



Centraal Planbureau

CPB Achtergronddocument | 8 maart 2018

Voorspellen van de werkloosheid

Kan het beter?

CPB Achtergronddocument

Voorspellen van de werkloosheid: kan het beter?

8 maart 2018

Yvonne Adema, Kees Folmer, Harro van Heuvelen, Sonny Kuijpers, Rob Luginbuhl en Bas Scheer

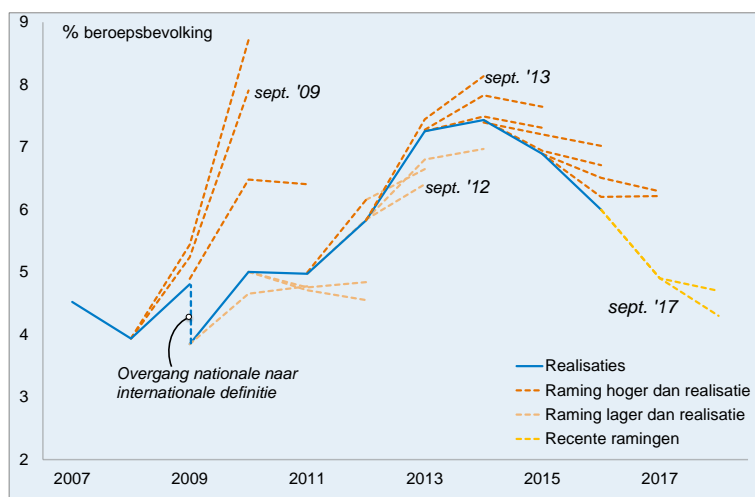
Inhoud

1	Inleiding.....	3
2	De voorspelkwaliteit van kortetermijnramingen.....	4
2.1	De gemiddelde voorspelfout per instituut.....	5
2.2	Voorspelfouten over de tijd.....	6
3	Methoden en technieken	7
4	Toepassingen van de wet van Okun.....	10
4.1	Toepassingen in de literatuur.....	10
4.2	Empirische uitwerking	11
4.3	Regressieresultaten	12
4.4	Voorspellen van de werkloosheid.....	13
4.4.1	Data	13
4.4.2	Resultaten van de voorspellingen	13
5	Tijdreeksmodellen.....	14
5.1	Benchmarkmodel.....	15
5.2	VAR- en BVAR-modellen.....	15
5.2.1	VAR-modellen.....	15
5.2.2	BVAR-modellen.....	16
5.3	Data.....	17
5.4	Resultaten	18
6	Derde alternatief: stromenanalyse.....	19
6.1	Methodiek in het kort.....	19
6.1.1	Empirische aanpak	20
6.2	Data.....	22
6.3	Resultaten	24
7	Combinatievoorspelling.....	26
8	Omslagpunten	27
9	Conclusies	29
	Appendix.....	31
A.1	Achtergrond bij Okun-regressies.....	31
A.2	Afleiding driestatenmodel in hoofdstuk 6.....	32
A.3	VAR-methode in hoofdstuk 6.....	33
	Literatuurlijst.....	37

1 Inleiding

Sinds het uitbreken van de financiële crisis in 2009 is het verschil tussen het geraamde en gerealiseerde werkloosheidspercentage groot, zie figuur 1.1.

Figuur 1.1 Werkloosheidsramingen CPB t.o.v. realisatie



Bron: MEV- en CEP-publicaties. Dit zijn realisaties ten tijde van de laatste raming gecorrigeerd voor latere CBS-revisies.

Het doel van dit document is het ontwikkelen van een aantal methodes om de kortetermijnramingen van de werkloosheid te ondersteunen.¹ Het doel is niet het voorspellen van omslagpunten, al laten we wel zien hoe de verschillende alternatieven presteren in de buurt van een omslagpunt.

Eerst passen we onderzoek van Gordon (2010) en Ball et al. (2017) toe op Nederland. Zij bouwen voort op de wet van Okun, die een verband formuleert tussen werkloosheid en bbp. Tijdreeksmodellen met (veel) meer verklarende variabelen zijn een alternatief. Naast de traditionele vector-autogressieve (VAR) modellen besteden we vooral aandacht aan het Bayesiaanse VAR-model. Ten slotte gebruiken we het werk van Barnichon en Nekarda (2012) en Barnichon en Garda (2016) om met VAR- en BVAR-methoden de werkloosheid te voorspellen uit stromen op de arbeidsmarkt. Bij alle modellen toetsen we de voorspelkwaliteit. Ten slotte combineren we de beste resultaten van alle methoden om te zien of dit de voorspelkwaliteit verbetert. Voor ramingen over een periode langer dan twee jaar zijn deze methoden minder geschikt. Het is beter om hiervoor een structureel model (zoals op dit moment in gebruik bij het CPB) te gebruiken.

¹ De werkloosheid wordt voorspeld met het structurele macromodel SAFFIER II, als het verschil tussen het arbeidsaanbod en de werkgelegenheid in personen. Het arbeidsaanbod volgt vooral trendmatige factoren, zoals demografische ontwikkelingen. De werkgelegenheid vloeit voort uit een CES-productiefunctie en hangt af van het productievolume, de winstgevendheid, de relatieve arbeidskosten, de arbeidsbesparende technologische vooruitgang en de arbeidstijd. Voor meer details zie Verbruggen et al. (2010). Vervolgens wordt er op basis van *expert opinion* bepaald of de modeluitkomsten bijgestuurd moeten worden.

We concluderen allereerst dat de geschatte Okun-modellen niet tot betere kortetermijnvoorspellingen leiden. Dit geldt wel voor het beste tijdreeksmodel (de zogenaamde directe BVAR-specificatie) en de combinatievoorspelling: deze verbeteren de voorspelling van de werkloosheid tot vijf kwartalen vooruit. Ook na een omslagpunt voorstellen deze methoden beter, maar vanaf zes kwartalen vooruit zijn de CPB-voorspellingen weer nauwkeuriger. Met deze nieuwe methoden beschikt het CPB over een extra bron van informatie om de kortetermijnramingen van de werkloosheid verbeteren.²

Hoofdstuk 2 laat zien dat de voorspelkwaliteit van de werkloosheid niet systematisch verschilt tussen de verschillende ramingsinstituten. In de daaropvolgende hoofdstukken staan we stil bij de alternatieve benaderingen. Hoofdstuk 3 behandelt de gevolgde methoden om tot de modelselectie te komen. Hoofdstuk 4 onderzoekt specificaties die steunen op de wet van Okun. Hoofdstuk 5 gaat over de voorspelkwaliteit van pure tijdreeksmodellen. Hoofdstuk 6 benadert de werkloosheid als het resultaat van stromen op de arbeidsmarkt, en daarna combineert hoofdstuk 7 de verschillende methodes om te analyseren of dit de voorspellingen (verder) verbetert. Ten slotte onderzoekt hoofdstuk 8 de voorspelkwaliteit van de verschillende methodes na omslagpunten. Het gaat hierbij dus niet om het voorspellen van de omslagpunten zelf. Hoofdstuk 9 concludeert en bespreekt de vervolgstappen van dit onderzoek.

2 De voorspelkwaliteit van kortetermijnramingen

Een trefzekerheidsanalyse laat zien dat de ramingen van het CPB niet systematisch beter of slechter zijn dan die van andere instituten, zoals De Nederlandsche Bank (DNB), de Organisatie voor Economische Samenwerking en Ontwikkeling (OESO) en de Europese Commissie (EC).³

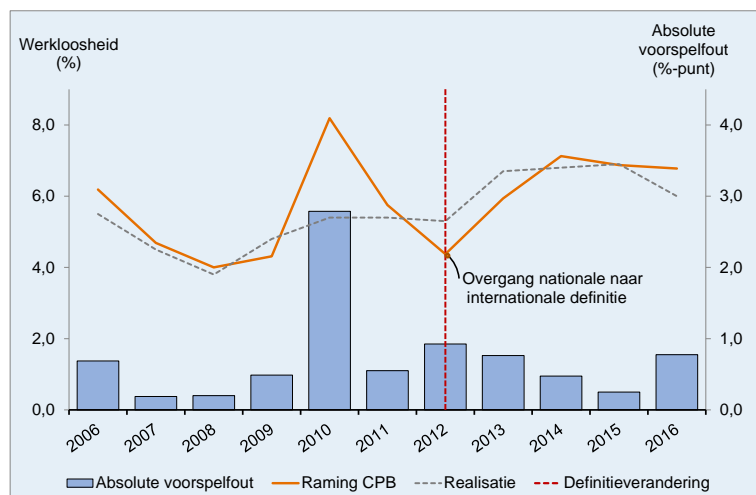
Zoals gezegd was het de afgelopen jaren lastig de werkloosheid goed te ramen. Tijdens de financiële crisis werd de verandering in de werkloosheid zowel over- als onderschat. Bij het uitbreken ervan in 2009 liep deze veel langzamer op dan verwacht en bij de tweede dip in 2012 steeg deze juist sneller. Tijdens de herstelfase vanaf 2015 daalde de werkloosheid sneller dan geraamd (zie figuur 2.1). De raming in 2009 voor 2010 zit er het verst naast: het verwachte werkloosheidspercentage was 8,2%, maar het kwam uit op 5,4%.

² Behalve de werkloosheid worden nog veel andere macro-economische variabelen geraamd. Omdat dit met elkaar een consistent beeld moet opleveren, zullen we de nieuwe methoden slechts als 'zijlicht' gebruiken. De voorspellingen van deze zijlichten zullen naast de uitkomsten van het macromodel gehouden worden en op basis van *expert judgement* wordt dan bepaald of de modeluitkomsten bijgestuurd moeten worden.

³ Met dank aan DNB voor het beschikbaar stellen van aanvullende data.

Figuur 2.1 toont de CPB-ramingen vanaf 2006. Het CPB publiceert jaarlijks vier ramingen: in het CEP in maart, de juniraming, in de MEV in september en ten slotte de decemberraming. De grafiek illustreert het gemiddelde werkloosheidspercentage uit deze publicaties. De absolute voorspelfout is berekend als het verschil tussen deze gemiddelde raming en de realisatie.

Figuur 2.1 Raming en realisatie werkloosheidspercentage (één jaar vooruit)



2.1 De gemiddelde voorspelfout per instituut

Tabel 2.1 laat de gemiddelde absolute voorspelfouten van de vier genoemde instituten zien. De verschillen tussen het CPB en de andere instituten zijn klein en statistisch insignificant. De fout is het kleinst bij DNB met 0,59%-punt; bij het CPB ligt deze 0,05%-punt hoger. De OESO en de EC maken gemiddeld iets grotere fouten dan het CPB en DNB. Dit komt waarschijnlijk doordat het CPB bij het opstellen van de MEV en de decemberraming over meer feitelijke informatie kan beschikken dan de buitenlandse instituten, omdat zij eerder uitkomen met hun raming.

Tabel 2.1 Absolute voorspelfouten in de gemiddelde ramingen (één jaar vooruit)

Instituut	Periode	Absolute voorspelfouten (%-punt)		
		Gemiddeld	Mediaan	Max
CPB	2003 – 2016	0,64	0,50	2,6
DNB	2003 – 2016	0,59	0,53	1,9
OESO	2003 – 2016	0,66	0,55	1,6
EC	2003 – 2016	0,67	0,50	1,5

Absolute voorspelfout (%-punt) = |Raming- Realisatie|.

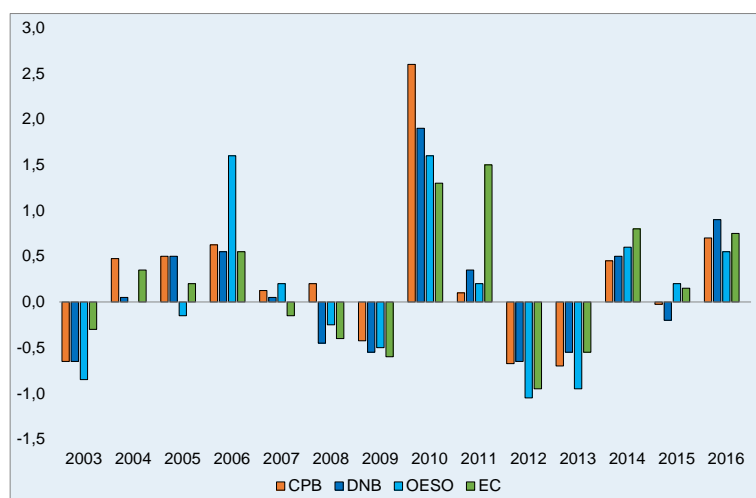
Het CPB heeft de grootste maximale voorspelfout (2,6%-punt in 2010). Deze uitschieter zorgt ervoor dat de gemiddelde fout van het CPB hoger uitkomt dan die van DNB. De mediaan van het CPB is lager dan van DNB en de OESO, want dit is de waarneming die precies in het midden ligt. Deze is dus minder gevoelig voor uitschieters.

2.2 Voorspelfouten over de tijd

We kunnen dus concluderen dat de gemiddelde voorspelfout van de vier instituten dicht bij elkaar ligt. Er zijn geen structurele verschillen tussen de instituten. De verschillen in afzonderlijke jaren zijn af en toe wel groot (zie figuur 2.2), maar er is geen instituut dat structureel de kleinste of grootste voorspelfouten heeft. Ook tijdens de financiële crisis maakten ze alle vier vergelijkbare fouten.

Zo overschatte de OESO het werkloosheidspercentage met meer dan 1%-punt in 2006. In 2010 raamde zowel het CPB als DNB de werkloosheid te hoog, met respectievelijk 2,6%-punt en 1,9%-punt. De EC zat er in 2011 met 1,5%-punt het verst naast. Omdat het CPB vier keer per jaar een voorspelling uitbrengt, loopt de gebruikte informatie niet synchroon met die van met name de OESO en de EC.

Figuur 2.2 Voorspelfout over de tijd (één jaar vooruit, %-punt)



Dit figuur toont de gemiddelde voorspelfout (raming minus realisatie) van het CPB en andere ramingsinstituten per jaar. DNB, OESO en EC publiceren allemaal twee ramingen per jaar: de eerste in mei/juni en de tweede in november/december. De voorspelfout per jaar is daarom het gemiddelde van deze twee publicaties per jaar. Voor het CPB gebruiken we ook het gemiddelde van de juni- en decemberraming.

