

# *PROJEKT ORIENTERET STUDIE*

*PRAKTIK HOS SDU ANALYTICS*



*Kilde: <https://www.sdu.dk/da>*

## Indholdsfortegnelse

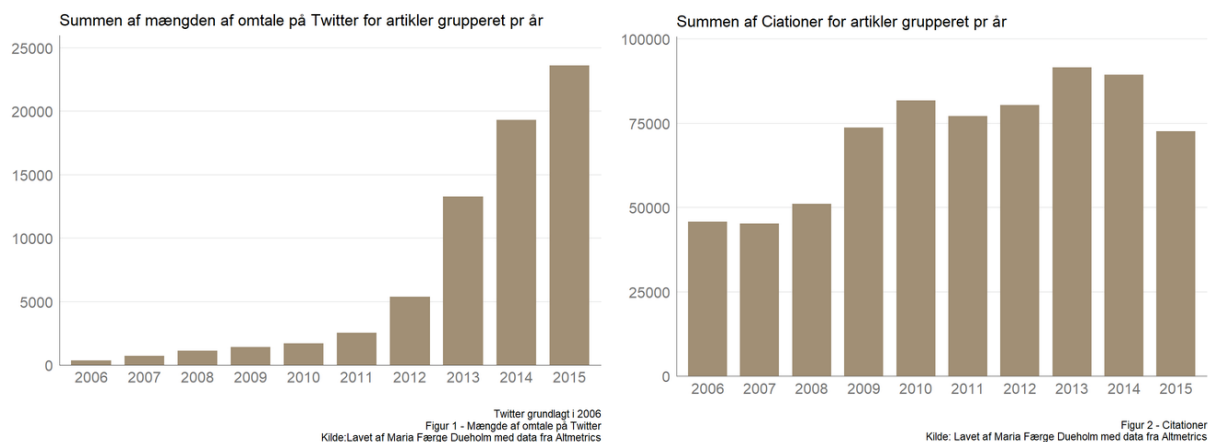
<b>1.1</b>	<b>Introduktion.....</b>	<b>3</b>
<b>1.2</b>	<b>Baggrund .....</b>	<b>4</b>
1.2.1	Generelt .....	4
1.2.2	Klassisk bibliometri og citationer .....	4
<b>1.3</b>	<b>Data.....</b>	<b>5</b>
<b>1.4</b>	<b>Empirisk strategi .....</b>	<b>9</b>
<b>1.5</b>	<b>Resultater .....</b>	<b>10</b>
<b>1.6</b>	<b>Begrænsninger .....</b>	<b>11</b>
<b>1.7</b>	<b>Perspektivering .....</b>	<b>13</b>
<b>1.8</b>	<b>Anbefaling af en alternativ fremgangsmåde .....</b>	<b>14</b>
<b>1.9</b>	<b>Diskussion.....</b>	<b>15</b>
<b>1.10</b>	<b>Konklusion .....</b>	<b>15</b>
<b>2</b>	<b><i>Bibliometri .....</i></b>	<b><i>17</i></b>
<b>3</b>	<b><i>Appendix oversigt.....</i></b>	<b><i>18</i></b>
3.1	Appendix 1 - Summen af Citationer for artikler grupperet pr år .....	18
3.2	Appendix 2 - Model af forholdet mellem sociale effekter og forsknings effekter .....	19
3.3	Appendix 3 - Koefficient estimater for alle variablerne X repræsenterer .....	20
3.4	Appendix 4 - Subgroups: Campus beliggenhed .....	23
3.5	Appendix 5 - Subgroups: BFI .....	24
3.6	Appendix 6 - Test for simultanitet .....	25

Analysen har til formål at undersøge om der er sammenhæng mellem mængde af omtale på sociale medier (Twitter) og citationer målt som klassisk akademisk bibliometri. Opgaven vil være bygget op som følgende: Introduktion, baggrund, data, empirisk strategi, resultater, begrænsninger, perspektivering, anbefaling af en alternativ fremgangsmåde, diskussion, konklusion på analysen og til sidst en refleksion til de kompetencer opgaven har givet mig.

## 1.1 Introduktion

Sociale medier er efter hånden blevet en stor del af hverdagen, derfor er det ikke uvæsentligt at undersøge om omtale af forskningspubliceringen på sociale medier fører til bedre performance. Jeg har udelukkende kigget på Twitter som socialt medie, da det er det medie som Altmetrics har mest data fra. Denne analyse er lavet ud fra den definition at en citation er følgende: citationer er et mål for bedre performance, målt som klassisk bibliometri. Mængde af omtale på Twitter er følgende: Antallet af Twitter opslag hvor i artiklen nævnes på dataindsamlingstidspunktet.

På Figur 1 kan der ses udviklingen af mængde af omtale på Twitter af videnskabelige artikler skrevet af forskere på SDU på Twitter fra år 2006 til 2015. Twitter blev grundlagt i 2006 ([Twitter](#)), siden da er omtalen af videnskabelige artikler skrevet af forskere på SDU steget. Mængden af omtale på Twitter af er for alvor begyndt at stige i 2013, det er også i 2013 der klart er den største stigning. På figur 2, kan der ses udviklingen af citationer, for de videnskabelige artikler, der er skrevet af forskerne på SDU. Det kan ses på figur 1 at twitteropslag med artikler er steget markant. Det kan ses på figur 2 med citationer er der er en generel stigning, men den er lidt mere svingende i udviklen.



En af grundene til at antallet af citationer i år 2014 og 2015 på Figur 2 er lavere end i 2013, kan være at det tager flere år at indsamle citationer, og derfor har de videnskabelige artikler der er publiceret i 2013 haft henholdsvis 1 og 2 år mere til at blive citeret end årene 2014 og 2015 (Catharina Rehn, 2014) [AU bibliotek](#). Som det kan ses i Appendix 1 er antallet af publikationer stigende hvert år. Det kan derfor ikke være en

forklaring til at summen af citationer for artikler grupperet pr år er større i 2013 end de to efterfølgende år. Dog kan der også være andre faktorer der spiller ind her.

Altmetrics kan blandt andet bruges af forskere, til at se hvordan deres research klarer sig, samt hvor de bliver omtalt. Altmetrics indsamler data fra sociale medier, research blogs, nyheder, dokumenter om offentlig politik, diskussioner på forskningsblogs og flere til ([Altmetrics](#)). Dette kan være til gavn for den enkelte forsker for at få informationer om, hvem der interesserer sig for deres forskning. Altmetrics er en anden måde at hente informationer om artiklen, ned på. De primære forskel på citationer og Altmetrics er at Altmetrics har flere informationer med, så som mængde af omtale på sociale medier, nyheder osv.

## 1.2 Baggrund

I dette afsnit, vil jeg komme ind på hvad tidligere videnskabelige artikler har under søgt. Jeg vil inddele dette afsnit i 2 under afsnit, generelt og til sidst Klassisk bibliometri og citationer

### 1.2.1 Generelt

Der er ikke lavet meget forskning inden for dette emne. Det meste af forskningen er lavet med fokus på hvilke faktorer, der giver citationer (Tahamtan et al., 2016), men ikke om et højere aktivitetsniveau på sociale medier giver flere citationer. Der er dog nogle enkelte der har undersøgt emnet, trods en lidt anden vinkel end denne opgave. Eysenbach (2011) kigger på om Tweets kan forudsige citationer. Han finder at mængden af omtale på Twitter og citationer er "noget" korreleret<sup>1</sup>. Samtidig fastslår han, at man ikke skal forvente perfekt korrelation, da mængden af omtale på Twitter og citationer måles på forskellige koncepter, samt de hver især henvender sig til forskelligt publikum. Mængden af omtale på Twitter skal ifølge Eysenbach (2011) ses som en måling for social effekt og viden, altså hvor hurtigt ny viden bliver observeret af offentligheden, samt hvad offentligheden er opmærksom på. Citationer derimod skal primært ses som en måling for videnskabelig effekt.

### 1.2.2 Klassisk bibliometri og citationer

Der er stor forskel på hvor mange citationer de forskellige videnskabelige artikler får. Det er langt største delen af de videnskabelige artikler, der er knap omtales, mens nogle få er meget citerede. Der er en general holdning til at artikler der er særdeles højt citeret, er af bedre kvalitet end de artikler der er lavt citeret (Garfield, 2006, Bornmann et al., 2012). Det er vigtigt at nævne, at citationer ikke altid er lig positiv omtale. Nogle artikler bliver citeret, hvis der er andre forskere, kritiserer den citerede artikel. Tahamtan et al. (2016) har undersøgt hvilke faktorer som påvirker hyppigheden af citationer. De har inddelt faktorerne i 3 kategorier: (1) Artikelrelaterede faktorer, (2) Journal relaterede faktorer samt (3) Person relaterede faktorer.

