

ERASMUS UNIVERSITEIT ROTTERDAM

Erasmus School of Economics

Bachelorscriptie [Economie en bedrijfseconomie]

Een wedloop tussen het voorspellend vermogen van impliciete- en historische volatiliteit

Abstract

Centraal in deze paper staat het onderzoeken van de mate waarin impliciete- en historische volatiliteit in staat zijn gerealiseerde volatiliteit te voorspellen, waarbij ik gebruik maak van 55 aandelen over de periode 1996-2020. In tegenstelling tot de literatuur van Christensen en Prabhala (1998) kies ik voor een aanpak waarbij het modelleren van historische volatiliteit met een AR model de basis vormt voor de voorspelde volatiliteit. Middels de mean square error stel ik vast in hoeverre de impliciete volatiliteit en de voorspelling op basis van historische volatiliteit afwijken van de gerealiseerde volatiliteit. Vervolgens onderzoek ik hoe de voorspelbaarheid verschilt over de 11 sectoren van de S&P-500 en of de Covid-19 pandemie de voorspelbaarheid beïnvloedt. Uit de resultaten blijkt dat de impliciete volatiliteit in zijn totaal een betere voorspeller is van gerealiseerde volatiliteit dan historische volatiliteit. Daarnaast verschilt de voorspelbaarheid aanzienlijk tussen sectoren en zorgt de aanwezigheid van de Covid-19 pandemie ervoor dat het moeilijker is om een accurate voorspelling te maken.

Naam student: Stijn Ouwehand

Studentnummer: 507490

Begeleider: R. Quaedvlieg

Tweede beoordelaar: R. de Blik

Datum definitieve versie: 19-7-2021

Het geschrevene in deze scriptie is de opvatting van de auteur en niet noodzakelijk die van de begeleider, tweede beoordelaar, Erasmus School of Economics of Erasmus Universiteit Rotterdam.

Inhoudsopgave

1	Introductie.....	3
2	Theoretisch kader.....	4
2.1	Een vergelijking tussen impliciete- en historische volatiliteit.....	4
2.2	Optie-prijs formules en impliciete volatiliteit.....	6
2.3	Modelleren van gerealiseerde volatiliteit.....	7
2.4	Hypotheses.....	7
3	Data.....	9
4	Methodologie.....	12
5	Resultaten.....	14
5.1	Hypothese 1.....	14
5.2	Hypothese 2.....	15
5.3	Hypothese 3.....	15
6	Conclusie en discussie.....	18
	Appendix A.....	20
	Appendix B.....	22
	Literatuurlijst.....	25

1 Introductie

De impliciete volatiliteit van een optie staat bekend als de verwachting van de markt van de gemiddelde volatiliteit tot aan de expiratie van een optie (Mayhew, 1995). De meest voor de hand liggende manier waarop de impliciete volatiliteit tot stand komt is door gebruik te maken van één van de vele verschillende optie-prijs modellen, waarvan elke variant zijn eigen implicaties kent. Het meest bekende model is het Black-Scholes model (Black en Scholes, 1973), die gekenmerkt wordt door een constante impliciete volatiliteit. Aangezien empirische waarnemingen het tegengestelde bewijzen, zijn door de tijd heen verschillende alternatieven ontwikkeld.

Een logische gedachte die volgt, is om te onderzoeken in hoeverre deze voorspelde (impliciete) volatiliteit voldoet aan de daadwerkelijk gerealiseerde volatiliteit. Veel onderzoekers maken hierbij een vergelijking tussen impliciete- en historische volatiliteit, waarvan bekende voorbeelden afkomstig zijn van Canina & Figlewski (1993) en Christensen & Prabhala (1998). De laatstgenoemde onderzoekers zetten binnen deze tak van de literatuur een standaard waarbij impliciete volatiliteit het wint ten opzichte van historische volatiliteit. Ook in het kader van historische volatiliteit, bestaan veel mogelijkheden om deze te gebruiken voor het voorspellen van toekomstige volatiliteit. Het meest bekende model is het HAR-RV model van Corsi (2008), dat uitblinkt door met een simpel model de meeste relevante kenmerken van financiële data te volgen.

De hiervoor genoemde (literaire) elementen schetsen het probleemgebied van dit onderzoek en vormen daarom de basis voor de volgende hoofdvraag: *In welke mate zijn de impliciete volatiliteit en historische volatiliteit in staat een accurate voorspelling te vormen van de gerealiseerde volatiliteit van een selectie bedrijven uit de S&P-500 over een periode van 1996 tot en met 2020?*

Door een antwoord te formuleren op de hoofdvraag, wordt er naar gestreefd een bijdrage te leveren aan de literatuur met betrekking tot het voorspellend vermogen van impliciete- en historische volatiliteit. Dit kan zijn in de vorm van een bevestiging van de huidige consensus of eventueel een inzicht dat aanleiding geeft tot het verrichten van nieuw onderzoek. Daarnaast heeft deze paper maatschappelijke relevantie aangezien investeerders in de aandelenmarkt hun werkwijze kunnen aanpassen of baseren op de uitkomsten van dit onderzoek.

In de hierop volgende sectie geef ik een overzicht van de relevante literatuur ter onderbouwing van de hypotheses. Hoofdstuk 3 beschrijft de gebruikte data en de wijze waarop deze geselecteerd is. Vervolgens wordt in hoofdstuk 4 genoemd welke methodiek gehanteerd is ter analyse van de data. Hoofdstuk 5 geeft alle resultaten weer, waaruit ik in hoofdstuk 6 een conclusie vorm in combinatie met een aanbeveling voor vervolgonderzoek.

2 Theoretisch kader

2.1 Een vergelijking tussen impliciete- en historische volatiliteit

Vermoedelijk één van de meest besproken aspecten uit de literatuur met betrekking tot het voorspellen van volatiliteit gaat over de afweging tussen impliciete volatiliteit en historische volatiliteit om toekomstige volatiliteit te voorspellen. De algemene conclusie die hieruit getrokken kan worden is dat impliciete volatiliteit hiervoor meer geschikt is, zoals wordt beschreven in het literatuuroverzicht van Mayhew (1995) (Latané en Rendleman, 1976; Schmalensee en Trippi, 1978; Chiras en Manaster, 1978; Beckers, 1981). Naast dat Latané en Rendleman (1976) tot deze conclusie komen, ondervinden zij ook dat impliciete volatiliteit gecorreleerd is met gerealiseerde volatiliteit. Schmalensee en Trippi (1978) voegen daar nog aan toe dat zij geen relatie vinden tussen de verwachte volatiliteit en historische data van aandelen. Echter wordt nog niet gecorrigeerd voor het uitkeren van dividenden, wat Chiras en Manaster (1978) ter sprake brengen. Zij tonen aan dat een aangepast Black-Scholes model corrigerend voor dividenden betere resultaten behaalt dan in vorige studies waarbij dit niet het geval is. Beckers (1981) benadrukt tevens het in overweging nemen van dividenden. Daarnaast wordt aangetoond dat de meest relevante informatie verwerkt zit in de optieprijsen die at-the-money zijn, in tegenstelling tot de bevindingen van Latané en Rendleman (1976), die de nadruk leggen op het gebruiken van alle opties. Tenslotte blijkt uit de bevindingen van Beckers (1981) dat niet alle informatie zichtbaar is in de huidige optieprijs, daarom wordt de suggestie gewekt dat optiemarkten mogelijk niet efficiënt zijn.