De ramingen van andere instituten zijn dus zeker informatief voor het CPB, maar ze bieden weinig houvast om de eigen voorspellingen systematisch te verbeteren. Daarom onderzoeken we in de volgende hoofdstukken alternatieve methoden om de ramingen te ondersteunen.

3 Methoden en technieken

Dit hoofdstuk bespreekt de methoden om de voorspelmodellen te selecteren.

Onze steekproef bestaat uit kwartaaldata vanaf het tweede kwartaal van 2003. We hebben dus een beperkt aantal waarnemingen om een model te schatten en de voorspelkwaliteit te testen. Tegelijkertijd is er een overdaad aan mogelijk relevante variabelen. Zo ontstaat het gevaar van *overfitting*; het kader aan het eind van dit hoofdstuk licht dit probleem verder toe. Het is daarom eerst nodig een optimaal model te selecteren om te kunnen voorspellen.

Statistical learning en modelselectie

Statistical-learning-technieken zijn nuttig bij het selecteren van een optimaal model. Een aantal technieken uit dit vakgebied zijn namelijk specifiek ontwikkeld voor de selectie van optimale voorspelmodellen. Een belangrijke stap hierbij is cross-validatie: deze selecteert het optimale model op basis van *out-of-sample*-voorspelfouten. Voorspellen is immers per definitie een *out-of-sample*-probleem, dus is het ook beter om bij de modelselectie een *out-of-sample*-selectiecriteria te gebruiken (Wooldridge, 2012).

De literatuur over *statistical learning* laat zien dat het creëren van een voorspelmodel op basis van *out-of-sample*-optimalisatie doorgaans bestaat uit de volgende drie stappen (James et al., 2013). Deze stappen verwijzen naar de indeling zoals figuur 3.1 hieronder die laat zien.

1. **Training:** Verschillende modellen worden geschat op een deel van de steekproef (dit is de trainingsperiode).
2. **Cross-validatie:** de schattingen uit de trainingsperiode worden gebruikt om de doelvariabele (in dit geval de werkloosheid) te voorspellen in de cross-validatieperiode. Op basis hiervan wordt het beste model gekozen, bijvoorbeeld het model met de kleinste gemiddelde voorspelfout.
3. ***Out-of-sample* (oos) test:** Het geselecteerde model wordt vervolgens gebruikt om de doelvariabele te voorspellen in de rest van de dataset: de testperiode. In dit deel wordt de voorspelkwaliteit van het model getest.

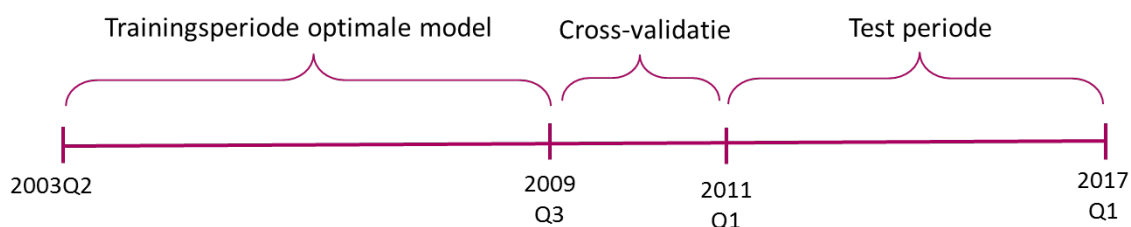
Gebruik van de steekproef

We schatten de verschillende specificaties van onze modellen op het eerste deel van de steekproef (het tweede kwartaal van 2003 tot en met kwartaal drie van 2009). Dit is onze trainingsperiode (zie figuur 3.1). Vervolgens voorspellen we de werkloosheid in de periode erna, tot vier kwartalen vooruit. De voorspelkwaliteit in deze cross-validatieperiode gebruiken we om het optimale model te selecteren.^{4,5} Met dit optimale model voorspellen we

⁴ De cross-validatieperiode (2009Q4 – 2010Q4) is een crisisperiode. Onze modellen zijn dus geoptimaliseerd op basis van verbanden die mogelijk verschillen van die in een "normale" periode. Een uitbreiding of het verschuiven van de cross-validatieperiode gaat ten koste van de lengte van de trainings- of testperiode omdat de data slechts beschikbaar zijn vanaf 2003Q2. We hebben voor de verdeling zoals beschreven in figuur 3.1 gekozen omdat we de voorspellingen van onze

de werkloosheid in de testperiode. De testperiode loopt van het eerste kwartaal 2011 tot en met het eerste kwartaal van 2017. Het begin van de testperiode is zo gekozen, omdat het CPB toen is overgegaan op het ramen van de werkloosheid volgens de internationale definitie.⁶ Zoals gezegd gebruiken we de testperiode om de voorspelkwaliteit van het optimale model te onderzoeken. Hierbij vergelijken we de verschillende voorspelmethoden met elkaar.

Figuur 3.1 Gebruik van de steekproef



Door het beperkte aantal observaties is onze cross-validatieperiode kort. Hierdoor kan *in-sample*-optimalisatie een kansrijk alternatief zijn. In dat geval kunnen we de trainingsperiode uitbreiden met vijf kwartalen; bij *in-sample*-optimalisatie is de trainingsperiode immers de gehele periode vóór de testperiode. Daarna selecteren we de specificatie met de beste *fit*. Dit kan het model zijn met het laagste Akaike of Bayesian informatie criterium (AIC of BIC). Deze twee informatie criteria proberen *overfitting* te verminderen door de score te verhogen wanneer er meer parameters worden geschat.

Beoordeling van de voorspellingsmethoden

In de testperiode gebruiken we de optimale modellen om de werkloosheid tot zeven kwartalen vooruit te voorspellen. Dit is de *out-of-sample*-toets. Telkens als er een reeks van voorspellingen is gemaakt,⁷ wordt het begin van de voorspelperiode een observatie naar voren geschoven,⁸ worden de parameters van de optimale modellen opnieuw geschat en voorspellen we weer tot (maximaal) zeven kwartalen vooruit. Zo berekenen we de voorspelfouten voor elke voorspellingstermijn. De gemiddelde voorspelfout meten we op basis van de *root mean squared error*: de wortel uit de som van de kwadraten van de afwijkingen.

verschillende methoden over een zo lang mogelijke periode *out-of-sample* willen vergelijken. In de toekomst, als er langere datareeksen beschikbaar zijn, kunnen we de cross-validatie- en testperiode aanpassen.

⁵ Bij de stromenanalyse in hoofdstuk 6 weegt ook de voorspelkwaliteit van de transitiekansen tussen de staten mee voor het bepalen van het optimale model.

⁶ Volgens de internationale definitie van de werkloosheid, die aansluit bij de richtlijnen van de International Labour Organization (ILO), wordt iedereen geteld die minstens 1 uur per week werkt of wil werken. Bij de nationale definitie ligt deze grens bij 12 uur per week. Zie ook het kader hierover in de MEV 2011 ([link](#)).

⁷ De eerste reeks voorspellingen gaat dus over de periode 2011Q1 ($t+0$) tot en met 2012Q4 ($t+7$).

⁸ Dit doen we zowel met een *recursive* als een *rolling window*. In beide gevallen voeg je er elke periode een observatie bij. Bij de *rolling window* verwijder je ook steeds de eerste observatie, zodat het totale aantal gelijk blijft.

Het probleem van *overfitting*

Het probleem van *overfitting* doet zich voor wanneer een model te veel verklarende variabelen bevat, waardoor het niet meer bruikbaar is. Een voorbeeld hiervan is het schatten van een regressiemodel met meer verklarende variabelen dan observaties. In dat geval is het simpelweg technisch onmogelijk om de coëfficiënten te schatten.

Maar ook wanneer het schatten technisch mogelijk is, kan *overfitting* slechte voorspellingen opleveren. Bij het schatten van een model kan de te verklaren variabele nooit perfect voorspeld worden. Het deel dat we niet kunnen voorspellen, gegeven de set van verklarende variabelen, noemen we ruis. Bronnen van deze ruis zijn bijvoorbeeld meetfouten, verklarende variabelen die we niet in onze dataset hebben of willekeurige onvoorspelbaarheden. Een model met te veel variabelen neemt de ruis in de trainingsperiode mee, waardoor de geschatte waarden van de parameters vaak onrealistisch groot of klein zijn. Hierdoor zal de *out-of-sample* voorspellingskwaliteit, waar de ruis anders is, niet goed zijn.

De *fit* in de trainingsperiode neemt toe met de complexiteit van het model. Echter, hoe meer variabelen er worden toegevoegd, des te meer vormt het model zich naar de steekproef. Dus, hoe ingewikkelder en flexibeler het model wordt, des te meer zal het zich vormen naar de specifieke trainingsperiode en des te meer zullen de parameters van het model veranderen als de trainingsperiode wordt vervangen door een andere steekproef. De mate waarin de parameters hierdoor veranderen noemen we ook wel de **variantie**. De variantie neemt dus toe met de flexibiliteit van het model (James et al., 2013, p. 34).

Met de **bias** van een model bedoelen we de voorspelfout die ontstaat wanneer we proberen om complexe processen te vatten met een simpel model. Een dergelijk probleem ontstaat bijvoorbeeld wanneer we proberen een lineair model te schatten, terwijl het datagenerend proces eigenlijk kwadratisch is. In tegenstelling tot de variantie neemt de bias van een model doorgaans af naarmate het flexibeler is (James et al., 2013, p. 35).

De totale verwachte voorspelfout van een model kan altijd worden verdeeld in drie componenten: de variantie, de bias en de variantie van de standaardfout:

$$E(y_0 - \hat{f}(x_0))^2 = \text{Var}(\hat{f}(x_0)) + [\text{Bias}(\hat{f}(x_0))]^2 + \text{Var}(\epsilon)$$

Hierbij is y_0 de daadwerkelijke uitkomst en $\hat{f}(x_0)$ de voorspelling van het model. $E(y_0 - \hat{f}(x_0))^2$ is dus de verwachte gekwadrateerde fout. $\text{Var}(\hat{f}(x_0))$ is de variantie van het model en $[\text{Bias}(\hat{f}(x_0))]^2$ de gekwadrateerde bias. $\text{Var}(\epsilon)$ representeert de niet te verklaren ruis: hieronder vallen alle elementen die niet modelmatig kunnen worden gevat. Dit is de minimale fout die we kunnen maken.

Bij het schatten van een voorspellingsmodel is het doel altijd het minimaliseren van de verwachte fout. We zoeken dus een model met zowel een lage variantie als bias. Aangezien bij toenemende flexibiliteit de variantie toeneemt en de bias afneemt, ontstaat hier een afruil. De technieken die we in dit hoofdstuk bespreken zijn erop gericht het optimale punt te bepalen in deze afruil. Veel *statistical-learning*-technieken gebruiken hiervoor een cross-validatieperiode, waarin het optimale model *out-of-sample* wordt geselecteerd.

Het is echter ook mogelijk om *in-sample*-statistieken en -technieken te gebruiken om het model te selecteren, zoals het model met het laagste Akaike of Bayesian informatie criterium (AIC of BIC). Deze twee informatiecriteria proberen *overfitting* te verminderen door de score te verhogen wanneer er meer parameters worden geschat. Ten slotte zijn er methoden die het optimale punt in de afruil proberen te bepalen door de parameters van het model in te krimpen (*shrinkage*), zoals bijvoorbeeld de Bayesiaanse vector-autoregressieve (BVAR) modellen in hoofdstukken 5 en 6.

4 Toepassingen van de wet van Okun

In 1962 beschreef Okun een empirische wetmatigheid: een toename van één procent van het bruto binnenlands product (bbp) ten opzichte van het potentiële bbp gaat samen met een afname van 0,5% van de werkloosheidsvoet. Hij baseerde deze regel op data van voor de jaren zestig in de VS. In onze analyse zijn we specifiek geïnteresseerd in een recente relatie voor Nederland.

We schatten een eenvoudig model dat het werkloosheidspercentage voorspelt op basis van het percentage in de periode ervoor en bbp-ontwikkelingen; zie paragraaf 4.3 voor details. Hierbij gebruiken we resultaten uit de empirische toepassing van de wet van Okun van Gordon (2010) en Ball et al. (2017).

De conclusie is dat dit type modellen niet beter voorspelt dan de conventionele CPB-instrumenten. Tot vier kwartalen vooruit zijn de voorspellingsfouten bijna even groot, maar daarna zijn de CPB-ramingen beduidend beter. Toch kan deze benadering van nut zijn, omdat er maar twee variabelen nodig zijn: het werkloosheidspercentage en het bbp. Daarnaast is het mogelijk dat de voorspelfouten van de methoden die we in dit document bestuderen tegen elkaar inwerken. Een combinatie van deze methoden kan de voorspelling van de werkloosheid dan verbeteren (zie hoofdstuk 7).

4.1 Toepassingen in de literatuur

De economische theorie maakt een negatief verband tussen economische productie en werkloosheid aannemelijk, omdat productie direct afhankelijk is van de productiefactoren kapitaal en arbeid. Hoe groter de inzet van arbeid, des te lager is de werkloosheid, en des te hoger de economische productie – gegeven dat de overige omstandigheden, zoals arbeidsproductiviteit en kapitaal, gelijk blijven. Arthur Okun vond inderdaad een sterk negatief verband: een 1% stijging in economische productie leidt gemiddeld tot een 0,5% daling van het werkloosheidspercentage.⁹

Het uitgangspunt is het werk van Gordon (2010) en Ball et al. (2017). Gordon schat het effect op verschillende maatstaven voor de werkgelegenheid met data voor de VS. Ball et al. schatten het effect van bbp op het werkloosheidspercentage met jaarlijkse data voor verschillende landen, waaronder Nederland. Onderstaand kader licht deze benaderingen in meer detail toe.

⁹ Okun schatte dit verband op basis van afwijkingen van het potentieel. Deze afwijkingen noemen we een 'gap'. De wet van Okun beschrijft dus het verband tussen de bbp- en werkloosheids gap .

De wet van Okun: schattingen van Gordon (2010) en Ball et al. (2017).

