

Jaaroverzicht FD-sentiments- indicator: 2021 in de achteruitkijkspiegel

December 2021

Jasper de Winter en Dorinth van Dijk

DeNederlandscheBank

EUROSYSTEM

Jaaroverzicht FD-sentimentsindicator: 2021 in de achteruitkijkspiegel
©2021 De Nederlandsche Bank n.v.

Auteurs: Jasper de Winter en Dorinth van Dijk. Met dank aan Tjerk Kroes, Olaf Sleijpen en Peter van Els voor hun commentaar. Alle overgebleven fouten zijn de onze.

Met de serie 'DNB Analyse' beoogt De Nederlandsche Bank inzicht te verschaffen in de analyses die DNB ten behoeve van actuele beleidsvraagstukken uitvoert. De tot uitdrukking gebrachte zienswijzen zijn voor rekening van de auteurs en komen niet noodzakelijkerwijs overeen met de officiële standpunten van De Nederlandsche Bank. Niets uit deze uitgave mag worden verveelvoudigd en/of openbaar gemaakt door middel van druk, fotokopie, microfilm of op welke andere wijze ook en evenmin in een retrieval system opgeslagen worden, zonder voorafgaande schriftelijke toestemming van De Nederlandsche Bank.

De Nederlandsche Bank n.v.
Postbus 98 1000 AB
Amsterdam
Internet: www.dnb.nl
Email: info@dnb.nl

Inhoudsopgave

Inhoudsopgave	3
Kernpunten	4
1 Inleiding	5
2 Sentimentsindicator herstelt in 2021	6
3 Sentiment sterk gestuurd door Corona	9
4 Financiële markten doorstaan coronastorm het best	12
5 Corona dominante thema binnen andere deelindicatoren	14
Geraadpleegde literatuur	15
Appendix	16

Kernpunten

- Deze analyse dient als verantwoording van methode en techniek bij een eindejaarsartikel in het Financieele Dagblad (verder, FD) en een DNBulletin dat over de FD-indicator op de DNB-website is verschenen. Beide stukken analyseren het beloop van de sentimentsindicator in 2021;
- Het krantensentiment is volgens de sentimentsindicator na de forse daling bij aanvang van de coronacrisis voor een belangrijk deel hersteld, maar de zorgen over de verspreiding van het virus blijven het sentiment drukken;
- De deelindicator die het sentiment op financiële markten weergeeft is eind 2021 zelfs positiever dan bij aanvang van de coronacrisis. Dit komt met name door het optimisme op de beurs. Artikelen over cryptomunten, meme-aandelen en de energietransitie en de coronapandemie droegen daarentegen relatief negatief bij aan het financieel marktsentiment in 2021;
- In de rest van de krant (bedrijven, economie, politiek) voert corona de boventoon. Het bedrijven-, economisch- en politiek sentiment is, hoewel verbeterd, nog niet terug op het niveau van voor de pandemie.

1 Inleiding

Op 26 maart 2021 is de door DNB op basis van artikelen in het Financieele Dagblad ontwikkelde FD-sentimentsindicator geïntroduceerd (Financieele Dagblad, 2021 en DNB, 2021). In een artikel in het vaktijdschrift Economische Statistische Berichten wordt de indicator in meer detail beschreven (de Winter en van Dijk, 2021). Sindsdien maakt de indicator een vast onderdeel uit van de variabelen die zijn opgenomen in het model dat DNB gebruikt om de groei van het bruto binnenlands product (verder, bbp) in het lopende en komende kwartaal te voorspellen. Zie voor meer informatie over dit zogenaamde nowcasting model, de DNB-website ([hier](#)).

