

Big Data en Financieel Toezicht Eindrapport Pilot

Oktober 2018

Inhoudsopgave

1. Inleiding
2. Concept
3. Dataset
4. Model
5. Resultaten
6. Informatiesysteem
7. Bevindingen provincies
8. Conclusies en aanbevelingen
9. Colofon

1. Inleiding

- De afdeling Financiële Regelgeving en Toezicht van het ministerie van BZK voert onderzoek uit naar het financieel toezicht op decentrale overheden (gemeenten, provincies en gemeenschappelijke regelingen). Het ministerie van BZK heeft ABF gevraagd om onderzoek te doen naar de mogelijkheden van het gebruik van (omvangrijke) databestanden bij financieel toezicht. Voorafgaand aan het onderzoek vond de 'Big Data Challenge' plaats, waarin vooronderzoek werd gedaan naar de haalbaarheid en wenselijkheid van het onderzoek. Op een deel van de input die uit deze Big Data Challenge kwam is in het onderzoek voortgeborduurd.
- Het doel van de pilot is om inzicht te krijgen in de verklarende en zelfs voorspellende waarde van financiële en niet-financiële variabelen voor de financiële stabiliteit van gemeenten. Hetzelfde pad is ook verkend voor provincies en gemeenschappelijke regelingen, maar dit heeft niet tot bruikbare resultaten geleid. De pilot bestaat uit twee delen: een verkennende fase in het najaar van 2017 en een validatiefase in het voorjaar van 2018.
- In de eerste fase van de pilot (najaar 2017) is onderzocht in hoeverre het mogelijk is om 'big data' in te zetten voor het verklaren of zelfs voorspellen van de financiële stabiliteit van gemeenten. Hiertoe is een conceptueel model opgezet, data verzameld, en geëxperimenteerd met verscheidene modelvormen. De globale conclusie uit de eerste fase is dat er significante verbanden zichtbaar zijn, maar dat de kracht van het model meer in risicoprofilering moet worden gezocht dan in het maken van accurate voorspellingen voor de toekomst.

1. Inleiding

- In de tweede fase van de pilot (voorjaar 2018) worden data en methodiek verder aangescherpt. Minder goed te duiden variabelen worden nader uitgelicht, vooral daar waar oorzaak-gevolgverbanden nog onscherp zijn, en de instellingen en parameters van het model worden nauwkeuriger afgesteld. Daarnaast wordt in Excel een informatiesysteem ingericht met daarin de gebruikte dataset en de uitkomsten van het model.
- Aan de hand van dit dashboard wordt het model vervolgens gevalideerd door provinciaal toezichthouders van de provincies Gelderland, Zeeland en Zuid-Holland. Zij gaan ieder aan de slag met het systeem voor hun 'eigen' gemeenten, dat naast de eigen methode en bevindingen wordt gelegd. Hierbij is het doel te onderzoeken of er een bruikbaar model ontwikkeld kan worden ter ondersteuning van het financieel toezicht op gemeenten door de toezichthouder.
- De pilot is een doorontwikkeltraject en daarmee in de eerste plaats een zoektocht naar wat wel en niet mogelijk is. De kracht van het model ligt op dit moment meer in risicoprofilering dan in voorspellen, en causale verbanden zijn nog niet altijd eenduidig. Echter, in korte tijd zijn forse stappen gemaakt in termen van modelprestatie. Ook blijkt uit de ervaringen met het dashboard van de provinciaal toezichthouders dat een zeker informatiesysteem van toegevoegde waarde kan zijn op het toezichtwerk.

2. Concept

- Een eerste uitdaging in de zoektocht naar een verklaring voor financiële stabiliteit is het vinden van een geschikte definitie voor financiële stabiliteit. Het ligt voor de hand om dit te doen aan de hand van de zes financiële kengetallen: ongecorrigeerde netto schuldquote, gecorrigeerde netto schuldquote, solvabiliteitsratio, belastingcapaciteit, grondexploitatie en structurele exploitatieruimte. De bruikbaarheid van deze kengetallen is verkend, waarna de keuze is gevallen op de drie kengetallen in onderstaande tabel. De netto schuldquote gecorrigeerd voor alle verstrekte leningen wordt gebruikt, omdat deze meer representatief lijkt voor de financiële positie dan de ongecorrigeerde netto schuldquote. Ook de solvabiliteit van een gemeente blijkt een goede graadmeter. In plaats het 'gewone' kengetal belastingcapaciteit is het OZB-tarief op woningen t.o.v. het landelijk gemiddelde gebruikt, dit omdat de andere twee componenten uit de belastingcapaciteit (riool- en afvallasten) minder variatie kennen en in mindere mate 'stuurbaar' zijn.

