

# CPB-doorrekeningen, een kaartenhuis op drijfzand

## *Notitie over de econometrie*

Op 26 maart 1998 bracht het Centraal Planbureau (CPB) *Keuzes in kaart* uit. Het bevat de resultaten van de doorrekeningen die PvdA, CDA, VVD, D66 en GroenLinks (GrL) met het oog op de Kamerverkiezingen van dat jaar, op 6 mei, van hun verkiezingsprogramma's door het CPB hadden laten doen.

*Keuzes in kaart* leek me een goede casus: is de econometrie een stevig gefundeerd gebouw of een kaartenhuis op drijfzand? Om dat uit te zoeken ben ik een schriftelijke discussie begonnen met de directeur van het CPB, prof.dr. F.J.H. Don, die zo sportief was de handschoen op te pakken. In vier rondes zijn er brieven uitgewisseld: op 27 maart (ondergetekende), 24 april 1998 (Don), 29 april (ondergetekende), 15 juni (Don), 3 december 1998 (ondergetekende), 19 januari (Don), 2 februari (ondergetekende) en 26 april 1999 (Don).

Na deze discussie heb ik een aantal conclusies getrokken die zijn samengevat in de titel van deze notitie: de CPB-doorrekeningen zijn een kaartenhuis op drijfzand.

### 1. Uitkomsten van de doorrekeningen

Wat is er zoal uit de doorrekeningen gekomen? Nemen we vier voorbeelden uit de tabellen IV.2.1 en IV.3.1 van *Keuzes in kaart*.

**Voorbeeld 1.** Het eerste voorbeeld betreft de groei van de werkgelegenheid per jaar boven de 250.000 arbeidsplaatsen waar de vijf betrokken partijen en het CPB binnen een behoedzaam scenario (waarover verderop meer) van uitgaan. Als eerste komt hier GrL uit de bus met een groei van 0,6% boven die 250.000 arbeidsplaatsen per jaar (150.000 extra banen); op een gedeelde tweede plaats volgen PvdA, VVD en D66 met 0,4% (100.000 extra banen) en als laatste eindigt het CDA met 0,1% (25.000 extra banen).

**Voorbeeld 2.** De vijf partijen en het CPB gaan uit van een jaarlijks stijgend Bruto Binnenlands Product (BBP) van 2%. Volgens *Keuzes in kaart* komt hier bij uitvoer van het VVD-programma nog eens 0,3% bovenop; op een gedeelde tweede plaats komen PvdA, D66 en GrL met 0,2% en opnieuw eindigt het CDA als laatste met 0,0%.

**Voorbeeld 3.** Op 26 maart 1998 zat 5,75% van de beroepsbevolking zonder een betaalde baan. Als een verkiezingsprogramma regeringsbeleid zou worden, zou dat cijfer met een bepaald cijfer dalen. Als eerste eindigt GrL met een daling van 1,3% tot 4,45%; dan komen D66 met een daling van 0,9% (wordt 4,85%), VVD met een daling van 0,8% (wordt 4,95%), PvdA met een daling van 0,7% (wordt 5,05%) en CDA met een daling van 0,2% (wordt 5,55%).

**Voorbeeld 4.** Het CPB berekende ook een toename van de koopkracht. Voor GrL komt *Keuzes in kaart* op een toename van 1,75%; daarna volgt PvdA met 1,5% en op een gedeelde derde plaats staan CDA, VVD en D66 met 1,25%.

#### In tabel:

	PvdA	CDA	VVD	D66	GrL
Voorbeeld 1.					
Extra groei werkgelegenheid:	0,4 %	0,1 %	0,4 %	0,4 %	0,6 %
Voorbeeld 2.					
Stijging BBP:	0,2 %	0,0 %	0,3 %	0,2 %	0,2 %
Voorbeeld 3.					
Daling werkloosheidspercentage:	0,7 %	0,2 %	0,8 %	0,9 %	1,3 %

Werkloosheidspercentage:	5,05%	5,55%	4,95%	4,85%	4,45%	
Voorbeeld 4.						
Koopkrachttoename:	1,5 %	1,25%	1,25%	1,25%	1,75%	

Allemaal mooi en wel, maar wat zijn deze cijfers waard? Zijn het reële voorspellingen of niet? Hoe groot zijn hun onzekerheidsmarges? Hoe zit het met het democratische gehalte van beleid dat gebaseerd is op die voorspellingen, wanneer mocht blijken dat ze op schijnberekeningen berusten?

Deze en andere vragen kunnen we alleen beantwoorden door na te gaan hoe bovenstaande en andere uitkomsten uit CPB-doorrekeningen ontstaan zijn. We meten die doorrekeningen als het ware door, precies zoals een controlleur de kwaliteit van een lamp bepaalt door alle stappen van het productieproces van die lamp langs te lopen.

## 2. Hoe komen de uitkomsten in een doorrekening tot stand?

Een uitkomst als 'Wanneer het D66-verkiezingsprogramma regeringsbeleid zou worden, zou het BBP de komende vier jaar met 0,2% per jaar extra stijgen' komt in drie stappen tot stand: het afspreken van een basisscenario, begrotingskundige berekeningen en econometrische berekeningen.

In de eerste stap spreken het CPB en de betrokken partijen met elkaar af welk gemeenschappelijk behoedzaam basisscenario ze voor de economische ontwikkeling van de komende vier jaar nemen. Dat basisscenario bestaat uit aannames als een jaarlijkse stijging van het BBP van 2%, een relevant wereldhandelsvolume van 4%, een dollarkoers van f 2,05 en een olieprijs van 16 dollar per vat.

In de tweede stap voert het CPB op basis van dat overeengekomen basisscenario begrotingskundige berekeningen uit. Stel dat de overheid in 1998 in een bepaalde sector 10.000 mensen in dienst had, die de overheid elk per jaar f 30.000,- kosten. Stel voorts dat de PvdA er voor 1999-2002 per jaar 2.000 mensen bij wil, maar de VVD per jaar 1.000 minder. Bij een gezamenlijk overeengekomen aannames van een loonkostenstijging van 2% per jaar kan men dan uitrekenen hoeveel de PvdA in elk van de komende jaren meer moet begroten en met hoeveel de VVD het minder kan doen.