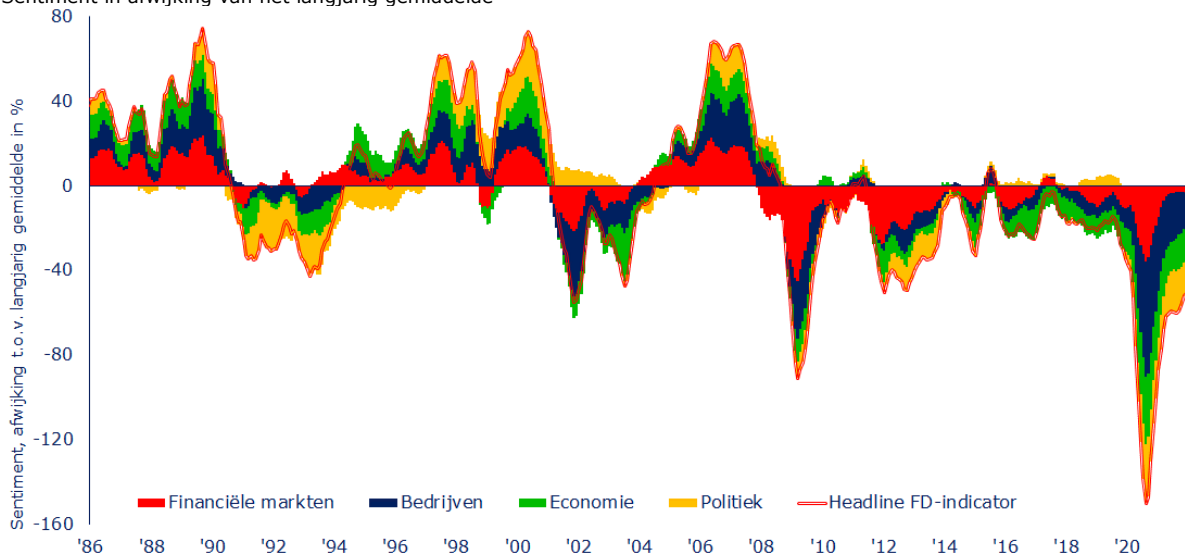
In deze analyse geven we een overzicht van het beloop van de sentimentsindicator in 2020 en 2021, en schetsen we welke de belangrijkste drijvende krachten achter deze ontwikkeling waren. De analyse dient als achtergrond bij een overzichtsartikel dat het FD onlangs schreef over de indicator en als onderbouwing voor uitingen over de indicator op de DNB-website. De opbouw van de analyse is als volgt: in hoofdstuk 2 beschrijven we de ontwikkeling van de FD-indicator sinds januari 2020. In hoofdstuk 3 gaan we in op de belangrijkste woorden onderliggend aan de sentimentsdynamiek. In hoofdstuk 4 zoomen we in op het deelonderwerp financiële markten. In het laatste hoofdstuk gaan we kort in op de ontwikkeling van de andere drie deelindicatoren van de FD-sentimentsindicator.

2 Sentimentsindicator herstelt in 2021

De sentimentsindicator is geconstrueerd op basis van ongeveer 1 miljoen krantenartikelen uit het Financieele Dagblad (uit de periode 1985 – 2021), en meet de stand van de conjunctuur. Figuur 2.1 geeft de ontwikkeling weer van de headline sentimentsindicator (rode lijn) en de vier deelindicatoren die de indicator vormen, te weten: financiële markten (rode staafjes), bedrijven (blauwe staafjes), economie (groene staafjes) en politiek (gele staafjes). De coronacrisis heeft over een breed front geleid tot een historisch sterke daling van de sentimentsindicator. De daling was circa twee maal zo groot als ten tijde van de financiële crisis. In de loop van 2020 verbeterde het vertrouwen, maar in de voorjaar van `2021 lijkt het vertrouwensherstel te stokken. In Box 1 wordt een beschrijving gegeven over de techniek achter de FD-sentimentsindicator.

Figuur 2.1 FD-sentimentsindicator naar deelindicator

Sentiment in afwijking van het langjarig gemiddelde



Bron: FD-database, berekeningen DNB.

Figuur 2 zoomt in op de ontwikkeling van de indicator in de afgelopen twee jaar. In de figuur is duidelijk te zien dat de verspreiding van het Covid-19-virus de FD-sentimentsindicator over een breed front deed kelderen. Het sentiment daalde met maar liefst 84 procent tussen januari 2020 en het dieptepunt in augustus 2020. Dit viel samen met de publicatie van de eerste inschatting van de initiële schok op het bbp als gevolg van de verspreiding van het Covid-19-virus, die op 15 augustus verscheen. Het bruto binnenlands product (bbp) kromp volgens die raming in het tweede kwartaal van 2020 met 8,5% ten opzichte van het eerste kwartaal van 2020, de grootste krimp van de economie ooit door het CBS gemeten. De krimp werd onder meer ingegeven door de sluiting van de horeca, de cultuur-evenementensector, scholen en kinderopvang en het verbod op het uitvoeren van contactberoepen. Daarnaast leidde het dringende advies thuis te werken tot productieverlies bij bedrijven en organisaties waarbij telewerken lastig was.

Box 1: Constructie sentimentsindicator in vogelvlucht¹

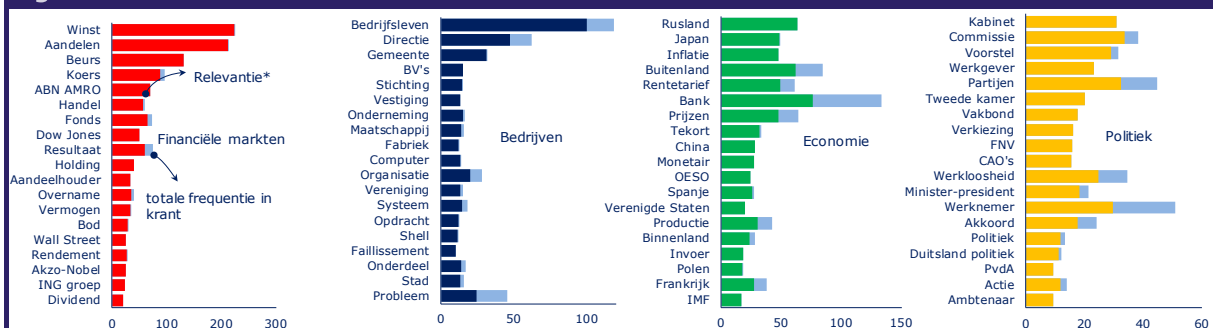
De vier deelindicatoren/hoofdonderwerpen worden uit de teksten in het Financieele Dagblad geëxtraheerd met gebruikmaking van een zogeheten topicmodel (Blei et al., 2003). Het idee van een topicmodel is dat een artikel uit een of meerdere onderwerpen bestaat, en dat het woordgebruik per onderwerp verschilt.