	Categorie A minst risicovol	Categorie B neutraal	Categorie C meest risicovol
Gecorrigeerde Netto Schuldquote	< 90%	90% - 130%	> 130%
Solvabiliteitsratio	> 50%	20% - 50%	< 20%
Belastingcapaciteit	< 95%	95% - 105%	> 105%

2. Concept

- Het kengetal structurele exploitatieruimte wordt niet meegenomen in de definitie. Deze blijkt te grillig en te onvoorspelbaar om van waarde te zijn in het model. Het kengetal grondexploitatie komt juist terug in de verklarende kant van het model, in plaats van in de definitie van de financiële positie. Het kengetal bleek van meer waarde te zijn als verklaarder van een verslechterde financiële positie, dan als component van de definitie. Hierover wordt in hoofdstuk 5 nader uitleg gegeven. Gedurende de gehele pilot is de discussie over de gewenste kengetallen gaande gebleven.
- Fundamenteel onderzoek hiernaar komt dan ook terug als één van de aanbevelingen uit de pilot.
- De drie kengetallen definiëren dus gezamenlijk de financiële positie van een gemeente. De tabel bevat de grenswaarden waarop elk van de losse kengetallen is ingedeeld in risicocategorieën. Een gemeente scoort overall een 'A' wanneer ze op tenminste twee kengetallen een 'A' scoort, een 'C' wanneer ze op tenminste twee kengetallen een 'C' scoort, en een 'B' in alle andere gevallen.

	Categorie A minst risicovol	Categorie B neutraal	Categorie C meest risicovol
Gecorrigeerde Netto Schuldquote	< 90%	90% - 130%	> 130%
Solvabiliteitsratio	> 50%	20% - 50%	< 20%
Belastingcapaciteit	< 95%	95% - 105%	> 105%

2. Concept

- In de vorige dia is de financiële positie van een gemeente gedefinieerd aan de hand van drie gangbare financiële kengetallen. Iedere gemeente scoort op basis van de drie kengetallen netto schuldquote, solvabiliteit en belastingcapaciteit overall een 'A', een 'B' of een 'C'. Uit de literatuur, beleidsstukken en de aan de pilot voorafgaande Big Data Challenge komt naar voren dat de relatie tussen deze kengetallen en de financiële situatie van een gemeente groot is. Daarom wordt als gegeven verondersteld dat een

'C'-gemeente een grotere kans heeft om onder verscherpt toezicht te worden gesteld of zelfs een artikel 12-status te krijgen. Dit verband wordt gesymboliseerd door de grijze pijl in onderstaand conceptueel model. Deze veronderstelling wordt ook bevestigd door de ervaringen van de stakeholders van de provincies: doorgaans scoren de gemeenten die 'bekend staan' financieel slechter te presteren (zoals de artikel 12-gemeenten) een 'C' in het model.



2. Concept

- De analyse in de pilot richt zich op de groene pijl: het verklaren en voorspellen van de financiële positie. Het idee hierachter is dat het toetsen van de objectief bepaalde financiële kengetallen meer geschikt is voor een generieke, landsdekkende analyse, dan het kenmerk of een gemeente onder toezicht is gesteld of een artikel 12 status heeft. Immers, dat laatste is vaak ook ingegeven door bestuurlijke context. Ook is het aantal gemeenten met een artikel 12-status door

de jaren heen te klein om betrouwbare analyses op uit te kunnen voeren. Het onderzoek behelst eerst het verzamelen van een set relevante historische gegevens. Hierbij wordt zowel naar financiële als niet-financiële variabelen gekeken en is het nuttig om te experimenteren met effecten door de tijd heen. Vervolgens wordt met (big) data-analyse de zoektocht gestart om verbanden tussen deze historische dataset en de gedefinieerde financiële positie bloot te leggen.



3. Dataset

- Volgens het conceptueel model is voor de data-analyse een grote set van gegevens nodig. Met het oog op het verklaren van de financiële positie van een gemeente wordt voornamelijk gekeken naar variabelen die intuïtief gerelateerd zijn aan financiën. Daarnaast krijgen ook minder voor de hand liggende variabelen een ‘kans’, om ook minder evidente verbanden te onthullen.
- Om een goede dataset samen te stellen is een verkenning gedaan naar verschillende thema's en bronnen. De data moet betrouwbaar en landsdekkend zijn. Daarnaast moet de data zijn teruggebracht naar door de tijd constante regionale indelingen, in dit geval de gemeentelijke indeling van 2017. Of ‘big data’ de juiste term is binnen dit onderzoek is betwistbaar. Deze term is normaal gesproken voorbehouden aan ongestructureerde datasets met miljoenen waarnemingen, waar in dit geval geen sprake van is. Er zijn echter wel parallellen: door het grote aantal inputvariabelen en waarnemingen zijn reguliere analysemethoden niet geschikt.
- Een reeks experimenten met verschillende databronnen en modelmethoden is uitgevoerd. Onderweg zijn enkele tientallen variabelen afgevallen die niet werkbaar bleken, zoals de lastenposten uit de jaarrekeningen van de gemeenten. Deze zorgden voor grote ruis en schommelingen in het model. Ook de financiële positie van de gemeente in voorgaande jaren wordt niet meegenomen als verklaarder. Deze hangt zeer sterk samen met de huidige financiële positie en is een goede voorspeller voor de toekomst. Echter, dit is in feite ‘het weer van morgen voorspellen aan de hand van het weer van vandaag’, terwijl het doel nu juist is om de invloed van ‘externe’ factoren op de financiële positie bloot te leggen.