Een tegengeluid wordt geboden vanuit Canina en Figlewski (1993) die concluderen dat de impliciete volatiliteit een slechte voorspeller is van gerealiseerde volatiliteit. Hierbij maken zij gebruik van S&P100 opties (OEX) over een periode van 1983 tot en met 1987 met behulp van een binomiaal model dat corrigeert voor dividenden en vroege uitoefening. Uit de resultaten blijkt dat de impliciete volatiliteit überhaupt niet gecorreleerd is aan de toekomstige volatiliteit en geen informatie opneemt uit historische volatiliteit. Canina en Figlewski (1993) beargumenteren dat externe factoren invloed uitoefenen op de volatiliteit (zonder dat deze gecorrigeerd worden door arbitrage) en dat daardoor de voorspelbaarheid van impliciete volatiliteit samen moet hangen met de mogelijkheden voor arbitrage. Ondanks dat de bevindingen aantonen dat historische volatiliteit enige variatie kan verklaren in de toekomstige volatiliteit zien zij historische volatiliteit niet gezien als een geschikte voorspeller.

Een reactie hierop volgt onder andere uit de studie van Christensen en Prabhala (1998). Dit onderzoek verschilt van Canina en Figlewski (1993) door het gebruik van een langer durende dataset en het hanteren van een minder frequente data observatie. In plaats van wekelijkse data wordt gekozen voor maandelijks, wat de onderzoekers in staat stelt een model te creëren zonder overlappende data. Zij

concluderen dat impliciete volatiliteit een betere voorspelling geeft van toekomstige volatiliteit dan historische volatiliteit, waarbij het resultaat te verklaren is door de genoemde aanpassingen. Daarnaast tonen de bevindingen aan dat impliciete volatiliteit een betere voorspeller is van toekomstige volatiliteit ná de aandelenmarktcrash uit 1987. Ze tonen aan dat er een structurele verandering plaats vond in het waarderen van aandelenopties volgend op deze gebeurtenis.

Christensen en Prabhala (1998) geven twee redenen als mogelijke verklaring voor de bevindingen van Canina en Figlewski (1993), namelijk dat de volatiliteit inefficiënt verwerkt wordt door index optiemarkten en het gebruikte Black-Scholes model niet geschikt zou zijn voor het prijzen van indexopties vanwege te hoge transactiekosten. Gezien de hoge liquiditeit in de OEX opties en het ontbreken van een directe relatie tussen transactiekosten en optieprijzen worden deze motivaties onaannemelijk geacht door de auteurs.

Verder wordt in de paper van Christensen en Hansen (2002) een uitbreiding genoemd op de bevindingen van Christensen en Prabhala (1998) door impliciete volatiliteit te berekenen door een gemiddelde te nemen van zowel in-the-money als out-of-the-money call- en put opties. Wederom komen de onderzoekers tot de conclusie dat impliciete volatiliteit een efficiënte voorspelling geeft van gerealiseerde volatiliteit (Christensen en Hansen, 2002).

Recenter onderzoek van Busch et al. (2011) probeert gerealiseerde volatiliteit te voorspellen in de valuta-, aandelen- en obligatiemarkten door een HAR model toe te passen en impliciete volatiliteit als extra verklarende variabele toe te voegen. Door op basis van de genomen sample een voorspelling te maken van de gerealiseerde volatiliteit wordt geconstateerd dat impliciete volatiliteit in alle onderzochte markten een belangrijke verklarende factor is. Opmerkelijk is dat Busch et al. (2011) concluderen dat het 'jump' onderdeel in de gerealiseerde volatiliteit in beperkte mate te voorspellen is.

Tenslotte betogen Bandi en Perron (2006) dat het niet mogelijk is om de relatie tussen impliciete- en gerealiseerde volatiliteit op korte termijn te onderzoeken middels een regressie. Dit komt omdat zij beargumenteren dat er met een regressie mogelijk sprake is van fractionele coïntegratie. Met behulp van een aangepaste benadering tonen de resultaten van Bandi en Perron (2006) aan dat impliciete volatiliteit op de lange termijn een unbiased voorspelling geeft van de gerealiseerde volatiliteit, consistent met de algemene consensus binnen de literatuur. Op het gebied van interpretatie van de resultaten verschilt dit onderzoek dus van de overige literatuur.

2.2 Optie-prijs formules en impliciete volatiliteit

Door de jaren heen zijn oneindig veel manieren ontwikkeld om de prijs van een optie te benaderen, welke allemaal verschillende waarden van impliciete volatiliteit tot stand brengen. Derhalve is het dus relevant een inzicht te krijgen in hoe bepaalde modellen functioneren. Over het algemeen wordt de impliciete volatiliteit berekend met ofwel het Black-Scholes model (Black en Scholes, 1973) of het Cox-Ross-Rubinstein binomial model (Cox, Ross en Rubinstein, 1979). Opvallend bij het Black-Scholes model is dat de volatiliteitsparameter wordt gezien als een constante. Deze is te interpreteren als de verwachting van de markt van de gemiddelde volatiliteit tot aan de expiratie van een optie (Mayhew, 1995). Empirisch gezien blijkt alleen, onder andere uit Rubinstein (1994), dat er helemaal geen sprake is van een invariabele impliciete volatiliteit. In het geval dat het Black-Scholes model een volledig accurate weergave geeft van de werkelijkheid, zou elke optie met overeenkomende onderliggende waarde en tijd tot expiratie namelijk dezelfde impliciete volatiliteit moeten geven.

Als reactie op het probleem van invariabele impliciete volatiliteit zijn veel onderzoekers aan de slag gegaan met het creëren van een model op basis van gewogen gemiddelden van de impliciete volatiliteit (Schmalensee en Trippi, 1978; Latané en Rendleman, 1976; Chiras en Manaster, 1978; Mayhew, 1995). Beckers (1981) deed vervolgens een studie waarbij het voorspellend vermogen van verschillende gewogen gemiddelde modellen getest wordt, waaruit blijkt dat simpelweg het gebruiken van de optie die zich at-the-money bevindt minstens even goed presteert als de gewogen gemiddelden. Als aanvulling hierop ondervinden Brenner en Galai (1981) dat het voorspellend vermogen verbetert indien het gewogen gemiddelde van de impliciete volatiliteit meerdere keren per dag wordt berekend en hiervan het gemiddelde wordt gehanteerd. Hieruit blijkt dat intraday verschillen mogelijk van invloed zijn op de schatting van de impliciete volatiliteit (Mayhew, 1995).

Met oogpunt op het invullen van de gebreken gerelateerd aan het Black-Scholes model onderzoeken Dumas, Fleming en Whaley (1998) het voorspellend vermogen van een *deterministic volatility function (DVF)*, ondersteund door de literatuur van Derman en Kani (1994), Dupire (1994) en Rubinstein (1994). De gedachte hierachter is dat de volatiliteit van de returns van onderliggende waarden afhankelijk zijn van de prijs en tijd en maken daarom gebruik van een binomiale of trinomiale waardering van optiepreizen. Uit de resultaten blijkt echter dat de DVF niet in staat is de volatiliteit beter te modelleren dan bij het Black-Scholes model. Dumas, Fleming en Whaley (1998) concluderen letterlijk “simpler is better”, refererend naar het Black-Scholes model.

2.3 Modelleren van gerealiseerde volatiliteit

Gegeven het brede scala aan verschillende opties voor het modelleren van gerealiseerde volatiliteit zal ik in deze sectie vooral aandacht besteden aan de paper van Corsi (2008), waarin een kort overzicht wordt gegeven van gerelateerde literatuur en een model wordt geïntroduceerd dat heden ten dage als vooraanstaand beschouwd wordt.