Met Bayesiaanse schattingstechnieken is vervolgens te bepalen welke verdeling qua onderwerp en artikel, en welke verdeling qua onderwerp en woord het meest waarschijnlijk zijn. Het enige dat nodig is om het algoritme op te starten zijn de aannames over het maximaal aantal onderwerpen in een artikel (in ons geval vier), de waarschijnlijkheid dat een artikel één, twee, drie of vier onderwerpen behelst en de kansverdeling van woorden over de vier door ons onderscheiden onderwerpen. We volgen de bestaande literatuur, en gaan ervan uit dat de kans klein is dat een artikel uit heel veel onderwerpen bestaat en dat de diverse woorden redelijk uniek zijn voor een onderwerp.

Voordat het algoritme de database in onderwerpen kan verdelen, moet de database worden opgeschoond. Het doel van deze stap is om alleen artikelen over te houden die betrekking hebben op relevant nieuws. Zo verwijderen we onder meer lifestyle-artikelen en Engelstalige artikelen. Daarnaast is het noodzakelijk om het aantal woorden terug te brengen tot de woorden die de artikelen onderling onderscheiden. Daartoe worden semantisch lege ('betekenisloze') woorden ('de', 'het', 'een', 'om' en 'dergelijke'), leestekens, getallen en HTML-codes verwijderd. Tevens is het noodzakelijk om werkwoorden terug te brengen tot hun stam. Zo worden werkte, gewerkt en werkten gereduceerd tot werk.

We extraheren voor deze DNB Analyse vier hoofdonderwerpen die de grote rubrieken in de krant goed blijken weer te geven, te weten: 'financiële markten', 'bedrijfsleven', 'economie' en 'politiek'. Het model labelt de geëxtraheerde hoofd- en deelonderwerpen niet. Dat is aan de onderzoeker. Om een indruk te geven waarom deze labels zijn gekozen, is in Figuur 2.2 weergegeven wat de relevantie is van de woorden in de vier hoofd-onderwerpen. Zo zien we dat binnen het onderwerp 'politiek' de woorden minister, kabinet en wetgeving van grote relevantie zijn. De woorden 'kabinet' en 'wetgeving' komen vrijwel alleen voor bij het onderwerp 'politiek'. Dat is te zien aan het feit dat de totale frequentie van deze woorden (de lichtblauwe balkjes) vrijwel gelijk is aan de relevantie (de gele balkjes), want de lichtblauwe en gele balkjes zijn even lang. Het uitgebreidere door DNB ontwikkelde topicmodel wordt n drie lagen geschat. Hierbij worden de hoofdonderwerpen verder verbijzonderd naar uiteindelijk 64 deelonderwerpen (zie de Winter en van Dijk, 2021 voor meer details).

Figuur 2.2 FD-sentimentsindicator naar deelindicator



Bron: FD-database, berekeningen DNB.

Toelichting: Relevantie is een maatstaf voor het belang en uniciteit van een woord binnen een onderwerp (zie Steyvers en Griffiths, 2007).

Een innovatie ten opzichte van andere onderzoeken is dat we het topicmodel tijdsvariërend schatten, dus op basis van het woordgebruik en onderwerpen die op dat moment worden gebruikt.

(wordt vervolgd op volgende pagina)

¹ Deze box is grotendeels gebaseerd op de Winter en van Dijk (2021).