---

<sup>1</sup> Se Appendix 2, hvor yderligere forklaring

Jeg vil beskrive disse nærmere i afsnittet vedrørende data. Tahamtan et al. (2016) finder at nogle faktorer er stærkere indikatorer for citationer end andre. Faktorer så som kvalitet af artiklen, faktorer for journalens indvirkning og antal af forskere er stærkere indikatorer end forskerens køn, alder og race. Bornmann et al. (2012) bakker op omkring antallet af forskere, journalens sprog samt fagområdet popularitet er stærke indikatorer.

### 1.3 Data

I dette afsnit, vil jeg komme ind på hvilke databaser jeg har brugt, samt hvordan jeg har tilpasset dataet.

Det data der er brugt til denne analyse, er mikrodata, fra Altmetrics og fra HR på Syddansk Universitet. Alt data fra Altmetrics er på artikel niveau, og er et øjebliks billede, af hvordan den enkelte artikel har klaret sig i forhold citationer frem til den dato jeg hentede dataene ned (d. 10-10-2020). Data fra Altmetrics indeholder informationer fra 58.765 artikler, og indeholder data fra 01-01-1810 til 15-03-2021.

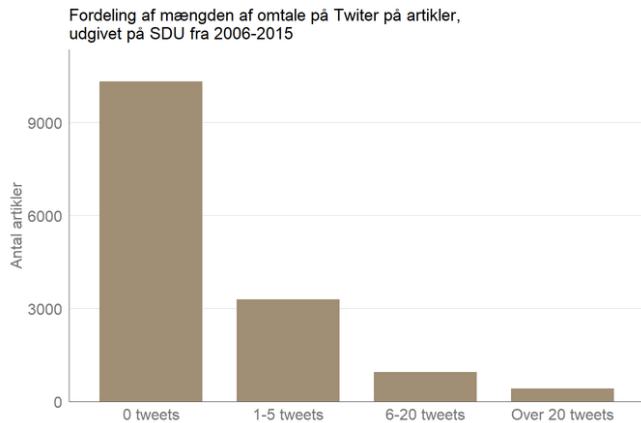
Datasættet fra HR på SDU indeholder informationer omkring 2658 personer. Jeg har desuden suppleret det nævnte data med data fra Uddannelses- og Forskningsministeriet, for at få Den Bibliometriske Forskningsindikator (BFI) for journalerne. BFI-listen indeholder de tidsskrifter, bogserier, konferenceserier og forlag som er pointgivende i Den Bibliometriske Forskningsindikator. Den pointgivende skala går fra 1-3, hvor 3 er bedst.

Jeg har været interesseret i at få informationer om artiklernes forfatter med ind. Nogle af artiklerne har der været flere forfattere bag, derfor har jeg opdelt variablen forfattere, så der udelukkende har været én forfatter per artikel, dette har resulteret i dupliketter af artikler. Jeg har været nødt til dette, for at få de personlige faktorer ind, som køn, alder osv for forfatterne. For igen at få det på artikel niveau har jeg efterfølgende grupperet per artikel og regnet et gennemsnit ud på tværs af forfatterne per artikel. Dette kommer ikke til at give et helt præcist billede, det er dog det mest retvisende, når det er på artikel niveau. Jeg har desuden også tilføjet en variablen for antal af forfattere pr artikel. Antal forfattere per artikel, er kreeret da det ifølge Bornmann et al. (2012) og Tahamtan et al. (2016) som tidligere nævnt er en faktor, der er en strækindikator. Data før 2006 er ikke anvendt, da Twitter først blev grundlagt i 2006. Data efter 2015 anvendes ikke, da citationer falder skævt i forhold publicering, altså det tager en del år for en database at indsamle citationer ([AU bibliotek](#), (Catharina Rehn, 2014)). Da jeg har fokuseret på de artikler, der er skrevet på SDU's fem fakulteter, har jeg skåret de artikler væk der er blevet udgivet under følgende "universitetsbiblioteket" (147 artikler) samt "fælles område" (4 artikler). Fordi de ikke indgår i SDU's fem fakulteter. Desuden har jeg ikke en alders variable på forfatterne, derfor har jeg lavet endnu en variabel "år siden første artikel" denne variable giver et acceptabelt estimat på alderen på forskeren. Grunden til at en

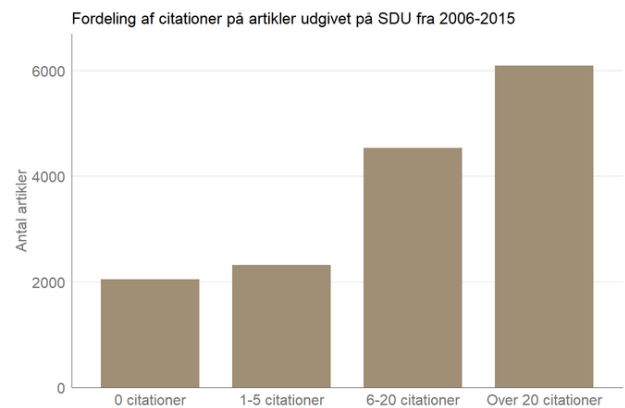
alders variable vil have været interessant, er Tahamtan et al. (2016) har alder med som i hans teori, omkring hvad der påvirker citationer. Da denne nye variable fortæller noget om, hvor lang tid forskeren har været i den videnskabelige verden, giver den et godt billede af hvor længe de har været i gang inden for deres felt. Dog skal der tages højde for at der kan komme bias, da den ikke giver det helt præcise billede af forskerens alder, på den enkelte artikel, da det som tidligere er nævnt, vil personlige faktorer, være angivet som en andel, altså et gennemsnit af forfatterens alder, hvis der har været flere. Barnett and Fink (2008) finder at yngre forfattere (under 36 år) får flere citationer end forfattere over 41 år. Derfor er dette en vigtig faktor i forhold til påvirkning af citationer. Den bias der kan opstå her, kaldes "Measurement error bias", denne bias opstår når den indsamlede data til variable er upræcis. Dog vil der være en risiko for at der kan opstå multikollinearitet mellem denne faktor ("Alder" / "år siden første artikel") og variabelen karrige stige, da jo højere oppe på karrierestigen jo større er sandsynligheden for at forfatteren også er ældre. Til sidst har jeg fjernet alle NA værdier.

I min analyse indledte jeg med at kreerer flere forskellige plots (Disse kan ses længere nede i dette afsnit). For at skabe et overblik over mit data, samt for at se om fordelingen af mængden af omtale på Twitter og citationer var ligeligt fordelt, har jeg lavet barcharts, for hen holdvis fordelingen af mængden af omtale på Twitter og citationer. Hvor jeg har inddelt antal af citationer og mængden af omtale på Twitter i de samme grupper, som er følgende: 0, 1-5, 6-20 og over 20. Dette har jeg gjort for at få et indblik i hvordan fordelingen ser ud i mit data, jeg bruger disse informationer til at have med i mine overvejelser når jeg går i gang med min analyse, samt konkludere der på.

Efterfølgende har jeg lavet et scatterplot, hvor jeg har citationer på y-aksen og mængden af omtale på Twitter på x-aksen. Dette har jeg gjort for at se om mit data var lineært eller eksponentielt. Jeg har desuden tegnet to lineære grafer på scatterplottet, en med outliers og en uden outliers. Dette gør en stor forskel for  $R^2$  værdien. Det er vigtigt at opdage og muligvis behandle outliers, da det kan give kraftige bias eller ændre fitted for estimerne. Der ville med fordel kunne laves en analyse, både med og uden outlier, for at se om der er stor forskel. Jeg har desuden også fjernet de værste outlier da dette kan give forringelse af mine regressioner. En outlier i denne opgave kan fx være en artikel, der har fået meget omtale på Twitter, og få citationer. For at få en bedre visualisering af mit scatterplot, har jeg taget den naturlige logaritme af både citationer og mængden af omtale på Twitter. Jeg har valgt at fjerne alle artikler der har enten 0 citationer, 0 omtale på Twitter eller begge dele. Dette er gjort, da jeg ønsker at undersøge de artikler, der enten har fået citationer eller omtale på sociale medier. Dette kan ikke undersøges med artikler, der har fået 0 citationer og / eller 0 omtale.

