

# De betrouwbaarheid van macro-economische voorspellingen

Een bijdrage van Grecon

De voorspelkracht van een econom(etr)isch model wordt in de praktijk veelal afgemeten aan de resultaten van een onderzoek naar de mate waarin het model in het verleden in staat is geweest of in staat zou zijn geweest realistische voorspellingen te genereren. Als deze resultaten gunstig uitvallen, wordt wel de conclusie getrokken dat de gebruikte voorspelprocedure bevredigend is en dat het gebruikte model voor de toekomst een vergelijkbare voorspelkracht zal vertonen. Een andere benaderingswijze om de betrouwbaarheid van voorspellingen aan te geven is het construeren van betrouwbaarheidsintervallen met behulp van de geschatte standaarddeviatie van de verstoringen. In dit artikel worden de resultaten gegeven van toepassing van beide criteria op voorspellingen van de zes belangrijkste macro-economische variabelen van het Grecon-model in de jaren 1977 – 1984. Wat betreft het eerstgenoemde criterium worden de resultaten tevens vergeleken met de resultaten van de voorspellingen die het Centraal Planbureau jaarlijks in zijn Macro Economische Verkenning publiceert alsmede met het gemiddelde van CPB- en Grecon-voorspellingen ex ante. Tevens wordt een overzicht gegeven van mogelijke oorzaken van voorspelfouten.

## DR. W. VOORHOEVE\*

Of een voorspelling betrouwbaar kan worden geacht, hangt af van de verwachting omtrent de mogelijke fluctuaties van de voorspelfout, het absolute verschil tussen voorspelling en realisatie. Een bekende maatstaf hiervoor is de geschatte standaarddeviatie van de verstoringen, die kan worden gebruikt om een betrouwbaarheidsinterval voor de voorspelling te geven in plaats van een puntvoorspelling. Indien op deze wijze voor een aantal opeenvolgende jaren intervallen zijn berekend, dan is het vervolgens mogelijk na te gaan in hoeveel gevallen de realisaties in het interval zijn terecht gekomen. Een nadeel van zulke voorspelintervallen is dat de lokatie afhankelijk is van de veronderstelde waarden van de gepredetermineerde variabelen. Wil men daarom uit een confrontatie van de realisaties met de voorspelintervallen conclusies trekken ten aanzien van de betrouwbaarheid van het model, dan dient men de intervallen opnieuw te berekenen, nu met gebruik van de gerealiseerde waarden van de gepredetermineerde variabelen.

Helaas worden niet voor alle macro-economische voorspellingen standaarddeviaties berekend om de betrouwbaarheid aan te geven. In dergelijke gevallen kan men realisaties slechts met puntvoorspellingen vergelijken. Niettemin kan een dergelijke confrontatie iets zeggen over de voorspelkracht in het verleden. Indien men vervolgens veronderstelt dat de beschouwde voorspelprocedure ook in de toekomst even goed of even slecht zal functioneren, geven de conclusies op grond van genoemde confrontatie enig inzicht in de betrouwbaarheid van te genereren voorspellingen. In de literatuur worden voor een dergelijke confrontatie diverse mogelijkheden aangegeven <sup>1)</sup>, die ruwweg in twee groepen kunnen worden onderscheiden. De eerste groep is gebaseerd op een regressie tussen realisaties en voorspellingen. De tweede groep kan worden samengevat onder het begrip 'gemiddelde voorspelfout',

waarbij veelal een kwadratisch of een absoluut gemiddelde wordt gehanteerd.

In dit artikel wordt de betrouwbaarheid van de Grecon-voorspellingen onderzocht. Nagegaan zal worden in hoeverre de standaarddeviaties van de geschatte herleidvormverstoringen (op basis waarvan de betrouwbaarheidsintervallen worden bepaald) kunnen worden gehanteerd als maatstaf voor de betrouwbaarheid van de voorspellingen. Tevens zullen de realisaties worden geconfronteerd met zowel de gepubliceerde voorspellingen ex ante als de voorspellingen ex post. Voorafgaand aan de presentatie van de resultaten volgen eerst enkele opmerkingen over het Grecon-model zelf, alsmede over de mogelijke oorzaken van voorspelfouten.

### Het Grecon-model

In 1976 werd aan het Econometrisch Instituut van de Rijksuniversiteit Groningen het econometrische model Grecon geconstrueerd, waarmee jaarlijks voorspellingen (inclusief de standaarddeviaties van de geschatte herleidvorm-verstoringen) worden gemaakt van de belangrijkste Nederlandse macro-economische variabelen.

\* Wetenschappelijk hoofdmedewerker aan het Econometrisch Instituut van de Rijksuniversiteit Groningen. De auteur is dank verschuldigd aan drs. H.W.A. Dietzenbacher en drs. A.G.M. Steerneman voor hun kritische opmerkingen op een eerdere versie van dit artikel.