Concreet: de aangenomen loonkosten per functionaris bedragen f 30.000 in 1998, f 30.600,- in 1999, f 31.212,- in 2000, f 31.836,24 in 2001 en f 32.472,96 in 2002. Als het PvdA-programma regeringsbeleid zou worden, zou het aantal functionarissen 12.000 in 1999 bedragen, 14.000 in 2000, 16.000 in 2001 en 18.000 in 2002. De aantallen zijn 9.000, 8.000, 7.000 respectievelijk 6.000 voor het geval het VVD-programma in beleid zou worden omgezet. De te begroten bedragen per jaar zijn dan met het PvdA-programma als regeringsbeleid f 367.200.000,- (1999), f 436.968.000,- (2000), f 509.379.840,- (2001) en f 584.513.280,- (2002) en met het VVD-programma als regeringsbeleid f 275.400.000,- (1999), f 249.696.000,- (2000), f 222.853.680,- (2001) en f 194.837.760,- (2002).

In de derde stap voert het CPB eveneens op basis van het overeengekomen basisscenario econometrische berekeningen uit met behulp van een drietal econometrische modellen, namelijk het MIMIC-, het ATHENA- en het JADE-model. Formule 23 uit het JADE-model, de consumptievergelijking, bijvoorbeeld luidt:

$$(c - c_{-1})/c_{-1} = A\Delta^2 \ln w + B\Delta r_{t-1} + C\{\Delta(Y_L/p_c)\}/c_{-1} + D\{\Delta(Y_0/p_c)\}/c_{-1} \\ + E(\Delta Y_Z/p_c)/c_{-1} + F(\Delta A_F/p_c)/c_{-1} + G\{\ln(c_{-1}/c^*_{-1})\}.$$

In de CPB-publikatie over het JADE-model, *Werkdocument, nr.99* (1997), zijn niet alle symbolen in die vergelijking duidelijk gedefinieerd; hun definitie heb ik in een telefoongesprek met een CPB-medewerker op 28 april 1998 gekregen:  $c$  = reële consumptie in een bepaald jaar,  $c_{-1}$  = *idem* in het

jaar ervoor,  $w$  = loonvoet,  $r_{t-1}$  = lange rente in het jaar ervoor,  $Y_L$  = inkomsten uit loon,  $p_c$  = prijs van de reële consumptie,  $Y_0$  = inkomsten uit overdracht van vermogen,  $Y_Z$  = inkomsten uit winst,  $A_F$  = huizenvoorraad.

In de formule kunnen we  $(c - c_{-1})/c_{-1}$  kort aangeven met  $Z$ ,  $\Delta^2 \ln w$  met  $P$ ,  $\Delta r_{t-1}$  met  $Q$ ,  $\{\Delta(Y_L/p_c)\}/c_{-1}$  met  $R$ ,  $\{\Delta(Y_0/p_c)\}/c_{-1}$  met  $S$ ,  $(\Delta Y_Z/p_c)/c_{-1}$  met  $T$ ,  $(\Delta A_F/p_c)/c_{-1}$  met  $U$  en  $\{\ln(c_{-1}/c_{-1}^*)\}$  met  $V$ . Dan luidt de consumptievergelijking:

$$Z = AP + BQ + CR + DS + ET + FU + GV$$

### 3. Waardevolle aspecten aan en vragen bij de doorrekeningen

In het geschetste driestapsproces is de derde stap mijns inziens de zwakke schakel. Voor we daarop ingaan in §§4-12, staan we eerst stil bij de waardevolle aspecten aan de doorrekeningen.

Om te beginnen de eerste stap. Dat de partijen van eenzelfde behoedzame basisscenario uitgaan maakt de vergelijkbaarheid van hun verkiezingsprogramma's voor de kiezers tot iets reëls. Dat lijkt me winst. Immers, de aannames erin zijn zonder meer onzeker en eventuele verschillen ertussen zouden een vergelijking nodeloos bemoeilijken. Als de PvdA bijvoorbeeld een loonkostenstijging van 1% had aangenomen en de VVD een van 3%, dan zouden de effecten van de beleidsvoornemens, '2.000 mensen per jaar erbij' respectievelijk '1.000 mensen per jaar minder', vertroebeld worden. Het meerdere bij de PvdA van '2.000 mensen erbij' zou immers gedeeltelijk worden opgeheven door de kleinere loonkostenstijging van 1% en het mindere bij de VVD van '1.000 mensen minder' zou deels aan het oog worden onttrokken door een grotere loonkostenstijging van 3%.

Ook aan de tweede stap kan ik vanuit democratisch standpunt geen nadelen ontdekken. Begrotingskundige berekeningen dwingen de partijen immers tot keuzen in plaats van leuzen. In plaats van 'Wij willen meer blauw op straat' is het 'Voor bedrag X willen wij elk jaar 2000 extra agenten' en in plaats van 'De klassen moeten kleiner' wordt het 'Onze partij wil bedrag Y uittrekken voor meer onderwijskrachten in het basisonderwijs'.

Het CPB werkt aan de gesprekken tussen de betrokken politieke partijen ten behoeve van de eerste stap mee en voert de berekeningen in de tweede stap voor die partijen uit. Voor beide is het CPB echter strikt genomen niet nodig. Zo'n basisscenario kunnen partijen ook onderling afspreken of onder leiding van een andere gespreksleider. En begrotingskundige berekeningen kan iedereen uitvoeren, die over een goede rekennaardigheid beschikt. Dat het CPB desondanks in die twee stappen meewerkt, is daarom iets waarvoor het alle lof verdient.

De eigenlijke betekenis van de betrokkenheid van het CPB bij de doorrekeningen zou tot uiting moeten komen in de derde stap. Niet dat het CPB zich uitsluitend met econometrie bezighoudt, maar het opstellen van econometrische modellen en het uitvoeren van econometrische berekeningen ten behoeve van de overheid behoren wel tot haar specifieke veld. Daarin zou dus de kracht van de doorrekeningen moeten liggen. Hoe staat het daarom met de econometrie?

We kunnen die vraag naar twee soorten vragen uitsplitsen. De eerste soort is: in hoeverre verwijzen econometrische modellen van het CPB naar de economische werkelijkheid? Immers, zoals we in §1 hebben gezien, het bouwen van zulke modellen en het rekenen ermee zou tot feitelijke kennis over de werkgelegenheidsgroei, de stijging van het BBP, de daling van het werkloosheidspercentage, de koopkrachttoename en dergelijke leiden, en dat zijn toch bepaald geen metafysische grootheden, maar aspecten van de economische werkelijkheid. Hierover gaan §§4-6.

