

ERASMUS UNIVERSITY ROTTERDAM  
ERASMUS SCHOOL OF ECONOMICS

De relatie tussen de R&D intensiteit en de  
financiële prestaties van gezondheidsbedrijven in  
de S&P 500

**Erasmus  
University  
Rotterdam**



<b>Auteur:</b>	P.R. Bakker
<b>Student nummer:</b>	471215
<b>Scriptiebegeleider:</b>	Dr. J.J.G. Lemmen
<b>Tweede lezer:</b>	Yashvir Gangaram-Panday
<b>Finale datum:</b>	Juli 2021

# Abstract

In dit onderzoek wordt over de periode van 2010 tot en met 2019 de relatie tussen de R&D intensiteit en financiële prestaties van gezondheidsbedrijven uit de S&P 500 onderzocht. Uit dit onderzoek blijkt dat de R&D een negatieve invloed heeft op de financiële prestaties van een bedrijf wanneer op basis van de boekwaarden wordt gemeten, waarbij de grootte van het bedrijf dit negatieve effect verzwakt. Echter, de R&D heeft een positieve invloed op de financiële prestaties van een bedrijf wanneer een combinatie van boekwaarden en marktwaarden wordt gebruikt als proxy voor de financiële prestaties.

**Sleutelwoorden:** S&P 500, R&D intensiteit, ROA, gezondheid

**JEL Codes:** O32, I11, C12

# Inhoudsopgave

<b>Abstract</b>	<b>2</b>
<b>Figurenlijst</b>	<b>4</b>
<b>Tabellenlijst</b>	<b>5</b>
<b>Hoofdstuk 1 Introductie</b>	<b>6</b>
<b>Hoofdstuk 2 Literatuur en hypothesen</b>	<b>9</b>
<b>Hoofdstuk 3 Data</b>	<b>12</b>
<b>Hoofdstuk 4 Methodiek</b>	<b>15</b>
4.1 PooledOLS, Fixed-effect model, random-effect model	15
4.2 PooledOLS of FE- en RE model	16
4.2.1 Normaalverdeling en uitschieters	16
4.2.2 Heteroscedasticiteit	18
4.2.3 Autocorrelatie	21
4.3 FE- of RE model	21
4.4 Multicollineariteit	29
<b>Hoofdstuk 5 – Resultaten</b>	<b>30</b>
<b>Hoofdstuk 6 Conclusie</b>	<b>35</b>
<b>Bibliografie</b>	<b>39</b>
<b>Appendix A Normaal Verdeling</b>	<b>44</b>
<b>Appendix B Breusch-Godfrey Regressie</b>	<b>46</b>
<b>Appendix C Multicollineariteit VIF test</b>	<b>50</b>

# Figurenlijst

Figuur 1 Medische technologie R&D en Omzet 2000-2019	6
Figuur 2 Het spreidingsdiagram van de foutenterm van de regressie (afhankelijk: R&D intensiteit; onafhankelijk: de bedrijfsomvang)	18
Figuur 3 Het spreidingsdiagram van de foutenterm van de regressie (afhankelijk: R&D intensiteit; onafhankelijk: de w_brutowinstmarge)	19
Figuur 4 Het spreidingsdiagram van de foutenterm van de regressie (afhankelijk: w_ROA; onafhankelijk: R&D intensiteit; controlevariabelen: de brutowinstmarge en bedrijfsomvang)	20
Figuur 5 Het spreidingsdiagram van de foutenterm van de regressie (afhankelijk: Ln Tobin's q; onafhankelijk: R&D intensiteit; controlevariabelen: de brutowinstmarge en bedrijfsomvang)	20

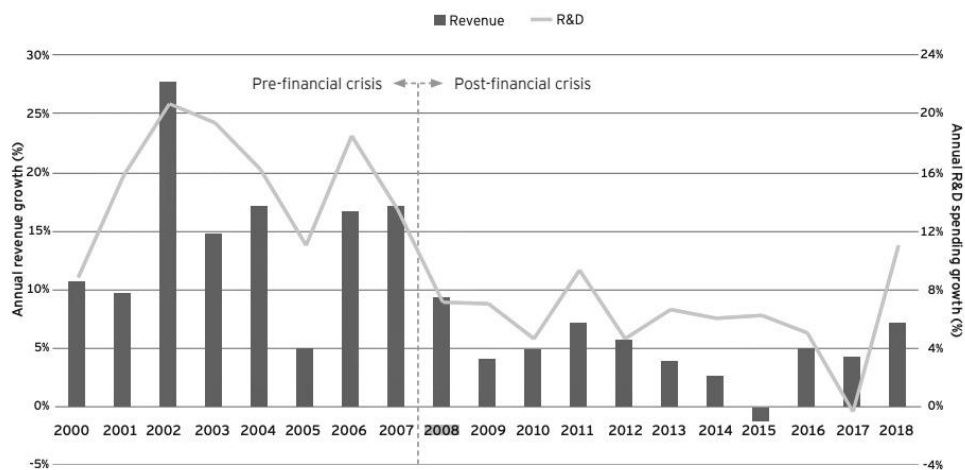
# Tabellenlijst

Tabel 1: Beschrijvende statistieken van gezondheidsbedrijven uit de S&P 500	14
Tabel 2: Regressie resultaat voor een RE (1) en FE (2) model van hypothese 1	22
Tabel 3: Regressie resultaat voor de Hausman-test van hypothese 1 met Ho: Verschil tussen de coëfficiënten is niet systematisch	23
Tabel 4: Regressie resultaat voor een RE (1) en FE (2) model van hypothese 2	23
Tabel 5: Regressie resultaat voor de Hausman-test van hypothese 2	24
Tabel 6: Regressie resultaat voor de Hausman-test van hypothese 3 met w_ROA als afhankelijke variabele en met de R&D intensiteit en de bedrijfsomvang als interactie-effect	25
Tabel 7: Regressie resultaat voor de Hausman-test van hypothese 3 met de natuurlijk logaritme van de Tobin's q als afhankelijke variabele	26
Tabel 8: Regressie resultaat voor een RE (1) en FE (2) model van hypothese 3 met de 98% winsorized ROA als afhankelijke variabele.	27
Tabel 9: Regressie resultaat voor een RE (1) en FE (2) model van hypothese 3 met de natuurlijk logaritme van de Tobin's q als afhankelijke variabele,	28
Tabel 10: Regressie resultaat voor een RE model van hypothese 1	30
Tabel 11: Regressie resultaat voor een RE model hypothese 2	31
Tabel 12: Regressie resultaat voor een FE model van hypothese 2 met w_ROA als afhankelijke variabele	33
Tabel 13: Regressie resultaat voor een RE model van hypothese 2 met Tobin's q als afhankelijke variabele	34

# Hoofdstuk 1 Introductie

De gezondheidskosten voor Amerikanen blijven stijgen. Er wordt verwacht dat de gemiddelde jaarlijkse groei van 2019 tot 2028 5.4% zal zijn. Als de kosten door blijven groeien in dit tempo, komt de groei uit boven de toename van het BBP van Amerika (Jones, 2020). De gezondheidsindustrie neemt ongeveer 1/5 in van het gehele BBP. Een deel van deze kosten wordt door gezondheidsbedrijven omgezet in *Research and Development* (R&D). R&D is nodig om tot nieuwe bevindingen te komen, maar ook essentieel om de concurrentie voor te blijven. R&D is een brede categorie die het geheel omschrijft van fundamenteel onderzoek, toegepast onderzoek en ontwikkelingsactiviteiten. In het algemeen staat R&D voor alle systematische activiteiten om kennis te vergroten en om deze kennis te gebruiken voor de ontwikkeling van nieuwe producten, processen of diensten (Kainulainen, 2014). In dit scriptieonderzoek wordt onderzocht wat de relatie is tussen de R&D intensiteit en de financiële prestaties van de gezondheidsbedrijven in de S&P 500. Financiële duurzaamheid is een brede definitie die wordt samengevat in drie elementen: economische prestatie, milieuprestatie en sociale prestatie (Bolívar, 2016). In dit onderzoek zal de economische prestatie centraal staan. Gezondheidsbedrijven zijn vaak grote bedrijven. De onderlinge concurrentie is zeer hoog. De gezondheidsmarkt is ook sterk verzadigd en door en het stelsel van wettelijke licenties en regels is het lastig om toe te treden. Door deze factoren speelt R&D een strategische rol. Er is al veel onderzoek gedaan naar de relatie tussen R&D en financiële prestaties van ondernemingen, echter nog niet op het gebied van de gezondheid. Dit scriptieonderzoek geeft verrassende inzichten en breidt de kennis uit over de relaties tussen de R&D en de financiële prestaties van bedrijven.

De aanleiding voor dit onderzoek is onder meer de coronapandemie die heeft laten zien dat R&D essentieel is in de gezondheidsindustrie om de bevolking te beschermen. Wereldwijd werden de financiële prestaties van bedrijven op de proef gesteld. Op grote schaal werd er geïnvesteerd in onderzoek naar een coronavaccin. Uit een onderzoek bleek dat overheden grofweg 93 miljard euro hebben uitgegeven aan onderzoek voor de ontwikkeling van het coronavaccin (Hoecklin, 2021).



Figuur 1 Medische technologie R&D en Omzet 2000-2019

In 2018 bedroegen de totale Amerikaanse R&D investeringen in de medische en de gezondheidsindustrie 194,2 miljard dollar. Er werd verwacht dat de gezondheidsindustrie in 2020 nummer 1 zou worden in de uitgaven aan R&D (Lagasse, 2019). In Figuur 1 is te zien dat de R&D uitgaven in de medische technologische sector in 2019 met 11.5% groeiden ten opzichte van 8.1% in 2018 (Welch, 2020). R&D uitgaven zijn tijdens de corona pandemie sterk gegroeid en een verklaring hiervoor is het financieel bestendig en daarmee duurzaam blijven tijdens deze crisis. Financiële prestatie speelt in deze sector een cruciale rol om sterker uit de crisis te komen. Daarnaast speelt R&D een belangrijke rol in het vergroten van de kennis en het ontdekken van nieuwe producten en diensten om zo een grotere toekomstige zekerheid te beiden. In deze studie wordt de vraag beantwoord: *Wat is de invloed van R&D op de financiële prestatie van bedrijven in de gezondheidsindustrie in de S&P 500?*

Er is eerder al onderzoek gedaan naar de impact van R&D op de financiële prestaties van een bedrijf. Echter door de brede definitie van R&D en financiële prestaties werden er verschillende maatstaven gebruikt. Belderbos et al. (2004) gebruikten arbeidsproductiviteit en de productiviteit om nieuwe innovatieve verkopen op de markt te brengen als maatstaaf voor de prestatie van een bedrijf. Artz et al. (2010) gebruikten het aantal patent aanvragen en het aantal nieuw aangekondigde producten als maatstaaf voor R&D investeringen. Vervolgens onderzochten ze hoe het bedrijf profiteert van deze investeringen door de *return on assets* (ROA) en verkoopgroei als maatstaaf te gebruiken voor de prestatie van het bedrijf. Sher en Yang (2005) benoemden de ROA, *return on sales* (ROS) en *return on equity* (ROE) als een maatstaaf voor de financiële prestaties van een bedrijf en gebruikten de ROA in hun onderzoek. Daarnaast gebruikten ze onder andere het aantal patent aanvragen, de R&D intensiteit<sup>1</sup> en de R&D mankracht<sup>2</sup> als maatstaf voor de innovatieve input van het bedrijf. Hung en Chou (2013) gebruiken de Tobin's q<sup>3</sup> in plaats van de ROA of ROS als maatstaaf voor de prestaties van een bedrijf.

Doordat er verschillende maatstaven werden gebruikt, zijn er ook verschillende uitkomsten gevonden. Shin et al. (2009) vonden dat een toename in de R&D uitgaven een significant positief effect heeft op de brutowinst, maar dat die toename niet resulteert in een hogere ROE of ROA. Artz et al. (2010) vonden een positieve relatie tussen de R&D-uitgaven en het aantal patentaanvragen, maar een negatieve relatie tussen het aantal patentaanvragen en ROA en de groei van verkopen. Sher en Yang (2005) vonden dat een hogere R&D intensiteit en R&D mankracht een positieve relatie heeft op de ROA. Daarnaast vonden ze een negatieve insignificante relatie tussen het aantal patentaanvragen en de ROA. Kotabe et al. (2002) vonden onder andere dat de R&D-intensiteit een positief effect heeft op de ROA. Connolly en Hirschey (2005) onderzochten de relatie tussen de grootte van het bedrijf en het effect van R&D op de Tobin's q. Zij vonden een duidelijk verschil in de effectiviteit van R&D uitgave

---

<sup>1</sup>  $R\&D\ intensiteit = \frac{R\&D\ uitgaven}{totale\ activa}$

<sup>2</sup>  $R\&D\ mankracht = \frac{voltijd\ R\&D\ werknemers}{totaal\ aantal\ werknemers}$

<sup>3</sup>  $Tobin's\ q = \frac{marktwaarde\ van\ het\ eigen\ vermogen + marktwaarde\ van\ de\ schulden}{boekwaarde\ van\ de\ totale\ activa}$

wanneer deze afhangt van de bedrijfsomvang. Een dollar uitgegeven aan R&D voegt significant meer toe aan de Tobin's  $q$  voor grote bedrijven dan voor kleine bedrijven. Bardhan et al. (2013) vonden een positieve relatie tussen de R&D-uitgaven en de Tobin's  $q$ . Pargaridis en Varsakelis (2010) vonden een sterkere relatie tussen de R&D-intensiteit en de Tobin's  $q$  voor kleinere bedrijven dan voor grotere. Uit deze resultaten blijkt een negatief interactie-effect tussen de variabelen bedrijfsomvang en de R&D intensiteit als effect op de Tobin's  $q$ . Een interactie-effect is het gelijktijdige effect van twee of meer onafhankelijke variabelen op ten minste een afhankelijke variabele waarbij hun gezamenlijke effect significant groter of kleiner is dan de som van de aparte onderdelen (Lavrakas, 2008).