Corsi (2008) beschrijft dat financiële data een aantal kenmerken kent die het creëren van een effectief model lastig kunnen maken. Onder deze kenmerken vallen onder andere een lang aanhoudende autocorrelatie van de rendementen, de aanwezigheid van fat tails en het bestaan van scaling behaviour. Zo wordt aangegeven dat de gebruikelijke GARCH en stochastische volatiliteitsmodellen niet in staat al deze kenmerken te respecteren. Met name zijn deze modellen niet geschikt om te gebruiken voor scaling, waardoor andere modellen gebruikt moeten worden voor het aggregeren over grotere periodes.

Volatiliteit over de lange termijn wordt normaliter bepaald met behulp van fractional difference operators, zoals in een ARFIMA of FIGARCH model (Corsi, 2008). Nadeel hiervan is alleen dat de economische interpretatie ontbreekt en een groot aantal observaties opgeofferd moet worden wegens een opbouwperiode. Als alternatief introduceert Corsi (2008) het HAR-RV model, dat de volatiliteit berekend over verschillende intervallen in overweging neemt. “Verrassend genoeg, ondanks de simpliciteit en het feit dat dit model formeel gezien niet onder de klasse lange-termijn volatiliteitsmodellen valt, is het HAR-RV model in staat dezelfde volharding te reproduceren als wordt waargenomen in de empirische data net als vele van de andere belangrijke kenmerken van financiële data.” (Corsi, 2008) Nadat het HAR-RV model in een out-of-sample performance test vergeleken wordt met zowel AR- als ARFIMA modellen, is de conclusie dan ook dat het HAR-RV model consistent betere resultaten behaalt dan de overige modellen.

2.4 Hypotheses

Uit het onderzoek van onder andere Christensen en Prabhala (1998) volgt de algemene consensus dat impliciete volatiliteit beter in staat is gerealiseerde volatiliteit te voorspellen dan historische volatiliteit. Aangezien het aannemelijk is dat deze conclusie eveneens voort komt uit dit onderzoek en grotendeels relevant is voor het beantwoorden van de hoofdvraag, vormt dit een voor de hand liggende eerste hypothese:

Vergeleken met historische volatiliteit is de impliciete volatiliteit beter in staat de gerealiseerde volatiliteit te voorspellen.

De eerste hypothese wekt interesse om in groter detail te bekijken hoe de voorspelbaarheid van volatiliteit is opgedeeld, wat mogelijk is door te kijken naar sectoren. Gezien de onderliggende aandelen en eigenschappen tussen sectoren aanzienlijk kunnen verschillen, is het bevattelijk dat in bepaalde sectoren in vergelijking met andere sectoren het beter mogelijk is een voorspelling te maken van de gerealiseerde volatiliteit. Om deze zienswijze te bevestigen volgt hypothese 2:

Er is een significant verschil waarneembaar tussen sectoren met betrekking tot de mate waarin impliciete- en historische volatiliteit in staat zijn gerealiseerde volatiliteit te voorspellen.

De onverwachtse situatie in de vorm van de Covid-19 pandemie geeft nieuwe mogelijkheden om de voorspelbaarheid van volatiliteit in een ander opzicht te bekijken. Gegeven de hoge volatiliteit die komt kijken bij een pandemie als deze is te verwachten dat dit een obstakel kan vormen om accuraat de gerealiseerde volatiliteit te voorspellen. Hieruit volgt hypothese 3:

Impliciete- en historische volatiliteit zijn in verband met de Covid-19-pandemie in het jaar 2020 minder goed in staat de gerealiseerde volatiliteit te voorspellen ten opzichte van voorafgaande periode.

Merk op dat hypothese 2 en 3 geen direct gerelateerde literatuur hebben ter onderbouwing. Ik ben niet in staat geweest enige vorm van literatuur te vinden die de voorspelbaarheid van volatiliteit behandelt en een afsplitsing maakt naar enerzijds sectoren of anderzijds de Covid-19 pandemie. De gerelateerde literatuur zou daarom van te lage relevantie zijn ter ondersteuning van dit onderzoek en is wegens deze reden logischerwijs weggelaten uit het theoretisch kader.

3 Data

De data benodigd om de gegeven hypotheses empirisch te onderzoeken is afkomstig uit de *Option Metrics* database, beschikbaar gesteld via *Wharton Research Data Services (WRDS)*. Deze database staat bekend om het leveren van hoogwaardige kwaliteit data met betrekking tot historische opties en impliciete volatiliteit, gebruikt door zowel academici als toonaangevende spelers in het bedrijfsleven. Binnen dit onderzoek maak ik gebruik van zowel de impliciete volatiliteit als de gerealiseerde volatiliteit. Voor beide variabelen hanteer ik een tijdsbestek van januari 1996 tot en met december 2020, wat de grootst mogelijke beschikbare periode is waarin de variabelen overlappen.

De impliciete volatiliteit is berekend op basis van gestandaardiseerde optiepreizen. Allereerst wordt de forward price vastgesteld met behulp van de zero curve en de daarbij behorende verwachte distributies. Door de volatility surface punten lineair te interpoleren op basis van de forward price en de expiratie, wordt een at-the-money impliciete volatiliteit verkregen (*IvyDB US Reference Manual*, 2021). Deze at-the-money strike price is in lijn met het eerder besproken onderzoek van Beckers (1981) waaruit volgt dat hier de meest relevante informatie van optiepreizen verwerkt zit. De gestandaardiseerde optiepreizen brengen als voordeel mee dat het minder waarschijnlijk is dat de dataobservaties overlappen, wat volgens de studie van Christensen en Prabhala (1998) zorgt voor meer accurate resultaten. De gerealiseerde volatiliteit is gedefinieerd als de gemiddelde volatiliteit van de rendementen behorende bij een specifiek aandeel over een geselecteerde periode uit het verleden en berekend met een standaarddeviatie over de logaritme van de dagelijkse totale rendementen (*IvyDB US Reference Manual*, 2021). Voor zowel de impliciete volatiliteit als de gerealiseerde volatiliteit heb ik besloten om deze te berekenen over een periode van 30 kalenderdagen. Dit resulteert in een dataset van maandelijks observaties over totaal 300 maanden.

Tabel 1: overzicht sectoren met bijbehorende ETF's gebaseerd op de SP-500

Sector	ETF Ticker	Top 5 holdings weight
Information technology	XLK	53,39%
Health Care	XLV	30,96%
Consumer Discretionary	XLY	53,19%
Communication Services	XLC	57,50%
Financials	XLF	40,77%
Industrials	XLI	23,1%
Consumer Staples	XLP	48,93%
Utilities	XLU	43,05%
Materials	XLB	40,15%

Real estate	XLRE	43,72%
Energy	XLE	57,53%

De derde kolom geeft het percentage van de 5 grootste holdings binnen een specifieke ETF.

De data heb ik geselecteerd op basis van de *S&P Dow Jones Indices (Sectors - S&P Dow Jones Indices, 2021)*, welke de S&P-500 onderverdeelt in 11 verschillende sectoren. Vervolgens zijn voor de sectoren representatieve ETF's gevonden, zoals zichtbaar in tabel 1 (Israelsen, 2021). Van elke sector worden de 5 aandelen meegenomen in de analyse die het meest meewegen in de sector ETF, wat resulteert in een totaal aantal aandelen van 55 (zie Appendix A). Aangezien de 5 grootste holdings van de geselecteerde ETF's gemiddeld 45% uitmaken van de totale ETF, is redelijkerwijs te verwachten dat dit een representatief beeld kan geven van de sector zonder elke holding specifiek op het nemen in het onderzoek. Vanuit het oogpunt van een grote hoeveelheid extra arbeid heb ik op basis van de voorafgaande verklaring besloten 5 aandelen per sector te hanteren. In het geval dat de dataperiode van een aandeel niet voldoet aan de tijdsperiode 1996-2020, vervang ik deze door een ander aandeel die dit wel heeft, geselecteerd op basis van het percentage waarin dit aandeel meeweegt in de ETF.