De tweede soort vragen ten aanzien van de econometrie heeft betrekking op de onzekerheidsmarges in de uitkomsten. Hoe groot zijn ze en publiceert het CPB ze altijd, bijvoorbeeld in *Keuzes in kaart*? Zie verder §§7-10.

### 4. De doorrekeningen en de economische werkelijkheid

De vraag in hoeverre econometrische modellen van het CPB naar de economische werkelijkheid

verwijzen, kunnen we in het kader van de doorrekeningen toespitsen op de vraag 'Naar welke economische werkelijkheid verwijst het JADE-model?' en die weer op 'Naar welke economische werkelijkheid verwijst de consumptievergelijking?'

Van de consumptievergelijking  $Z = AP + BQ + CR + DS + ET + FU + GV$  dient dus aangetoond te zijn dat ze feitelijk houdbaar is alvorens ze wordt toegepast, bijvoorbeeld in de doorrekeningen. Via allerlei berekeningen waar niet alleen de formule voor deze vergelijking in voorkomt, maar ook allerlei andere formules uit het JADE-, MIMIC- en ATHENA-model, komen immers uiteindelijk de getallen als die van de tabel in §1 tot stand.

De vraag naar de feitelijke houdbaarheid is een gebruikelijke in de wetenschap. In de natuurkunde en de geneeskunde bijvoorbeeld wordt iets pas toegepast als er voldoende onderzoek is geweest dat die geldigheid en betrouwbaarheid van het toegepaste aantoont.

In beginsel moet de toetsing van de consumptievergelijking niet al te ingewikkeld zijn. Ze bestaat immers uit zeven termen die bij elkaar worden opgeteld. In feite houdt de formule in: als de som van de laatste zes termen constant is, dan geldt  $Z = AP$  (voor zover  $P$  onafhankelijk is van  $Q, R, S, T, U$  en  $V$ ; en ook dat zou aangetoond dienen te zijn). De gehele formule zelf is opgesteld aan de hand van gegevens uit de periode 1958-1995. Wellicht is iemand nagegaan of er in die 38 jaren bijvoorbeeld 8 jaren zijn waarin de som van de laatste zes termen inderdaad een (min of meer) gelijke waarde heeft.

Feitelijk komt  $Z = AP$  er dan op neer dat het CPB aanneemt dat de veranderingen van  $Z$  deels door die van  $P$  worden verklaard, met  $A$  als constante: als  $P$  stijgt, stijgt  $Z$  ook; en als  $P$  daalt, daalt  $Z$  ook. Dat dat geen raar vermoeden is, zal iedereen onderschrijven die zich realiseert dat  $P$  staat voor  $\Delta^2 \ln w$ , met  $\ln w$  de logaritme uit de loonvoet, terwijl  $\Delta^2$  verband houdt met de inflatie. Bij een negatieve waarde van  $A$  betekent  $Z = AP$  dus dat wordt aangenomen dat de consumptie daalt als de inflatie uit de logaritme van de loonvoet stijgt en dat de eerste stijgt als de tweede daalt.

Of en in hoeverre  $Z = AP$  apart is onderzocht is me onduidelijk gebleven zoals we in §§5-6 nog zullen zien. Het blijft dus ook onduidelijk of  $P$  echt wel  $\Delta^2 \ln w$  moet bevatten. Weet het CPB echt zo zeker dat het niet  $\Delta$  of  $\Delta^3$  en dat het beslist  $\ln w$  moet zijn en niet  $(w - 1)/w$  of  $1/(w^2 + 1)$ ? Wat de gegevens over de consumptievergelijking echter wel laten zien is dat coëfficiënt  $A$  allerm minst constant is. Zijn waarde is  $-0,253 \pm 0,08$  (in het algemeen betekent  $10 \pm 2$  dat een waarde tussen  $10 - 2$  en  $10 + 2$  ligt, dus tussen 8 en 12), wat neerkomt op een relatieve onzekerheidsmarge van 31,6%.

Wat ik hierboven over  $Z = AP$  schreef, is slechts een mogelijke toetsing van een term uit de consumptievergelijking. Waarschijnlijk zijn er andere toetsingsmogelijkheden en dus was en ben ik nog steeds benieuwd naar onderzoek waaruit blijkt dat de consumptievergelijking inderdaad beschrijft hoe het er in (een deel van) de economische werkelijkheid toe gaat. En ik concludeer nu wel uit de informatie die mij ter beschikking staat dat  $A$  geen constante is, maar wellicht is er sedert 1997 onderzoek bekend geworden waarin de constantheid van  $A$  wel is aangetoond - *Keuzes in kaart* stamt immers uit 1998.

Wat voor  $Z = AP$  is geopperd, geldt *mutatis mutandis* ook voor  $Z = BQ, Z = CR, Z = DS, Z = ET, Z = FU$  en  $Z = GV$ . Uit welk onderzoek blijkt dat het aannemen van die lineaire afhankelijkheden feitelijk verantwoord is? En hoe zit het met de aangenomen constantheiden van  $B, C, D, E, F$  en  $G$ ? Hun waarden zijn:  $B = -0,458 \pm 0,35, C = 0,881 \pm 0,13, D = 0,871 \pm 0,20, E = 0,225 \pm 0,11, F = 0,071 \pm 0,03$  en  $G = -0,462 \pm 0,16$ . Oftewel, hun relatieve onzekerheidsmarges zijn: 76,4%, 14,8%, 23,0%, 48,9%, 42,3% respectievelijk 34,6%. In hoeverre is het gerechtvaardigd grootheden met zo'n grote relatieve onzekerheidsmarges nog als constanten aan te merken en niet als grootheden die (fors) in de tijd fluctueren?

## 5. Enig heen en weer

Ik heb Don herhaaldelijk gevraagd naar bewijzen voor de economische werkelijkheid achter de consumptievergelijking en naar het al dan niet constant zijn van de coëfficiënten A-G.

Beginnen we met dat laatste. In zijn eerste brief, van 24 april 1998, blijkt Don A-G inderdaad als constanten op te vatten: 'De gepubliceerde [onzekerheidsmarges] bij de coëfficiënten dienen te worden opgevat als een indicatie van de *meetnauwkeurigheid* van constanten' (punt 5).