De uitkomsten van dit scriptieonderzoek zijn verschillend. De eerste hypothese, waarbij verondersteld wordt dat de bedrijfsomvang een negatief effect heeft op de R&D intensiteit, wordt niet verworpen. De tweede hypothese, waarbij verondersteld wordt dat de brutowinstmarge een positief effect heeft op de R&D intensiteit, wordt wel verworpen. De derde hypothese en tevens het antwoord op de hoofdvraag, gaf een tweezijdig antwoord. Hier wordt verondersteld dat de R&D intensiteit een positief effect heeft op de ROA en Tobin's  $q$ . De bedrijfsomvang bleek een significant negatief effect te hebben op de R&D intensiteit. Dat is te verklaren doordat grote bedrijven niet investeren in specifieke R&D, maar eerder een bedrijf zullen overnemen dat al beschikt over deze technologie of innovatieve kennis. Daarnaast bleek de brutowinstmarge een negatief effect te hebben op de R&D. Een bedrijf maakt namelijk een afweging tussen de verdeling van kosten. Een afname van de brutowinstmarge gaat vaak gepaard met een toename van andere kosten. In dit geval zal dit een toename van R&D-uitgaven betekenen. Het resultaat hiervan is een negatieve relatie tussen de brutowinstmarge en de R&D-uitgaven. De hoofdvraag werd beantwoord met een tweezijdig antwoord. De R&D heeft een negatieve invloed op de financiële prestaties van een bedrijf wanneer op basis van de boekwaarden wordt gemeten (ROA), maar niet op basis van marktwaarden en boekwaarden (Tobin's  $Q$ ). R&D heeft echter een positieve invloed op de financiële duurzaamheid van een bedrijf wanneer een combinatie van boekwaarden en marktwaarden wordt gebruikt als proxy voor de financiële duurzaamheid.

Allereerst zullen alle hypothesen in dit onderzoek onderbouwd worden. Daarna worden de databron en bewerking van de data besproken. Vervolgens worden de methodiek en het testen van de hypothesen behandeld. Daarna volgen de resultaten en de conclusies. Tot slot zijn bewerkingen en aanbevelingen voor dit onderzoek opgenomen.



## Hoofdstuk 2 Literatuur en hypothesen

Dit onderzoek bouwt voort op verschillend empirisch onderzoek en sluit hierbij aan bij de voorgaande benoemde maatstaven. In dit onderzoek zal daarom gebruik worden gemaakt van de ROA en de Tobin's q als maatstaf voor de financiële prestaties van een bedrijf. Daarnaast zal de R&D intensiteit, met als controlevariabelen de bedrijfsomvang en de brutowinstmarge, de maatstaf zijn voor de mate van R&D.

Ten eerste zal onderzocht worden wat de invloed is van de omvang van het bedrijf op de R&D intensiteit. Cohen et al. (1987) vonden dat de bedrijfsomvang een klein insignificant effect op de R&D intensiteit van een bedrijf heeft. Choi en Lee (2017) vonden een positief significant effect tussen de bedrijfsomvang en de uitgaven aan R&D. Daarentegen vond Doi (1994) een negatieve relatie tussen de omvang van het bedrijf en zijn R&D uitgaven. Shefer en Frenkel (2005) vonden in een empirisch onderzoek een sterke negatieve relatie tussen de bedrijfsomvang en de mate van R&D investeringen, maar benadrukken dat uit eerder onderzoek bleek dat er juist een positieve relatie werd gevonden. Shefer en Frenkel (2005) vonden echter een negatieve relatie die volgens hun juist te danken is aan een grote toename van het aantal kleine bedrijven aan innovatieve activiteiten. Een gezondheidsbedrijf moet jaarlijks ongeveer 2 tot 4 miljard investeren in R&D om op een 'normaal' tempo mee te groeien met de markt (Alvaro et al., 2010). Veel grote gezondheidsbedrijven beschikken vaak over meer financiële middelen om mee te investeren en investeren niet in specifieke R&D, maar zullen eerder een bedrijf overnemen dat al beschikt over de technologie of innovatieve kennis. Alleen al in de periode van 1995 tot 2015 werden er bijna 50 fusies en overnamedeals gesloten met een waarde van meer dan 10 miljard dollar. In de periode van 2000 tot 2010 werden ongeveer 1.350 fusies en overname transacties gesloten in de gezondheidsindustrie met een totale waarde van ongeveer 700 miljard dollar (Irving Levin Associates, 2010). Daarom wordt in dit onderzoek ook verwacht dat een groter bedrijf een kleinere R&D intensiteit heeft. In dit onderzoek zal de volgende eerste hypothese worden onderzocht: *De bedrijfsomvang heeft een significant negatief effect op de R&D intensiteit.*

Ten tweede wordt de invloed van de brutowinstmarge op de R&D-uitgaven onderzocht. Dave et al. (2013) vonden een positieve relatie tussen de brutowinstmarge en de R&D-intensiteit. Wel benadrukten zij dat de vertraging van de impact van de R&D-investering verschillend is per bedrijf en dat dit voor complicaties zorgt. Toch is de positieve relatie waarschijnlijk te danken aan het feit dat bedrijven investeren in onderzoek dat de brutowinstmarge van een bedrijf verbetert. De brutowinstmarge wordt verbeterd door kostenreductie door het toepassen van innovatieve technologieën. Chen et al. (2005) vonden een sterke positieve relatie tussen de R&D-intensiteit en de toename in winst. Grabowski en Mueller (1978) vonden in hun onderzoek ook een positieve relatie tussen de mate van R&D-investeringen en de winstgevendheid van het bedrijf. Echter, Akben-Selcuk (2016) vond een negatieve relatie tussen de brutowinstmarge en R&D uitgaven. Hij benadrukt dat een

toename van de brutowinstmarge vaak ook gepaard gaat met een afname in andere uitgaven. Echter, wordt in dit onderzoek verondersteld dat een toename in de brutowinstmarge meer financiële ruimte biedt voor innovatie. In dit scriptieonderzoek worden de volgende tweede hypothese onderzocht: *De brutowinstmarge heeft een significant positief effect op de R&D intensiteit van een bedrijf.*

Ten derde wordt de hoofdvraag onderzocht: wat is de invloed van de R&D intensiteit op de ROA en op de Tobin's q? De ROA is een maatstaf voor de winstgevendheid van een bedrijf ten opzichte van zijn totale activa en wordt berekend door het netto-inkomen te delen door de boekwaarde van de totale activa. De ROA wordt berekend op basis van de boekwaarde van het bedrijf en de Tobin's q daarentegen is een combinatie van marktwaarde en boekwaarde componenten. Volgens Sauaia & Castro (2014) laat de Tobin's q de verhouding zien tussen de marktwaarde van de totale activa van een bedrijf en de vervangingswaarde van die activa. Door het gebruik van de Tobin's q in economie en *finance* kan niet alleen onderzocht worden wat de prestaties in het bedrijf zijn (op basis het verleden), maar kan deze variabele ook wijzen op groeimogelijkheden. Als de Tobin's q een waarde boven de één aanneemt, dan waardeert de markt het bedrijf hoger dan de boekwaarde laat zien en kan het bedrijf in waarde toenemen door meer kapitaal aan te nemen en daarmee investeringen te doen. Wanneer de Tobin's q kleiner is dan één, dan is het bedrijf op de markt minder waard dan de vervangingskosten van de totale activa en zal het bedrijf zijn afgeschreven activa niet vervangen en daarmee een afname in kapitaal realiseren. Connolly en Hirschey (2005), zoals eerder benoemd, vonden een grotere effectiviteit van R&D uitgaven voor grotere bedrijven dan voor kleinere bedrijven. De verklaring hiervoor zijn eventuele schaalvoordelen in de productie of marketing van R&D intensieve goederen, geografisch bereik of superieure financiële middelen van grote bedrijven ten opzichte van kleine bedrijven. Bardhan et al. (2013) vonden een positieve relatie tussen de R&D uitgaven en de Tobin's q. Zij beargumenteren dat de Tobin's q een toekomstgerichte maatstaf is voor prestaties van een bedrijf. Het is niet een schatting op basis van de boekwaarde, maar ook van de markt en kan daarmee al een deel van het effect van een R&D investering laten zien. Vithessonthi en Racela (2016) vonden dat de R&D intensiteit op korte termijn een negatief effect heeft op de prestaties van een bedrijf. Bij een gegeven budget moet een bedrijf een afweging maken tussen uitbuitende en verkennende activiteiten. Op de korte termijn zal een investering in R&D een negatief effect geven op de ROA van een bedrijf, omdat dit effect pas op de lange termijn merkbaar zijn. Yeh et al. (2010) vonden een positieve relatie tussen de R&D intensiteit en de ROA voor een bepaalde intensiteitsgrens en een negatieve relatie na die grens. Zij benadrukken dat investeren in R&D ook risico's met zich meebrengt, omdat investeren in innovatie een kostbare activiteit is die niet altijd garant staat voor potentiële winst. Kotabe et al. (2002) onderzochten ook de relatie tussen de R&D intensiteit en de ROA als basis voor hun verdere bevindingen. Zij vonden een R&D intensiteitsgrens van 3.06% vanaf wanneer internationale bedrijven een positieve relatie vinden tussen de R&D intensiteit en de ROA. Pantagakakis et al. (2012) vonden een positieve correlatie tussen de R&D investering en de prestaties van een bedrijf op de markt, echter geldt dit niet wanneer gekeken wordt naar de relatie met de ROA. Dan werd een negatieve relatie gevonden en de verklaring hiervoor

was vooral te vinden in de vertraging van het effect van de R&D investering. In dit onderzoek zal de derde hypothese worden onderzocht: *De R&D intensiteit van een bedrijf heeft een significant positief effect op de ROA en Tobin's q van een bedrijf.*

## Hoofdstuk 3 Data

In dit onderzoek worden data gebruikt van gezondheidsbedrijven in de Standard & Poor's 500 Index (S&P 500). De S&P 500 is een gewogen index<sup>4</sup> op basis van de marktkapitalisatie van de 500 grootste beursgenoteerde bedrijven in de Verenigde Staten die verhandeld worden op de New York Stock Exchange (NYSE) of de NASDAQ. De 10 zwaarstwegende aandelen in de S&P 500 zijn in totaal goed voor 20% van de totale index (Santen, 2021). De S&P 500 is een verzameling van verschillende industrieën en de gezondheidsindustrie staat op nummer twee als een na grootste portie met 12.40% (62 van de 500 bedrijven) van de totale index na informatietechnologie (IT). Johnson & Johnson, UnitedHealth Group inc en Pfizer Inc zijn hierbij de drie grootste bedrijven.

De data voor dit onderzoek is verzameld door middel van de balans en de winst- en verliesrekening van 62 gezondheidszorgbedrijven uit de S&P 500 met behulp van Compustat data. Per bedrijf zijn de ROA, totale activa, R&D uitgaven, de brutowinstmarge van 31/12/2010 tot en met 31/12/2019 verzameld. Deze data zijn afkomstig van een Compustat database die verkregen is via *Wharton Research Data Service* (WRDS). Compustat Xpressfeed is een wereldwijd systeem van Standard & Poor's dat data levert (Standard & Poor's, 2011). Specifiek werd voor dit onderzoek de Noord Amerikaanse Compustat database gebruikt. Dit is een database voor Noord Amerikaanse en Canadese bedrijven die fundamentele en marktinformatie over actieve en inactieve beursgenoteerde bedrijven biedt. De database biedt per kwartaal 100 en jaarlijks 300 resultatenrekeningen, balansen, kasstroomoverzichten en aanvullende gegevens. Om alle variabelen te vinden werden twee datasets met elkaar gecombineerd. De *Financial Ratios Firm Level* dataset werd gebruikt om de ROA en de brutowinstmarge van alle bedrijven te verzamelen. Vervolgens werd de *Compustat Daily Updates – Fundamental Annual* dataset gebruikt om de totale activa, het totale eigen vermogen, de uitstaande aandelen, de slotkoers van de aandelen en de R&D uitgaven per bedrijf te verzamelen.

De *Financial Ratios Firm Level* bestaat uit maandelijks data en de *Fundamental Annual* data bestaat uit jaarlijkse data elke keer op 31 dec. Deze data is aangepast naar jaarlijkse data en vervolgens zijn de bedrijven verwijderd met ontbrekende data. Viatris is niet aanwezig in de dataset en daarnaast zijn Zoetis, Iqvia Holdings, Catalent, UnitedHealth Group, Steris Corporation, Medtronic, Anthem, Perrigo Company, Zimmer Biomet Holdings, DexCom, AbbVie, HCA Healthcare, Centene Corporation en Viatic verwijderd uit de dataset, omdat er te weinig data beschikbaar was. Na het aanpassen van de WRDS Financial Ratios Firm level dataset werden vervolgens de data van Compustat Daily updates aan die dataset toegevoegd. Hierdoor ontstond er één dataset bestaande uit alle benodigde variabelen. Daar bleek vervolgens dat voor LabCorp, CarinalHealth, Humana Inc., Cigna Corporation, Teleflex Incorporation er geen data beschikbaar was voor de R&D uitgaven. Daarnaast hebben Henry Schein, DaVita incorporation, AmerisourceBergen, Universal Health Services en Cvs Group geen R&D

---

<sup>4</sup> S&P 500 gewogen index =  $\frac{\text{Marktkapitalisatie van het bedrijf}}{\text{Totale marktkapitalisatie van alle bedrijven uit de S\&P 500}}$

uitgaven en zijn deze bedrijven ook verwijderd uit de dataset.