Vervolg Box 1: Constructie FD-sentimentsindicator in vogelvlucht

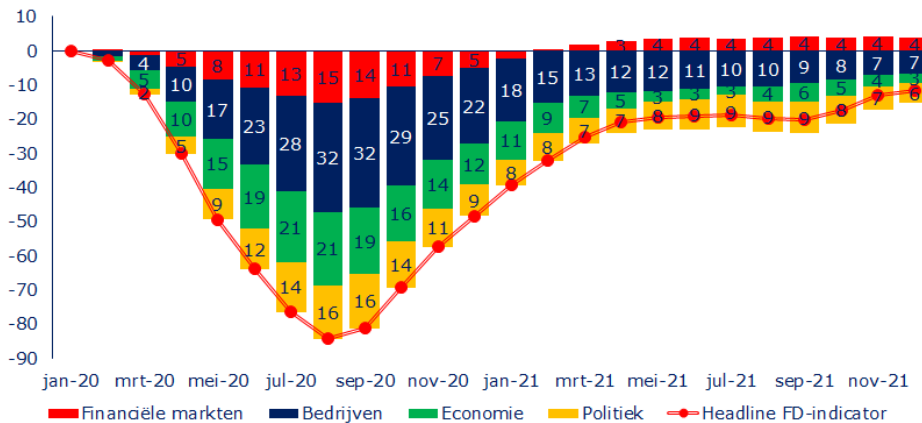
Ter illustratie: deze tijdsvariatie geeft het model de ruimte om aan het woord 'Griekenland' een relatief laag gewicht toe te kennen in de jaren tachtig en negentig, maar juist een groot gewicht tijdens de financiële crisis en de Europese schuldencrisis. Indien hier geen rekening mee wordt gehouden en het model over de hele periode wordt geschat, wordt de relevantie van het woord 'Griekenland' te laag ingeschat in de laatste periode en te hoog in de eerste periode. Dit kan ertoe leiden dat belangrijke ontwikkelingen, zoals de schuldencrisis, over het hoofd gezien worden. Het uiteindelijke resultaat is een kansverdeling over 4 onderwerpen per artikel die het aandeel van elk van deze onderwerpen in het FD weergeven.

Vervolgens meten we het sentiment per artikel door het aantal positieve en negatieve woorden te tellen. De basis van de lijst die de sentimentsscore per woord berekent is het veelgebruikte Engelstalige sentiment lexicon van Loughran en McDonald (2011). Deze op financiële berichtgeving toegespitste lijst vertalen we naar het Nederlands. We kiezen de vertaling die het best past binnen het economische domein. We vullen de lijst aan met stemmingswoorden die specifiek zijn voor de Nederlandse taal en onze toepassing. Zo dwingen we af dat woordcombinaties als 'krimp' afgenomen en 'werkloosheid gedaald' ondanks de negatieve gevoelswaarde van de afzonderlijke woorden toch worden meegeteld als een positief sentiment. De stemming van een artikel is de som van het aantal woorden met een positief sentiment (+1) en negatief sentiment (-1). We normaliseren het sentiment per artikel door de som van positief en negatief sentiment te delen door het totaal aantal woorden in het artikel. Het sentiment per maand wordt berekend door het gemiddelde te nemen van de stemming in alle artikelen van die maand. Het berekende sentiment per maand is erg volatiel. Om een signaal uit de data te extraheren, berekenen we een simpel zesmaands voortschrijdend gemiddelde.

Door de meting van sentiment per artikel en het topicmodel te combineren, kan de stemming worden toegekend aan de 4 onderwerpen die ons model onderscheidt. Dat zit zo: elk artikel heeft een sentimentsscore en door het topicmodel weten we van elk artikel uit welke onderwerpen het bestaat. We delen het gemeten sentiment per artikel toe aan de onderwerpen in dat artikel conform het aandeel van de onderwerpen in het artikel. Stel het gemeten sentiment is 10. Volgens het topicmodel gaat het voor negentig procent over economie en voor tien procent over politiek, dan wordt een 9 toegekend aan economie, en een 1 aan politiek. Zo kunnen we de stemming in o.a. Figuur 2.1 onderverdelen in de headline FD-indicator en vier onderwerpen of deelindicatoren, te weten: 'financiële markten', 'bedrijfsleven', 'economie' en 'politiek'.

Figuur 2.3 FD-sentimentsindicator naar deelindicator

Sentiment in afwijking van het sentiment in januari 2021 in procenten (headline) en procentpunten (deelonderwerpen)



Bron: FD-database, berekeningen DNB.

De meeste conjunctuurnvorschers gingen in augustus 2020 ervanuit dat de coronacrisis kort zou duren en de economische activiteit in het derde kwartaal van 2020 weer zou aantrekken. Dit komt ook tot uiting in

sentiment.³ Figuur 3.1B geeft de *trending* woorden binnen de artikelen die het meest bij hebben gedragen aan de sentimentsdynamiek. Trending woorden zijn woorden die relatief vaak voorkomen in 2021 én weinig in de vier voorafgaande jaren (2017–2020). De grootte geeft wederom de frequentie aan waarin het woord voorkwam; de kleur het gemiddelde sentiment. Naast de situatie rondom corona – met trending woorden als ‘pandemie’, ‘lockdown’, ‘vaccin’, en ‘RIVM’ – was de kabinetsformatie ook trending. Dit komt in Figuur 3.1B naar voren met woorden als ‘kabinet’ en ‘Rutte’. Uit woorden als ‘klimaat’, ‘groen’ en ‘elektrisch’ blijkt dat duurzaamheid ook een hot topic was. De 10 woorden met de hoogste trendscore voor 2021 waren: ‘demissionair’, ‘vaccin’, ‘prikken’, ‘gevaccineerd’, ‘avondklok’, ‘Evergrande’, ‘Joe Biden’, ‘Afghanistan’, ‘Astrazeneca’ en ‘variant’. Dit onderstreept de grote aandacht voor de vaccinatie campagne, de lockdown en de opkomst van corona-varianten. De terugtrekking van Amerikaanse troepen uit Afghanistan komt ook duidelijk naar voren als een van de belangrijke nieuwsitems in 2021.

