

CPB Document

No 178

Februari, 2009

Trefzekerheid CPB-prognoses 1971-2007

Henk Kranendonk, Jasper de Jong en Johan Verbruggen

Centraal Planbureau
Van Stolkweg 14
Postbus 80510
2508 GM Den Haag

Telefoon (070) 338 33 80
Telefax (070) 338 33 50
Internet www.cpb.nl

ISBN 978-90-5833-392-6

Korte samenvatting

In dit document wordt de trefzekerheid van korte-termijnramingen van het CPB onderzocht. Gekeken wordt naar voorspellingen voor de jaren 1971-2007. De belangrijkste doelen van dit onderzoek zijn enerzijds het verkrijgen van inzichten die kunnen leiden tot betere ramingen in de toekomst, anderzijds het illustreren van de onzekerheid waarmee economische voorspellingen noodzakelijkerwijs zijn omgeven.

Hoewel prognoses van de ontwikkeling van diverse economische grootheden aan bod komen, gaat de meeste aandacht uit naar de BBP-groei. In iedere trefzekerheidspublicatie van het CPB worden één of enkele specifieke onderwerpen nader bekeken. Dit keer worden de CPB-ramingen van de BBP-groei vergeleken met die van andere ramingsinstituten, zowel nationale als internationale. Uit dit 'vergelijkend warenonderzoek' komt naar voren dat het maken van een nauwkeurige voorspelling van de economische groei voor alle instellingen een moeilijke opgave blijkt. Er blijft ruimte voor verbetering.

Steekwoorden: Trefzekerheid, voorspelfouten, korte-termijnramingen, modelonzekerheid:

Abstract

This report analyses the accuracy of short-term economic forecasts by the CPB. The years 1971-2007 are taken into account. The goals of this study are twofold. First of all, knowledge of the quality of CPB predictions and of sources of prediction errors may help to improve future forecasts. Secondly, we like to illustrate the uncertainties which necessarily surround economic projections.

A number of economic variables are being analyzed, but the main focus is on GDP growth. This enables us to compare CPB forecasts to those performed by other institutions, both national and international ones. Since all forecasters make sizeable prediction errors, an important lesson from this study is that making accurate forecasts is a difficult task. Nevertheless, this also indicates that there is room for improvement.

Key words: Accuracy, forecast errors, short-term forecasts, model uncertainty

A comprehensive summary is available from www.cpb.nl.

Inhoud

Ten geleide	7
Samenvatting	9
1 Inleiding	13
2 Ramingen en realisaties	15
2.1 De totstandkoming van CPB-ramingen	15
2.2 Selectie van de relevante CBS-realisaties	16
2.3 Kengetallen om trefzekerheid te beoordelen	19
3 Trefzekerheid korte-termijnramingen	23
3.1 Economische groei	23
3.2 Andere economische grootheden	28
3.3 Zijn de ramingen zuiver en efficiënt?	29
4 Analyse voorspelfouten	31
4.1 Onzekerheid in CPB-ramingen: een uitsplitsing in determinanten	31
4.2 Betekenis buitenlandexogenen voor de trefzekerheid	34
5 Trefzekerheid tussen instellingen vergeleken	37
5.1 Methodiek	37
5.2 Vergelijkbaarheid	39
5.3 Resultaten	41
6 Trefzekerheid in internationaal perspectief	47
7 Evaluatie	51
Bijlage 1 Specifieke onderwerpen in vorige trefzekerheidspublicaties	53
Bijlage 2 Trefzekerheid van CPB-ramingen	55
Literatuur	59

Ten geleide

De korte-termijnramingen van het CPB spelen een belangrijke rol bij de voorbereiding van het financieel-economische beleid van het kabinet en bij cao-onderhandelingen. In woord en geschrift waarschuwt het CPB regelmatig dat deze ramingen onvermijdelijk met grote onzekerheidsmarges zijn omgeven. Een van de manieren waarop het CPB de gebruikers van zijn prognoses hierover informeert, is door regelmatig de ramingen te vergelijken met de realisaties. Ook zijn deze analyses behulpzaam bij het opsporen van de zwakke plekken in het ramingsproces en op die manier medebepalend voor de onderzoeksprioriteiten. Omwille van een zo groot mogelijke transparantie zijn alle in dit document gebruikte reeksen elektronisch via de CPB-website beschikbaar.

Eerdere publicaties over de trefzekerheid van de korte-termijnramingen bevatten steeds analyses van bepaalde deelaspecten van de trefzekerheid. In deze publicatie zijn de belangrijkste onderwerpen opnieuw opgenomen, op basis van geactualiseerde berekeningen. Dat betreft bijvoorbeeld de gevoeligheid voor de keuze van het CBS-realisatiecijfer, de duiding van conjuncturomslagen, de efficiency van de prognoses, de vergelijking met ramingen van andere Nederlandse instituten en het belang van de uitgangspunten ten aanzien van de internationale conjunctuur. Deze keer is de analyse uitgebreid met een vergelijking van de trefzekerheid van internationale instellingen (Europese Commissie, IMF en OESO) voor andere landen.

De publicatie is geschreven door Henk Kranendonk, Jasper de Jong en Johan Verbruggen. Jelte Haagsma en Paul Westra hebben geassisteerd bij het maken van de grafieken en het samenstellen van de bijlagen.

Coen Teulings

Directeur

Samenvatting

In dit document wordt verslag gedaan van onderzoek naar de trefzekerheid van de kortetermijnramingen van het CPB. De meeste aandacht gaat uit naar de prognoses voor de BBP-groei, maar ook andere economische grootheden komen, zij het beperkt, aan bod. Allereerst worden enkele kengetallen berekend die een cijfermatig beeld geven van de nauwkeurigheid van de ramingen. Daarna wordt besproken welke factoren de gemaakte voorspelfouten kunnen verklaren. Om de prestaties van het CPB in het juiste perspectief te kunnen plaatsen, worden de CPB-ramingen vervolgens vergeleken met de ramingen voor Nederland van enkele nationale en internationale instellingen. Ook wordt bekeken hoe het CPB het in internationaal opzicht doet. Een evaluerende paragraaf bespreekt hoe het CPB bij het maken van prognoses omgaat met onzekerheid.

De trefzekerheid van de CPB-projecties wordt geanalyseerd aan de hand van een aantal maatstaven die ieder een ander aspect van de betrouwbaarheid van de ramingen belichten. De *gemiddelde voorspelfout* van de BBP-groei in het komende jaar is in de periode 1971-2007 nul. Voor het lopende jaar is de gemiddelde fout licht negatief (-0,3%-punt). Achter een geringe gemiddelde voorspelfout kunnen forse over- en onderschattingen schuilgaan, die elkaar compenseren. Zo werd de BBP-groei in jaren met lage of zelfs negatieve groeicijfers vaak overschat, terwijl in jaren met hoogconjunctuur de economische groei doorgaans werd onderschat. Deze onnauwkeurigheid komt wel tot uiting in de *gemiddelde absolute voorspelfout*. Deze maatstaf varieert van 1,2%-punt voor de raming voor het komende jaar in de Macro Economische Verkenning (MEV) tot 0,7%-punt voor het lopende jaar in de MEV. Dit toont aan dat de ramingen beter worden naarmate er meer informatie beschikbaar komt. De *ongelijkheids- of Theilcoëfficiënt*, die bij de BBP-groei voor alle ramingsmomenten onder de één uitkomt, geeft aan dat de ramingen een meerwaarde hebben boven alternatieve, naïeve ramingen. Het voorspellen van een *versnelling of vertraging* van de economische groei blijkt lastig. Wanneer een ondergrens van 0,5%-punt wordt genomen om een groeiverandering aan te duiden als versnelling of vertraging, wordt slechts in 40 tot 50% van de gevallen de richting voor het komende jaar juist voorspeld. Dit percentage loopt op tot ongeveer 60% voor het lopende jaar. In de *bijstellingen* van de BBP-groeiraming tussen opeenvolgende edities van het Centraal Economisch Plan (CEP) en de MEV blijkt niet of nauwelijks systematiek te zitten.

Naast de BBP-groei is ook een bredere set van economische grootheden, waaronder de relevante wereldhandel, particuliere consumptie, contractloon, consumentenprijsindex en bedrijfsinvesteringen, bekeken. Voor alle variabelen geldt dat in de periode 1971-2007 de gemiddelde voorspelfout in de MEV-raming voor het komende jaar tussen -0,4%-punt en 0,6%-punt ligt. De gemiddelde absolute voorspelfouten komen uiteraard hoger uit. Met name variabelen die nadrukkelijk afhangen van buitenlandse ontwikkelingen worden gekenmerkt

door relatief grote absolute voorspelfouten. De bedrijfsinvesteringen blijken eveneens moeilijk te voorspellen. Toch hebben nagenoeg alle variabelen een Theilcoëfficiënt kleiner dan één. De particuliere consumptie is met een Theilcoëfficiënt van 1,0 de uitzondering. Voor de particuliere consumptie had in principe dus even goed de raming voor het lopend jaar gebruikt kunnen worden als raming voor het komende jaar. Een eerste statistische toets wijst voorts uit dat voor alle variabelen de gemiddelde voorspelfout niet significant van nul verschilt (geen 'bias'). De tweede toegepaste toets geeft aan dat, behalve voor de particuliere consumptie, de ramingen 'efficiënt' zijn, wat wil zeggen dat alle beschikbare informatie in de raming is opgenomen.

Welke factoren liggen ten grondslag aan de voorspelfouten? Er zijn in principe vier bronnen te noemen. De eerste bron betreft bijstellingen in de Nationale rekeningen door het CBS. Dit betekent dat er bij het maken van een raming is uitgegaan van een 'onjuist' voorlopig cijfer voor het meest recente jaar dat in de Nationale rekeningen is gepubliceerd. Dit beïnvloedt de raming voor het lopend jaar zowel via de overloop als via allerlei vertragingen in het gebruikte model. De tweede, en belangrijkste, bron van onzekerheid zit in de aannames omtrent de ontwikkeling van diverse exogene, meestal internationale, variabelen. Met name de ontwikkeling van de relevante wereldhandel is belangrijk voor de Nederlandse economie, klein en open als ze is. De ontwikkeling van internationale economische grootheden is op zichzelf echter al lastig te ramen, maar wordt daarnaast ook nog beïnvloed door niet te voorziene gebeurtenissen als terroristische aanslagen en geopolitieke spanningen. Voorts brengt het gebruik van een macro-econometrisch model noodzakelijkerwijs onzekerheid met zich mee. Modelparameters zijn geschat, zodat de exacte waarde onzeker is. Bovendien zijn ze geschat met data over het verleden en het is niet altijd bekend of relaties uit het verleden zonder meer ongewijzigd toegepast kunnen worden op de toekomst. Daarnaast is een model een gestileerde weergave van de werkelijkheid, zodat ook de residuen van de gedragsvergelijkingen onzekerheid met zich meebrengen. Alles wat het model niet kan verklaren, loopt in het residu en de waarde daarvan is daarom inherent onzeker. 'Expert opinion' is de vierde bron van onzekerheid. Modelramingen worden soms aangepast op basis van inzichten van experts. Dit is internationaal gebruikelijk en kan er bijvoorbeeld toe dienen om relevante beschikbare informatie te benutten. Zo wordt de ontwikkeling van contractlonen vaak aangestuurd, omdat een deel hiervan al vastligt door middel van cao's.

Van deze bronnen van onzekerheid is onderzocht hoe belangrijk ze zijn in het verklaren van voorspelfouten. In dit Document worden de uitkomsten van dit onderzoek besproken. Hieruit volgt dat met name de uitgangspunten over de exogene variabelen een grote rol spelen. Zo'n driekwart van de voorspelfout van de BBP-groei wordt hierdoor verklaard. De voorspelfouten bij de prijs en het volume van de goederenuitvoer komen voor ongeveer 90% voor rekening van de exogenen. Modelonzekerheid is met name bij de werkgelegenheid en het contractloon

belangrijk, terwijl bijstellingen in de CBS-cijfers het grootste deel van de voorspelfout bij de bedrijfsinvesteringen verklaren. Bijsturingen van modeluitkomsten zijn doorgaans vrij beperkt en leiden meestal tot een verkleining van de voorspelfouten. Om de betekenis van de buitenlandexogenen voor de trefzekerheid te illustreren, is bekeken hoe de ramingsuitkomsten zouden zijn geweest indien de juiste wereldhandelsgroei en wisselkoers waren ingezet. Het blijkt dat in geval van de goederenuitvoer de gemiddelde absolute voorspelfout zou zijn beperkt tot 2,8%-punt, tegenover 3,3%-punt in de officiële raming. Bij de BBP-raming zou de gemiddelde absolute voorspelfout worden gereduceerd van 1,2%-punt tot 0,8%-punt.

Een vergelijking van ramingen voor de BBP-groei in Nederland van een selectie van instellingen, inclusief het CPB, leert dat de onderlinge verschillen gering zijn. Geen enkel instituut steekt in de periode 1998-2007 ten aanzien van de onderzochte maatstaven van trefzekerheid in zowel het lopende als het komende jaar duidelijk boven de andere instanties uit. De trefzekerheid, afgemeten aan de gemiddelde absolute voorspelfout, voor de BBP-groeiraming voor het *lopende* jaar was in de periode 1998-2007 het grootst bij Rabobank, op de voet gevolgd door NIBC, ABN-AMRO, CPB, DNB en Consensus Forecasts. De Europese Commissie (EC) en de OESO scoren qua trefzekerheid wat minder dan gemiddeld. Hierin komt tot uiting dat deze internationale instituten het nadeel hebben van een langdurig ramingsproces, bestaande uit diverse iteratierondes, en niet beschikken over specifieke op de Nederlandse economie toegesneden actuele en unieke informatie. De gemiddelde absolute voorspelfout van de ramingen voor de BBP-groei in het *komende* jaar was in de onderzochte periode het kleinst bij de Rabobank, gevolgd door de ABN-AMRO en het CPB. Naar verhouding zijn de verschillen in trefzekerheid bij de raming voor het lopend jaar groter dan bij de raming voor het komende jaar. Dit is opmerkelijk, aangezien de onzekerheden omtrent de economische ontwikkeling in het komende jaar relatief groot zijn. Wellicht dat men juist om die reden bij het maken van de raming voor het komende jaar minder vaart op eigen waarnemingen en informatie en juist meer naar elkaar kijkt. Wat betreft het voorspellen van versnellingen of vertragingen van de economische groei, gedurende de jaren 1999-2007, valt op dat het CPB als enige instituut meer dan de helft 'juist' scoort.

Vervolgens wordt in deze studie bekeken hoe het CPB het er in internationaal perspectief vanaf brengt. Een gemiddelde voorspelfout van ongeveer nul, zoals het CPB die heeft voor de meeste variabelen, is tamelijk bijzonder. Het IMF, de OESO en de EC hebben voor veel landen en regio's gemiddelde voorspelfouten van enkele tienden van procentpunten of meer. De gemiddelde absolute voorspelfouten van het CPB en internationale instituten zijn daarentegen wel vergelijkbaar. Dat Nederland als kleine, open economie in relatief grote mate afhankelijk is van de moeilijk te ramen internationale conjunctuur, is een mogelijke verklaring voor het feit dat de absolute voorspelfouten vergelijkbaar zijn, ondanks gemiddelde voorspelfouten van het CPB die dicht bij nul liggen.

Deze studie sluit af met een evaluatie, waarin wordt besproken hoe het CPB zijn ramingen kan verbeteren, maar ook hoe het CPB met onzekerheid omgaat. Ramingen profiteren van een betere inschatting van de buitenlandexogenen, al is dat maar tot op zekere hoogte mogelijk, en van het up-to-date houden of verbeteren van het gebruikte macromodel. Daarnaast wordt er steeds meer gebruik gemaakt van 'leading indicators' en VAR-analyses om te bekijken wanneer er eventueel moet worden bijgestuurd. Toch zal er altijd onzekerheid blijven bestaan omtrent de uitkomsten. Daarom geeft het CPB in de betreffende publicaties, in aanvulling op de centrale projectie, vrijwel altijd een aantal onzekerheidsvarianten. Deze varianten bieden een goede mogelijkheid om op specifieke risico's in gaan en de gevolgen te laten zien indien de vooraf ingeschatte risico's bewaarheid worden.

1 Inleiding

Sinds 1994 publiceert het CPB elk kwartaal korte-termijnprognoses voor de Nederlandse economie die betrekking hebben op de verwachte macro-economische ontwikkelingen in het lopende en komende jaar.¹ De meeste aandacht gaat gewoonlijk uit naar de ramingen in het *Centraal Economisch Plan (CEP)* en de *Macro Economische Verkenning (MEV)*. Het CPB vergelijkt regelmatig de in het verleden gepubliceerde prognoses met de realisaties, teneinde lessen te trekken die tot verbetering van de ramingen kunnen leiden en om de onzekerheid omtrent de prognoses te illustreren. Gebeurde dat voorheen onregelmatig,² sinds 2002 gebeurt dit bijna jaarlijks.³ In iedere publicatie werd een specifiek aspect besproken. Bijlage 1 bevat een overzicht hiervan.

Dit jaar is ervoor gekozen een aantal onderwerpen waarover eerder is gepubliceerd, in geactualiseerde vorm opnieuw te analyseren. Daarbij wordt vooral ingezoomd op de meest relevante conjunctuurvariabele, te weten de volumegroei van het bruto binnenlands product (BBP). In de bijlagen is meer informatie over de andere variabelen waarover in het verleden werd gerapporteerd beschikbaar. Nieuw in dit Document is de vergelijking van de trefzekerheid van de BBP-groeiramingen van het CPB voor Nederland met die van internationale instellingen voor andere landen.

De opzet van dit Document is als volgt. Voordat we de trefzekerheid zelf bespreken, wordt in paragraaf 2 ingegaan op de ramingen van het CPB, de realisaties van het CBS en de maatstaven die we hanteren bij de analyse. In paragraaf 3 komt de trefzekerheid van de korte-termijnramingen aan de orde. Een analyse van de voorspelfouten en de grote rol daarbij van de buitenlandexogenen is te vinden in paragraaf 4. Vervolgens vergelijken we in paragraaf 5 de trefzekerheid van het CPB met de trefzekerheid van andere instellingen, terwijl we deze in paragraaf 6 in internationaal perspectief plaatsen. Paragraaf 7 bevat een afsluitende evaluatie. Ten slotte bevatten de bijlagen een overzicht van de specials die in eerdere publicaties over de trefzekerheid zijn opgenomen en een aantal extra tabellen met berekende kengetallen. De data die bij deze trefzekerheidsanalyses zijn gebruikt, zijn beschikbaar via de website van het CPB.

¹ In de periode 1991-1994 bevatten de zogenoemde '*Halfjaarlijkse Tussenrapportages*' alleen een kwalitatieve bijstelling (+, = of -) ten opzichte van de eerder in het CEP en de MEV gegeven prognoses.

² Zie bijvoorbeeld Van den Berg (1986), CPB (1999, 2000) en Don (2001).

³ Zie Kranendonk en Verbruggen (2002, 2003, 2005, 2006a en 2006c).