In totaal bestaat de dataset nu uit 380 data punten en 38 bedrijven. In dit onderzoek zullen volgende variabelen worden meegenomen:

- ROA
- Tobin's q
- Brutowinstmarge
- R&D intensiteit
- Bedrijfsomvang

Om de bedrijfsomvang te berekenen wordt een natuurlijk logaritme genomen van de totale activa van een bedrijf (1).

$$\text{Bedrijfsomvang} = \ln(\text{totale activa}) \quad (1)$$

Net als in eerder onderzoek gedaan op dit gebied zal de R&D intensiteit worden berekend door de R&D uitgaven te delen door de totale activa van een bedrijf (Padgett & Galan, 2010; Lin et al., 2006; Hitt et al., 1989; Munster, 2017) (3).

$$\text{R\&D intensiteit} = \frac{\text{R\&D uitgaven}}{\text{totale activa}} \quad (2)$$

De Tobin's q zal berekend worden aan de hand van een formule van Chung en Pruitt (1994). De Tobin's q is een combinatie van marktwaarde en boekwaardecomponenten en wordt berekend door de marktwaarde de totale activa te delen door de boekwaarde van de totale activa. Echter, is het een veelal gebruikte aanname om te veronderstellen dat de boekwaarde van het vreemd vermogen gelijk is aan de marktwaarde van het vreemd vermogen. Daarom wordt de Tobin's q in dit onderzoek berekend door de marktwaarde van het eigen vermogen en de boekwaarde van het vreemd vermogen bij elkaar op te tellen en te delen door de boekwaarde van de totale activa (3). De boekwaarde van het vreemd vermogen wordt berekend door de boekwaarde van het eigenvermogen af te trekken van de boekwaarde van de totale activa. De marktwaarde van het eigen vermogen wordt berekend door de uitstaande aandelen te vermenigvuldigen met de slotkoers van het aandeel.

$$\text{Tobin's } q = \frac{(\text{uitstaande aandelen} * \text{aandelenkoers}) + \text{boekwaarde vreemd vermogen}}{\text{boekwaarde totale activa}} \quad (3)$$

Tabel 1: Beschrijvende statistieken van gezondheidsbedrijven uit de S&P 500

	Gemiddelde	Mediaan	Min	Max	Scheefheid	Standaardafwijking
<i>ROA</i>	,161	,154	-,364	,626	-,391	,094
<i>Tobin's Q</i>	3,774	2,704	,83	23,964	2,632	3,087
<i>Brutowinstmarge</i>	,633	,658	-,657	,964	-1,944	,231
<i>R&amp;D intensiteit</i>	,082	,054	,001	,637	2,935	,094
<i>Bedrijfsomvang</i>	22,938	2,748	18,695	25,996	,014	1,525
<i>Totale activa</i>	2.5640,148	7.574,55	131,588	195.014	2,368	38.733,938
<i>R&amp;D * bedrijfsomvang</i>	1,841	1,24	,03	12,49	2,686	1,984
<i>R&amp;D * Brutowinstmarge</i>	,046	,035	-,0262	,384	1,353	,056

\* Totale activa in miljoenen US Dollars

# Hoofdstuk 4 Methodiek

## 4.1 PooledOLS, Fixed-effect model, random-effect model

De hypothesen zullen aan de hand van een paneldata regressieanalyse beantwoord worden. Allereerst wordt in STATA 16.1 per variabele worden onderzocht of die voldoet aan alle aannames voor een juiste paneldata regressieanalyse. Een paneldataset is een dataset die verschillende individuen over een bepaalde periode observeren en daardoor verschillende observaties van hetzelfde individu bevat (Hsiao, 2014). Daarnaast zijn er ook meerdere variabelen voor ieder individu (bedrijf) geobserveerd. Er is sprake van een combinatie tussen cross-sectionele data en tijdreeks data. Meerdere variabelen voor individuen over een bepaalde tijdreeks. Er zijn drie soorten technieken die kunnen worden toegepast op paneldata: *PooledOLS* regressie, *fixed-effect* (FE) model, *random-effect* (RE) model.

De eerste techniek is het toepassen van een *Ordinary Least Squared* (OLS) regressie op paneldata. Dit is een veel gebruikte regressie voor cross-sectionele data. Echter wanneer een OLS regressie wordt toegepast op paneldata worden de tijdeffecten en individuele eigenschappen van de observaties buiten beschouwing gelaten en wordt er vooral gefocust op de relatie tussen de variabelen. Wanneer een enkele variabele voor een individu over een langere tijd wordt geobserveerd zal er hoogstwaarschijnlijk sprake zijn van autocorrelatie in de foutterm. De data van een jaar ervoor zal samenhangen met de huidige data van een bedrijf. *PooledOLS* zal daarom grote standaardfouten hebben, omdat er verschillende aannames geschonden worden en hierdoor bestaat er een kans dat er verkeerde conclusies worden getrokken. Naast *PooledOLS* heb je nog een FE of RE model.

Een FE model beschouwt de individuele effecten van niet observeerbare en onafhankelijke variabelen als constant over tijd. Binnen een FE model mag er wel een relatie bestaan tussen de niet observeerbare onafhankelijke variabele en de endogene variabele. Het tijdsinvariantie gedeelte uit de foutterm wordt dan opgenomen in de regressie als constante ( $a_i$ ). Dit constante gedeelte zou hetzelfde effect en waarde nu moeten hebben als dat het een aantal tijdsperiodes geleden had. Als je de gemiddelde waarden van elke vergelijking aftrekt dan zou die constante nul moeten zijn. Het enige dat dan overblijft is een idiosyncratische foutterm die volledig tijdsinvariant is. Het effect is dan constant (*fixed*). Hierdoor is het model robuust tegen heterogeniteit aanwezig in het model. Echter beschouwt het model alle individuele effecten als constant en kan daardoor alleen afhankelijkheid aantonen per individu en niet tussen de individuen.

Een RE model beschouwt individuele effecten van niet observeerbare en onafhankelijke variabelen als willekeurig over tijd. Het RE model is een afweging tussen een OLS model en FE model. Net als bij het FE model is er sprake van een constante ( $a_i$ ) die het tijd invariante gedeelte van de foutterm weergeeft. Vervolgens kijkt het model naar de covariantie tussen de constante en de endogene variabele. Wanneer de covariantie groter is dan nul zal het FE model worden toegepast.

Wanneer die covariantie heel klein of nul is dan is er geen correlatie tussen de variabelen en zal een OLS regressie worden toegepast. Echter wanneer een OLS regressie wordt toegepast zal er nog steeds sprake zijn van autocorrelatie in de foutterm. Het model maakt gebruik van de term  $\lambda$  en die berekent simpelweg hoe groot de variantie is van de niet observeerbare heterogeniteit, dus de variantie van de  $a_i$ . Als die nul is dan geldt dat de variantie van  $a_i$  nul is en *PooledOLS* toegepast kan worden. Het model kan ‘wisselen’ tussen beide soorten regressies en daardoor niet alleen de afhankelijkheid aantonen per individu, maar ook tussen de individuen.

Om te kiezen welk model het meest geschikt is, wordt een Hausman-test gebruikt. De Hausman-test is een test voor endogeniteit en test of de geschatte coëfficiënten van een FE model en een RE model significant van elkaar verschillen. Hierbij wordt de nulhypothese getest dat de covariantie tussen de endogene variabele en de tijd invariante constante  $a_i$  gelijk is aan nul. Als de nulhypothese niet wordt verworpen dan wordt een RE model geprefereerd en wanneer deze wel wordt verworpen een FE model. Zoals eerder al is genoemd, kan een FE model alleen de afhankelijkheid aantonen per individu en het RE model daarnaast ook de afhankelijkheid aantonen tussen de individuen. Een FE model houdt rekening met alle biases van weggelaten variabelen op een groepsniveau, dus in dit geval bedrijfsniveau. Een RE model kijkt ook naar de afhankelijkheid tussen de individuen en houdt niet automatisch rekening met alle biases van weggelaten variabelen. Wanneer alle significante covarianten op groepsniveau in het RE model worden opgenomen, levert dat bij benadering dezelfde resultaten, maar de standaardfouten zijn dan kleiner bij een RE model. Indien de coëfficiënten van een FE- en RE model significant niet van elkaar verschillen, dan wordt een RE model geprefereerd, omdat deze kleinere standaardfouten heeft. Als dit niet het geval is, dan moet worden een FE model geprefereerd, want gezien het feit dat ze verschillen betekent dit dat er een weggelaten bias van een variabele is op een hoger niveau waar een RE model geen rekening mee houdt.

## 4.2 PooledOLS of FE- en RE model

Er zijn vijf aannames waaraan moet worden voldaan bij een juiste OLS regressie: willekeurige steekproef, lineair verband, exogeniteit, homoscedasticiteit van de variantie van de foutterm (geen autocorrelatie) en geen multicollineariteit. Wanneer de aannamen van exogeniteit en die van homoscedasticiteit van de variantie van de foutterm (autocorrelatie) worden geschonden, dan kan er beter gebruik worden gemaakt van een FE model of van een RE model. Deze modellen zijn robuust wanneer deze aannames geschonden worden. Om te bepalen of PooledOLS of FE- en RE-model toegepast zal worden zijn er dus twee aannames essentieel om te bepalen welk model gekozen zal worden en deze aannames zullen allereerst getest worden. De aanname voor homoscedasticiteit van de foutterm zal bij paneldata worden uitgebreid met de aanname dat er geen sprake is van autocorrelatie in de fouttermen. Normaal zal hier geen sprake van zijn bij cross-sectionele data, omdat er geen



tijdreeks data aanwezig is voor dezelfde individuen. Daarnaast zal moeten gelden dat de controlevariabelen normaal verdeeld zijn, anders zal er geen juiste conclusie getrokken worden.

#### 4.2.1 Normaalverdeling en uitschieters

Ten eerste moeten alle controlevariabelen normaal verdeeld zijn. Bij de controlevariabelen “Totale activa” en de “Brutowinstmarge” zal gekeken worden of ze normaal verdeeld zijn. Wanneer een histogram wordt geplot van de variabele “Totale activa” dan is te zien dat hij een rechtsscheefheid vertoont. Door de natuurlijk logaritme te nemen van de variabele “Totale activa” ontstaat er een normalere verdeling. Echter door de natuurlijk logaritme te nemen van de variabele “Totale activa” bereken je de variabele (zie Appendix A). “Bedrijfsomvang” en daarom zullen deze twee variabelen niet samen in een regressie worden gebruikt. Wanneer een histogram wordt geplot van de variabele “Brutowinstmarge” dan is te zien dat hij een linksscheefheid vertoont (zie Appendix A). De variabele kan getransformeerd worden naar een normale verdeling door het kwadraat te nemen van de variabele. Dit kan echter niet wanneer een variabele ook negatieve waardes bevat, omdat het kwadraat van een negatief getal positief is en daardoor een verkeerd beeld geeft. Er is een aantal bedrijven waarbij dit het geval is, dus zal de variabele *gewinsorized* worden. Het *winsorizen* van data is de invloed van uitschieters in de dataset te verminderen door niet de uitschieters te verwijderen, maar deze te vervangen voor een waarde die dicht bij de andere datapunten in de set zit (Lusk et al., 2011). Een 95% *winsorize* betekent een transformatie van de 2,5% onderste en 2,5% bovenste waardes in de dataset. De variantie neemt hierbij af, maar de observaties verminderen niet. Na een 90% *winsorize* van de variabele brutowinstmarge lijkt die normaler verdeeld (zie Appendix A).

Ten tweede moeten de afhankelijke en onafhankelijke variabelen geen last hebben van extreme uitschieters. De afhankelijke variabelen in dit onderzoek zijn de variabelen “ROA” en “Tobin’s q”. Wanneer een histogram geplot wordt van de variabele “ROA” dan laat deze een sterke verdeling in het midden zien met al snel een afvlakking bij de uiterste zijden van het histogram. Na een 98% *winsorize* van de variabele “ROA” lijkt die normaler verdeeld. Wanneer een histogram geplot wordt van de variabele “Tobin’s q”, laat deze een sterke rechtsscheefheid zien. Het *winsorizen* van deze variabele zal het probleem van scheefheid niet verhelpen, omdat de uitschieters zich vooral aan de rechterzijde van de datapunten bevinden. De Tobin’s q is echter wel al een ratio en daarom is het nemen van een logaritme van deze variabele niet een logische stap. De sterke rechtsscheefheid van de variabele kan echter wel worden verklaard. Tijdens de coronapandemie hebben enkele gezondheidsbedrijven hiervan geprofiteerd. Dit is terug te zien in de Tobin’s q. De marktwaarde van deze bedrijven is sterk toegenomen, mede dankzij het ontwikkelen van een vaccin. Om een juiste regressie te kunnen uitvoeren, zal daarom toch de natuurlijk logaritme worden genomen van de Tobin’s q in dit onderzoek.

De onafhankelijke variabele in dit onderzoek is de variabele “R&D intensiteit”. Wanneer een histogram wordt geplot van deze variabele blijkt die een rechtsscheefheid te volgen en zou van de

variabele “R&D-intensiteit” de natuurlijk logaritme moeten worden genomen. Bijna alle waarden van de variabele zijn echter kleiner dan nul. Wanneer een natuurlijk logaritme van deze variabele wordt genomen, zal dat resulteren in negatieve waarden en een verkeerd verband aantonen. Wanneer deze variabele echter *gewinsorized* wordt, zal dit een vertekend beeld geven. Bij weinig bedrijven is er sprake van een hoge R&D intensiteit. Omwille van de waarde voor dit onderzoek, worden deze data niet uitgesloten.

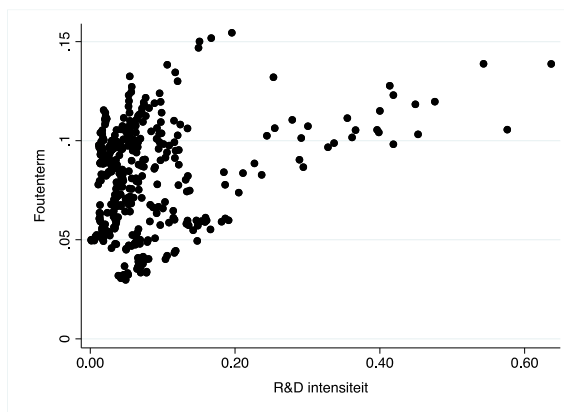
## 4.2.2 Heteroscedasticiteit

Ten tweede moet de variantie van de foutenterm homoscedastisch zijn om een juiste conclusie te kunnen trekken aan de hand van de resultaten. Door de foutenterm te plotten kan er gekeken worden of er sprake is van homogeniteit in de foutterm. Wanneer een spreidingsdiagram niet direct een duidelijk beeld geeft, kan een Breusch-Pagan test voor heteroscedasticiteit worden uitgevoerd. Deze test de nulhypothese of de variantie van de foutterm homogeen is. Er is sprake van heterogeniteit indien bij een significantieniveau van 0,05 de nulhypothese wordt verworpen.