In Box 2 wordt een uitgebreidere toelichting gegeven op de meting en weging van de sentiment-, frequentie-, uniciteits- en trendscore van een woord. In de figuren in deze nota kennen we een vier keer zo groot gewicht toe aan ‘uniciteit’ als aan de ‘frequentie’ van een woord. In Figuur A.1 in de Appendix zijn de wordclouds weergegeven indien we uniciteit respectievelijk gelijk en twee keer zo zwaar laten meewegen als de frequentie. Het algemene beeld is vergelijkbaar, maar het is lastiger te onderscheiden welke woorden in 2021 trending waren.

Box 2: Frequentie, uniciteit, trend en sentiment van een woord

Elk artikel in de FD-indicator bestaat voor een bepaald percentage uit de onderwerpen ‘financiële markten’, ‘bedrijfsleven’, ‘economie’ en ‘politiek’ (zie ook Box 1). Om de wordclouds te kunnen maken selecteren we binnen de vier hoofdonderwerpen die de indicator onderscheidt, de 5000 artikelen die het meest bijdroegen aan het sentiment van de hoofdonderwerpen. Binnen de 5000 artikelen voor de vier hoofdonderwerpen berekenen we vervolgens voor elk woord –uitgezonderd stopwoorden, de meeste werkwoorden en sentimentwoorden– vier kengetallen:

Frequentie woord

De frequentie geeft weer welke woorden het meest werden gebruikt. We berekenen voor elk woord hoe vaak het woord voorkomt en maken een rangorde van alle woorden van hoogfrequent naar laagfrequent, en bepalen het percentiel van de woorden in de frequentieverdeling. Dit kengetal loopt dus van 1 tot 100.

Uniciteit woord

De uniciteit geeft weer hoe uniek het woord is voor het jaar 2021. We delen de frequentie van een woord door de gemiddelde frequentie van het woord in de afgelopen vijf jaar (2017-2021) en maken een rangorde van alle woorden van hoogfrequent naar laagfrequent, en bepalen het percentiel van elk woord in de frequentieverdeling. Dit kengetal loopt dus van 1 tot 100. We bepalen de uniciteit van elke woord binnen elk onderwerp.

Trend woord

Berekent hoe ‘trending’ een woord in 2021 is door de frequentie op te tellen bij de uniciteit. We construeren een woordenlijst van trending woorden door de frequentie en uniciteit op drie manieren te definiëren: gelijke weging frequentie en uniciteit; twee maal zo grote weging uniciteit als frequentie, vier maal zo grote weging uniciteit als frequentie.

(wordt vervolgd op volgende pagina)

³ Merk hierbij op dat woordsynoniemen zijn samengevoegd. Bijvoorbeeld ‘AEX-index’, ‘AEX’ en ‘Beursplein 5’ worden allen ‘Euronext’ en ‘inflatietempo’ en ‘geldontwaarding’ worden ‘inflatie’.

In de figuur zijn verschillende twee duidelijke clusters te zien die gerelateerd zijn aan de coronacrisis. Een felrood cluster met woorden als 'covidpatiënten', 'spreiding' en 'oversterfte', en een veel groener cluster rondom de (goedkeuring van) vaccins. Aan dit laatste cluster verbonden zien we duidelijk een cluster rondom het kabinetsbeleid. Het sentiment daaromtrent is gemiddeld (niet duidelijk groen of rood), met woorden als 'versoepelingen' en 'terrassen' etc. De hard door de coronacrisis geraakte horeca en reisbranche kennen hun eigen cluster met respectievelijk woorden als 'horeca', 'lockdown' en 'evenementen' en 'D-reizen' en 'reisbranche'. Het sentiment werd in 2021 ook deels gedreven door (inter)nationale conflicten, zoals de coronarelschoppers (cluster 'rel', 'relschoppers'), de terugtrekking van Amerikaanse troepen uit Afghanistan en de daaropvolgende machtsovername door de Taliban ('Kaboel', 'Afghanistan', 'Taliban'), onrust in Israël, en de verkiezingen en migrantencrisis in Wit-Rusland. Tot slot was er de nodige aandacht voor Credit Suisse dat verliezen leed op faillissementen van financieel dienstverlener 'Greensill' en hedgefonds 'Archeos' en het tekort aan computerchips door aanbodbeperkingen die met de heropleving van de mondiale economie ontstonden ('chipbedrijf', 'chiptekort' en 'halfgeleiders').⁵ In de Appendix is in Figuur A.2 de netwerkgrafiek weergegeven indien we lijnen trekken tussen de woorden waarbij de correlatie behoort tot de bovenste 10 procent van de woord-woord correlaties. Het beeld is in deze figuur vergelijkbaar.

