

Keuze voor economisch voorspelmodel hangt af van stand conjunctuur

Een accurate raming van de economische ontwikkeling in de nabije toekomst is van belang omdat mensen hun gedrag hierop kunnen aanpassen. Welk voorspelmodel presteert wanneer het beste? Een analyse van de werkloosheidsvoorspellingen op basis van Nederlandse data.

IN HET KORT

- De voorspelfouten zijn groter in tijden van economische neergang en herstel, maar er zijn verschillen tussen modellen.
- ‘BVAR’ voorspelt het beste tijdens recessies, vooral wat betreft de ontwikkeling over anderhalf à twee jaar.
- De combinatievoorspelling doet het beter tijdens perioden van herstel, vooral op een periode van een à anderhalf jaar.

MARK KATTENBERG

Programmaleider
Data Science bij het
Centraal Planbureau
(CPB)

BAS SCHEER

Wetenschappelijk
medewerker bij het
CPB

Her voorspellen van economische ontwikkelingen is makkelijker gezegd dan gedaan. Modellen wijken in meer of mindere mate altijd wat af van de werkelijkheid. Neem bijvoorbeeld werkloosheid. Die blijkt erg moeilijk te voorspellen (Askitas en Zimmermann, 2009; Barnichon en Nekarda, 2012; Galbraith en Van Norden, 2019; Clark et al., 2020). Dit is helemaal het geval als de te voorspellen duur een recessieperiode betreft, of een periode van economisch herstel. Dat bleek bijvoorbeeld tijdens de coronapandemie, toen de contactbeperkende maatregelen voor een forse economische schok zorgden. Dankzij de coronasteunmaatregelen liep het gemiddelde werkloosheidspercentage over 2020 uiteindelijk op tot ‘slechts’ 4,9 procent (CBS, 2022). Dat was nog altijd fors hoger dan het gemiddelde werkloosheidspercentage van 3,7 procent dat het Centraal Planbureau raamde vóór het uitbreken van de coronapandemie (CPB, 2019).

Dát modellen recessie- of herstelperioden minder goed voorspellen, is op zich niet verrassend (Butler et al., 2017). Een recessie of economisch herstel kan immers het gevolg zijn van ontwikkelingen buiten de economie om, zoals het voorbeeld van de coronapandemie laat zien. Maar op een gegeven moment worden de recessies en herstelperioden wel voor het eerst in de data zichtbaar. Het ene model kan dit misschien sneller oppikken dan het andere.

De vraag is of er modellen zijn die systematisch relatief beter voorspellen tijdens economische recessies of herstel (Gelfer, 2019; Paccagnini, 2019; Foroni et al., 2020; Primiceri en Tambalotti, 2020). In dit artikel analyseren we voor tien modellen hoe hun voorspelprestaties aangaande de werkloosheid in Nederland veranderen tijdens economische recessies of herstel. We kijken daarbij niet naar één specifieke crisis, maar gebruiken een nieuwe methode om voorspelfouten te wegen naar de stand van de conjunctuur (zie Scheer (2022) voor een toepassing op Amerikaanse werkloosheidsdata).

Kennis van de prestaties van modellen met betrekking tot het voorspellen van recessies en herstelperioden is van belang omdat mensen juist in dergelijke turbulente tijden houvast zoeken door het economische nieuws meer te volgen, vooral tijdens recessies (Damstra en Boukes, 2021). Dat zien we ook, nu er vanwege de oorlog in Oekraïne en de daaruit volgende inflatie een recessie in de lucht hangt (Dicou, 2022).

Data en methoden

We maken gebruik van de maandelijkse reeks van het Nederlandse werkloosheidspercentage die beschikbaar is vanaf januari 2003 (figuur 1). De periode 2003–2012 gebruiken we om de modellen te schatten en te trainen. De periode 2012–2021 gebruiken we om de voorspellingen van de verschillende modellen te analyseren.

De geanalyseerde voorspelmodellen zijn deels afkomstig uit de econometrie en deels uit *machine learning*. De modellen zijn gebaseerd op schatters die vaker worden toegepast bij onderzoeken naar voorspelmodellen voor werkloosheid (Coulombe et al., 2022): *Lasso*, *Random Forest* (RF), lineair *Support Vector Regression* (SVR-lin) en *Support Vector Regression* met een *Radial Basis Function*-kernel (SVR-rbf), een *Auto-Regression* (AR), en een *Bayesian Vector Auto-Regression* (BVAR). Alle schatters zijn op twee manieren toegepast: op het niveau van het werkloosheidspercentage, en op de verandering in het werkloosheidspercentage – met uitzondering van de AR en Lasso. In totaal telt dit op tot tien voorspelmodellen. Ten slotte voegen we een gecombineerde voorspelling toe, namelijk de gemiddelde voorspelling van alle tien voorspelmodellen samen.