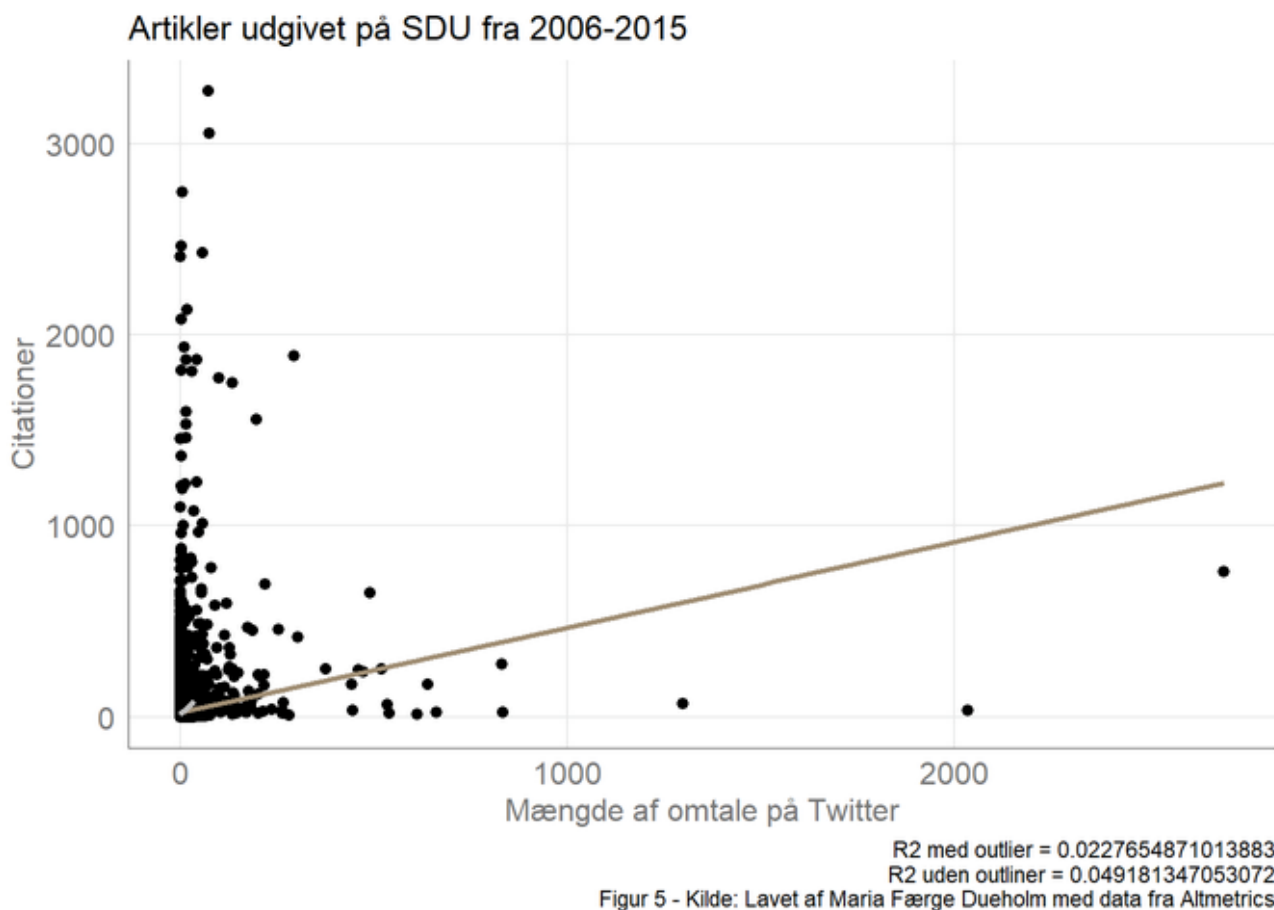


Figur 3 - Kilde: Lavet af Maria Færge Dueholm med data fra Altmetrics



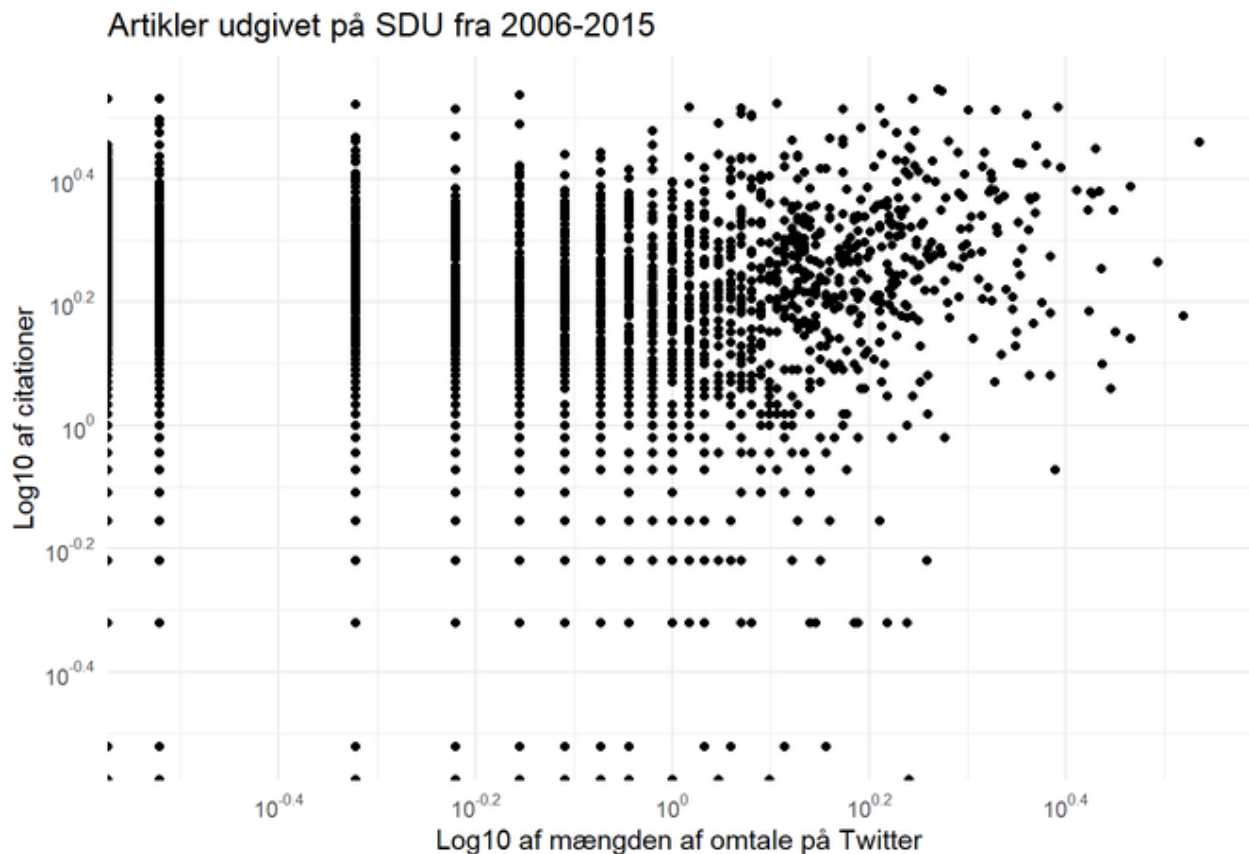
Figur 4 - Kilde: Lavet af Maria Færge Dueholm med data fra Altmetrics

Det kan ses ud fra figur 3 og figur 4, at fordelinger over mængden af omtale på Twitter og citationer ikke er ligeligt fordelt. På figur 3 med mængden af omtale på Twitter er der klart flest artikler i gruppen med 0 Tweets, og færrest i gruppen med over 20 tweets. Dette er ikke gældende for figur 4 grafen med citationer. Her er det gruppen med over 20 citationer, der er størst og gruppen med 0 citationer er mindst.



På scatter plottet oven over kan det ses, at langt største delen af artiklerne klumper sammen med i venstre hjørne omkring 0,0. Dog er der væsentligt flere artikler som får 0 tweets og mange citationer end der er

artikler som får 0 citationer og mange tweets. Jeg har opsat to lineære ligninger på scatter plottet, for at se, hvad der sker med  $R^2$  værdien når jeg medtager outlier og fra vælger outlier. Den lineære ligning uden outlieres er svær at se, men nederst i venstre hjørne er der en lille hvid streg. Jeg har lavet denne graf for at se hvordan jeg skal estimere mit data, og for at se om der er nogen funktionel form der skal på mine funktioner. Jeg kan ud fra denne graf se at mit data ikke er lineært, det er også derfor at  $R^2$  ikke er høj.