4 Financiële markten doorstaan coronastorm het best

In dit hoofdstuk gaan we in meer detail in op de deelindicator financiële markten. Deze indicator daalde in 2020 het minste van de vier deelindicatoren. Op het dieptepunt in augustus 2020 droeg de daling van de deelindicator slechts 15 procentpunt bij aan de daling van het headline vertrouwen. Ter vergelijking: voor het sentiment in het bedrijfsleven en de economie was deze negatieve bijdrage respectievelijk 32 en 21 procentpunt. Het herstel van vertrouwen op de financiële markten na het dieptepunt was sterk, in lijn met de stijging van de beursgraadmeters. De deelindicator lag in december 2021 als enige deelindicator boven het pre-crisis niveau (zie ook Figuur 2.3). Het artikel dat in 2021 het meest positief bijdroeg aan de deelindicator financiële markten was 'Als het zo doorgaat, halen Europese aandelen de Amerikaanse nog in' ([link](#)) dat eind juni op de FD-website verscheen, en op 1 juli in de papieren editie. Het artikel met de meest negatieve bijdrage was het FD-artikel 'Citi krijgt foutief overgemaakte \$500 mln voorlopig niet terug' ([link](#)) dat 17 februari op de FD-website verscheen.

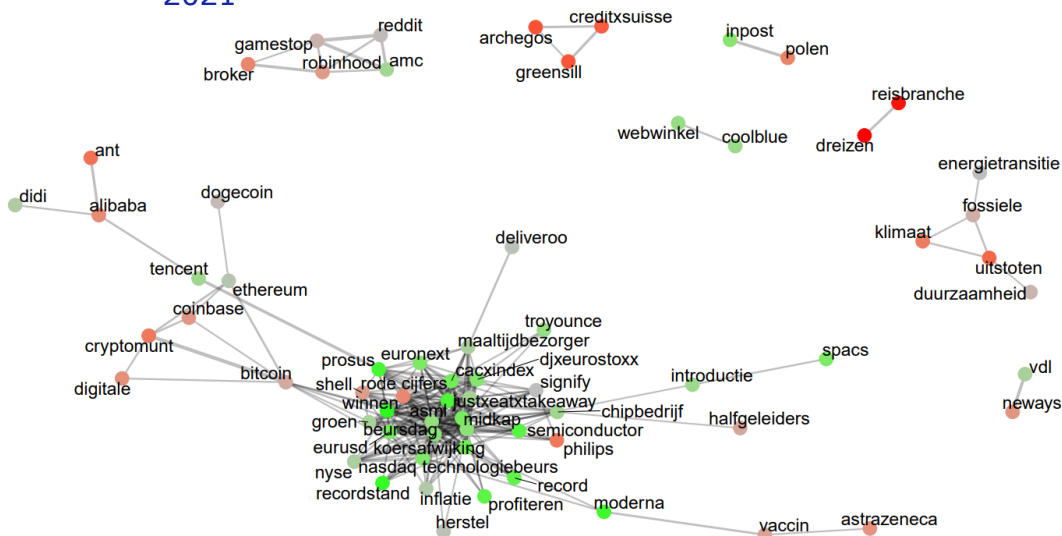
Figuur 4.1 biedt inzicht in de woorden die FD-journalisten in 2021 het meest gebruikte en welke woorden 2021 het best typeren in vergelijking met eerdere jaren binnen het onderwerp financiële markten. Eerst bepalen we wederom welke 5000 artikelen in 2021 het meest bijdroegen aan het door de deelindicator gemeten sentiment op financiële markten. Figuur 4.1A geeft vervolgens de 100 meest gebruikte woorden weer binnen deze groep van 5000 artikelen. In 2021 ging het ook op financiële markten uiteraard vaak over het Covid-19-virus en de pandemie. De woorden 'Corona' en 'pandemie' zijn rood en kennen - logischerwijs - een sterk negatief sentiment. Andere rode woorden zijn 'China', 'probleem', 'financieel', 'toezichthouder' en 'rechtspraak'. De woorden 'beurs' en 'Euronext' zijn groen van kleur, omdat deze

⁵ Zie voor meer informatie onze DNB Analyse over de impact van aanbodbeperkingen op de Nederlandse economie (Hemmerlé en Schaik, 2021).

cluster van de zogenoemde meme-aandelen. De koers van deze aandelen wordt vooral gestuurd door hypes op sociale media en online platforms, met name onder jonge beleggers.⁷ Dit komt tot uitdrukking in het woordnetwerk bestaande uit het internetforum 'Reddit', online broker 'Robinhood' en de aandelen 'AMC' en 'Gamestop'.