2 Ramingen en realisaties

2.1 De totstandkoming van CPB-ramingen

Het maken van prognoses voor de korte en middellange termijn behoort tot de kerntaken van het CPB. Sinds de oprichting van het CPB in 1945 verschijnt in het voorjaar het CEP. Daarnaast is het een lange traditie om op de derde dinsdag van september de MEV te publiceren. Aanvankelijk werd in het CEP alleen een prognose voor het lopende jaar gepubliceerd. In 1989 is de horizon verruimd en sindsdien staat er in het CEP ook een raming voor het komende jaar.⁴ In de loop der jaren zijn er twee andere publicatiemomenten bijgekomen. Sinds 1994 wordt er zowel in juni als december, in beperktere opzet, een actualisatie van de verwachtingen opgesteld en gepubliceerd. In het eerste deel van deze studie beperken we ons tot de analyse van de trefzekerheid van CEP en MEV.⁵ In paragraaf 5, waar de trefzekerheid van CPB-ramingen voor de Nederlandse economie wordt vergeleken met die van andere ramingsinstituten, nemen we ook de juni- en decemberramingen mee.

Het basisinstrument voor de korte-termijnramingen van het CPB is op dit moment het grootschalige macro-econometrische kwartaalmodel SAFFIER⁶, wat overigens geenszins betekent dat deze ramingen rechtsreeks uit dit model voortvloeien. De keuze om van een dergelijk model gebruik te maken komt voort uit het feit dat alleen ramingen die zijn gebaseerd op een grootschalig econometrisch model kunnen voldoen aan de eisen die de gebruikers aan de korte-termijnramingen van het CPB stellen, teneinde deze ramingen op een zinvolle wijze te kunnen gebruiken bij de sociaal-economische beleidsvoorbereiding.⁷

Wat zijn deze eisen? In de eerste plaats dienen de ramingen conceptueel helder en boekhoudkundig consistent te zijn. In de praktijk betekent dit dat de ramingen dienen aan te sluiten bij de data en definities van de Nationale rekeningen. In de tweede plaats dienen de ramingen inzicht te verschaffen in een groot aantal variabelen waarin de politiek en andere spelers in de sociaal-economische beleidsvoorbereiding geïnteresseerd zijn. Ten derde dient het mogelijk te zijn om op een consistente wijze de raming aan te vullen met onzekerheidsvarianten. Die varianten kunnen als doel hebben om de onzekerheidsmarges rondom de projectie te illustreren, maar kunnen ook zijn bedoeld om de economische effecten op korte termijn van beleidsvoorstellen in beeld te brengen. Ten slotte is het van groot belang

⁴ Dit extra ramingsjaar werd aanvankelijk apart gepubliceerd, maar is sinds 1992 integraal opgenomen in het CEP.

⁵ Voor de middellange termijn wordt er ook met enige regelmaat ook een 'verkenning' opgesteld. Deze heeft betrekking op een horizon van zo'n vier jaar en is meestal gekoppeld aan een nieuwe kabinetsperiode. De trefzekerheid van middellange-termijnverkenningen is besproken in Kranendonk en Verbruggen (2006c).

⁶ Zie Kranendonk en Verbruggen (2006b)

⁷ Voor een beoordeling van de verschillende modeltypen aan de hand van een aantal criteria waaraan een goede voorspelling moet voldoen, zie Don (2001, pag. 155-175). Recentelijk heeft het CPB een studie gedaan naar de voorspelkwaliteit van VAR-modellen in vergelijking met de gepubliceerde ramingen voor de Nederlandse economie, zie Elbourne (2008).

dat het CPB in staat is de ramingen zelf alsmede de verschillen met eerdere ramingen economisch te verklaren en te verantwoorden, wat het noodzakelijk maakt om de veronderstelde economische verbanden en de gebruikte exogenen te expliciteren. Zonder een uitgebreid economisch model is het niet mogelijk om ramingen te maken die aan alle bovengenoemde eisen voldoen.

Het macromodel wordt gevoed met data voor het verleden en projecties ten aanzien van het economisch beleid en de gang van zaken in het buitenland. De benodigde gegevens voor het verleden worden zoveel mogelijk ontleend aan het CBS. Voor zover voor het recente verleden nog geen realisatiecijfers beschikbaar zijn, maakt het CPB zelf ramingen. De beleidsveronderstellingen worden gebaseerd op de beleidsvoornemens van het kabinet, voor zover die bij het CPB bekend zijn. De buitenlandse economische ontwikkeling wordt door het CPB geraamd. Hierbij maakt het bureau onder andere gebruik van prognoses en analyses van internationale organisaties, zoals de Europese Commissie en de Organisatie voor Economische Samenwerking en Ontwikkeling (OESO), en zusterinstituten uit andere landen.

Ter verbetering van de CPB-prognoses wordt echter ook van diverse andere informatiebronnen gebruikgemaakt. De modeluitkomsten zijn geenszins 'heilig'; bij het maken van de prognoses realiseert het CPB zich terdege dat het gehanteerde model de gemiddelde economische situatie in het verleden weerspiegelt en dat de toekomst daar op sommige punten van kan afwijken. Daarom worden de modeluitkomsten altijd geconfronteerd met de inzichten van deskundigen op diverse terreinen, bijvoorbeeld op gebied van de investeringen en de overheidsinkomsten- en uitgaven. Ook wordt gebruikgemaakt van een stelsel van voorlopende conjunctuurindicatoren.⁸ Dit kan ertoe leiden dat de modeluitkomsten beargumenteerd worden aangepast.⁹ In het kader 'Bijsturing modelramingen levert meestal verbetering op' in paragraaf 4.1 wordt toegelicht dat recent onderzoek heeft uitgewezen dat deze bijstellingen kwantitatief niet ingrijpend zijn en voor de meeste variabelen tot een verbetering van de raming leidt.¹⁰

2.2 Selectie van de relevante CBS-realisaties

De trefzekerheid wordt beoordeeld door na te gaan of de ramingen ook zijn uitgekomen. Dat vereist dus een criterium als beoordelingsmaatstaf. Hier gaan we in paragraaf 2.3 nader op in. Wat er in werkelijkheid 'echt' gebeurd is, is op zich echter al een lastige vraag. Het CBS past de cijfers over het verleden namelijk regelmatig aan op basis van nieuwe informatie en nieuwe inzichten. In deze paragraaf geven we een overzicht van de informatie die het CBS publiceert

⁸ Zie Kranendonk, Bonenkamp en Verbruggen (2003).

⁹ Technisch gebeurt het verwerken van andere informatiebronnen door het bijsturen van de modeluitkomsten via autonome termen.

¹⁰ Zie Franses, Kranendonk en Lanser (2007).

en lichten we toe welke cijfers in deze studie zijn gehanteerd om de ramingen mee te vergelijken.

Door het CBS worden, behoudens revisies van de Nationale rekeningen, achtereenvolgens vijf schattingen gepubliceerd van de economische groei¹¹:

- Half februari na afloop van het jaar: ‘flashraming’ van het vierde kwartaal en een eerste schatting over het gehele voorbije jaar;
- Begin april na afloop van het jaar: reguliere Kwartaalrekeningen;
- Half juli na afloop van het jaar: Nationale rekeningen (‘voorlopig’);
- Half juli, anderhalf jaar na het verslagjaar: Nationale rekeningen (‘nader voorlopig’);
- Half juli, tweeënhalf jaar na het verslagjaar: Nationale rekeningen (‘definitief’).

Vijf opeenvolgende schattingen betekenen vier bijstellingen. Figuur 2.1 laat zien dat in de periode 1991-2005 de definitieve groei in twaalf gevallen hoger uitkwam dan de eerste flashraming en slechts drie maal gelijk of lager. Anders gezegd, het CBS heeft voor de meeste jaren de economische groei in eerste instantie onderschat. De gemiddelde totale opwaartse bijstelling van de groeicijfers op jaarbasis was 0,45%-punt.¹² De rechterfiguur in 2.1 geeft weer hoe de bijstelling is verdeeld over de vier aanpassingsmomenten. Het is daarbij opmerkelijk dat de grootste aanpassing gemiddeld genomen pas plaatsvindt bij de vaststelling van het definitieve cijfer. Overigens gebeurt het ook wel dat de groei tussentijds neerwaarts wordt herzien, maar slechts in vier van vijftien jaren is het definitieve cijfer niet tevens het hoogste cijfer dat voor een kalenderjaar is gepubliceerd. De gemiddelde absolute aanpassing is met 0,54%-punt vanzelfsprekend iets groter. De bijstellingen zijn in recentere jaren niet afgenomen: de aanpassing tussen de flash-raming en de definitieve raming bedroeg in de jaren 2003, 2004 en 2005 gemiddeld 1,0%-punt.¹³

In de gepresenteerde grafieken en tabellen in de volgende paragrafen worden de CPB-ramingen beoordeeld op basis van de indertijd gepubliceerde ‘definitieve’ CBS-cijfers.¹⁴ Deze werkwijze, die gerust als ‘streng’ kan worden gekwalificeerd, heeft in het algemeen een opwaarts effect op de ramingsfouten. Bij het maken van de ramingen spelen de voorlopige CBS-cijfers immers een belangrijke rol. Wanneer deze voorlopige cijfers naderhand worden bijgesteld, dan zijn bij het

¹¹ Zie Dankmeyer (2008).

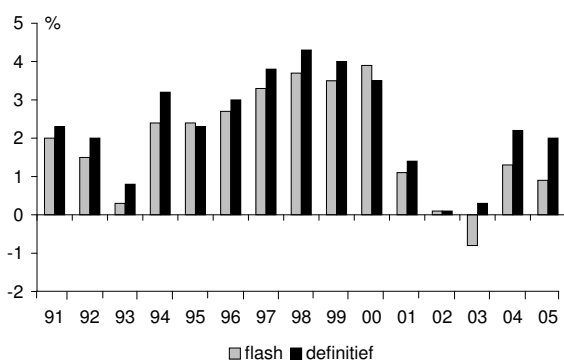
¹² Voor de periode 1991-2000 was de gemiddelde bijstelling volgens Nijmeijer en Hijman (2004) nog 0,35%-punt. Een uitgebreidere beschrijving van dat onderzoek is te vinden in Kazemier, Nijmeijer en Hijman (2004).

¹³ Tijdens het ramingsproces maakt het CPB gebruik van de ‘flashramingen’ en reguliere kwartaalramingen die het CBS publiceert over de meest recente kwartalen. De bijstellingen van afzonderlijke kwartaalcijfer zijn soms nog groter dan de besproken aanpassing van de jaarcijfers, omdat bij de jaarraming over- en onderschattingen van kwartalen binnen het jaar elkaar kunnen compenseren.

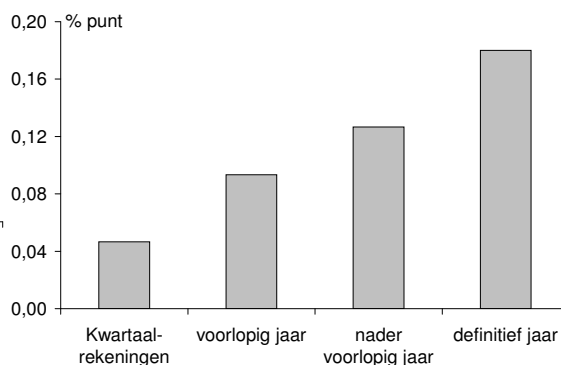
¹⁴ Voor de jaren 2006 en 2007 zijn nog geen definitieve cijfers beschikbaar. Voor deze jaren zijn dan ook respectievelijk de nader voorlopige en voorlopige CBS-cijfers gebruikt.

Figuur 2.1 Bijstelling economische groei

Economische groei: flash en definitief



Gemiddelde bijstelling economische groei ten opzichte van voorgaande schatting, 1991-2005



Bron: CBS.

ramen verkeerde realisatiecijfers gehanteerd, wat in het algemeen de trefzekerheid niet ten goede komt. Indien de ramingen worden vergeleken met voorlopige CBS-cijfers resulteren in het algemeen kleinere voorspelfouten voor de BBP-groei, zoals in paragraaf 3.1 in een kader wordt geïllustreerd.

Zowel door het CBS als het CPB zou overwogen kunnen worden de flashraming van de BBP-groei met een vaste opslag (van bijvoorbeeld 0,4%-punt) te verhogen. Dit lijkt geen goed idee, alleen al omdat de gemiddelde bijstelling in de komende perioden geenszins gelijk hoeft te zijn aan die in de afgelopen jaren. Wel ziet het CPB in de onzekerheid van de 'flashraming' reden om deze kwartaalcijfers niet altijd exact over te nemen. Het CPB kan en wil echter niet anticiperen op de aanpassingen die naderhand in de Nationale rekeningen worden doorgevoerd. Ter voorkoming van verwarring worden door het CPB de jaarcijfers uit de Nationale rekeningen wel ongewijzigd overgenomen.

Naast deze reguliere aanpassingen van de cijfers voor het recente verleden herzielt het CBS de cijfers ook nog vanwege grootschalige revisies. Dat is het gevolg van internationale afspraken over definities en meetmethoden van de economische grootheden die binnen het kader van de Nationale rekeningen worden gehanteerd. Circa eens in de vijf jaar moet het CBS daarom ook de cijfers ver terug aanpassen aan de internationale afspraken. Daarnaast kan het CBS op dergelijke momenten ook nieuwe methoden om de cijfers te berekenen doorvoeren. Deze aanpassingen van de cijfers zijn niet relevant voor ons onderzoek. De ramingen van het CPB

zijn immers gebaseerd op de op dat moment geldende systematiek van de Nationale rekeningen en pretenderen niet dergelijke revisies te voorspellen.¹⁵

2.3 Kengetallen om trefzekerheid te beoordelen

Om de trefzekerheid van ramingen te beoordelen zijn in de literatuur meerdere criteria voorhanden. Deze betreffen de zuiverheid, nauwkeurigheid en efficiëntie van voorspellingen. Enkele belangrijke criteria waar we in deze studie naar kijken, zijn:¹⁶

- a. Voorspellingen moeten 'unbiased' zijn: ramingen mogen niet systematisch over- of onderschatten;
- b. Voorspelfouten moeten klein zijn;
- c. De variantie van de voorspelfouten moet kleiner worden naarmate de voorspelhorizon korter wordt;
- d. Voorspellingen moeten een meerwaarde hebben boven alternatieve voorspellingen;
- e. Bijstellingen tussen opeenvolgende ramingen moeten onvoorspelbaar zijn;
- f. Voorspellingen moeten een versnelling of vertraging van de economische groei juist aangeven.

Concrete maatstaven om de criteria a, b en c te beoordelen zijn de gemiddelde voorspelfout, de gemiddelde absolute voorspelfout, de Root Mean Square Error en een statistische toets. Voor criterium d is de Theilcoëfficiënt beschikbaar. Voor criterium e is een specifieke statistische toets beschikbaar. Het laatste criterium kan worden beschouwd door na te gaan in hoeveel procent van de ramingen een versnelling of vertraging van de economische groei juist is aangegeven. Hieronder geven we aan welke beoordelingsmaatstaven we in deze studie gebruiken en hoe die zijn gedefinieerd.

2.3.1 Maatstaven voor criteria a, b en c

De *gemiddelde voorspelfout* (V) geeft voor een variabele i aan of sprake is geweest van systematische over- of onderschatting van een variabele. Deze grootheid wordt als volgt berekend:¹⁷

$$V_i = \frac{1}{n} \sum_{t=1}^n (P_{i,t} - R_{i,t})$$

¹⁵ In het algemeen hebben revisies een opwaarts effect op de groeicijfers. Voor de jaren 1990-2001 liggen de huidige CBS-groeicijfers voor het BBP-volume gemiddeld 0,25%-punt hoger dan de destijds gepubliceerde 'definitieve' cijfers. Voor een uitvoeriger bespreking van de problemen die revisies opleveren voor trefzekerheidsstudies wordt verwezen naar paragraaf 2.1 in Kranendonk en Verbruggen (2006a).

¹⁶ Deze zijn gebaseerd op Vogel (2007).

¹⁷ Met deze definitie volgen wij de OESO en de Europese Commissie. Andere instituten, waaronder het IMF, berekenen de voorspelfout als raming minus prognose.

met

n = aantal onderzochte ramingen

$P_{i,t}$ = prognose van variabele i voor periode t

$R_{i,t}$ = realisatie van variabele i voor periode t

Hoe dichter V_i bij nul ligt, des te kleiner is de gemiddelde over- of onderschatting geweest. We maken ook gebruik van een statistische toets om te zien of de voorspelfouten significant van nul afwijken. Als dit niet het geval is, is er dus geen sprake van een bias en kunnen we de raming *zuiver* noemen. We schatten hiervoor de volgende vergelijking: $V_{i,t} = \alpha_i + \varepsilon_{i,t}$. Vervolgens wordt op basis van een Wald-toets beoordeeld of α_i gelijk is aan nul.

Een gemiddelde voorspelfout van nul betekent echter nog niet dat alle afzonderlijke ramingen juist waren. Over- en onderschattingen kunnen elkaar immers compenseren. Daarom wordt ook de *gemiddelde absolute voorspelfout (A)* berekend. Die geeft aan wat de gemiddelde absolute afwijking is tussen prognose en realisatie:

$$A_i = \frac{1}{n} \sum_{t=1}^n |P_{i,t} - R_{i,t}|$$

Bij deze maatstaf worden grote en kleine fouten even zwaar gewogen. Internationaal wordt er echter vaak gebruik gemaakt van een andere maatstaf, waarbij grotere fouten zwaarder worden afgerekend, te weten de *Root Mean Square Error (RMSE)*. De RMSE wordt berekend als:

$$RMSE_i = \sqrt{\frac{1}{n} \sum_{t=1}^n (P_{i,t} - R_{i,t})^2}$$

Ook deze maatstaf kent zijn beperkingen. Een gemiddelde absolute voorspelfout van 1%-punt bij de investeringen wordt bijvoorbeeld even zwaar aangerekend als een even grote fout bij de particuliere consumptie, terwijl de investeringen veel grotere uitslagen van jaar op jaar laten zien dan de consumptie. De voorspelfout bij de investeringen is in dit geval dus relatief kleiner dan die bij de consumptie. Om hiervoor te corrigeren, berekenen we ook de R^2 . De R^2 schaalt de gemiddelde gekwadrateerde fout (=kwadraat van de RMSE) op de variantie van de onderliggende reeks en wordt als volgt berekend:

$$R_i^2 = 1 - \frac{\sum_{t=1}^n (P_{i,t} - R_{i,t})^2}{\sum_{t=1}^n (R_{i,t} - R_{gemiddeld})^2}$$

In de ideale situatie is R^2 gelijk aan 1. In dat geval is de RMSE namelijk nul, ofwel, de prognoses waren volledig gelijk aan de realisaties. Zolang de gemiddelde gekwadrateerde fout

kleiner is dan de variantie van de reeks, blijft de maatstaf groter dan nul. Als de voorspelfouten groter worden, kan de maatstaf negatief worden.