1) Zie b.v. L.R. Klein, *An essay on the theory of economic prediction*, Chicago, 1971; M.D. Intriligator, *Econometric models, techniques and applications*, Amsterdam, 1978; D.J. Smyth, Short-run macroeconomic forecasting: the OECD performance, *Journal of Forecasting*, 1983, nr. 2, blz. 37 – 49.

Deze voorspellingen worden gepubliceerd als alternatief voor de officiële voorspellingen van het Centraal Planbureau (CPB). Het Grecon-model is een lineair simultaan econometrisch model, waarvan de gedragsvergelijkingen zijn gespecificeerd op grond van zowel economisch relevante als statistisch significante causale verbanden. Het model is relatief klein (circa tien gedragsvergelijkingen) en de parameters zijn geschat met de methode van de kleinste kwadraten in twee ronden, op basis van officieel, voor ieder toegankelijk cijfermateriaal vanaf 1952, dat grotendeels is ontleend aan de Nationale Rekeningen van het Centraal Bureau voor de Statistiek. Vanaf 1977 worden jaarlijks de Grecon-voorspellingen gepubliceerd en wordt een rapport uitgebracht waarin een zo volledig mogelijke verantwoording wordt afgelegd over onder meer het gebruikte cijfermateriaal, van de wijzigingen in de modelspecificatie en van de veronderstellingen die ten grondslag liggen aan de voorspellingen.

## Oorzaken van voorspelfouten

Indien de voorspellingen zijn gemaakt met behulp van een econometrisch model, kunnen voorspelfouten (de verschillen tussen de voorspelde en de gerealiseerde waarden) het gevolg zijn van vier onzekerheden: de gekozen modelspecificatie; de waarden van de parameters; de waarde van de storingsterm in de te voorspellen periode; en de waarden van de gepredetermineerde variabelen. We gaan nu nader in op de genoemde oorzaken.

*De model-specificatie.* Een model is in wezen niet meer dan een benadering van het systeem zoals dit in werkelijkheid functioneert. Deze benadering kan de volgende afwijkingen ten opzichte van de werkelijkheid vertonen:

- het is niet mogelijk rekening te houden met alle relaties binnen een economisch systeem. Zowel de relaties als de variabelen in het model kunnen niet beschouwd worden als een volledige beschrijving van de werkelijkheid;
- het is mogelijk dat een model niet-relevante relaties of variabelen bevat, bij voorbeeld doordat gehanteerde economische theorieën in werkelijkheid niet opgaan voor het beschouwde systeem of de beschouwde periode;
- zelfs indien een model beschouwd kan worden als een bevredigende beschrijving van het systeem in het verleden, behoeft dit niet op te gaan voor de te voorspellen periode. Er kan een structuurverandering optreden tussen verleden en toekomst;
- ook indien verondersteld kan worden dat de in het model opgenomen causaliteiten en variabelen relevant zijn, bestaat de mogelijkheid dat de relaties op een andere wijze dienen te worden gedefinieerd. Men kan wat dit betreft bij voorbeeld denken aan een in feite niet-lineaire reactie die wordt benaderd door een lineaire functie.

*De waarden van de model-parameters.* Het feit dat een model geschatte waarden in plaats van de echte waarden voor de parameters bevat, is op zich al een vanzelfsprekende bron van voorspelfouten. Daarnaast kunnen onderstaande omstandigheden van invloed zijn op de geschatte waarden zelf:

- toepassing van verschillende (al of niet simultane) schattingsmethoden resulteert in verschillende sets van schattingen van parameter waarden en daardoor in verschillende sets van voorspellingen;
- het gebruik van verschillende steekproefperiodes voor het schatten van de parameters kan eveneens tot alternatieve uitkomsten leiden. Wat dit betreft kan men zich afvragen of het redelijk is cijfermateriaal uit de jaren vijftig te gebruiken voor het schatten van een model dat vervolgens wordt gebruikt voor het genereren van voorspellingen in de jaren tachtig;
- de betrouwbaarheid van de schatting van de parameters hangt vanzelfsprekend af van de betrouwbaarheid

en van de beschikbaarheid van adequaat cijfermateriaal. Voor Nederland is men aangewezen op door het CBS in de Nationale Rekeningen gepubliceerde jaarcijfers. Wil men in het najaar van jaar  $t$  een model schatten ten behoeve van voorspellingen voor jaar  $t+1$  op basis van de recentste Nationale Rekeningen zoals Grecon doet, dan heeft men de beschikking over cijfermateriaal voor de jaren  $t-1$  en eerder. Bovendien hebben de gegevens voor jaar  $t-1$  en  $t-2$  een voorlopig karakter.

