

Sentimentsindicator op basis van financieel-economisch nieuws

Macro-economische cijfers worden met een vertraging van één of enkele maanden gepubliceerd. Maar nieuws is veel sneller. Zijn krantenberichten te gebruiken om de stand van de conjunctuur te monitoren, en deze beter en eerder te duiden? Een analyse op basis van artikelen in *het Financieele Dagblad* vanaf 1985.

IN HET KORT

- De sentimentsindicator op basis van berichten in de krant loopt vrijwel synchroon met de Nederlandse conjunctuur.
- In tegenstelling tot bestaande metingen, kan deze indicator dagelijks worden geactualiseerd, zonder revisies.
- De stemming rondom COVID-19 was in het najaar van 2020 negatiever dan tijdens het dieptepunt van de financiële crisis.

JASPER DE WINTER

Econoom bij De Nederlandsche Bank (DNB)

DORINTH VAN DIJK

Econoom bij DNB

Aan het eind van de jaren negentig van de vorige eeuw publiceerde *The Economist* (1998) een grafiek met de dagelijkse telling van het woord 'recessie' in de Britse kwaliteitskranten. Dat bleek destijds een goede voorspeller te zijn van recessies in het Verenigd Koninkrijk. Is het mogelijk om deze methode te verfijnen en te gebruiken voor het verbeteren van economische projecties?

Tekst als voorspeller economie

Sinds de publicaties in *The Economist* is het gebruik van tekstanalyse om het verloop van de economie te voorspellen sterk toegenomen. Zo vond Tetlock (2007) dat de teneur in de dagelijkse analisten-column van *The Wall Street Journal* voorspellende waarde had voor de beursbewegingen op de dag dat de column gepubliceerd werd.

In de afgelopen tien jaar heeft de methodologie voor het verrichten van grootschalige tekstanalyses zich sterk ontwikkeld, vooral vanwege de exponentiële toename van computerrekenkracht. Recent doen onder meer de Noorse centrale bank (Thorsrud, 2020) en de Europese Commissie (Barbaglia et al., 2020) interessant onderzoek naar *machine-learning*-algoritmes die uit krantenteksten betekenis en stemming kunnen distilleren.

In dit artikel zijn we geïnteresseerd in de toegevoegde waarde van tekstanalyse voor het voorspellen van de economische groei, gemeten via het bruto binnenlands pro-

duct (bbp). We ontwikkelen een sentimentsindicator op basis van krantenartikelen, met als doel om de groei van het bbp op korte termijn sneller en beter te kunnen voorspellen. Dit wordt, analoog aan het Engelse *nowcasting*, 'nuspellen' genoemd. Daarmee komen we tot projecties van de Nederlandse bbp-groei voor zowel het lopende kwartaal (waarvoor er nog geen realisatiecijfer beschikbaar is) als het eerstvolgende kwartaal.

Een accurate inschatting van de economische groei op de korte termijn is belangrijk als uitgangspunt bij het opstellen van voorspellingen voor de economische groei op de langere termijn. Dit geldt vooral wanneer de economische groei abrupt afneemt, zoals ten tijde van de coronacrisis. Een goede inschatting van de groeivoorzichten is ook voor beleidsmakers cruciaal om goed onderbouwde beleidsbeslissingen te kunnen nemen wat betreft bijvoorbeeld de begroting.

Naast het feit dat een sentimentsindicator op basis van krantenartikelen extra informatie kan verschaffen, is een belangrijk bijkomstig voordeel dat deze dagelijks kan worden geactualiseerd. Let wel, het Centraal Bureau voor de Statistiek (CBS) komt pas 45 dagen na afloop van een kwartaal met een eerste inschatting van de bbp-groei. Stel, er moet een beleidsbeslissing genomen worden in de eerste maand van een kwartaal. Zonder een *nowcasting*-model zou de beleidsmaker meer dan 100 dagen moeten wachten op de eerste CBS-publicatie over de bbp-groei in dat kwartaal.

Proefproject

De Nederlandsche Bank (DNB) heeft in een proefproject het volledig gedigitaliseerde tekstbestand van *het Financieele Dagblad* (FD) gebruikt om te analyseren welk signaal de teksten in de krant afgeven over het conjuncturele verloop. Dit is gedaan door uit de teksten een interpreteerbare sentimentsindicator af te leiden. Daartoe combineren we twee onderzoekstromingen in de econometrische literatuur. De eerste houdt zich bezig met het meten van de stemming in een tekst (Rambaccussing en Kwiatkowski, 2020). De tweede stroom betreft het extraheren van onderwerpen uit teksten op basis van econometrische technieken (Bybee et al., 2020; Hansen et al., 2018; Thorsrud, 2020).