Gordon schat het effect op verschillende werkgelegenheidsmaatstaven aan de hand van de volgende regressievergelijking:

$$\Delta \tilde{x}_t = \alpha + \sum_{j=-4}^{-1} \beta_j \Delta \tilde{x}_{t+j} + \sum_{j=-4}^0 \gamma_j \Delta \tilde{y}_{t+j} + \theta \tilde{x}_{t-1} + \sum_{k=1}^6 \gamma_k D_k + \varepsilon_t$$

Hierbij staat \tilde{x}_t voor de *gap* in een werkgelegenheidsmaatstaf x in kwartaal t . De *gap* is het verschil tussen de feitelijke en potentiële waarde ($\tilde{x}_t = X_t - \bar{X}_t$); \tilde{y}_t is dan de bbp-*gap*: $\tilde{y}_t = Y_t - \bar{Y}_t$. Daarbij meet Δx_t de verandering in x_t ten opzichte van de vorige periode en staat Y voor $\log(BBP)$. Gordon neemt ook een aantal dummyvariabelen op die toelaten dat de verandering in \tilde{x}_t zich anders kan gedragen wanneer het eind van een hoogconjunctuur is bereikt (*end-of-expansion effect*).

De vergelijking verklaart de verandering in de *gap* van een werkgelegenheidsmaatstaf ($\Delta \tilde{x}_t$) uit de verandering daarvan in de vier kwartalen ervoor ($\sum_{j=-4}^{-1} \beta_j \Delta \tilde{x}_{t+j}$), de verandering in de bbp-*gap* in het huidige kwartaal en de vier ervoor ($\sum_{j=-4}^0 \gamma_j \Delta \tilde{y}_{t+j}$) en het niveau van de maatstaf in het vorige kwartaal (\tilde{x}_{t-1}). Die laatste term fungeert als een foutcorrectiemechanisme: de verwachte waarde van θ is negatief, waardoor de waarde van x neigt naar zijn potentiële waarde: $\Delta \tilde{x}_t$ is kleiner als \tilde{x}_{t-1} groot is en vice versa.

Ball et al. bestuderen specifiek de relatie tussen werkloosheid en bbp in verschillende landen, op basis van jaarlijkse data. Hierbij gebruiken zij alleen de huidige waarde van werkloosheid en bbp. Dit doen ze op twee manieren:

1. Vergelijking in niveaus:

$$\tilde{u}_t = \alpha + \beta \tilde{y}_t + \varepsilon_t$$

Hierbij is \tilde{u} de werkloosheids*gap*: het verschil tussen de feitelijke werkloosheid en de evenwichtswaarde op lange termijn.

2. Vergelijking in eerste verschillen:

$$\Delta U_t = \alpha + \beta \Delta Y_t + \varepsilon_t$$

Bij deze methode ligt de focus alleen op werkloosheid, terwijl Gordon zich juist richt op maatstaven voor de werkgelegenheid. Verder speelt informatie uit voorgaande perioden geen rol, terwijl de analyse van Gordon variabelen tot vier voorgaande kwartalen meeneemt. Bovendien houden Ball et al. geen rekening met een mogelijk *end-of-expansion effect* of een foutcorrectiemechanisme.

Omdat ze de vergelijking in eerste verschillen niet schatten in termen van *gaps*, is het niet nodig de potentiële waarde van bbp en de evenwichtswerkloosheid te bepalen. Dit betekent echter ook dat de schattingen van de parameters incorrect zijn als het potentiële bbp en/of de evenwichtswerkloosheid verandert over de tijd. Aan de andere kant kan dit model juist correcter zijn omdat het niet mogelijk is deze grootheden nauwkeurig te meten. Van tevoren kunnen we dus niet vaststellen welke van de twee methoden de meest nauwkeurige schattingen oplevert.

4.2 Empirische uitwerking

Om te werkloosheid te voorspellen schatten we een eenvoudig model, dat informatie tot maximaal vier voorgaande kwartalen meeneemt. We schatten drie verschillende modellen. Het eerste relateert het niveau van de werkloosheids*gap* aan dat van het eigen verleden en de *gap* in het bbp. Het tweede formuleert alle variabelen in eerste verschillen en de derde formulering gebruikt de eerste verschillen van de *gaps*:

1. Een model in niveau van de *gaps*:

$$\tilde{u}_t = \alpha + \sum_{j=-4}^{-1} \beta_j \tilde{u}_{t+j} + \sum_{j=-4}^0 \gamma_j \tilde{y}_{t+j} + \varepsilon_t$$

2. Een model in eerste verschillen:

$$\Delta U_t = \alpha + \sum_{j=-4}^{-1} \beta_j \Delta U_{t+j} + \sum_{j=-4}^0 \gamma_j \Delta Y_{t+j} + \varepsilon_t$$

3. Een model in eerste verschillen van de *gaps*:

$$\Delta \tilde{u}_t = \alpha + \sum_{j=-4}^{-1} \beta_j \Delta \tilde{u}_{t+j} + \sum_{j=-4}^0 \gamma_j \Delta \tilde{y}_{t+j} + \theta \tilde{u}_{t-1} + \varepsilon_t$$

Bij alle formuleringen kiezen we de modelspecificatie op basis van *backward selection*. Eerst wordt het complete model geschat. Daaruit verwijderen we de variabele met de hoogste p-waarde, waarna we het model opnieuw schatten. Dit gaat door totdat er alleen variabelen over zijn met een p-waarde onder de 0,10.

Om deze modellen te kunnen schatten, moeten we het niveau van het potentiële bbp en de evenwichtswerkloosheid bepalen. Dit doen we op twee manieren. De eerste optie is om de CPB-reeksen over te nemen. Omdat deze alleen beschikbaar zijn op jaarlijkse basis, zijn ze omgezet naar kwartaaldata¹⁰.

Een andere manier is een Hodrick-Prescott (HP) filter,¹¹ dat een tijdreeks opdeelt in een conjunctureel en een structureel deel. Het structurele deel functioneert dan als de potentiële waarde. Om dit filter toe te passen moeten we wel een zogenaamde *smoothing*-parameter (λ) kiezen. Deze bepaalt hoeveel van de variatie in de reeks wordt toebedeeld aan het conjuncturele deel. We kiezen deze parameter zo dat het verschil tussen de CPB-reeks en HP-reeks zo klein mogelijk is, zodat we in feite de potentiële reeks uit het HP-filter iken op de bijbehorende CPB-reeks. Hierbij leggen we op dat eenzelfde waarde van λ van toepassing is voor zowel het potentiële bbp als de evenwichtswerkloosheid. Uit onze schattingen blijkt dat $\lambda = 51.200$ hiermee het best overeenkomt. Zie ook figuur A.1 in appendix A.1 voor een grafische toelichting.

4.3 Regressieresultaten

De resultaten van onze schattingen staan in tabel 4.1.¹² In de eerste twee kolommen (\tilde{u}_t) staan de resultaten van het model in niveaus. Na de selectieprocedure blijft alleen de bbp-*gap* van het huidige kwartaal over (\tilde{y}_t). Dit is precies de specificatie van Okun. We vinden een

¹⁰ Hiervoor gebruiken we een *cubic spline*: dit is een interpolatie op basis van derdegraads polynomen.

¹¹ Een HP-filter is een relatief eenvoudige methode om de trend van een reeks te bepalen. Er zijn ook meer geavanceerde manieren, maar toepassing hiervan voert te ver voor dit onderzoek.

¹² We hebben ook *forward selection* uitgevoerd. Deze komt in alle gevallen uit op dezelfde modelspecificatie als de *backward selection*.

coëfficiënt van tussen de -0,33 en -0,42, afhankelijk van de manier waarop de *gap* is berekend. Dit komt overeen met de bevindingen van Ball et al.: zij vonden namelijk een coëfficiënt van -0,34 met behulp van een HP-filter over de periode 1995-2011. Dit betekent dat een stijging van de *bbp-gap* met 1%-punt gepaard gaat met een daling van de *werkloosheidsgap* met 0,3-0,4 %-punt.

Tabel 4.1 Regressieresultaten van de geselecteerde specificaties (2003 – 2017)

	Afhankelijke variabele				
	\tilde{u}_t		ΔU_t	$\Delta \tilde{u}_t$	
Regressoren	CPB <i>gap</i>	HP <i>gap</i>		CPB <i>gap</i>	HP <i>gap</i>
\tilde{y}_t	-0,424*** (0,0385)	-0,329*** (0,0418)			
ΔY_t			-0,137*** (0,0424)		
$\Delta \tilde{u}_{t-1}$				0,527*** (0,105)	0,595*** (0,0987)
$\Delta \tilde{y}_{t-1}$				-0,115*** (0,0324)	-0,109*** (0,0325)
Constante	-0,001 (0,0008)	-0,000 (0,0008)	0,001* (0,0003)	0,000 (0,0002)	-0,000 (0,0002)
N	57	57	56	52	52
R ²	0,688	0,530	0,163	0,561	0,604

*: $p < 0,1$; **: $p < 0,05$; ***: $p < 0,01$.

Ook bij het model dat is gespecificeerd in eerste verschillen blijft alleen de *bbp*-verandering van het huidige jaar overeind. In dit geval is de coëfficiënt ongeveer -0,14. De modellen in eerste verschillen van de *gaps* (laatste twee kolommen, $\Delta \tilde{u}_t$), nemen de verandering in zowel de *bbp-gap* ($\Delta \tilde{y}_{t-1}$) als de *werkloosheidsgap* ($\Delta \tilde{u}_{t-1}$) van het voorgaande kwartaal mee. De coëfficiënt van de *bbp-gap* ligt rond de -0,11.

4.4 Voorspellen van de werkloosheid

4.4.1 Data

De CPB-reeksen van het potentiële *bbp* zijn gebaseerd op CBS-tijdreeksen, inclusief alle herzieningen en revisies. Daarom kiezen we noodgedwongen voor een duale benadering: een onrealistische en een realistische. In de onrealistische voorspelling gebruiken we een complete *bbp*-reeks, inclusief alle herzieningen en revisies, om de modellen te schatten. De realistische voorspelling maakt gebruik van de *real-time* *bbp*-reeks: de reeks die op het moment van ramen beschikbaar is. Deze methode is eerlijker omdat op dat moment de gevolgen van de revisies nog niet bekend zijn. In dit realistische scenario hebben we zoals gezegd geen CPB-reeksen van het potentiële *bbp* beschikbaar, dus we gebruiken een HP-filter om deze te bepalen.

4.4.2 Resultaten van de voorspellingen

Uit tabel 4.2 blijkt dat de Okun-modellen niet beter voorspellen dan de conventionele CPB-ramingen. Tot vier kwartalen vooruit zijn de fouten bijna even groot, maar daarna zijn de

CPB-ramingen beduidend beter. Van de onrealistische ramingsmodellen is het model in niveau van de *gaps* en met de CPB-reeksen voor het potentiële bbp en de evenwichtswerkloosheid het beste: het is de specificatie met de kleinste fout in de cross-validatie. Bij de realistische ramingsmodellen is dit het model in eerste verschillen van de *gaps*. Verschillen tussen de voorspelfouten van alternatieve specificaties zijn klein, zowel in de cross-validatie als in de *out-of-sample*-test.

Ook zijn de fouten in de cross-validatie beduidend kleiner dan in de *out-of-sample*-test. Mogelijk is de cross-validatie periode te kort, of hangt de optimale modelspecificatie te sterk af van de gekozen periode. Ook blijkt de keuze van het optimale model af te hangen van de gebruikte data (realistisch versus onrealistisch). De beste onrealistische voorspelling is vanaf t+4 zelfs beter dan de optimale realistische. Mogelijke opties zijn het verlengen van de periode van de cross-validatie of het uitbreiden van het model met andere voorspellers van werkloosheid.

Tabel 4.2 Voorspelfouten van de Okun-regressies, root mean squared error (RMSE)

		Cross- validatie	Out- of -sample test							
			t+0	t+1	t+2	t+3	t+4	t+5	t+6	t+7
CPB		-	0,20	0,38	0,57	0,74	0,86	0,98	1,04	1,14
Realistische voorspellingen (Flash bbp en revisies)										
Eerste verschillen		0,060	0,21	0,42	0,64	0,88	1,11	1,33	1,51	1,70
Eerste verschillen <i>gaps</i>	HP filter	0,056	0,20	0,40	0,63	0,86	1,09	1,29	1,46	1,65
Niveau <i>gaps</i>	HP filter	0,059	0,19	0,42	0,69	0,99	1,28	1,56	1,81	2,05
Onrealistische voorspellingen										
Eerste verschillen		0,056	0,20	0,39	0,59	0,81	1,02	1,22	1,39	1,58
Eerste verschillen <i>gaps</i>	CPB	0,066	0,21	0,42	0,63	0,87	1,08	1,28	1,45	1,63
	HP filter	0,054	0,19	0,38	0,58	0,80	1,03	1,24	1,41	1,59
Niveau <i>gaps</i>	CPB	0,038	0,19	0,38	0,61	0,82	1,01	1,19	1,33	1,44
	HP filter	0,059	0,19	0,42	0,69	0,99	1,28	1,56	1,81	2,05

De modelspecificaties staan in tabel A.1 in appendix A.1. Er is bij de Okun-voorspelmodellen telkens een *recursive window* gebruikt. Het CPB publiceert alleen jaarcijfers, de voorspelfouten van het CPB zijn gebaseerd op interne kwartaalvoorspellingen.

5 Tijdreeksmodellen

In dit hoofdstuk behandelen we voorspelmethodes die alleen gebaseerd zijn op de correlatie van de werkloosheid met zijn eigen verleden, of op de correlaties tussen *leading* indicatoren en de werkloosheid over de tijd. Dit zijn de tijdreeksbenaderingen. De modellen die wij hier gebruiken, bevatten weinig of geen economische theorie, maar toch leveren sommige van deze modellen de beste voorspelresultaten op.

Het eerste onderzochte model is een eenvoudig benchmarkmodel, dat alleen het verleden van de werkloosheid gebruikt om deze te voorspellen. Dit is een univariaat tijdreeksmodel.

Als andere modellen niet beter voorspellen dan dit benchmarkmodel, is de conclusie dat onze analyses geen beter voorspelmodel voor de werkloosheid hebben opgeleverd.