*De waarde van de storingsterm.* Een voorspelling op basis van een model is in wezen een puntschatting van de betreffende variabele in een bepaald jaar, waarbij de storingsterm gelijk aan 0 wordt gesteld. Per definitie is de waarde van de storingsterm onvoorspelbaar.

*De waarden van de gepredetermineerde variabelen.* Voorspellingen gemaakt met een simultaan model, zijn conditioneel, gegeven de waarden van de gepredetermineerde variabelen in het te voorspellen jaar. In de praktijk worden deze waarden niet door het model geschat, maar op grond van additionele kwantitatieve informatie over de economische situatie geraamd. Het is duidelijk dat de kwaliteit van deze raming van invloed is op de kwaliteit (en dus op de betrouwbaarheid) van de voorspelprocedure als geheel. Dit geldt te meer voor variabelen die zich moeilijk laten ramen. Ashley merkt hierover op: „An explanatory variable may appear to be quite significant, and be highly useful even outside the sample when actual values of it are available. Yet the errors in forecasting it may nevertheless overshadow its explanatory power”. 2)

Op basis van het Grecon model, versie 79-D, werden een aantal experimenten uitgevoerd met het doel na te gaan hoe groot de invloed op de kwaliteit van de voorspellingen is van het voorlopig cijfermateriaal uit de Nationale Rekeningen, van de geraamde waarden van de exogene invloeden die op het model inwerken en van de onzekerheden die voortvloeien uit het werken met modellen als benaderende beschrijving van de werkelijkheid 3). Hiertoe is het model met cijfermateriaal voor de periode 1952 – 1976 geschat op basis van drie verschillende datasets: de laatste twee jaren voorlopig, het laatste jaar voorlopig en ten slotte alle cijfers definitief. Met elk van deze drie zijn vervolgens voor 1976 (het recentste jaar waarvoor ten tijde van de experimenten definitieve gegevens beschikbaar waren) voorspellingen berekend op basis van eveneens drie verschillende sets waarden voor de gepredetermineerde variabelen. In totaal dus 9 voorspellingen voor elke variabele. Deze voorspellingen zijn vervolgens geconfronteerd met wederom drie verschillende sets van realisaties voor 1976. Behalve deze confrontaties is voor elke variabele de standaarddeviatie van de herleide-vormverstorings berekend. Onder andere werden de volgende conclusies uit het onderzoek getrokken:

- het effect op de voorspellingen van het gebruik van voorlopige gegevens voor de laatste twee jaren van de schattingsperiode bleek te kunnen worden verwaarloosd;
- niet verwaarloosbaar klein bleken de verschillen ten gevolge van de onzekerheid met betrekking tot de waarden van de gepredetermineerde variabelen;
- de grootste bron van onzekerheid waren de herleide-vormverstorings, gezien de relatief hoge waarden van de standaarddeviaties hiervan.

Het hierboven aangehaalde onderzoek is op twee punten beperkt te noemen. Enerzijds is slechts een versie van het Grecon-model in beschouwing genomen, anderzijds is slechts de invloed onderzocht van het al of niet voorlopige karakter van de cijfers voor 1975 en 1976. In het hierna volgende wordt evenwel verondersteld dat de eerstgenoemde

2) R. Ashley, On the usefulness of macroeconomic forecasts as inputs to forecasting models, *Journal of Forecasting*, 1983, nr. 2, blz. 211 – 223.

3) De resultaten van de hier bedoelde experimenten zijn gepubliceerd in *ESB*, 12 maart 1980, blz. 306 – 311.

**Tabel 1. Gemiddelde, standaarddeviatie, maximum- en minimum-waarde van de jaarlijkse procentuele toename van enkele belangrijke variabelen van het Grecon-model, 1952 – 1976**

Variabele	Gemiddelde	Standaarddeviatie	Maximum	Minimum
Particuliere consumptie	5,0	2,3	7,8	0,3
Invoer van goederen	8,4	8,1	24,0	-10,6
Afzet van goederen	6,2	3,5	10,9	-1,2
Consumptieprijs	4,5	3,4	11,5	-1,0
Loonvoet	9,9	3,6	15,7	2,4
Particuliere werkgelegenheid	1,3	1,6	3,8	-1,6
Uitvoer van goederen	9,4	5,0	16,7	-4,7

de conclusie eveneens voor andere modelversies en voor andere jaren opgaat. Hierdoor is het mogelijk dat de verschillen voorspellingen ex ante (dat zijn de in de praktijk gegenereerde voorspellingen, dat wil zeggen inclusief het deels voorlopige karakter van het cijfermateriaal en inclusief de onzekerheid ten aanzien van de waarden van de gepredetermineerde variabelen) en voorspellingen ex post (in dit artikel te definiëren als voorspellingen gegenereerd met een model waarvan de parameters zijn geschat met definitief cijfermateriaal en waarbij voor de gepredetermineerde variabelen definitieve waarden zijn gesubstitueerd) hoofdzakelijk kunnen worden toegeschreven aan het gebruik van onjuiste waarden voor de gepredetermineerde variabelen.