Figur 6 - Kilde: Lavet af Maria Færge Dueholm med data fra Altmetrics

Det interessante ved log scatter plot er at, der ikke er nogle observationer nede i højre hjørne. Altså er der ikke nogen artikler med meget omtale på Twitter og lave citationer. Formålet med dette plot er at få mere spredning på mit data, der gør at jeg får et bedre overblik, i forhold til et "normal" scatter plot. Jeg har lavet dette plot for at forstå mit data, da jeg på det forrige scatter plot kunne se at langt det meste af mit data var samlet nede i venstre hjørne, jeg vil gerne kunne se om der er en forskel på "klumpen" i hjørnet. Jeg kan altså med dette plot se om klumpen reelt er samlet eller om der er forskel, hvilket der er til en vis grad. De er stadig samlet, men der er som sagt mere spredning mellem punkterne.



## 1.4 Empirisk strategi

I dette afsnit vil jeg præsentere min empiriske strategi.

For at vise hvorvidt omtale af forsknings publicering på sociale medier (Twitter) efterfølgende fører til bedre performance mål som klassisk akademisk bibliometri, har jeg først brugt en Ordinary least squares regression (OLS) til at analysere min citations variable på variablen for omtale på Twitter. Dette opfanger den overordnede effekt af tweets. Efterfølgende har jeg tilføjet en række af kontrol variabler, i grupper. Disse kontrol variabler og grupper følger teorien beskrevet i Tahamtan et al. (2016), som tidligere er nævnt i mit afsnit baggrund..

Den endelige ligning:

$$Citationer_i = \beta_0 + \beta_1 tweet_i + \sum \beta X_i + \epsilon_i$$

Hvor  $\epsilon$  er et fejllid og  $i$  er artiklen

Citationer er antal citationer artiklen har fået på dataindsamlingstidspunktet. leveret af Altmetrics, via dimensions-citations data. Hvor de måler, hvor mange citationer den enkelte artikel har fået, dette bliver hele tiden opdateret via deres hjemmeside.

Jeg har efterfølgende tilføjet en mængde kontrol variabler. Som der er grupperet i 3 grupper. Disse tre grupper følger Tahamtan et al. (2016) teori. Som nævnt tidligere i afsnittet, baggrund, vil jeg nu nærmere beskrive, hvilke variabler jeg har valgt at medtage inden for hver af de tre grupper. Jeg har følgende faktorer med: Kategori 1 "artikel relaterede faktorer": Kvalitet af artiklen<sup>2</sup>, Fagområde<sup>3</sup>, Alder på artiklen. Kategori 2: "journal relaterede faktorer": Form for publikation. Kategori 3 "Person relaterede faktorer": Antal forfattere, forfatterens akademiske rang, forfatterens land<sup>4</sup>, køn, alder<sup>5</sup>, produktivitet<sup>6</sup>, finansiering. Som det fremgår, har jeg klart flest faktorer fra kategori 3, med i min analyse. Altså er det specielt kategorierne "artikel relaterede faktorer" og "Journal relaterede faktorer", der halter efter. Jeg har ikke haft muligheden for at have alle de faktorer teorien bygger på.

---

<sup>2</sup> Målt på BFI

<sup>3</sup> Som institut

<sup>4</sup> Målt som dansk / ikke dansk

<sup>5</sup> Målt som år siden første artikel

<sup>6</sup> Målt som antal af skrevne artikler

## 1.5 Resultater

I dette afsnit vil jeg kort opridse de resultater jeg har fået fra min analyse. Her vil jeg præsentere mine regressioner.

	Tabel 1: Afhængig variable			
	Citationer			
	(1)	(2)	(3)	(4)
Mængde af omtale på Twitter	0.403*** (0.035)	0.374*** (0.033)	0.373*** (0.033)	0.360*** (0.033)
Artikel Relaterede Faktorer		X	X	X
Journal Relaterede Faktorer			X	X
Person Related Factors				X
Observations	4,545	4,545	4,545	4,545
R <sup>2</sup>	0.029	0.167	0.172	0.187

Note:

$p < 0.1$ ;  $p < 0.05$ ;  $p < 0.01$

Appendix 3, viser alle koefficientestimerne for alle variablerne X repræsenterer i regressionerne  
Kilde: Lavet af Maria Færge Dueholm, med data fra Altmetrics, HR-data fra SDU samt data fra Uddannelses- og Forskningsministeriet

I tabel 2, kan der ses de OLS regressioner jeg har lavet. I kolonne (1) har vi den overordnede effekt af mængden af omtale på Twitter. I kolonne (2) tilføjer jeg de artikel relaterede faktorer, det kan ses at effekten af mængden af omtale på Twitter falder. Dette er også gældende når der tilføjes journal relaterede faktorer og igen når de personlige relaterede faktorer tilføjes. Hver af de tre faktorer kategorier påvirker stort set mængden af omtale på Twitter lige meget. Vi kan se at estimatet falder, i takt med der kontrolleres for flere ting, dette fortæller at jo tættere vi kommer på den sandeværdi, jo lavere bliver estimatet. I kolonne (1) kan vi se at ved en ekstra omtale på Twitter vil citationer stige med 0.403. Det kan ses jo flere variabler der kontrolleres for, jo lavere bliver påvirkningen af mængden af omtale på Twitter.

Jeg har desuden lavet to subgroups, den ene er opdelt efter placering, da SDU har 6 forskellige campusser: Odense, København, Kolding, Slagelse, Esbjerg og Sønderborg. Jeg har gjort dette for at se, om der er forskel og hvor stor en forskel der er på hvor en artikel bliver udgivet. Den anden subgroup er efter niveau på journal, dette er efter BFI<sup>7</sup>. Resultaterne for disse subgroups kan ses i appendix 4 og 5. Jeg har lavet subgroups analyse, fordi dette gør det muligt at undersøge, hvordan de forskellige subgroups klarer sig.

<sup>7</sup> Niveau på journal bliver målt via BFI. BFI er en pointgivende skala går fra 1-3, hvor 3 er bedst.

## 1.6 Begrænsninger

I dette afsnit vil jeg belyse de begrænsninger, der har været igennem denne analyse, samt hvad det har gjort for mine resultater.

Der har været omitted variable bias i denne analyse. Som det kunne ses i mit afsnit omkring Empirisk strategi, mangler jeg en hel del kontrol variabler i forhold til at kunne følge Tahamtan et al. (2016) teori om, hvilke faktorer der påvirker citationer. På tabellen neden under fremgår det hvilke faktorer jeg har kunne medtage og hvilke jeg ikke har haft muligheden for at medtage. Dem med streg under er de variabler jeg har med i min analyse. Som det kan ses på tabellen neden under, er der en lang række faktorer jeg ikke har haft muligheden for at medtage i min analyse. Jeg har en klar forventning at mine resultater havde været mere troværdige, hvis jeg havde haft flere brugbare kontrol variabler. Det vil have en effekt på mine estimater, da for eksempel ved at undlade nyhed værdi, vil give mit estimat negativ bias. Ved at have et estimat med negativ bias, vil det være muligt at estimatet er underestimeret.

Kategori 1	Kategori 2	Kategori 3
Artikel relaterede faktorer	Journal relaterede faktorer	Person relaterede faktorer
<ul style="list-style-type: none"> <li>- <u>Kvalitet af artiklen</u></li> <li>- Nyhedsværdi</li> <li>- Fagområdet popularitet</li> <li>- <u>Fagområde</u></li> <li>- Metodologi</li> <li>- Dokumenttype</li> <li>- Studie design</li> <li>- Karakteristik af resultater og diskussion</li> <li>- Brugen af figurer og appendix</li> <li>- Længde på artiklen</li> <li>- <u>Alder på artiklen</u></li> <li>- Tidlig citation af artiklen</li> </ul>	<ul style="list-style-type: none"> <li>- Faktorer for journal indvirkning</li> <li>- Journalens sprog</li> <li>- Område af Journal</li> <li>- <u>Form for publikation</u></li> </ul>	<ul style="list-style-type: none"> <li>- <u>Antal af forfattere</u></li> <li>- Forfatterens omdømme</li> <li>- <u>Forfatterens akademiske rang</u></li> <li>- Selvcitation</li> <li>- Internationalt og nationalt samarbejde</li> <li>- <u>Forfatterens Land</u></li> <li>- <u>Køn</u></li> <li>- Race</li> <li>- <u>Alder</u></li> <li>- <u>Produktivitet</u></li> <li>- Organisatoriske træk</li> <li>- <u>Finansiering</u></li> </ul>

- Tilgængelighed og papirets synlighed.		
---	--	--

Der er to variabler, hvis jeg havde muligheden, gerne ville have tilføjet til min analyse. En netværks variable, der viste hvor mange steder en forsker har arbejdet, jeg har en teori om: jo flere jobs og bosteder, jo større netværk. Dette må om alt andet give større mulighed for at få sine publikationer ud blandt offentligheden. Den anden variable jeg gerne ville have tilføjet til min analyse, er en variabel, der fortæller noget den enkelte forskers egen aktivitet på sociale medier, med henblik på forskerens egne artikler. Dette kunne have været interessant i forhold til at kigge på om, hvis en forsker øger sin social medieaktivitet, om forskeren så også får mere omtale både på sociale medier men også om deres artikler bliver mere citeret. Dette behøves nødvendigvis ikke at være med henblik på om de omtaler deres egne artikler. Dog vil disse to skulle gøres i andele i min analyse, da de begge er personlige faktorer. Dette bringer mig videre til den næste begrænsning.