3. Dataset

De onderstaande lijst met variabelen wordt meegenomen in de schatting van het model. Uit het model zal vervolgens blijken welke

gegevens grote verklarende kracht hebben op de financiële positie van een gemeente, en welke gegevens juist minder van belang zijn.

Bron: CBS

- Bevolking totaal
- Gemiddelde leeftijd bevolking
- Bevolkingsdichtheid
- Groene druk: -19 t.o.v. 20-64 jaar
- Grijs druk: 65+ t.o.v. 20-64 jaar
- Natuurlijke aanwas per 1.000 inw.
- Saldo migratie per 1.000 inw.
- Gemiddelde huishoudensgrootte
- Aandeel alleenstaanden
- Aandeel paren zonder kinderen
- Werkloosheidspercentage
- Bruto arbeidsparticipatie
- Netto arbeidsparticipatie
- Gemiddelde WOZ-waarde woningen
- Aandeel huishoudens met een laag inkomen
- Aantal bedrijfsoprichtingen per 1.000 vest.
- Aantal bedrijfsopheffingen per 1.000 vest.

Bron: BAG

- Aandeel eengezinswoningen
- Aandeel woningen bouwjaar < 1960
- Gemiddelde oppervlakte woningen

Bron: LISA

- Aantal banen per 1.000 inwoners
- Aantal bedrijfsvestigingen per 1.000 inw.

Bron: DUO

- Aantal basisscholen per 10.000 inw.
- Aantal middelbare scholen per 10.000 inw.

Bron: Apollo

- Aanwezigheid HBO/WO ja/nee

Bron: DFG

- Kengetal grondexploitatie

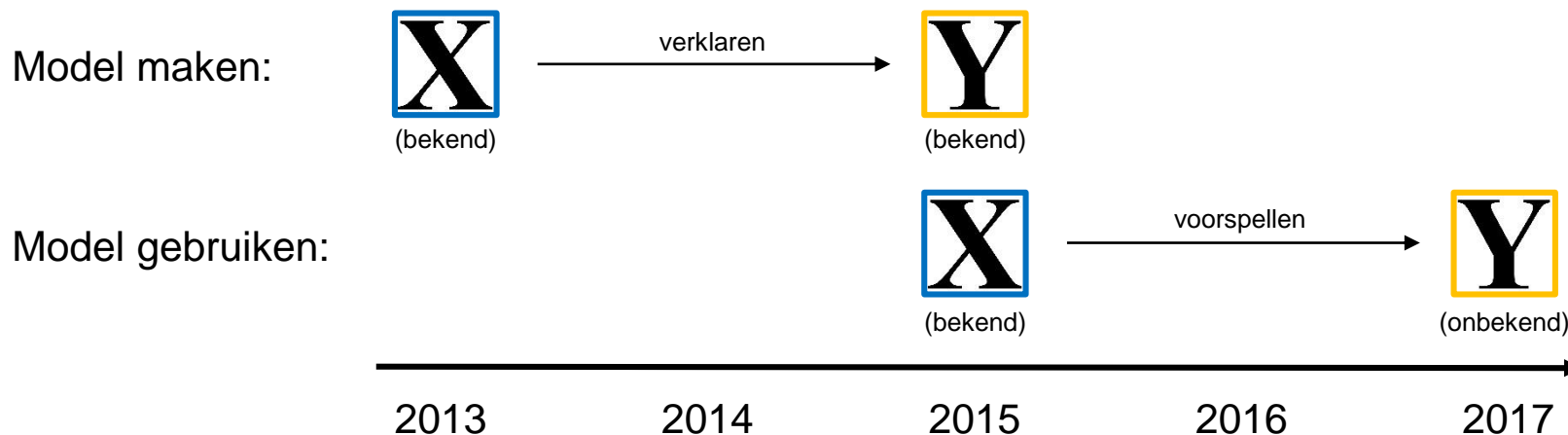
4. Model

- Nadat de dataset is samengesteld is de zoektocht naar verbanden tussen de 26 verzamelde variabelen en de financiële positie gestart. Hierbij wordt regressie-analyse gebruikt, een veelgebruikte statistische analysemethode die zich richt op het schatten van relaties tussen variabelen. Met verscheidene methoden is geëxperimenteerd.
- De LASSO, een Machine Learning techniek, is als beste uit de bus gekomen. Net als klassieke regressie heeft LASSO als doel om relaties tussen de verklarende en de te verklaren variabelen bloot te leggen. Bijzonderheid is dat variabelen met (te) weinig verklarende kracht automatisch buiten beschouwing worden gelaten. Het meenemen van zowel (meerdere) stand- als ontwikkelingsvariabelen van de gegevens is verkend. Het model blijkt het best te presteren wanneer de verklarende kenmerken op één moment worden meegenomen. Wanneer voor een kenmerk, bijvoorbeeld de werkloosheid, voor elk jaar een aparte variabele wordt meegenomen, is de onderlinge samenhang tussen deze variabelen te groot. Het model is niet in staat om verbanden ‘door de tijd heen’ te identificeren, vandaar de keuze om elk kenmerk op één moment te beschouwen.
- Het voorspelmodel richt zich door middel van (logistische) LASSO regressie op het verklaren van een ja / nee kenmerk, namelijk: ‘belandt een gemeente in categorie C?’ Hiermee ligt de focus dus niet meer op het onderscheid tussen categorie A en B, maar richten we ons meer op de kenmerken die leiden tot een financieel zwakke(re) positie.