De inputs voor de modellen bestaan uit informatie die beschikbaar is tot aan het moment van het ramen. Als voorspellende variabelen gebruiken we het werkloosheids-

percentage zelf, het consumenten- en producentenvertrouwen, de vacature-indicator, het aantal faillissementen, en de instroom, uitstroom, en standcijfers van de WW-uitkeringen. Van elk van deze indicatoren geven we de modellen de meest recente zes maanden aan gegevens mee, om een voorspelling op te baseren. De meest recente realisatie wordt meegenomen in niveau, de overige vijf maanden in eerste verschillen.

Om na te gaan of de prestaties van de voorspelmodellen samenhangen met de conjunctuur, meten we de stand van de conjunctuur met de mate van verandering van het werkloosheidspercentage. Deze cyclusindicator is gelijk aan de afwijking van de huidige werkloosheid ten opzichte van X maanden geleden en X maanden later. De cyclusindicator meet zowel de richting waarin de werkloosheid zich beweegt, als de kracht van de beweging. Het is dus een continue maatstaf voor de snelheid waarmee de werkloosheid van richting verandert – wat het een goede indicator maakt voor turbulente economische tijden.

Figuur 1 illustreert wanneer de cyclusindicator uitslaat. De cyclusindicator is vooral hoog tijdens en vlak voor recessies, zoals in 2009, 2011 en 2019. De cyclusindicator is laag in herstelperioden, zoals het herstel van de dotcom-crisis in 2005, een kortstondig herstel van de kredietcrisis in 2010, en het krachtige herstel van de Europese schulden-crisis in 2014.

Resultaten

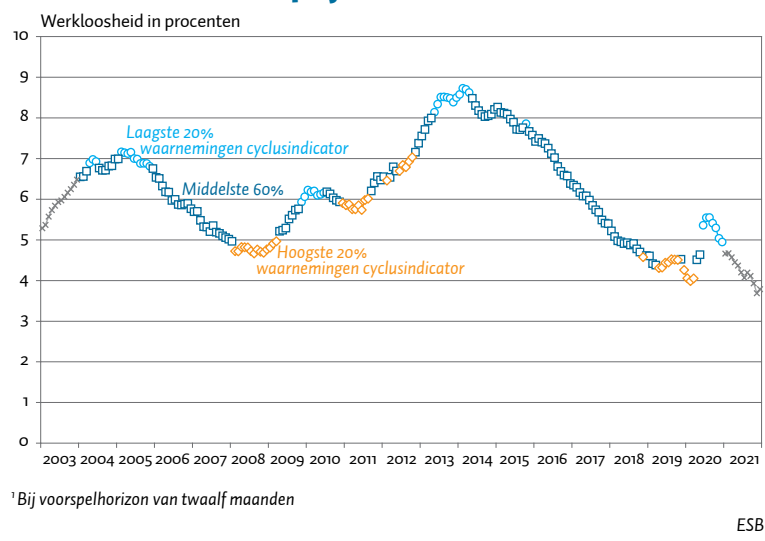
Figuur 2 toont de afwijking van de modelschatting ten opzichte van de waargenomen waarde, gemeten aan de hand van de *root-mean-square error* (RMSE) van de verschillende voorspelmodellen wanneer de conjunctuurindicator laag, gemiddeld of hoog is. Bij een voorspelhorizon van zes maanden zijn de voorspelfouten doorgaans het grootst in tijden van economische neergang, maar ze zijn ook groot bij herstel (figuur 2a). De voorspelfouten zijn het kleinst in de tussenperioden waarin de verandering in de werkloosheid stabiel is. Tot twaalf maanden vooruit is dit patroon duidelijk te zien, maar daarna wordt het minder sterk. Bij een langere horizon van 24 maanden vooruit is er geen patroon meer te zien, en zijn de voorspelfouten tijdens perioden van neergang of herstel regelmatig kleiner dan die tijdens de tussengelegene perioden (figuur 2b).