Voorts toetsen we de prognose op *efficiëntie*. Dit doen we door de gerealiseerde groei te regresseren op de prognose en een constante: $R_{i,t} = \alpha_i + \beta_i * P_{i,t} + \varepsilon_{i,t}$. Hierbij wordt getest of tegelijkertijd geldt dat $\alpha_i = 0$, $\beta_i = 1$ en dat de residuen *white noise*, ofwel niet serieel gecorreleerd, zijn. Als dit het geval is, dan kunnen we concluderen dat de raming alle beschikbare informatie benut. Zo niet, dan kunnen de ramingen nog verbeterd worden door bijvoorbeeld een constante bij de raming op te tellen, door de raming met een bepaalde factor te vermenigvuldigen of door gebruik te maken van het niet-willekeurige patroon van de residuen. Overigens is de eerder besproken toets op zuiverheid een bijzondere vorm van de toets op efficiëntie. Bij de toets op zuiverheid wordt op voorhand verondersteld dat $\beta_i = 1$.

2.3.2 Maatstaf voor criterium d

De hiervoor beschreven maatstaven beoordelen alleen de voorspelfouten zelf. Een andere manier van evalueren is om te kijken hoe de voorspelfout van de raming met de gehanteerde methode zich verhoudt tot de voorspelfout van een raming die op een alternatieve manier is opgesteld. Daarbij wordt als alternatief vaak gekozen voor een ‘naïeve’ raming: de realisatie van voorgaande periode wordt dan gebruikt als raming voor de huidige periode. Ook in deze trefzekerheidsstudie wordt deze naïeve raming als alternatieve raming gehanteerd. De maatstaf die hieruit volgt is de *Theil- of Ongelijkheidscoëfficiënt (OC)* en wordt als volgt berekend:

$$OC_i = \sqrt{\frac{\sum_{t=1}^n (P_{i,t} - R_{i,t})^2}{\sum_{t=1}^n (R_{i,t-1} - R_{i,t})^2}}$$

Hoe lager de ongelijkheidscoëfficiënt is, hoe beter. Indien alle ramingen precies juist zijn, is de ongelijkheidscoëfficiënt gelijk aan nul. Zolang de waarde maar onder de 1 blijft, heeft de werkelijke raming een meerwaarde boven de naïeve raming. Een waarde van 1 houdt in dat de voorspelfouten van de werkelijke en de naïeve raming een gelijke variantie hebben. De ongelijkheidscoëfficiënt kan zelfs boven 1 uitkomen.

2.3.3 Maatstaf voor criterium e

Is er systematiek waarneembaar in de bijstellingen van prognoses? Idealiter is dit niet het geval, immers, in dat geval waren de bijstellingen te voorspellen en deze informatie had meteen bij de eerste prognose meegenomen kunnen worden. Een eerste vereiste is dan ook dan de bijstellingen geen significante opwaartse of neerwaartse bias hebben. Dit testen we door de bijstellingen te regresseren op een constante, α_i , waarbij α_i gelijk aan nul is indien er geen bias is: $\Delta P_{i,t} = \alpha_i + \varepsilon_{i,t}$. Daarnaast zouden opeenvolgende bijstellingen van de raming voor jaar t

onafhankelijk van elkaar moeten zijn. Dit toetsen we door te onderzoeken of de coëfficiënt β_i in de vergelijking $\Delta P_{i,t}^f = \alpha_i + \beta_i \Delta P_{i,t}^{f-1} + \varepsilon_{i,t}$ gelijk is aan 0. Indien de geschatte waarde van β_i groter dan nul blijkt te zijn, dan wijst dit op het (te) geleidelijk doorvoeren van nieuwe inzichten. Een geschatte β_i kleiner dan nul wijst op te fors bijstellen van ramingen.

2.3.4 Maatstaf voor criterium f

In aanvulling op maatstaven voor de grootte van voorspelfouten en toetsen voor de zuiverheid en efficiëntie van prognoses, gaat het laatste criterium in op het juist voorspellen van vertragingen en versnellingen. De bijbehorende maatstaf is simpelweg het percentage van de gevallen waarin een vertraging of versnelling juist is voorspeld.

We spreken van een versnelling, indien de economische groei in een jaar hoger is dan in het voorgaande jaar. Omgekeerd, als de groei in een jaar lager ligt dan in het voorgaande jaar, is er sprake van een vertraging. Echter, gezien de onzekerheid van de groeicijfers zowel ten aanzien van de ramingen als de realisaties, lijkt een ondergrens voor het aanmerken van een groeiverandering als versnelling of vertraging nuttig. Een mutatie van bijvoorbeeld slechts 0,1%-punt wijst eerder op een stabiele groei dan op een groeiversnelling. Daarom maken we in deze studie onderscheid tussen een groeiversnelling c.q. -vertraging als de groeivoet van het ene op het andere jaar 0,5%-punt of meer verandert. Een kleinere mutatie kenschetsen we als stabiele groei.

We bekijken zowel voor ramingen als realisaties of er sprake is van versnelling, stabiele groei of vertraging. In theorie zijn dan negen combinaties denkbaar. Als exact gebeurt wat was voorspeld, kunnen we de richting van de voorspelling als 'juist' typeren. Als een versnelling c.q. vertraging is geraamd, terwijl de realisaties het tegenovergestelde aangeven, classificeren we raming als 'geheel onjuist'. De overige combinaties noemen we 'enigszins onjuist'. Tabel 2.1 vat de mogelijke combinaties samen.

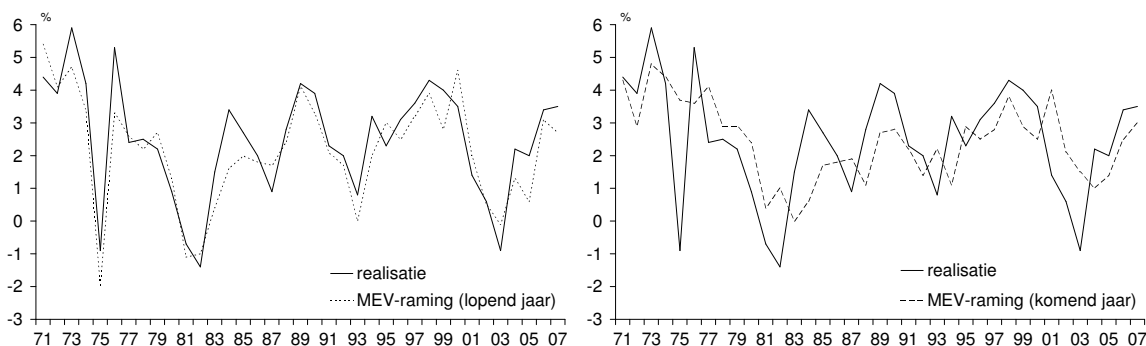
	Raming		
Realisatie	Versnelling	Stabiel	Vertraging
Versnelling	juist	enigszins onjuist	geheel onjuist
Stabiel	enigszins onjuist	juist	enigszins onjuist
Vertraging	geheel onjuist	enigszins onjuist	juist

3 Trefzekerheid korte-termijnramingen

3.1 Economische groei

De volumeontwikkeling van het BBP doet op dit moment dienst als dé maatstaf voor de economische groei. Daarom staat in deze paragraaf de trefzekerheid van de BBP-raming centraal. Figuur 3.1 geeft voor het lopende en komende jaar MEV-ramingen weer en vergelijkt deze met de realisaties.¹⁸ Daaruit blijkt dat de raming voor het lopende jaar beter is dan die voor het komende. Tevens is te zien dat voor sommige jaren de mate van op- en neergang fors is onderschat, met name voor het komende jaar.

Figuur 3.1 MEV-ramingen van de BBP-groei voor het lopende en komende jaar vergeleken met de realisaties, 1971-2007



Tabel 3.1 toont enkele maatstaven aan de hand waarvan de trefzekerheid van de CPB-prognoses kan worden beoordeeld. Daarbij worden twee perioden onderscheiden: de gehele analyseperiode (1971-2007) en de laatste tien jaar (1998-2007). Een vergelijking van deze perioden leert in hoeverre de trefzekerheid in het meer recente verleden beter of slechter is dan in de hele periode.

Gemiddeld heeft het CPB in de periode 1971-2007 in de MEV de BBP-groei voor het komend jaar goed voorspeld. Voor zowel de CEP- als MEV-raming voor het lopende jaar wordt de BBP-groei met 0,3%-punt onderschat. Dit hangt samen met het feit dat de voorlopige realisatiecijfers van het CBS naderhand gemiddeld opwaarts worden bijgesteld, zoals in paragraaf 2.2 is toegelicht. Voor de meest recente tienjaarsperiode komt de gemiddelde voorspelfout eveneens dicht bij nul uit. Deze resultaten suggereren dat bij de CEP- en MEV-ramingen geen sprake is van systematische onder- of overschatting. Een dergelijke uitkomst is in internationaal perspectief opmerkelijk, omdat voor veel landen de groeiramingen vaak systematisch te optimistisch zijn, zoals in paragraaf 6 wordt toegelicht. Als de CPB-ramingen

¹⁸ De voorspelfouten kunnen overigens nog wijzigen indien ook voor 2006 en 2007 de definitieve realisatiecijfers beschikbaar zijn.

Tabel 3.1 Trefzekerheid volumegroei van het bruto binnenlands product, 1971-2007

	Komend jaar		Lopend jaar	
	CEP	MEV	CEP	MEV
Gemiddelde voorspelfout^a				
1971-2007	-	0,0	- 0,3	- 0,3
1998-2007	0,1	0,1	- 0,1	- 0,3
Gemiddelde absolute voorspelfout				
1971-2007	-	1,2	1,0	0,7
1998-2007	1,4	1,2	1,0	0,8
RMSE				
1971-2007	-	1,5	1,2	0,8
1998-2007	1,6	1,4	1,2	0,9
R²				
1971-2007	-	0,5	0,6	0,7
1998-2007	0,4	0,5	0,6	0,7
Ongelijkheidscoëfficiënt				
1971-2007	-	0,8	0,7	0,4
1998-2007	0,9	0,8	0,7	0,6

^a Minteken duidt op onderschatting.

van de BBP-groei in het lopende jaar worden vergeleken met de voorlopige in plaats van definitieve NR-cijfers, dan is de gemiddelde onderschatting nagenoeg nul (zie kader).

In aanvulling op de gemiddelde voorspelfout, is ook de gemiddelde *absolute* voorspelfout relevant. Daarbij wordt gekeken naar de afwijking tussen raming en realisatie, los van de vraag of het om een over- of onderschatting gaat. De gemiddelde absolute ramingsfout van de BBP-groei blijkt af te nemen met het beschikbaar komen van meer informatie. Voor een komend jaar is de absolute fout voor de BBP-raming groter dan 1%-punt. Wanneer het jaar al deels gerealiseerd is (bij de MEV-raming), bedraagt de absolute fout nog zo'n 0,7%-punt. Hetzelfde patroon is te zien bij de RMSE, waarbij grotere fouten wat zwaarder worden meegewogen. De resultaten laten eveneens zien dat de trefzekerheid in de meest recente periode nagenoeg gelijk is aan die voor de gehele periode. Ook de R^2 -maatstaf, die corrigeert voor volatiliteit, geeft aan dat de trefzekerheid in de afgelopen tien jaar vergelijkbaar is gebleven.

De ongelijkheidscoëfficiënten liggen voor alle ramingsronden onder de één en laten een duidelijke verbetering zien met het beschikbaar komen van nieuwe informatie. Dat betekent dat de gepubliceerde ramingen geringere voorspelfouten hebben opgeleverd dan wanneer de realisatie van de voorafgaande periode was gebruikt als alternatieve ('naïeve') raming. De ongelijkheidscoëfficiënten voor de laatste tien jaar zijn vrijwel gelijk aan die voor de gehele analyseperiode.

Gevoeligheid gebruik 'realisatiecijfers' van het CBS voor trefzekerheid

In paragraaf 2.2 is toegelicht dat het CBS vijf realisatiecijfers op rij publiceert. De bijstellingen tussen deze cijfers zijn aanzienlijk. Bij het maken van de korte-termijnramingen maakt het CPB gebruik van de voorlopige cijfers zoals deze door het CBS naar buiten worden gebracht. Wanneer deze voorlopige cijfers naderhand worden bijgesteld, is het CPB van verkeerde cijfers uitgegaan, wat de trefzekerheid van de CPB-ramingen kan beïnvloeden. In de tabellen in de hoofdtekst van dit document zijn ter bepaling van de trefzekerheid de CPB-ramingen afgezet tegen de indertijd gepubliceerde definitieve schattingen. In de MEV is door het CPB de BBP-groei voor het lopende jaar in de periode 1971-2007 gemiddeld met 0,3%-punt onderschat (zie tabel 3.1). Bij deze raming beschikt het CPB over informatie van de flash-ramingen van het CBS over de eerste helft van lopende jaar. Uit de analyse in paragraaf 2.2 blijkt dat het CBS de cijfers gemiddeld genomen achteraf opwaarts bijstelt. Dat roept de vraag op of de gemiddelde voorspelfouten lager uitkomen wanneer de CPB-ramingen voor het lopend jaar uit de MEV worden afgezet tegen de voorlopige cijfers zoals het CBS deze een half jaar na afloop van het verslagjaar in de Nationale rekeningen publiceert.

Uit onderstaande tabel blijkt dat de gemiddelde voorspelfout van de BBP-groei nagenoeg nul is wanneer de CPB-ramingen worden vergeleken met de voorlopige cijfers van het CBS, terwijl vergeleken met de definitieve cijfers sprake is van een gemiddelde onderschatting. Ook de gemiddelde absolute voorspelfout komt beter uit. De gemiddelde opwaartse bijstelling van de BBP-cijfers door het CBS wordt blijkbaar niet door het CPB in de raming verdisconteerd. Het CPB lijkt beter in staat de voorlopige dan de definitieve realisatiecijfers van de BBP-groei te ramen. Daarnaast komt ook de gemiddelde (absolute) voorspelfout van de raming van de consumptiegroei dichterbij nul te liggen wanneer deze wordt afgezet tegen de voorlopige CBS-cijfers. Bij de goederenuitvoer en de investeringen van bedrijven neemt de gemiddelde absolute voorspelfout echter toe wanneer gekeken wordt naar voorlopige cijfers.

Trefzekerheid MEV-ramingen voor het lopend jaar, gerelateerd aan 'definitieve' en 'voorlopige' realisatiecijfers, 1971-2007

	Gemiddelde voorspelfout ^a		Gemiddelde absolute voorspelfout	
	T.o.v. 'voorlopig'	T.o.v. 'definitief'	T.o.v. 'voorlopig'	T.o.v. 'definitief'
Volume bestedingen en productie				
Bruto binnenlands product (%)	- 0,1	- 0,3	0,5	0,7
Goederenuitvoer (%)	0,1	- 0,1	1,6	1,5
Particuliere consumptie (%)	0,1	- 0,4	0,6	0,8
Investeringen bedrijven (%) ^b	- 1,0	- 0,5	3,0	2,9
Arbeidsmarkt				
Werkgelegenheid bedrijven (arb. jr.) (%)	- 0,3	- 0,3	0,4	0,5

^a Minteken duidt op onderschatting.

^b Exclusief woningen en voorraden.

Bij het maken van korte-termijnprognoses maakt het CPB veel en dankbaar gebruik van voorlopige CBS-cijfers. Toch worden door het CPB, in tegenstelling tot bijvoorbeeld De Nederlandsche Bank en de Europese Commissie, niet alle voorlopige cijfers van het CBS letterlijk overgenomen. Dit geldt vooral voor de cijfers uit de Kwartaalrekeningen met betrekking tot de economische ontwikkeling in recente kwartalen, die weliswaar serieus worden gezien, maar waarvan soms bewust wordt afgeweken. Pas wanneer het CBS een half jaar na afloop van het verslagjaar in de Nationale rekeningen jaarcijfers publiceert, worden deze onverkort overgenomen.

In hoeverre is het CPB in staat om een *versnelling* of *vertraging* van de economische groei goed te voorzien? Zoals eerder aangegeven onderscheiden we situaties met een BBP-groeimutatie van 0,5%-punt of meer, die we aanduiden als versnelling dan wel vertraging, en situaties met een min of meer stabiele groei. Tabel 3.2 laat zien dat in 40 tot 50% van de gevallen de groeimutatie voor het komend jaar juist is voorspeld. Dat percentage neemt toe tot ongeveer 60% voor zowel de CEP- als de MEV-raming in het lopende jaar. Het percentage ‘geheel onjuist’ is voor de eerste raming nog 22%, maar voor latere ramingen beduidend lager.¹⁹

Tabel 3.2 Trefzekerheid groeiversnellingen/-vertragingen, 1971-2007

	Ramingen	Groeiversnellingen/-vertragingen		
		% juist	% enigszins onjuist	% geheel onjuist
CEP komend jaar (1990-2007)	18	50	28	22
MEV komend jaar	37	41	52	7
CEP lopend jaar	37	62	27	11
MEV lopend jaar	37	59	38	3

Ten slotte bekijken we of er sprake is van systematiek in de bijstellingen van de ramingen. Voor de jaren 1971-1989 zijn er voor elk jaar drie CEP- of MEV-ramingen, en dus twee bijstellingen, beschikbaar: de bijstelling tussen de CEP-raming voor het lopend jaar en de daaraan voorafgaande MEV-raming voor volgend jaar én die tussen de MEV-raming voor het lopend jaar en de daaraan voorafgaande CEP-raming voor het lopend jaar. Vanaf 1990 wordt er vier keer een CEP- of MEV-prognose voor een bepaald jaar gemaakt. Dat betekent dat vanaf dat jaar ramingen drie keer (kunnen) worden herzien.

Op- en neerwaartse bijstellingen wisselen elkaar willekeurig af. In de jaren 1971-1989 werd in 7 van de 19 gevallen de raming voor een bepaald jaar tweemaal in dezelfde richting bijgesteld, terwijl in de overige 12 ramingen een opwaartse bijstelling werd gevolgd door een neerwaartse of omgekeerd. In de periode 1990-2007, met drie mogelijke bijstellingen van de ramingen, was voor 3 van 17 onderzochte jaren sprake van drie bijstellingen in dezelfde richting: naar boven voor het jaar 1996 en neerwaarts voor de jaren 1993 en 2005. Voor de overige ramingsjaren wisselen op- en neerwaartse aanpassingen en ongewijzigde ramingen elkaar willekeurig af.