4. Model

- Met behulp van de LASSO methode wordt een model gemaakt waarin de financiële positie (Y: 'scoort een gemeente een C?') van gemeenten in 2015 verklaard wordt vanuit de 26 kenmerken (X) in 2013, dus twee jaar daarvoor. Uit het model komen de 14 meest relevante kenmerken naar voren. Met de kenmerken, en de vastgelegde

verbanden tussen deze kenmerken en de financiële positie van gemeenten – kunnen voorspellingen worden gemaakt voor de latere jaren. Om de financiële posities van de gemeenten in 2017 te voorspellen worden de waarden van de kenmerken twee jaar daarvoor, dus in 2015, aan het model gegeven als input.



5. Resultaten

- Met de LASSO zijn de 14 kenmerken geselecteerd met de meeste verklarende kracht voor de financiële positie van gemeenten. Deze staan in de tabel hieronder.
 - In de linker kolom staan de variabelen met een goede invloed op de financiële positie (zoals WOZ-waarde: een hoge gemiddelde WOZ-waarde verlaagt de kans op een 'C').
 - In de rechter kolom staan de variabelen met een slechte invloed op de financiële positie (zoals lage inkomens: een hoog werkloosheidspercentage verhoogt de kans op een 'C').

Houdt verband met betere financiële positie (hoge waarde → kleinere kans op "C")	Houdt verband met slechtere financiële positie (hoge waarde → grotere kans op "C")
Gemiddelde WOZ-waarde woningen	Kengetal grondexploitatie
Aantal bedrijfsvestigingen per 1.000 inw.	Aanwezigheid HBO/WO instelling
Gemiddelde huishoudensgrootte	Aandeel huishoudens met een laag inkomen
Saldo migratie per 1.000 inw.	Aantal basisscholen per 10.000 inw.
Aantal bedrijfsoprichtingen per 1.000 vest.	Aandeel alleenstaanden
Aandeel paren zonder kinderen	Werkloosheidspercentage
Netto arbeidsparticipatie	Aandeel woningen bouwjaar < 1960

5. Resultaten

- Het is uiteraard wenselijk dat de gevonden verbanden zoveel mogelijk verklaarbaar zijn. Echter, de kracht van de methode zit juist ook in het vinden van verbanden die niet direct 'logisch' zijn. Mogelijk is een verklarende variabele een 'proxy' voor een ander, niet onthuld fenomeen, en kan nader onderzoek tot meer inzichten leiden.
- De meeste gevonden verbanden tussen de verklarende variabelen en de financiële positie worden door de stakeholders van de provincies als plausibel ervaren (zie hoofdstuk 7). Veel van deze kenmerken hebben van doen met werk, participatie en zelfredzaamheid.
- Enkele gevonden verbanden roepen meer vragen op, en behoeven nadere duiding. Zo toont de hoogte van het kengetal grondexploitatie een sterk negatief verband

met de financiële positie. Een grote grondexploitatie verhoogt dus de kans op een 'C'-score. Dit zet aan het denken, omdat het al dan niet gunstig zijn van een hoge grondexploitatie vooral afhankelijk is van de conjunctuur. Uit het model blijkt dat het 'risico' dat grondexploitatie met zich meeneemt, gemiddeld genomen meer slecht dan goed doet. Belangrijk om hierbij dus te beseffen is dat deze uitkomst gekleurd is door de afgelopen periode. Wanneer gemeenten haar gronden op lange termijn gemakkelijk en tegen een goede prijs kunnen verkopen, zal het verband juist positief uitvallen. Door een ander gebruik van de grondexploitatie, waarin rekening wordt gehouden met de economische omstandigheden, kan de afhankelijkheid van deze economische context weggenomen worden.

6. Informatiesysteem

- Op basis van de dataset en modeluitkomsten is een interactief informatiesysteem in Excel opgezet. Het informatiesysteem bestaat uit twee tabbladen, 'Samenvatting' en 'Verdieping', die hier kort, en in de bijlage in meer detail worden toegelicht. De bijlage bevat ook screenshots van het systeem naar het voorbeeld van een fictieve provincie.
- In het tabblad '**Samenvatting**' wordt een samenvatting van de gegevens en modeluitkomsten van alle gemeenten in de provincie weergegeven. Voor elke gemeente is de financiële positie in 2015 getoond, de waarden van de 14 belangrijkste kenmerken in 2015, en het daaruit bepaalde voorspelde risico voor een relatief slechte financiële positie in 2017. In het tabblad '**Verdieping**' worden voor de geselecteerde gemeente de gegevens en modeluitkomsten nader uitgewerkt aan de hand van een overzichtstabel en een profielkaart. Ook is hier de mogelijkheid de gemeente te vergelijken met een andere gemeente of de provincie.
- Het informatiesysteem biedt op twee manieren inzicht in de financiële positie anno 2017. Enerzijds via de realisatie van de financiële positie in 2015, anderzijds via de voorspelling voor 2017 aan de hand van de variabelen. Deze voorspelling is op zichzelf namelijk bijna nooit volledig toereikend; er kunnen altijd 'ongrijpbare' zaken spelen (bijv. bestuurlijke onrust) waarvan het effect niet in de voorspelling wordt meegenomen. Het is daarom belangrijk om naast de voorspelling voor 2017 ook de gerealiseerde financiële positie in 2015 in beschouwing te nemen.