*H1: De bedrijfsomvang heeft een significant negatief effect op de R&D intensiteit.*

$$\mathbf{R\&D\ intensiteit} = \beta_0 + \beta_1 \mathbf{Bedrijfsomvang} + \mu \quad (4)$$

De afhankelijke variabele in de eerste hypothese is de “R&D intensiteit” en de onafhankelijke variabele is de “bedrijfsomvang” (4). Het spreidingsdiagram van de foutenterm van de regressie lijkt homoscedasticiteit aan te tonen, echter is er sprake van een opwaartse trend. Daarnaast geeft de Breusch-Pagan test een p-waarde van 0,000 en daarmee wordt de nulhypothese verworpen bij een significantieniveau van 0,05 en is er sprake van heteroscedasticiteit.

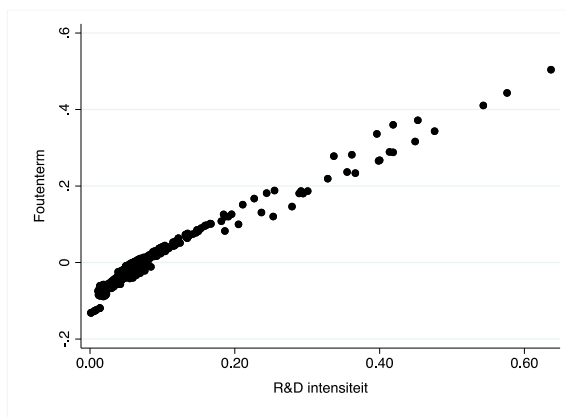


*Figuur 2 Het spreidingsdiagram van de foutenterm van de regressie (afhankelijk: R&D intensiteit; onafhankelijk: de bedrijfsomvang)*

H2: De brutowinstmarge heeft een significant positief effect op de R&D intensiteit van een bedrijf.

$$R\&D\ intensiteit = \beta_0 + \beta_1 w\_Brutowinstmarge + \mu \quad (5)$$

De afhankelijke variabele in de tweede hypothese is de “R&D intensiteit” en de onafhankelijke variabele is de “90% winsorized brutowinstmarge” (5). Het spreidingsdiagram van de foutenterm van de regressie lijkt heteroscedasticiteit aan te tonen, omdat hij een positieve lineaire relatie aan lijkt te nemen. Wanneer je een Breusch-Pagan test uitvoert, geeft de test een p-waarde van 0,000 en daarmee is er een significante aanleiding gevonden om de nulhypothese te verwerpen bij een significantieniveau van 0,05.



Figuur 3 Het spreidingsdiagram van de foutenterm van de regressie (afhankelijk: R&D intensiteit; onafhankelijk: de w\_brutowinstmarge)

H3: De R&D intensiteit van een bedrijf heeft een significant positief effect op de ROA en Tobin's q van een bedrijf.

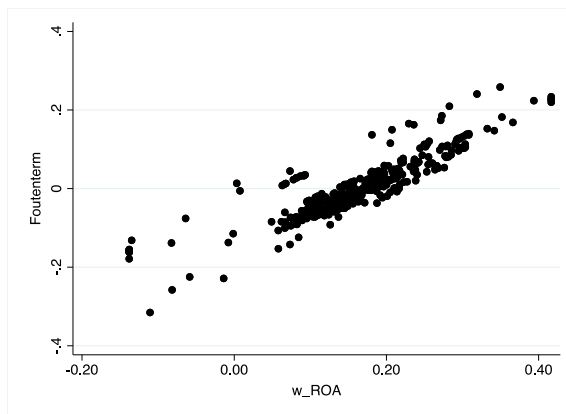
$$w\_ROA = \beta_0 + \beta_1 R\&D\ intensiteit + \beta_2 Bedrijfsomvang + \beta_3 w\_Brutowinstmarge + \mu \quad (6)$$

$$\ln\ Tobin's\ q = \beta_0 + \beta_1 R\&D\ intensiteit + \beta_2 Bedrijfsomvang + \beta_3 w\_Brutowinstmarge + \mu \quad (7)$$

De afhankelijke variabele in de eerste regressie is de “98% winsorized ROA” (6) en in de tweede regressie de natuurlijk logaritme van de “Tobin's q” (7) en de onafhankelijke variabele in beide

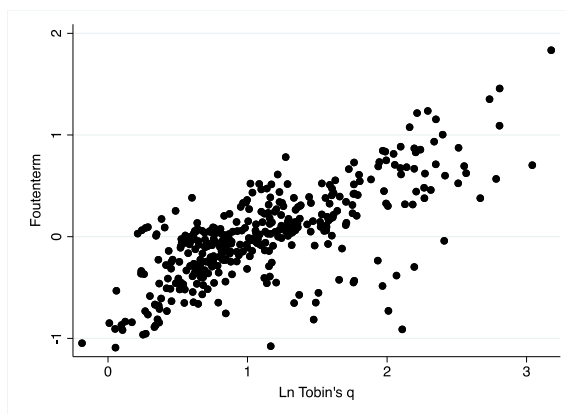
regressies is de “R&D intensiteit”. Daarnaast zijn de controlevariabelen de “bedrijfsomvang” en de “90% winsorized brutowinstmarge”. Het spreidingsdiagram voor de regressie met de variabele “98% winsorized ROA” als afhankelijke variabele (6) lijkt een lineair positieve vorm aan te nemen. Wanneer je een Breusch-Pagan test uitvoert, geeft de test een p-waarde van 0,000. Er is significante aanleiding gevonden om de nulhypothese te verwerpen.

Het spreidingsdiagram voor de regressie met de natuurlijk logaritme van de variabele “Tobin’s



q” als afhankelijke variabele lijkt een lineair positieve relatie aan te nemen. Wanneer je een Breusch-Pagan test uitvoert, geeft de test een p-waarde van 0,000. Daarmee is er een significante aanleiding gevonden om de nulhypothese te verwerpen voor beide testen.

*Figuur 4 Het spreidingsdiagram van de foutenterm van de regressie (afhankelijk: w\_ROA; onafhankelijk: R&D intensiteit; controlevariabelen: de brutowinstmarge en bedrijfsomvang)*



*Figuur 5 Het spreidingsdiagram van de foutenterm van de regressie (afhankelijk: Ln Tobin's q; onafhankelijk: R&D intensiteit; controlevariabelen: de brutowinstmarge en bedrijfsomvang)*

Doordat er hoogstwaarschijnlijk sprake is van heteroscedasticiteit in de variantie van de foutenterm van alle regressies zal in dit onderzoek bij alle hypothesen gebruik worden van *robust standard error* met behulp van Stata. Deze standaardfouten zijn robuuster en meer bestendig tegen het

probleem van heteroscedasticiteit en geven hierdoor een accuratere schatting van de werkelijke standaardfout van een coëfficiënt in een regressie. De dataset in dit onderzoek is een paneldataset. Zoals hier gevisualiseerd wordt, is er een positieve trend te vinden in de foutentermen van de regressies (4), (5), (6) en (7). Dit kan verklaard worden door het feit dat deze paneldataset ook een tijdreeks bevat per variabele. Deze positieve trend is te verklaren doordat de winstgevendheid en financiële prestaties vaak afhangt van het jaar ervoor. Een paneldata regressie door middel van een FE- of RE model is dan ook bestemd tegen heteroscedasticiteit. De eerste aanname die wij getest hebben is hierbij geschonden en geeft meer aanleiding om een FE- of RE model te gebruiken. Nu zullen wij testen of er sprake is van autocorrelatie.

### 4.2.3 Autocorrelatie

Ten derde is de aanname voor homoscedasticiteit van de foutenterm uitgebreid met de aanname dat er geen sprake is van autocorrelatie in de foutentermen. Dit zal onderzocht worden door middel van een Breusch-Godfrey test. Deze test de nulhypothese dat er geen sprake is van autocorrelatie in de foutenterm tot een bepaald aantal jaren terug. Als de nulhypothese wordt verworpen bij een significantieniveau van 0,05 is er sprake van autocorrelatie in de foutenterm. Per regressie worden de foutentermen berekend. Vervolgens wordt een regressie uitgevoerd met de foutenterm als afhankelijke variabelen en als onafhankelijke variabelen vertragingen van die foutenterm. Met vertraging wordt het effect van die variabele op het aantal periodes terug bedoeld, dus een vertraging van 4 van de foutenterm is het effect van de foutenterm 4 jaar geleden. Foutentermen zijn niet te voorspellen, omdat ze buiten de voorspellende waarde van het model liggen. Wanneer er sprake is van autocorrelatie dan zegt de foutenterm van een aantal periodes terug wel iets over een bepaalde foutenterm die minder ver in het verleden ligt. Bij de Breusch-Godfrey test wordt de nulhypothese getest dat in de regressie met vertraagde waardes van de foutenterm de coëfficiënten gelijk zijn aan nul (zie Appendix B). Wanneer de coëfficiënten gelijk zijn aan nul is er geen voorspellende informatie te vinden in het verleden. Wanneer die niet gelijk is aan nul, dan is er significant bewijs gevonden om te veronderstellen dat het verleden wel iets zegt over de toekomst. Uit de Breusch-Godfrey test blijkt voor de eerste hypothese p-waarde van 0,000. Bij de test voor de tweede hypothese blijkt een p-waarde van 0,000. Bij de test voor de derde hypothese met als afhankelijke variabele de “w\_ROA” een p-waarde van 0,006. Bij de test voor de derde hypothese met als afhankelijke variabele de “Tobin’s q” een p-waarde van 0,005. Op een significantieniveau van 0,05 worden bijna alle hypothesen verworpen en dit betekent dat er genoeg significant bewijs is om aan te nemen dat er waarschijnlijk sprake is van autocorrelatie van de foutenterm. De laatste regressie wordt echter pas bij een significantieniveau van 0.1 verworpen. Er is genoeg waarschijnlijkheid dat ook bij de laatste hypothese sprake is van autocorrelatie. Daarom wordt ook deze aanname voor geen autocorrelatie verworpen.

### 4.3 FE- of RE model

De aannames voor homoscedasticiteit van de foutenterm en geen autocorrelatie van de foutenterm zijn beide verworpen en daarom zal er geen gebruik worden gemaakt van PooledOLS, maar wel van een FE model of een RE model. Per hypothese zal nu de Hausman-test worden uitgevoerd en gekeken of een FE- of RE model toegepast moet worden. Hierbij wordt de nulhypothese getest dat de covariantie tussen de endogene variabele en de tijd invariante constante  $a_i$  gelijk is aan nul.

*H1: De bedrijfsomvang heeft een significant negatief effect op de R&D intensiteit.*

$$R\&D\ intensiteit = \beta_0 + \beta_1\ Bedrijfsomvang + \mu + \varepsilon_t \quad (8)$$

In regressie (8) (zie tabel 2) is een RE model te zien met als constante  $\mu$  de foutenterm van de variantie “tussen” de bedrijven en het tijds invariantie gedeelte van de foutenterm. De  $\varepsilon_t$  laat hier de foutenterm zien “binnen” de bedrijven. De afhankelijke variabele in de regressie is de “R&D intensiteit” en de onafhankelijke variabele is de “bedrijfsomvang”. Uit de Hausman-test voor endogeniteit blijkt een p-waarde van 0,297 (zie tabel 3). Er is geen significant bewijs om de nulhypothese te verwerpen bij een significantieniveau van 0,05 en daarom zal de voorkeur uitgaan naar een RE model bij elke hypothese.

*Tabel 2: Regressie resultaat voor een RE (1) en FE (2) model van hypothese 1*

	RE model	FE model
	R&D intensiteit	R&D intensiteit
<i>Bedrijfsomvang</i>	-,026*** (,003)	-,027*** (,003)
<i>Constante</i>	,675*** (,076)	,702*** (,079)
<i>Observaties</i>	380	380
<i>R kwadraat algemeen</i>	,078	,077
<i>R kwadraat tussenin</i>	,072	,072
<i>R kwadraat binnenin</i>	,153	,153

*Robuuste standaardfouten tussen de haakjes*

\*\*\*  $p < ,01$ ; \*\*  $p < ,05$ ; \*  $p < ,1$

Tabel 3: Regressie resultaat voor de Hausman-test van hypothese 1 met  $H_0$ : Verschil tussen de coëfficiënten is niet systematisch

	Coëf.
Chi-kwadraat test waarde	1,089
P-waarde	,297

H2: De brutowinstmarge heeft een significant positief effect op de R&D intensiteit van een bedrijf.

$$R\&D \text{ intensiteit} = \beta_0 + \beta_1 w\_Brutowinstmarge + \mu + \varepsilon_t \quad (9)$$

In regressie (9) (zie tabel 4) is een RE model te zien met als constante  $\mu$  de foutenterm van de variantie “tussen” de bedrijven en het tijds invariantie gedeelte van de foutenterm. De  $\varepsilon_t$  laat hier de foutenterm zien “binnen” de bedrijven. De afhankelijke variabele in de regressie is de “R&D intensiteit” en de onafhankelijke variabele is de “90% winsorized brutowinstmarge”. Uit de Hausman-test voor endogeniteit blijkt een p-waarde van 0,998 (zie tabel 5). Er is geen significant bewijs om de nulhypothese te verwerpen bij een significantieniveau van 0,05 en daarom zal de voorkeur uitgaan naar een RE model bij elke hypothese.