For de personlige faktorer, har jeg været nødsaget til at beregne en afdel, hvis der har været flere forskere til at skrive artiklen. Dette har været medvirkende til at mine resultater kan være misvisende. Da langt de fleste andele for personlige faktorer, ikke vil være retvisende. Hvis jeg havde lavet analysen på individ niveau, kunne jeg have arbejdet endnu mere i dybden, dog skal det nævnes, hvis jeg havde lavet det på individ niveau ville der være dubletter af artiklerne, hvilket kunne give mine resultater bias.

En anden begrænsning denne analyse har, at Twitter er det eneste sociale medie jeg kigger på, da er det eneste medie Altmetrics har fuld adgang til. Google+, Pinterest, LinkedIn har de ikke længere adgang til, da de respektive medier har lukket for deres data strøm, og dermed kan Altmetrics ikke indsamle nye data på de 3 nævnte medier ([Altmetrics](#)). Dette er en begrænsning, da mere data fra flere sociale medier, vil have givet et mere retvisende resultat.

Der er simultanitet i denne analyse, da citationer i et hvis omfang også vil påvirke mængden af omtale på Twitter. Og ikke kun at omtale på Twitter vil påvirke citationer, effekten går begge veje. Simultanitet er også en stor begrænsning, instrumental variables estimation er den mest brugte metode til at kontrollere for dette på. Jeg vil komme nærmere ind på dette i mit afsnit omkring anbefaling af en alternativ fremgangsmåde. Ved at der er simultanitet gør det at mine variabler er endogene.

Den sidste begrænsning jeg er stødt på, er følgende: Det data jeg har er udelukkende på nuværende ansatte på SDU. Dette fjerner højst sandsynligt en hel del artikler, for forskere, der førhen har arbejdet på SDU. Grunden til dette er at det HR-data jeg har, er en opdateret liste, for at få det sammenkoblet med data fra

Altmetrics, har jeg været nødsaget til udelukkende at fokusere på de forskere jeg har i mit HR-data. Da jeg ellers ikke ville kunne få de personlige faktorer ind i mit samlede datasæt. Denne begrænsning gør at jeg får en mindre stikprøvestørrelse, hvilket gør jeg får en mindre "statistisk power".

## 1.7 Perspektivering

I dette afsnit vil jeg belyse, hvad der kan arbejdet videre med, for at undersøge dette emne, endnu dybere og hvad dette kan bruges til.

Som nævnt i mit afsnit omkring begrænsninger, ville en anden variable eller flere variabler for "socialt medie" have været mere optimalt. Jeg tror bestemt et andet socialt medie, så som LinkedIn, vil have større sammenhæng mellem citationer målt som klassisk akademisk bibliometri og mængden af omtale på sociale medier. Da LinkedIn er verdens største faglige netværk på internettet, det er en platform for personer, der ønsker at fremme deres karriere ([LinkedIn](#)). Hvor Twitter er et socialt netværk og mikro-blogging-værktøj, der giver brugerne mulighed for at sende og læse beskeder ([Twitter](#)).

For at få et andet overblik og tilgang til denne analyse, kunne der laves en tidslinje over hvor længe der går før en artikel får den første omtale på Twitter fra publikationsdatoen. Altså tidsserie data, vil have givet et måske bedre overblik, da jeg på den måde også kunne tage højde for tiden. Der kunne eventuelt laves en oversigt over hvor meget omtale hver enkelt artikel har efter 14 dage på Twitter. Dette kunne holdes op mod citationer, målt som klassisk bibliometri. Da det burde tage flere år fra en artikel er publiceret til den før sin første målte citation ([AU bibliotek](#), (Catharina Rehn, 2014)). Dette ville have været interessant at dykke ned i, for at se om omtale på Twitter forholder sig lige ledes, eller om der kommer omtale på Twitter (Tweets) allerede dage efter publikationsdatoen. Dette kunne meget nemt blive påvirket, af hvor aktiv de enkelte forskere er på de sociale medier, da hvis de selv smider deres publikation op på et socialt medie, vil der jo med det samme komme 1 omtale (tweet) kort efter publikationsdatoen.

De personer der omtaler en artikel, bruger de også hashtags eller ej? Da brugen af hashtags burde øge algoritmen i forhold til at artiklerne kan opnå mere opmærksomhed og dermed også mere omtale på Twitter. Der kunne ses på hashtagget #SDUDK og #SDU, for at se om de hashtags også fanger mængden af omtale på artiklerne og omvendt. Det kan være, det har været en fejlagtig fremgangs metode at kigge på hvor mange gange en artikel bliver omtalt, frem for hvor mange #SDUDK og #SDU i forbindelse med artiklerne bliver omtalt. Der kunne også være en lang række andre hashtags, der kunne undersøges.

En anden analyse der kunne laves, vil være en analyse på forsker niveau og sammenligne med analysen i denne opgave. Dette skulle gøres for at se om min teori, om at angivende andele for de personlige faktorer, ikke er helt ret visende, og derfor kan give misledende resultater. Dog skal der i den nye analyse tages højde

for at de artikler med flere forskere bag vil blive duplikeret med det antal forfattere (forskere) der er bag artiklen. Denne nye analyse kunne også udføres på en anden måde. Der kunne laves to forskellige analyser. En analyse, hvor det udelukkende er artikler, hvor der kun er én forsker bag, dog vil dette fjerne en stærk indikator for citationer ifølge (Tahamtan et al., 2016, Bornmann et al., 2012). Og en anden analyse hvor det udelukkende er artikler med flere forfattere. Den sidste analyse kan også deles op i to, en med andele og en med dupletter. Disse 3 nye forslag til at undersøge dette emne, vil give det fuldstændige overblik, over hvilken metode der er mest retvisende.

### 1.8 Anbefaling af en alternativ fremgangsmåde

Dette afsnit vil kort belyse, hvilke andre metoder der kunne være brugt, samt andre alternative fremgangsmåder.

Andre analyser der kunne have været interessante at kigge på i forhold til denne analyse kun blandt andet være Instrumental variables estimation (IV), dette vil hjælpe på omvendt kausalitet. I denne analyse vil det betyde at mængden af omtale på Twitter og citationer målt som klassisk bibliometri er associeret, men ikke kun som jeg forventede. Min forventning til denne analyse var at mængden af omtale på Twitter vil påvirke citationer. Men ud fra Appendix 6 kan det ses at citationer også påvirker mængden af omtale på Twitter. Det kan ses at citationer har en kausal effekt på Twitter, hvilket betyder at der er simultan effekt. Det kan også ses i mine resultater på figur 6, er der meget få artikler med høj mængde af omtale på Twitter og lave citationer. Altså ifølge mit data kan det ses at der muligvis er en sammenhæng, mellem at en artikel får meget omtale på Twitter også har en del citationer. For at køre en IV regression vil det kræve at jeg skulle bruge en variabel som påvirker mængden af omtale på Twitter (x) men ikke Citationer (y), hvor denne variable så skulle være mit instrument. Muligvis vil et hashtag kunne være et IV instrument. Da hashtag godt kan påvirke omtale på Sociale medier, men nødvendigvis ikke citationer, dog vil omtale påvirke citationer senere hen. Det vil bestemt være interessant at undersøge om det er muligt at skaffe en variabel der benytter hashtag, så der kunne tages højde for endogenitet bias. Dog har dette ikke været muligt for mig at få indsamlet data til denne hashtag variable.