De gebruikte database omvat ruim één miljoen krantenartikelen verschenen in de periode van 1 januari 1985 tot 1 januari 2021. Naast de ruwe teksten bevat de database diverse achtergrondkenmerken, zoals de datum van publi-

catie, het katern waarin het artikel verscheen, de weblink van het artikel en de titel ervan. Omdat het motto van het FD is dat ze schrijven over *“Al het geene tot het financieele eenige betrekking heeft”* is de krant een uitermate geschikte informatiebron voor een indicator die bedoeld is om de stand van de conjunctuur aan te geven.

Onderwerpen in 1,1 miljoen krantenartikelen

De onderwerpen worden uit de teksten 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 er in artikelen met andere onderwerpen andere woorden worden gebruikt.

Met Bayesiaanse schattingstechnieken is vervolgens per artikel 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.

Figuur 1 presenteert een infographic die – in niet-technische termen – weergeeft hoe we komen van de ruwe FD-database tot een indeling in 64 onderwerpen. Voordat het algoritme de database in onderwerpen kan verdelen, moet deze worden opgeschoond (de grijze cirkel in figuur 1a). 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*.

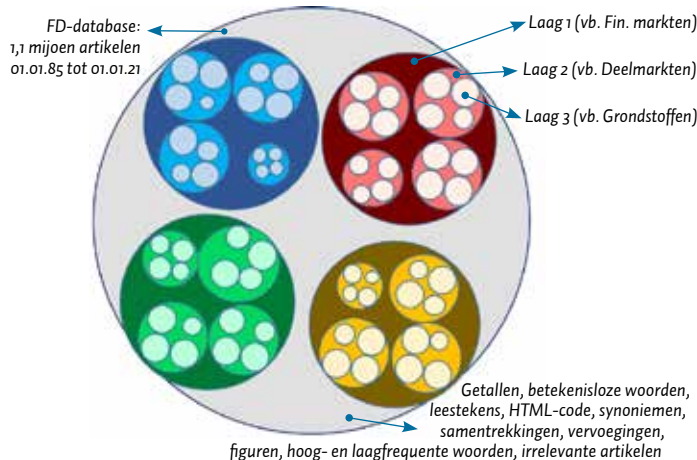
Het topicmodel schatten we vervolgens in drie stappen (lagen). Uit de eerste laag extraheren we vier hoofdonderwerpen die de grote rubrieken in de krant goed blijken weer te geven, te weten: ‘financiële markten’, ‘bedrijfsleven’, ‘economie’ en ‘politiek’. Het model labelt de geëxtraheerde hoofd- een deelonderwerpen niet – dat is aan de onderzoeker. Om een indruk te geven waarom deze labels zijn gekozen, is in figuur 1b weergegeven wat de relevantie is van de woorden in de vier hoofdonderwerpen. 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.

In de tweede laag verdelen we de artikelen per hoofdonderwerp weer op in vier deelonderwerpen, en dit wordt in de derde laag nog een keer gedaan. Zo onderscheiden we binnen het hoofdonderwerp ‘financiële markten’ in de

Extractie onderwerpen uit krantenberichten

FIGUUR 1

1a. Krantendatabase van 1,1 miljoen artikelen naar 64 onderwerpen



Financiële markten	Bedrijfsleven	Economie	Politiek
Deelmarkten	Infrastructuur	Verkiezingen	Tweede Kamer
Grondstoffen	Chemie & farma	Verkiezingen	Partijpolitiek
Valutamarkten	Kengetallen	Oost-Europa	Overheidsbegroting
Buitenland	Mobiliteit	Afrika & Azië	Kabinetten
Monetair beleid	Bedrijfsresultaten	Vereenigde Staten	Ministeries
Financials	Multinationals	Indicatoren	Nationaal
Financiering	Telecom	Internationaal	Rechtspraak
Internationale financials	Klanten	Europa	Pensioen & zorg
Banken	Big tech	Handelspartners	Toezichthouder & bestuur
Verzekeraars	Media	Begrotingsbeleid	Onderwijs & onderzoek
Beursnieuws	Industrie	Grondstoffen	Lagere overheden
Aandelenmarkt	Bouw	Azië	Volkshuisvesting
Fusies & overnames	Logistiek	Olie & gas	Publiek/privaat
Handel	Energie	Conflicten	Landbouw & visserij
Verzekeraars	Industrie	Opkomende economieën	Verkeer & vervoer
Fin. Kengetallen	Demografie	Europese Unie	Sociale partners
Aandelenmarkt	Detailhandel	Duitsland	CAO
Euronext	Failissementen	Europese Unie	Arbeidsmarkt
Analisten	Beursgenoteerd	Italië & Spanje	Ondernemers
Resultaten	Internationaal	Frankrijk	Sociale zekerheid