Tabel 4: Regressie resultaat voor een RE (1) en FE (2) model van hypothese 2

	RE model	FE model
	R&D intensiteit	R&D intensiteit
<i>w_Brutowinstmarge</i>	-,088*** (,017)	-,088*** (,017)
<i>Constante</i>	,139*** (,018)	,139*** (,011)
<i>Observaties</i>	380	380
<i>R kwadraat algemeen</i>	,036	,036
<i>R kwadraat tussenin</i>	,032	,032
<i>R kwadraat binnenin</i>	,036	,036

Robuuste standaardfouten tussen de haakjes

\*\*\*  $p < ,01$ ; \*\*  $p < ,05$ ; \*  $p < ,1$

Tabel 5: Regressie resultaat voor de Hausman-test van hypothese 2

	Coëf.
<i>Chi-kwadraat test waarde</i>	,000
<i>P-waarde</i>	,998

*H3: De R&D intensiteit van een bedrijf heeft een significant positief effect op de ROA en Tobin's q van een bedrijf.*

$$w\_ROA = \beta_0 + \beta_1 R\&D\ intensiteit + \beta_2 Bedrijfsomvang + \beta_3 w\_Brutowinstmarge + \beta_4 R\&D\ intensiteit * Bedrijfsomvang + R\&D\ intensiteit * Brutowinstmarge + \mu + \varepsilon_t \quad (10)$$

$$\ln\ Tobin's\ q = \beta_0 + \beta_1 R\&D\ intensiteit + \beta_2 Bedrijfsomvang + \beta_3 w\_Brutowinstmarge + \beta_4 R\&D\ intensiteit * Bedrijfsomvang + \beta_4 w\_Brutowinstmarge * Bedrijfsomvang + \mu + \varepsilon_t \quad (11)$$

In regressie (10) en (11) zijn RE modellen te zien met als constante  $\mu$  de foutenterm van de variantie “tussen” de bedrijven en het tijds invariantie gedeelte van de foutenterm. De  $\varepsilon_t$  laat hier de foutenterm zien “binnen” de bedrijven. De afhankelijke variabele in regressie (10) is de “98% winsorized ROA” en in de tweede regressie de natuurlijk logaritme van de “Tobin's q” (11). De onafhankelijke variabele in beide regressies is de “R&D intensiteit”. De controlevariabelen zijn de “bedrijfsomvang” en de “90% winsorized brutowinstmarge”. Daarnaast zijn er nog twee interactie-effecten toegevoegd tussen de R&D intensiteit en de bedrijfsomvang en ook tussen de R&D intensiteit met de brutowinstmarge. Deze zijn toegevoegd, omdat in hypothese 1 en 2 significante relaties zijn gevonden en daarom is er aanleiding gevonden om deze als interactie-effect mee te nemen in de regressies (10) en (11) van hypothese 3. Beide effecten zijn negatief, dus voor beide effecten wordt een negatief interactie-effect verwacht. Allereerst is een regressie uitgevoerd zonder interactie-effect en vervolgens is eerst alleen het interactie-effect tussen de R&D intensiteit en de bedrijfsomvang uitgevoerd. Daarna is alleen het interactie-effect tussen de R&D intensiteit en de brutowinstmarge toegevoegd. Als laatste zijn beide interactie-effecten toegevoegd om te kijken of er een verschil is tussen de significantie van de coëfficiënten en de R kwadraten. De  $R^2$  laat zien hoeveel van de variantie van de afhankelijke variabele in dit model wordt verklaard door de onafhankelijke variabele.

Uit regressie (10) (zie tabel 8) blijkt dat de significantie toeneemt wanneer het interactie-effect tussen de variabele R&D intensiteit en de bedrijfsomvang wordt toegevoegd. Er wordt dan wel een significante relatie gevonden tussen de bedrijfsomvang en de variabele “98% winsorized ROA”. Daarnaast is het interactie-effect ook significant. Daarnaast neemt de R kwadraat toe van 18,90% naar 28,80% bij een RE model. Echter, het toevoegen van het interactie-effect tussen de R&D intensiteit en de brutowinstmarge resulteert niet in het significant worden van de variabele bedrijfsomvang, maar geeft wel een significant interactie-effect. Wanneer beide interactie-effecten worden toegevoegd, dan



wordt de variabele bedrijfsomvang significant en het interactie-effect tussen de R&D intensiteit en de bedrijfsomvang ook. Het interactie-effect tussen de R&D intensiteit en de brutowinstmarge wordt dit echter niet. Daarnaast is er geen significant verschil tussen de regressies met beide interactie-effecten toegevoegd en die met alleen het interactie-effect tussen de R&D intensiteit en de bedrijfsomvang. Daarom zal regressie (10) worden uitgevoerd met daaraan alleen het interactie-effect tussen de R&D intensiteit en de bedrijfsomvang toegevoegd. Uit de Hausman-test voor endogeniteit blijkt een p-waarde van 0,034 (zie tabel 6). Er is significant bewijs gevonden om de nulhypothese te verwerpen bij een significantieniveau van 0,05 en daarom zal de voorkeur uitgaan naar een FE model bij regressie (10).

*Tabel 6: Regressie resultaat voor de Hausman-test van hypothese 3 met w\_ROA als afhankelijke variabele en met de R&D intensiteit en de bedrijfsomvang als interactie-effect*

	Coëf.
<i>Chi-kwadraat test waarde</i>	10,444
<i>P-waarde</i>	,034

Uit regressie (11) (zie tabel 9) blijkt dat de significantie niet toeneemt wanneer beide interactie-effecten aan de regressie worden toegevoegd. De variabele “90% *winsorized* brutowinstmarge” is alleen bij de eerste RE model regressie significant. Er is echter wel een significant verband waarneembaar tussen het interactie-effect tussen de R&D intensiteit en de bedrijfsomvang en de natuurlijk logaritme van de “Tobin’s q” bij de RE modellen. Daarnaast is te zien dat geen enkele variabele significant is wanneer een FE model wordt uitgevoerd wanneer beide interactie-effecten worden toegevoegd aan de regressie. Het R kwadraat tussen de regressie zonder interactie-effect en met beide interactie-effecten neemt niet significant toe. Daarom is geen aanleiding gevonden om een interactie-effect toe te voegen en zal de regressie worden uitgevoerd met alleen de R&D intensiteit als onafhankelijke variabele en de “90% *winsorized* brutowinstmarge” en bedrijfsomvang als controlevariabelen. De coëfficiënt van het interactie-effect tussen de variabele R&D intensiteit en bedrijfsomvang neemt echter een positieve waarde aan. Dit gaat niet tegen de verwachtingen in, maar werd er wel een negatieve coëfficiënt verwacht voor het interactie-effect. Echter, werd er ook verwacht dat de R&D intensiteit een positief effect zou hebben op de Tobin’s q, maar dat effect bleek negatief te zijn wanneer het interactie-effect wordt toegevoegd. Het negatieve effect van R&D op de op de Tobin’s q neemt af in de toename van de

bedrijfsomvang. Dit is te verklaren doordat grote bedrijven minder investeren in R&D en juist meer investeren in overnames. Het interactie-effect tussen de R&D intensiteit en de brutowinstmarge blijkt negatief niet significant te zijn bij het RE model met het interactie-effect tussen de R&D intensiteit en de brutowinstmarge aan het model toegevoegd. Het positieve effect van de R&D op de Tobin's  $q$  is kleiner voor bedrijven met een hoger brutowinstmarge. Dit is te verklaren doordat bedrijven met een hoger brutowinstmarge vaak minder R&D kosten maken en een lagere R&D intensiteit hebben, omdat ze die kosten besparen en daarmee hun marge vergroten. De R&D intensiteit heeft daarom ook minder effect op hun Tobin's  $q$ , omdat hun intensiteit lager is. Daarnaast blijkt uit de Hausman-test voor endogeniteit een  $p$ -waarde van 0,000. Er is significant bewijs gevonden om de nulhypothese te verwerpen bij een significantieniveau van 0,05 en normaal gesproken zou de voorkeur uitgaan naar een FE model. Bell et al. (2019) benoemen echter dat het niet altijd beter is om een FE model te gebruiken als een Hausman-test suggereert. Een FE model is niet goed in staat om voor *biasesness* robuust te zijn wanneer de normaliteit aanname wordt geschonden. De R&D intensiteit aanname is rechtsscheef en kan daarom een verkeerde uitkomst geven. Het probleem bij een FE model is dat er wordt aangenomen dat de variantie in de afhankelijke variabele door een gemiddelde waarde van de bedrijven verklaard kan worden. Dit impliceert dat de schatters bij een FE model niet *biased* zijn. Als de schattingen van de parameters van een RE model significant verschillen van die van een FE model zal de Hausman-test een FE model aanraden. Ook al heeft die lagere verklarende waarde. Bij regressie (11) heeft een RE model (algemene R kwadraat van 45,90%) significant meer verklarende waarde dan een FE model (algemene R kwadraat van 0,10%). Daarnaast zijn alle variabelen significant bij een RE model en is de variabele "90% *winsorized* brutowinstmarge" niet significant bij een FE model. Daarom zal een RE model gebruikt worden bij regressie (11).

Tabel 7: Regressie resultaat voor de Hausman-test van hypothese 3 met de natuurlijk logaritme van de Tobin's  $q$  als afhankelijke variabele

	Coëf.
<i>Chi-kwadraat test waarde</i>	49,979
<i>P-waarde</i>	,000

Tabel 8: Regressie resultaat voor een RE (1) en FE (2) model van hypothese 3 met de 98% winsorized ROA als afhankelijke variabele.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	RE model	FE model	RE model	FE model	RE model	FE model	RE model	FE model
	w_ROA	w_ROA	w_ROA	w_ROA	w_ROA	w_ROA	w_ROA	w_ROA
<i>R&amp;D intensiteit</i>	-,385***	-,543***	-4,764***	-4,398***	-,465***	-,595***	-4,155***	-3,903***
	(,076)	(,097)	(,702)	(,737)	(,077)	(,096)	(,779)	(,836)
<i>w_Brutowinstmarge</i>	,235***	,244***	,177***	,198***	,1**	,101**	,126***	,148***
	(,027)	(,031)	(,027)	(,031)	(,041)	(,05)	(,04)	(,05)
<i>Bedrijfsomvang</i>	-,005	-,002	-,024***	-,021***	-,004	-,001	-,021***	-,018**
	(,005)	(,007)	(,006)	(,007)	(,005)	(,007)	(,006)	(,008)
<i>R&amp;D intensiteit * Bedrijfsomvang</i>			,206***	,182***	,38***		,175***	,158***
			(,033)	(,035)			(0,037)	(,04)
<i>R&amp;D intensiteit * Brutowinstmarge</i>					,38***	,349***	,17*	,136
					(,088)	(,097)	(,096)	(,109)
<i>Constante</i>	,158	,093	,603***	,548***	,214*	,143	,564***	,506***
	(,116)	(,155)	(,13)	(,172)	(,115)	(,153)	(,132)	(,175)
<i>Observaties</i>	380	380	380	380	380	380	380	380
<i>R kwadraat algemeen</i>	0,189	0,166	0,288	0,257	0,205	0,167	0,289	0,253

Robuuste standaardfouten tussen de haakjes

\*\*\*  $p < ,01$ ; \*\*  $p < ,05$ ; \*  $p < ,1$

Tabel 9: Regressie resultaat voor een RE (1) en FE (2) model van hypothese 3 met de natuurlijk logaritme van de Tobin's q als afhankelijke variabele,

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	RE model	FE model	RE model	FE model	RE model	FE model	RE model	FE model
	Ln Tobin's q	Ln Tobin's q	Ln Tobin's q	Ln Tobin's q	Ln Tobin's q	Ln Tobin's q	Ln Tobin's q	Ln Tobin's q
<i>R&amp;D intensiteit</i>	2,695*** (,448)	2,278*** (,695)	-9,979* (5,23)	-5,266 (5,573)	2,773*** (,507)	2,251*** (,721)	-13,439** (5,679)	-7,448 (6,436)
<i>w_Brutowinstmarge</i>	,327** (,137)	,012 (,155)	,181 (,148)	-,062 (,164)	,401 (,259)	-,035 (,366)	,514** (,261)	,179 (,391)
<i>Bedrijfsomvang</i>	-,091*** (,028)	,153*** (,049)	-,137*** (,035)	,116** (,056)	-,091*** (,029)	,154*** (,05)	-,155*** (,037)	,098 (,062)
<i>R&amp;D intensiteit * Bedrijfsomvang</i>			,59** (,243)	,357 (,261)			,769*** (,269)	,467 (,308)
<i>R&amp;D intensiteit * Brutowinstmarge</i>					-,274 (,804)	,142 (1,007)	-1,37 (,886)	-,804 (1,183)
<i>Constante</i>	2,769*** (,654)	-2,587** (1,119)	3,866*** (,809)	-1,736 (1,28)	2,735*** (,66)	-2,589** (1,121)	4,088*** (,821)	-1,459 (1,344)
<i>Observaties</i>	380	380	380	380	380	380	380	380
<i>R kwadraat algemeen</i>	0,459	0,001	0,440	0,004	0,465	0,000	0,460	0,001

Robuuste standaardfouten tussen de haakjes

\*\*\*  $p < ,01$ ; \*\*  $p < ,05$ ; \*  $p < ,1$

## 4.4 Multicollineariteit

Ten derde mag er geen sprake zijn van multicollineariteit. Er is dan sprake van hoge correlatie tussen twee variabelen waardoor ze bijna geen unieke of onafhankelijke informatie bevatten. Als er sprake is van multicollineariteit kan het teken van de coëfficiënt verkeerd zijn, maar bovenal kan er sprake zijn van grote standaardfouten en grote variabiliteit van de steekproef. Hierdoor zijn de coëfficiënten niet meer betrouwbaar en neemt de precisie van de voorspellingen sterk af (Alin, 2010). In dit scriptieonderzoek zal de multicollineariteit getest worden met behulp van de *variance inflation factor* (VIF)<sup>5</sup>. De VIF meet de correlatie en de sterkte van de correlatie tussen twee verklarende variabelen in een regressie. Vanaf een VIF van vijf is er sprake van zorgelijke correlatie tussen de twee variabelen. Wanneer er sprake is multicollineariteit zullen de variabelen worden samengevoegd of zal één gelijkwaardige variabele verwijderd worden. Appendix C laat de VIF van drie soorten regressies met alle variabelen erin, maar met verschillende afhankelijke variabelen. De drie regressies bevatten in deze volgorde de 98% *winsorized* ROA, de R&D intensiteit en de natuurlijke logaritme van de Tobin's q als afhankelijke variabele. Als onafhankelijke variabelen zijn dat, naast de afhankelijke variabelen die op dat moment als onafhankelijke variabele functioneren, de bedrijfsomvang en de 90% *winsorized* brutowinstmarge. Alle variabelen hebben een VIF die niet groter is dan twee en daarmee is er geen sprake van zorgelijke correlatie tussen de variabelen.