## Grecon-voorspellingen 1977 – 1984

Het onderhavige onderzoek heeft betrekking op de zes belangrijkste endogene variabelen van het Grecon-model, die in alle tot nu toe geconstrueerde modelversies dezelfde definitie hebben. Het betreft de volgende variabelen:

- particuliere consumptie  $c$ ;
- invoer van goederen  $m$ ;
- afzet van bedrijven  $v'$ ;
- consumptieprijs  $p_c$ ;
- loonvoet  $l$ ;
- particuliere werkgelegenheid  $a$ .

Voorafgaand aan de presentatie van de uitkomsten volgt een korte samenvatting van het verloop van de beschouwde zes variabelen en van de uitvoer van goederen in de periode 1977 – 1984 in vergelijking met de daaraan voorafgaande jaren.

In een aantal opzichten verschilt de periode 1977 – 1984 van de daaraan voorafgaande jaren 1952 – 1976. In 1976 leek de invloed van de oliecrisis op de Nederlandse economie nagenoeg te zijn verdwenen. De uitvoer van goederen nam in 1976 met bijna 12% toe, na een daling van 5% in 1975. Hierdoor leek zich eveneens het verloop van de afzet en de invoer van goederen te herstellen: de afzet nam in 1976 toe met meer dan 6%, een groei die vergelijkbaar is met de voorafgaande jaren; de invoer steeg zelfs met bijna 9%, een hogere toename dan tevoren. De cijfers voor 1977 laten evenwel zien dat sprake was van een zeer incidenteel herstel: de uitvoer daalde met ruim 2%; de afzet steeg slechts met 3% en de invoer met circa 2,5%. In tabel 1 en tabel 2 is een en ander weergegeven.

Niet zelden bedraagt de jaarlijkse procentuele verandering in de jaren na 1976 meer dan twee maal de standaarddeviatie van het gemiddeld in de periode 1952 – 1976. Voor de particuliere consumptie is dit bij voorbeeld het geval voor alle jaren 1980 – 1984; voor de afzet van goederen voor de jaren 1981 en 1982; en voor de particuliere werkgelegenheid voor de jaren 1981 – 1983. Bovendien blijkt de jaarlijkse procentuele verandering na 1976 relatief vaak te liggen beneden het minimum van de gehele periode 1952 – 1976.

**Tabel 2. Gemiddelde jaarlijkse procentuele toename van enkele belangrijke variabelen van het Grecon-model, perioden 1977 – 1979, 1980 – 1982 en 1983 – 1984**

Variabele	Periode		
	1977 – 1979	1980 – 1982	1983 – 1984
Particuliere consumptie	3,9	-1,2	-0,1
Invoer van goederen	5,1	-2,7	4,7
Afzet van goederen	3,8	-1,1	2,5
Consumptieprijs	4,7	6,2	2,6
Loonvoet	7,2	5,4	2,5
Particuliere werkgelegenheid	0,9	-1,6	-1,6
Uitvoer van goederen	3,7	0,4	5,8

Samenvattend kan worden gesteld dat de economische situatie in de jaren 1977 – 1984 aanzienlijk afwijkt van de daaraan voorafgaande jaren. Indien het voorspellen van de economische ontwikkeling in deze periode gebaseerd is op een modelspecificatie en op schatting van de parameters vanuit de jaren 1952 – 1976 (zoals met het Grecon – model het geval is), is het duidelijk dat het model moeite zal hebben de jaren 1977 – 1984 te verklaren. Daarom zullen tevens confrontaties worden gepresenteerd van realisaties met alle puntvoorspellingen ex post, berekend met de diverse specificaties van de Grecon-modellen. Op deze wijze kan worden nagegaan of de voorspelkracht van het model beter wordt naarmate de specificatie is gebaseerd op meer informatie betreffende de periode 1977 – 1984.