Het tweede soort tijdreeksmodellen beschrijft de werkloosheid en andere *leading* indicatoren samen. Dit zijn vector-autoregressieve modellen, ofwel VAR-modellen. VAR-modellen die een Bayesiaanse schattingsmethode gebruiken (BVAR) zijn het meest succesvol. Sectie 5.1 bespreekt het eenvoudige benchmarkmodel; in 5.2 gaan we in op VAR- en BVAR-modellen.

5.1 Benchmarkmodel

Een eenvoudige voorspelling op basis van de correlatie van de werkloosheid met zichzelf over de tijd is een autoregressief model, ofwel een AR-model. Als de variabele x_t wordt gemodelleerd als een functie van de laatste p realisaties, spreken we van een AR(p)-model:

$$x_t = a + \sum_{j=1}^p \phi_j x_{t-j} + \varepsilon_t \quad (1)$$

Hier zijn de ϕ_j de te schatten parameters (samen met de constante a en de variantie σ_ε^2 van de storingsterm ε_t .) De eenvoudigste versie is een AR(1)-model. Dit gebruiken we als benchmark.

Een AR(1)-model is een van de meest eenvoudige modellen met maar drie parameters. Dit model heeft daarom geen last van *overfitting*. Zoals uitgelegd in hoofdstuk 3, doet dit probleem zich voor bij het voorspellen met complexere modellen met veel geschatte parameters.

5.2 VAR- en BVAR-modellen

VAR-modellen zijn complexere modellen met veel parameters, waardoor ze al snel last hebben van *overfitting*. Een type model uit de literatuur dat succesvol is gebleken om te voorspellen is het Bayesiaanse VAR-model. Dit model gebruikt extra informatie (zogenaamde “priors”) om het probleem van *overfitting* tegen te gaan. Wij bespreken VAR- en BVAR modellen in deze paragraaf.

5.2.1 VAR-modellen

Een VAR-model lijkt veel op een AR-model, het verschil is dat een VAR-model meerdere reeksen modelleert. De vergelijking voor het VAR-model is:

$$\vec{x}_t = \vec{a} + \sum_{j=1}^p \Phi_j \vec{x}_{t-j} + \vec{\varepsilon}_t \quad (2)$$

Hier bestaat de vector \vec{x}_t uit n reeksen. In ons geval is $\vec{x}_t = (u_t, \vec{z}_t)$, waar de vector \vec{z}_t $n-1$ *leading* indicatoren bevat en u_t het werkloosheidspercentage is. De parameters Φ_j zijn matrices met parameters die de autocorrelaties en de correlaties tussen de reeksen in \vec{x}_t en de vertragingen \vec{x}_{t-j} bepalen. De storingsterm $\vec{\varepsilon}_t$ en constante term a zijn ook vectoren.

Het voordeel van een VAR-model ten opzichte van een AR-model is dat we in potentie de informatie in de *leading* indicatoren kunnen gebruiken om het verloop van het werkloosheidspercentage te verklaren, en daardoor ook te voorspellen. Het nadeel is dat een VAR-model al gauw veel parameters heeft: elke van de p parametermatrices Φ_j heeft $(n^2 + n)/2$ parameters, terwijl de covariantiematrix van de storingsterm ook nog $(n^2 + n)/2$ parameters heeft. Behalve bij VAR-modellen van lage orde (met $p=1$ of 2) en met slechts een paar variabelen (kleine n), bevat een VAR-model dus al snel weinig waarnemingen per geschatte parameter. Hierdoor hebben VAR-modellen van enige omvang snel last van *overfitting*, waardoor ze slechte voorspellingen produceren.

We beschikken over ongeveer twintig mogelijke *leading* indicatoren. Het blijkt echter dat zelfs een VAR(1)-model met maar zes reeksen niet goed kan voorspellen. Vanwege deze slechte resultaten zien we af van het presenteren van de uitkomsten.

5.2.2 BVAR-modellen

Uit de literatuur blijkt dat een pure tijdreeksbenadering die het werkloosheidspercentage rechtstreeks voorspelt via een BVAR-model goede uitkomsten oplevert. Een BVAR gebruikt Bayesiaanse schattingstechnieken en vereist de specificatie van een prior. Deze prior bevat extra informatie (die niet in de steekproef zit) over de onbekende parameters. Het is de laatste tijd gebruikelijk om de prior te gebruiken om *overfitting* van grote VAR-modellen tegen te gaan. Door de juiste soort prior hoeft een BVAR-model geen gebruik te maken van een *out-of-sample*-modelselectieprocedure.

Er zijn verschillende priors die voor dit doel worden ingezet. De Minnesota-prior stuurt de geschatte parameterwaarden van het VAR-model terug in de richting van een univariaat AR(1)-model met $\phi_1 = 1$ voor elke reeks in het model.¹³ Alleen de aspecten van het VAR-model die voor elke reeks afwijken van dit eenvoudige AR(1)-model komen tot uiting in de schattingen ("posteriors"), mits de data die aspecten sterk ondersteunen. Uit de literatuur blijkt dat deze BVAR-modellen weinig last hebben van *overfitting* en daardoor goed kunnen voorspellen. We hebben ook geëxperimenteerd met een vergelijkbare prior die veronderstelt dat alle reeksen in het model univariate AR(1)-processen zijn met een AR-coëfficiënt $\phi=0.8$. Als deze prior wordt gebruikt, mag het BVAR-model dus alleen stationaire reeksen bevatten.

¹³ Zo'n model wordt ook wel een *random-walk*-model genoemd. In de Minnesota-prior kunnen de elementen van de storingsterm nog wel met elkaar gecorreleerd zijn. Hierdoor wijkt deze prior wel enigszins af van een univariaat *random-walk*-model.

Ten slotte gebruiken we ook de zogenaamde *combination*-prior (Sims, 1998). Deze prior combineert drie verschillende mogelijkheden. De eerste component volgt de Minnesota-prior. De tweede prior veronderstelt dat er geen cointegratie tussen de reeksen in het model is. De derde prior heet de *single-unit-root*-prior en laat de mogelijkheid van cointegratie in het model wel toe. De *combination*-prior is hierdoor consistent met een aantal verschillende typen eenvoudige VAR-modellen. In onze studie blijkt deze prior iets betere voorspellingen te genereren dan de andere twee afzonderlijke priors.

Zowel de Minnesota- als de *combination*-prior kan worden gecombineerd met niet-stationaire reeksen. Omdat we in onze studie ook niet-stationaire *leading* indicatoren meenemen, sluiten deze priors iets beter aan bij onze analyse.

5.3 Data

We hebben verschillende *leading* indicatoren gebruikt om de BVAR-modellen te schatten. De datareeksen staan in tabel 5.1. In de analyses gebruiken we zogenaamde *quasi-real-time*-data. Voor de meeste reeksen zijn namelijk geen *real-time*-data beschikbaar. Dit zijn data die beschikbaar waren ten tijde van de raming en niet herzien zijn. In plaats daarvan gebruiken we de laatst beschikbare reeksen, waar tussentijdse herzieningen wel in zitten. Door bij de voorspellingen alleen de reeksen tot aan het kwartaal vóór de raming te gebruiken, houden we wel zo goed mogelijk rekening met de timing van de beschikbaarheid van de data (vandaar de term *quasi-real-time*). Dit betekent echter dat de vergelijking met de CPB-voorspelfouten niet helemaal eerlijk is. Maar doordat het aantal herzieningen in de testperiode beperkt is, heeft dit niet veel invloed op de resultaten.¹⁴

¹⁴ Elbourne et al. (2015) hebben onderzocht wat het effect is van de CBS-bijstellingen in cijfers voor het bbp op de ramingsfouten van het CPB. Zij concluderen dat deze bijstellingen niet de belangrijkste bron van ramingsonzekerheid zijn.

Tabel 5.1 *Leading indicatoren in BVAR-modellen*

Variabelen	Bron	
Stroomkansen tussen de drie staten (zie hoofdstuk 6) <i>E</i> (werkend), <i>U</i> (werkloos) en <i>N</i> (niet-participerend)	CBS	$U \rightarrow E, U \rightarrow N, E \rightarrow U, E \rightarrow N, N \rightarrow U, N \rightarrow E$
$\Delta \ln(\text{ontslag_aan}), \ln(\text{ontslag_aan})$	UWV	Ontslagaanvragen
$\Delta \ln(\text{faill}), \ln(\text{faill})$	CBS	Uitgesproken faillissementen
u_lang	CBS	% langdurige werkloosheid
$\Delta \ln(\text{instr}), \ln(\text{instr})$	UWV	Totale instroom ww
$\Delta \ln(\text{ww}), \ln(\text{ww})$	UWV	Niveau ww
uitstroomkans_ww_limiet $\ln(\text{uitstroom_ww_limiet})$	UWV	Uitstroomkans ww (limiet) Niveau uitstroom ww (limiet)
uitstroomkans_ww_werk $\ln(\text{uitstroom_ww_werk})$	UWV	Uitstroomkans ww (werkhervatting) Niveau uitstroom ww (werkhervatting)
$\Delta \ln(\text{vac}), \ln(\text{vac})$	CBS	Vacatures
$\Delta \ln(\text{vac_ont}), \ln(\text{vac_ont})$	CBS	Ontstane vacatures
cons_vertrouw	CBS	Consumentenvertrouwen
$\Delta \ln(\text{bbp}), \ln(\text{bbp})$	CBS	Reëel bbp
uitzend , uitzend	CBS	Index uren uitzendsector

Alle reeksen zijn seizoensgecorrigeerd. De stationaire reeksen zijn dikgedrukt.

5.4 Resultaten

Tabel 5.2 laat uitkomsten van BVAR-modellen zien op basis van de verschillende priors en verschillende *leading* indicatoren.

Tabel 5.2 *Voorspelfouten BVAR modellen*

RMSE	t+0	t+1	t+2	t+3	t+4	t+5	t+6	t+7
CPB	0,20	0,38	0,57	0,74	0,86	0,98	1,04	1,14
AR(1)	0,28	0,54	0,81	1,07	1,33	1,59	1,86	2,13
BVAR								
<i>Combination-prior</i>	0.17	0.33	0.49	0.64	0.77	0.93	1.10	1.27
<i>Combination+flows</i>	0.17	0.33	0.50	0.65	0.79	0.96	1.14	1.32
Stationaire Minnesota-prior	0.18	0.34	0.51	0.67	0.82	1.03	1.24	1.47
<i>Combination-prior, nsdata</i>	0.17	0.34	0.51	0.67	0.84	1.02	1.20	1.39
Minnesota-prior, nsdata	0.17	0.35	0.57	0.80	1.03	1.31	1.60	1.91

Het CPB publiceert alleen jaarcijfers, de voorspelfouten van het CPB zijn gebaseerd op interne kwartaalvoorspellingen.

Onder de kop 'BVAR' staan de uitkomsten op basis van voorspellingen uit de directe BVAR-modellen. In de rij '*Combination-prior*' staan de resultaten van de BVAR op basis van de *combination-prior*. Deze gebruikt stationaire reeksen, maar niet de data van de stromen op arbeidsmarkt. In de rij '*Combination+flows*' is te zien hoe goed de BVAR voorspelt als de stromendata wel worden meegenomen in de BVAR. Daaronder, in de rij 'Stationaire Minnesota-prior', staan de uitkomsten op basis van de prior die veronderstelt dat alle reeksen stationaire univariate AR(1)-processen zijn. De laatste twee rijen in de tabel geven aan hoe de BVAR-modellen met *combination-* en Minnesota-prior voorspellen als we de niet-stationaire *leading* indicatoren gebruiken.

De BVAR met de *combination*-prior doet het beter dan de BVAR met de Minnesota-prior. Ook blijkt dat de BVAR-modellen met niet-stationaire data minder goed voorspellen dan de stationaire varianten. In het volgende hoofdstuk zullen we zien dat de beste BVAR-uitkomsten voorspelfouten opleveren die vergelijkbaar zijn met die van het twee-staten-stromenmodel met BVAR op basis van de CBS-duurdata. Zoals blijkt uit hoofdstuk 7, is de samengestelde voorspelling van de drie best presterende modellen uiteindelijk de meest nauwkeurige.

6 Derde alternatief: stromenanalyse

Dit hoofdstuk richt zich op het kwantificeren van stromen op de arbeidsmarkt. Deze markt is namelijk voortdurend in beweging: mensen worden werkloos, besluiten om toe te treden tot de arbeidsmarkt, vinden een (nieuwe) baan, of trekken zich terug. De werkloosheid is het resultaat van deze bewegingen. Het hoofdstuk richt zich op het toelichten van de methode en het bespreken van de uitkomsten.

De analyse leidt tot een specificatie die de werkloosheid tot vier kwartalen vooruit systematisch beter voorspelt dan de gebruikelijke CPB-ramingen. Het succes hangt af van de tijdige beschikbaarheid van geschikte data.

6.1 Methodiek in het kort

Barnichon heeft met verschillende coauteurs laten zien dat de werkloosheid goed is te voorspellen met behulp van stroomgegevens; zie bijvoorbeeld Barnichon en Nekarda (2012) en Barnichon en Garda (2016). Deze analyses maken onderscheid tussen twee benaderingen. De eerste gaat uit van twee staten op de arbeidsmarkt: mensen zijn óf werkzaam óf werkloos. De tweede benadering richt zich op drie staten en dus ook drie verschillende groepen: de werkloze beroepsbevolking (U), de werkzame beroepsbevolking (E) en de niet-beroepsbevolking (N). De figuur hieronder brengt deze tweede benadering in beeld.

Figuur 6.1 Een stroommodel met 3 staten



De omvang van stromen tussen de verschillende staten kan gebruikt worden om de werkloosheid te voorspellen. Het idee is als volgt. De werkloosheid is vergelijkbaar met de hoeveelheid water in een badkuip. Er stroomt voortdurend water in en uit. Het water dat eruit stroomt zijn mensen die werk vinden (de stroom van U naar E) of mensen die niet meer actief op zoek zijn naar werk (de stroom van U naar N). Het water dat in het bad stroomt, zijn de mensen die werkloos worden, dat kan vanuit werk (de stroom van E naar U) of vanuit niet-participatie (de stroom van N naar U). Als je goed kunt voorspellen hoeveel water er in en uit de badkuip stroomt (de stromen in en uit de werkloosheid), dan kan je daarmee ook goed voorspellen wat er, gegeven de beginstand, gebeurt met de hoeveelheid water in de badkuip (de werkloosheid). In dit hoofdstuk doen we analyses waar we zowel twee staten (U en E) als drie staten (U , E en N) onderscheiden. Het kader hieronder laat wiskundig zien hoe het mogelijk is de werkloosheid te voorspellen met behulp van stromen in een tweestatenmodel. In appendix A.2 leiden we dit af voor het driestatenmodel.