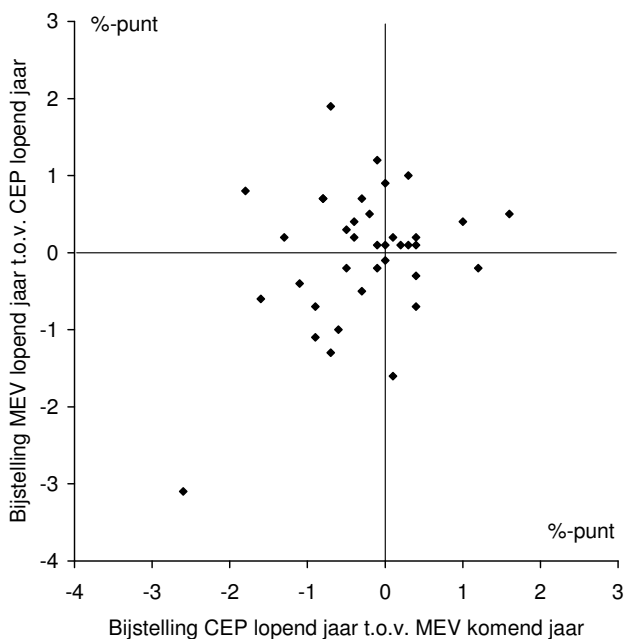
Van alle afzonderlijke aanpassingen was ongeveer 52% in neerwaartse en 48% in opwaartse richting. Gemiddeld werd de raming stapsgewijs iets neerwaarts aangepast. De gemiddelde bijstelling tussen de CEP-raming voor het komend jaar en de MEV-raming voor het komend jaar bedraagt -0,12%-punt. Tussen de MEV-raming voor het komende jaar en de CEP-raming

¹⁹ De percentages juiste voorspellingen nemen toe wanneer een lagere ondergrens om een groeimutatie aan te duiden als versnelling of vertraging wordt gekozen. In het extreme geval, ofwel een ondergrens van 0%-punt, wordt elke groeimutatie gezien als versnelling of vertraging. In dat geval wordt in het CEP voor het komend jaar in 67% van de gevallen een versnelling of vertraging juist voorspeld. Voor de ramingen MEV komend jaar, CEP lopend jaar en MEV lopend jaar liggen deze percentages op respectievelijk 62, 73 en 76%.

voor het lopend jaar was de wijziging van de raming gemiddeld forser: $-0,28\%$ -punt. De in paragraaf 2.3.3 beschreven statistische toets die test of de geschatte waarde van α in de vergelijking $\Delta P_t = \alpha + \varepsilon_t$ significant van nul verschilt, levert een p-waarde op van 0,044, wat betekent dat deze bijstelling, ten aanzien van het veel als kritieke grens gebruikte significantieniveau van 5% (ofwel een p-waarde van 0,05), significant negatief is. Overigens waren de herzieningen met name in het verleden relatief groot: tussen 1971-1989 bedroeg de gemiddelde bijstelling $-0,42\%$ -punt (met een p-waarde van 0,059), terwijl deze in de jaren 1990-2007 slechts $-0,12\%$ -punt was (p-waarde 0,441). De derde bijstelling, tussen CEP- en MEV-raming voor het lopend jaar, is gemiddeld $-0,02\%$ -punt en verschilt niet significant van nul.

Om de relatie tussen bijstellingen te onderzoeken, is voor de gehele periode 1971-2007 de vergelijking $\Delta P_t = \alpha + \beta \Delta P_t^{t-1} + \varepsilon_t$ geschat. Dit houdt, gegeven de beschikbare data, in dat de bijstelling in de MEV-raming voor het lopend jaar wordt geresseerd op een constante en op de bijstelling die reeds in het CEP daarvoor had plaatsgevonden. De geschatte waarde van α verschilt niet significant van nul, maar voor β wordt een waarde van 0,36 geschat (p-waarde 0,045). Aangezien deze groter dan nul is, suggereert dit resultaat dat het CPB in de onderzochte periode zijn ramingen gemiddeld iets te langzaam heeft bijgesteld. Het gevonden resultaat hangt echter nauw samen met de waarneming voor 1975, het punt linksonder in het spreidingsdiagram dat in figuur 3.2 wordt getoond. De raming voor dat jaar werd tweemaal met meer dan 2,5%-punt neerwaarts bijgesteld. Wanneer deze observatie, die als outlier gekwalificeerd kan

Figuur 3.2 Spreidingsdiagram van ramingsbijstellingen, 1971-2007



worden²⁰, wordt weggelaten, resteert een schatting voor β van 0,07. Deze verschilt niet significant van nul.

3.2 Andere economische grootheden

In de vorige paragraaf hebben we ons beperkt tot de volumegroei van het BBP. Kleine fouten bij het voorspellen van de BBP-groei kunnen echter het gevolg zijn van grotere fouten bij de componenten waaruit het BBP is opgebouwd. In tabel 3.3 zijn voor de MEV-raming voor het komende jaar de uitkomsten van de drie meest gebruikte maatstaven opgenomen voor een grotere set variabelen, waaronder de belangrijke uitgangspunten bij de ramingen omtrent de internationale conjunctuur ('externe data'), de ramingen van enkele relevante prijzen en de werkgelegenheid. Vergelijkbare tabellen van de drie andere ramingsmomenten zijn te vinden in bijlage 2.

De gemiddelde voorspelfout van ongeveer nul bij het BBP blijkt het saldo te zijn van overschattingen van de uitvoer, de invoer en de investeringen en een onderschatting van de particuliere consumptie. Voor wat betreft de goederenhandel hangt dat in belangrijke mate samen met de overschatting van de wereldhandel. Aan de prijzenkant is er vooral bij de internationale handel sprake van een onderschatting van de prijsontwikkelingen door niet-geraamde fluctuaties in wisselkoersen en/of grondstoffenprijzen, waaronder de olieprijs. De absolute voorspelfouten zijn bij de uit- en invoer en de investeringen beduidend groter dan bij het BBP en de particuliere consumptie. Dat hoeft op zich niet te verbazen, omdat de volatiliteit van deze reeksen veel groter is. De maatstaf die corrigeert voor de volatiliteit (R^2) vertoont een aanzienlijk kleinere spreiding en ligt voor deze vijf variabelen tussen de 0,3 tot 0,5. De fouten bij de binnenlandse prijzen zijn relatief gezien veel geringer. De R^2 ligt in de buurt van de 0,8, wat omgerekend betekent dat de voorspelfouten slechts 20% van de volatiliteit zijn. De ongelijkheidscoëfficiënten wijzen vooral op problemen bij de particuliere consumptie, zeker voor de meer recente jaren. Dit hangt met name samen met aanzienlijke onderschattingen in de periode van hoogconjunctuur (rond 1999-2000) en de overschattingen in de jaren daarna (2001-2003). In beide perioden zijn de vermogenseffecten onderschat, eerst bij de stijgende huizen- en aandelenprijzen en later bij de gevolgen van de daling van de aandelenkoersen.²¹

²⁰ De gestandaardiseerde waardes (berekend als (observatie - gemiddelde bijstelling) / (standaarddeviatie van de bijstellingen)) van de twee bijstellingen voor 1975 bedragen 2,9 voor de eerste bijstelling en 3,5 voor de tweede.

²¹ Voor een uitvoeriger analyse van de voorspelfouten bij de particuliere consumptie, zie Kranendonk en Verbruggen (2006a, pag. 22). Daar wordt ook toegelicht hoe de modellering van vermogenseffecten in de consumptiefunctie is aangepast.

Tabel 3.3 Trefzekerheid van de MEV-ramingen voor het komende jaar, 1971-2007

	Gemiddelde voorspelfout ^a		Gem. absolute voorspelfout		Ongelijkheidcoëfficiënt	
	1971-2007	1998-2007	1971-2007	1998-2007	1971-2007	1998-2007
Externe data						
Relevant wereldhandelsvolume (%)	0,3	0,7	2,6	2,7	0,6	0,8
Prijs relevante wereldhandel (%)	0,5	- 0,9	4,1	2,5	0,7	0,7
Prijs goedereninvoer (%)	- 0,4	- 1,6	4,7	2,5	0,7	0,8
Lonen en prijzen						
Contractloon marktsector (%)	- 0,2	0,0	0,9	0,3	0,9	0,4
Consumentenprijsindex (CPI) (%)	- 0,2	- 0,3	0,8	0,5	0,7	0,6
Prijs uitvoer goederen (%)	- 0,3	- 1,4	4,0	2,3	0,8	0,7
Prijs bruto binnenlands product (%)	- 0,2	- 0,3	0,8	0,8	0,6	0,6
Volume bestedingen en productie						
Bruto binnenlands product (%)	0,0	0,1	1,2	1,2	0,8	0,8
Goederenuitvoer (%)	0,6	0,7	3,4	3,3	0,8	0,8
Goedereninvoer (%)	0,5	0,5	3,3	3,1	0,7	0,9
Particuliere consumptie (%)	- 0,4	- 0,1	1,4	1,6	1,0	1,1
Investerings bedrijven (%) ^b	0,1	1,0	4,0	3,1	0,6	0,8
Arbeidsmarkt						
Werkgelegenheid bedr. (arb. jr.) (%)	- 0,2	- 0,1	0,9	0,8	0,8	0,6

^a Minteken duidt op onderschatting.

^b Exclusief woningen en voorraden.

3.3 Zijn de ramingen zuiver en efficiënt?

Is er sprake van systematische voorspelfouten bij het BBP en de andere variabelen? In tabel 3.3 zijn de gemiddelde (absolute) voorspelfouten van de MEV-raming voor het komende jaar weergegeven. Mag daaruit bijvoorbeeld worden afgeleid dat het CPB de prijsontwikkeling systematisch te laag inschat en de buitenlandse handel systematisch te hoog? Of liggen de kengetallen zo dicht bij nul dat, statistisch gezien, de ramingen niet op een systematische onder- of overschatting wijzen? Keereman (2003) presenteert twee in dit kader relevante toetsen.

De eerste toets heeft betrekking op de zuiverheid van een voorspelling en zegt iets over de vraag of op statistische gronden moet worden geconcludeerd dat de gemiddelde voorspelfout afwijkt van nul. Als de kans dat de voorspelfouten afkomstig zijn uit een proces met een gemiddelde van nul kleiner dan 5% is, moet geconcludeerd worden dat de raming 'onzuiver' is. In dat geval is sprake van een systematische onder- of overschatting. Uit onderstaande tabel blijkt dat voor alle onderzochte variabelen de gemiddelde voorspelfout niet significant van nul verschilt.

De tweede toets analyseert de efficiency van de voorspellingen door de realisatie te regresseren op de raming en een constante: $R_{i,t} = \alpha_i + \beta_i * P_{i,t} + \varepsilon_{i,t}$. Efficiëntie vereist dat α_i gelijk is aan nul, terwijl β_i gelijk aan 1 moet zijn. Bovendien dienen de residuen niet serieel gecorreleerd te zijn. Als aan deze vereisten is voldaan, is het niet mogelijk de raming op basis

van de beschikbare informatie te verbeteren. Een waarde van $\beta_i = 0$ zou daarentegen betekenen dat de raming geen enkele informatieve waarde heeft. In feite is de test op efficiëntie een algemene versie van de eerdergenoemde toets op zuiverheid, toen impliciet werd verondersteld dat β_i een waarde van 1 had. Als de raming niet efficiënt is, betekent dit dat er een samenhang bestaat tussen de voorspelling en de voorspelfout. Indien een dergelijke samenhang bestaat, zou de voorspelling verbeterd kunnen worden door hiermee rekening te houden. In tabel 3.4 is weergegeven hoe waarschijnlijk het is dat de voorspelfouten afkomstig zijn uit een proces waarin aan beide eisen, $\alpha_i = 0$ en $\beta_i = 1$, is voldaan. Als die kans kleiner is dan 5%, concluderen we dat de voorspeller niet efficiënt is. Uit onderstaande tabel blijkt dat dit alleen het geval is bij de consumptieraming. Voor de overige variabelen kan de nulhypothese die stelt dat de ramingen efficiënt zijn, niet verworpen worden.

Tabel 3.4 Zuiverheid en efficiency van de MEV-ramingen voor komend jaar, 1971-2007

	Zuiverheid	Zuiverheid en efficiency
Specificatie	$P_{i,t} - R_{i,t} = \alpha_i + \varepsilon_{i,t}$	$R_{i,t} = \alpha_i + \beta_i * P_{i,t} + \varepsilon_{i,t}$
Nulhypothese	$\alpha = 0$	$\alpha = 0$ en $\beta = 1$
Kans dat voorspelfouten komen uit een proces waarin de nulhypothese geldt^a		
Relevant wereldhandelsvolume (%)	0,59	0,24
Prijs relevante wereldhandel (%)	0,60	0,62
Prijs goedereninvoer (%)	0,74	0,91
Contractloon marktsector (%)	0,36	0,16
Consumentenprijsindex (CPI) (%)	0,22	0,21
Prijs uitvoer goederen (%)	0,80	0,96
Prijs bruto binnenlands product (%)	0,18	0,41
Bruto binnenlands product (%)	0,87	0,42
Goederenuitvoer (%)	0,39	0,13
Goedereninvoer (%)	0,43	0,28
Particuliere consumptie (%)	0,18	0,00
Investeringen bedrijven, exclusief woningen en voorraden (%)	0,90	0,48
Werkgelegenheid bedrijven (arbeidsjaren) (%)	0,20	0,12

^a Op basis van de Wald-toets (p-waarde).

4 Analyse voorspelfouten

Een economische raming schetst een toekomstbeeld van de economie en is zodoende per definitie onzeker. Dat blijkt ook wel uit de hiervoor besproken resultaten omtrent de trefzekerheid. In paragraaf 2.1 is besproken dat CPB-ramingen worden opgesteld met behulp van een macro-econometrisch model. De verschillende onderdelen van het ramingsproces dragen ieder bij aan de onzekerheid van de ramingen. Om een beter beeld te krijgen van de bijdrage van de diverse onderdelen, is onlangs een gevoeligheidsanalyse uitgevoerd met behulp van stochastische simulaties.²² Dergelijk onderzoek geeft inzicht in de kwantitatieve betekenis van de factoren die mede verantwoordelijk zijn voor de onzekerheid. Dat inzicht kan helpen bij de prioritering voor toekomstig onderzoek.

4.1 Onzekerheid in CPB-ramingen: een uitsplitsing in determinanten

Uit de eerdere beschrijving van het ramingsproces valt af te leiden dat er bij het maken van een CPB-raming vier bronnen van onzekerheid zijn. Deze vier bronnen worden hier achtereenvolgens besproken.

De eerste bron betreft de onzekerheid in de voorlopige data uit de Nationale rekeningen. Voordat het CBS de definitieve cijfers over een kalenderjaar publiceert, neemt het CPB de voorlopige cijfers reeds op als uitgangspunt van zijn raming. Dit beïnvloedt de trefzekerheid van de raming op twee manieren. Allereerst betekent een bijstelling dat er sprake is van een andere realisatie dan waarvan bij het maken van de raming is uitgegaan, zoals besproken is in paragraaf 3.1. Dat heeft een direct effect via de overloop van het gerealiseerde jaar naar het eerste ramingsjaar. Daarnaast hebben andere realisaties ook een indirect effect op de ramingen van andere variabelen, vanwege de vertragingen die een grote rol spelen in de vergelijkingen in het macromodel waarmee de ramingen worden gemaakt. Een voorbeeld daarvan is de vraag naar arbeid en kapitaal (investeringen), die afhangen van de productie in het vorige jaar.

De tweede bron van onzekerheid, de exogene variabelen, beschrijft het onzekere toekomstbeeld van economische variabelen buiten het model, zoals de wereldhandel en de olieprijs. De invloed van dergelijke buitenlandexogenen op de raming is onderzocht. In paragraaf 4.2 wordt hier dieper op ingegaan. Daarnaast zijn er ook nog exogenen die betrekking hebben op het binnenland, waaronder het (voorgenomen) overheidsbeleid. Als de regering het beleid wijzigt, door bijvoorbeeld extra te bezuinigen of een aangekondigde belastingmaatregel uit te stellen, zal dat meestal ook resulteren in een aanpassing van de raming voor het economisch beeld. Om de onzekerheid rondom het beleid in beeld te brengen, heeft het CPB beleidsvarianten opgesteld die aangeven wat effecten zijn van een aantal specifieke

²² Zie Lanser en Kranendonk (2008).

maatregelen.²³ Naast de beleidsexogenen bestaan er ook andere binnenlandexogenen, zoals het arbeidsaanbod, de afzet van aardgas of de consumptie van woningdiensten.

Voorts is er onzekerheid die samenhangt met het model dat gebruikt wordt. De modelparameters zijn veelal geschat over het verleden. De exacte waarde is al onzeker en bovendien is het de vraag of de geschatte waarde ongewijzigd toegepast mag worden voor de toekomst. Daarnaast geven de residuen van de gedragsvergelijkingen in het model een indicatie van de overige onzekerheid in het model, die bijvoorbeeld het gevolg is van een onvolledige weergave van de werkelijkheid. Een model is immers altijd een gestileerde weergave van de realiteit. De residuen meten in welke mate de vergelijkingen niet in staat zijn de werkelijkheid, afgemeten aan de realisaties van het CBS, te verklaren. De onzekerheid in de modelparameters en in de residuen worden in tabel 4.1 samengenomen onder de noemer ‘modelonzekerheid’.

De vierde en laatste bron van onzekerheid wordt gevormd door de ‘expert opinion’. Voor de toekomst worden de modelramingen soms aangepast. In Don en Verbruggen (2006) is toegelicht waarom dat internationaal gebruikelijk is. Onderstaand kader gaat hier op in. In het kader worden ook de uitkomsten van onderzoek van Franses, Kranendonk en Lanser (2007) naar de omvang van bijstellingen van modeluitkomsten en het succes ervan getoond.

Lanser en Kranendonk (2008) hebben, om een beter beeld te krijgen van de bijdrage van de diverse bronnen van onzekerheid aan de voorspelfout bij diverse variabelen, een gevoeligheidsanalyse uitgevoerd met behulp van stochastische simulaties. Ze hebben daarbij niet alleen de korte-termijnramingen onderzocht, maar ook de ramingen voor vier jaar vooruit. Deze ramingen worden gemaakt wanneer er een verkenning voor de middellange termijn wordt opgesteld. Tabel 4.1 geeft weer in welke mate de drie hiervoor genoemde bronnen bijdragen aan de onzekerheid omtrent de prognoses voor de belangrijkste macro-economische variabelen.²⁴ Het aandeel van de verschillende bronnen van onzekerheid in de voorspelfout wordt gemeten aan de hand van hun aandeel in de totale variantie van deze fout.

Voor de meeste variabelen zijn de uitgangspunten over de exogene variabelen de belangrijkste bron van onzekerheid. In het geval van exportprijzen en het uitvoervolume heeft deze bron van onzekerheid zelfs een aandeel van zo’n 90% in de totale variantie, waarbij vooral de buitenlandexogenen relevant zijn. De onzekerheid in de voorlopige data vormt vooral voor de korte-termijnraming van de investeringen een belangrijke bron van onzekerheid. Het aandeel van de modelonzekerheid is voor meerdere variabelen belangrijk, maar met name voor de werkgelegenheid, het contractloon en de particuliere consumptie.

De resultaten suggereren dat de grootste winst wat betreft nauwkeurigheid van de CPB-voorspellingen te halen is bij het bepalen van de exogene variabelen. Helaas is dit gemakkelijker gezegd dan gedaan.

²³ Voorbeelden van beleidsvarianten zijn te vinden in hoofdstuk 5 van Kranendonk en Verbruggen (2006b).

²⁴ Voor een verantwoording van dit onderzoek, zie Lanser en Kranendonk (2008).