## Voorspelintervallen

Hiervoor is reeds vermeld dat bij de jaarlijkse Grecon-voorspellingen tevens de standaarddeviaties van de geschatte herleide-vormverstorings worden gepubliceerd. Tabel 3 geeft hiervan de samenvatting. De interpretatie van deze cijfers is als volgt. Als deze kunnen worden beschouwd als een redelijke benadering van de standaarddeviaties van de voorspellingen 4) en als het achterliggende model als juist mag worden beschouwd, en ten slotte als de gesubstitueerde waarden van de gepredetermineerde variabelen juist zijn, zal de realisatie gemiddeld in circa 95% binnen een interval van twee maal de standaarddeviatie rondom de voorspelling liggen; binnen een interval van één maal de standaarddeviatie zal gemiddeld circa 68% liggen. Het is duidelijk dat de voorspellingen des te betrouwbaarder zijn naarmate de standaarddeviatie kleiner is. Op grond van deze constatering kan uit tabel 3 de conclusie worden getrokken dat de voorspellingen van alle beschouwde variabelen (uitgezonderd de particuliere werkgelegenheid) in de loop van de periode 1977 – 1986 steeds betrouwbaarder blijken te worden. Niettemin is het zo, dat de cijfers relatief (beschouwd ten opzichte van de waarde van de voorspelling zelf) nog steeds aan de hoge kant zijn. Om een voorbeeld te noemen: voor 1986 is door Grecon een stijging voorspeld van de particuliere consumptie van 2,0% met een standaarddeviatie van 1,3. Dit betekent dat (onder de reeds genoemde veronderstellingen) de kans circa 68% bedraagt dat de stijging tussen de 0,7% en de 3,3% zal uitkomen. De kans bedraagt voorts circa 95% dat de stijging zal uitkomen tussen de -0,6% en de +4,6%.

Voor een confrontatie van realisaties en voorspelinter-

4) Formeel is de variantie van de voorspelfout gelijk aan de som van de varianties van respectievelijk de verstoring en de voorspelling. De grootste variabiliteit wordt evenwel veroorzaakt door de verstoring. Voor een numerieke uitwerking van deze grootheden zie R.H. Ketelapper, B. Bos, M.A. Kooyman en W. Voorhoeve, A simultaneous econometric model for the Dutch economy, *Statistica Neerlandica*, jg. 31, 1977, nr. 3, blz. 141 – 159.

Tabel 3. Standaarddeviaties van de geschatte herleide-vormverstorings (ex ante) van de zes beschouwde variabelen, 1977 - 1986

Jaar	Variabele					
	c	m	v'	p <sub>c</sub>	l	a
1977	2,1	5,0	2,2	2,7	2,4	0,9
1978	1,9	4,1	1,5	2,7	2,3	0,8
1979	1,6	3,4	1,1	3,0	2,5	0,6
1980	1,5	4,2	1,5	2,8	2,4	0,7
1981	1,5	2,5	1,0	2,9	2,3	0,7
1982	1,5	2,5	1,0	2,1	2,0	0,8
1983	1,5	2,4	0,9	2,0	1,4	0,8
1984	1,4	3,0	0,9	1,9	1,2	1,0
1985	1,3	2,8	0,8	1,8	1,6	1,0
1986	1,3	2,5	0,9	1,9	1,6	1,0

vallen kan gebruik worden gemaakt van de grootheid die aangeeft het aantal malen dat een realisatie ten hoogste 1, 2 of 3 maal de standaarddeviatie van de voorspelling af ligt. Theoretisch (indien het achterliggende model juist zou zijn) vindt dit plaats in respectievelijk 68%, 95% en 99% van de gevallen. Tabel 4 geeft hiervan een overzicht voor de beschouwde variabelen in de jaren 1977 - 1984. Eerder onderzoek heeft uitgewezen dat met namen de eerste modelspecificatie (versie 77-B) enigszins te wensen overliet wat betreft de voorspelkracht 5). Om deze reden is tevens tabel 5 opgenomen die hetzelfde cijfermateriaal bevat als tabel 4, met weglating van de cijfers voor 1977.

De cijfers dienen als volgt te worden geïnterpreteerd, aannemende dat de achterliggende modelspecificatie juist is. Per variabele mag men verwachten dat in de jaren 1977 - 1984 (tabel 4) 5 à 6 keer de voorspelfout ten hoogste 1 maal de standaarddeviatie bedraagt; 7 à 8 keer ten hoogste 2 maal de standaarddeviatie en 8 keer ten hoogste 3 maal de standaarddeviatie. Voor de beschouwde zes variabelen als totaal zijn deze getallen respectievelijk 32 à 33, 45 à 46 en 47 à 48. Voor de jaren 1978 - 1984 (tabel 5) mag men per variabele verwachten dat de voorspelfout ten hoogste 1, 2 of 3 maal de standaarddeviatie bedraagt in respectievelijk 4 à 5, 6 à 7 en 7 gevallen. Voor het totaal zijn deze getallen respectievelijk 28 à 29, 40 en 41 à 42.

Indien het er om gaat de praktische voorspelkracht van het model te beoordelen, kunnen de frequenties ex ante worden vergeleken met de theoretische frequenties. Immers, ex ante voorspellen is wat men in de praktijk doet. De hieruit te trekken conclusies zeggen iets over de voorspelprocedure als geheel. Wil men evenwel een oordeel uitspreken over de juistheid van de modelspecificatie, dan dient men de frequenties ex post te bezien, aangezien het verschil tussen ex ante en ex post in de beschouwde situatie hoofdzakelijk kan worden toegeschreven aan het gebruik van onjuiste waarden van de gepredetermineerde variabelen (zie de hierboven aangehaalde experimenten met het model 79-D).