Op de korte tot middellange termijn maken alle modellen, tijdens perioden van economische neergang of economisch herstel, grotere voorspelfouten, maar niet in dezelfde mate. Sommige modellen presteren bijvoorbeeld relatief goed voor recessies. Door deze verschillen in het foutenpatroon hangt de keuze van het beste voorspelmodel af van het doel van de voorspellingen: de rangschikking van voorspelmodellen verandert als we meer gewicht toekennen aan de juistheid van voorspellingen in recessies, perioden van herstel of normale perioden (Scheer, 2022). Zo blijkt de BVAR het beste te voorspellen als we meer gewicht toekennen aan recessieperioden (vooral op de langere termijn: 18 tot 24 maanden), terwijl de combinatievoorspelling dit het beste doet wanneer we meer gewicht toekennen aan herstelperioden (vooral op de termijn van 12 tot 18 maanden).

We hebben gezien dat alle modellen, bij een horizon van zes maanden, grotere voorspelfouten maken tijdens

Ontwikkeling in de Nederlandse werkloosheid en momenten waarop cyclusindicator' uitslaat

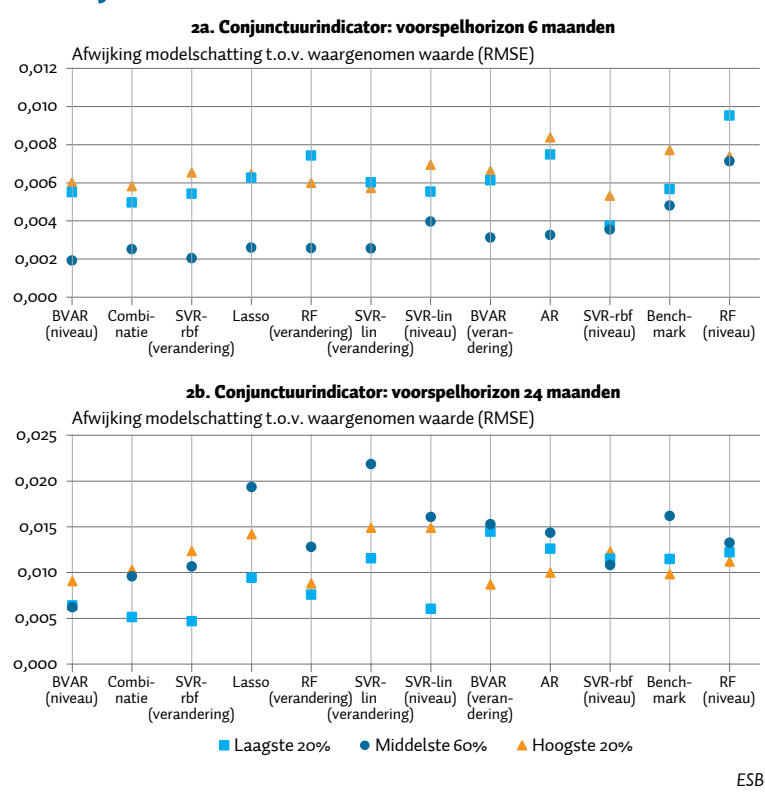
FIGUUR 1



ESB

Gemiddelde voorspelfout naar stand van de conjunctuur

FIGUUR 2



ESB

perioden van economische neergang of economisch herstel, maar niet in dezelfde mate. Over het algemeen geldt dus dat voorspelmodellen grotere fouten maken in turbulenteren tijden. Dit is echter geen wetmatigheid, maar hangt af van de voorspelhorizon. Zo laat het rechterpaneel van figuur 2 zien dat, bij een horizon van 24 maanden, voorspelfouten in turbulente tijden dichterbij die van normale tijden liggen. Soms liggen de voorspelfouten in turbulente tijden zelfs onder die van normale tijden, zoals bijvoorbeeld bij de Lasso en RF diff.

Conclusie en implicaties

Sommige voorspelmodellen doen het relatief goed tijdens recessieperioden (de BVAR), terwijl andere het juist weer goed doen tijdens herstelperioden (de combinatieschatter). Een beperking van de analyse is dat de uitkomsten gebaseerd zijn op weinig datapunten. Scheer (2022) komt, in vergelijkbaar onderzoek op langere Amerikaanse tijdreeksdata met meerdere recessie- en herstelperioden, echter tot dezelfde hoofdconclusie: de prestatie van voorspelmodellen hangt af van de stand van de economie en de voorspelhorizon.

Het zijn niet altijd dezelfde modellen die relatief goed voorspellen. De betrouwbaarheid hangt bijvoorbeeld af van de hoeveelheid data – zo doen SVR en RF het beter op de langere Amerikaanse dataset. Het is dus van belang om de voorspelkwaliteit te blijven monitoren.