7. Bevindingen provincies

- In de tweede fase van de pilot is het informatiesysteem in gebruik genomen door de provinciaal toezichthouders van de provincies Gelderland, Zeeland en Zuid-Holland. Van elke provincie hebben één of meerdere toezichthouders een door ABF opgezette vragenlijst ingevuld, waaruit moest blijken in hoeverre er behoefte is aan een dergelijk systeem, hoe de ervaringen met het model en achterliggende ideeën zijn, en hoe het informatiesysteem zelf bevalt. In de zomer van 2018 zijn ABF en BZK op bezoek geweest bij de provincies om de ervaringen te bespreken, mede aan de hand van de vragenlijst. De conclusies en lessen hieruit worden in dit hoofdstuk uiteengezet.
- **Behoeftte ‘big data’ informatiesysteem:** uit de vragenlijsten en gesprekken komt naar

voren dat alle drie de provincies reeds gebruik maken van een (of ander) informatiesysteem bij het toezichtwerk. Dit varieert van eenvoudige Excelsheets met kengetallen tot vrij uitgebreide online databases. Alle ondervraagden geven aan dat een informatiesysteem zoals in de pilot ontwikkeld – nog los van de inhoudelijke uitkomsten – nuttig zouden zijn bij het toezichtwerk. Daarbij wordt vaak de kanttekening gemaakt dat een dergelijk systeem altijd ter ondersteuning van en als aanvulling op de huidige werkwijze zou dienen, en dus nooit als vervanging. Het ideale informatiesysteem bevat eenduidige data die langdurig voorhanden zijn, in de vorm van een interactief Excelbestand (zoals in de pilot) of in een doorklikbare online database.

7. Bevindingen provincies

- **Ervaringen pilot model en achterliggende ideeën:** de benaderde provinciaal toezichthouders geven allemaal aan dat ze zich in redelijke mate kunnen vinden in de opzet en achterliggende theorie van het model. Voor de meeste gemeenten sluit het 'big data' totaalplaatje aan bij het provinciaal oordeel over deze gemeenten: de gemeenten die als 'A' geclassificeerd zijn, worden herkend als gemeenten waarover de provincie weinig zorgen heeft, en vice versa voor de 'C'-gemeenten. Wanneer de toezichthouder zich niet kan vinden in het 'big data' oordeel, blijkt dit meestal een bijzondere, vaak incidentele reden te hebben die niet door de kengetallen en/of het model bevat (kunnen) worden. Een terugkerend punt van discussie vormen de verklarende variabelen in het model, waarbij de behoefte om de gevonden verbanden logisch te kunnen verklaren groot is. De samenhang tussen bepaalde X-variabelen is groot – denk aan werkloosheid en het aandeel lage inkomens – en dit leidt tot onduidelijkheid. Desondanks ziet men de verklarende variabelen als waardevolle 'nieuwe' contextinformatie die inzichten kunnen bieden. Het model houdt rekening met sterke onderlinge samenhang tussen variabelen, echter dit zou duidelijker uitgelegd moeten worden.
- Ondanks dat de definitie van de financiële positie, die nu bestaat uit de kengetallen netto schuldquote, solvabiliteit en belastingcapaciteit, goed aansluit bij de praktijkervaring, leeft het gevoel dat deze verder kan worden verbeterd. De geconsulteerde toezichthouders hebben niet direct concrete oplossingen, maar terugkomende suggesties voor 'toevoegingen' aan de definitie zijn (langdurige) structurele exploitatieruimte, weerstandsvermogen en begrotingsflexibiliteit.

7. Bevindingen provincies

- **Ervaringen pilot informatiesysteem:** het informatiesysteem wordt over het algemeen ervaren als van toegevoegde waarde, al is het aanvankelijk door de hoeveelheid informatie vaak lastig te doorgronden. Los van het onderliggend model, dat inhoudelijk op meerdere punten verbeterd zou kunnen worden, zien de meeste toezichthouders in de klankbordgroep het informatiesysteem als een prima start. De tabel met het overzicht van alle gemeenten in de provincie, ondersteund door het kaartje van de provincie, is helder. De verdiepende profielkaart is volgens sommigen wat moeilijk te interpreteren, maar uiteindelijk wel nuttig om in één oogopslag te kunnen zien hoe het 'oordeel' over een gemeente tot stand komt.
- Een suggestie voor verbetering is om een meer uitgebreide vergelijkingsfunctionaliteit te maken. Het is nu alleen mogelijk om een gemeente te vergelijken met een andere gemeente, of met de (gewogen) gemiddelden van de heel de provincie. Het zou zinvol zijn om gemeenten ook te kunnen vergelijken met bijvoorbeeld andere gemeenten in dezelfde gemeentegrootteklasse, of met andere gemeenten die vergelijkbaar zijn op bepaalde kenmerken. In het algemeen kan het informatiesysteem worden verbeterd door meer mogelijkheden in te bouwen om zelf met de data en presentatiewijze te 'spelen'. Een systeem zoals gebruikt in de Databank Financiën Gemeenten zou in deze wensen kunnen voorzien.
- Een laatste suggestie voor verbetering van het dashboard is om trends van de kengetallen over de jaren te kunnen inzien. Een praktisch knelpunt hierbij is op dit moment nog de beschikbaarheid van de kengetallen, die pas sinds enkele jaren zo voorgeschreven en daarom beschikbaar zijn.