Deze strategie waarbij sectordata wordt gebruikt, is geschikt voor dit onderzoek omdat dit een goede manier is om met één dataset alle hypothesen te analyseren. Ondanks dat de data bestaat uit verschillende sectoren is het mogelijk om na te gaan welke variabele in totaal een betere voorspeller is en te zien hoe dit verschilt tussen een jaar in combinatie met de Covid-19 pandemie en de periode daarvoor.

Tabel 2 geeft de beschrijvende statistieken van de impliciete- en gerealiseerde volatiliteit, waarbij de tabel onderscheid maakt tussen de 11 geselecteerde sectoren. Wanneer gekeken wordt naar het gemiddelde van de twee variabelen, valt op dat in een meerderheid van de sectoren de impliciete volatiliteit een hogere waarde aanneemt dan de gerealiseerde volatiliteit. Daarnaast is de gemiddelde impliciete- en gerealiseerde volatiliteit het hoogst in de Information Technology sector met 0,351, gevolgd door Financials, Consumer Discretionary en Communication Services. De Utilities sector is gemiddeld gezien met 0,204 het minst volatiel, gevolgd door Consumer Staples, Health care en Real Estate. Opmerkelijk is dat de standaarddeviatie van gerealiseerde volatiliteit in alle sectoren hoger is dan de impliciete volatiliteit, met als uitschieter de Financials sector. Tenslotte is gemiddeld gezien over de sectoren in hoge mate sprake van een scheve frequentieverdeling en positieve kurtosis.

Tabel 2: beschrijvende statistieken van de impliciete- en gerealiseerde volatiliteit

Sector	Variabele	Gemiddelde	Standaarddeviatie	Skewness	Kurtosis
<i>Information technology</i>	IV	0,351	0,134	1,074	3,678
	RV	0,347	0,173	1,608	6,075
<i>Health Care</i>	IV	0,252	0,080	0,987	4,035
	RV	0,246	0,116	2,109	10,682
<i>Consumer Discretionary</i>	IV	0,315	0,109	0,858	2,781
	RV	0,312	0,131	0,966	3,387
<i>Communication Services</i>	IV	0,307	0,109	1,264	4,296
	RV	0,295	0,141	1,810	7,745
<i>Financials</i>	IV	0,320	0,169	2,852	14,735
	RV	0,336	0,257	3,296	16,159
<i>Industrials</i>	IV	0,276	0,098	1,691	7,266
	RV	0,280	0,144	3,002	17,369
<i>Consumer Staples</i>	IV	0,222	0,084	1,035	3,233
	RV	0,216	0,114	2,087	9,367
<i>Utilities</i>	IV	0,204	0,070	1,725	6,639
	RV	0,200	0,113	4,096	27,730
<i>Materials</i>	IV	0,281	0,078	1,115	3,667
	RV	0,271	0,101	1,858	8,514
<i>Real Estate</i>	IV	0,258	0,061	1,045	3,894
	RV	0,258	0,084	1,200	3,824
<i>Energy</i>	IV	0,290	0,096	2,286	12,211
	RV	0,290	0,147	4,100	30,202

Impliciete volatiliteit en gerealiseerde volatiliteit zijn respectievelijk afgekort als 'IV' en 'RV'

4 Methodologie

De meest bekende en voor de hand liggende methode om het voorspellende vermogen van impliciete- en historische volatiliteit te vergelijken, is om een simpele regressieanalyse toe te passen. Onder andere in de literatuur van Christensen en Prabhala (1997) is hier sprake van. Echter heb ik binnen dit onderzoek besloten om de individuele waarden van zowel de impliciete volatiliteit als de voorspelde volatiliteit op basis van een model middels MSE te vergelijken met de gerealiseerde volatiliteit. Voor de voorspelling van de volatiliteit maak ik een out of sample forecast op basis van een AR(3) model in de vorm van

$$RV_t = c + \varphi_1 RV_{t-1} + \varphi_2 RV_{t-2} + \varphi_3 RV_{t-3} + \varepsilon_t \quad (1)$$

Hierbij staat RV voor de gerealiseerde maandelijkse volatiliteit, is c een constante en ε_t de error term in de voorspelde waarde van de RV.

Gezien het doel om de voorspelde volatiliteit zo goed mogelijk overeen te laten komen met de gerealiseerde volatiliteit heb ik ervoor gekozen de optimale vorm van het model te selecteren op basis van de laagste MSE, in plaats van significantie of information criteria zoals AIC en BIC. Uit eigen empirische waarnemingen blijkt dat een AR(1) en AR(2) model ver onder de prestaties van een AR(3) model zitten. Vanaf een AR(4) model worden de MSE's eveneens groter en neemt de performance dus af. Daarnaast hanteer ik voor het gegeven model in vergelijking 1 een rolling window van 60 maanden. In het geval van een groter rolling window neemt het voorspellende vermogen slechts minimaal toe in vergelijking met de hoeveelheid observaties die hiervoor weggelaten moeten worden. Dit weegt daarom niet meer tegen elkaar op.

Ondanks dat het geselecteerde AR model een geschikte optie is voor het voorspellen van volatiliteit, bestaan er betere alternatieven. Voor de hand liggende mogelijkheden bestaan uit onder andere het HAR-RV model van Corsi (2008) en een ARFIMA model. Vanwege de hoge datafrequentie benodigd voor het HAR-RV model en complexiteit van een ARFIMA model heb ik echter besloten een AR model te hanteren.

Nadat middels het model een voorspelling is gemaakt, bereken ik de MSE over:

- Het verschil tussen de impliciete volatiliteit en de gerealiseerde volatiliteit
- Het verschil tussen de voorspelde volatiliteit en de gerealiseerde volatiliteit

De variabele die hierbij resulteert in de laagste waarde van de MSE is het best in staat een weergave te geven van de gerealiseerde volatiliteit. Het verschil tussen de voorspellingen zal op significantie getoetst worden met behulp van een Diebold-Mariano (1995) test.

Voor het analyseren van de derde hypothese heb ik de data opgesplitst in twee panels die verschillende tijdsbestekken betreffen:

- Panel A: 1 Januari 2019 – 31 Januari 2020
- Panel B: 1 Februari 2020 – 31 December 2020

Panel B start aan het begin van de Covid-19 pandemie en met behulp van Panel A als referentiekader kan ik een vergelijking maken tussen de voorspelbaarheid van volatiliteit gedurende de Covid-19 pandemie en de voorafgaande periode. De gehanteerde startdatum van de Covid-19 pandemie is 22 Januari 2020 en afkomstig uit het onderzoek van Engelhardt et al. (2021). Aangezien sprake is van maandelijkse datafrequentie wordt de datum 1 Februari aangehouden. Nadeel van de maandelijkse data is dat Panel A nu slechts bestaat uit 11 observaties per aandeel. Het zou daarom beter geweest zijn om de volatiliteit op dagelijks niveau te kunnen voorspellen. Echter neemt dit niet weg dat op de huidige manier een indicatie van voldoende kwaliteit gegeven kan worden met betrekking tot deze hypothese.

Tenslotte vind ik het belangrijk nog een tweetal kanttekeningen te plaatsen bij de gebruikte aanpak. Zo heb ik geen rekening gehouden met het feit dat corrigeren voor dividenden zorgt voor betere onderzoeksresultaten (Chiras en Manaster, 1978; Beckers, 1981). Daarnaast kan er een geringe dataoverlap plaatsvinden in het geval dat specifieke maanden niet gelijk lopen met de gehanteerde periode van 30 dagen.