1b. Woordrelevantie¹ per onderwerp



¹Relevantie is een maatstaf voor het belang en uniciteit van een woord binnen een onderwerp (Steyvers en Griffiths, 2007)

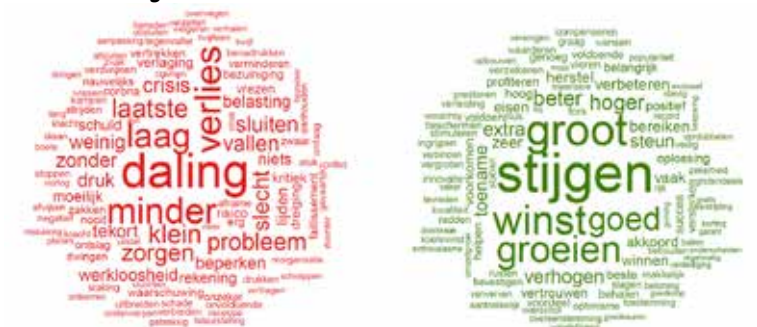
Data: FD | ESB

Frequentie positieve en negatieve sentimentwoorden²

FIGUUR 2

2a. Negatief sentiment

2b. Positief sentiment

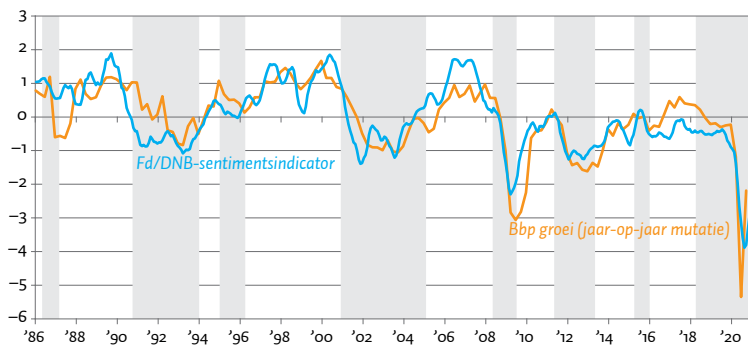


²Tijdperiode 1 januari 1985 tot 1 januari 2021; de figuur geeft de honderd positieve en negatieve sentimentwoorden met de hoogste frequentie en de grootte van een woord geeft de frequentie van het woord aan

Data: FD | ESB

Sentimentsindicator en bbp-groei

FIGUUR 3

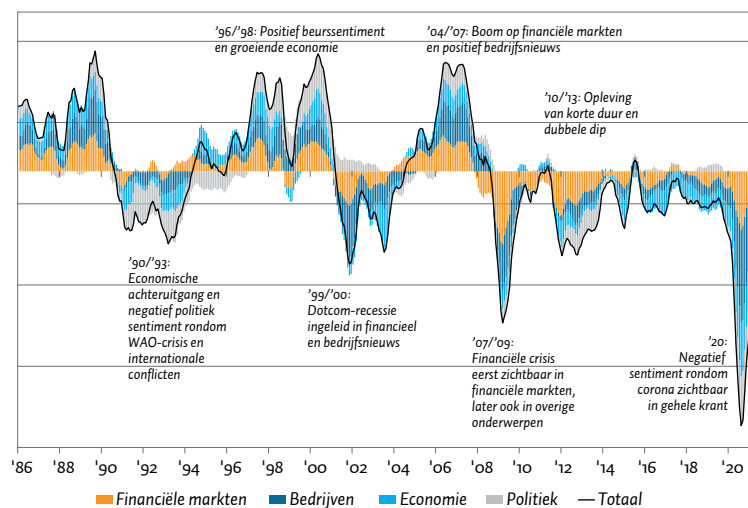


Noot: Genormaliseerd zesmaands voortschrijdend gemiddelde en genormaliseerde jaar-op-jaar groei bbp. Recessies, volgens de OESO-datering, zijn grijs gearceerd en markeren de periode tussen een conjuncturele piek en conjunctureel dal.