8. Conclusies en aanbevelingen

- De algemene conclusie van de pilot is dat de data-analyse aan de hand van de LASSO techniek waardevolle inzichten biedt. De modeluitkomsten bieden een indicatie van de verklarende waarde van verscheidene financiële en niet-financiële variabelen voor de financiële stabiliteit van een gemeente. De meeste gevonden significante verbanden worden door de stakeholders van de provincies als plausibel ervaren, andere vragen om meer duiding. Er zijn enkele concrete suggesties voor inhoudelijke verbeteringen gedaan.
- De bruikbaarheid van de uitkomsten moet meer worden gezocht in het kader van risicoprofilering van gemeenten dan in daadwerkelijk voorspellen hoe gemeenten zich financieel gaan ontwikkelen. Het model is logischerwijs niet in staat tot het maken van perfecte voorspellingen. We hebben geen ‘heilige graal’ in handen, en het blijft zaak om goed na te denken over de manier van inzetten: door wie, wanneer en hoe?
- Modelmatige mogelijkheden tot verbetering liggen zowel in de definitie van de financiële positie als in de verklarende kant van het model. De hoeveelheid wetenschappelijk onderzoek gedaan naar gemeentefinanciën is beperkt, dus er was weinig om in de pilot op voort te borduren. Veel van de verklarende indicatoren kunnen worden verbeterd wanneer meer aandacht uitgaat naar de juiste definitie ervan. Een voorbeeld is de grondexploitatie: wanneer een variant wordt meegenomen waarin rekening wordt gehouden met de economische omstandigheden, kan waarschijnlijk een beter en helderder verband met de financiële positie worden blootgelegd. Het lag buiten de scope van de pilot om elke variabele zodanig individueel uit te lichten dat hierin de beste keuze kon worden gemaakt.

8. Conclusies en aanbevelingen

- Ook de gewenste definitie van de financiële positie is een belangrijk discussiepunt. De definitie zoals uiteindelijk gebruikt in de pilot wordt ervaren als incompleet en niet volledig toereikend. Terugkomende suggesties voor 'toevoegingen' aan de definitie zijn weerstandsvermogen, begrotingsflexibiliteit en (langdurige) structurele exploitatieruimte. Het gebruik van de structurele exploitatieruimte is verkend en bleek niet geschikt in de pilot. Een 'slimmere' variant, waarin de exploitatieruimte niet voor één jaar maar voor meerdere historische jaren wordt beschouwd, zou wel van waarde kunnen zijn. Daarnaast is de vraag of de grondexploitatie, die nu als verklarende variabele wordt gebruikt, niet juist deel zou moeten uitmaken van de kengetallen die de financiële positie definiëren.
- Een andere, belangrijke aanbeveling betreft het uitvoeren van fundamenteel onderzoek naar zowel de y-kant (het definiëren van de financiële positie van een gemeente) en de x-kant (welke indicatoren hebben hierop invloed) van het verhaal. Resultaten uit dergelijk onderzoek kunnen op termijn dienen als nieuwe input voor het model.
- In algemene zin lijkt de branche in zekere zin niet toe aan (big) data-analyse zoals verkend in deze pilot, omdat de behoefte aan duiding en het vinden van 'logische', intuïtieve verbanden heel groot is.

8. Conclusies en aanbevelingen

- Wat betreft de gevonden verbanden tussen de verklarende variabelen en de financiële positie is het in algemene zin van belang om correlatie niet te verwarren met causaliteit. Wanneer een variabele, bijvoorbeeld de gemiddelde WOZ-waarde, positief blijkt te samen te hangen met de financiële positie, wil dit nog niet zeggen dat deze variabele daadwerkelijk een betere positie *veroorzaakt*. Het is alleen bekend dat een hogere gemiddelde WOZ-waarde vaak *gepaard gaat* met een betere financiële positie. Bij het verklaren en voorspellen van de financiële positie van een gemeente op zich is het niet van belang om vast te stellen of een verband niet alleen correlatie, maar ook causaliteit betreft. Of er sprake is van causaliteit is natuurlijk wel interessant, omdat daaruit kan blijken of het zin heeft om op een bepaalde variabele te 'sturen'.
- Echter, het onderzoek in deze pilot behelsde niet het per indicator achterhalen of er al dan niet sprake is van een causaal verband.
- Meerdere indicatoren in de dataset zijn tevens maatstaf in het gemeentefonds, bijvoorbeeld het aandeel alleenstaanden. Dit bemoeilijkt het vinden van (sterke) verbanden tussen deze indicatoren en de financiële positie van een gemeente, omdat deze ook al opgenomen zijn in het gemeentefonds. Mede om deze reden is er behoefte aan gemeentekennmerken die geen rol spelen in het gemeentefonds. Hierbij kan gedacht worden aan 'soft controls'. Voorbeelden hiervan zijn het ziekteverzuim bij de gemeente of een bepaalde indicator voor bestuurlijke onrust. In de pilot was dit soort data niet voorhanden.