Ik begrijp wat Don bedoelt, maar de vraag is natuurlijk of men iets als constant mag blijven aanmerken als de aanname van constantheid niet nauwkeuriger bepaald kan worden dan met een onzekerheidsmarge tussen 14,8% en 76,4%. Ik heb dat diverse keren naar voren gebracht, maar zonder succes. In mijn laatste brief, van 2 februari 1999, bijvoorbeeld schrijf ik: 'Ik herhaal in andere bewoordingen de vraag die ik al eerder in mijn vragen 5, 14 en 23 heb gesteld en die u onaangeroerd en dus onbeantwoord heeft gelaten: Hoe kunt u de constantheidsaanname van de coëfficiënten [...] garanderen bij zulke grote onzekerheidsmarges?' (punt 28).

Ook in zijn laatste brief, van 26 april 1999, beantwoordt Don die voor de hand liggende en cruciale vraag niet. Ik blijf daarom bij mijn aanname dat de coëfficiënten A-G geen constanten zijn maar grootheden die fors in de tijd fluctueren (mijn brief van 2 februari 1999, punt 28).

Dan bewijzen voor de economische werkelijkheid achter de consumptievergelijking. Al in mijn eerste brief, van 27 maart 1998, vraag ik Don daarnaar: 'Kunt u [...] de economisch-feitelijke betekenis van elk van de termen [in de consumptievergelijking] aangeven?' (punt 3).

Don antwoordt in zijn brief van 24 april 1998: 'Voor een bespreking van de feitelijke economische betekenis van de consumptievergelijking verwijzen wij naar pagina's 33-37 van [CPB Werkdocument nr.99]' (punt 3).

Aangezien er in dat Werkdocument niets over het gevraagde te vinden is, ook niet op p.33-37 vraag ik in mijn brief van 29 april 1998: 'Over het JADE-model en de erin voorkomende formules vroeg ik onder meer: "Uit welk onderzoek blijkt dat deze econometrische formules een economische werkelijkheid dekken en niet speculatief zijn?" (vraag 3). Kunt u die vraag alsnog beantwoorden, bijvoorbeeld voor de consumptievergelijking uit het JADE-model? (U schreef dat de "feitelijke economische betekenis" van die vergelijking op p.33-37 van de publikatie over het JADE-model zou staan. Dat is slechts zeer gedeeltelijk het geval. Eén van uw medewerkers was zo vriendelijk me gisteren de ontbrekende uitleg te verschaffen.)' (punt 12).

In zijn brief van 15 juni 1998 luidt Dons antwoord: 'De consumptievergelijking in JADE borduurt voort op internationaal empirisch en theoretisch onderzoek van de macro-economische consumptievergelijking dat al meer dan een halve eeuw plaats vindt. Zie onder andere de verwijzingen naar de literatuur in het werkdocument over het JADE model' (punt 12).

Mijn reactie daarop in mijn brief van 3 december 1998 is: 'In uw antwoord 12 schrijft u dat de consumptie-vergelijking voortborduurde op allerlei onderzoek. Of er "onderzoek" in verband met die vergelijking wordt gedaan, was echter niet mijn vraag. (Ik betwijfel overigens in hoge mate dat er in en rond bovenstaande consumptie-vergelijking "al meer dan een halve eeuw" onderzoek wordt gedaan, zoals u stelt; mijns inziens is die formule veel jonger en is er - wat wetenschapshistorisch wel een noodzakelijke voorwaarde is - geen complete inhoudelijke continuïteit met voorgangers die eventueel diezelfde naam dragen.) Ik vroeg of u onderzoek kon aanhalen waaruit blijkt dat die vergelijking naar een economische werkelijkheid verwijst. Uit uw antwoord 13 maak ik op dat het JADE-model slechts 30 econometrische vergelijkingen zou bevatten. Het aantal doet er echter niet toe: als een uitkomst is verkregen op basis van vele [begrotingskundige] berekeningen en slechts één econometrische, dan is die uitkomst waardeloos geworden als mijn kritiek op de econometrie hout snijdt. Ik herhaal daarom mijn vraag: *naar welke economische werkelijkheid verwijst de consumptie-vergelijking* en uit welk onderzoek blijkt dat dat zo is?' (punt 20).

## **6. Op welk macro-economisch onderzoek wordt gedoeld? Een eerste conclusie**

Zonder mijn vragen te beantwoorden of mijn bezwaren te trachten te weerleggen blijft Don naar *Werkdocument nr.99* verwijzen, waar de beloofde verwijzingen over een periode van 'meer dan een halve eeuw' zouden staan maar feitelijk niet te vinden zijn. Don in zijn brief van 19 januari 1999: 'Ik herhaal onze verwijzing naar de macro-economie, in het bijzonder de consumptie-theorie. De voor het JADE-model belangrijkste verwijzingen naar de literatuur staan in Werkdocument 99' (punt 20).

Bedoelt Don dat er al meer dan een halve eeuw sprake is van een of meerdere theorieën met de naam 'consumptietheorie'? Misschien, maar dat was niet mijn vraag: ik wilde en wil nog steeds weten uit welk onderzoek blijkt dat de consumptievergelijking in de gedaante van §2 naar de economische werkelijkheid verwijst. In mijn brief van 2 februari 1999 heb ik daarom zelf maar een antwoord geformuleerd bij wijze van conclusie die ik aanhoud totdat Don of iemand anders van het CPB haar daadwerkelijk weerlegt:

**De modellen van het CPB verwijzen niet naar de economische werkelijkheid. De zogenaamde consumptievergelijking bijvoorbeeld beschrijft geen daadwerkelijk consumptiegedrag.**

Als toelichting hierop voeg ik het volgende toe. Wat Don 'verwijzing naar de macro-economie' noemt (zijn punt 20), verwijst in werkelijkheid *niet* naar iets empirisch, maar naar een rekenkundig opgetuigd model dat onaangetoond is gebleven. Als hij werkelijk meent dat de verwijzingen in *Werkdocument nr.99* literatuur bevatten die laat zien naar welk reël (in plaats van verbaal benoemd) consumptiegedrag de consumptievergelijking verwijst, dan zou hij niet ongespecificeerd een hele literatuurlijst dienen te noemen, maar concreet dienen aan te geven om welke pagina's en/of hoofdstukken in welke publicaties het gaat en wat daarin aan feitelijke economische inzichten omtrent consumptiegedrag te vinden is. In de onmiddellijke nabijheid van de consumptievergelijking staan in elk geval geen opmerkingen die zelfs maar in de richting van een empirische verwijzing gaan.