---

<sup>5</sup>  $VIF = \frac{1}{1-R^2}$

# Hoofdstuk 5 – Resultaten

H1: De bedrijfsomvang heeft een significant positief effect op de R&D intensiteit.

$$R\&D\text{ intensiteit} = \beta_0 + \beta_1 \text{Bedrijfsomvang} + \mu + \varepsilon_t \quad (8)$$

Uit het RE model blijkt dat de bedrijfsomvang een negatieve significante relatie heeft met de R&D intensiteit van een bedrijf (zie tabel 10). De regressie laat een effect van -0,026 zien met een p-waarde van 0,000. Deze verandering is het gemiddelde van de effecten tussen de bedrijven en de effecten door de jaren heen van de bedrijven zelf. De  $\rho$  laat zien welk deel van de variantie van de afhankelijke variabele door de tijd invariante constante wordt verklaard. In dit model wordt ongeveer 91% van de variantie in het model verklaard door tijd invariante variantie en variantie tussen de bedrijven zelf. Dit laat zien dat veel van de veranderingen te verklaren zijn door cross-sectionele effecten. De  $R^2$  laat zien hoeveel van de variantie van de variabele “R&D intensiteit” in dit model wordt verklaard door de variabele “bedrijfsomvang”. De binnenin  $R^2$  laat zien hoeveel van de variantie in het model binnen de individuele bedrijven wordt verklaard en de tussenin  $R^2$  hoeveel van de variantie van het model tussen de bedrijven wordt verklaard. De binnenin  $R^2$  is twee keer zo hoog als de tussenin  $R^2$ , echter zoals de  $\rho$  al liet zien weegt de variantie tussen de bedrijven zwaarder mee en is de gemiddelde verklarende waarde van het model 7.78%. Dit is geen grote verklarende waarde, echter is er slechts een onafhankelijke variabele en verklaard die genoeg om iets te zeggen over de afhankelijke variabele.

Tabel 10: Regressie resultaat voor een RE model van hypothese 1

	RE model
	R&D intensiteit
<i>Bedrijfsomvang</i>	-,026*** (,003)
<i>Constante</i>	,675*** (,076)
<i>Observaties</i>	380
<i>R kwadraat algemeen</i>	,078
<i>R kwadraat tussenin</i>	,072
<i>R kwadraat binnenin</i>	,153
$\sigma_u$	,088
$\sigma_e$	,028
$\rho$	,908

Robuuste standaardfouten tussen de haakjes

\*\*\*  $p < ,01$ ; \*\*  $p < ,05$ ; \*  $p < ,1$

H2: De brutowinstmarge heeft een significant positief effect op de R&D intensiteit van een bedrijf.

$$\mathbf{R\&D\ intensiteit} = \beta_0 + \beta_1 w\_Brutowinstmarge + \mu + \varepsilon_t \quad (9)$$

Uit het RE model blijkt dat de 90% *winsorized* brutowinstmarge een negatieve significante relatie heeft met de R&D intensiteit van een bedrijf (zie tabel 11). De regressie laat een effect zien van -0,088 met een p-waarde van 0,000. Daarnaast wordt ongeveer 90% van de variantie in het model verklaard door tijd invariante variantie en de variantie tussen de bedrijven zelf. De gemiddelde verklarende waarde van het model is 7.76%.

Tabel 11: Regressie resultaat voor een RE model hypothese 2

	RE model
	R&D intensiteit
<i>w_Brutowinstmarge</i>	-.088*** (.017)
<i>Constante</i>	.139*** (.018)
<i>Observaties</i>	380
<i>R kwadraat algemeen</i>	,036
<i>R kwadraat tussenin</i>	,032
<i>R kwadraat binnenin</i>	,073
$\sigma_u$	,089
$\sigma_e$	,029
$\rho$	,904

Robuuste standaardfouten tussen de haakjes

\*\*\*  $p < ,01$ ; \*\*  $p < ,05$ ; \*  $p < ,1$

H3: De R&D intensiteit van een bedrijf heeft een significant positief effect op de ROA van een bedrijf.

$$w\_ROA = \beta_0 + \beta_1 R\&D\ intensiteit + \beta_2 Bedrijfsomvang + \beta_3 w\_Brutowinstmarge + \beta_4 R\&D\ intensiteit * Bedrijfsomvang + \mu + \varepsilon_t \quad (12)$$

$$Ln\ Tobin's\ q = \beta_0 + \beta_1 R\&D\ intensiteit + \beta_2 Bedrijfsomvang + \beta_3 w\_Brutowinstmarge + \mu + \varepsilon_t \quad (13)$$

Uit het FE model voor regressie (12) blijkt dat de R&D intensiteit een negatieve significante relatie heeft met de 98% *winsorized* ROA (zie tabel 12). De regressie laat een effect zien van -4,398 met een p-waarde van 0,000. Daarnaast blijkt de bedrijfsomvang een kleine negatieve significante relatie te hebben met de 98% *winsorized* ROA. De regressie laat een effect zien van -0,021 met een p-waarde van 0,005. De 90% *winsorized* brutowinstmarge een positief significant effect zien van 0,198 en een p-waarde van 0,000. Het interactie-effect tussen de R&D intensiteit en de bedrijfsomvang laat een positief significant effect zien van 0,182 met een p-waarde van 0,000. Dit gaat niet tegen de verwachtingen in, echter werd er wel een negatieve coëfficiënt verwacht voor het interactie-effect. Echter, werd er ook verwacht dat de R&D intensiteit een positief effect zou hebben op de ROA, maar dat effect bleek negatief te zijn. Het negatieve effect van R&D op de op de 98% *winsorized* ROA neemt af in de toename van de bedrijfsomvang. Dit is te verklaren doordat grote bedrijven minder investeren in R&D en juist meer investeren in overnames. R&D uitgaven zijn kosten voor een bedrijf die een negatieve invloed hebben op de ROA en dit wordt versterkt voor grotere bedrijven. Dit zal door hun omvang namelijk nog minder opleveren en zullen hierdoor sneller overnames willen doen. Het negatieve effect van de R&D intensiteit neemt daarom af voor grotere bedrijven, omdat die bedrijven minder in R&D investeren en er daarom ook minder last van hebben.

Uit het RE model voor regressie (13) blijkt dat de R&D intensiteit een significante positieve relatie heeft met de natuurlijk logaritme van de variabele Tobin's q. De regressie laat een effect zien van 2,695 met een p-waarde van 0,000. De 90% *winsorized* brutowinstmarge heeft een positief significant effect zien van 0,327 en een p-waarde van 0,017. Daarnaast blijkt de bedrijfsomvang een negatieve significante relatie te hebben met de natuurlijk logaritme van de Tobin's q. De regressie laat een effect zien van -0,910 met een p-waarde van 0,001. Daarnaast wordt ongeveer 75% van de variantie in het model verklaard door tijd invariante variantie.



Tabel 12: Regressie resultaat voor een FE model van hypothese 2 met w\_ROA als afhankelijke variabele

	FE model
	w_ROA
<i>R&amp;D intensiteit</i>	-4.398*** (.737)
<i>w_Brutowinstmarge</i>	.198*** (.031)
<i>Bedrijfsomvang</i>	-.021*** (.007)
<i>R&amp;D intensiteit *</i>	.182*** (.035)
<i>Constante</i>	.548*** (.172)
<i>Observaties</i>	380
<i>R kwadraat algemeen</i>	,257
<i>R kwadraat tussenin</i>	,244
<i>R kwadraat binnenin</i>	,340
$\sigma_u$	,067
$\sigma_e$	,047
$\rho$	,665

Robuuste standaardfouten tussen de haakjes

\*\*\*  $p < ,01$ ; \*\*  $p < ,05$ ; \*  $p < ,1$

Tabel 13: Regressie resultaat voor een RE model van hypothese 2 met Tobin's q als afhankelijke variabele

	RE model
	Ln Tobin's q
<i>R&amp;D intensiteit</i>	2.695*** (.448)
<i>w_Brutowinstmarge</i>	.327** (.137)
<i>Bedrijfsomvang</i>	-.091*** (.028)
<i>Constante</i>	2.769*** (.654)
<i>Observaties</i>	380
<i>R kwadraat algemeen</i>	,459
<i>R kwadraat tussenin</i>	,703
<i>R kwadraat binnenin</i>	,033
$\sigma_u$	,254
$\sigma_e$	,352
$\rho$	,341

Robuuste standaardfouten tussen de haakjes

\*\*\*  $p < ,01$ ; \*\*  $p < ,05$ ; \*  $p < ,1$

## Hoofdstuk 6 Conclusie

Het doel van dit onderzoek was om antwoord te geven op de hoofdvraag: *Wat is de invloed van R&D op de financiële prestaties van bedrijven in de gezondheidsindustrie in de S&P 500?* Deze hoofdvraag werd beantwoord met twee hypothesen en een derde hypothese die functioneerde als antwoord op de hoofdvraag.

Ten eerste werd de hypothese onderzocht dat de bedrijfsomvang een significant negatief effect heeft op de R&D intensiteit. Uit dit onderzoek bleek een negatieve significante relatie en daarmee wordt de eerste hypothese niet verworpen. De resultaten sluiten hierbij aan bij eerder onderzoek (Doi, 1994; Shefer & Frenkel, 2005). Eerder onderzoek liet echter ook een positieve relatie zien en zoals Shefer en Frenkel (2005) benadrukten, maar het gevonden negatieve verband is vooral te verklaren door de vertraging van het effect van een investering in R&D. Daarnaast is het te verklaren doordat grote bedrijven niet investeren in specifieke R&D, maar deze zullen eerder andere bedrijven overnemen die al beschikken over deze technologie of innovatieve kennis. Een gezondheidsbedrijf moet ongeveer 2 tot 4 miljard jaarlijks investeren in R&D om op een normaal tempo mee te groeien met de markt (Alvaro et al., 2010). Veel grote gezondheidsbedrijven investeren daarom in fusies en overnames en daarom is er een negatieve relatie tussen omvang van een bedrijf en de R&D intensiteit.

Ten tweede werd de hypothese onderzocht dat de brutowinstmarge een significant positief effect heeft op de R&D intensiteit van een bedrijf. Uit dit onderzoek bleek een significant negatieve relatie en daarmee wordt de tweede hypothese verworpen. Uit eerder onderzoek bleek vooral een positieve relatie, maar de resultaten worden onderbouwd door onderzoek van Akben-Selcuk (2016). Een bedrijf maakt namelijk een afweging tussen de verdeling van kosten. Een afname van de brutowinstmarge gaat vaak gepaard met een toename van andere kosten. In dit geval zal dit een toename van R&D-uitgaven betekenen en dit resulteert in een negatieve relatie tussen de brutowinstmarge en de R&D-uitgaven.

Ten derde werd de hypothese onderzocht dat de R&D-intensiteit een significant positief effect heeft op de ROA en Tobin's q van een bedrijf. Daarmee wordt ook antwoord gegeven op de hoofdvraag: *Wat is de invloed van R&D op de financiële prestaties van bedrijven in de gezondheidsindustrie in de S&P 500?* Uit dit onderzoek bleek een negatieve significante relatie tussen de R&D intensiteit en de ROA. Hierbij wordt de hypothese verworpen, maar dit wordt onderbouwd door eerder onderzoek (Vithessonthi & Racela, 2016; Pantagakis et al., 2012; Yeh et al., 2010; Kotabe et al., 2002). Het effect van R&D investering zal op de korte termijn een negatief effect hebben op de ROA, omdat het effect op de lange termijn pas zichtbaar is. Een huidige investering in R&D zal niet gelijk resulteren in positief resultaat en een positieve ROA, maar dit effect zal vertraagd zijn en later pas zichtbaar worden. Daarnaast wordt in eerder onderzoek vaak een R&D-intensiteitsgrens gevonden vanaf wanneer de relatie pas positief wordt. Een combinatie tussen de mate van de R&D-intensiteit en de vertraging van het effect kan een verklaring zijn voor de negatieve relatie tussen de R&D-intensiteit en de ROA in dit

onderzoek. Daarnaast bleek uit de resultaten dat een toename van de bedrijfsomvang een verzwakte werking heeft op het negatieve effect van de R&D-intensiteit. R&D-uitgaven zijn kosten voor een bedrijf die een negatieve invloed hebben op de ROA en dit effect wordt verminderd bij grotere bedrijven. Dit is te verklaren doordat grote bedrijven minder investeren in R&D en juist meer investeren in overnames. Uit het onderzoek bleek echter een positieve relatie tussen de R&D intensiteit en de Tobin's q en daarmee wordt de hypothese niet verworpen. Dit is te verklaren doordat de Tobin's q niet alleen de prestaties in het bedrijf op basis van de boekwaarde weerspiegelt, maar dat deze variabele ook wijst op groeimogelijkheden, omdat die ook huidige marktwaarden meeneemt in zijn berekening. Bardhan et al. (2013) vonden een positieve relatie tussen de R&D uitgaven en de Tobin's q en deze sluit aan bij de bevindingen in dit onderzoek. Zij beargumenteren dat de Tobin's q een toekomstgerichte maatstaf is voor prestaties van een bedrijf. Het is niet een schatting op basis van de boekwaarde, maar ook van de markt en kan daarmee al een deel van het effect van een R&D investering laten zien.