De cijfers in tabel 4 en tabel 5 dienen met voorzichtigheid te worden gehanteerd aangezien zij de resultaten zijn van een zeer beperkt aantal waarnemingen: 6 variabelen over 7 of 8 jaar. Wil men evenwel aan de cijfers enige betekenis ontleenen, dan lijken de volgende conclusies getrokken te kunnen worden:

- bezien wij de totaalkolommen in beide tabellen, dan blijkt dat in 60% à 70% van de gevallen de realisatie (zowel ex ante als ex post) ten hoogste een keer de standaarddeviatie van de voorspelling ligt. Ten hoogste twee keer de standaarddeviatie ligt 80% à 90%. Met name wat betreft de voorspellingen ex post zijn de resultaten met weglating van 1977 beter te noemen, een bevestiging van reeds eerdere conclusie dat de voorspelkracht van het model 77-B te wensen overlaat;
- de frequenties per variabele laten een wisselend beeld zien. Met name blijken de voorspellingen van de parti-

Tabel 4. Aantal malen dat de voorspelfout ten hoogste k maal de standaarddeviatie van de geschatte herleide-vormverstoring bedraagt in de periode 1977 - 1984

	Variabele						Totaal	
	c	m	v'	p <sub>c</sub>	l	a	absoluut	procentueel
Ex ante								
- k=1	6	4	3	8	8	3	32	67
- k=2	7	7	6	8	8	6	42	88
- k=3	8	8	8	8	8	8	48	100
Ex post								
- K=1	6	6	6	5	5	3	31	65
- k=2	6	8	7	6	6	4	37	77
- k=3	8	8	8	8	7	6	45	94

Tabel 5. Aantal malen dat de voorspelfout ten hoogste k maal de standaarddeviatie van de geschatte herleide-vormverstoring bedraagt in de periode 1978 - 1984

	Variabele						Totaal	
	c	m	v'	p <sub>c</sub>	l	a	absoluut	procentueel
Ex ante								
- k=1	5	3	2	7	7	2	26	62
- k=2	6	6	5	7	7	5	36	86
- k=3	7	7	7	7	7	7	42	100
Ex post								
- K=1	6	5	6	5	5	3	30	71
- k=2	6	7	7	6	6	4	36	86
- k=3	7	7	7	7	7	5	40	95

culiere werkgelegenheid (a) in alle gevallen slecht te zijn. Voor de overige variabelen komen de frequenties voor wat betreft de voorspellingen ex post in grote lijnen overeen met wat men mag verwachten, zeker indien het jaar 1977 buiten beschouwing wordt gelaten. Ex ante laten de voorspellingen voor de invoer (m) en de afzet van goederen (v') te wensen over; dit wordt kennelijk veroorzaakt door een onjuiste substitutie van de waarden van de gepredetermineerde variabelen; - de resultaten zijn aanzienlijk beter te noemen indien de slechte voorspellingen van de particuliere werkgelegenheid (a) buiten beschouwing worden gelaten. Zeker als tevens het jaar 1977 niet in aanmerking wordt genomen, kan op basis van de frequenties in tabel 5 worden berekend dat in dat geval 69% van de realisaties niet meer dan een keer de standaarddeviatie van de voorspellingen ex ante afwijkt; voor de voorspellingen ex post bedraagt dit percentage zelfs 77. Ten hoogste twee keer de standaarddeviatie zijn deze percentages respectievelijk 89 en 91.

## Puntvoorspellingen ex ante

Een beoordeling van puntvoorspellingen wordt in dit artikel uitgevoerd met behulp van de ongelijkheidscoëfficiënt van Theil (6). Een nadeel van het gebruik van deze maatstaf is dat geen objectief oordeel kan worden gegeven of een voorspelling als redelijk kan worden beschouwd of niet. Een perfecte voorspelling levert een waarde gelijk aan 0; naïeve voorspellingen (geen verandering) leveren een coëfficiënt van 1. Op grond hiervan wordt wel als criterium gebruikt dat een waarde tussen 0 en 1 redelijk is en dat een waarde groter dan 1 slecht is, omdat dan beter naïef voorspeld had kunnen worden. Op deze wijze wordt

5) Zie b.v. W. Voorhoeve, *Short-term economic forecasting in practice: the Grecon case*, proefschrift, Rijksuniversiteit Groningen, 1986.  
6) Zie H. Theil, *Applied economic forecasting*, Amsterdam, 1966.