Dons reactie op die eerste conclusie luidt onder meer (zijn brief van 26 april 1999): 'Ik volsta met de constatering dat het CPB de "state of the art" van het relevante vakgebied toepast. De economische inhoud staat in de analyses voorop; econometrische technieken zijn behulpzaam bij het kwantificeren van het relatieve belang van de verschillende economische samenhangen'. En zo is de cirkel rond, want in de gehele briefwisseling stonden nu juist de 'state of the econometrical art' en econometrische kwantificeringen ter discussie.

## 7. Onzekerheidsmarges in de doorrekeningen

Laten we nu toch eens veronderstellen dat de consumptievergelijking *wel* naar een economische werkelijkheid verwijst. Nemen we ook even aan dat de uitkomsten van de doorrekeningen van 1998 alleen op die ene formule zijn gebaseerd zodat ze feitelijke kennis bevatten in plaats van schijnkennis. (In werkelijkheid zijn die uitkomsten op vele formules gebaseerd waar eenzelfde verhaal voor gehouden zou kunnen worden als dat in §§4-6 voor de consumptievergelijking. Het JADE-model alleen al telt ongeveer 30 econometrische formules.) Dan luidt de vervolgvraag: hoe nauwkeurig zijn die uitkomsten? Het maakt immers nogal wat uit of een onzekerheidsmarge 10% bedraagt of 150%.

Stel dat het ene partijprogramma volgens de doorrekeningen in 100.000 extra banen resulteert en het andere in 50.000 extra banen. Bij een onzekerheidsmarge van 10% spreken die getallen klare taal. Immers, 10% van 100.000 is 10.000, zodat het voorspelde aantal van 100.000 extra banen naar alle waarschijnlijkheid feitelijk ergens tussen 90.000 en 110.000 ligt:  $100.000 \pm 10.000$ . Voor die andere partij bedraagt die 10% 5.000 extra banen en komt het aantal op  $50.000 \pm 5.000$ . Het informatieve van de getallen 100.000 en 50.000 zit hem hierin dat de twee intervallen elkaar niet overlappen: de ondergrens van het hoogste interval, 90.000 extra banen, is groter dan de bovengrens van het laagste interval, 55.000 extra banen.

Bij 150% ligt dat anders. In het ene geval gaat het dan om een onzekerheid van 150.000 extra banen en in het andere om een onzekerheid van 75.000. En deze grote onzekerheden hebben tenminste twee gevolgen. In de eerste plaats, de intervallen  $100.000 \pm 150.000$  en  $50.000 \pm 75.000$  overlappen elkaar: het eerste interval bevat ook de voorspelling 50.000 voor de tweede partij en het tweede interval ook de voorspelling 100.000 voor de eerste partij. De voorspellingen zijn dus in hoge mate inwisselbaar en dus nietszeggend geworden. Maar in de tweede plaats, terwijl beide voorspellingen extra banen lijken aan te geven, komen ze door hun onzekerheidsmarge van meer

dan 100% ook in de negatieve getallen terecht, wat juist op minder banen wijst. De ondergrens van het eerste interval is immers -50.000 en die van het tweede interval -25.000. In theorie is het dus denkbaar dat de partij met een extra banengroei van 100.000 in werkelijkheid uitkomt op een *verlies* van 50.000 banen, terwijl de andere partij met zijn voorspelde 50.000 extra banen op een *groei* van 75.000 uitkomt.

'Ja maar', zo zou men kunnen tegenwerpen, 'dat is niet eerlijk. Als je aanneemt dat het voor de ene partij +150% is, dan moet je voor de andere ook +150% aannemen'. Dat zou inderdaad waar zijn als we maar met één beleidsvoornemen te maken hadden of als we de effecten van de overige beleidsvoornemens precies zouden kennen. Geen van beide is het geval en dus kan niemand, ook niemand van het CPB, ons garanderen dat de feitelijke uitkomst voor de ene partij wel eens op -150% ten opzichte van de voorspelling zou kunnen uitkomen en bij de andere op +150%.

## 8. Publiceert het CPB zijn onzekerheidsmarges altijd? Een tweede conclusie

Hoe zit het dus met de onzekerheidsmarges in de doorrekeningen? Met die voor de coëfficiënten van de consumptievergelijking hebben we al kennisgemaakt. Ze bedragen in oplopende grootte: 14,8% (voor C), 23,0% (voor D), 31,6% (voor A), 34,6% (voor G), 42,3% (voor F), 48,9% (voor E) en 76,4% (voor B).

In zijn brief van 26 april 1999 geeft Don van P tot en met V ook de gemiddeldes en getallen die in de econometrische optiek slechts hun spreidingen zijn, maar in de algemene meetfoutentheorie (waarover meer in §9) hun onzekerheidsmarges:  $P = -0,0001376 \pm 0,0312$ ,  $Q = 0,0001243 \pm 0,00576$ ,  $R = 0,0222 \pm 0,0253$ ,  $S = 0,0167 \pm 0,0107$ ,  $T = 0,0145 \pm 0,0910$ ,  $U = -0,00326 \pm 0,0209$  en  $V = 0,000233 \pm 0,0146$ . Oftewel, in oplopende grootte: 64% (voor S), 114% (voor R), 628% (voor T), 641% (voor U), 4.634% (voor Q), 6.266% (voor V) en 22.674% (voor P). Voor Z zijn deze cijfers  $0,0348 \pm 0,0275$  (79%).

Publiceert het CPB deze (en andere) spreidingen en/of onzekerheidsmarges? Een algemeen aanvaarde wetenschappelijke code is dat spreidingen en/of onzekerheidsmarges bij getalsmatig onderzoek vermeld worden. Ook het CPB onderschrijft die code. Zo stelt Don in *Forecast uncertainty in economics* (1994, p.568): 'Het hoofddoel van de voorspellende activiteiten van het CPB is een vast punt te verschaffen voor de voorbereiding van de economische politiek. De onzekerheid die gepaard gaat met de voorspelling moet uiteraard ook aan de beleidsmakers worden meegedeeld'.