De R&D heeft een negatieve invloed op de financiële prestaties van een bedrijf wanneer op basis van de boekwaarden wordt gemeten, waarbij de grootte van het bedrijf dit negatieve effect verzwakt. De mate van R&D investeringen nemen af in de toename van de omvang van een bedrijf en daarom is het effect op de financiële prestaties van een bedrijf dan minder negatief effect. Echter, de R&D heeft een positieve invloed op de financiële prestaties van een bedrijf wanneer een combinatie van boekwaarden en marktwaarden wordt gebruikt als proxy voor de financiële prestaties.

## Hoofdstuk 7 Beperkingen en aanbevelingen

Dit onderzoek is een paneldataset oorspronkelijk bestaande uit 62 bedrijven. Voor een aantal bedrijven bleek de dataset incorrect of incompleet. 38 bedrijven bleven over na filtering. Dit verkleint echter wel de dataset en daarmee de betrouwbaarheid van het onderzoek. Dit zijn daarnaast bedrijven uit de S&P 500. Wanneer wordt getest of de bedrijfsomvang invloed heeft op de R&D intensiteit van een bedrijf, zal deze test zeer kleine effecten liet zien. Een mogelijkheid zou zijn om gebruik te maken van de S&P 1500 met een grotere variantie in omvang van de bedrijven.

Het onderzoek zou uitgebreid kunnen worden door een langere tijdreeksreeks te nemen van de bedrijven. Een grote beperking van de resultaten en daarmee ook verwerping van de hypotheses, kwam vooral doordat het effect van de R&D nog niet goed te zien was. Bij een lange tijdreeksreeks zal het effect hopelijk meer zichtbaar worden en kan er ook meer worden gezegd over de lange termijneffecten van R&D op de financiële prestaties van een bedrijf. Daarnaast zou het ook een toevoeging zijn voor het onderzoek om andere kosten mee te nemen zodat de kostenverdeling binnen de bedrijven ook meer zichtbaar is. Dan kan er ook een verband onderzocht worden tussen de kostenverdeling binnen een bedrijf en de mate waarin wordt geïnvesteerd in R&D. Om een allocatie te vinden die optimaal zou zijn voor de R&D intensiteit en winstgevendheid van een bedrijf.

Daarnaast kan er ook sprake zijn van endogeniteit. Robert en Whited (2013) zeggen dat afhankelijke variabele bij exogeniteit afhangt van de foutterm en van de onafhankelijke variabele. Daarnaast betekent het ook dat de onafhankelijke variabele niet mag afhangen van de afhankelijke variabele, in andere woorden mag die niet gecorreleerd zijn met de foutterm. Het tegenovergestelde van exogeniteit is endogeniteit. De regressiecoëfficiënt zal verkeerd geschat worden als hier sprake van is. Deze aanname is vooral gebaseerd op correlatie met de foutterm, echter kan men de foutterm niet observeren en daarom is deze aanname onmogelijk om te testen. Er zijn vier welbekkende oorzaken voor het schenden van de exogeniteit aanname: vergeten variabelen, simultaniteit, meetfouten en *selection bias*.

Wanneer een variabele niet wordt opgenomen in de regressie, maar gecorreleerd is met de onafhankelijke variabele en ook een relatie heeft met afhankelijke variabele is er sprake van een *omitted variable bias* (OVB). Dit is een weggelaten variabele die voor een vertekend beeld zorgt. Dit probleem speelt vaak een grote rol in *corporate finance*. De onderwerpen van verschillende studies zijn vaak heterogeen op verschillende gebieden. Bepaalde fricties zoals informatie asymmetrie en management drijfveren zijn lastig om te kwantificeren en te observeren. Daarnaast zijn veel beslissingen door een bedrijf gebaseerd op publieke en private informatie. Specifiek in dit onderzoek zijn er verschillende andere factoren die invloed op de R&D van een bedrijf hebben. In de eerste hypothese zal een variabele gevonden moeten worden die gecorreleerd is met de R&D intensiteit van een bedrijf en met de grootte van het bedrijf. Eén variabele is bijvoorbeeld de leeftijd van het bedrijf. Hoe ouder het bedrijf des te

meer R&D nodig is om te concurreren tegen de markt en om innovatieve technologieën op de markt te brengen. Daarnaast zal de leeftijd van het bedrijf ook invloed hebben op de omvang van het bedrijf, omdat het langer de tijd heeft gehad om te groeien. Het probleem van OVB is lastig te verhelpen en een bepaalde *bias* zal in dit onderzoek aanwezig zijn. Dit kan verholpen worden door een *general-to-specific* aanpak te gebruiken waarbij eerst veel variabelen worden toegevoegd aan een model en daarna worden er steeds meer variabelen verwijderd die te weinig verklarende of significante waarde toevoegen aan het model.

Daarnaast kan er is sprake zijn van simultaniteit wanneer de afhankelijke variabele de onafhankelijke variabele beïnvloedt, maar de onafhankelijke tegelijkertijd ook afhankelijke variabele beïnvloedt. Wanneer dit het geval is, zal het toenemen van de observaties of het toevoegen van een verklarende variabele het probleem niet oplossen. In dit onderzoek kan dit het geval zijn tussen de financiële prestaties van een bedrijf enerzijds en de R&D anderzijds. Beide variabelen kunnen invloed op elkaar hebben, omdat wanneer het goed gaat met het bedrijf vaak meer financiële ruimte is voor onderzoek en daarnaast meer onderzoek kan leiden tot een beter financiële prestaties. Om dit probleem op te lossen zal een *two-stage least squares* (2SLS) toe worden gepast. In de eerste stap zal de regressie de daadwerkelijke exogene variabele (instrumentele variabele) gebruiken om een kunstmatige variabele te creëren. In de tweede stap zal die kunstmatige variabele in plaats van de endogene variabele aan de rechterkant van de regressie gebruikt worden. 2SLS is een consistente schatter van een parameter in een simultane vergelijking (Hendry & Morgan, 1997). Daarnaast kan ook de onafhankelijke variabele vertraagd worden met een aantal periodes om zo het directe effect op elkaar weg te halen. Naast OVB en simultaniteit kan er allerlaatst ook sprake zijn van meetfouten in de onafhankelijke variabele.

Bij *corporate finance* wordt vaak een proxy gebruikt voor niet observeerbare of moeilijke kwantificeerbare variabelen. Wanneer er een groot verschil is tussen de werkelijke variabele en de proxy is er sprake van een meetfout. De resultaten zullen dan bevooroordeeld zijn richting 0. Het daarom belangrijk om juiste data te gebruiken van een betrouwbare bron. In dit onderzoek zijn er twee proxies gebruikt voor het meten van de financiële prestaties van een bedrijf en beide gaven twee verschillende effecten. De verklaring werd gevonden in de aard van de basis van de proxys, omdat die werd berekend op basis van de boekwaarde en de ander een combinatie was van boek- en marktwaarde. De keuze van de proxy kan van invloed zijn op de uitkomst van het onderzoek. Dit kan verholpen worden door veel theoretische onderbouwing te zoeken voor de gekozen variabelen om daarmee verkeerde keuzes te voorkomen.

Tot slot kan er ook sprake van een selectie *bias*, omdat een groot deel van de oorspronkelijk dataset is gefilterd, of omdat er te weinig data of observaties beschikbaar waren. Dit kan ertoe leiden dat er op een verkeerde manier gefilterd is. Dit kan worden voorkomen door van tevoren een juiste onderzoeksopzet te kiezen.

# Bibliografie

- Akben-Selcuk, E. (2016). Factors Affecting Firm Competitiveness: Evidence from an Emerging Market. *International Journal of Financial Studies*, 4(2), 1-10.
- Alin, A. (2010). Multicollinearity. *Wiley Interdisciplinary Reviews Computational Statistics*, 2(3), 370–374.
- Alvaro, D., Challener, C., & Branch, E. (2010, 20 Maart). M&A: Fundamental to Pharma Industry Growth. *Pharma's Almanac*.  
<https://www.pharmasalmanac.com/articles/ma-fundamental-to-pharma-industry-growth>
- Artz, K. W., Norman, P. M., Hatfield, D. E., & Cardinal, L. B. (2010). A longitudinal study of the impact of R&D, patents, and product innovation on firm performance. *Journal of Product Innovation Management*, 27(5), 725-740.
- Ashmarina, S. I., Zotova, A. S., & Smolina, E. S. (2016). Implementation of financial sustainability in organizations through valuation of financial leverage effect in Russian practice of financial management. *International Journal of Environmental and Science Education*, 11(10), 3775-3782.
- Bardhan, I., Krishnan, V., & Lin, S. (2013). Business Value of Information Technology: Testing the Interaction Effect of IT and R&D on Tobin's Q. *Information Systems Research*, 24(4), 1147-1161.
- Beck, T., Demirguc-Kunt, A., Laeven, L., & Levine, R. (2008). Finance, firm size and growth. *Journal of Money, Credit and Banking*, 40(7), 1379-1405.
- Belderbos, R., Carree, M., & Lokshin, B. (2004). Cooperative R&D and firm performance. *Research Policy Journal*, 33(10), 1477–1492.
- Bell, A., Fairbrother, M., & Jones, K. (2019). Fixed and Random effects models: making an informed choice. *Springer*, 53(2), 1051-1074.

- Bolívar, M. P. (2016). *Global Encyclopedia of Public Administration, Public Policy, and Governance*. Springer.
- Chen, M., Cheng, S.J., & Hwang, Y. (2005). An empirical investigation of the relationship between intellectual capital and firms' market value and financial performance. *Journal of Intellectual Capital*, 6(2), 159-175.
- Choi, J., & Lee, J. (2017). Firm Size and Compositions of R&D Expenditures: Evidence from a Panel of R&D Performing Manufacturing Firms. *Industry and Innovation*, 25(5), 459-481.
- Chung, K.H., & Pruitt, S.W. (1994). A simple approximation of Tobin's Q. *Financial Management*, 23(2), 70-74.
- Cohen, M., Richard, C., & David, C. (1987). Firm Size and R&D intensity: A Re-Examination. *Journal of Industrial Economics*, 35(4), 543-565.
- Connolly, R. A., & Hirschey, M. (2005). Firm Size and the Effect of R&D on Tobin's. *R&D Management*, 35(2), 217-223.
- Dave, P., Wadhwa, V., Aggarwal, S., & Seetharman, A. (2013). The Impact of Research and Development on the Financial Sustainability of Information Technology (IT) Companies Listed on the S&P 500 Index. *Journal of Sustainable Development*, 6(11), 122-138.
- Doi, N. (1994). Firm Size and R&D Activity in Japanese Manufacturing Industries. *Science and Public Policy*, 21(3), 295-308.
- Grabowski, H., & Mueller, D. (1978). Industrial Research and Development, Intangible Capital Stocks, and Firm Profit Rates. *The Bell Journal of Economics*, 9(2), 328-343.
- Hendry, D., & Morgan, M. (1997). *The Foundations of Econometric Analysis*. Cambridge University Press, Cambridge.
- Hitt, M. A., Hoskisson, R. E., Ireland, R. D., & Harrison, J. D. (1989). Acquisitive growth strategy and relative R&D intensity: The effects of leverage, diversification and size. *Academy of Management Best Paper Proceedings*, 22-26.
- Hoeklin, M. (2021). €93 Billion Spent By Public Sector On COVID Vaccines and Therapeutics in 11 Months, Research Finds. *Health Policy Watch*. <https://healthpolicy-watch.news/81038-2/>



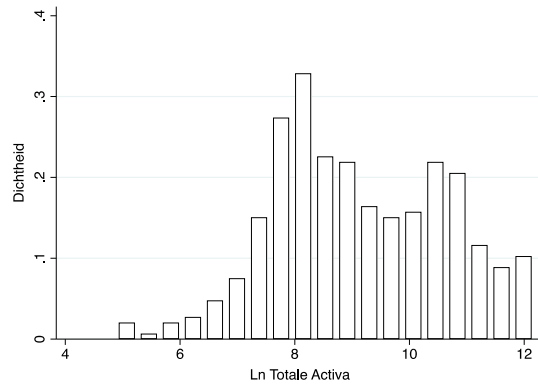
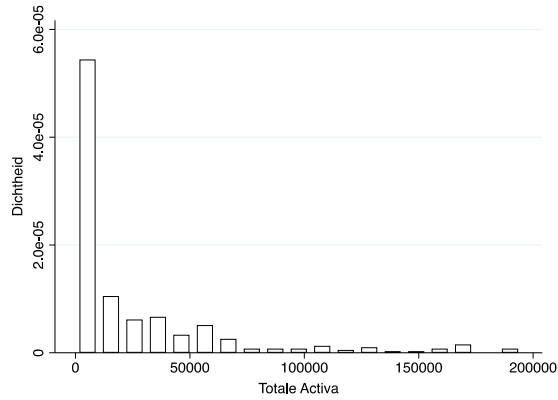
- Hsiao, C. (2014). *Analysis of panel data*. University of Southern California, Los Angeles.
- Hung, K. P., & Chou, C. (2013). The impact of open innovation on firm performance: The moderating effects of internal R&D and environmental turbulence. *Technovation*, 33(10), 368–380.
- Iezza, P. (2010). *Financial Sustainability of Microfinance Institutions (MFIs): An Empirical Analysis Master's Thesis*. Copenhagen Business School, Frederiksberg.
- Irving Levin Associates. (2010, 26 Maart). Ten-Year Data on Pharmaceutical Mergers and Acquisitions. *Fierce Biotech*. <https://www.fiercebiotech.com/biotech/ten-year-data-on-pharmaceutical-mergers-and-acquisitions-from-dealsearchonline-com-reveals>
- Jones, B. (2020). A Significant Rise in Health Care R&D Provides Investors Opportunity in this Sector. *Nasdaq*. <https://www.nasdaq.com/articles/a-significant-rise-in-health-care-rd-provides-investors-opportunity-in-this-sector-2020-07>
- Kainulainen, S. (2014). *Encyclopedia of Quality of Life and Well-Being Research*. Springer, Prince George.
- Kotabe, M., Srinivasan, S., & Aulakh, P. (2002). Multinationality and Firm Performance: The Moderating Role of R&D and Marketing Capabilities. *Journal of International Business Studies*, 33(2), 79–97.
- Lagasse, J. (2019, 31 December). Investment in medical and health R&D not keeping up with needs of nation, report finds. *HealthCare Finance*. <https://www.healthcarefinancenews.com/news/investment-medical-and-health-rd-not-keeping-needs-nation-report-finds>
- Lavrakas, P. J. (2008). *Encyclopedia of survey research methods*. Sage, Los Angeles.
- Lin, B., Lee, Y., & Hung, S. (2006). R&D intensity and commercialization orientation effects on financial performance. *Journal of Business Research*, 59(6), 79–85.
- Lusk, E. J., Halperin, M., & Heiling, F. (2011). A note of power differentials in data preparation between trimming and winsorizing. *Business Management Dynamics*, 1(2), 23–31.