6.1.1 Empirische aanpak

De methode draait om het modelleren van de overgangen tussen de verschillende staten. De overgangskans is de kans dat iemand van toestand wisselt. We gebruiken drie verschillende manieren om het tijdspad hiervan te beschrijven. De zogenaamde Bayesiaanse VAR (BVAR) levert de beste voorspellingen op.

1. De eerste methode houdt de overgangskansen constant over de voorspelperiode. In deze methode kiezen we hiervoor de laatstbekende waarde.
2. In het vector-autoregressieve model (VAR) hangen de overgangskansen af van het eigen verleden en van andere economische variabelen (voor details zie appendix A.3). Dit is de methode gebruikt door Barnichon en Nekarda (2012) en Barnichon en Garda (2016).
3. Het Bayesiaanse VAR-model (BVAR) maakt ook gebruik van een VAR-model om de overgangskansen te voorspellen, alleen gebruikt het een Bayesiaanse schattingsmethode om het model te schatten (zie ook hoofdstuk 5).

Het tweestatenmodel van Barnichon en Garda (2016)

Voor de analyse van twee staten volgen we zoveel mogelijk Barnichon en Garda (2016).^a In het tweestatenmodel kunnen individuen alleen werkzaam of werkloos zijn. Er zijn dus ook slechts twee stromen: van werkzaam naar werkloos en van werkloos naar werkzaam. De werkloosheid verandert dan als volgt:

$$u_{t+1} - u_t = s_{t+1}(1 - u_t) - f_{t+1}u_t$$

Hier is u_t u_t het werkloosheidspercentage op moment t , waar t de periode aangeeft (een kwartaal).

Tussen t en $t + 1$ vinden werklozen een baan volgens een Poissonproces met (constante) verwachtingswaarde f_{t+1} en verliezen werknemers hun baan volgens een Poissonproces met (constante) verwachtingswaarde s_{t+1} .

Oplossen naar u_{t+1} geeft de volgende vergelijking voor het transitiepad, de zogenaamde *law of motion*:

$$u_{t+1} = \beta_{t+1}u_{t+1}^* + [1 - \beta_{t+1}]u_t .$$

Hier is u^* de *steady-state*-werkloosheidsvoet die gelijk is aan:

$$u_{t+1}^* \equiv \frac{s_{t+1}}{s_{t+1} + f_{t+1}}$$

en de convergentiesnelheid (β) naar die evenwichtstoestand is:

$$\beta_{t+1} \equiv 1 - e^{-(s_{t+1} + f_{t+1})} .$$

De *law of motion* laat zien dat het werkloosheidspercentage van volgend kwartaal een gewogen gemiddelde is van het huidige percentage en het *steady-state*-werkloosheidspercentage van het volgende kwartaal. Hierbij bepalen de stromen op de arbeidsmarkt zowel het *steady-state*-percentage als het gewicht hiervan (de convergentiesnelheid β). Het werkloosheidspercentage kan dus worden voorspeld door de stroomkansen te voorspellen en deze vervolgens te voeden aan de *law of motion*.

Door de *law of motion* iteratief door te rekenen, kan de werkloosheid ook recursief meerdere (j) kwartalen vooruit worden voorspeld:

$$\hat{u}_{t+j|t} = \hat{\beta}_{t+j|t}\hat{u}_{t+j|t}^* + (1 - \hat{\beta}_{t+j|t})\hat{u}_{t+j-1|t}$$

met

$$\hat{u}_{t+j|t}^* = \frac{\hat{s}_{t+j|t}}{\hat{s}_{t+j|t} + \hat{f}_{t+j|t}}$$

en

$$\hat{\beta}_{t+j|t} = 1 - e^{-(\hat{s}_{t+j|t} + \hat{f}_{t+j|t})} ,$$

waar $\hat{f}_{t+j|t}$ en $\hat{s}_{t+j|t}$ voorspellingen van de stroomkansen j perioden vooruit zijn. Dan is de voorspelling van de werkloosheid in $t + j$ dus een gewogen gemiddelde van de voorspelde *steady-state*-werkloosheid in $t + j$ en de voorspelling van de werkloosheid in $(t + j - 1)$ (of de gerealiseerde werkloosheid als $j=1$).

(a) De vergelijkingen in dit kader zijn een discrete benadering van de differentiaalvergelijkingen.

6.2 Data

Het doel van de stromenanalyse is om, per kwartaal, het werkloosheidspercentage volgens de internationale definitie te voorspellen. Hiervoor gebruiken we seizoensgecorrigeerde kwartaaldata.¹⁵ Data over de internationale werkloosheid zijn afkomstig van het CBS, net als de stromendata, nodig voor de constructie van de transitiekansen in het driestatenmodel. De twee transitiekansen s en f uit het tweestatenmodel moeten echter geconstrueerd worden met behulp van werkloosheidsduurdata van Eurostat of het CBS. Zie het kader op de volgende pagina voor details over de precieze aanpak bij de verschillende varianten.

Een belangrijk nadeel van de duurdata van Eurostat is het feit dat deze meerdere reeksbreuken¹⁶ bevatten. Dit heeft zijn weerslag op de betrouwbaarheid van de uitkomsten van het tweestatenmodel op basis van deze data. De CBS-duurdata zijn echter veel minder gedetailleerd, met als belangrijk bijkomend nadeel dat de geconstrueerde transitiekansen voor een beperktere periode beschikbaar zijn.

De transitiekansen kunnen in de VAR en BVAR worden voorspeld aan de hand van diverse *leading* indicatoren, zoals het aantal faillissementen, het aantal vacatures en het bruto binnenlands product. In tabel 6.1 staan alle onderzochte voorspelvariabelen met hun bron.¹⁷ Net als bij de directe BVAR gebruiken we ook hier *quasi-real-time-data* (zie sectie 5.3).

Tabel 6.1 Voorspelvariabelen met bron

Variabelen	Bron	
$\ln(\text{ontslag_aan}), \Delta \ln(\text{ontslag_aan})$	UWV	Ontslagaanvragen
$\ln(\text{faill}), \Delta \ln(\text{faill})$	CBS	Uitgesproken faillissementen
$u_lang, \Delta u_lang$	CBS	% langdurige werkloosheid
$\ln(\text{instr}), \Delta \ln(\text{instr})$	UWV	Totale instroom ww
$\text{instroomkans_ww}, \Delta \text{instroomkans_ww}$	CBS, UWV	Instroomkans ww
$\text{uitstroomkans_ww_limiet}, \Delta \text{uitstroomkans_ww_limiet}$	UWV	Uitstroomkans ww (limiet)
$\text{uitstroomkans_ww_werk}, \Delta \text{uitstroomkans_ww_werk}$	UWV	Uitstroomkans ww (werkhervatting)
$\ln(\text{vac}), \Delta \ln(\text{vac})$	CBS	Vacatures
$\ln(\text{vac_ont}), \Delta \ln(\text{vac_ont})$	CBS	Ontstane vacatures
$u, \Delta u$	CBS	% werkloosheid
$\ln(\text{bbp}), \Delta \ln(\text{bbp})$	CBS	Reëel bbp
$\text{uitzend}, \Delta \text{uitzend}$	CBS	Index uren uitzendsector

Ten behoeve van de VAR zijn de variabelen gegroepeerd in zes categorieën (zie appendix A.3).

¹⁵ Voor alle variabelen gebruiken we seizoensgecorrigeerde reeksen. Voor variabelen waarvoor deze niet beschikbaar zijn, zijn de reeksen (waar nodig) gecorrigeerd met behulp van de Census X-13-methode. Reeksen die alleen op maandbasis beschikbaar zijn, zijn omgezet naar kwartaalreeksen (door de som of het gemiddelde te nemen).

¹⁶ In de onderzochte periode komen vier reeksbreuken voor, te weten in 2005Q1, 2010Q1, 2010Q2 en 2011Q1.

¹⁷ Dit komt grotendeels overeen met de gebruikte data bij de directe BVAR in hoofdstuk 5.

Constructie transitiekansen

Tweestatenmodel

Voor het construeren van transitiekansen op basis van duurdata gebruiken we de methode van Elsby et al. (2013). Werkloosheidsduurdata zijn beschikbaar via Eurostat en het CBS. Een belangrijk verschil tussen de twee databronnen zit in het detailniveau van de verschillende duurcategorieën. Dit heeft ook gevolgen voor de precieze uitvoering van de Elsby-methode.

Eurostat

De Eurostat-data bevatten (onder meer) gegevens over het aantal werklozen met een werkloosheidsduur van minder dan één kwartaal, $U^{<1}$. Hiermee kunnen dus transities tussen kwartalen worden opgepikt. Allereerst wordt per kwartaal de uitstroom-*hazard-ratio* (f) berekend op basis van informatie over $U^{<1}$ en het totaal aantal werklozen, U .^a De kans dat een werkloze tussen kwartaal t en $t + 1$ de werkloze staat verlaat is gelijk aan:

$$F_{t+1} = 1 - \frac{U_{t+1} - U_{t+1}^{<1}}{U_t}$$

Hieruit volgt de *hazard-ratio* $f_{t+1} = -\ln(1 - F_{t+1})$. De instroom-*hazard-ratio* is dan gelijk aan de waarde van s_{t+1} die de onderstaande vergelijking oplost:

$$U_{t+1} = \frac{[1 - e^{-(f_{t+1} + s_{t+1})}]s_{t+1}}{f_{t+1} + s_{t+1}}(U_t + E_t) + e^{-(f_{t+1} + s_{t+1})}U_t$$

waar E_t het aantal werkzame personen op tijdstip t is (afkomstig van Eurostat).

CBS

De CBS-duurdata zijn minder gedetailleerd, waardoor informatie over het aantal werklozen met een werkloosheidsduur van minder dan één jaar (vier kwartalen), $U^{<4}$, moet worden gebruikt. De *hazard-ratio* f tussen kwartaal t en $t + 1$ ziet er in dat geval als volgt uit:

$$f_{t+1} = -\frac{1}{4}\ln\left(\frac{U_{t+1} - U_{t+1}^{<4}}{U_{t-3}}\right)$$

De uitstroom-*hazard-ratio* op kwartaalbasis is dus eigenlijk gebaseerd op jaarlijkse transities (in bovenstaande vergelijking is dat tussen $t - 3$ en $t + 1$). s_{t+1} kan vervolgens weer verkregen worden door de bovenstaande vergelijking voor U_{t+1} op te lossen.^b

Driestatenmodel

De zes transitiekansen uit het driestatenmodel kunnen eenvoudig direct worden berekend op basis van stromendata van het CBS. De stroomkans van E naar U is bijvoorbeeld gedefinieerd als $\lambda_{t+1}^{EU} = \frac{N_{t+1}^{EU}}{E_t}$, waar N_{t+1}^{EU} de stroom vanuit werkzaam naar werkloos tussen kwartaal t en $t + 1$ weergeeft en E_t het aantal personen in de werkzame beroepsbevolking in periode t weergeeft.

(a) U is hier gedefinieerd als de som over alle duurcategorieën (inclusief de categorie 'onbekend') van Eurostat.

(b) Met dien verstande dat data over E (net als die over U) hier afkomstig zijn van het CBS.

Pas vanaf het tweede kwartaal van 2003 zijn alle data beschikbaar (bij gebruik van duurdata van het CBS is dit zelfs 2004Q1). Dit staat in schril contrast met de data die Barnichon en Garda (2016) tot hun beschikking hebben voor andere landen; die gaan soms tientallen jaren verder terug. In vergelijking hiermee zijn er dus relatief weinig observaties beschikbaar. De dataset loopt door tot en met het eerste kwartaal van 2017.

Gegevens over de gebruikte variabelen komen vertraagd beschikbaar. Voor de meeste variabelen bedraagt de publicatievertraging maximaal 1,5 maand. Uitgaande van de

ramingsmomenten van het CPB, ligt het dus voor de hand dat op het moment van voorspellen de data van één kwartaal geleden ($t - 1$) bekend zijn. Bij de raming van september 2017 (MEV 2018), loopt de beschikbare dataset dan bijvoorbeeld tot en met 2017Q2. De duurdata van Eurostat zijn op het moment van voorspellen echter pas met twee kwartalen vertraging beschikbaar. Bij het tweestatenmodel op basis van Eurostat-duurdata moeten we dus rekening houden met het feit dat variabelen verschillende publicatievertragingen hebben. Dit doen we door s_{t-1} en f_{t-1} steeds te vervangen door de op dat moment laatst bekende waarden (respectievelijk s_{t-2} en f_{t-2}).¹⁸ Bij de BVAR worden de ontbrekende f - en s -waarden voorspeld met een Kalmanfilter.

6.3 Resultaten

Zowel bij het twee- als driestatenmodel leveren de varianten waarin de transitiekansen worden voorspeld met een BVAR de beste resultaten op. Tot twee kwartalen vooruit presteert de BVAR met CBS-duurdata (twee staten) het beste, daarna is dat de BVAR op stromendata (drie staten).

Tabel 6.2 vergelijkt de gemiddelde kwadratische voorspelfout (RMSE) van verschillende voorspelmodellen met de fouten in de ramingen van het CPB. Zoals blijkt uit de tabel, hebben de varianten met constante transitiekansen (dit zijn de \bar{u}^* -modellen) over het algemeen een grotere gemiddelde voorspelfout dan het CPB. Een uitzondering is de variant waar we gebruik maken van de duurdata van het CBS. Deze presteert tot drie kwartalen vooruit beter dan het CPB.

De twee-staten-VAR-modellen zijn alleen geschat op duurdata van Eurostat.¹⁹ De tabel toont voor deze modellen uitkomsten voor twee afzonderlijke situaties. In het eerste geval zijn de data met twee kwartalen vertraging beschikbaar. Dit noemen we ook wel de ‘*eerlijke*’ variant. In het alternatief gaan we er vanuit dat de data, net als de andere variabelen, beschikbaar zijn met één kwartaal vertraging; dit is dan de ‘*oneerlijke*’ variant. Dit onderscheid heeft gevolgen voor de ramingen: de *oneerlijke* projectie presteert doorgaans beter dan de *eerlijke*.