Men zou dus verwachten dat het CPB keurig netjes in *Keuzes in kaart* allerlei onzekerheidsmarges meedeelt: in tabellen, in een aanhangsel of hoe ook maar. Dat is echter *niet* het geval. Op p.7 wijst het CPB slechts op 'de nodige onzekerheid' en op 'de grote onzekerheid die economische *prognoses* kenmerkt' en stelt het relativerend dat 'de verschillen in beleidseffecten tussen de partijprogramma's vaak kleiner zijn dan de onzekerheidsmarges in de ramingen', maar voor het overige blijft het gissen.

Onzekerheidsmarges van coëfficiënten tussen 14,8% en 76,4% en spreidingen en/of onzekerheidsmarges van variabelen tussen 64% en 22.674% (inderdaad: meer dan tweeëntwintigduizend!): dat lijken me nog eens gegevens die niet alleen beleidsmakers graag zouden willen kennen, maar ook hun uiteindelijke opdrachtgevers (de kiezers) en hun uiteindelijke financiers (de belastingbetalers). En uiteraard werken spreidingen en onzekerheidsmarges door in de onzekerheidsmarges van de uitkomsten als die in §1. Daarover meer in §10.

Groot of klein: onzekerheidsmarges in voorspellingen van economische grootheden dienen vermeld te worden en dat is in *Keuzes in kaart* niet gebeurd. Op die van A-G na zijn ze ook niet vermeld in *Werkdocument nr.99*. Daarom heb ik er Don om verzocht in mijn eerste brief, van 27 maart 1998 (punt 7). In zijn brief van 24 april 1998 antwoordt hij: '[Een onderzoek naar de onzekerheid van de uitkomsten] is erg bewerkelijk en kostbaar, en zal voor een deel ook op ongetoetste veronderstellingen moeten berusten. Een wetenschappelijk verantwoorde kwantificering van de onvoorwaardelijke onzekerheidsmarges rond de uitkomsten is daardoor niet mogelijk' (punt 7).

Voor sommige onzekerheidsmarges bestaan dus zelfs geen econometrisch verantwoorde schattingen! Dan ligt het voor de hand dat het CPB ze niet kan publiceren, maar dan zou het die onmogelijkheid in elk geval dienen te publiceren of de voorspellingen zelf niet publiceren.

Al met al formuleer ik in mijn brief van 2 februari 1999 als mijn tweede conclusie:

**Het CPB publiceert onzekerheidsmarges niet (steeds), wat in strijd is met de wetenschappelijke code die het CPB verbaal wel onderschrijft.**

In zijn reactie hierop van 26 april 1999 noemt Don slechts paragraaf IV.1 uit *Centraal Economisch Plan 1999*. Die tekst bevat echter geen wetenschappelijke verantwoording voor het niet vermelden van onzekerheidsmarges in publicaties in het algemeen en in doorrekeningen in het bijzonder. Blijft dus overeind dat het CPB slechts lippendienst bewijst aan de noodzaak tot meedelen van onzekerheidsmarges bij voorspellingen.

**9. Wijkt het CPB af van de algemene meetfoutentheorie?**

In mijn tweede brief, van 29 april 1998, schrijf ik het vermoeden dat de algemene theorie over onzekerheidsmarges, de zogeheten meetfoutentheorie, afwijkt van de meetfoutentheorie in de econometrie (punt 14). Omdat Dons reacties weinig duidelijkheid boden, heb ik de algemene meetfoutentheorie vanuit mijn natuurkundige achtergrond uiteengezet in mijn brief van 2 februari 1999 (punt 27). Op die manier hoopte ik meer helderheid te krijgen over overeenkomsten en verschillen tussen beide meetfoutentheorieën en over de vraag of de econometrische variant wel geoorloofd is.

Omdat geen enkele waarneming perfect is, is in de natuurkunde de meetfoutentheorie ontwikkeld. Deze is bedoeld om onzekerheidsmarges te schatten en dus om de relatieve waarde van een meting aan te geven. Stel, theorie I voorspelt voor een bepaalde grootte het getal 8 en theorie II het getal 4 (beide in dezelfde eenheid). Stel verder dat de meting 7,5 oplevert, dan lijkt theorie I bevestigd te zijn, daar 7,5 dicht bij 8 ligt dan bij 4. Als de onzekerheidsmarge van de meting echter 4,5 bedraagt, ligt de uitkomst tussen 3 en 12 en is theorie II nog steeds te handhaven, want de voorspelde waarde daarvan, 4, ligt ook tussen 3 en 12. De meetfoutentheorie is er dus om de waarde van gedane beweringen aan te geven; we zagen iets dergelijks ook al in §7.

Al in zijn brief van 19 januari 1999 had Don mijn vermoeden bevestigd dat de econometrie een andere meetfoutentheorie hanteert dan wetenschappen als de natuurkunde: 'Een belangrijk verschil tussen econometrisch onderzoek en natuurkundig onderzoek is dat de econometrie vrijwel altijd moet werken met non-experimentele gegevens, omdat gecontroleerde experimenten meestal onmogelijk zijn. Over de consequenties hiervan voor de econometrische modelbouw, en in het bijzonder voor de toepasbaarheid van statistiek en meetfoutentheorie, zie het boek van E.E. Leamer "Specification searches: ad hoc inference with nonexperimental data" (punt 23).

Om te beginnen, Dons voorstelling van zaken als zou de meetfoutentheorie iets van doen hebben met experimentele versus niet-experimentele waarnemingen is onjuist. Metingen zijn metingen en omdat die altijd onvolmaakt zijn, zijn onzekerheidsmarges altijd vereist. Sterker nog, de meetfoutentheorie is ontwikkeld in de sterrenkunde, niet bepaald een experimentele wetenschap.

Vervolgens, sedert de Duitse wiskundige Gauss (1777-1855) neemt men in de algemene meetfoutentheorie onder meer aan dat er *veel* onzekerheden in een meting een rol spelen en dat ze allemaal *klein* zijn ten opzichte van het gemiddelde. Gelden die twee aannames ook in de econometrische meetfoutentheorie? De eerste klopt waarschijnlijk wel al was het alleen maar vanwege de vele subjectieve en sociale factoren die in de economie een rol spelen; zie verder §11. De tweede aanname daarentegen lijkt me volstrekt ongeoorloofd te zijn, te oordelen aan de groottes van de onzekerheidsmarges in A-G en aan de groottes van de spreidingen en/of onzekerheidsmarges in P-V.

