een dergelijk experiment ook voor andere datasets moeten worden uitgevoerd. Het kan immers zo zijn dat deze *fminsearch* methode niet voor iedere dataset even efficiënt werkt. Met andere woorden: indien de *fminsearch* procedure blijft hangen in een *lokaal* minimum (in plaats van een *globaal* minimum) kan het voordeel van een lage SSR na stap 5a zeer waardevol zijn.

STAP 5A	40-15-26-16		80-15-26-16		120-15-26-16	
	Lineair <i>df</i>	Exp. <i>df</i>	Lineair <i>df</i>	Exp. <i>df</i>	Lineair <i>df</i>	Exp. <i>df</i>
30-exp52-32	-	0.17 (43s)	-	0.12 (42s)	-	0.12 (30s)
30-exp26-32	-	0.26 (17s)	-	0.12 (14s)	-	0.12 (12s)
30-52-32	0.17 (29s)	-	0.24 (11s)	-	0.28 (11s)	-
30-26-32	0.17 (11s)	-	0.29 (6s)	-	0.31 (6s)	-

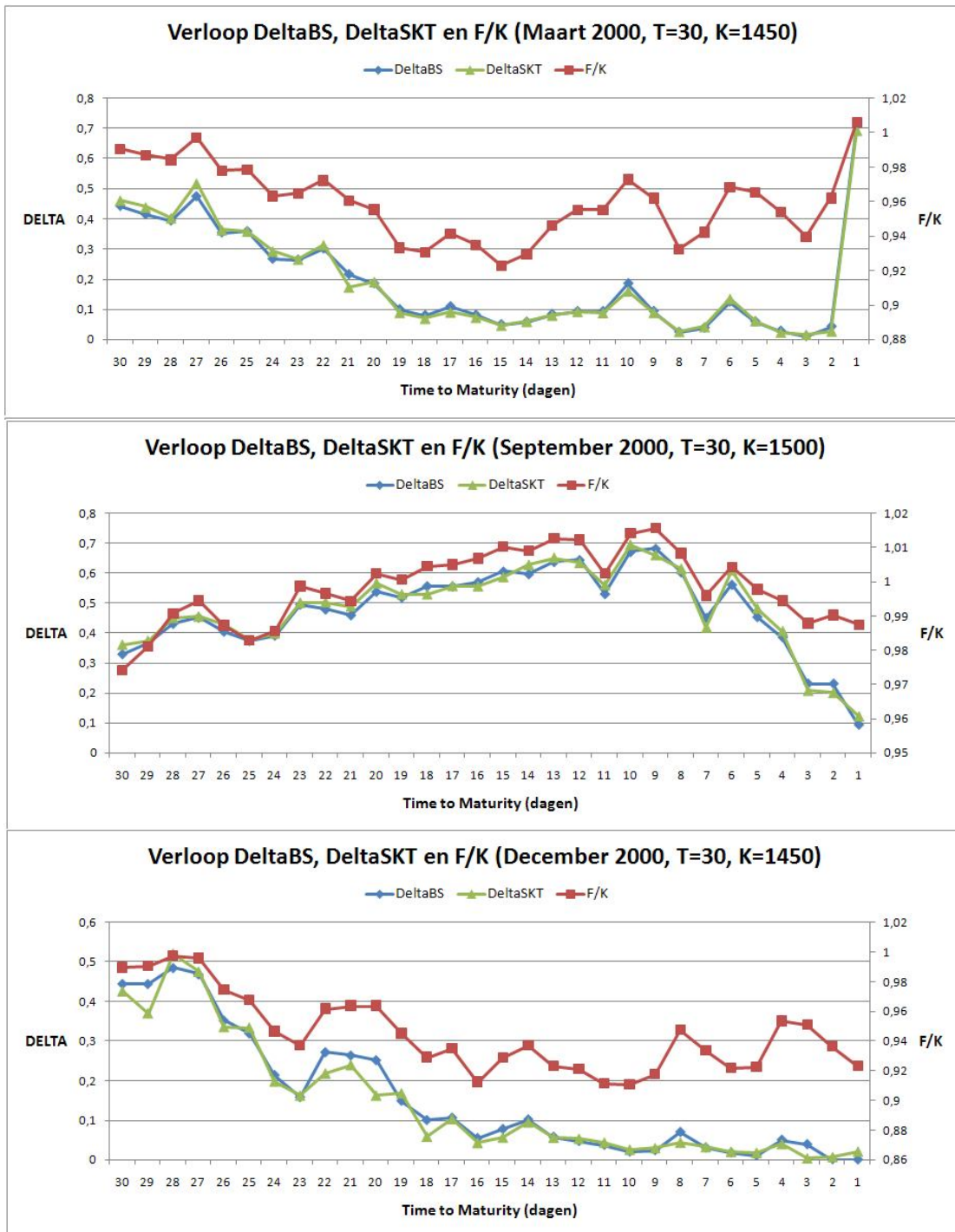
Tabel 17 Uitkomsten van vervolgonderzoek van *exponentiële df*-grid analyse t/m stap 5a. In de kolommen staan verfijningen van F_0/K in de *database* (resp 40, 80 en 120 stappen), in de rijen staan de verschillende grids voor het real time model (van zowel het lineaire als het *exponentiële df*-grid). De combinaties van een lineair grid in de database en een exponentieel grid in het real time model (en vice versa) zijn hier niet van toepassing.

STAP 5B	40-15-26-16		80-15-26-16		120-15-26-16	
	Lineair <i>df</i>	Slim <i>df</i>	Lineair <i>df</i>	Slim <i>df</i>	Lineair <i>df</i>	Slim <i>df</i>
30-exp52-32	-	0.15 (241s)	-	0.09 (73s)	-	0.08 (51s)
30-exp26-32	-	0.15 (253s)	-	0.09 (35s)	-	0.08 (35s)
30-52-32	0.13 (214s)	-	0.10 (23s)	-	0.08 (31s)	-
30-26-32	0.13 (119s)	-	0.10 (18s)	-	0.08 (21s)	-

Tabel 108 Uitkomsten van vervolgonderzoek van *exponentiële df*-grid analyse t/m stap 5a. De interpretatie van het model is analoog aan die van tabel 17.

Appendix 10 Delta-Hedging resultaten: Verloop van Delta

Analoog aan figuur 6 worden in figuur 11 vergelijkbare grafieken gepresenteerd van het verloop van de verschillende delta's en bijbehorende ratio's van F/K.



Figuur 11 Verloop van DeltaBS, DeltaSKT en F/K over de tijd van het Delta-Hedging experiment voor maart, september en december (van boven naar beneden). In de primaire verticale as staan de 2 delta's weergegeven, in de secundaire verticale as staat de ratio F/K.