8. Conclusies en aanbevelingen

- Het opgezette pilot informatiesysteem, dat getest is door provinciaal toezichthouders uit drie verschillende provincies, wordt over het algemeen ervaren als een goede start. Het dashboard dient voor de benaderde toezichthouders voornamelijk als contextinformatie. Een dergelijk systeem zou altijd ter ondersteuning van en als aanvulling op de huidige werkwijze dienen, en dus niet als vervanging.
- Nog los van het presteren van het huidige model is duidelijk dat een dergelijk systeem wenselijk is naast het huidige instrumentarium, al is het maar in de vorm van een database met tijdreeks van de variabelen die als relevant uit het model naar voren komen. Het is hierbij belangrijk om goed na te denken over niet alleen de kansen, maar ook de bedreigingen van een informatiesysteem met verklaringen en voorspellingen van de financiële situatie van gemeenten.
- Hoe de pilot ‘big data en financieel toezicht’ eventueel kan worden uitgerold in de praktijk moet de komende tijd worden bezien. Mogelijk kan een informatiesysteem vergelijkbaar met het systeem uit de pilot worden opgezet, dat door elke provincie naar eigen inzien kan worden gebruikt als aanvulling op hun huidige werkwijze.

- In het tabblad ‘**Samenvatting**’ wordt een samenvatting van de gegevens en modeluitkomsten van alle gemeenten in de provincie weergegeven. Voor elke gemeente wordt de financiële positie in 2015 getoond: een ‘A’, ‘B’ of ‘C’ die bepaald is aan de hand van de drie kengetallen in dat jaar. De middelste kolommen bevatten de waarden in 2015 van de 14 belangrijkste verklarende (X-)variabelen in het model. Wanneer een gemeente goed scoort op een variabele wordt een groen vinkje getoond, bij een slechte score een rood kruisje. In de laatste kolom wordt het relatieve risico op een ‘C’ – dus een relatief slechte financiële positie – in 2017 getoond. Deze is bepaald in het model aan de hand van de scores op de X-variabelen in 2015. Een score van 1 staat voor een gemiddeld risico, een score van 2 voor een risico twee keer zo groot als landelijk gemiddeld, en een score van 0,5 voor een risico twee keer zo klein als landelijk gemiddeld. De vinkjes en kruisjes op de variabelen in de voorgaande kolommen geven een globale indicatie van de ‘oorzaken’ van het voorspelde risico: veel rode kruisjes leiden tot een hoog risico, veel groene vinkjes tot een laag risico.

Bijlage: informatiesysteem 'Samenvatting'

Gemeente	Positie 2015	Hoge waarde is goed							Lage waarde is goed							Risico 2017
		WOZ-waarde	Bedrijfsvestigingen	Huishoudensgrootten	Migratiesaldo	Bedrijfsoprichtingen	Paren zonder kinderen	Netto arbeidsparticipatie	Grondexploitatie	Aanwezigheid HBO/WO	Lage inkomens	Basisscholen	Alleenstaanden	Werkloosheidspercentage	Bouwjaar < 1960	
Abbestad	B		✓	✓	✗		✓	✓	✓		✓	✗		✓	✗	0,4
Baalderdam	B		✓	✗								✗				0,8
Berkendaal	A	✗		✗			✓	✗							✗	0,7
Dombergen	B		✓	✓			✓	✓	✗		✓		✓	✓		0,6
Dinkelo	C	✗		✗			✗		✗	✗	✗	✓	✗			3,7
Jonkersfeen	A	✗	✓	✗	✗		✓	✗				✗		✗		0,8
Keteldorp	A	✗		✓	✗			✓			✗		✓	✗		0,7
Kinkeldam	B		✓	✗	✗		✓				✗		✓	✗		0,5
Lenten	B	✗	✓	✗	✗		✓	✗			✗	✗	✓	✗		0,7
Nederpoort	C	✗		✗			✓	✗	✗							2,0
Stellendijk	A	✗	✓	✓	✗						✗		✓	✗		0,9
Uitkerken	A	✓	✓		✗		✓			✓	✗		✓	✗		0,4
Zorghoek	C	✗	✗	✗			✗	✗	✗	✗	✓	✗	✗	✗		3,5

- In het tabblad ‘**Verdieping**’ worden voor de geselecteerde gemeente de gegevens en modeluitkomsten nader uitgewerkt. Links in het tabblad wordt een overzichtstabel getoond. Hierin wordt de realisatie van de financiële positie in 2015 uiteengezet aan de hand van de drie financiële kengetallen in dat jaar. Daaronder worden de 14 belangrijkste verklarende variabelen uitgelicht. Waar het tabblad ‘Samenvatting’ voor elke gemeente alleen vinkjes of kruisjes staan bij de variabelen, worden in