Ten slotte is het wel voorstelbaar dat econometristen behoefte hadden aan een andere meetfoutentheorie dan de algemene meetfoutentheorie. Achter berekeningen van Gauss en anderen



met waarnemingen aan planeten, planetoïden en andere hemellichamen zaten allerlei getoetste natuurkundige wetten en theorieën, zoals Keplers wetten, Newtons mechanicawetten en Newtons zwaartekrachttheorie. In §§4-6 daarentegen hebben we gezien dat er achter econometrische berekeningen nu juist *geen* getoetste economische theorie zit. En dus doet zich in de econometrie de situatie simpelweg niet voor, dat een relatieve waarde van een meting gewogen moet worden of dat er tussen twee economische theorieën I en II beslist moet worden of dat de relatieve waarde van een economische bewering aangegeven moet worden. Vandaar ook dat  $\pm 0,0312$  (bij P) een onzekerheidsmarge is in de algemene meetfoutentheorie, maar kennelijk slechts een niet op enige onzekerheid duidende spreidingsmaat in de econometrie. Tussen Gauss en Leamer moet er iets ontspoord zijn, maar wat en waar?

De vraag rijst of wat econometristen meetfoutentheorie en onzekerheidsmarges noemen, die naam wel verdienen. Waarschijnlijk zijn die benamingen onterecht. Leamers boek *Specification searches*, dat Don noemt, geeft alle aanleiding dat te veronderstellen. Het blijkt dat de econometrische meetfoutentheorie er niet is om de grenzen van de kennis van de onderzoeker aan te geven, maar - zonder kwade opzet, naar ik aanneem - feitelijk het effect heeft dat uitkomsten van econometrische berekeningen geflatteerd worden. Zie §10.

## 10. Kunnen 'econometrische zonden' soms 'wenselijk' zijn? Een derde conclusie

In zijn voorwoord schetst Leamer een alarmerend beeld, waarvan ik vermoed dat het in 1999 nog steeds van kracht is. Hij constateert een brede kloof tussen econometrische theorie en econometrische praktijk: 'We verdelen ons comfortabel in een celibatair priesterschap van statistische theoretici aan de ene en een legioen van onverbeterlijke zondaar-data analisten ("legion of inveterate sinner-data analysts") aan de andere kant'. Weliswaar vindt Leamer dat de 'zondaars' hun zonden openlijk zouden moeten bekennen, maar hij gaat er ondertussen vanuit dat 'onvermijdbare zonden eigenlijk geen zonden kunnen zijn'...

Hoe ver gaat Leamer met die 'econometrische zonden': ze vergeven of ze ook nog eens legitimeren? Dat laatste! 'De voornaamste uitkomst [van dit boek] is een waardering van waarom bepaalde zonden onvermijdbaar, zelfs wenselijk zijn'.

Wenselijke, wetenschappelijke zonden!?!? Ik meen dat Leamers methode gemakkelijk kan leiden tot een vervaging van wetenschappelijke normen. Ik beschuldig het CPB er niet van te fozelen met feiten of te morrelen aan metingen, maar wel vermoed ik dat de meetfoutentheorie die het CPB hanteert, op zijn zachtst gezegd gepubliceerde uitkomsten te kleine onzekerheidsmarges meegeeft.

In dat vermoeden word ik gesterkt als Leamer schrijft dat hij zijn algemene methode heeft uitgewerkt naar zes concrete methodes. Eén daarvan noemt hij 'postdata-model-constructie'. Over de onzekerheidsmarges van de andere vijf schrijft hij: 'Met uitzondering van de postdata-model-constructie produceren de andere doorzoekingen een vergelijking die ertoe neigt de onzekerheid te onderschatten' (p.12).

Van dat onderschatten heb ik een aardige indruk kunnen krijgen door de algemene meetfoutentheorie toe te passen op de consumptievergelijking. (Om de reden die in §9 is vermeld, is dit strikt genomen niet terecht.) Daar krijg ik een onzekerheidsmarge van 0,0357 uit. Voor de gemiddelde waarde van Z voor de periode 1958-1995, 0,0348, is dat dus een relatieve onzekerheidsmarge van 103%. In zijn brief van 26 april 1999 schrijft Don daarentegen: 'De parameteronzekerheid [...] leidt slechts tot een standaardfout van 14% [in Z]'. In een begeleidend memo bij die brief en uit een telefoongesprek van 6 mei 1999 blijkt overigens dat die 14% ook niet de totale onzekerheidsmarge van Z is volgens de meetfoutentheorie van het CPB: die bedraagt uiteindelijk 48%, toch nog fors minder dan de 103% volgens de algemene meetfoutentheorie en ook minder dan de 79% in de spreiding van Z in de periode 1958-1995 (dat een voorspelling nauwkeuriger zouden kunnen zijn dan de metingen waarop ze is gebaseerd is me een raadsel).

En dan hebben we het nog maar over uitkomsten van één econometrische formule. Bij een uitkomst waar twee of meer van zulke formules een rol in speelt, zijn de onzekerheden alleen maar nog groter. Dat is echter niet na te gaan: het CPB publiceert ook niet welke formules voor elk der

voorspelde grootheden gebruikt zijn.

Over Leamers methode heb ik in mijn brief van 2 februari 1999 diverse vragen aan Don gesteld, onder meer of Leamers vijf methodes die onzekerheidsmarges onderschatten, ook in de doorrekeningen zijn gebruikt (punt 31). In zijn brief van 26 april 1999 antwoordt hij slechts: 'Gelet op de eerdere discussie en het kennelijke misverstand inzake de becijfering van onzekerheidsmarges, zie ik af van een nadere reactie op uw punten 27 t/m 31' (de punten 27-31 gingen allemaal over het onderwerp meetfoutentheorie).

Wellicht vindt Don oprecht dat hij eerder op het onderwerp meetfoutentheorie is ingegaan en meent hij heus dat er geen twee verschillende meetfoutentheorieën in het spel zijn maar slechts een misverstand bij mij, maar ondertussen luidt mijn derde conclusie:

**Zelfs als de modellen van het CPB werkelijkheidswaarde zouden hebben, zijn de onzekerheidsmarges in de uitkomsten van de berekeningen van het CPB te groot om er reële richtlijnen aan te kunnen ontleen, als men die marges met de algemene meetfoutentheorie uitrekt. Het CPB hanteert echter een meetfoutentheorie die daarvan afwijkt zonder dat duidelijk is waarom dat geoorloofd is, maar die onzekerheidsmarges wel onderschat.**

## 11. Het falen van de econometrie

Het CPB gaat in zijn econometrische berekeningen niet uit van een economische theorie (§§4-6). Het is dan ook geen wonder dat het CPB er doorgaans (fors) naast zit bij zijn voorspellingen.