Bijsturing modelramingen levert meestal verbetering op

Om de kwaliteit van de ramingen te verhogen, worden modeluitkomsten soms aangevuld op basis van expert-informatie. Deze bijsturingen komen tot stand met behulp van experts die nieuwe informatie over de economische ontwikkeling verwerken in de ramingen. Voorbeelden van modelbijsturing zijn het verwerken van cao-afspraken over de loonontwikkeling en het opnemen van gegevens rond de oplevering van grote investeringsprojecten bij de industrie of de transportsector. Bijsturing kan ook nodig zijn als onderdelen van het model niet langer representatief zijn voor de ramingsperiode.^a

Het CPB stuurt slechts in beperkte mate zijn modellen bij. De gevolgen van het bijsturen zijn onderzocht voor de periode 1997-2006.^b Voor acht van de elf vergelijkingen waarbij is ingegrepen, is de aanpassing gemiddeld minder dan 0,5%-punt geweest. Alleen bij enkele kleinere onderdelen van de in- en uitvoer van goederen waren de aanpassingen gemiddeld groter. De grootste absolute bijsturingen vonden plaats bij de investeringen en de wederuitvoer.

Met name bij de raming van de prijsontwikkelingen heeft deze bijsturing van de voorlopige modeluitkomsten de gemiddelde voorspelfouten dichter bij nul gebracht. Het verschil tussen de uiteindelijk gerealiseerde cijfers en de gepubliceerde ramingen in die periode is voor de meeste variabelen kleiner dan zonder aanpassing het geval zou zijn geweest. Zoals blijkt uit de tabel, zijn in de periode 1997-2006, afgemeten aan de vermindering van de gemiddelde voorspelfouten, de grootste verbeteringen gerealiseerd bij de raming van de investeringen en het contractloon.

Verder blijkt de gemiddelde absolute voorspelfout van de niet-bijgestuurde modeluitkomsten het grootst bij de volumegroei van de investeringen en het kleinst bij de inflatie (CPI). Bijsturing resulteerde meestal in een afname van de gemiddelde absolute voorspelfouten, oftewel in een verbetering van de trefzekerheid, vooral bij de contractloonontwikkeling en de inflatie. Alleen voor de investeringen geldt dat de absolute voorspelfout is toegenomen. Dat betekent dat, achteraf bezien, de trefzekerheid van de investeringsraming iets beter was geweest als er niet was bijgestuurd.

Voorspelfouten van bijgestuurde en niet-bijgestuurde modeluitkomsten, 1997-2006

	Gemiddeld		Gemiddeld absoluut	
	Niet-bijgestuurd	Bijgestuurd	Niet-bijgestuurd	Bijgestuurd
Volumeontwikkeling (%)				
BBP	0,1	0,3	1,2	1,2
Export	1,0	0,7	3,4	3,1
Consumptie	0,0	0,2	2,0	1,8
Investerings	-2,2	-1,0	4,7	4,9
Prijsontwikkeling (%)				
Contractloon	-1,0	-0,1	1,2	0,4
Consumentenprijsindex (CPI)	-0,8	-0,4	1,1	0,7
Exportprijs	-1,8	-1,7	2,8	2,7

^a Don en Verbruggen (2006).

^b Franses, Kranendonk en Lanser (2007).

Tabel 4.1 Aandeel van de bronnen van onzekerheid in de variantie van voorspelfouten, in %

	Voorlopige data		Exogene variabelen		Modelonzekerheid	
	Jaar 1	Jaar 4	Jaar 1	Jaar 4	Jaar 1	Jaar 4
BBP (% , volume)	15	5	73	83	12	12
Particuliere consumptie (% , volume)	15	5	38	66	47	29
Investerings (% , volume)	55	16	16	44	29	40
Goederenuitvoer (% , volume)	10	3	84	92	6	5
Werkgelegenheid (% , arb. jaren)	32	2	4	56	64	42
Consumentenprijsindex (%)	24	9	48	53	28	38
Goederenuitvoer prijs (%)	4	2	93	94	3	4
Contractloon (%)	18	12	28	35	54	53

4.2 Betekenis buitenlandexogenen voor de trefzekerheid

In de vorige paragraaf is aangegeven dat de uitgangspunten over de buitenlandexogenen een belangrijke verklarende factor voor de mate van trefzekerheid zijn. In deze paragraaf gaan we daar nader op in. Om de invloed van de buitenlandexogenen op de voorspelfouten te illustreren, zijn alternatieve ramingen opgesteld die gebaseerd zijn op de daadwerkelijk gerealiseerde ontwikkeling van twee belangrijke exogenen, te weten de wereldhandelsgroei en de wisselkoers. Deze alternatieve ramingen zijn samengesteld door de oorspronkelijke raming te amenderen met behulp van standaardvarianten voor de effecten van wijzigingen in de wereldhandel en de wisselkoers.²⁵

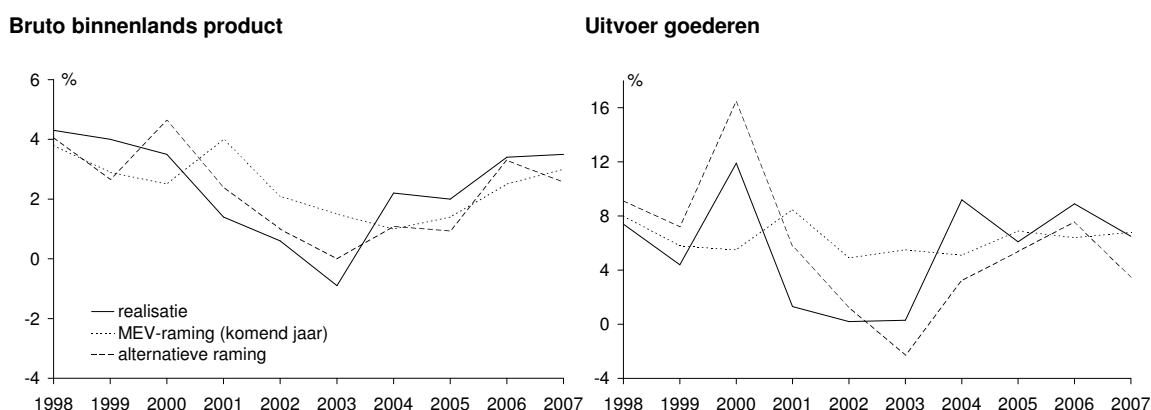
In figuur 4.1 zijn voor de groei van het BBP en de uitvoer van goederen de oorspronkelijke raming voor het komend jaar, de alternatieve raming voor het komend jaar en de realisaties weergegeven voor de periode 1998-2007. Hieruit blijkt dat in de meeste jaren de alternatieve raming van de uitvoergroei (veel) dichterbij de realisatie ligt dan de officiële raming. Voor de periode 1998-2007 ligt de gemiddelde absolute voorspelfout van de alternatieve raming 0,5%-punt onder die van de gepubliceerde raming: 2,8%-punt vs. 3,3%-punt. Ook met betrekking tot de BBP-groei scoort de alternatieve raming beter. Voor deze grootte wordt de gemiddelde absolute voorspelfout met een derde gereduceerd, van 1,2%-punt tot 0,8%-punt. De resterende voorspelfout in de periode 1998-2007 hangt in belangrijke mate samen met ramingsfouten bij de particuliere consumptie in jaren met grote vermogens-effecten.²⁶

Om een beter inzicht te krijgen in de betekenis van goede buitenlandexogenen voor de trefzekerheid is onderzocht of er een relatie is tussen de voorspelfouten van een tweetal buitenlandexogenen, namelijk de groei van de relevante wereldhandel en de ontwikkeling van

²⁵ Deze standaardvarianten zijn ontleend aan Kranendonk en Verbruggen (2006b)

²⁶ Een uitvoeriger analyse van voorspelfouten in de periode 1998-2003 is opgenomen in paragraaf 3 van Kranendonk en Verbruggen (2005).

Figuur 4.1 Realisaties, MEV-ramingen en alternatieve ramingen, 1998-2007



de invoerprijs, en enkele kerngrootheden van de Nederlandse economie. Daarbij wordt de voorspelfout bij de groei van de wereldhandel voor het komend jaar gekoppeld aan enkele reëel-economische variabelen, terwijl voor de voorspelfout van de ontwikkeling van de invoerprijs, zowel die voor het lopend als het komend jaar, wordt bekeken wat de samenhang is met de voorspelfout van diverse prijsvariabelen. De geschatte samenhang staat vermeld in tabel 4.2.

Deze schattingen bevestigen dat foutieve ramingen van de wereldhandelsgroei een aanzienlijke en significante factor zijn bij de verklaring van ramingsfouten van de volumeontwikkelingen. De ramingsfouten bij de wereldhandelsgroei kunnen ongeveer 70% van

Tabel 4.2 Samenhang tussen ramingsfouten van buitenlandvariabelen (komend en lopend jaar) en kerngrootheden van de Nederlandse economie (komend jaar), 1971-2007^a

Samenhang met ramingsfout van relevante wereldhandelsgroei in het:	Komend jaar	Constante	R^2_{gec}	
Volume bestedingen en productie				
Bruto binnenlands product (%)	0,33 (5,7)	- 0,14 (0,7)	0,48	
Goederenuitvoer (excl. energie) (%)	1,17 (9,9)	0,30 (0,8)	0,74	
Goedereninvoer (%)	1,06 (8,7)	0,23 (0,6)	0,68	
Particuliere consumptie (%)	0,22 (2,7)	- 0,44 (1,7)	0,17	
Investeringsbedrijven (%) ^b	0,57 (2,3)	- 0,06 (0,1)	0,13	
Arbeidsmarkt				
Werkgelegenheid bedrijven (arb. jr.) (%)	0,13 (2,9)	- 0,24 (1,7)	0,19	
Samenhang met ramingsfout van invoerprijsontwikkeling in het:	Komend jaar	Lopend jaar	Constante	R^2_{gec}
Lonen en prijzen				
Contractloon marktsector (%)	0,09 (2,8)	0,23 (1,3)	- 0,27 (1,3)	0,29
Consumentenprijsindex (CPI) (%)	0,07 (3,3)	0,15 (1,2)	- 0,24 (1,6)	0,34
Prijs uitvoer goederen (%)	0,85 (33,1)	-	0,07 (0,4)	0,97
Prijs bruto binnenlands product (%)	0,06 (2,6)	0,20 (1,5)	- 0,29 (1,8)	0,29

^a Tussen haakjes staan de absolute t-waarden.

^b Exclusief woningen en voorraden.

ramingsfouten bij de in- en uitvoergroei van goederen verklaren. Maar ook voor de CPB-ramingsfouten van de BBP-groei bedraagt dit percentage nog bijna 50%. De geschatte coëfficiënten komen nagenoeg overeen met de met het model SAFFIER berekende effecten op de betreffende variabelen van een wereldhandelsimpuls.²⁷

In de sfeer van lonen en prijzen is er een duidelijk verband met de voorspelfouten van de invoerprijzen. Met name de voorspelfout van de uitvoerprijsontwikkeling voor het komend jaar hangt zeer sterk samen met die van de invoerprijsontwikkeling. Het is opmerkelijk dat de trefzekerheid van de MEV-ramingen voor de contractloonstijging, de CPI en de BBP-prijsontwikkeling in het komende jaar sterker samenhangt met de in de MEV gemaakte ramingsfout van de invoerprijsstijging voor het lopende jaar (Pm_{t-1}) dan met de in diezelfde MEV gemaakte ramingsfout voor het komende jaar (Pm_t). Toch is dit minder vreemd dan het op het eerste gezicht lijkt. Immers, op het moment dat de MEV wordt gemaakt zijn er (meer of minder voorlopige) realisatiecijfers voor de eerste twee kwartalen van het lopende jaar beschikbaar. De ramingsfout die in de MEV voor de invoerprijsstijging in het lopende jaar wordt gemaakt heeft dan ook grotendeels betrekking op onvoorziene ontwikkelingen in de tweede helft van het jaar, die in belangrijke mate 'overlopen' naar de raming voor het komende jaar. Daar komt nog bij dat internationale prijsontwikkelingen met vertraging worden doorberekend in afzetprijzen en lonen.

Voor het lopende jaar is er niet of nauwelijks samenhang gevonden tussen de CPB-voorspelfouten van de externe data en die van de kernvariabelen. Dat komt waarschijnlijk omdat de MEV-raming pas in augustus wordt afgesloten en er dan al een redelijk goed beeld is voor het lopende jaar. De invloed van de gedragsrelaties in het voor de korte-termijnramingen gehanteerde model SAFFIER is voor de ramingen voor het komende jaar aanzienlijk groter dan voor die voor het lopende jaar.

Een sterke correlatie tussen de internationale uitgangspunten en de raming voor de binnenlandse conjunctuur geldt niet alleen voor Nederland. Uit onderzoek van Keereman (2003) naar de trefzekerheid van ramingen van de Europese Commissie blijkt echter wel dat de voorspellingen voor de kleine en open Nederlandse economie veel gevoeliger zijn voor deze uitgangspunten dan in andere Europese landen het geval is. Van alle twaalf door Keereman onderzochte landen heeft Nederland verreweg de hoogste correlatie (0,80) tussen de voorspelfouten van de BBP-groei voor het komend jaar en die van de externe veronderstellingen. Omdat de internationale conjunctuur, zowel reëel als nominaal, heel grillig kan zijn door onverwachte schokken, bijvoorbeeld als gevolg van geopolitieke ontwikkelingen, is het relatief moeilijk om de ontwikkeling van de kerngrootheden van de Nederlandse economie goed te ramen.

²⁷ Zie paragraaf 5.2 van Kranendonk en Verbruggen (2006b).

5 Trefzekerheid tussen instellingen vergeleken

5.1 Methodiek

Niet alleen het CPB maar ook diverse andere nationale en internationale instellingen maken prognoses voor de Nederlandse economie. In deze paragraaf maken we een vergelijking van de trefzekerheid van een aantal instellingen die periodiek hun ramingen publiceren. Het onderling vergelijken van de trefzekerheid van de ramingen van meerdere instanties, waaronder die van het CPB zelf, ligt gevoelig. Daarom gaan we eerst uitvoerig in op de gemaakte keuzes en de gebruikte cijfers.²⁸ De keuzes die moeten worden gemaakt omvatten:

- Welke instanties worden in de analyse betrokken?
- Welke ramingen van deze instanties worden beschouwd?
- Naar de ramingen van welke kernvariabele(n), voor welke tijdsperiode, wordt gekeken?
- Welke indicatoren van trefzekerheid worden gebruikt?
- Waartegen worden de ramingen ‘afgerekend’?

Welke instanties worden in de analyse betrokken?

Ramingen voor de Nederlandse economie worden door vele instanties gemaakt. Er is gekozen voor instanties die dat al geruime tijd en op regelmatige basis doen, dat wil zeggen ieder jaar op ongeveer dezelfde momenten, en waarvan de ramingen bovendien openbaar zijn.

Gerennomeerde internationale instituten die ramingen voor de Nederlandse economie maken zijn de OESO, de Europese Commissie (EC) en het Internationaal Monetair Fonds (IMF).²⁹ Daarnaast zijn het vooral financiële instanties die ramingen afgeven. Rekening houdend met de zojuist genoemde keuzecriteria is gekozen voor De Nederlandsche Bank (DNB) en de private banken ABN-AMRO, ING, NIBC Bank en Rabobank, wetende dat daar in ieder geval voldoende capaciteit aanwezig is om uitgebreide ramingen voor de Nederlandse economie te maken. Er zijn echter nog veel meer financiële instellingen die ramingen maken. Om het aantal instanties dat wordt vergeleken enigszins te beperken, is ervoor gekozen ook de maandelijksse ramingen van Consensus Forecasts (CF), het gemiddelde van de ramingen van een tiental financiële instellingen, in de analyse te betrekken.³⁰ In totaal wordt dus de trefzekerheid van tien instanties geanalyseerd.

Welke ramingen van deze instanties worden beschouwd?

Omwille van de vergelijkbaarheid van de ramingen is het van belang dat de instanties op het moment van ramen over dezelfde informatie beschikken. In de praktijk is dit echter niet het

²⁸ De gebruikte data zijn beschikbaar via de CPB-website.

²⁹ De Europese Centrale Bank (ECB) publiceert geen ramingen voor de Nederlandse economie en is daarom niet in deze analyse betrokken.

³⁰ De ramingen van de private banken zijn overgenomen uit de Consensus Forecasts.

geval. Hierdoor wordt vooral de vergelijkbaarheid van de ramingen voor het lopende jaar beperkt. Realisatiecijfers over het komende jaar zijn per definitie voor geen enkel instituut beschikbaar, zodat de vergelijkbaarheid van die ramingen veel minder wordt gehinderd door verschillen in beschikbaarheid van informatie.³¹

Er zijn vier periodes van twee maanden waarin de ramingen verschijnen, te weten voorjaar (maart/april), zomer (mei/juni), najaar (augustus/september) en winter (november/december). Strikt genomen moet echter niet gekeken worden naar de publicatiedatum, maar naar de datum waarop de exogenen worden vastgesteld en de datum waarop de berekeningen voor de raming zijn afgesloten. De datum waarop de exogenen worden vastgesteld is bepalend voor de vraag welke recente informatie over de internationale economie en het overheidsbeleid nog is meegenomen. De afsluitingsdatum is bepalend voor de vraag welke recente conjunctuurinformatie men nog in de raming heeft kunnen verwerken. In een enkel geval worden deze data aangegeven, maar meestal is dat niet zo. Dit is een berucht lastig punt in het 'vergelijkend warenonderzoek'. Daarom wordt hier in paragraaf 5.2 uitgebreider op ingegaan.

Naar de ramingen van welke kernvariabele(n) wordt gekeken?

De analyse beperkt zich tot de macro-economische kernvariabele die verreweg het meest in de belangstelling staat, namelijk de volumegroei van het BBP, ofwel de economische groei. Het gaat daarbij om de raming voor het lopende jaar en die voor het komende jaar. Er is gekozen voor de langste periode waarover voor alle gekozen instanties ramingen beschikbaar zijn, te weten 1998-2007. Deze periode omvat zowel jaren met hoge economische groei als jaren met lage groei.

Welke trefzekerheidsindicatoren worden gebruikt?

Als indicator voor de omvang van de voorspelfout is gekozen is voor de gemiddelde absolute voorspelfout. De gemiddelde voorspelfout is berekend als maatstaf voor eventuele structurele onder- of overschatting. Daarnaast hebben we onderzocht hoe goed de verschillende instituten in staat zijn groeiversnellingen en -vertragingen te ramen. De ongelijkheidscoëfficiënt is niet in de analyse betrokken, omdat deze indicator voor de 'komend jaar'-ramingen minder makkelijk interpreteerbaar is bij de onderlinge vergelijking van de verschillende instituten. De ongelijkheidscoëfficiënt voor de 'komend jaar'-ramingen hangt immers ook af van de kwaliteit van de raming voor het lopende jaar, wat een verstoring element vormt als we de kwaliteit van de ramingen voor het komende jaar willen beoordelen.

³¹ In beperkte mate geldt het echter ook voor de ramingen voor het komende jaar, omdat via zogenoemde overloopeffecten de ontwikkeling in het lopende jaar mede van invloed zijn op de ontwikkeling in het komende jaar. Bovendien is op gebied van bijvoorbeeld de contractloonontwikkeling via reeds afgesloten cao's al het een en ander bekend over de loonontwikkeling in het komende jaar.