Tidsserie data, vil være bedre end "øjeblikksdata" som jeg har brugt i denne analyse. Da mit data udelukkende viser, hvor mange citationer og omtale den enkelte artiklen havde da jeg trak mit data ned fra Altmetrics. Hvis jeg havde haft tidsserie data, vil jeg kunne have kigget på hvordan artiklerne havde udviklet sig over tid, om citationer kom øjeblikkeligt, samt hvor lang tid der går før en artikel får citationer og omtale på sociale medier. Jeg vil også kunne have haft informationer om artikler bliver ved med at få citationer og omtale på sociale medier, eller om det udelukkende er inden for de første år at en artikel bliver omtalt på sociale medier. Hvis der havde været en mulighed for at få tidsserie data, over citationer og mængden af omtale på sociale

medier, vil der kunne køres en Random Effekt model med Fixed effekt for år. Fixed effect for år, ville blive implementeret ved hjælp af en dummy variable for hvert år. Ved at køre en Random effekt med fikserede år, ville jeg have taget højde for at jeg har mange artikler, men kun få observationer per artikel. Af samme grund vil en ren fixed effect være ineffektiv. Dog kunne Random Effekt for artikler være lidt problematisk, grundet nogle artikler kun er med én gang. Ved at køre Random Effekt fjerner jeg noget bias, dette kommer fra u observeret heterogenitet. Denne bias kommer fra omitted variables bias, da u observeret heterogenitet er variationen mellem de artikler der ikke er målt, ved ikke at tage højde for dette, kommer mine resultater til at blive misvisende.

### 1.9 Diskussion

I dette afsnit vil jeg kort holde mine egen resultater op mod tidligere litteratur.

Jeg inddelte denne opgave med at skrive, at der ikke er lavet meget forskning inden for dette emne. Dog vil jeg sammenligne de få tidligere undersøgelsers resultater jeg har med i afsnittet baggrund, med mine egne resultater. Eysenbach (2011) finder at omtale på Twitter og citationer er "noget" korreleret, det er de til en vis grad også i mine resultater. Dog er mængden af omtale på Twitter og citationer svagt positivt korreleret i mit data.

Tahamtan et al. (2016), kigger på hvilke faktorer, der giver citationer. Han finder at faktorer som kvalitet af artiklen, faktorer for journalens indvirkning og antal af forskere er stærkere indikatorer end forskerens køn, alder og race. Mine resultater er ikke helt sammenfaldende med Tahamtan et al. (2016). Med dette skal der tages højde for at Tahamtan et al. (2016) udelukkende undersøger, hvilke faktorer som påvirker hyppigheden af citationer, og ikke har social medie påvirkning med, hvilket jeg har. Derfor giver der god mening at vores resultater ikke er helt sammenfaldende. Hvis der kigges på resultaterne i denne analyse, finder jeg at ud af de 4 stærke faktorer er det journalens indvirkning der har langt den største påvirkning på hvor meget 1 omtalte på Twitter påvirker citationer, denne påvirkning er positiv. Hvor forskerens køn har den næst største påvirkning, køn har en negativ påvirkning. Hvor efter de sidste 2 faktorer kommer i følgende rækkefølge: Alder (år siden første artikel), der har en positiv påvirkning og til sidst antallet af forskere, der har en negativ påvirkning.

### 1.10 Konklusion

Denne analyse havde til formål at undersøge om der er sammenhæng mellem mængden af omtale på sociale medier (Twitter) og citationer målt som klassisk akademisk bibliometri. Jo flere variabler jeg kontrollerer for, jo lavere bliver påvirkningen af omtale på Twitter. Dog ender 1 "omtale på Twitter" med at påvirke citationer med 0.36. Altså vil en artikel der er blevet nævnt 10 gange på Twitter have 3,6 flere citationer. Den

sandeværdi vises mere og mere jo flere variabler der kontrolleres for. Mine resultater er lider af meget bias, derfor kan de ikke bruges til meget. Med mere tid, mere og bedre data, ville det måske være muligt at skabe bedre resultater. Resultater der kan bruges af forskerne og ledelsen på Syddansk Universitet.



## 2 Bibliometri

ALTMETRIC A beginner's guide to altmetrics. youtube.com.

ALTMETRICS. *What are Altmetrics?* [Online]. [www.altmetric.com](http://www.altmetric.com). Available: [https://www.altmetric.com/about-altmetrics/what-are-altmetrics/?fbclid=IwAR3aW39iXI7-BL4gZ3pYkAH3hwUvUypQnE\\_LrBppbANHaMBCWcyVNphbJ\\_4#prettyPhoto](https://www.altmetric.com/about-altmetrics/what-are-altmetrics/?fbclid=IwAR3aW39iXI7-BL4gZ3pYkAH3hwUvUypQnE_LrBppbANHaMBCWcyVNphbJ_4#prettyPhoto) [Accessed 11-01-2021 2021].

ALTMETRICS. *What outputs and sources does Altmetric track?* [Online]. help.altmetrics.com: Altmetrics. Available: <https://help.altmetric.com/support/solutions/articles/6000060968-what-outputs-and-sources-does-altmetric-track-> [Accessed 15-01-2021 2021].

AU-LIBRARY. *Citations* [Online]. AU-Library. Available: <https://library.au.dk/en/researchers/bibliometrics/citations/> [Accessed 12-01-2021 2021].

BARNETT, G. A. & FINK, E. L. 2008. Impact of the internet and scholar age distribution on academic citation age. *Journal of the American Society for Information Science and Technology*, 59, 526-534.

BORNMANN, L., SCHIER, H., MARX, W. & DANIEL, H.-D. 2012. What factors determine citation counts of publications in chemistry besides their quality? *Journal of Informetrics*, 6, 11-18.

CATHARINA REHN, D. W., CARL GORNITZIK & AGNE LARSON 2014. Bibliometric Indicators - Definitions And Usage At Karaolinska Institutet *Karolinska Institutet*.

EYSENBACH, G. 2011. Can Tweets Predict Citations? Metrics of Social Impact Based on Twitter and Correlation with Traditional Metrics of Scientific Impact. *Journal of Medical Internet Research*, 13, e123.

GARFIELD, E. 2006. The History and Meaning of the Journal Impact Factor. *JAMA*, 295, 90.

LINKEDIN. *Hvad er LinkedIn, og hvordan kan jeg bruge det?* [Online]. LinkedIn. Available: <https://www.linkedin.com/help/linkedin/answer/111678/hvad-er-linkedin-og-hvordan-kan-jeg-bruge-det-?lang=da> [Accessed 15-01-2021 2021].

SDU.DK/ANALYTICS. *SDU Analytics* [Online]. sdu.dk. Available: [https://www.sdu.dk/da/om\\_sdu/faellesomraadet/sdu\\_analytics](https://www.sdu.dk/da/om_sdu/faellesomraadet/sdu_analytics) [Accessed 08-01-2021 2021].

TAHAMTAN, I., SAFIPOUR AFSHAR, A. & AHAMDZADEH, K. 2016. Factors affecting number of citations: a comprehensive review of the literature. *Scientometrics*, 107, 1195-1225.