5 Resultaten

5.1 Hypothese 1

Allereerst toets ik Hypothese 1, die stelt dat de impliciete volatiliteit vergeleken met historische volatiliteit beter in staat is de gerealiseerde volatiliteit te voorspellen. Uit de resultaten in tabel 3 (zie Appendix B voor een compleet overzicht inclusief alle onderliggende aandelen) blijkt dat op basis van de totale sample de MSE van impliciete volatiliteit en historische volatiliteit respectievelijk 0,0094 en 0,0194 bedragen. De MSE van de impliciete volatiliteit is daardoor 0,0102 lager, wat meer dan 50% bedraagt van MSE op basis van historische volatiliteit. Dit verschil is significant op basis van een significantieniveau van 5%. Slechts bij één aandeel is MSE HV met een bijna verwaarloosbare waarde lager dan MSE IV, namelijk in het geval van het bedrijf Public Storage binnen de sector Real Estate. Aangezien in de overige 54 aandelen dit niet van toepassing is en er een aanzienlijk significant verschil (gemiddeld 0,0102) bestaat tussen de twee MSE's, vormt dit reden om de eerste hypothese niet te verwerpen. De impliciete volatiliteit ligt consistent veel dichterbij de gerealiseerde volatiliteit dan de volatiliteit geschat op basis van historische volatiliteit.

Tabel 3: de mean square error van de impliciete- en historische volatiliteit verdeeld over sectoren

Sector	MSE IV	MSE HV	Vershil
<i>Information Technology</i>	0,0119	0,0190	<u>-0,0071*</u>
<i>Health Care</i>	0,0090	0,0141	<u>-0,0064*</u>
<i>Consumer Discretionary</i>	0,0085	0,0129	<u>-0,0045*</u>
<i>Communication Services</i>	0,0076	0,0156	<u>-0,0080*</u>
<i>Financials</i>	0,0210	0,0511	<u>-0,0301*</u>
<i>Industrials</i>	0,0076	0,0202	<u>-0,0126*</u>
<i>Consumer Staples</i>	0,0041	0,0093	<u>-0,0052*</u>
<i>Utilities</i>	0,0062	0,0171	<u>-0,0109</u>
<i>Materials</i>	0,0066	0,0135	<u>-0,0069*</u>
<i>Real estate</i>	0,0124	0,0162	<u>-0,0038</u>
<i>Energy</i>	0,0083	0,0249	<u>-0,0166</u>
Totaal	0,0094	0,0194	<u>-0,0102*</u>

Overzicht van de mean square error (MSE) van zowel de impliciete- als historische volatiliteit ten opzichte van de gerealiseerde volatiliteit. De derde kolom is tot stand gekomen door MSE IV minus MSE HV te berekenen. De onderste regel geeft een totaal overzicht over alle meegenomen aandelen in het onderzoek. De significantieniveaus 5%, 1% en 0,1% voor de Diebold-Mariano (1995) toets in de laatste kolom zijn respectievelijk aangegeven met *, ** en ***.

5.2 Hypothese 2

In deze sectie maak ik een analyse van Hypothese 2. Deze bouwt voort op de eerste hypothese door te kijken of er onderscheid gemaakt kan worden tussen verschillende sectoren in de mate waarin impliciete- en historische volatiliteit in staat zijn gerealiseerde volatiliteit te voorspellen. Hiervoor zal ik wederom gebruik maken worden van tabel 3, waarin alle resultaten vermeld staan, verdeeld over 11 sectoren. De Financials sector is de grootste uitschieter, waarbij het verschil van 0,0301 tussen de MSE van impliciete volatiliteit en historische volatiliteit het hoogst is over alle sectoren. Naast dat het verschil enorm is, geldt dit ook voor de MSE's zelf. De MSE van de impliciete- en historische volatiliteit bedragen respectievelijk 0,0210 en 0,0511 ten opzichte van het gemiddelde van 0,0094 en 0,0194. Verder is het verschil tussen de MSE's groot bij de Energy sector met 0,0166 en de Industrials sector met 0,0126. Naast uitschieters aan de bovenkant is ook sprake van een aantal sectoren waarbij de MSE's en het verschil hiertussen relatief laag zijn. In de sector Real Estate is het verschil het kleinst, met een waarde van 0,0038. Van een vergelijkbare situatie is sprake bij de sector Consumer Discretionary met 0,0045 en Consumer Staples met 0,0052. Beide MSE's zijn het laagst in de Consumer Staples sector, wat aangeeft dat de voorspelling hier het meest accuraat is. Het verschil tussen de MSE's is significant op een significantieniveau van 5% bij 8 van de 11 sectoren.

Op basis van het voorafgaande en de overige resultaten in tabel 3, stel ik vast dat de MSE van impliciete- en historische volatiliteit in hoge mate verschilt tussen sectoren. Dat wil zeggen dat de mate waarin de impliciete- en historische volatiliteit in staat zijn een accurate weergave te geven van de daadwerkelijke volatiliteit, verschilt per sector. Ondanks dat er niet bij alle sectoren voldoende significantie aanwezig is, vormt dit voldoende bewijs om deze hypothese niet te verwerpen.

5.3 Hypothese 3

Tenslotte veronderstelt Hypothese 3 dat de impliciete- en historische volatiliteit in verband met de Covid-19 pandemie in het jaar 2020 minder goed in staat zijn de gerealiseerde volatiliteit te voorspellen ten opzichte van de voorafgaande periode. Zoals te zien in tabel 4 wordt hier onderscheid gemaakt tussen twee panels, waarbij panel A geheel 2019 en 2020 bedekt en panel B het tijdsbestek hanteert vanaf de start van de Covid-19 pandemie. Duidelijk is af te lezen dat de gemiddelde MSE van de impliciete- en historische volatiliteit in panel B veel hoger ligt dan in panel A, wat erop wijst dat de voorspelbaarheid van volatiliteit gedurende COVID-19 minder is dan in 2019. Ook in combinatie met de gemiddelde MSE's uit tabel 3 blijkt dat de waarden gedurende de pandemie veel hoger liggen.

Indien ik de MSE's direct met elkaar vergelijk, is er een interessant inzicht te ontleen uit tabel 4. Gedurende de COVID-19 pandemie presteert de impliciete volatiliteit namelijk aanzienlijk beter dan de voorspelling op basis van de historische volatiliteit. Het verschil tussen de MSE's bedraagt in panel

B 0,0801, wat 50 keer zoveel is als de 0,0016 uit panel A. Indien ik deze vergelijking maak met de totale steekproef uit tabel 3, is het verschil grofweg 8 keer zo groot. Het lijkt er dus op dat in deze periode van uitzonderlijk hoge volatiliteit, het model op basis van historische volatiliteit in nog mindere mate in staat is de accuraatheid van de impliciete volatiliteit bij te benen. Het verschil tussen de voorspellingen in panel A en B is significant tegen een significantieniveau van respectievelijk 5% en 1%. Op basis van de gemiddeld hogere MSE's en het aanzienlijke significante verschil tussen de voorspellingen in panel B verwerp ik de derde hypothese niet.