Tabel 6. Ongelijkheidscoëfficiënt van Theil voor de periode 1977 – 1984 per variabele voor het totaal van de voorspellingen ex ante van CPB, Grecon en het gemiddelde van beide

Bron van de voorspelling	Variabele						Totaal
	c	m	v'	p <sub>c</sub>	l	a	
CPB	0,66	0,78	0,63	0,16	0,27	0,77	0,59
GRECON	0,58	0,82	0,61	0,25	0,25	0,69	0,57
Gemiddelde	0,56	0,79	0,60	0,16	0,22	0,70	0,56

de genoemde maatstaf dan ook gehanteerd als vergelijkingsmaat tussen verschillende voorspellingsuitkomsten.

Ook in dit artikel zal de ongelijkheidscoëfficiënt als vergelijkingsmaat worden gehanteerd. Daartoe zullen de door het Centraal Planbureau jaarlijks in haar Macro Economische Verkenning gepubliceerde voorspellingen ex ante als vergelijking dienen. Behalve de CPB- en Grecon-voorspellingen ex ante zullen tevens worden gebruikt de voorspellingen die worden berekend als een gewoon rekenkundig gemiddelde van beide. Deze soort voorspelling is gekozen naar aanleiding van het feit dat empirisch onderzoek laat zien dat een gemiddelde van diverse voorspellingen vaak een betere voorspelling blijkt te zijn dan elk van de beschouwde voorspellingen 7).

Uit tabel 6 blijkt dat het verschil in voorspelkracht tussen CPB en Grecon klein is. De ongelijkheidscoëfficiënt laat een iets betere voorspelkracht zien van Grecon wat betreft de particuliere consumptie en de werkgelegenheid, terwijl het omgekeerde kan worden gezegd van de invoer van goederen en de consumptieprijs. Over het geheel lijken de Grecon-voorspellingen enigszins beter dan de CPB-voorspellingen, hoewel het verschil uiterst klein is 8). Op een enkele kleine uitzondering na blijkt ook hier het gemiddelde van beide voorspellingen een lagere ongelijkheidscoëfficiënt op te leveren dan elk van beide.

## Puntvoorspellingen ex post

Jaarlijks wordt de specificatie van het Grecon-model aangepast op basis van nieuw voorhanden cijfermateriaal en op basis van nieuw verworven inzichten. De in tabel 3 gepresenteerde standaarddeviaties van de geschatte herleide-vormverstorings laten voor de beschouwde variabelen zien dat deze aanpassingen hebben geleid tot een lagere standaarddeviatie, met uitzondering van die van de particuliere werkgelegenheid. Een lagere standaarddeviatie suggereert een betere modelspecificatie met als gevolg betrouwbaardere voorspellingen (indien de waarden van de gepredetermineerde variabelen juist zijn). Wij kunnen deze conclusie nader onderbouwen door niet alleen het verloop van de standaarddeviaties te bezien, maar door opnieuw met elk van de tien versies van het Grecon-model voorspellingen te doen en deze uitkomsten te vergelijken met de realisaties. Vervolgens kunnen de resultaten van de modelversies met elkaar worden vergeleken.

Hiertoe hebben wij de volgende experimenten uitgevoerd. Met elk van de modelversies 77-B tot en met 86-B 9) zijn voor de periode 1977 – 1984 voorspellingen ex post berekend voor de beschouwde zes variabelen. De voorspelling ex post voor jaar t op basis van modelversie X is berekend door dit model X te herschatten op basis van de jaren 1952 t/m t-1. Vervolgens zijn de gerealiseerde waarden van de gepredetermineerde variabelen gesubstitueerd. Op deze wijze kunnen de verschillen tussen de uitkomsten van verschillende modelversies niet worden toegeschreven aan het gebruikte cijfermateriaal (steekproef en/of gepredetermineerde variabelen), maar uitsluitend aan de modelspecificatie.

De uitkomsten van de experimenten, wederom uitgedrukt in de ongelijkheidscoëfficiënt van Theil, zijn weerge-

Tabel 7. Ongelijkheidscoëfficiënt van Theil voor de gehele periode 1977 – 1984 per variabele en totaal: Grecon voorspellingen ex post met de verschillende modelversies

Voorspelmodel	Variabele						Totaal
	c	m	v'	p <sub>c</sub>	l	a	
77-B	1,16	0,67	0,97	0,54	0,81	0,78	0,85
78-A	0,87	0,58	0,56	0,45	0,79	0,73	0,68
79-D	0,91	0,90	0,42	0,63	0,68	0,72	0,73
80-C	0,51	0,72	0,27	0,50	0,57	0,85	0,60
81-A	0,62	0,28	0,28	0,46	0,54	0,90	0,56
82-B	0,34	0,32	0,30	0,51	0,61	0,76	0,50
83-B	0,34	0,32	0,33	0,55	0,57	0,77	0,51
84-D	0,46	0,47	0,31	0,27	0,32	0,83	0,48
85-B	0,43	0,42	0,29	0,21	0,32	0,81	0,45
86-B	0,41	0,32	0,28	0,23	0,30	0,77	0,42

geven in tabel 7. Hierin wordt in grote lijnen de conclusie bevestigd dat, met uitzondering van de particuliere werkgelegenheid, alle beschouwde variabelen op bevredigender wijze worden voorspeld naarmate de modelversie van recentere datum is. Ook de ongelijkheidscoëfficiënt voor het totaal (laatste kolom) vertoont een steeds lagere waarde.