Het ging in de financiële katernen van het FD in 2021 vaak over duurzaamheid en de energietransitie ('fossiele', 'klimaat', 'uitstoten', 'duurzaamheid' en 'energietransitie'). Dit reflecteert dat bedrijven en financiële instellingen zelf steeds meer aandacht aan het thema duurzaamheid besteden. Ook de klimaatconferentie in Glasgow zorgde voor extra aandacht in de financiële katernen in het FD. Net als bij het netwerk voor de gehele krant in Figuur 3.2 is een aantal clusters te onderscheiden rondom individuele bedrijven waarover in 2021 rumoer was. Zie het elektronicabedrijf 'Neways', het Poolse pakketkuisjesbedrijf 'InPost' en de bank 'Credit Suisse' die verliezen leed op faillissementen van financieel dienstverlener 'Greensill' en hedgefonds 'Archeos'. Tot slot was er veel aandacht voor de hard door de coronacrisis geraakt reisbranche ('D-reizen', 'reisbranche'). In Figuur A.4 in de Appendix is de netwerkgrafiek weergegeven indien we lijnen trekken tussen de woorden waarbij de correlatie behoort tot de bovenste 10 procent van de woord-woord correlaties.

Figuur 4.2 Netwerkgrafiek trending woorden deelindicator financiële markten, 2021



Bron: FD-database, berekeningen DNB.

Toelichting: woord-woord correlaties voor de woorden uit de wordcloud in Figuur 3B in top 5% woord-woord correlaties. In de figuur zijn de minimale, mediane, gemiddelde en maximale correlaties respectievelijk 0,29, 0,38, 0,41 en 0,99. Rood: relatief negatief sentiment (t.o.v gemiddelde sentiment in 2021), Groen: relatief positief sentiment (t.o.v gemiddelde sentiment in 2021).

5 Corona dominante thema binnen andere deelindicatoren

Voor de overige deelindicatoren zijn – net als voor de deelindicator financiële markten – wordclouds en netwerkgrafieken gemaakt. Met het oog op de lengte van de Analyse zijn deze in de Appendix

⁷ De term 'meme' refereert hierbij aan grappen of parodieën in de vorm van beeld (foto, video) of tekst, die zich heel snel via internet kunnen verspreiden (en vaak ook weer snel verdwijnen).

opgenomen (Figuur A.5 t/m Figuur A.10). We bespreken de resultaten hier kort. Opvallend is dat de clusters rondom corona en vaccinaties bij de onderwerpen `bedrijven`, `economie` en `politiek` in de wordclouds een stuk dominanter zijn dan bij het onderwerp financiële markten. In alle drie de deelonderwerpen is centraal een groot negatief cluster te zien rondom besmettingen en een groot positief cluster te zien rondom de vaccins. Dit geeft aan hoe groot de impact van de coronacrisis op de economie, politiek en het bedrijfsleven in 2021 was.

Verder zijn er bij de deelindicator die het sentiment in het bedrijfsleven beschrijft o.a. clusters te zien rondom het Suezkanaal en het chiptekort. Het schip Ever Given blokkeerde in maart 2021 een week lang het Suezkanaal waardoor leverketens wereldwijd werden verstoord. Het chiptekort beperkte de productie van veel bedrijven gedurende 2021. Bij `economie` is een klimaatcluster met woorden als `klimaat`, `fossiele` en `Glasgow` trending. Dit heeft te maken met de Klimaattop die in oktober en november plaatsvond in de Schotse stad Glasgow. Bij `politiek` is een duidelijk cluster zichtbaar rondom de kabinetsformatie met de namen van bewindslieden `Kaag` en `Rutte`, het kamerlid `Omtzigt` en `formatie`. De sterke link tussen `ouders` en `kinderen` slaat waarschijnlijk op de toeslagenaffaire.

Geraadpleegde literatuur

Blei, D.M., A.Y. Ng en M.I. Jordan (2003) Latent dirichlet allocation. *Journal of Machine Learning Research*, (3), 993–1022, [link](#).

Winter, de J.M. en D.W. van Dijk (2021), Sentimentsindicator op basis van financieel-economisch nieuws, *Economische Statistische Berichten* 106(4799), 340-343, [link](#).

Dnb.nl (2021) DNB meet stand economie met woorden uit de krant, Nieuwsbericht op www.dnb.nl, 26 maart, [link](#).