- Munster, W. (2017). *WRDS: Data Items*. WWU Münster. [https://www.wiwi.uni-muenster.de/uf/sites/uf/files/2017\\_10\\_12\\_wrds\\_data\\_items.pdf](https://www.wiwi.uni-muenster.de/uf/sites/uf/files/2017_10_12_wrds_data_items.pdf)
- Padgett, R. C., & Galan, J. I. (2010). The effect of R&D intensity on corporate social responsibility. *Journal of Business Ethics*, 93(3), 407–418.
- Pantagakis, E., Terzakis, D., & Arvanitis, S. (2012). R&D investments and firm performance: an empirical investigation of the high technology sector (software and hardware) in the E.U. *Electronic Journal*.
- Parcharidis, E.G., & Varsakelis, C. V. (2010). R&D and Tobin's q in an Emerging Financial Market: The Case of the Athens Stock Exchange. *Managerial and Decision Economics*, 31(5), 353–61.
- Roberts, M.R., & Whited, T.M. (2013). Endogeneity in empirical corporate finance1. *Handbook of the Economics of Finance*, 493–572.
- Santen, R. V. (2021). De S&P 500 index; alles over de populairste index op Wall Street. *Lynx*. <https://www.lynx.nl/kennis/artikelen/sp-500-index-alles-populairste-speler/>
- Sauaia, A.C.A., & Castro, F.H.F. (2002). Is the Tobin's q a Good Indicator of a Company's Performance. *Association for Business Simulation and Experiential Learning*.
- Science & Engineering Indicators. (2018). Production Patterns and Trends of Knowledge- and Technology-Intensive Industries. *Nation Science Foundation*. <https://nces.nsf.gov/pubs/nsb20205/production-patterns-and-trends-of-knowledge-and-technology-intensive-industries>
- Shefer, D. & Frenkel, A. (2005). R&D, firm size and innovation: an empirical analysis. *Technovation*, 25(1), 25–32.
- Sher, P.J., & Yang, P.Y. (2005). The effects of innovative capabilities and R&D clustering on firm performance: the evidence of Taiwan's semiconductor industry. *Technovation*, 25(2), 33–43.
- Shin, N., Kraemer, K. L., & Dedrick, J. (2009). R&D, value chain location and firm performance in the global electronics industry. *Industry & Innovation*, 16(3), 315-330.

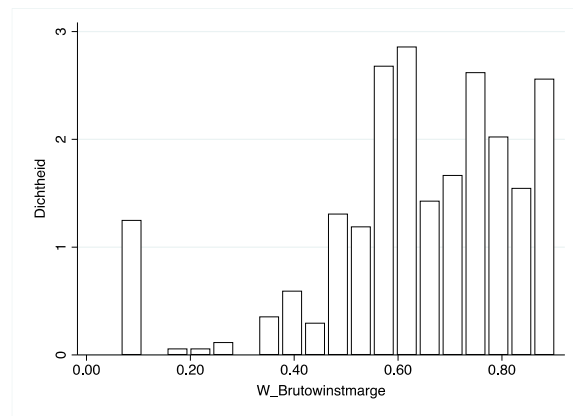
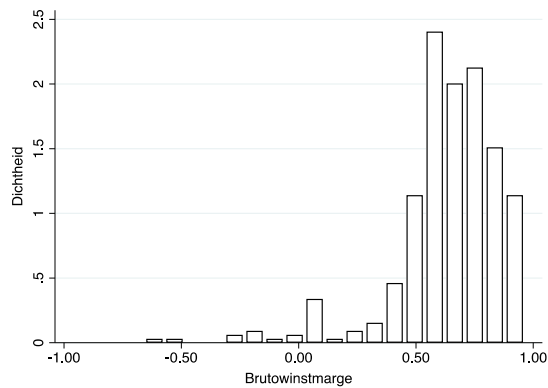
- Soppe, A. (2004). Sustainable corporate finance. *Journal of Business Ethics*, 53(3), 213– 224.
- Standard & Poor's. (2011). Standard & Poor's Compustat Xpressfeed Understanding the Data. *The McGraw-Hill Companies*. [https://wrds-www.wharton.upenn.edu/documents/841/Understanding\\_The\\_Data\\_57jARh2.pdf](https://wrds-www.wharton.upenn.edu/documents/841/Understanding_The_Data_57jARh2.pdf)
- Vithessonthi, C., & Racela, O. C. (2016). Short-and long-run effects of internationalization and R&D intensity on firm performance. *Journal of Multinational Financial Management*, 34(2), 28–45.
- Welch, J. (2020). *Pulse of the industry: Medical technology report 2020* (Report no. 006754-20Gbl). EY company Report.
- Yeh, M. L., Chu, H. P., Sher, P. J., & Chiu, Y. C. (2010). R&D Intensity, Firm Performance and the Identification of the Threshold: Fresh Evidence from the Panel Threshold Regression Model. *Applied Economics*, 4(2), 389–401.

# Appendix A Normaal Verdeling

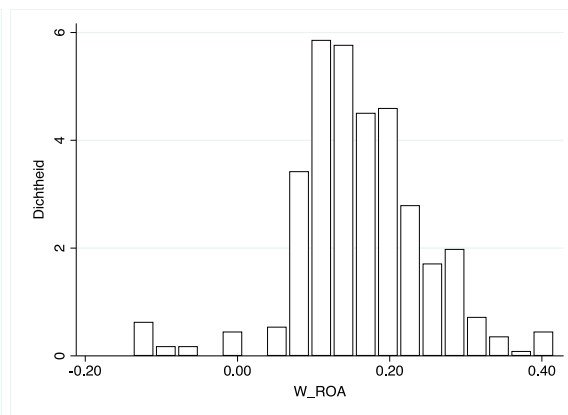
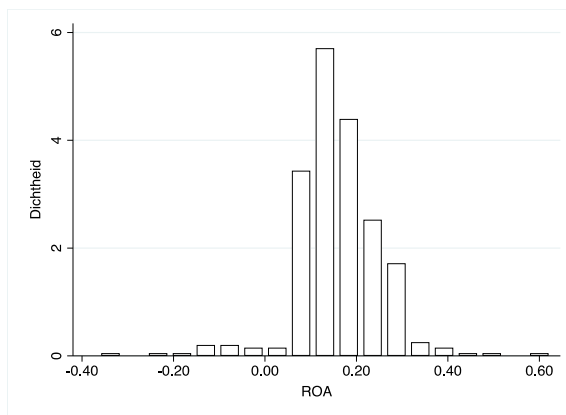
Verdeling van de variabele "Totale activa"



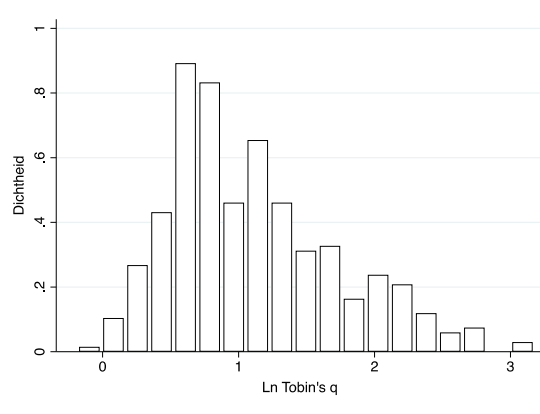
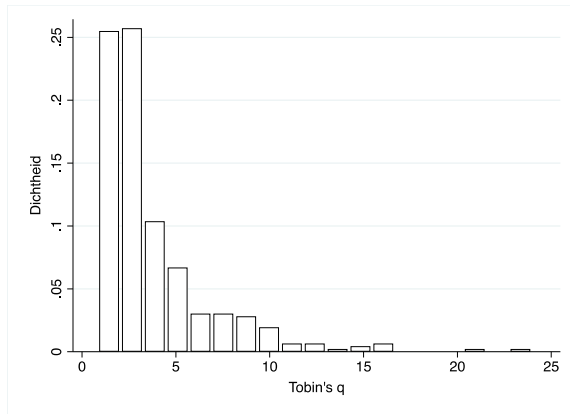
Verdeling van de variabele "Brutowinstmarge"



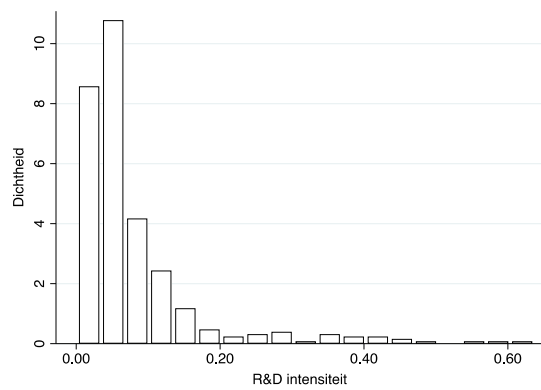
Verdeling van de variabele "ROA"



*Verdeling van de variabele Tobin's q*



*Verdeling van de variabele R&D intensiteit*



# Appendix B Breusch-Godfrey Regressie

*Lineaire regressie Breusch Godfrey hypothese 1: met fouteterm U als afhankelijke variabele en de vertragingen van U als onafhankelijke variabele*

	Fouteterm U
Vertraging 1 van U	,395*** (,173)
Vertraging 2 van U	,815 (,151)
Vertraging 3 van U	-,039* (0,157)
Vertraging 4 van U	-,041** (,144)
Vertraging 5 van U	,559 (,197)
Vertraging 6 van U	-,377*** (,257)
Vertraging 7 van U	-,218*** (,129)
Bedrijfsomvang	,000 (,001)
Constante	,004 (,013)
Observaties	114
R kwadraat	,914

*Robuuste standaardfouten tussen de haakjes*

\*\*\*  $p < ,01$ ; \*\*  $p < ,05$ ; \*  $p < ,1$

*Lineaire regressie Breusch Godfrey hypothese 2: met fouteterm U als afhankelijke variabele en de vertragingen van U als onafhankelijke variabele*

	Fouteterm U
Vertraging 1 van U	,876*** (,194)
Vertraging 2 van U	,468 (,193)
Vertraging 3 van U	-,337* (,208)
Vertraging 4 van U	,211** (,171)
Vertraging 5 van U	-,200 (,305)
Vertraging 6 van U	,624*** (,402)
Vertraging 7 van U	-,607*** (,127)
Bedrijfsomvang	,016 (,017)
Constante	-,011 (,012)
Observaties	114
R kwadraat	,986

*Robuuste standaardfouten tussen de haakjes*

\*\*\*  $p < ,01$ ; \*\*  $p < ,05$ ; \*  $p < ,1$

*Lineaire regressie Breusch Godfrey hypothese 3: met fouteterm U als afhankelijke variabele en de vertragingen van U als onafhankelijke variabele voor de regressie met w\_ROA*

	Fouteterm U
Vertraging 1 van U	,940*** (,095)
Vertraging 2 van U	-,317** (,122)
Vertraging 3 van U	,281*** (,105)
Vertraging 4 van U	-,270*** (,088)
Vertraging 5 van U	,307*** (,074)
Vertraging 6 van U	-,057 (,098)
Vertraging 7 van U	-,048 (,058)
R&D intensiteit	-,043 (,038)
w_Brutowinstmarge	,026 (,017)
Bedrijfsomvang	-,005** (,002)
Constante	,109** (,054)
Observaties	114
R kwadraat	,791

*Robuuste standaardfouten tussen de haakjes*

\*\*\*  $p < ,01$ ; \*\*  $p < ,05$ ; \*  $p < ,1$



*Lineaire regressie Breusch Godfrey hypothese 3: met fouteterm U als afhankelijke variabele en de vertragingen van U als onafhankelijke variabele voor de regressie met Ln Tobin's q*

	Fouteterm U
Vertraging 1 van U	,976*** (,098)
Vertraging 2 van U	-,057 (,149)
Vertraging 3 van U	,336** (,157)
Vertraging 4 van U	-,241* (,141)
Vertraging 5 van U	-,177 (,127)
Vertraging 6 van U	,239* (,123)
Vertraging 7 van U	-,036 (,09)
R&D intensiteit	-1,022*** (,302)
w_Brutowinstmarge	,138 (,111)
Bedrijfsomvang	-,046*** (,017)
Constante	1,049*** (,393)
Observaties	114
R kwadraat	,797

*Robuuste standaardfouten tussen de haakjes*

\*\*\*  $p < ,01$ ; \*\*  $p < ,05$ ; \*  $p < ,1$

# Appendix C Multicollineariteit VIF test

*VIF test met winsorized ROA als afhankelijke variabele*

	VIF	1/VIF
Ln Tobin's Q	1,907	0,524
R&D intensiteit	1,499	0,667
Bedrijfsomvang	1,476	0,677
w_Brutowinstmarge	1,139	0,878
Gemiddelde VIF	1,505	

*VIF test met R&D intensiteit als afhankelijke variabele*

	VIF	1/VIF
Ln Tobin's Q	1,547	0,646
Bedrijfsomvang	1,476	0,678
w_ROA	1,346	0,743
w_Brutowinstmarge	1,278	0,783
Gemiddelde VIF	1,412	

*VIF test met de natuurlijk logaritme van de Tobin's q als afhankelijke variabele*

	VIF	1/VIF
w_Brutowinstmarge	1,281	0,781
w_ROA	1,267	0,790
R&D intensiteit	1,145	0,874
Bedrijfsomvang	1,144	0,874
Gemiddelde VIF	1,209	