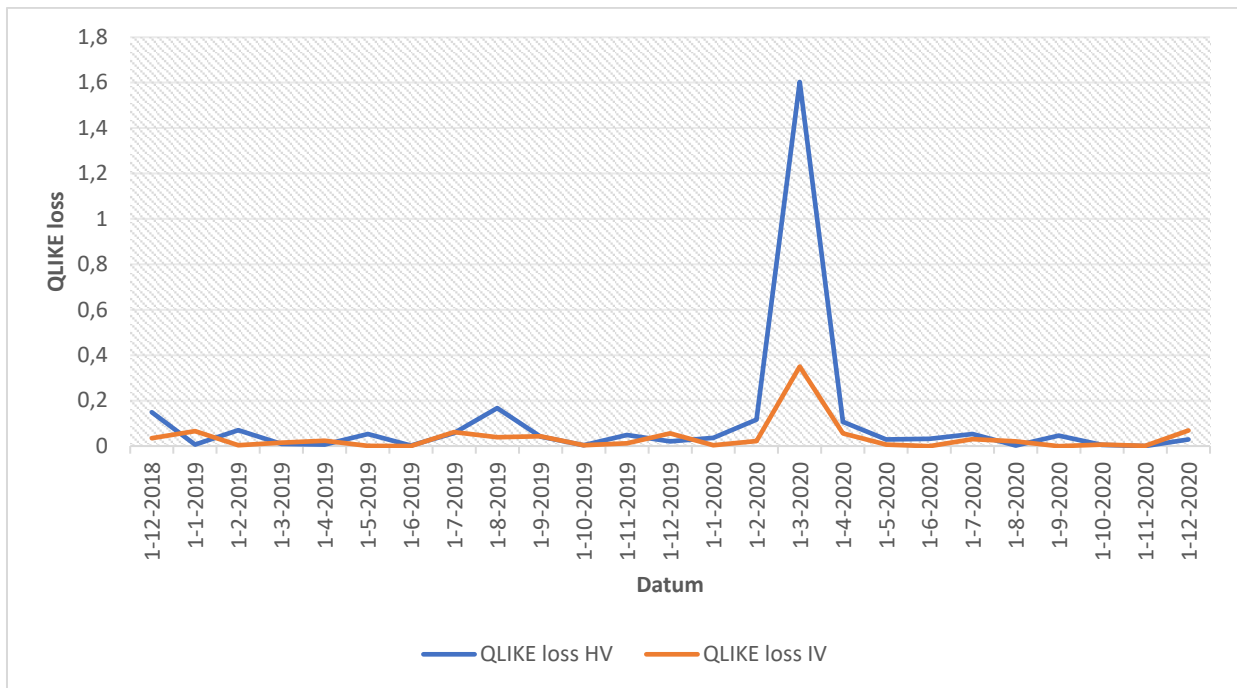
In figuur 1 heb ik een visuele weergave gecreëerd van de voorspellingen gedurende 2019 en 2020. Hier is goed te zien dat de QLIKE loss op basis van historische volatiliteit veel hoger doorschiet dan de QLIKE loss van impliciete volatiliteit. Mogelijk is de verklaring hiervoor te vinden in het korte termijngeheugen van het onderliggende AR(3) model.

Tabel 4: de mean square error van de impliciete- en historische volatiliteit over alle aandelen, waarbij de periode gedurende de Covid-19 pandemie vergeleken wordt met de periode 2019-2020

Sector	Panel A			Panel B		
	MSE IV	MSE HV	Vershil	MSE IV	MSE HV	Vershil
<i>Information technology</i>	0,0081	0,0111	<u>-0,0030</u>	0,0407	0,0881	<u>-0,0474*</u>
<i>Health Care</i>	0,0037	0,0053	<u>-0,0016*</u>	0,0274	0,0740	<u>-0,0466***</u>
<i>Consumer Discretionary</i>	0,0067	0,0085	<u>-0,0018*</u>	0,0481	0,1111	<u>-0,0631**</u>
<i>Communication Services</i>	0,0084	0,0093	<u>-0,0009</u>	0,0221	0,0628	<u>-0,0408*</u>
<i>Financials</i>	0,0021	0,0040	<u>-0,0019*</u>	0,0579	0,1563	<u>-0,0983***</u>
<i>Industrials</i>	0,0075	0,0091	<u>-0,0016*</u>	0,0467	0,1636	<u>-0,1169**</u>
<i>Consumer Staples</i>	0,0043	0,0056	<u>-0,0013*</u>	0,0201	0,0647	<u>-0,0447***</u>
<i>Utilities</i>	0,0008	0,0018	<u>-0,0009*</u>	0,0442	0,1559	<u>-0,1117***</u>
<i>Materials</i>	0,0053	0,0068	<u>-0,0015</u>	0,0287	0,0783	<u>-0,0496**</u>
<i>Real estate</i>	0,0086	0,0099	<u>-0,0013</u>	0,0273	0,0963	<u>-0,0690*</u>
<i>Energy</i>	0,0028	0,0055	<u>-0,0027*</u>	0,0598	0,2534	<u>-0,1926**</u>
Totaal	0,0043	0,0060	<u>-0,0016*</u>	0,0385	0,1186	<u>-0,0801**</u>

Overzicht van de mean square error (MSE) van zowel de impliciete- als historische volatiliteit ten opzichte van de daadwerkelijk gerealiseerde volatiliteit. Onderscheid is gemaakt tussen panel A: 1 Januari 2019 – 31 Januari 2020 en panel B: 1 Februari 2020 – 31 December 2020 (Covid-19 pandemie). Het verschil is tot stand gekomen

door MSE IV minus MSE HV te berekenen. De onderste regel geeft een totaal overzicht over alle meegenomen aandelen in het onderzoek. De significantieniveaus 5%, 1% en 0,1% voor de Diebold-Mariano (1995) toets in de laatste kolom zijn respectievelijk aangegeven met *, ** en ***.



In deze figuur is een QLIKE loss functie gehanteerd, welke kijkt naar de ratio tussen de voorspelde- en gerealiseerde volatiliteit in plaats van het verschil zoals bij MSE. Het voordeel van deze loss functie is dat de voorspellingen voor en na de piek in begin 2020 beter te interpreteren zijn.

Figuur 1: de QLIKE loss van de impliciete- en historische volatiliteit over alle aandelen, 2019-2020

6 Conclusie en discussie

In dit onderzoek heb ik getracht een antwoord te vinden op de volgende vraag: 'In welke mate zijn de impliciete volatiliteit en historische volatiliteit in staat een accurate voorspelling te vormen van de gerealiseerde volatiliteit van een selectie bedrijven uit de S&P-500 over een periode van 1996 tot en met 2020?' Hiervoor heb ik een analyse gemaakt van de MSE's van de impliciete- en historische volatiliteit, waarbij zowel de individuele waarden als de verhouding ten opzichte van elkaar van belang zijn geweest om tot de bevindingen te komen.

De resultaten geven aan dat de gemiddelde MSE van impliciete volatiliteit grofweg 50% lager is dan die van historische volatiliteit. Daarnaast is naar voren gekomen dat het verschil tussen de MSE's van impliciete- en historische volatiliteit het grootst is in de Financials sector en de individuele MSE's het laagst zijn in de Consumer Staples sector. Tenslotte heb ik waargenomen dat de gemiddelde MSE's hoger zijn gedurende de Covid-19 pandemie in vergelijking met de periode daarvoor. Vergeleken met de totale sample is het verschil tussen de MSE's ongeveer 8 keer groter gedurende de Covid-19 pandemie. Op basis van de Diebold-Mariano (1995) test zijn de verschillen tussen de voorspellingen over alle aandelen significant tegen een significantieniveau van 5%.

Met betrekking tot de hoofdvraag concludeer ik dat de impliciete volatiliteit op basis van deze sample in zijn totaal een betere voorspeller is van gerealiseerde volatiliteit dan historische volatiliteit. De mate waarin deze twee in staat zijn accuraat gerealiseerde volatiliteit te voorspellen verschilt aanzienlijk per sector. Dit geldt ook voor de prestatie van twee gebruikte methoden ten opzichte van elkaar. Tenslotte zorgt de aanwezigheid van de Covid-19 pandemie ervoor dat het moeilijker is de gerealiseerde volatiliteit te voorspellen in vergelijking met voorafgaande periode. Het verschil tussen de accuraatheid van voorspelling van de impliciete- en historische volatiliteit is daarnaast aanzienlijk hoger, in het voordeel van de impliciete volatiliteit.