¹⁸ Het ‘voorspellen’ op basis van de laatst bekende waarde komt neer op het aannemen van een *random walk*.

¹⁹ Door de beperkte beschikbaarheid van de CBS-duurdata kregen we (waarschijnlijk door *overfitting*) grote voorspelfouten bij de optimale VAR-modellen met CBS-duurdata. De resultaten zijn op aanvraag beschikbaar.

Tabel 6.2 Resultaten stromenmodellen

RMSE	t+0	t+1	t+2	t+3	t+4	t+5	t+6	t+7
CPB	0,20	0,38	0,57	0,74	0,86	0,98	1,04	1,14
Constante transitiekansen								
\bar{u}^* (twee staten)	0,22	0,46	0,71	0,93	1,12	1,30	1,49	1,68
\bar{u}^* (twee staten, CBS-data)	0,16	0,31	0,51	0,72	0,95	1,14	1,32	1,53
\bar{u}^* (drie staten)	0,21	0,42	0,64	0,86	1,08	1,27	1,44	1,62
Duurdata (twee staten)								
<i>Out-of-sample</i> VAR	0,27	0,41	0,63	0,80	0,95	1,19	1,47	1,75
<i>Out-of-sample</i> VAR (zonder vertraging)	0,23	0,34	0,50	0,65	0,79	0,92	1,15	1,43
<i>In-sample</i> VAR	0,29	0,44	0,68	0,84	0,98	1,20	1,45	1,67
<i>In-sample</i> VAR (zonder vertraging)	0,25	0,39	0,61	0,83	1,00	1,23	1,45	1,68
BVAR	0,22	0,43	0,65	0,77	0,89	1,01	1,16	1,29
BVAR (zonder vertraging)	0,18	0,31	0,49	0,67	0,79	0,92	1,04	1,20
BVAR (CBS-data)	0,16	0,29	0,47	0,68	0,87	1,06	1,25	1,44
Stromen (drie staten)								
<i>Out-of-sample</i> VAR	0,19	0,35	0,56	0,82	1,08	1,39	1,71	1,99
<i>In-sample</i> VAR	0,20	0,37	0,57	0,78	0,96	1,16	1,37	1,56
BVAR	0,17	0,32	0,48	0,64	0,80	0,98	1,16	1,33

Alle BVAR-varianten in combinatie met stromenmodellen maken gebruik van de *combination*-prior. Het CPB publiceert alleen jaarcijfers, de voorspelfouten van het CPB zijn gebaseerd op interne kwartaalvoorspellingen.

De *out-of-sample*-methode presteert gemiddeld iets beter dan de *in-sample*-methode. Maar de ‘eerlijke’ varianten, waar we rekening houden met de vertraging van twee kwartalen bij de duurdata, leiden niet tot betere voorspellingen dan het CPB. De ‘oneerlijke’ *out-of-sample*-variant presteert voor de kwartalen $t + 1$ tot en met $t + 5$ wel beter dan het CPB. Dit betekent dat we met de methode van Barnichon en Garda (2016) de werkloosheidsvoorspelling voor de korte termijn zouden kunnen verbeteren als de Eurostat-duurdata met een in plaats van twee kwartalen vertraging beschikbaar zouden zijn.

Op het moment dat we een BVAR gebruiken om de stroomkansen in het tweestatenmodel te voorspellen, verbeteren de voorspellingen. De BVAR die gebruik maakt van de CBS-duurdata doet het tot en met drie kwartalen vooruit beter dan het CPB en voorspelt ook wat beter dan de \bar{u}^* -variant met dezelfde data.

Ook bij het driestatenmodel presteert de variant waar de BVAR wordt gebruikt om de transitiekansen te voorspellen het beste. Deze methode presteert zelfs tot vier kwartalen vooruit beter dan het CPB.

En nu?

In dit hoofdstuk en de vorige hebben we diverse modellen beoordeeld op hun voorspelkwaliteit. Op basis daarvan is steeds een beste specificatie gekozen. In het volgende hoofdstuk onderzoeken we of we nog een extra slag kunnen maken door die verschillende voorspelmodellen handig te combineren.

7 Combinatievoorspelling

Een combinatie van verschillende ramingen kan een beter resultaat opleveren dan de beste individuele voorspelling. Dit geldt zeker als voorspelfouten van de verschillende modellen voor eenzelfde kwartaal niet allemaal hetzelfde teken hebben. We combineren onze methodes door simpelweg het gemiddelde te nemen van de best presterende modellen van de verschillende methoden. Dit zijn er vier:

1. Het Okun model met als te verklaren variabele het eerste verschil in de werkloosheids*gap*, waarbij de evenwichtswaarden zijn bepaald met een HP-filter.
2. De BVAR die de werkloosheid direct schat, zonder eerst de stroomkansen te berekenen.
3. De BVAR-variant van het tweestatenmodel, geschat met duurdata van het CBS.
4. De BVAR-variant van het driestatenmodel.

We hebben ook een gemiddelde voorspelling berekend exclusief het Okun-model, omdat deze specificatie duidelijk de slechtste voorspellingen oplevert. Om het probleem van *overfitting* te voorkomen hebben we verder geen andere gemiddelden berekend. Tabel 7.1 zet de resultaten van de vier best presterende modellen en de twee combinatiemodellen op een rijtje en vergelijkt deze met de CPB-ramingen.

Tabel 7.1 Resultaten beste presterende modellen

RMSE	t+0	t+1	t+2	t+3	t+4	t+5	t+6	t+7
CPB	0,20	0,38	0,57	0,74	0,86	0,98	1,04	1,14
Okun ($\Delta gaps$, HP filter)	0,20	0,40	0,63	0,86	1,09	1,29	1,46	1,65
BVAR direct (<i>combination-prior</i>)	0,17	0,33	0,49	0,64	0,77	0,93	1,10	1,27
BVAR twee staten (<i>CBS-duurdata</i>)	0,16	0,29	0,47	0,68	0,87	1,06	1,25	1,44
BVAR drie staten (<i>stromendata</i>)	0,17	0,32	0,48	0,64	0,80	0,98	1,16	1,33
Combinatie	0,16	0,31	0,48	0,65	0,80	0,96	1,13	1,29
Combinatie (excl. Okun)	0,15	0,29	0,45	0,62	0,78	0,95	1,12	1,29

Het CPB publiceert alleen jaarcijfers, de voorspelfouten van het CPB zijn gebaseerd op interne kwartaalvoorspellingen.

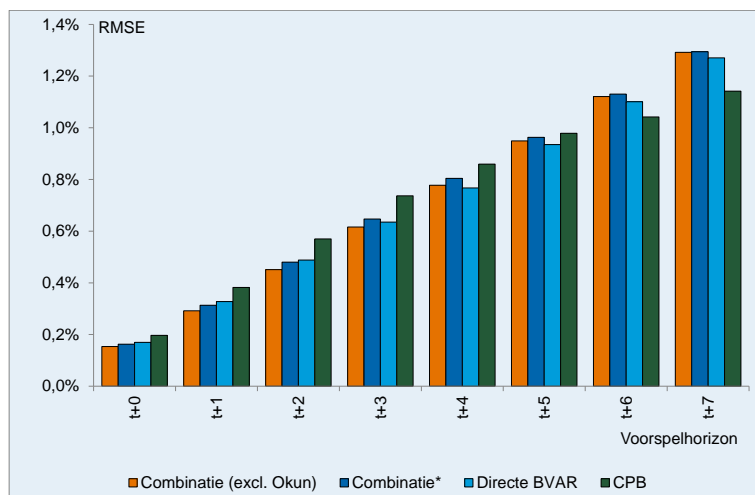
Daarnaast vergelijken we in figuur 7.1 de twee combinatievoorspellingen met de voorspellingen van de directe BVAR en de CPB-ramingen volgens de huidige methodiek. Op basis van de tabel en de figuur kunnen we de volgende conclusies trekken:

1. Over alle tijdshorizons is de combinatievoorspelling beter als we het Okun-model niet meenemen.
2. Tot en met vijf kwartalen vooruit zijn de combinatievoorspellingen zonder Okun beter dan de voorspellingen van het CPB.
3. Op de korte termijn (tot en met $t + 3$) geldt dat deze combinatievoorspelling gemiddeld ook de laagste voorspelfout heeft.
4. Op de langere termijn, vanaf zes kwartalen vooruit, is de CPB-raming de beste.

5. In de tussenperiode is het beeld minder duidelijk. Voor de raming van vier en vijf kwartalen vooruit is de gemiddelde voorspelfout van de directe BVAR het kleinst.

De keuze van een beter instrument hangt niet alleen af van de gemiddelde voorspelkwaliteit gedurende de hele testperiode. Zoals de inleiding al aangaf, is het ook belangrijk hoe de voorspelkwaliteit is rond omslagpunten. Daarover gaat het volgende hoofdstuk.

Figuur 7.1 Voorspelkwaliteit van de combinatievoorspelling



De combinatie-raming is de gemiddelde raming van: Okun (dgap HP), Duurdata (BVAR, met CBS-data), drie staten (BVAR) en de directe BVAR.

8 Omslagpunten

Een van de doelen van dit onderzoek is ook om te kijken hoe de verschillende methodes presteren rond zogenaamde omslagpunten. We hanteren daarbij de volgende definitie: periode $t - 1$ is een omslagpunt als de werkloosheid in $t - 1$ steeg (daalde) maar in periode t daalt (stijgt).²⁰

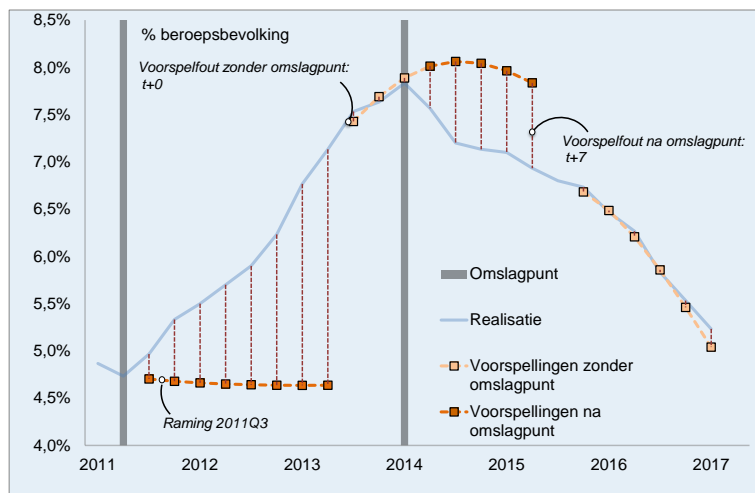
Onze steekproef met voorspellingen loopt van het eerste kwartaal in 2011 tot en met het tweede kwartaal van 2017. Dit is de testperiode van alle voorspelmodellen, zie ook hoofdstuk 3. In deze periode komen twee omslagpunten voor: het tweede kwartaal van 2011 en het eerste kwartaal van 2014. Figuur 8.1 laat dit zien.

Als een omslagpunt op het ramingsmoment nog niet bekend is, kan dit de voorspelkwaliteit voor de periode na dit punt negatief beïnvloeden. Daarom kijken we in dit hoofdstuk specifiek naar de voorspelfouten van ramingen die zijn gedaan vóórdát bekend was dat er sprake was van een omslagpunt.

²⁰ Het maakt niet uit hoe lang de periode voor en na een omslagpunt wordt gedefinieerd, omdat de werkloosheid een duidelijk dal (2011Q2) of een duidelijke piek (2014Q1) heeft, zie figuur 8.1.

In een *real-time* context is pas na twee kwartalen bekend dat een omslagpunt heeft plaatsgevonden. Een knik in een reeks is per definitie pas een kwartaal later observeerbaar, omdat dan pas blijkt dat de werkloosheid is gedaald (gestegen) terwijl deze in de periode daarvóór steeg (daalde). Daarnaast komen alle gegevens met één kwartaal vertraging beschikbaar. Het omslagpunt is dus met twee kwartalen vertraging bekend: we weten bijvoorbeeld pas in 2011Q4 dat 2011Q2 een omslagpunt was. De tussenliggende raming die is gemaakt in 2011Q3 is dus een voorbeeld van een reeks voorspellingen die negatief beïnvloed kan worden door een omslagpunt. Figuur 8.1 laat voor nog een tweetal andere reeksen zien welke voorspellingen wel en niet betrekking hebben op een omslagpunt.

Figuur 8.1 Omslagpunten en voorspellingen werkloosheid

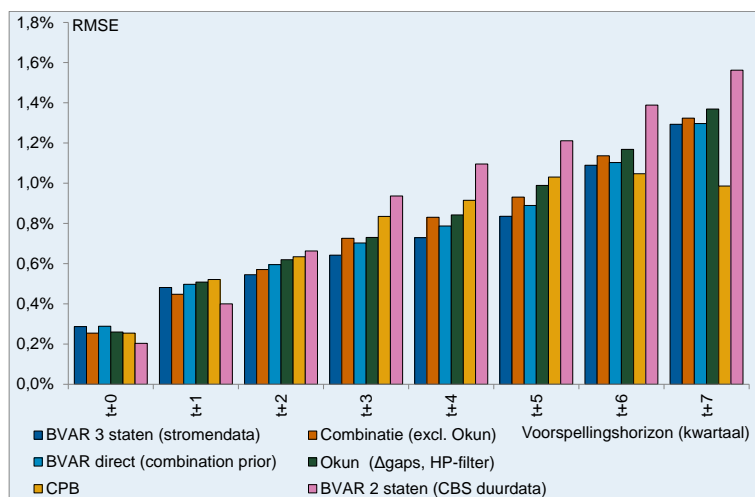


Voor iedere onderzochte methode meten we de voorspelfouten na een omslagpunt. Door deze met elkaar te vergelijken, kunnen we beoordelen hoe goed hun voorspelkwaliteit is. Figuur 8.2 laat de gemiddelde voorspelfouten na een omslagpunt zien van de best presterende modellen uit de vorige hoofdstukken. Hier is de gemiddelde voorspelfout voor $t + 0$ bijvoorbeeld de gemiddelde fout van de voorspelling van het lopend kwartaal na het omslagpunt; in ons geval zijn dit de voorspellingen voor $t + 0$ van de ramingen gemaakt in 2011Q3 en 2014Q2. En de gemiddelde voorspelfout voor $t + 2$ laat de gemiddelde fout zien van de voorspellingen twee kwartalen vooruit, waarbij $t + 1$, $t + 0$ of $t - 1$ een omslagpunt was.²¹ Voorwaarde is dus dat de raming is gemaakt vóórdat het omslagpunt bekend was.