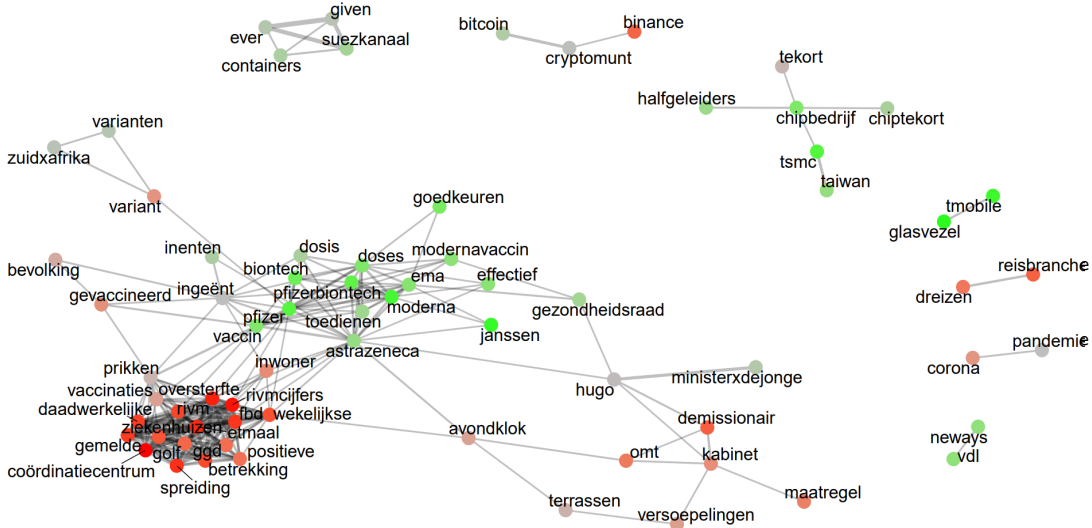
Financieele Dagblad (2021) DNB-onderzoek: het FD voorspelt de economische trend in Nederland, *Financieele Dagblad*, 27 maart, 8-9, [link](#).

Hemmerlé, Y en Schaik, van I. (2021), *Impact Aanbodbeperkingen op de Nederlandse economie in internationaal perspectief*, DNB Analyse. De Nederlandsche Bank.

Loughran, T. en B. McDonald (2011) When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks, *The Journal of Finance*, 66(1), 35–65, [link](#).

Steyvers, M. en T. Griffiths (2007), Probabilistic topic models. In: T. Landauer, D. McNamara, S. Dennis, en W. Kintsch (red.), *Latent semantic analysis: a road to meaning*. Mahwah, NJ: Lawrence Erlbaum, p. 424–440, [link](#).

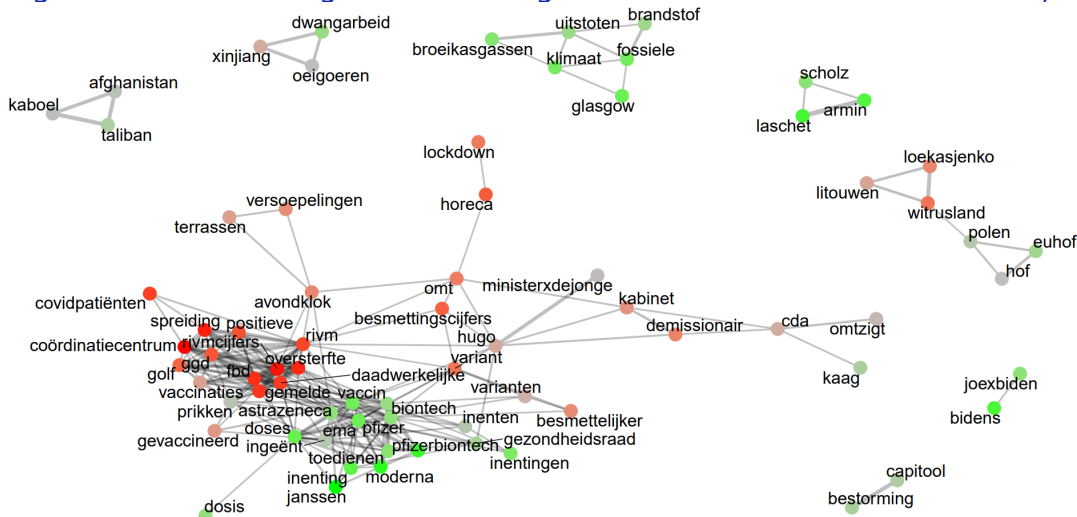
Figuur A.8 Netwerkgrafiek trending woorden deelindicator bedrijven, 2021



Bron: FD-database, berekeningen DNB.

Toelichting: woord-woord correlaties voor de woorden uit de wordcloud in Figuur 3B in top 5% woord-woord correlaties. In de figuur zijn de minimale, mediane, gemiddelde en maximale correlaties respectievelijk 0,33, 0,42, 0,48 en 0,99. Rood: relatief negatief sentiment (t.o.v gemiddelde sentiment in 2021), Groen: relatief positief sentiment (t.o.v gemiddelde sentiment in 2021).

Figuur A.9 Netwerkgrafiek trending woorden deelindicator economie, 2021



Bron: FD-database, berekeningen DNB.

Toelichting: woord-woord correlaties voor de woorden uit de wordcloud in Figuur 3B in top 5% woord-woord correlaties. In de figuur zijn de minimale, mediane, gemiddelde en maximale correlaties respectievelijk 0,28, 0,36, 0,42 en 0,98. Rood: relatief negatief sentiment (t.o.v gemiddelde sentiment in 2021), Groen: relatief positief sentiment (t.o.v gemiddelde sentiment in 2021).