De hoofdbevinding dat de impliciete volatiliteit beter in staat is gerealiseerde volatiliteit te voorspellen dan de historische volatiliteit is lijn met de gerelateerde literatuur van onder andere Christensen en Prabhala (1998). Echter betekent dit niet dat de gevonden resultaten volledig correct zijn. Zo is het voor de afsplitsing naar sectoren en de Covid-19 pandemie niet mogelijk om een vergelijking te maken met bestaande literatuur, wat het trekken van een veelzeggende conclusie moeilijk maakt. Enkele belangrijke kritiekpunten van dit onderzoek bestaan uit: niet corrigeren voor dividenden, geringe dataoverlap, gebruik maken van een relatief simpel model, insignificantie over bepaalde aandelen/sectoren en een beperkte dataset. Voor vervolgonderzoek raad ik dan ook aan de complete S&P-500 in overweging te nemen, een vorm van het HAR-RV model (of een ander model van vergelijkbaar kaliber) te hanteren in combinatie met een hogere datafrequentie en een correctie te

maken voor de overige genoemde gebreken. Daarnaast is het een optie huidig onderzoek uit te breiden door verschillende crises en periodes van hoge volatiliteit te vergelijken met de Covid-19 pandemie om na te gaan in hoeverre deze dezelfde kenmerken vertonen. In tegenstelling tot dit onderzoek is het dan van belang meer observaties mee te nemen, aangezien de Diebold-Mariano (1995) test op basis van 12 observaties mogelijk snel een vertekend beeld geeft.

In praktisch opzicht kan dit onderzoek informatie verstrekken met betrekking tot de vermogensallocatie van particuliere- en professionele beleggers. Zo kunnen de bevindingen aanleiding geven bepaalde sectoren te vermijden waarin het lastig is volatiliteit te voorspellen. Daarnaast biedt dit onderzoek bewijs dat het mogelijk beter is gedurende een periode van hoge volatiliteit (zoals de Covid-19 pandemie) niet actief te zijn in de aandelenmarkt.

Theoretisch gezien vormt deze paper een drijfveer tot onderzoek naar het voorspellend vermogen van impliciete- en historische volatiliteit in groter detail. Door de genoemde suggesties te onderzoeken en dit verder uit te breiden, kan een betrouwbaarder beeld worden geschetst van verschillen tussen sectoren en periodes van hoge volatiliteit. Kortgezegd geef ik met dit experiment aanleiding tot het onderzoeken van een vooralsnog bescheiden en mogelijk ontbrekende tak binnen de literatuur.

Appendix A

Tabel met geselecteerde aandelen per sector

Sector	Aandeel
Information Technology	Apple
	Microsoft
	Adobe Systems
	Intel
	Cisco Systems
Health Care	Johnson & Johnson
	UnitedHealth Group Incorporated
	Pfizer Inc.
	Merck & Co. Inc.
Consumer Discretionary	Eli Lilly and Company
	Amazon
	Home Depot
	McDonald's Corporation
	Nike
Communication Services	Lowe's Companies
	Comcast Corporation Class A
	Walt Disney Company
	AT&T Inc.
	Electronic Arts Inc.
Financials	Omnicom Group Inc.
	JPMorgan Chase & Co
	Bank of America Corp
	Wells Fargo & Company
	Citigroup Inc.
Industrials	Morgan Stanley
	United Parcel Service
	Boeing Company
	Raytheon Technologies Corporation
	Caterpillar Inc.
Consumer Staples	General Electric Company
	Procter & Gamble Company
	Coca-Cola Company
	PepsiCo

	Walmart Inc.
	Costco Wholesale Corporation
Utilities	Duke Energy Corporation
	Southern Company
	Dominion Energy Inc.
	Exelon Corporation
	American Electric Power Company
Materials	Sherwin-Williams Company
	Air Products and Chemicals
	PPG Industries
	Ecolab Inc.
	Newmont Corporation
Real Estate	Public Storage
	Simon Property Group
	Welltower
	AvalonBay Communities
	Weyerhaeuser Company
Energy	Exxon Mobil Corporation
	Chevron Corporation
	EOG Resources
	Schlumberger NV
	ConocoPhillips

Appendix B

Tabel met de mean square error van de impliciete- en historische volatiliteit verdeeld over sectoren inclusief onderliggende aandelen

Aandeel	MSE IV	MSE HV	Vershil
<i>Sector 1: Information Technology</i>	<i>0,0119</i>	<i>0,0190</i>	<i><u>-0,0071*</u></i>
Apple	0,0121	0,0189	-0,0068*
Microsoft	0,0077	0,0148	-0,0072*
Adobe Systems	0,0143	0,0227	-0,0084***
Intel	0,0120	0,0204	-0,0084*
Cisco Systems	0,0134	0,0180	-0,0046
<i>Sector 2: Health Care</i>	<i>0,0090</i>	<i>0,0141</i>	<i><u>-0,0064*</u></i>
Johnson & Johnson	0,0041	0,0079	-0,0039**
UnitedHealth Group Incorporated	0,0106	0,0246	-0,0140*
Pfizer Inc.	0,0056	0,0102	-0,0046**
Merck & Co. Inc.	0,0111	0,0160	-0,0049*
Eli Lilly and Company	0,0071	0,0118	-0,0048*
<i>Sector 3: Consumer Discretionary</i>	<i>0,0085</i>	<i>0,0129</i>	<i><u>-0,0045*</u></i>
Amazon	0,0208	0,0251	-0,0043
Home Depot	0,0080	0,0168	-0,0088*
McDonald's Corporation	0,0046	0,0097	-0,0051
Nike	0,0081	0,0146	-0,0065**
Lowe's Companies	0,0095	0,0182	-0,0087*
<i>Sector 4: Communication Services</i>	<i>0,0076</i>	<i>0,0156</i>	<i><u>-0,0080*</u></i>
Comcast Corporation Class A	0,0079	0,0156	-0,0077
Walt Disney Company	0,0072	0,0166	-0,0094**
AT&T Inc.	0,0036	0,0096	-0,0059*
Electronic Arts Inc.	0,0138	0,0213	-0,0075*
Omnicom Group Inc.	0,0054	0,0150	-0,0096**
<i>Sector 5: Financials</i>	<i>0,0210</i>	<i>0,0511</i>	<i><u>-0,0301*</u></i>
JPMorgan Chase & Co	0,0164	0,0259	-0,0095**
Bank of America Corp	0,0229	0,0471	-0,0242*
Wells Fargo & Company	0,0131	0,0278	-0,0147*