Voor de combinatievoorspelling, het drie-staten-BVAR-model en de directe BVAR voorspellen relatief goed na omslagpunten. Voor een langere voorspelhorizon, zes tot zeven kwartalen vooruit, voorspelt de huidige CPB methodiek juist relatief goed. Onze analyse steunt op slechts twee omslagpunten; we kunnen dus slechts voorlopige conclusies trekken.

²¹ Voor $t + 2$ zijn dit zijn in totaal zes voorspellingen, namelijk de $(t + 2)$ -ramingen gemaakt in 2011Q1, 2011Q2, 2011Q3, 2013Q4, 2014Q1 en 2014Q2.

Figuur 8.2 Voorspelkwaliteit na een omslagpunt



9 Conclusies

Dit document beschrijft verschillende routes om de CPB-kortetermijnramingen van de werkloosheid te verbeteren. Op basis van de recente wetenschappelijke literatuur zijn een aantal instrumenten ontwikkeld. Hun voorspelkwaliteit is onderzocht voor de periode vanaf het eerste kwartaal van 2011 tot en met het eerste kwartaal van 2017. Een tweede aandachtspunt was de kwaliteit van voorspellingen in de buurt van omslagpunten.

De belangrijkste conclusie is dat de zogenaamde Bayesiaanse tijdreeksmodellen (BVAR) en de voorspelling die de beste resultaten van alle onderzochte technieken combineert, de werkloosheid tot vijf kwartalen vooruit beter voorspellen dan de huidige CPB systematiek. Dit geldt voor 'normale' perioden, maar ook rond de twee onderzochte omslagpunten.²² Vanaf zes kwartalen vooruit zijn de CPB-voorspellingen weer beter dan die van de alternatieven. Deze methoden zullen bij het CPB als zijlicht gebruikt worden bij de werkloosheidsramingen.²³ Op deze manier hopen we de kortetermijnramingen van de werkloosheid in de toekomst te verbeteren.

Deze algemene conclusie volgt uit de analyse van drie verschillende methoden om de werkloosheid te voorspellen. Als eerste hebben we geanalyseerd of de aanpak van Gordon (2010) en Ball et al. (2017) op basis van de wet van Okun tot betere kortetermijnvoorspellingen van de werkloosheid leidt. De tweede lijn richtte zich op pure tijdreeksanalyses. Daarbij maakten we onder meer gebruik van een specifieke (Bayesiaanse)

²² Omdat we in deze studie slechts twee omslagpunten konden meenemen is de conclusie over de voorspelkwaliteit van de modellen rondom omslagpunten enigszins onzeker.

²³ Omdat er naast de werkloosheid nog veel andere macro-economische variabelen geraamd worden en dit met elkaar een consistent beeld moet opleveren, zullen we de nieuwe methoden slechts als 'zijlicht' gebruiken. De voorspellingen van de zijlichten zullen naast de uitkomsten van het macromodel gehouden worden en op basis van *expert judgement* wordt dan bepaald over de modeluitkomsten bijgestuurd moeten worden.

techniek om het grote aantal parameters van een standaard vector-autoregressief (VAR) model in te perken. De derde methode probeerde de werkloosheid te voorspellen als de resultante van stromen op de arbeidsmarkt. In navolging van Barnichon et al. hebben we onderzocht of de tijdreekstechnieken uit de tweede onderzoekslijn de voorspellingen van deze stromen, en daarmee de werkloosheid, verbeteren. Tot slot hebben we combinatievoorspellingen gemaakt; deze combineert de best presterende modellen van elke methode. Dit zijn de modellen met de kleinste gemiddelde voorspelfout.

De eerste conclusie over de bruikbaarheid van de verschillende technieken is dat de Okun-modellen niet leiden tot betere voorspellingen dan die volgens de huidige methodiek van het CPB.

Het Bayesiaanse VAR-model (BVAR), dat de werkloosheid direct voorspelt met een groot aantal *leading* variabelen, leidt wel tot betere voorspellingen. Tot vijf kwartalen vooruit zijn de voorspelfouten kleiner dan die van het CPB.

Ten derde komen uit de stromenanalyse twee specificaties die de werkloosheid tot vier kwartalen vooruit beter voorspellen dan de gebruikelijke CPB-ramingen. Het eerste model kent slechts twee staten: iemand werkt of werkt niet. De beste specificatie van dit type is een tweestatenmodel geschat met CBS-data over de duur van werkloosheid. Deze gebruikt een BVAR om de stroomkansen in het stromenmodel te voorspellen. Een alternatief voor deze data zijn de meer gedetailleerde gegevens van Eurostat. Gebruik hiervan leidt echter niet tot betere voorspellingen dan die volgens de huidige CPB-methodiek. Doordat de CBS-duurdata tijdiger beschikbaar komen verwachten we wel dat het met het beschikbaar komen van meer gedetailleerde CBS-data mogelijk zou zijn de voorspelkwaliteit van het tweestatenmodel verder te verbeteren. Het tweede modeltype uit de stromenanalyse onderscheidt drie verschillende staten: werkend, werkloos en niet-participerend. De beste specificatie is hier een driestatenmodel dat de stromen voorspelt met een BVAR.

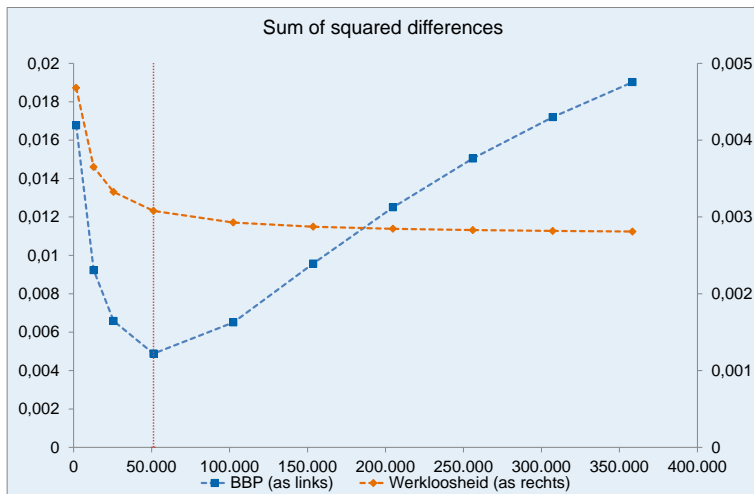
Een combinatievoorspelling op basis van de best presterende stromenmodellen en de directe BVAR verbetert de kwaliteit van de voorspelling van de werkloosheid nog verder.

Voorspellingen van de werkloosheid die vlak voor een omslagpunt worden gemaakt zijn vaak van mindere kwaliteit omdat het meestal onmogelijk is zo'n omslagpunt zelf goed te voorspellen. Toch verbeteren de combinatievoorspelling, het drie-staten-BVAR-model en de directe BVAR de voorspelling na een omslagpunt van een tot vijf kwartalen vooruit.

Appendix

A.1 Achtergrond bij Okun-regressies

Figuur A.1 *Smoothing parameters: verschil tussen CPB-reeks en Hodrick-Prescott filter*



Tabel A.1 **Modelspecificaties Okun-regressies uit hoofdstuk 4**

Specificatie		
Realistische voorspellingen (Flash bpb en revisies)		
Eerste verschillen		$\Delta U_{t+k} = \beta_1 \Delta U_{t-1} + \beta_2 \Delta U_{t-2} + \beta_3 \Delta U_{t-3} + \beta_4 \Delta U_{t-4} + \gamma_1 \Delta Y_{t-1} + \gamma_2 \Delta Y_{t-4} + \theta_1 U_{t-1} + \varepsilon_t$
Eerste verschillen <i>gaps</i>	HP filter	$\Delta \tilde{u}_{t+k} = \beta_1 \Delta \tilde{u}_{t-2} + \beta_2 \Delta \tilde{u}_{t-3} + \gamma_1 \Delta \tilde{y}_{t-4} + \varepsilon_t$
Niveau <i>gaps</i>	HP filter	$\tilde{u}_{t+k} = \beta_1 \tilde{u}_{t-1} + \beta_2 \tilde{u}_{t-3} + \varepsilon_t$
Onrealistische voorspellingen		
Eerste verschillen		$\Delta U_{t+k} = \beta_1 \Delta U_{t-2} + \gamma_1 \Delta Y_{t-1} + \gamma_2 \Delta Y_{t-3} + \gamma_3 \Delta Y_{t-4} + \varepsilon_t$
Eerste verschillen <i>gaps</i>	CPB	$\Delta \tilde{u}_{t+k} = \beta_1 \Delta \tilde{u}_{t-2} + \beta_2 \Delta \tilde{u}_{t-3} + \beta_3 \Delta \tilde{u}_{t-4} + \gamma_1 \Delta \tilde{y}_{t-1} + \varepsilon_t$
	HP filter	$\tilde{u}_{t+k} = \beta_1 \tilde{u}_{t-2} + \beta_2 \tilde{u}_{t-3} + \beta_3 \tilde{u}_{t-4} + \gamma_1 \tilde{y}_{t-1} + \gamma_2 \tilde{y}_{t-4} + \varepsilon_t$
Niveau <i>gaps</i>	CPB	$\tilde{u}_{t+k} = \beta_1 \tilde{u}_{t-1} + \beta_2 \tilde{u}_{t-2} + \beta_3 \tilde{u}_{t-4} + \gamma_1 \tilde{y}_{t-2} + \gamma_2 \tilde{y}_{t-3} + \varepsilon_t$
	HP filter	$\tilde{u}_{t+k} = \beta_1 \tilde{u}_{t-1} + \beta_2 \tilde{u}_{t-3} + \varepsilon_t$

A.2 Afleiding driestatenmodel in hoofdstuk 6

In tegenstelling tot het tweestatenmodel onderscheidt het driestatenmodel een derde categorie: niet-participerend, zie ook figuur 6.1. Er zijn dan in totaal zes stromen. Startpunt vormen de differentiaalvergelijkingen voor de drie arbeidsmarktvoorraden gedurende kwartaal t , met $\tau \in [0,1]$:

$$\dot{U}_{t+\tau} = -(\lambda_t^{UN} + \lambda_t^{UE})U_{t+\tau} + \lambda_t^{EU}E_{t+\tau} + \lambda_t^{NU}N_{t+\tau} \quad (3)$$

$$\dot{E}_{t+\tau} = \lambda_t^{UE}U_{t+\tau} - (\lambda_t^{EN} + \lambda_t^{EU})E_{t+\tau} + \lambda_t^{NE}N_{t+\tau} \quad (4)$$

$$\dot{N}_{t+\tau} = \lambda_t^{UN}U_{t+\tau} + \lambda_t^{EN}E_{t+\tau} - (\lambda_t^{NU} + \lambda_t^{NE})N_{t+\tau} . \quad (5)$$

Hier is U_t het aantal mensen dat in kwartaal t werkloos is, E_t het aantal mensen dat in kwartaal t werkzaam is en N_t het aantal mensen dat niet actief is op de arbeidsmarkt. Dit stelsel kan geschreven worden als:

$$\frac{d}{d\tau} \bar{y}_{t+\tau} = A_t \bar{y}_{t+\tau} \quad (5)$$

$$\bar{y}_{t+\tau} = \begin{pmatrix} U_{t+\tau} \\ E_{t+\tau} \\ N_{t+\tau} \end{pmatrix} \quad (5)$$

$$A_t = \begin{bmatrix} -(\lambda_t^{UE} + \lambda_t^{UN}) & \lambda_t^{EU} & \lambda_t^{NU} \\ \lambda_t^{UE} & -(\lambda_t^{EU} + \lambda_t^{EN}) & \lambda_t^{NE} \\ \lambda_t^{UN} & \lambda_t^{EN} & -(\lambda_t^{NU} + \lambda_t^{NE}) \end{bmatrix}. \quad (6)$$

De lambda's λ_t^{ab} geven de transitiekansen tussen de drie posities weer. De stroomkans van E naar U is bijvoorbeeld gedefinieerd als $\lambda_t^{EU} = \frac{EU_t}{E_{t-1}}$, waar EU_t de stroom vanuit werkzaam naar werkloos tussen t en $t-1$ weergeeft en E_{t-1} het aantal personen in de werkzame beroepsbevolking aan het begin van periode $t-1$ is.

De oplossing voor de differentiaalvergelijkingen ziet er als volgt uit:

$$\bar{y}_{t+\tau} = \sum_{i=1}^3 c_{it} \bar{q}_{it} e^{r_{it}\tau} . \quad (7)$$

Hier zijn de \bar{q}_{it} , $i=1, \dots, 3$ de eigenvectoren van A_t en de bijbehorende eigenwaarden van A_t zijn r_{it} , $i=1, \dots, 3$.²⁴ De constanten c_{it} , $i=1, \dots, 3$ worden bepaald aan de hand van de

²⁴ Voor een matrix A_t van deze vorm is bekend dat $-1 < r_{1i} < 0$, $i=1$ en 2 , en $r_{3i}=0$, waar de nummering van de eigenwaarde arbitrair is.

initiële condities $\tau = 0$ waarvoor de waarden van \bar{y}_t bekend zijn, en het feit dat als $\tau \rightarrow \infty$, $\bar{y}_{t+\tau}$ naar zijn *steady-state*-waarde convergeert, waarvan de som gelijk aan de som van de elementen van \bar{y}_t moet zijn.²⁵

Wij gebruiken vergelijking (9) om de werkloosheid U_{t+1} en het aantal werkenden E_{t+1} te voorspellen en hieruit de voorspelling van het werkloosheidspercentage af te leiden. Net als bij het tweestatenmodel zullen we eerst de transitiekansen in A_t voorspellen, waarmee we vervolgens de werkloosheid kunnen voorspellen. En net als bij de tweestatenmodellen doen we dat op verschillende manieren. Bij de variant met constante stroomkansen itereren we op basis van vergelijking (9) voor $\tau = 1, \dots, v$, waar v het aantal kwartalen vooruit is waar we voorspellingen voor willen produceren. Daarnaast zullen we ook hier de *in-sample*-VAR, *out-of-sample*-VAR en Bayesiaanse VAR (BVAR) gebruiken om de transitiekansen te voorspellen.