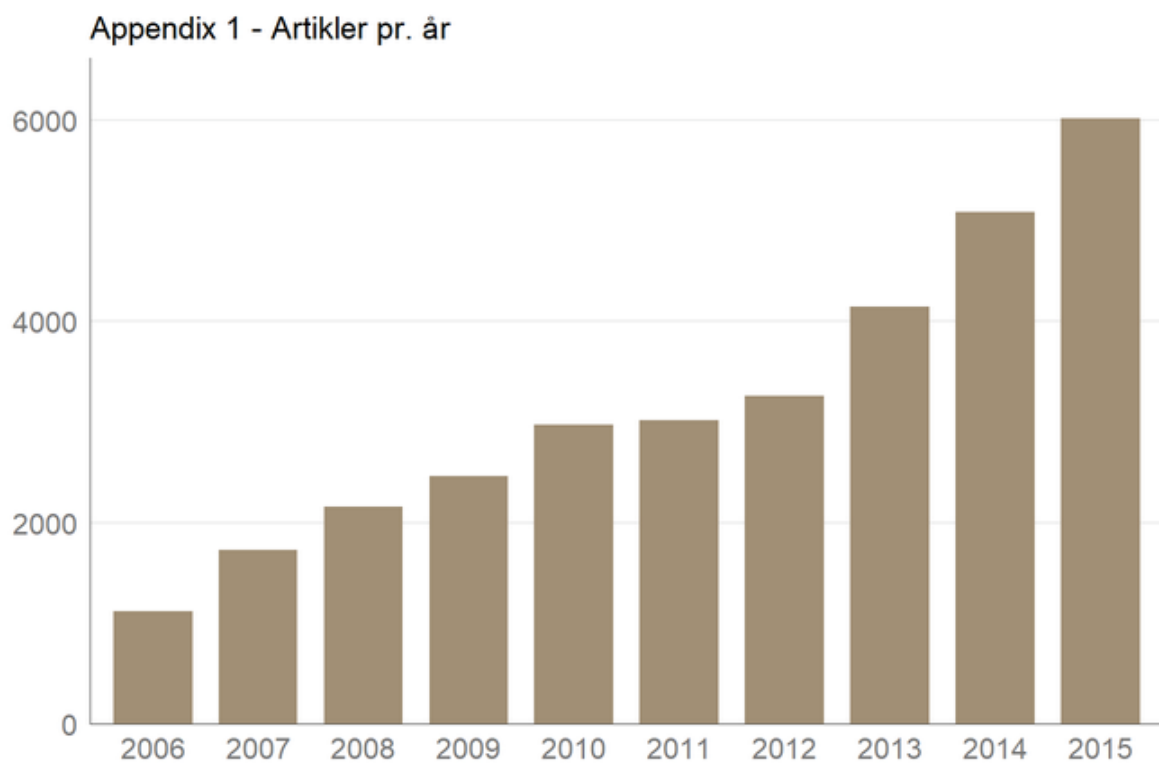
TWITTER. *Twitter is* [Online]. twitter.com. Available: about.twitter.com [Accessed 15-01-2021 2021].

TWITTER. *Twitter turns six* [Online]. twitter.com. Available: [https://blog.twitter.com/official/en\\_us/a/2012/twitter-turns-six.html](https://blog.twitter.com/official/en_us/a/2012/twitter-turns-six.html) [Accessed 11-01-2021 2021].

### 3 Appendix oversigt

#### 3.1 Appendix 1 - Summen af Citationer for artikler grupperet pr år

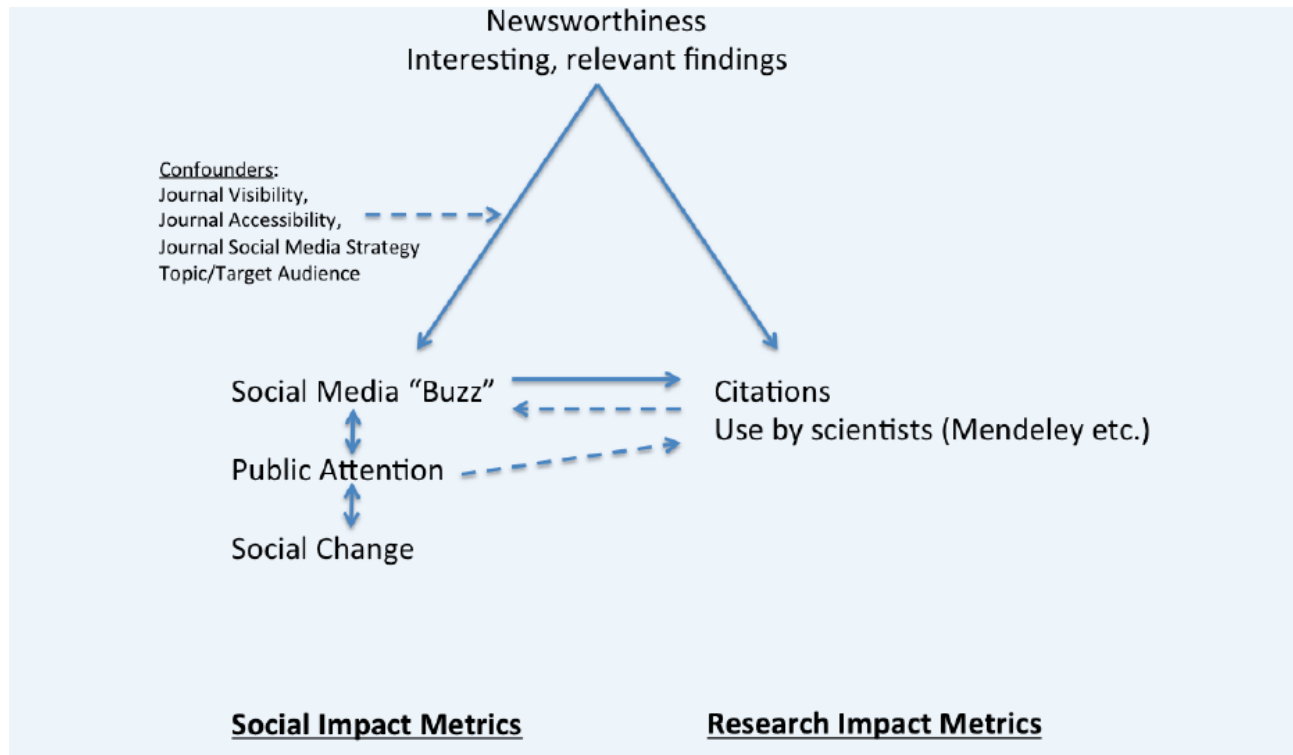
I dette appendix kan der ses en graf over antal af udgivende artikler på SDU fra årene 2006 til 2015. Det er tydeligt at se at disse er stigende, hen over årene.



Appendix 1,  
Kilde: Lavet af Maria Færge Dueholm med data fra Altmetrics

### 3.2 Appendix 2 - Model af forholdet mellem sociale effekter og forsknings effekter

Dette appendix forklarer hvordan (Eysenbach, 2011) definerer at mængde af omtale på Twitter og citationer er "noget" korreleret.



Appendix 2 - Model af forholdet mellem sociale effekter og forskning effekter. Kilde: (Eysenbach, 2011)

Som det kan ses på figuren oven over, er korrelationen og gensidig interaktion mellem disse to målgrupper med tovejspile, der peger fra "social Media Buzz" til "Citations" og fra "use by scientists" (forskere der er påvirket af sociale medier) til "Social media buzz" (Forskere der laver buzz on Twitter"), der illustrer den gensidige indflydelse af disse målgrupper og målinger. (Eysenbach, 2011)

## 3.3 Appendix 3 - Koefficient estimer for alle variablerne X repræsenterer

Dette appendix viser alle koefficientestimerne for alle variablerne X repræsenterer i tabel 1

	Appendix 3: Afhængig variable			
	Citationer			
	(1)	(2)	(3)	(4)
Mængde af omtale på Twitter	49.964*** (1.989)	-156.890* (87.242)	-165.500* (87.061)	-174.605** (87.105)
tweet	0.403*** (0.035)	0.374*** (0.033)	0.373*** (0.033)	0.360*** (0.033)
age_of_paper		12.801*** (0.883)	14.089*** (0.924)	14.329*** (0.935)
institutBiologi		52.134 (87.347)	57.708 (87.154)	54.214 (86.604)
institutBiomedicinsk Laboratorium		24.888 (122.882)	26.820 (122.599)	52.479 (121.644)
institutCenter for Industriel Elektronik (CIE)		32.854 (122.888)	35.432 (122.605)	45.928 (121.712)
institutCenter for Maritim Sundhed og Samfund (CMSS)		5.706 (95.197)	31.107 (95.127)	35.808 (95.138)
institutCentre for Quantum Mathematics		-0.674 (112.172)	-1.487 (111.913)	5.483 (111.140)
institutCPOP		59.158 (89.565)	58.210 (89.359)	41.868 (89.026)
institutInst. for Sundhedstjenesteforskning (IST)		29.100 (87.069)	30.872 (86.869)	35.723 (86.261)
institutInstitut for Biokemi og Molekylær Biologi		51.814 (87.315)	53.214 (87.116)	50.959 (86.528)
institutInstitut for Design og Kommunikation		3.132 (87.915)	16.750 (87.775)	33.415 (87.104)
institutInstitut for Entreprenørskab og Relationsledelse		12.238 (90.295)	30.791 (90.170)	55.067 (89.512)
institutInstitut for Fysik, Kemi og Farmaci		15.531 (87.361)	19.800 (87.164)	22.593 (86.590)
institutInstitut for Historie		-26.543 (87.588)	0.159 (87.567)	14.988 (86.965)
institutInstitut for Idræt og Biomekanik		41.183 (87.125)	41.524 (86.924)	49.472 (86.294)
institutInstitut for Kulturvidenskaber		0.070 (88.322)	22.831 (88.248)	36.081 (87.562)
institutInstitut for Marketing	Management	-4.860 (88.302)	12.942 (88.178)	12.942 (87.592)
institutInstitut for Molekylær Medicin		75.332 (87.679)	75.456 (87.477)	76.023 (86.852)
institutInstitut for Psykologi		45.412 (87.831)	48.584 (87.631)	57.354 (86.965)
institutInstitut for Regional Sundhedsforskning		52.163 (87.113)	54.245 (86.914)	66.153 (86.248)
institutInstitut for Sociologi, Miljø- og Erhvervsøkonomi		31.893 (89.408)	39.668 (89.216)	54.536 (88.532)
institutInstitut for Sprog og Kommunikation		32.249 (88.103)	61.877 (88.119)	79.335 (87.595)
institutInstitut for Statskundskab		21.520 (87.630)	37.687 (87.493)	58.460 (86.929)
institutInstitut for Trombose, Esbjerg		459.969*** (96.084)	458.567*** (95.863)	459.669*** (95.091)
institutInstitut for Virksomhedsledelse og Økonomi		3.185	9.443	26.928