Data: OESO, FD | ESB

Sentimentsindicator naar onderwerp

FIGUUR 4



Noot: Genormaliseerd zesmaands voortschrijdend gemiddelde

ESB

tweede laag de deelonderwerpen 'deelmarkten', 'financials', 'beursnieuws' en 'financiële kengetallen'. De lagen passen, zoals in de infographic is weergegeven, als een matroesjkapop in elkaar: alle deelonderwerpen van de derde laag tezamen vormen de tweede laag, en alle deelonderwerpen van de tweede laag de eerste laag. Hierdoor kan elk artikel in onze database tot op de derde laag worden ontleed, en kunnen de onderwerpen die voor een artikel relevant zijn, worden afgeleid.

Het voordeel van het schatten van het topicmodel in lagen is dat het zorgt voor een logische ordening van de onderwerpen. Het is, voor zover wij weten, de eerste keer dat er een gelaagd topicmodel geschat is.

Een tweede innovatie is dat we het topicmodel tijdsvariërend schatten, dus op basis van het woordgebruik en onderwerpen die op dat moment worden gebruikt. Ter illustratie: dit 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. In de economische literatuur is tot nu toe geen rekening gehouden met deze 'tijdsvariatie', waarschijnlijk door de grote rekenkracht die dit model vergt. Het uiteindelijke resultaat is een kansverdeling over 64 onderwerpen per artikel die het aandeel van elk van deze onderwerpen in het FD weergeven.

Metten van de stemming op basis van woorden

Om de stemming oftewel het sentiment te bepalen, kennen we een sentiment score (-1 is negatief en $+1$ is positief) toe aan de woorden in de artikelen. De basis van onze stemmingslijst die de sentiment score per woord bevat is de 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.

Figuur 2a geeft de frequentie van de 100 meestvoorkomende negatieve woorden weer, en figuur 2b de positieve woorden. Hoe groter het woord, des te hoger de frequentie.

Vervolgens meten we het sentiment per artikel door het aantal positieve en negatieve woorden te tellen. De stemming van een artikel is de som van het aantal woorden met een positief sentiment ($+1$) en negatief sentiment (-1). We normaliseren dit per artikel door de som te delen door het totaal aantal woorden in het artikel.

Het sentiment per maand wordt berekend door het gemiddelde te bepalen van het 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.

Sentiment volgt bbp-groei zeer precies

Indien we het berekende sentiment per maand afzetten tegen de jaar-op-jaar groei van het bbp valt direct de zeer sterke samenhang tussen beide op (figuur 3). De jaar-op-jaar mutatie van het bbp wordt algemeen gezien als goede maatstaf voor de stand van de conjunctuur, en vormt de referentiereeks in de veelgebruikte conjunctuurindicatoren van de OESO. De sentimentsindicator volgt het verloop van de bbp-groei zeer nauwkeurig, met uitzondering van de episode rondom de abrupte beurskrach in 1987 (correlatie: 0,78). De indicator lijkt dan ook een zeer tijdig, betrouwbaar signaal af te geven over de conjunctuurbeweging.

De daling van de bbp-groei in de aanloop naar de financiële crisis van 2008–2009 wordt al in 2007 ingezet door de sentimentsindicator. Ook de coronacrisis wordt tijdig opgepikt door de indicator. Er was eind 2019 al een lichte neergang zichtbaar, maar de neergang in maart 2020 en de daaropvolgende maanden is spectaculair te noemen. Het sentiment was in het najaar van 2020 negatiever dan tijdens het dieptepunt van de financiële crisis.

Sentiment per onderwerp

Door de meting van sentiment per artikel en het topicmodel te combineren, kan de stemming worden toegekend aan de 64 onderwerpen die ons model onderscheidt. Dat zit zo: elk artikel heeft een sentimentscore 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 uit figuur 3 onderverdelen in de hoofdonderwerpen ‘politiek’, ‘bedrijfsleven’, ‘economie’ en ‘financiële markten’. Dit kan verder worden onderverdeeld in 16 en uiteindelijk 64 deelonderwerpen. Indien nodig kan er zelfs worden afgedaald naar artikelniveau. Zodoende kan het verhaal achter het verloop van de stemming en daarmee de economische ontwikkeling verteld worden. Hiermee kunnen we de kwaliteit van de conjunctuuranalyse en voorspellingen verbeteren.

Figuur 4 laat zien waar de combinatie van de sentimentsindicator en het topicmodel toe kan leiden. De figuur deelt het sentiment in het belangrijkste nieuws van het FD op wat betreft de vier hoofdonderwerpen in de eerste laag. Bij verschillende episoden is in de figuur een interpretatie gegeven. De bijdrages van de onderwerpen aan het totale gevoel appelleren aan de intuïtie.