A.3 VAR-methode uit hoofdstuk 6

De stroomkansen kunnen voorspeld worden met een VAR-model. Omdat er veel potentieel relevante voorspelvariabelen zijn en het aantal op te nemen vertragingen flexibel is, zijn er heel veel specificaties mogelijk. Het selecteren van een optimaal model kan op basis van grofweg twee criteria, *in-sample* en *out-of-sample*. De uitkomst kan bovendien verschillen voor het twee- en driestatenmodel. Verder zou de keuze voor Eurostat- of CBS-duurdata uit kunnen maken voor de optimale tweestatenmodellen. Door de beperkte beschikbaarheid van de CBS-duurdata zullen we voor de tweestaten-VAR echter alleen de duurdata van Eurostat gebruiken.

In-sample-optimalisatie

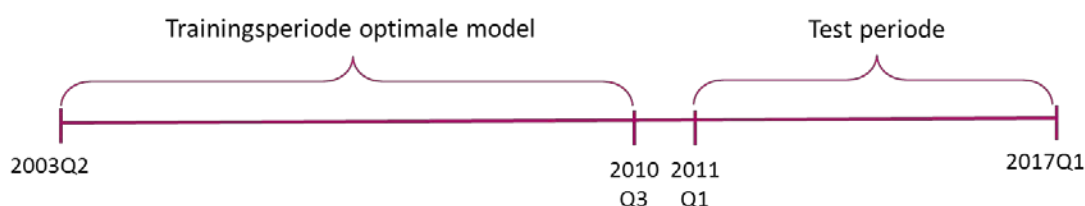
We maken onderscheid tussen de periode die gebruikt wordt voor de selectie van het optimale VAR-model en de periode die we gebruiken om het model te testen, zodat we de resultaten van de verschillende methodes naast elkaar kunnen leggen. Het CPB voorspelt vanaf het eerste kwartaal van 2011 de werkloosheid op basis van de internationale definitie. Dit is dus een natuurlijk moment om de testperiode te laten starten.

²⁵ In de limiet vallen de termen met $c_{it}=1$ en 2 weg, en hebben we dat $U_t + E_t + N_t = C_{3t}(Q_{13t} + Q_{23t} + Q_{33t})$. De steady state van y_t is dan $c_{3t}Q_{3t}$. Hierdoor nemen we aan dat wij de bevolkingsgroei kunnen negeren. Voor c_{1t} en c_{2t} moeten we de initiële condities gebruiken van twee van de drie voorraden en het twee bij twee stelsel voor twee van de elementen van $y_t - C_{3t}Q_{3t}$ oplossen.

Twee staten

Bij het tweestatenmodel kan de periode 2003Q2-2010Q3 dus gebruikt worden om het optimale VAR model te bepalen,²⁶ dit noemen we de trainingsperiode. Zie ook figuur A.2.

Figuur A.2 Optimale VAR model (*in sample*)



Door de relatief korte trainingsperiode van 30 kwartalen is het niet mogelijk om elke denkbare VAR-specificatie te schatten: wanneer er te veel voorspelvariabelen en/of vertragingen worden meegenomen, is het risico op *overfitting* groot. Naast de natuurlijke logaritmes van de transitiekansen s_t en f_t , zullen we maximaal zes verschillende variabelen in de VAR-specificatie opnemen. Dit is steeds maximaal één voorspelvariabele uit elke categorie uit tabel 6.1. Daarnaast is het maximale aantal vertragingen op twee gesteld. In totaal schatten we dus 14.850 unieke VAR-modellen. Vervolgens wordt op basis van het Akaike informatiecriterium (AIC) bepaald welke model het beste presteert.

De optimale *in-sample* VAR-specificatie (die met de kleinste AIC) is dan:

$$\mathbf{y}_t = \mathbf{c} + \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \boldsymbol{\varepsilon}_t, \quad (8)$$

met $\mathbf{y}_t = (\ln s_t, \ln f_t, \ln faill_t, \ln instr_t, \Delta \ln vac_ont_t, u_t, \ln bbp_t)'$.²⁷

Drie staten

De *in-sample*-optimale VAR-specificatie voor het driestatenmodel wordt op een vergelijkbare wijze geselecteerd. Maar omdat dit type model niet twee, maar zes transitiekansen beschrijft, is het maximale aantal vertragingen op één gesteld om *overfitting* te voorkomen. Verder loopt de trainingsperiode tot en met 2010Q4, aangezien de stromendata van het CBS geen extra publicatievertraging hebben.

De VAR-specificatie met de kleinste AIC is dan:

$$\mathbf{y}_t = \mathbf{c} + \Phi_1 \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \quad (9)$$

met $\mathbf{y}_t =$

$(\ln \lambda_t^{EU}, \ln \lambda_t^{UE}, \ln \lambda_t^{EN}, \ln \lambda_t^{NE}, \ln \lambda_t^{NU}, \ln \lambda_t^{UN}, \ln faill_t, \ln instr_t, \ln vac_t, u_t, \ln bbp_t)'$.

²⁶ Dit komt door de publicatievertraging van twee kwartalen van de duurdata.

²⁷ Zie tabel 6.1 voor een toelichting op de afgekorte variabelennamen. Wanneer wordt uitgegaan van duurdata die wel tijdig beschikbaar zijn, en de trainingsperiode dus tot en met 2010Q4 loopt, is het optimale model nagenoeg hetzelfde.. Het enige verschil is dat dan niet $\Delta \ln vac_ont_t$, maar $\ln vac_ont_t$ in de \mathbf{y}_t -vector wordt opgenomen.

Out-of-sample-optimalisatie

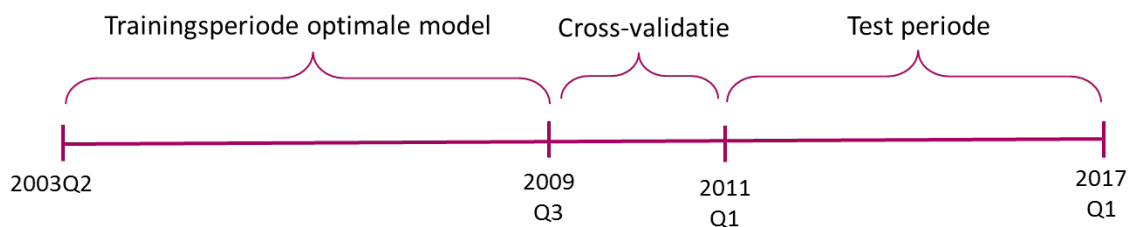
Bij *out-of-sample*-optimalisatie wordt onderzocht welk model het beste voorspelt tijdens de zogenaamde cross-validatieperiode. Het gaat hierbij om de *out-of-sample*-prestaties van het VAR-model in de periode direct na de trainingsperiode. Het model met de kleinste voorspelfouten in de cross-validatieperiode wordt aangemerkt als het optimale model.

Omdat het optimale model nog steeds in 2011Q1 bekend moet zijn, is er sprake van een afruil tussen de lengte van de trainings- en cross-validatieperiode. Door het relatief kleine aantal beschikbare observaties hebben we ervoor gekozen de trainingsperiode zo min mogelijk in te korten en gebruik te maken van een validatieperiode van (slechts) vijf kwartalen.

Twee staten

In een situatie met tijdig beschikbare duurdata werkt de methode als volgt (zie figuur A.3).

Figuur A.3 Optimale VAR model (*out-of-sample*)



De periode tot en met 2009Q3 wordt gebruikt voor het schatten van alle 14.850 modellen. Met elk VAR-model worden vervolgens de transitiekansen en daarmee (met behulp van de *law of motion*) het werkloosheidspercentage voorspeld. Dit gebeurt tot vijf kwartalen vooruit (dus voor de periode 2009Q4-2010Q4). De vijf voorspellingen van s , f en u worden ten slotte naast hun realisaties gelegd. Het beste model heeft de kleinste gewogen²⁸ gemiddelde gekwadraterde gestandaardiseerde²⁹ voorspelfout over deze drie variabelen. Dit is de volgende specificatie:

$$\mathbf{y}_t = \mathbf{c} + \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \boldsymbol{\varepsilon}_t, \quad (10)$$

met

$$\mathbf{y}_t = (\ln s_t, \ln f_t, \Delta \ln ontslag_aan_t, \ln instr_t, \ln vac_ont_t, \Delta u_t, \ln bbp_t, \Delta uitzend_t)'$$

Door de publicatievertraging van twee kwartalen bij de duurdata zijn er bij deze methode twee aanpassingen noodzakelijk. Ten eerste is er, net als bij *in-sample*-optimalisatie, een kwartaal minder beschikbaar voor de selectie van het optimale model doordat de gerealiseerde transitiekansen van 2010Q4 nog niet bekend zijn op het moment waarop het optimale model bekend moet zijn (2011Q1). De cross-validatieperiode loopt dus maximaal

²⁸ De gemiddelde voorspelfouten van s en f krijgen elk een gewicht van $\frac{1}{2}$, en de voorspelfout van u krijgt een gewicht van $\frac{1}{2}$. Er is dus sprake van een gelijk gewicht voor de transitiekansen en de werkloosheid.

²⁹ Omdat er een niveau- en variatieverschil is tussen de drie variabelen, worden de voorspelfouten gedeeld door de standaarddeviatie van de betreffende variabelen (over de periode tot en met de laatste trainingsobservatie).

tot en met 2010Q3, waardoor de trainingsperiode moet worden ingekort tot 2003Q2-2009Q2. Ten tweede moeten we rekening houden met het feit dat in de *real-time* context ook in die trainingsperiode de laatste realisaties van de transitiekansen op het moment van schatten ontbreken. Dit doen we door s_{2009Q2} en f_{2009Q2} te vervangen door de op dat moment laatst bekende waarden (respectievelijk s_{2009Q1} en f_{2009Q1}).³⁰

De optimale VAR-specificatie over de periode 2009Q3-2010Q3 ziet er dan als volgt uit:

$$\mathbf{y}_t = \mathbf{c} + \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \boldsymbol{\varepsilon}_t, \quad (11)$$

met

$$\mathbf{y}_t = (\ln s_t, \ln f_t, \Delta \text{uitstroomkans_ww_limiet}_t, \Delta \ln \text{vac_ont}_t, \Delta u_t, \ln \text{bbp}_t, \Delta \text{uitzend}_t)'$$

Drie staten

De *out-of-sample*-optimale VAR-specificatie voor het driestatenmodel komt op een vergelijkbare wijze tot stand als die voor het tweestatenmodel. Hier is echter geen sprake van een extra publicatievertraging, en is het maximale aantal vertragingen (wederom) op één gesteld. Bij de weging van de voorspelfouten krijgen de gemiddelde voorspelfouten van de λ 's allemaal een gewicht van $1/12$ en die van u een gewicht van $1/2$, waardoor er ook hier sprake is van een gelijk gewicht voor de transitiekansen en de werkloosheid.

De optimale VAR-specificatie over de periode 2009Q4-2010Q4 ziet er als volgt uit:

$$\mathbf{y}_t = \mathbf{c} + \Phi_1 \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \quad (12)$$

$$\text{met } \mathbf{y}_t = (\ln \lambda_t^{EU}, \ln \lambda_t^{UE}, \ln \lambda_t^{EN}, \ln \lambda_t^{NE}, \ln \lambda_t^{NU}, \ln \lambda_t^{UN}, \Delta \ln \text{faill}_t, \ln \text{instr}_t, \ln \text{vac_ont}_t, u_t)'$$

³⁰ Het 'voorspellen' op basis van de laatst bekende waarde komt neer op het aannemen van een *random walk*. Dit is een conservatieve aanpak: het is bijvoorbeeld ook mogelijk om de voorspelvariabelen (die wel tijdig beschikbaar komen) te gebruiken om de ontbrekende transitiekansen af te leiden. Een veelgebruikte methode om met dit soort *ragged edge data* om te gaan is het Kalmanfilter; deze is gebruikt bij de BVAR. De aanname dat *alle* data slechts beschikbaar zijn tot en met 2009Q1 is mogelijk nog conservatiever en kort de trainingsperiode met nog een kwartaal in.

Literatuur

Ball, L., D. Leigh en P. Loungani, 2017, Okun's law: Fit at 50?, *Journal of Money, Credit and Banking*, vol. 49(7): 1414-1440.

Barnichon, R. en P. Garda, 2016, Forecasting unemployment across countries: The ins and outs, *European Economic Review*, vol. 84: 165-183.

Barnichon, R. en C. Nekarda, 2012, The Ins and Outs of Forecasting Unemployment: Using Labor Force Flows to Forecast the Labor Market, *Brookings Papers on Economic Activity*, Fall 2012: 83-131.

Elbourne, A., K. Grabska, H. Kranendonk en J. Rhuggenaath, 2015, Effect van bijstelling economische groei door CBS op raming CPB, *CPB Achtergronddocument*, 11 december 2015.

Elsby, M.W.L., B. Hobijn en A. Şahin, 2013, Unemployment dynamics in the OECD, *Review of Economics and Statistics*, vol. 95(2): 530-548.

Gordon, R.J., 2010, Okun's Law and Productivity Innovations. *The American Economic Review*, vol. 100(2): 11-15.

James, G., D. Witten, T. Hastie en R. Tibshirani, 2013, *An Introduction to Statistical Learning*, New York: Springer.

Sims, C.A., 1998, Bayesian Methods for Dynamic Multivariate Models, *International Economic Review*, vol. 39(4): 949-968.

Verbruggen, J., H. Kranendonk en B. Smid, 2010, SAFFIER II: 1 model voor de Nederlandse economie, in 2 hoedanigheden, voor 3 toepassingen, *CPB Document*, no. 217.

Wooldridge, J.M., 2012, *Introductory Econometrics: A Modern Approach*. Mason, OH: South-Western.

Dit is een uitgave van:

Centraal Planbureau
Bezuidenhoutseweg 30
Postbus 80510 | 2508 GM Den Haag
T (088) 984 60 00

info@cpb.nl | www.cpb.nl

Maart 2018