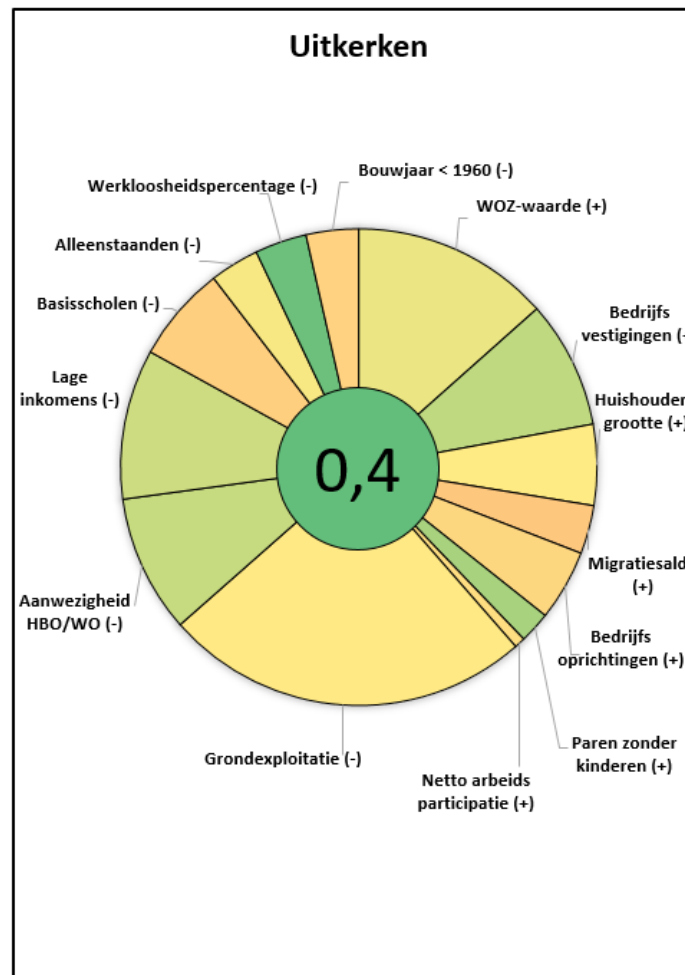
het tabblad ‘Verdieping’ de daadwerkelijke waarden in 2015 getoond. Ook geven pijltjes aan of de waarde van het gegeven is gedaald of gestegen ten opzichte van 2013. Naast de overzichtstabel wordt een ‘profieltaart’ getoond, die aan de hand van vlakken en kleurtjes inzicht biedt in hoe de voorspelling voor 2017 is opgebouwd. Tot slot is er de mogelijkheid de gemeente te vergelijken met een andere gemeente of de provincie.

Bijlage: informatiesysteem 'Verdieping'

Selecteer hier de gewenste gemeente
Uitkerken

Realisatie financiële positie in 2015	A
Solvabiliteitsratio	B 35%
Netto schuldquote (gecorrigeerd)	A 43%
OZB-tarief woningen	A 10%

Risico op slechte financiële positie in 2017	0,4
Gemiddelde WOZ-waarde woningen in €1000 (2015)	✓ ↑ 272
Aantal bedrijfsvestigingen per 1000 inwoners (2015)	✓ ↑ 116
Gemiddelde huishoudensgrootte (2015)	↑ 2,4
Saldo migratie per 1000 inwoners (2015)	✗ ↓ 2,2
Aantal bedrijfsopr. per 1000 vestigingen (2015)	↑ 160
Aandeel paren zonder kinderen (2015)	✓ ↑ 42%
Netto arbeidsparticipatie (2015)	↑ 65%
Grondexploitatie (2015)	↑ 19%
Aanwezigheid HBO/WO instelling (2015)	⇒ Nee
Aandeel lage inkomens (2015)	✓ ↑ 4,1%
Aantal basisscholen per 10000 inwoners (2015)	✗ ↓ 7,1
Aandeel alleenstaanden (2015)	↓ 29%
Werkloosheidspercentage (2015)	✓ ↓ 4,0%
Aandeel woningen bouwjaar < 1960 (2015)	✗ ↑ 32%



9. Colofon

- Dit rapport is samengesteld in opdracht van het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties door Berry Blijie en Denise Ligthart. Voor meer informatie kunt u contact opnemen met berry.blijie@abf.nl of denise.ligthart@abf.nl.
- Naast nauwe samenwerking met de projectgroep van BZK is tijdens het eerste deel van de pilot een drietal sessies georganiseerd met een klankbordgroep van experts op verschillende gebieden. In de tweede fase zijn de toezichthouders van de provincies Gelderland, Zeeland en Zuid-Holland onderdeel geweest van de testgroep. Onze dank gaat uit naar beide groepen voor de zeer bruikbare bijdragen die zij hebben geleverd aan het eindresultaat van de pilot.
- *Copyright ABF Research 2018: De informatie in dit rapport is met de grootste zorg samengesteld. ABF Research aanvaardt geen aansprakelijkheid voor eventuele fouten, onnauwkeurigheden of onvolledigheden. Het gebruik van (onderdelen van) dit rapport is toegestaan mits de bron duidelijk wordt vermeld.*