Citigroup Inc.	0,0297	0,0751	-0,0453
Morgan Stanley	0,0229	0,0798	-0,0569
<hr/>			
<i>Sector 6: Industrials</i>	<i>0,0076</i>	<i>0,0202</i>	<i><u>-0,0126*</u></i>
United Parcel Service	0,0045	0,0101	-0,0055*
Boeing Company	0,0110	0,0357	-0,0251
Raytheon Technologies Corporation	0,0080	0,0224	-0,0146**
Caterpillar Inc.	0,0070	0,0133	-0,0064*
General Electric Company	0,0069	0,0185	-0,0116*
<hr/>			
<i>Sector 7: Consumer Staples</i>	<i>0,0041</i>	<i>0,0093</i>	<i><u>-0,0052*</u></i>
Procter & Gamble Company	0,0030	0,0081	-0,0050*
Coca-Cola Company	0,0035	0,0084	-0,0049
PepsiCo	0,0044	0,0124	-0,0080
Walmart Inc.	0,0042	0,0078	-0,0036*
Costco Wholesale Corporation	0,0055	0,0099	-0,0043**
<hr/>			
<i>Sector 8: Utilities</i>	<i>0,0062</i>	<i>0,0171</i>	<i><u>-0,0109</u></i>
Duke Energy Corporation	0,0062	0,0158	-0,0095
Southern Company	0,0047	0,0160	-0,0112*
Dominion Energy Inc.	0,0058	0,0191	-0,0134
Exelon Corporation	0,0066	0,0145	-0,0078*
American Electric Power Company	0,0078	0,0204	-0,0126
<hr/>			
<i>Sector 9: Materials</i>	<i>0,0066</i>	<i>0,0135</i>	<i><u>-0,0069*</u></i>
Sherwin-Williams Company	0,0097	0,0166	-0,0070**
Air Products and Chemicals	0,0065	0,0141	-0,0076*
PPG Industries	0,0055	0,0140	-0,0140**
Ecolab Inc.	0,0084	0,0155	-0,0071
Newmont Corporation	0,0074	0,0198	-0,0124
<hr/>			
<i>Sector 10: Real estate</i>	<i>0,0124</i>	<i>0,0162</i>	<i><u>-0,0038</u></i>
Public Storage	0,0150	0,0145	0,0005
Simon Property Group	0,0239	0,0846	-0,0604
Welltower	0,0216	0,0269	-0,0046
AvalonBay Communities	0,0136	0,0205	-0,0069
Weyerhaeuser Company	0,0084	0,0230	-0,0145*

<i>Sector 11: Energy</i>	<i>0,0083</i>	<i>0,0249</i>	<i><u>-0,0166</u></i>
Exxon Mobil Corporation	0,0053	0,0173	-0,0120*
Chevron Corporation	0,0090	0,0291	-0,0201
EOG Resources	0,0111	0,0244	-0,0133
Schlumberger NV	0,0069	0,0284	-0,0214
ConocoPhillips	0,0094	0,0256	-0,0162*
Totaal	<i>0,0094</i>	<i>0,0194</i>	<i><u>-0,0102*</u></i>

Overzicht van de mean square error (MSE) van zowel de impliciete- als historische volatiliteit ten opzichte van de gerealiseerde volatiliteit. Onder elke sector staan de 5 geselecteerde aandelen uit de desbetreffende sector. De derde kolom is tot stand gekomen door MSE IV minus MSE HV te berekenen. De onderste regel geeft een totaal overzicht over alle meegenomen aandelen in het onderzoek.

Literatuurlijst

- Bandi, F. M., & Perron, B. (2006). Long Memory and the Relation Between Implied and Realized Volatility. *Journal of Financial Econometrics*, 4(4), 636–670. <https://doi.org/10.1093/jjfinec/nbl003>
- Beckers, S. (1981). Standard deviations implied in option prices as predictors of future stock price variability. *Journal of Banking & Finance*, 5(3), 363–381. [https://doi.org/10.1016/0378-4266\(81\)90032-7](https://doi.org/10.1016/0378-4266(81)90032-7)
- Black, F., & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3), 637–654. <https://doi.org/10.1086/260062>
- Brenner, Menachem, en Dan Galai. (1981). "The Properties of the Estimated Risk of Common Stocks Implied by Option Prices." Working paper #112, University of California-Berkeley.
- Busch, T., Christensen, B. J., & Nielsen, M. R. (2011). The role of implied volatility in forecasting future realized volatility and jumps in foreign exchange, stock, and bond markets. *Journal of Econometrics*, 160(1), 48–57. <https://doi.org/10.1016/j.jeconom.2010.03.014>
- Canina, L., & Figlewski, S. (1993). The Informational Content of Implied Volatility. *Review of Financial Studies*, 6(3), 659–681. <https://doi.org/10.1093/rfs/5.3.659>
- Chiras, D. P., & Manaster, S. (1978). The information content of option prices and a test of market efficiency. *Journal of Financial Economics*, 6(2–3), 213–234. [https://doi.org/10.1016/0304-405x\(78\)90030-2](https://doi.org/10.1016/0304-405x(78)90030-2)
- Christensen, B. J., & Hansen, C. S. (2002). New evidence on the implied-realized volatility relation. *The European Journal of Finance*, 8(2), 187–205.
- Christensen, B. J., & Prabhala, N. R. (1998). The relation between implied and realized volatility. *Journal of financial economics*, 50(2), 125–150.
- Corsi, F. (2008). A Simple Approximate Long-Memory Model of Realized Volatility. *Journal of Financial Econometrics*, 7(2), 174–196. <https://doi.org/10.1093/jjfinec/nbp001>
- Cox, J. C., Ross, S. A., & Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of Financial Economics*, 7(3), 229–263. [https://doi.org/10.1016/0304-405x\(79\)90015-1](https://doi.org/10.1016/0304-405x(79)90015-1)
- Diebold, F. X., & Mariano, R. S. (1995). Comparing Predictive Accuracy. *Journal of Business & Economic Statistics*, 13(3), 253. <https://doi.org/10.2307/1392185>
- Derman, E., & Kani, I. (1994). Riding on a smile. *Risk*, 7(2), 32–39.
- Dumas, B., Fleming, J., & Whaley, R. E. (1998). Implied Volatility Functions: Empirical Tests. *The Journal of Finance*, 53(6), 2059–2106. <https://doi.org/10.1111/0022-1082.00083>
- Dupire, B. (1994). Pricing with a smile. *Risk*, 7(1), 18–20.

- Engelhardt, N., Krause, M., Neukirchen, D., & Posch, P. N. (2021). Trust and stock market volatility during the COVID-19 crisis. *Finance Research Letters*, 38, 101873. <https://doi.org/10.1016/j.frl.2020.101873>
- Israelsen, C. (2021, 27 februari). *Sector by sector in the S&P 500 with ETFs*. etf.com. <https://www.etf.com/sections/etf-strategist-corner/sector-sector-sp-500?nopaging=1>
- IvyDB US Reference Manual. (2021). Wharton Research Data Services. https://wrds-www-wharton-upenn-edu.eur.idm.oclc.org/login/?next=/documents/1504/IvyDB_US_Reference_Manual_rn2hAXz.pdf
- Latane, H. A., & Rendleman, R. J. (1976). Standard Deviations of Stock Price Ratios Implied in Option Prices. *The Journal of Finance*, 31(2), 369. <https://doi.org/10.2307/2326608>
- Mayhew, S. (1995). Implied Volatility. *Financial Analysts Journal*, 51(4), 8–20. <https://doi.org/10.2469/faj.v51.n4.1916>
- Option Metrics. (2021). *Historical Options Market Data Provider | OptionMetrics*. <https://optionmetrics.com/about-us/>
- Rubinstein, M. (1994). Implied Binomial Trees. *The Journal of Finance*, 49(3), 771–818. <https://doi.org/10.1111/j.1540-6261.1994.tb00079.x>
- Schmalensee, R., & Trippi, R. R. (1978). COMMON STOCK VOLATILITY EXPECTATIONS IMPLIED BY OPTION PREMIA. *The Journal of Finance*, 33(1), 129–147. <https://doi.org/10.1111/j.1540-6261.1978.tb03394.x>
- Sectors - S&P Dow Jones Indices. (2021). Spglobal.Com. <https://www.spglobal.com/spdji/en/landing/investment-themes/sectors/>