institutKlinisk Institut	56.272	57.866	68.573
	(86.953)	(86.753)	(86.122)
institutMatematik og Datalogi	25.245	38.893	52.505
	(87.539)	(87.445)	(86.910)
institutRetsmedicinsk Institut	14.352	15.541	41.864
	(93.863)	(93.647)	(93.006)
institutSamfundsvidenskab fælles	-8.286	12.978	28.967
	(94.464)	(94.352)	(94.187)
institutSDU Biorobotics	-13.667	2.245	13.817
	(88.988)	(88.943)	(88.404)
institutSDU Biotechnology	10.579	11.278	14.549
	(95.182)	(94.963)	(94.271)
institutSDU Centre for Photonics Engineering	242.421"	241.912"	264.721"
	(122.888)	(122.605)	(121.715)
institutSDU Centre for Sustainable Supply Chain Engineering	212.703"	213.720"	224.734"
	(94.458)	(94.241)	(93.826)
institutSDU Chemical Engineering	-5.340	12.039	19.640
	(90.166)	(90.032)	(89.328)
institutSDU Electrical Engineering	-10.004	-8.581	-17.614
	(150.504)	(150.157)	(148.987)
institutSDU Energy Informatics	18.823	32.851	43.345
	(102.812)	(102.617)	(101.881)
institutSDU Engineering Operations Management	371.186""	373.287""	389.349""
	(112.180)	(111.922)	(111.103)
institutSDU Game Development and Learning Technology	13.807	16.402	41.583
	(98.536)	(98.310)	(97.523)
institutSDU Health Informatics and Technology	-8.952	2.521	22.622
	(91.127)	(90.949)	(90.350)
institutSDU Life Cycle Engineering	97.385	98.008	122.635
	(102.817)	(102.579)	(101.926)
institutSDU Mechanical Engineering	12.797	15.508	41.123
	(150.512)	(150.166)	(149.133)
institutSDU Mechatronics	82.426	81.920	88.815
	(122.888)	(122.604)	(121.661)
institutSDU Nano Optics	167.292`	178.101"	166.900`
	(90.587)	(90.406)	(90.143)
institutSDU NanoSyd	-17.273	-20.195	-16.851
	(93.869)	(93.655)	(93.091)
institutSDU Robotics	90.374	91.800	83.260
	(122.891)	(122.608)	(121.774)
institutSDU Software Engineering	-41.932	-30.340	-27.266
	(92.532)	(92.351)	(91.708)
institutSDU UAS Centre	24.652	23.872	38.774
	(122.889)	(122.605)	(121.654)
institutSDU universitetspædagogik	-13.138	-13.004	10.345
	(112.183)	(111.924)	(110.993)
institutStatens Institut for Folkesundhed	42.294	42.180	45.336
	(87.442)	(87.240)	(86.559)
institutSUND Fakultetssekretariat	27.636	29.313	32.466
	(112.176)	(111.918)	(111.026)
niveau_of_journal	51.358""	51.087""	47.109""
	(3.531)	(3.523)	(3.524)
output_typebook		-32.698""	-23.520""
		(6.866)	(6.956)
output_typechapter		-4.218	1.891
		(17.462)	(17.355)

authors_per_article				-0.321 (0.283)
sex				-15.160*** (5.049)
nationalitet				9.456* (5.442)
years_scince_first_article				0.338 (0.330)
number_of_articels				0.053** (0.027)
karriere_stige				-2.207 (3.173)
funder				33.458*** (4.285)
Artikel Relaterede Faktorer	X		X	X
Journal Relaterede Faktorer			X	X
Person Related Factors				X
Observations	4,545	4,545	4,545	4,545
R <sup>2</sup>	0.029	0.167	0.172	0.187
Note: <span style="float: right;"><math>p &lt; 0.1</math>; <math>p &lt; 0.05</math>; <math>p &lt; 0.01</math></span>				

Appendix 3

Kilde: Lavet af Maria Færge Dueholm, med data fra. Altmetrics, HR-data fra SDU samt data fra Uddannelses- og Forskningsministeriet

## 3.4 Appendix 4 - Subgroups: Campus beliggenhed

Dette Appendix viser resultaterne for min subgroup Lokation. Der skal tages højde for at der i regressionen med København, er blevet fjernet 2 variabler, institut og output type. Dette er gjort, da de hen holdvis udelukkende indeholdte, Statens institut for folkesundhed og artikel som output type. Når der ingen variation er i en variabel kan der ikke køres en regression. Det er tydeligt at se, at der er forskel på, hvor artiklerne er udgivet, Esbjerg er den campus hvor omtale på Twitter påvirker citationer højest.

## Subgroup: Lokation

	Dependent variable:					
	Citationer					
	Odense (1)	København (2)	Kolding (3)	Slagelse (4)	Esbjerg (5)	Sønderborg (6)
Mængde af omtale på Twitter	0.35*** (0.03)	0.67** (0.26)	-1.96* (1.08)	-0.12*** (0.00)	4.06** (1.93)	0.20 (0.41)
Observations	4,155	156	90	9	79	56
R <sup>2</sup>	0.16	0.30	0.38	1.00	0.70	0.96

Note:

 $p < 0.1$ ;  $p < 0.05$ ;  $p < 0.01$ 

Appendix 4

Kilde: Lavet af Maria Færge Dueholm, med data fra Altmetrics, HR-data fra SDU samt data fra Uddannelses- og Forskningsministeriet

## 3.5 Appendix 5 - Subgroups: BFI

Dette Appendix viser resultaterne for min subgroup Niveau af journal. Der skal tages højde for at der i regressionen med BFI3, er blevet fjernet 1 variabel, output type. Dette er gjort, da den udelukkende indeholdte, Artikler som output type Når der ingen variation er i en variabel kan der ikke køres en regression. Det er tydeligt at se, at der er forskel på, de tre niveauer. BFI er det niveau hvor omtale på Twitter påvirker citationer højest.

Subgroup: Niveau af Journal

	Dependent variable:		
	BFI 1 (1)	Citationer BFI 2 (2)	BFI 3 (3)
Mængde af omtale på Twitter	0.35*** (0.05)	0.26*** (0.03)	2.31** (1.02)
Observations	2,428	2,032	85
R <sup>2</sup>	0.16	0.16	0.43

Note:

 $p < 0.1$ ;  $p < 0.05$ ;  $p < 0.01$ 

Appendix 5

Kilde: Lavet af Maria Færge Dueholm, med data fra. Altmetrics, HR-data fra SDU samt data fra Uddannelses- og  
Forskningsministeriet



## 3.6 Appendix 6 - Test for simultanitet

	Afhængig variable: Citationer	Afhængig variable: Omtale på Twitter
(Intercept)	49.96 *** (1.99)	6.27 *** (0.88)
tweet	0.40 *** (0.03)	
citations		0.07 *** (0.01)
R <sup>2</sup>	0.03	0.03
Adj. R <sup>2</sup>	0.03	0.03
Num. obs.	4545	4545
*** p < 0.001; ** p < 0.01; * p < 0.05		

Kilde: Lavet af Maria Færge Dueholm, med data fra Altmetrics, HR-data fra SDU samt data fra Uddannelses- og forskningsministeriet.

Dette appendix viser i Kolonne 1 (Afhængig variable: Citationer) den overordnede effekt af mængden af omtale på Twitter på citationer og i kolonne 2 (Afhængig variable: Omtale på Twitter) den overordnede effekt af citationer på mængden af omtale på Twitter. Det kan ses at effekten i kolonne 1 er større end effekten i kolonne 2. Alle estimaterne er signifikante.