Experimenten zoals het zojuist beschrevene kunnen pas worden uitgevoerd wanneer cijfermateriaal voorhanden is omtrent de realisaties van zowel de gepredetermineerde variabelen als de voorspelde variabelen, dus pas achteraf. Voor de praktijk, waarbij voorspeld moet worden in onzekerheid, heeft dit vanzelfsprekend minder waarde. Niettemin wijzen de ruwe overeenkomsten tussen het verloop van de cijfers in tabel 3 (welke gegevens wel ex ante voorhanden kunnen zijn) en de cijfers in tabel 7 in de richting dat de standaarddeviaties kunnen worden gezien als aanwijzing vooraf over de betrouwbaarheid van de voorspellingen.

## Slotopmerkingen

Tot slot volgt een samenvatting van de belangrijkste conclusies die hierboven van de experimenten zijn getrokken:

- de standaarddeviatie van de geschatte herleide-vormverstorings is een redelijke maat voor de (on)betrouwbaarheid van de Grecon-voorspellingen (uitgezonderd de voorspellingen voor de particuliere werkgelegenheid) wat betreft de modelversies 78-A en later;
- de jaarlijkse aan het Grecon-model aangebrachte specificatiewijzigingen hebben geleid tot een betere voorspelkracht van het model als geheel. Enerzijds blijkt dit uit het verloop van de standaarddeviaties van de geschatte herleide-vormverstorings van bijna alle zes hier beschouwde variabelen; anderzijds kan dit worden afgeleid uit de ongelijkheidscoëfficiënt van Theil voor de voorspellingen ex post berekend met de diverse modelspecificaties;

7) Zie b.v. S. Makridakis, A. Andersen (e.a.), The accuracy of extrapolation (times series) methods: results of a forecasting competition, *Journal of Forecasting*, jg. 1, 1982, blz. 111-153.

8) Van Schaaijk komt tot andere conclusies. (M. van Schaaijk Raming en realisatie van CPB-prognoses, *ESB*, 23 juli 1986, blz. 727 – 729). Ten onrechte vergelijkt hij echter de Grecon-voorspellingen met de CPB-voorspellingen uit het CEP. Wij verwijzen naar de diverse publicaties van de Grecon-voorspellingen waarin onder meer wordt vermeld dat de voorspellingen voor jaar t in het najaar van jaar t-1 worden berekend op basis van gegevens in de MEV en van de jaarlijkse CBS-publicatie „Het jaar 19.. in cijfers”.

9) Elke modelversie van het Grecon-model wordt aangeduid met twee cijfers en een letter. De twee cijfers verwijzen naar het jaar waarvoor met het betreffende model voorspellingen ex ante zijn berekend. De letter is een interne code.

- 
- het gemiddelde van de CPB- en Grecon-voorspellingen ex ante blijkt in de hier beschouwde periode een redelijk alternatief voor beide, indien de ongelijkheidscoëfficiënt van Theil als maatstaf wordt genomen;
  - de zwakke schakel van het Grecon-model is de particuliere werkgelegenheid. Met betrekking tot de standaarddeviatie van de geschatte herleide-vormverstoringen valt op te merken dat deze voor alle modelversies relatief hoog is ten opzichte van het niveau van de variabele zelf, en bovendien geen goede maatstaf blijkt te zijn voor de betrouwbaarheid van de voorspellingen. Ook het verloop van de ongelijkheidscoëfficiënt voor de voorspellingen ex post duidt niet op een verbetering van de specificaties voor de recentere modelversies.

Als men met name aan de laatste conclusie een aanwijzing zou willen ontleenen omtrent de punten waarop men

---

de voorspelprocedure zou kunnen verbeteren, dan zal dit in de richting gaan van een verlaging van de standaarddeviatie van de geschatte herleide-vormverstoringen. Dit kan geschieden door verbetering van de specificatie van de betreffende gedragsvergelijking. Een andere mogelijkheid zou zijn (indien de reeks van geschatte verstoringen een systematisch gedrag lijkt te vertonen) te trachten tevens de waarde van de verstoring te voorspellen. Een simulatie met modelversie 86-B, waarin over de periode 1971 – 1984 aan de verstoringen een ARIMA-patroon wordt aangepast, had als resultaat dat de ongelijkheidscoëfficiënt van de particuliere werkgelegenheid wordt verlaagd van 0,82 tot 0,65 10).

**W. Voorhoeve**

---

10) Zie Voorhoeve, op.cit.