Wartegen worden de ramingen 'afgerekend'?

Net als bij de trefzekerheidsanalyse van de CPB-ramingen in de vorige paragraaf worden de ramingen in beginsel vergeleken met de *indertijd* gepubliceerde *definitieve* CBS-cijfers. Zoals aangegeven in paragraaf 2.2 kan deze werkwijze als 'streng' worden gekwalificeerd en zullen de voorspelfouten van alle instanties hierdoor opwaarts worden beïnvloed. Doordat dit echter voor alle instanties geldt, zullen de uitkomsten van de vergelijkende analyse hierdoor niet of nauwelijks worden beïnvloed.

5.2 Vergelijkbaarheid

Uit eerdere ervaringen is gebleken dat de uitkomsten van de analyse van de relatieve trefzekerheid van verschillende instanties gevoelig zijn voor de keuze van de ramingen.³² Voor een eerlijke vergelijking is het noodzakelijk dat de opstellers van de ramingen tijdens het ramingsproces over dezelfde informatie kunnen beschikken. In de praktijk is dit, zoals vermeld, echter niet het geval. Dit wordt vooral veroorzaakt doordat tussen de instanties grote verschillen bestaan ten aanzien van (de doorlooptijd van) het ramingsproces en het moment waarop de ramingen worden gemaakt. Dat beïnvloedt in belangrijke mate welke informatie over realisaties van het CBS en eventuele nieuwe beleidsvoornemens van het kabinet zijn verwerkt in de ramingen. Tabel 5.1 vat de belangrijkste verschillen de in ramingsprocessen van de onderzochte instituten samen.³³

Het lijkt waarschijnlijk dat de internationale instituten, DNB en CPB enigszins in het voordeel zijn bij het ramen van de economische groei in het komende jaar, omdat zij kennis hebben van het (voorgenomen) overheidsbeleid, terwijl ze ook over een uitgebreider instrumentarium beschikken. Omgekeerd lijken de banken wat in het voordeel te zijn bij het maken van de ramingen voor het lopende jaar, omdat ze een kortere doorlooptijd kennen en dus meer mogelijkheden hebben om tot het laatste moment nieuwe informatie te verwerken.

Overigens moet worden aangetekend dat het gebruikmaken van informatie over toekomstig beleid, zoals aangekondigd in de Miljoenennota, ook nadelen oplevert. Het is immers niet zelden het geval dat deze beleidsvoornemens door de Tweede Kamer worden gewijzigd dan wel om andere redenen niet tot uitvoering worden gebracht. Het is gebruik dat

³² Zo volgt uit de analyse van Blix et al. (2001) dat de trefzekerheid van de OESO-ramingen voor een aantal grote landen (Verenigde Staten, Japan, Frankrijk, Duitsland, Italië en Zweden) gemiddeld minder goed is dan de raming die in Consensus Forecasts van juni en december wordt gepubliceerd. Lenain (2002) toont aan dat die uitkomst sterk afhankelijk is van de keuze omtrent de publicatiemaand van de Consensus Forecasts waar de ramingen van andere instituten uit afkomstig zijn. Worden de OESO-ramingen, die weliswaar in juni en december in boekformaat worden gepubliceerd maar in mei en november al via een persbericht bekend worden gemaakt, vergeleken met de ramingen van andere instituten zoals gepubliceerd in de Consensus Forecasts van mei en november (in plaats van juni en december), dan scoort de OESO gemiddeld juist iets beter.

³³ Deze informatie is gebaseerd op informatie over de situatie in het najaar 2005. Een uitvoeriger toelichting op deze tabel staat in paragraaf 3.2 van Kranendonk en Verbruggen (2006a).

Tabel 5.1 Enkele eigenschappen van het ramingsproces per instantie^a

	CPB uitgebreide raming ^b	CPB beperkte raming ^c	OESO, EC, IMF	DNB	ABN- AMRO	Rabo- bank	ING	NIBC
Aantal reguliere ramingen per jaar	2	2	2	2	6	4	12	12
Overheidsbeleid gegeven	ja	ja	ja	ja	nee	nee	nee	nee
Gebruik van grootschalig econometrisch model	ja	ja	ja	ja	nee	nee	nee	nee
Doorlooptijd in weken								
Berekeningen	5 tot 9	2 tot 3	6 tot 9	7 tot 8	± 1	± 1	± 1	< 1
Tussen afsluitingsdatum en eerste publicatie	3 tot 4	± 2	1 tot 2	2 tot 3	± 1	± 2	± 2	< 1
Totale doorlooptijd	8 tot 12	4 tot 5	7 tot 11	9 tot 11	± 2	3 tot 4	± 3	< 1

^a Consensus Forecasts is het gemiddelde van een tiental financiële instellingen en daarom hier niet opgenomen.

^b CEP en MEV.

^c Juni-raming (Koninginne-MEV) en Decemberraming.

CPB, DNB en internationale instituten bij het maken van hun ramingen ervan uitgaan dat het aangekondigde overheidsbeleid ook daadwerkelijk wordt uitgevoerd. Private banken hoeven daar niet van uit te gaan en kunnen anticiperen op beleidswijzigingen. Met betrekking tot de OESO formuleert Lenain (2002) het als volgt: *“The mandate given to the OECD is to prepare projections based on announced economic policies, while private forecasts usually attempt to predict how policymakers will change their stance based on future developments.”*

Hoewel het duidelijk is dat het ideaal van identieke informatie op het moment van ramen niet van toepassing is, is ernaar gestreefd de ideale situatie zo goed mogelijk te benaderen. Bij de keuze van de ramingen die met elkaar worden vergeleken is in dit onderzoek specifiek gekeken of kort voor het afsluitingsmoment van de berekeningen het CBS een ‘flash-raming’ heeft gepubliceerd. Er verschijnt weliswaar dagelijks informatie over de (internationale) conjuncturele situatie, maar de ‘flash-raming’ van de BBP-groei door het CBS bevat relatief veel relevante informatie over de economische ontwikkeling in het lopende jaar. Uitgaande van het genoemde criterium en rekening houdend met het bijzondere karakter van de MEV zijn voor de ‘lopend jaar’-raming van de banken de ramingen gebruikt van respectievelijk april, juni, augustus en december.³⁴ Alleen voor de ABN-AMRO is daarbij het juli-cijfer genomen in plaats van het juni-cijfer, omdat deze bank heeft aangegeven in juni meestal geen raming te publiceren. De flash-raming voor de BBP-groei in het eerste kwartaal is voor deze bank dus pas verdisconteerd in de raming die in juli wordt afgegeven. Voor de ‘komend jaar’-ramingen zijn voor de banken de ramingen uit respectievelijk april, juni, september en december genomen.

³⁴ Hierbij is gebruik gemaakt van de ramingen van de banken zoals vermeld in de maandelijkse publicatie van Consensus Forecasts.

5.3 Resultaten

5.3.1 Trefzekerheid van groeiramingen lopend jaar

Allereerst is bepaald welke ramingen onderling goed vergelijkbaar zijn. Voor de voorjaars- en najaarsraming zijn dat er zes, voor de zomer- en winterraming acht. Bij deze indeling vormt het IMF een probleem, omdat dit instituut de publicatiedatum van zijn ramingen in de loop der tijd heeft vervroegd.³⁵ Om het IMF niet te benadelen (door de IMF-ramingen te beschouwen als zomer- en winterramingen) of te bevoordelen (door de ramingen in te delen bij de voorjaars- en najaarsramingen) is ervoor gekozen de IMF-ramingen een 'status aparte' te geven: de gemiddelde (absolute) voorspelfouten van de betreffende ramingen worden wel in de tabellen vermeld, maar ze tellen niet mee voor het gemiddelde van alle instanties.

Vervolgens zijn per instituut en per ramingsronde de gemiddelde absolute voorspelfouten van de ramingen voor de BBP-groei in het lopende jaar berekend. Deze cijfers, die de basis vormen van het vergelijkende warenonderzoek, zijn vermeld in tabel 5.2. Bij het bepalen van de voorspelfouten zijn dezelfde conventies ten aanzien van de realisatiecijfers gehanteerd als in de vorige paragrafen.

De tabel laat zien dat de trefzekerheid in de loop van het jaar gestaag toeneemt. Dat is niet vreemd, aangezien er steeds meer informatie over het voorspeljaar beschikbaar komt. Toch bedraagt de gemiddelde absolute voorspelfout in de ramingen die in de twee laatste maanden van het jaar worden gepubliceerd nog altijd circa 0,6%-punt. Dit wordt vooral veroorzaakt doordat op dat moment de meest recente informatie betrekking heeft op het derde kwartaal van het voorspeljaar en de cijfers over de eerste drie kwartalen bovendien nog een voorlopig karakter hebben. Deze cijfers worden weliswaar 'voor waar' aangenomen, maar zoals in paragraaf 2.2 werd aangegeven bedraagt de bijstelling van cijfers tussen 'flash' en 'definitief' circa 0,45%-punt.

De instanties ontlopen elkaar niet veel. Uit de kolom met genormeerde gemiddelde absolute voorspelfouten volgt dat de trefzekerheid van de BBP-groeiraming voor het lopende jaar in de periode 1998-2007 het grootst was bij de Rabobank, gevolgd door NIBC, ABN-AMRO en

³⁵ Zo publiceerde het IMF in de periode 1997-2000 zijn World Economic Outlook in de eerste helft van mei, gedurende 2001 en 2002 in de tweede helft van april en de afgelopen twee jaren in de eerste helft van april. Ook het CPB heeft de publicatiedatum van het CEP in de loop der tijd wat vervroegd. Gemiddeld genomen verschijnt het CEP ongeveer drie weken vòòr de World Economic Outlook van het IMF.

Tabel 5.2 Gemiddelde absolute voorspelfout per instituut van de BBP-groeiramingen voor lopend jaar, 1998-2007

	Maart-april (voorjaar)	Mei-juni (zomer)	Aug.-sept. (najaar)	Nov.-dec. (winter)	Gemiddelde (genormeerd)	Rang
CPB	1,03	0,92	0,76	0,56	0,98	4
OESO		1,16		0,66	1,16	9
EC		1,09		0,63	1,10	8
IMF ^a		0,97		0,64		
DNB		0,91		0,61	0,99	5
Consensus	1,03	0,92	0,76	0,61	1,00	6
Forecasts						
ABN-AMRO	0,90	0,87	0,79	0,61	0,96	2
Rabobank	1,01	0,77	0,69	0,55	0,91	1
ING	1,10	0,94	0,89	0,60	1,06	7
NIBC	0,88	0,90	0,78	0,59	0,96	2
Gemiddeld (excl. IMF)	0,99	0,94	0,78	0,60	1	

^a Vanwege een verschuiving in de tijd van het ramingsproces bij het IMF vallen de IMF-ramingen tussen de voorjaars- en zomerraming respectievelijk de najaars- en winterraming in.

CPB.³⁶ De oorzaak van de relatief slechte scores van de OESO en de EC ligt waarschijnlijk mede besloten in het feit dat deze internationale instituten het nadeel hebben van een langdurig ramingsproces, bestaande uit diverse iteratierondes, en niet het voordeel hebben van specifieke op de Nederlandse economie toegesneden, actuele en unieke informatie.

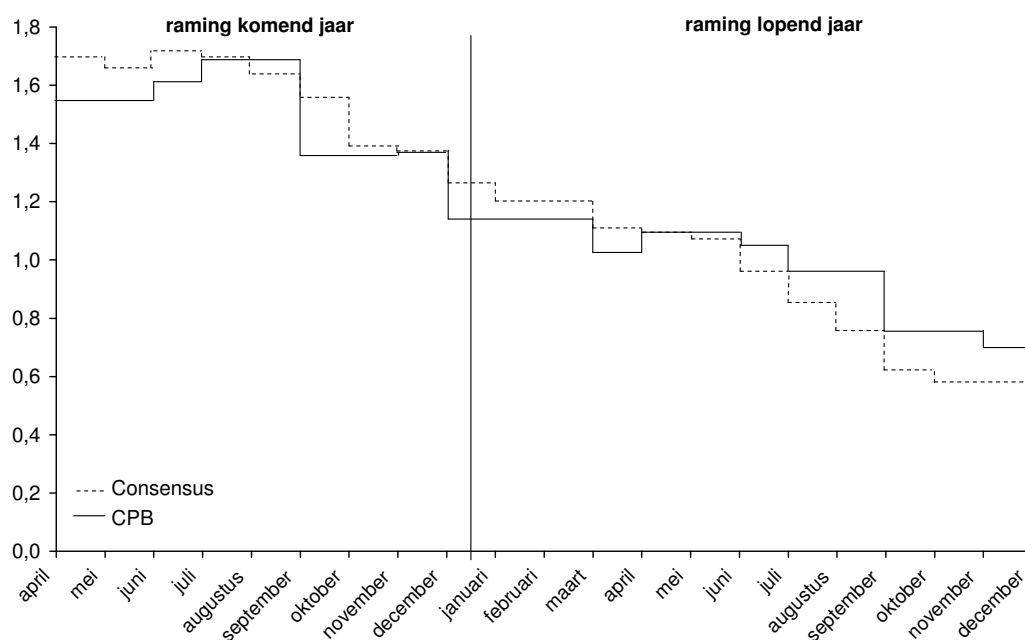
In figuur 5.1 is voor de Consensus Forecasts weergegeven hoe de trefzekerheid, afgemeten aan de gemiddelde absolute voorspelfout, van maand op maand verbetert met het beschikbaar komen van meer informatie. Ter vergelijking is ook de trefzekerheid van de CPB-ramingen opgenomen. Daarbij zijn de CPB-ramingen gedateerd op de maanden dat het CPB een nieuwe raming publiceert, april, juni, september en december.³⁷

De figuur illustreert dat de trefzekerheid van beide ramingen van de BBP-groei in het komende jaar in grote mate overeenkomen. In de maanden dat het CPB een nieuwe raming publiceert is de gemiddelde absolute voorspelfout in de meeste gevallen wat geringer dan die

³⁶ Om een eerlijke rangordering te verkrijgen is de gemiddelde voorspelfout van elk instituut per ramingsmoment geschaald op het gemiddelde van dat ramingsmoment. Vervolgens is op basis van deze 'genormeerde' voorspelfouten een rangorde bepaald. Zou dit niet gebeuren, dan zouden OESO, EC en DNB bevoordeeld worden, omdat zij slechts twee ramingen per jaar uitbrengen en dat relatief laat in het jaar doen. Dat levert een informatievoordeel op. Ter illustratie van het normeren van de gemiddelde absolute voorspelfout: het CPB komt nu uit op een genormeerde gemiddelde absolute voorspelfout van 0,98. Dit cijfer is als volgt berekend. De gemiddelde absolute voorspelfout in het voorjaar, 1,03, is geschaald op de gemiddelde fout van alle instellingen, in het voorjaar, 0,99. Dit levert een genormeerde voorspelfout op van 1,04. Op vergelijkbare wijze leveren de overige drie ramingen genormeerde voorspelfouten op van respectievelijk 0,98, 0,97 en 0,93. Het gemiddelde van deze vier genormeerde voorspelfouten is 0,98.

³⁷ Vanwege het afsluiten van de berekeningen is alleen de MEV-raming voor het lopende jaar gedateerd in augustus. De extra 'knik' in november hangt samen met de tussentijdse ramingen die het CPB in het najaar van 1998 en 2001 heeft uitgebracht in verband met het uitbreken van de Azië-crisis respectievelijk de terroristische aanslagen in de Verenigde Staten op 11 september 2001.

Figuur 5.1 Gemiddelde absolute fout van Consensus Forecasts en CPB



van de Consensus Forecasts. Voor het lopende jaar lijkt de verbetering van de trefzekerheid van het CPB wat na te ijlen op de maandelijkse ramingen van de Consensus Forecasts. Dat geldt met name ten tijde van de MEV-raming; om het begrotingsproces niet onnodig te verstoren, wijzigt het CPB tussen de concept-MEV van begin augustus en de echte MEV van de derde dinsdag in september haar uitgangspunten alleen nog in geval van majeure nieuwe informatie.

De gemiddelde voorspelfouten geven aan of sprake is van een structurele over- of onderschatting. Aan het begin van het jaar (maart-april) ligt de gemiddelde voorspelfout voor de

Tabel 5.3 Gemiddelde voorspelfout per instituut van de BBP-groeiramingen voor lopend jaar, 1998-2007^a

	Maart-april (voorjaar)	Mei-juni (zomer)	Augustus-september (najaar)	November-december (winter)
CPB	- 0,05	- 0,18	- 0,26	- 0,38
OESO		- 0,20		- 0,38
EC		- 0,11		- 0,45
IMF ^b		- 0,11		- 0,48
DNB		- 0,53		- 0,49
Consensus Forecasts	- 0,19	- 0,34	- 0,36	- 0,47
ABN-AMRO	- 0,08	- 0,27	- 0,37	- 0,47
Rabobank	0,01	- 0,27	- 0,27	- 0,41
ING	- 0,36	- 0,46	- 0,43	- 0,50
NIBC	0,00	- 0,22	- 0,36	- 0,39
Gemiddeld (excl. IMF)	- 0,11	- 0,29	- 0,34	- 0,44

^a Minteken duidt op onderschatting.

^b Vanwege een verschuiving in de tijd van het ramingsproces bij het IMF vallen de IMF-ramingen tussen de voorjaars- en zomerraming respectievelijk de najaars- en winterraming in.

meeste instituten redelijk dicht in de buurt van de nul. Naarmate het jaar vordert en er steeds meer informatie beschikbaar zou je verwachten dat de gemiddelde voorspelfout steeds dichterbij nul zou komen te liggen. Het tegendeel is echter het geval. Dit hangt waarschijnlijk voor een groot deel samen met de opwaartse bijstelling van de realisatiecijfers die moeilijk is te voorspellen door de instituten.

5.3.2 Trefzekerheid van groeiramingen komend jaar

Deze paragraaf gaat in op de trefzekerheid van de BBP-groeiramingen voor het komende jaar. Ook nu zijn eerst de gemiddelde absolute voorspelfouten berekend. Ten behoeve van een zo zuiver mogelijke vergelijking, wordt bij de 'komend jaar'-raming de MEV vergeleken met de septemberramingen van de private banken en Consensus Forecasts.³⁸ Uit tabel 5.4 komt naar voren dat de trefzekerheid van de ramingen van de BBP-groei bij alle instanties in de loop van de tijd toeneemt, met uitzondering van de mei-juni raming.