Zo is het negatieve gevoel rond de coronaperiode vanaf maart 2020 sterk, en direct zichtbaar in elk hoofdonderwerp. Dit komt overeen met het beeld dat de coronacrisis de gehele maatschappij raakt, en dus ook zichtbaar is in alle rubrieken van de krant. Dit verschilt duidelijk met de stand van de indicator rondom de financiële crisis. De aanloop naar de crisis is in 2007 hoofdzakelijk zichtbaar in het negatieve sentiment rond financiële markten. Pas later – in 2008 en 2009 – is de stemmingsomslag ook zichtbaar in de andere hoofdonderwerpen.

Betere voorspelling bbp-groei

De sterke samenhang tussen de stemming in de hoofdartikelen en de bbp-groei in figuur 3 doet vermoeden dat de indicator ook voorspelkracht heeft voor de bbp-groei. De cijfers over de bbp-ontwikkeling zijn nog niet bekend als het krantensentiment al wel bekend is. Het idee is dat de indicatoren, ruim voordat het bbp verschijnt, de ontwikkeling in delen van de economie meten.

Om te bepalen of de voorspelling van de bbp-groei daadwerkelijk verbeterd kan worden, worden de reeksen van het krantensentiment per onderwerp toegevoegd aan de gebruikelijke indicatoren in een nowcasting-model. We gebruiken het model dat DNB heeft ontwikkeld om een inschatting te maken van de actuele stand van de conjunctuur. Dit zogenoemde DFROG-model maakt projecties van de Nederlandse bbp-groei in het lopende kwartaal (waarvoor nog geen realisatiecijfer beschikbaar is) en het eerstvolgende kwartaal en is uitgebreid getest (Jansen en De Winter, 2018). DFROG is een zogenaamd dynamisch factormodel en bevat ongeveer tachtig economische reeksen die maandelijks beschikbaar zijn. Dit zijn bijvoorbeeld

cijfers over het vertrouwen, productie, verkopen en marktprijzen. Sommige van deze reeksen komen – net zoals het cijfer over de bbp-groei – met enige vertraging uit, terwijl de stemming op basis van krantenartikelen dagelijks beschikbaar is. De selectie van de reeksen in het model is gebaseerd op het feit dat ze (veel) eerder bekend zijn dan het bbp-cijfer, en op hun historische ‘nuspelkracht’ voor het bbp.

Het opnemen van de naar onderwerp verdeelde sentimentsindicator leidt tot een significante afname in de voorspelfout van DFROG. Bij het voorspellen van de groei in het lopende en eerstvolgende kwartaal neemt de voorspelfout met zo’n tien tot vijftien procent af.

Conclusie

Berichten in het FD vormen een relevante nieuwe informatiebron om de actuele stand van de conjunctuur te monitoren en te voorspellen. De uitkomst voor Nederland komt overeen met de uitkomst van soortgelijk onderzoek in Noorwegen (Thorsrud, 2020; Bybee et al., 2020) en Engeland (Rambaccussing en Kwiatkowski, 2020), waarbij ook werd aangetoond dat nieuwssentiment een duidelijke meerwaarde heeft.

Literatuur

- Barbaglia, L., S. Consoli en S. Manzan (2020) *Forecasting with economic news*. mimeo.
- Blei, D.M., A.Y. Ng en M.I. Jordan (2003) Latent dirichlet allocation. *Journal of Machine Learning Research*, (3), 993–1022.
- Bybee, L., B.T. Kelly, A. Manela en D. Xiu (2020) *The structure of economic news*. NBER Working Paper, 26648.
- Hansen, S., M. McMahon en A. Prat (2018) Transparency and deliberation within the FOMC: a computational linguistics approach. *The Quarterly Journal of Economics*, (33), 801–870.
- Jansen, W.J. en J.M. de Winter (2018) Combining model-based near-term GDP forecasts and judgmental forecasts: a real-time exercise for the G7 countries. *Oxford Bulletin of Economics and Statistics*, 80(6), 1213–1242.
- 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.
- Rambaccussing, D. en A. Kwiatkowski (2020) Forecasting with news sentiment: evidence with UK newspapers. *International Journal of Forecasting*, 36(4), 1501–1516.
- 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.
- Tetlock, P.C. (2007) Giving content to investor sentiment: the role of media in the stock market. *The Journal of Finance*, 62(3), 1139–1168.
- The Economist (1998) The recession index. *The Economist*, 10 december.
- Thorsrud, L.A. (2020) Words are the new numbers: a newsy coincident index of the business cycle. *Journal of Business & Economic Statistics*, 38(2), 393–409.