De resultaten verplichten economen om na te denken over welk model ze gebruiken. Er is geen *one-size-fits-all*-model. Welk model geprefereerd wordt, hangt af van het belang dat er wordt toegekend aan de voorspelfouten tijdens perioden van economische neergang of herstel.

Wat betekent dit in de praktijk? De economische situatie in Nederland is de afgelopen maanden verslechterd, waardoor velen, waaronder het Centraal Planbureau, een terugval in economische groei verwachten (CPB, 2022). Ons onderzoek geeft aan dat professionele ramers nu kunnen besluiten om meer gewicht te geven aan de voorspellingen van de BVAR. De BVAR voorspelt de werkloosheid immers relatief goed tijdens recessieperioden op de langere termijn.

Het afstellen van modellen op de conjunctuur, vereist wel een keuze wat betreft het gewicht dat aan zulke perioden gegeven moet worden. Daarbij dienen voorspellers tevens de maatschappelijke gevolgen mee te wegen. Zo zijn er bijvoorbeeld economen die beargumenteren dat voorspellingen van economische groei vooral niet te hoog moeten worden ingezet, omdat te optimistische ramingen van het bbp negatieve economische gevolgen kunnen hebben (Beaudry en Willems, 2022).

Door bij het afstellen van modellen extra rekening te houden met recessie- of herstelperioden, kunnen de relatieve prestaties van modellen veranderen. Een Random Forest afgesteld op basis van een normale RMSE, kan er bijvoorbeeld anders uit zien dan een Random Forest afgesteld op basis van een RMSE waarin meer gewicht wordt gegeven aan recessieperioden. Toekomstig onderzoek kan dus bekijken hoe de relatieve prestatie van modellen verandert als gevolg van de afstelling van modellen op de conjunctuur.

Literatuur

- Askitas, N. en K.F. Zimmermann (2009) *Google econometrics and unemployment forecasting*. IZA Discussion Paper, 4201.
- Barnichon, R. en C.J. Nekarda (2012) *The ins and outs of forecasting unemployment: Using labor force flows to forecast the labor market*. Brookings Papers on Economic Activity, najaar, 83–131.
- Beaudry, P. en T. Willems (2022) *On the macroeconomic consequences of over-optimism*. *American Economic Journal: Macroeconomics*, 14(1), 38–59.
- Butler, B., D. van Dijk en A. Stokman (2017) *Voorspellers onderschatten cyclische economie*. ESB, 102(4755), 552–555.
- CBS (2022) *Arbeidsdeelname en werkloosheid per maand*. CBS Statline, 29 maart.
- Clark, T.E., M.W. McCracken en E. Mertens (2020) *Modeling time-varying uncertainty of multiple-horizon forecast errors*. *The Review of Economics and Statistics*, 102(1), 17–33.
- Coulombe, P.G., M. Leroux, D. Stevanovic en S. Surprenant (2022) *How is machine learning useful for macroeconomic forecasting?* *Journal of Applied Econometrics*, 37(5), 920–964.
- CPB (2019) *Juniraming 2019*. CPB Raming, juni.
- CPB (2022) *Macro Economische Verkenning 2023*. CPB Raming, september.
- Damstra, A. en M. Boukes (2021) *The economy, the news, and the public: A longitudinal study of the impact of economic news on economic evaluations and expectations*. *Communication Research*, 48(1), 26–50.
- Dicou, D. (2022) *Geen glazen bol, wel een prisma*. Column op www.cpb.nl, 31 maart.
- Foroni, C., M. Marcellino en D. Stevanović (2020) *Forecasting the Covid-19 recession and recovery: Lessons from the financial crisis*. ECB Working Paper, 2468.
- Galbraith, J.W. en S. van Norden (2019) *Asymmetry in unemployment rate forecast errors*. *International Journal of Forecasting*, 35(4), 1613–1626.
- Gelfer, S. (2019) *Data-rich DSGE model forecasts of the great recession and its recovery*. *Review of Economic Dynamics*, 32, 18–41.
- Paccagnini, A. (2019) *Did financial factors matter during the Great Recession?* *Economics Letters*, 174, 26–30.
- Primiceri, G.E. en A. Tambalotti (2020) *Macroeconomic forecasting in the time of COVID-19*. Northwestern University, Paper, juni. Te vinden op faculty.wcas.northwestern.edu.
- Scheer, B. (2022) *Addressing unemployment rate forecast errors in relation to the business cycle*. CPB Discussion Paper, maart.