Tabel 5.4 Gemiddelde absolute voorspelfout per instituut van de BBP-groeiramingen voor komend jaar, 1998-2007

	Maart-april (voorjaar)	Mei-juni (zomer)	Aug.-sept. (najaar)	Nov.-dec. (winter)	Gemiddelde (genormeerd)	Rang
CPB	1,38	1,51	1,23	1,07	0,97	3
OESO		1,41		1,23	1,03	8
EC		1,43		1,15	1,00	5
IMF ^a		1,34		1,29		
DNB		1,43		1,12	0,99	4
Consensus Forecasts	1,46	1,47	1,37	1,15	1,02	7
ABN-AMRO	1,38	1,32	1,26	1,08	0,95	2
Rabobank	1,29	1,29	1,28	1,08	0,93	1
ING	1,62	1,65	1,45	1,24	1,12	9
NIBC	1,46	1,43	1,35	1,08	1,00	5
Gemiddeld (excl. IMF)	1,43	1,44	1,32	1,13	1	

^a Vanwege een verschuiving in de tijd van het ramingsproces bij het IMF vallen de IMF-ramingen tussen de voorjaars- en zomerraming respectievelijk de najaars- en winterraming in.

De onderlinge verschillen in trefzekerheid zijn zeer gering. De beste score wordt ook voor deze raming gehaald door de Rabobank, gevolgd door de ABN-AMRO en CPB. Vergelijken we de uitkomsten in tabel 5.4 met die in tabel 5.2 dan blijkt dat de relatieve verschillen in trefzekerheid van de ramingen voor het komende jaar nog iets kleiner zijn dan die van de BBP-groeiramingen in het lopende jaar. Dit is opmerkelijk, aangezien de onzekerheden omtrent de

³⁸ Via 'uitgelekte' informatie over de Miljoenennota zijn in augustus en september niet alleen het CPB maar ook andere instanties op de hoogte van de hoofdlijnen van het begrotingsbeleid in het komende jaar. Het CPB beschikt weliswaar over meer gedetailleerde informatie omtrent de beleidsvoornemens, maar daar staat tegenover dat vanwege de bijzondere rol van de MEV in het begrotingsproces het internationale beeld al in een zeer vroegtijdig stadium moet worden vastgesteld. Zie kader in paragraaf 7.2 van Kranendonk en Verbruggen (2006c).

economische ontwikkeling in het komende jaar relatief groot zijn. Wellicht dat men juist om die reden bij het maken van de raming voor het komende jaar minder vaart op eigen waarnemingen en informatie en meer naar elkaar kijkt. Daarnaast is het ook denkbaar dat bij het maken van ‘komend jaar’-ramingen sterker op modeluitkomsten wordt geleund. Wanneer deze modellen de belangrijkste economische relaties binnen de Nederlandse economie op vergelijkbare wijze weergeven, is het niet zo verwonderlijk dat ook de ramingen op elkaar lijken, tenzij de ramingen voor de zogenoemde ‘buitenlandexogenen’ sterk van elkaar verschillen.

Voor de ramingen voor het komende jaar zijn ook de gemiddelde voorspelfouten uitgerekend (zie tabel 5.5). De meeste instanties hebben in de periode 1998-2007 de BBP-groei in het komende jaar gemiddeld iets overschat. Dit is het saldo van een gemiddeld genomen te optimistische raming voor de neergaande conjunctuur in de periode 2001 t/m 2004 met vele opeenvolgende schokken (terroristische aanslagen, geopolitieke spanningen in het Midden Oosten, epidemieën en boekhoudschandalen) en het onderschatten van het herstel in de latere jaren.

In aanvulling op de kwantitatieve vergelijking van voorspelfouten, is ook een kwalitatieve analyse uitgevoerd van de mate waarin instituten verschillen wanneer het gaat om het voorspellen van groeiversnellingen en -vertragingen. Hierbij bekijken we aan de hand van de bijbehorende maatstaf, die in paragraaf 2.3.4 is toegelicht, hoe goed de instellingen scoren met

Tabel 5.5 Gemiddelde voorspelfout per instituut van BBP-groeiramingen voor komend jaar, 1998-2007^a

	Maart-april (voorjaar)	Mei-juni (zomer)	Augustus-september (najaar)	November-december (winter)
CPB	0,14	0,13	0,07	- 0,19
OESO		0,27		- 0,03
EC		0,25		- 0,09
IMF ^b		0,20		0,03
DNB		0,01		- 0,22
Consensus Forecasts	0,22	0,15	0,11	- 0,17
ABN-AMRO	0,10	0,18	0,24	- 0,10
Rabobank	0,31	0,21	0,16	0,00
ING	0,00	- 0,05	- 0,07	- 0,40
NIBC	0,18	0,13	0,17	0,06
Gemiddeld (excl. IMF)	0,16	0,14	0,11	- 0,13

^a Minteken duidt op onderschatting.

^b Vanwege een verschuiving in de tijd van het ramingsproces bij het IMF vallen de IMF-ramingen tussen de voorjaars- en zomerraming respectievelijk de najaars- en winterraming in.

betrekking tot criterium f. In tabel 5.6 zijn de relevante uitkomsten vermeld voor de ramingen voor het komend jaar voor de periode 1999-2007.³⁹

Een vertraging of versnelling van de groei werd gemiddeld in 41% van de ramingen juist geraamd, terwijl in 14% het tegenovergestelde is gebeurd van wat werd verwacht. In iets minder dan de helft van de situaties was de realisatie net anders dan voorzien. Het CPB scoort als enige instituut meer dan de helft 'juist'. Deze uitkomst voor de periode 1999-2007 is voor het CPB iets beter dan de uitkomst voor de langere periode 1971-2007 (zie paragraaf 3.1). Wanneer we inzoomen op de achterliggende ramingen, blijkt dat met name voor de jaren 1999-2003 de ontwikkeling van de groei vaak onjuist werd voorspeld. Uitschieter is daarbij 2003: alle instituten voorspelden in 2002, in al hun ramingen in dat jaar, dat de economische groei in 2003 0,5%-punt of meer boven de groei in 2002 zou liggen. Het tegendeel kwam echter uit: de groei vertraagde.⁴⁰

Tabel 5.6 Voorspelkwaliteit versnelling/vertraging van de groei voor komend jaar, 1999-2007

	Versnelling/vertraging		
	% juist	% enigszins onjuist	% geheel onjuist
CPB	56	33	11
OESO	44	44	11
EC	39	50	11
IMF	33	50	17
DNB	33	56	11
Consensus Forecasts	44	42	14
ABN-AMRO	42	44	14
Rabobank	42	39	19
ING	47	42	11
NIBC	25	58	17
Gemiddeld	41	46	14

³⁹ Omdat niet bekend is van welke groeicijfers van het afgelopen jaar is uitgegaan, is er geen vergelijkbare analyse gemaakt voor het raming van het lopende jaar.

⁴⁰ Het CBS heeft in 2005 een revisie van de Nationale rekeningen doorgevoerd. Na deze revisie bleek de economische groei in 2003 0,2%-punt hoger te zijn dan in 2002. Zoals in paragraaf 2.2 is uitgelegd, gaan we in deze trefzekerheidsstudie echter uit van de realisatiecijfers, zoals die destijds golden. Vóór de revisie van 2005 was er nog sprake van een forse afname van de economische groei in 2003. Het kader 'Economische ontwikkeling Nederland in afgelopen tien jaar gunstiger na herziening CBS-cijfers' uit de MEV 2007 gaat hier op in (CPB, 2006).

6 Trefzekerheid in internationaal perspectief

In dit hoofdstuk wordt de trefzekerheid van de groeiramingen van het CPB voor de Nederlandse economie in international perspectief geplaatst. Hoe scoort het CPB als we de hiervoor beschreven resultaten vergelijken met de trefzekerheid van grote internationale instellingen als de OESO, het IMF en de EC? Voor de analyse maken we gebruik van rapporten van deze instellingen over hun eigen trefzekerheid:

1. De OESO heeft in 2007 een studie gepubliceerd waarin de ramingen uit de Economic Outlook voor de G7 worden geanalyseerd voor de jaren 1991-2006.⁴¹ Bijzonder element in deze studie is dat ook uitvoerig wordt geanalyseerd hoe trefzekerheid van de OESO-ramingen zich verhoudt tot die van de ramingen die Consensus Forecasts maandelijks rapporteert.
2. Het IMF analyseert ook met enige regelmaat de kwaliteit van zijn eigen ramingen. De meest recente studie stamt uit 2007 en doet verslag van de trefzekerheid van de ramingen voor de periode 1990-2003 in de World Economic Outlook voor 178 landen, waar eerder alleen ramingen voor de G7 waren onderzocht.⁴² Van de zeven regio's die het IMF onderscheidt, worden in deze studie alleen de resultaten voor de (29) ontwikkelde landen ('advanced economies') bekeken.
3. Ten slotte gebruiken we een recente studie van de EC, waarin de trefzekerheid van de ramingen van de Commissie voor alle aangesloten landen wordt besproken op basis van enkele kengetallen en toetsen.⁴³

Evenals bij de vergelijking van de trefzekerheid van voorspellingen voor de Nederlandse economie in het vorige hoofdstuk zijn er factoren die de vergelijking bemoeilijken en tot voorzichtigheid manen bij het trekken van conclusies. Het eerste probleem is dat het moment van ramen, en daarmee de dan beschikbare informatie over het verloop van de conjunctuur en het overheidsbeleid, verschilt tussen de instituten. Daarvoor verwijzen we naar paragraaf 5.2, waar ook de relevante informatie over de ramingen van de OESO, het IMF en de EC is weergegeven. Een ander probleem betreft de gebruikte realisatiecijfers. Waar we in de vorige paragraaf voor alle instituten konden rekenen met dezelfde realisaties, is dat hier niet mogelijk. De OESO, het IMF en de EC gebruiken bij hun onderzoek als realisatiecijfer het cijfer dat ze destijds hebben gepubliceerd in het voorjaar direct na het verslagjaar. Dat betreft eigen cijfers, die niet noodzakelijk hetzelfde zijn bij de instituten, maar vermoedelijk dicht in de buurt komen van wat de statistische bureaus hebben gerapporteerd op basis van de vier kwartaalcijfers. De instituten laten de gebruikelijke bijstellingen in latere jaren buiten beschouwing, al geven ze wel expliciet aan dat de uitkomsten gevoelig zijn voor deze keuze. Dat raakt vooral de eventuele

⁴¹ Zie Vogel (2007). De G7-landen zijn de Verenigde Staten, Canada, Groot-Brittannië, Duitsland, Frankrijk, Italië en Japan.

⁴² Zie Timmerman (2007).

⁴³ Zie Melander, Sismanidis en Grenouilleau (2007).

bias (gemiddelde voorspelfout), maar kan ook invloed hebben op de richting waarin de groei zich ontwikkelt. Een derde obstakel voor een goede vergelijking wordt gevormd door het feit dat de perioden waarop de ramingen betrekking hebben verschillen. Onderstaande tabel is, vanwege de genoemde problemen met de vergelijkbaarheid van cijfers, dan ook vooral bedoeld als indicatie van de omvang van de voorspelfouten.

Tabel 6.1 geeft een vergelijking van de statistische kengetallen van de prognoses van de instituten zelf. Omwille van een zo goed mogelijke vergelijkbaarheid hebben we ook voor de CEP- en MEV-ramingen van het CPB de kengetallen berekend voor de periode 1990-2007.

Tabel 6.1 Internationale vergelijking van de trefzekerheid van BBP-groeiramingen voor Nederland en diverse landengroepen

	OESO G7	IMF Ontwikkelde landen	EC Lidstaten	EC Nederland	CPB Nederland
	1991-2006	1990-2003	1969-2005	1969-2005	1990-2007
Gemiddelde voorspelfout					
Komend jaar, raming in januari-juni T-1	0,7	0,6			0,0
Komend jaar, raming in juli-december T-1	0,3	0,4	0,3	0,0	-0,1
Lopend jaar, raming in januari-juni T	0,1	0,0	0,1	0,0	-0,3
Absolute voorspelfout					
Komend jaar, raming in januari-juni T-1					1,1
Komend jaar, raming in juli-december T-1			0,9	1,1	1,1
Lopend jaar, raming in januari-juni T			0,5	0,7	0,9
RMSE					
Komend jaar, raming in januari-juni T-1	1,6	2,1			1,3
Komend jaar, raming in juli-december T-1	1,3	2,0	1,2	1,4	1,3
Lopend jaar, raming in januari-juni T	0,7	1,4	0,7	0,9	1,1
Versnelling/vertraging juist geraamd (%)					
Komend jaar, raming in januari-juni T-1	63				67
Komend jaar, raming in juli-december T-1	74		77	57	61
Lopend jaar, raming in januari-juni T	86		83	75	67

Bij de ramingen voor het komende jaar is het beeld gevarieerd. Terwijl de gemiddelde voorspelfout van prognoses van het CPB dicht bij nul ligt, overschatten met name de OESO en het IMF de economische ontwikkeling. Door beide instituten wordt de groei voor volgend jaar gemiddeld meer ca. 0,5%-punt te hoog ingeschat. Deze overschatting neemt wel af naarmate meer informatie beschikbaar komt. Voor het lopende jaar hebben alle instituten een gemiddelde voorspelfout van rond de nul en is er dus geen of nauwelijks sprake over- of onderschatting van de groei. De EC laat bij de raming voor het komende jaar overigens gemengde resultaten zien: op het niveau van de lidstaten lijkt er sprake te zijn van een systematische overschatting van 0,3%-punt, maar voor Nederland afzonderlijk zijn de voorspelfouten gemiddeld nul. Als er rekening mee gehouden wordt dat de EC een voorlopig realisatiecijfer gebruikt, is er geen

Bias en onafhankelijkheid

In een afzonderlijke studie van de EC is onderzoek gedaan naar mogelijke systematische onder- dan wel overschatting van de economische groei.^a Daaruit komt naar voren dat de prognoses van de meeste westerse landen 'last' hebben van een bias. Dat wordt mede toegeschreven aan het feit dat de landen soms belang kunnen hebben bij een optimistische of wat voorzichtiger prognose. De economische groei is een belangrijke determinant voor de raming van de overheidsfinanciën. Een te optimistische raming kan daarom leiden tot een rooskleurig beeld van de overheidsfinanciën en het onvoldoende (h)erkennen van het treffen van (aanvullend) beleid om binnen de grenzen te blijven van de financiële afspraken die binnen Europa gemaakt zijn. In de studie wordt ook gemeld dat in de drie Europese landen waar geen systematische bias voorkomt bij de prognoses die door de landen zelf als basis wordt gebruikt voor de overheidsfinanciën - België, Nederland en Oostenrijk - de prognoses door onafhankelijke instituten worden opgesteld en niet door de regering zelf. Hoewel instituten als OESO en IMF zelf verantwoordelijk zijn voor hun ramingen, is niet uit te sluiten dat ze toch ook soms beïnvloed worden door kennis te nemen van de prognoses van de aangesloten landen en daarmee ook zelf last hebben van een bias. De OESO noemt als verklaring voor een bias in de ramingen dat met name in perioden van neergaande conjunctuur, achteraf bezien, de ramingen te laat neerwaarts zijn aangepast, omdat men te optimistisch bleek te zijn over de trendmatige groei.^b Ook het IMF zoekt de verklaring in die richting, wijzend op de veel gebruikte veronderstelling dat 'the output gap will be eliminated over a reasonably short span of time'.^c Vooral bij neergaande conjunctuur was men dan te optimistisch over het herstel van de economische groei in met name West-Europa.

^a Jonung and Larch (2004).

^b Vogel (2007, pagina 10).

^c Timmermann (2007, pagina 33).

verschil met de raming van het CPB.⁴⁴ Daarbij moet bedacht worden dat de EC-raming enkele maanden later wordt afgerond dan de CPB-raming (juni vs. maart en november vs. augustus). De EC kan dus kennis nemen van de inzichten die verwerkt zijn in de CPB-raming en bovendien recentere CBS-informatie verwerken.

Bij de analyse van de trefzekerheid kan niet volstaan worden met de gemiddelde voorspelfout, maar zijn ook de gemiddelde absolute voorspelfout en de RMSE relevant. Die geven inzicht in de grootte van afzonderlijke voorspelfouten. De RMSE van de IMF-ramingen ligt wat hoger dan die van de OESO en de EC. Daarnaast blijkt dat de EC bij de ramingen voor Nederland gemiddeld een iets grotere absolute voorspelfout maakt dan bij de ramingen voor de lidstaten, terwijl dat voor de gemiddelde fout andersom is. Vermoedelijk komt dit doordat de prognose voor de Nederlandse economie gevoeliger is voor de uitgangspunten voor de internationale conjunctuur en voor een minder groot deel gebaseerd kan worden op bekende binnenlandse factoren.

Voor alle hier vergeleken instellingen geldt dat de voorspelfout kleiner wordt naarmate er in de loop van de tijd meer informatie beschikbaar komt. Dit volgt zowel uit de ontwikkeling van de gemiddelde absolute voorspelfout als uit de ontwikkeling van de RMSE. De mate waarin een voorspelling verbetert bij het voortschrijden van de tijd, verschilt echter wel per instelling.

⁴⁴ Het effect van een ander realisatiecijfer voor de gemiddelde voorspelfout, zeker voor het lopend jaar, is toegelicht in paragraaf 3.1.

De vergelijking van de trefzekerheid van de prognoses ten aanzien van versnellingen en vertragingen van de economische groei laat zien dat de internationale instituten in meer dan 50% van de prognoses de richting waarin de groei zich beweegt juist ramen en dat dit percentage in het algemeen toeneemt met het beschikbaar komen van meer informatie. Omwille van een goede vergelijkbaarheid is hier voor wat betreft de CPB-ramingen voor de Nederlandse economie de ondergrens van 0,5%-punt om een groeimutatie als versnelling of vertraging te kenschetsen verlaagd naar 0%-punt.⁴⁵

Tot slot de efficiëntie van de prognoses. Uit de analyse in paragraaf 3.3 blijkt dat van de daar onderzochte variabelen alleen voor de particuliere consumptie niet gezegd kan worden dat de prognoses van het CPB efficiënt waren. Hoe scoren de internationale instituten op dit criterium? De OESO rapporteert dat de prognoses van zowel het voorjaar als het najaar de efficiëntietoets niet kunnen doorstaan. Voor het lopende jaar zijn de prognoses wel efficiënt. De EC heeft vergelijkbare toetsen uitgevoerd en deze tonen aan dat de EC-prognoses voor de meeste landen efficiënt zijn. Er zijn geen resultaten voor het IMF beschikbaar.

De belangrijkste les die uit bovenstaande analyse getrokken kan worden, is dat de verschillen tussen de internationale ramingsinstituten wat betreft trefzekerheid betrekkelijk gering zijn. De gemiddelde absolute voorspelfout en de RMSE van de diverse ramingen ontlopen elkaar niet veel, evenals de score bij het voorspellen van groeiversnellingen of -vertragingen. Het meest opmerkelijk is dat het CPB vanaf het eerste ramingsmoment een gemiddelde voorspelfout heeft die dichtbij nul ligt, terwijl deze gemiddelde fout bij de OESO en het IMF pas in de buurt van nul komt wanneer het ramingsjaar nadert. Dat een kleine en open economie als Nederland in relatief grote mate afhankelijk is van de moeilijk te ramen internationale conjunctuur, is een mogelijke verklaring voor het feit dat de absolute voorspelfouten van de verschillende instellingen vergelijkbaar zijn, ondanks het feit dat de gemiddelde voorspelfout van het CPB dicht bij nul ligt.

⁴⁵ Ook een toename van de economische groei met 0,1%-punt wordt hier dus beschouwd als een versnelling van de groei, dit in tegenstelling tot de analyse in paragraaf 3.1.

7 Evaluatie

Deze studie laat zien dat de uitkomsten van economische voorspellingen met de nodige onzekerheid omgeven zijn. Dat zal weinigen verbazen. In de periode 1971-2007 heeft het CPB de BBP-groei voor het komende jaar gemiddeld weliswaar juist geschat, maar de gemiddelde absolute voorspelfout bedraagt zelfs in de MEV voor het lopende jaar nog altijd 0,7%-punt. Dat lijkt vrij hoog, maar is desalniettemin niet slecht vergeleken met andere, zowel nationale als internationale, instituten. Behalve bij de bedrijfsinvesteringen, die notoir lastig te ramen zijn, ligt de absolute voorspelfout voor andere binnenlandse economische variabelen in dezelfde orde van grootte.

Uit de analyse komt naar voren dat een belangrijk deel van de voorspelfout voortkomt uit aannames omtrent internationale ontwikkelingen die niet uit blijken te komen. Zo komt ongeveer driekwart van de voorspelfout van de BBP-groei voor rekening van deze zogenoemde buitenlandexogenen. Omdat de Nederlandse economie klein en open is, wordt de kwaliteit van de ramingen relatief sterk beïnvloed door de situatie in het buitenland.

De tweede belangrijke bron van onzekerheid is het macro-econometrische model SAFFIER, dat wordt gebruikt bij het opstellen van ramingen. Dit model is een onmisbaar hulpmiddel bij het opstellen van ramingen. Enerzijds zorgt het model ervoor dat altijd aan de noodzakelijke boekhoudkundige identiteiten wordt voldaan. Anderzijds biedt het de mogelijkheid om consistent over de macro-economie, waar ‘alles met alles samenhangt’, na te denken en economische voorspellingen te maken. Een model is echter noodzakelijkerwijs een gestileerde weergave van de werkelijkheid. Bovendien zijn parameters van gedragsvergelijkingen geschat aan de hand van data over het verleden. Niet alleen is er onzekerheid omtrent de juiste waarden van de geschatte parameters voor het verleden, het is ook nog eens mogelijk dat deze waarden naar de toekomst toe veranderen.

Twee vragen komen op: 1) hoe kan de onzekerheid rondom ramingsuitkomsten worden verkleind en 2) wat is de meest geschikte manier om de onzekerheid weer te geven? Het antwoord op de eerste vraag is eenvoudig te geven, maar daarom nog niet eenvoudig uitvoerbaar. Een verbetering van de prognoses van externe factoren, zoals de groei van de relevante wereldhandel, de olieprijs of de wisselkoersen, zal leiden tot betere voorspellingen van de Nederlandse economische ontwikkeling. Deze grootheden zijn echter notoir lastig goed te voorspellen. Bovendien zullen zich vaak gebeurtenissen voordoen, die simpelweg niet te voorzien zijn, zoals terroristische aanslagen, boekhoudschandalen of de opbouw van geopolitieke spanningen. Daarnaast zal een beter model betere ramingen afleveren. Het is derhalve zaak het gebruikte macro-econometrische model SAFFIER zo up-to-date mogelijk te houden. In de praktijk gebeurt dit door regelmatig nieuwe theoretische inzichten te verwerken in het model en belangrijke gedragsvergelijkingen te herschatten met nieuwere data. Veranderingen in de economische structuur komen zo tot uiting in veranderende modelparameters. Ten slotte is het belangrijk altijd het gezonde verstand te gebruiken. Dit

klinkt vanzelfsprekend, maar een potentieel gevaar bij het gebruik van een model is dat de door het model gegenereerde uitkomsten 'blind' worden overgenomen. De wetenschap dat het model niet de volledige werkelijkheid beschrijft en dat het geschat is over het verleden, dwingt modelgebruikers de uitkomsten kritisch tegen het licht te houden en eventueel bij te sturen waar nodig. Het CPB heeft op dit gebied de afgelopen 60 jaar de nodige deskundigheid en ervaring opgebouwd.⁴⁶ Om te bepalen waar en wanneer ingegrepen moet worden, wordt er - naast inzichten van experts - ook gebruik gemaakt van hulpmiddelen als 'leading indicators' en VAR-modellen.

Maar al zouden de externe ontwikkelingen beter worden geraamd, al zou SAFFIER de juiste parameters bevatten en al zouden modeluitkomsten gegeven de inzichten optimaal worden aangestuurd, dan nog zullen onvoorziene gebeurtenissen zich voordoen en blijven de ramingen met onzekerheid omgeven. De vraag blijft dan hoe deze onzekerheid het best in beeld gebracht kan worden. Twee veelgebruikte mogelijkheden zijn het geven van onzekerheidsintervallen voor belangrijke economische variabelen en het geven van uitkomsten van varianten. In het eerste geval wordt er een bandbreedte, gebaseerd op historische data, gegeven waarbinnen de realisatie van een variabele met een bepaalde waarschijnlijkheid zal uitkomen. In het tweede geval worden puntschattingen van de uitkomsten gepresenteerd, die gebaseerd zijn op andere aannames omtrent exogene variabelen dan in de centrale raming.

Het CPB kiest om meerdere redenen voor het presenteren van varianten. Ten eerste hebben afnemers van ramingen van het CPB, zoals diverse ministeries, behoefte aan een puntschatting. Voor het maken van een begroting heeft bijvoorbeeld het ministerie van Financiën weinig aan onzekerheidsintervallen. Ten tweede, het geven van onzekerheidsintervallen verbetert de voorspellingen niet. Een substantiële internationale schok kan de realisaties gemakkelijk buiten het interval doen belanden. Ten derde, het is niet eenvoudig om onzekerheidsmarges te berekenen en de uitkomsten zijn mede afhankelijk van de gemaakte veronderstellingen. Ten slotte is het de vraag of marges voor het verleden wel representatief zijn voor de toekomst. Varianten bieden een betere mogelijkheid om op de specifieke risico's in gaan en de gevolgen te laten zien indien de vooraf ingeschatte risico's bewaarheid worden.⁴⁷

⁴⁶ Don en Verbruggen (2006).

⁴⁷ Zie ook het kader 'Puntschattingen of onzekerheidsmarges?' in de MEV 2006 (CPB, 2005).

Bijlage 1 Specifieke onderwerpen in vorige trefzekerheidspublicaties

In iedere trefzekerheidspublicatie wordt in aanvulling op de standaardtabellen speciale aandacht besteed aan één of meerdere specifieke onderwerpen. Onderstaand kader bevat een overzicht van de specifieke aandachtspunten die in voorgaande publicaties aan bod zijn gekomen.

Overzicht van specifieke onderwerpen in voorgaande trefzekerheidspublicaties

Centraal Economisch Plan 1999

- Vergelijking van de trefzekerheid van de najaarsramingen van de Nederlandse BBP-groei tussen CPB, EC en OESO;
- BBP-ramingen van CPB, NYFER en Consensus Forecasts vergeleken.

Centraal Economisch Plan 2000

- Trefzekerheid van CPB-ramingen voor buitenlandse grootheden.

CPB Memorandum 53 (december 2002)

- Raming en realisatie van contractlonen en consumentenprijsindex;
- Effect van ramingsfouten voor wereldhandel en wisselkoers op trefzekerheid;
- Vergelijking van de trefzekerheid van de voor- en najaarsramingen van de Nederlandse BBP-groei tussen CPB, OESO, NYFER en Consensus Forecasts.

CPB Memorandum 82 (december 2003)

- Trefzekerheid werkloosheids- en werkgelegenheidsramingen;
- Effect van ramingsfouten voor wereldhandel en wisselkoers op trefzekerheid.

CPB Document 77 (februari 2005)

- Welke realisatiecijfers van het CBS raamt het CPB?;
- Is sprake van systematische voorspelfouten?;
- Gevolgen van hectische bewegingen in aandelen- en dollarkoers op trefzekerheid;
- Invloed van de kwaliteit van externe veronderstellingen op trefzekerheid in 12 EU-landen.

CPB Document 106 (januari 2007)

- Vergelijking trefzekerheid CPB met trefzekerheid van drie instituten (OESO, Europese Commissie en IMF), enkele private banken (ABN-AMRO, ING, NIB Capital Bank en Rabobank) en Consensus Forecasts.

CPB Document 131 (oktober 2007)

- Hoe luiden de gehanteerde formules?;
 - Betekenis van trefzekere buitenlandexogenen;
 - Trefzekerheid van verkenningen voor de middellange termijn (MLT);
 - Onzekerheid in CPB-ramingen: een uitsplitsing in determinanten.
-

Bijlage 2 Trefzekerheid van CPB-ramingen

Deze bijlage bevat voor de gebruikelijke kerngrootheden de gemiddelde voorspelfout, de gemiddelde absolute voorspelfout en de ongelijkheidscoëfficiënt van de CPB-ramingen die niet in paragraaf 3.2 zijn opgenomen.

Tabel 1 Trefzekerheid van de CEP-ramingen voor het komende jaar, 1990-2007

	Gemiddelde voorspelfout ^a		Gem. absolute voorspelfout		Ongelijkheidcoëfficiënt	
	1990-2007	1998-2007	1990-2007	1998-2007	1990-2007	1998-2007
Externe data						
Relevant wereldhandelsvolume (%)	0,9	1,1	2,8	2,9	0,8	0,8
Prijs relevante wereldhandel (%)	0,6	- 1,1	3,4	2,9	0,9	0,7
Prijs goedereninvoer (%)	0,2	- 2,3	3,3	3,0	0,9	0,8
Lonen en prijzen						
Contractloon marktsector (%)	- 0,2	- 0,4	0,6	0,6	0,9	1,0
Consumentenprijsindex (CPI) (%)	- 0,2	- 0,6	0,7	0,7	0,9	0,9
Prijs uitvoer goederen (%)	- 0,1	- 2,2	3,1	3,2	1,0	0,8
Prijs bruto binnenlands product (%)	- 0,4	- 0,9	1,0	1,3	1,1	1,1
Volume bestedingen en productie						
Bruto binnenlands product (%)	0,0	0,1	1,1	1,4	0,9	0,9
Goederenuitvoer (%)	0,8	1,0	2,9	3,7	0,8	0,8
Goedereninvoer (%)	0,3	0,9	3,3	3,7	0,9	0,9
Particuliere consumptie (%)	- 0,2	0,1	1,4	1,9	1,0	1,0
Investeringsbedrijven (%) ^b	0,3	1,3	3,7	4,2	0,9	0,9
Arbeidsmarkt						
Werkgelegenheid bedr. (arb. jr.) (%)	- 0,2	0,0	0,8	0,8	0,7	0,7

^a Minteken duidt op onderschatting.

^b Exclusief woningen en voorraden.

Tabel 2 Trefzekerheid van de CEP-ramingen voor het lopende jaar, 1971-2007

	Gemiddelde voorspelfout ^a		Gem. absolute voorspelfout		Ongelijkheidcoëfficiënt	
	1971-2007	1998-2007	1971-2007	1998-2007	1971-2007	1998-2007
Externe data						
Relevant wereldhandelsvolume (%)	- 0,8	- 0,2	1,9	1,6	0,5	0,5
Prijs relevante wereldhandel (%)	- 1,1	- 1,4	2,8	2,1	0,5	0,5
Prijs goederenimport (%)	- 1,2	- 1,9	2,3	2,5	0,3	0,6
Lonen en prijzen						
Contractloon marktsector (%)	- 0,1	0,0	0,4	0,2	0,4	0,3
Consumentenprijsindex (CPI) (%)	- 0,2	- 0,2	0,4	0,3	0,5	0,4
Prijs uitvoer goederen (%)	- 1,0	- 1,6	2,1	2,1	0,4	0,5
Prijs bruto binnenlands product (%)	- 0,1	- 0,2	0,5	0,5	0,4	0,4
Volume bestedingen en productie						
Bruto binnenlands product (%)	- 0,3	- 0,1	1,0	1,0	0,7	0,7
Goederenexport (%)	- 0,9	- 0,4	2,3	2,0	0,5	0,5
Goederenimport (%)	- 0,7	- 0,4	2,3	2,3	0,5	0,6
Particuliere consumptie (%)	- 0,3	0,0	0,9	1,3	0,8	1,0
Investeringen bedrijven (%) ^b	- 0,5	0,4	3,5	3,3	0,6	1,0
Arbeidsmarkt						
Werkgelegenheid bedr. (arb. jr.) (%)	- 0,3	- 0,2	0,7	0,6	0,7	0,5

^a Minteken duidt op onderschatting.

^b Exclusief woningen en voorraden.

Tabel 3 Trefzekerheid van de MEV-ramingen voor het lopende jaar, 1971-2007

	Gemiddelde voorspelfout ^a		Gem. absolute voorspelfout		Ongelijkheidcoëfficiënt	
	1971-2007	1998-2007	1971-2007	1998-2007	1971-2007	1998-2007
Externe data						
Relevant wereldhandelsvolume (%)	- 0,7	- 0,7	1,3	0,9	0,4	0,3
Prijs relevante wereldhandel (%)	0,5	- 0,3	2,1	0,9	0,5	0,2
Prijs goederenimport (%)	0,4	0,1	1,0	0,9	0,1	0,2
Lonen en prijzen						
Contractloon marktsector (%)	0,0	0,0	0,3	0,1	0,3	0,2
Consumentenprijsindex (CPI) (%)	0,1	0,0	0,2	0,1	0,2	0,2
Prijs uitvoer goederen (%)	0,3	- 0,2	1,1	1,0	0,2	0,2
Prijs bruto binnenlands product (%)	- 0,1	- 0,2	0,5	0,5	0,4	0,5
Volume bestedingen en productie						
Bruto binnenlands product (%)	- 0,3	- 0,3	0,7	0,8	0,4	0,6
Goederenexport (%)	- 0,1	- 0,7	1,5	1,6	0,4	0,4
Goederenimport (%)	- 0,3	- 0,7	1,5	1,6	0,3	0,5
Particuliere consumptie (%)	- 0,4	- 0,3	0,8	0,9	0,6	0,6
Investeringen bedrijven (%) ^b	- 0,5	0,1	2,9	3,4	0,5	0,9
Arbeidsmarkt						
Werkgelegenheid bedr. (arb. jr.) (%)	- 0,3	0,0	0,5	0,4	0,5	0,4

^a Minteken duidt op onderschatting.

^b Exclusief woningen en voorraden.

Literatuur

- Berg, P.J.C.M. van den, 1986, De betrouwbaarheid van macro-economische voorspellingen, *Economisch Statistische Berichten*, nr. 3583, 26 november.
- Blix, M., J. Wadefjord, U. Wienecke en M. Adahl, 2001, How good is the forecasting performance of major institutions?, *Economic Review* (Penning - Och Valutapolitik), 3/2001, pag. 37-68.
- CBS, 2005a, Nationale rekeningen 2004 –Revisie 2001, Beschrijving en uitkomsten revisiejaar 2001, augustus.
- CBS, 2005b, Revisie Nationale rekeningen: Bijstellingen 2001-2004 (www.cbs.nl).
- CPB, 1999, *Centraal Economisch Plan 1999*, Sdu Uitgevers, Den Haag
- CPB, 2000, *Centraal Economisch Plan 2000*, Sdu Uitgevers, Den Haag
- CPB, 2005, *Macro Economische Verkenning 2006*, Sdu Uitgevers, Den Haag
- CPB, 2006, *Macro Economische Verkenning 2007*, Sdu Uitgevers, Den Haag
- Dankmeyer, B., 2008, *Bijstellingen volumemutaties Nationale Rekeningen 2005*, CBS, Den Haag.
- Don, F.J.H., 2001, Forecasting in macroeconomics: a practitioner's view, *De Economist*, vol. 149, nr. 2.
- Don, F.H.J. en J.P.V. Verbruggen, 2006, Models and methods for economic policy: 60 years of evolution at CPB, *Statistica Neerlandica*, vol. 60, nr. 2, pag. 145-170.
- Elbourne, A.C., H.C. Kranendonk, R.E. Luginbuhl, B.C. Smid en M.W.A.M. Vromans, 2008, Evaluating CPB's published GDP growth forecasts, CPB Document 172, oktober.
- Franses, P. H., H.C. Kranendonk en D. Lanser, 2007, On the optimality of expert-adjusted forecasts, CPB Discussion Paper 92, december.
- Jonung, L. and M. Larch, 2004, Improving fiscal policy in the EU: the case for independent forecasts, *Economic Paper*, nr. 210, Europese Commissie, juli.

Kazemier, B., H. Nijmeijer en R. Hijman, 2004, How to judge the reliability of provisional National Accounts, Paper prepared for the European Conference on Quality and Methodology in Official Statistics, Mainz, 24-26 mei.

Keereman, F., 2003, External assumptions, the international environment and the track record of the Commission Forecasts, *Economic Paper*, nr. 189, Europese Commissie, Directorate-General for Economic and Financial Affairs, september.

Kranendonk, H.C. en J.P. Verbruggen, 2002, Trefzekerheid van CPB-prognoses voor de jaren 1971-2002, CPB Memorandum 53, december.

Kranendonk, H.C. en J.P. Verbruggen, 2003, Trefzekerheid van CPB-prognoses voor de jaren 1971-2003, CPB Memorandum 82, december.

Kranendonk, H.C., J.P.M. Bonenkamp en J.P. Verbruggen, 2003, De CPB-conjunctuurindicator geactualiseerd en gereviseerd, CPB Document 40, november.

Kranendonk, H.C. en J.P. Verbruggen, 2005, Trefzekerheid van CPB-prognoses voor de jaren 1971-2003, CPB Document 77, februari.

Kranendonk, H.C. en J.P. Verbruggen, 2006a, Trefzekerheid van korte-termijnramingen van het CPB voor de jaren 1971-2004, CPB Document 106, januari.

Kranendonk, H.C. en J.P. Verbruggen, 2006b, SAFFIER, Een 'multi purpose'-model van de Nederlandse economie voor analyses op korte en middellange termijn, CPB Document 123, juni.

Kranendonk, H.C. en J.P. Verbruggen, 2006c, Trefzekerheid van korte-termijnramingen en middellange-termijnverkenningen, CPB Document 131, oktober.

Lanser, D. and H.C. Kranendonk, 2008, Investigating uncertainty in macroeconomic forecasts by stochastic simulation, CPB Discussion Paper 112, september.

Lenain, P., 2002, What is the track record of OECD Economic Projections?, Economics Department, OECD, Paris.

Melander, A., G Sismanidis and D. Grenouilleau, 2007, The track record of the Commission's forecasts- an update, *Economic Paper*, nr. 291, Europese Commissie, oktober.

Nijmeijer, H. en R. Hijman, 2004, Bijstellingen ramingen economische groei, Centraal Bureau voor de Statistiek, De Nederlandse conjunctuur - 2004, aflevering 1.

Timmerman, A., 2007, An Evaluation of the World Economic Outlook Forecasts. IMF Working Paper 59, maart.

Vogel, L., 2007, How do the OECD Growth Projections for the G7 Economies Perform?: A Post-Mortem, OECD Economics Department Working Papers, nr. 573, OECD Publishing.