Een kras voorbeeld van dat falen is het voorspellen van de zogeheten inkomensplaatjes. Drie economen van het Ministerie van Economische Zaken zijn voor de periode 1971-1995 eens nagegaan in hoeverre die zijn uitgekomen. Het bleek dat het CPB er tussen 0,7 en 1,3% naast zat (*Openbare uitgaven*, juli 1997). Dat is om twee redenen te hoog. a. De voorspelde waarden zelf zijn enige procenten zodat de relatieve fouten groot zijn; zo heeft een voorspelling van 2,5% een relatieve fout van 40% als het in werkelijkheid 3,5% is. b. Vaak is een verschil van 0,1% voor de overheid aanleiding om bepaalde voornemens al dan niet uit te voeren.

Een ander voorbeeld is hoofdstuk VI uit *Economische verkenning voor de volgende kabinetsperiode* (1997). Het schetst onder meer dat het aantal banen met 2% per jaar is gegroeid (tegen een CPB-voorspelling van 1,3%); dat 'de groei van de bestedingen en de productie [hoger is] uitgekomen' dan de CPB-voorspelling; dat de groei van de werkgelegenheid 400.000 arbeidsjaren bedroeg ('meer dan het dubbele' van de CPB-raming van 190.000 arbeidsjaren); dat de werkgelegenheid in personen met 465.000 is gestegen (tegen een CPB-voorspelling van 275.000); en dat 'de werkloosheid 170.000 personen lager [uitkomt]' dan door het CPB was voorspeld. Kortom, naast enkele goede voorspellingen, zoals voor lonen en prijzen, zat het CPB er met de meeste van zijn voorspellingen fors naast.

Deze en talrijke andere voorbeelden brengen ook econometristen ertoe hun vak terug te willen trekken uit de macro-economie. Prof.dr. Ph.H.B.F. Franses bijvoorbeeld stelde in zijn oratie van 24 september 1998 onder meer: 'De hooggespannen verwachtingen, die men destijds van de econometrie als vakgebied had, zijn niet waargemaakt' (p.11) en 'Het is dan ook mijn mening dat het vakgebied econometrie [...] zich beter kan richten op andere terreinen dan de macro-economie' (p.23).

Waarom faalt de econometrie? Daar zitten mijns inziens twee redenen achter.

De eerste reden is dat subjectieve en sociale factoren een grote maar onvoorspelbare rol spelen in de economie: gevoelens om iets al dan niet te willen kopen, mate van vertrouwen in de marktsector, enzovoort respectievelijk wetten, regelingen, internationale spanningen, invoering van de euro, enzovoort. Allemaal zaken die niet te behandelen zijn met wiskundige vergelijkingen, laat staan dat ze daarmee te voorspellen zouden zijn.

Maar dat laatste probeert het CPB nu juist wel, en daarin zit de tweede reden. Volgens de grondleggers van de econometrie, de Nederlander Tinbergen en de Noor Frisch, staat hun vak voor wat betreft de manier van modellenbouwen en rekenen op voet van gelijkheid met de natuurkunde.

Precies zoals natuurkundige modellen en berekeningen tot allerlei vormen van feitelijke natuurkundige kennis en toepassingen hebben geleid (variërend van thermometers naar wereldomvattende direct uitgezonden kleurentelevisie-uitzendingen), zo zouden econometrische modellen en berekeningen tot feitelijke economische kennis en toepassingen leiden. Het draait hierin dus niet om natuurkundige, economische of wat voor kennis ook, maar om kennisverwerving. Het is echter simpelweg niet waar dat die natuurkundige kennis tot stand gekomen is op de wijze waarlangs econometristen tot economische kennis proberen te komen. Daarover gaat §12.

## 12. De grondslag van de econometrie. Een vierde conclusie

Frisch stelt: 'Het is niet nodig te benadrukken wat de kwantitatieve formulering van begrippen en verbanden in de natuurwetenschappen heeft betekend. Voor meer dan een generatie is het mijn diepste overtuiging geweest dat de econometrische kwantificatiepoging even belangrijk is in de economie' (*Van utopische theorie naar praktische toepassingen; de casus econometrie*, 1970, p.12). Zeker, zonder economische theorie probeert de econometrie via kwantificaties en formules tot iets te komen, maar in de natuurkunde gaat het andersom. Daarin gaat een inzicht steeds aan een kwantificatie of formule vooraf. Zo was er eerst de golftheorie voor het licht en dan pas het opmeten van kleuren in golflengtes, en niet andersom zoals Frisch ons wil doen geloven en voor de econometrie bepleit.

En hoe worden modellen in de natuurkunde gebouwd? Niet zoals in de econometrie waarin men vaak tientallen formules tegelijk poneert voor een deelgebied van de economie. In de natuurkunde daarentegen is een model voor één van haar deelgebieden in de eerste plaats een geïntegreerde reconstructie. Zo vatte Rutherford's atoommodel van 1911 samen wat er toen zoal over het atoom bekend was. Bohrs atoommodel van 1913 deed hetzelfde met de atoomfysische kennis van toen. Enzovoort. Dat wil zeggen, ter verklaring van allerlei atoomfysische verschijnselen zijn er allerlei wetten en theorieën geopperd waarvan slechts een deel de empirische toets heeft doorstaan. Die houdbare wetten en theorieën in één systeem integreren heet dan het atoommodel van dat moment. Daar kunnen meer verbanden, wetten en vergelijkingen in voorkomen dan in het vorige atoommodel, maar ook kunnen bepaalde vergelijkingen die eerder slechts vier termen hadden, er nu zeven hebben. Hetzelfde geldt voor het model van een ster, voor het model van de aardkern, enzovoort.

De term 'model' wordt in de natuurkunde ook gebruikt in de betekenis van 'verklaringspoging'. Vaak krijgt zo'n verklaringspoging een metaforische naam. In het druppelmodel voor de atoomkern bijvoorbeeld nam Bohr aan dat een atoomkern een radioactief deeltje zou absorberen zoals een loshangende waterdruppel damp- en stoomdeeltjes absorbeert; hij deed dat onder meer om te verklaren waarom sommige stoffen radioactieve straling niet doorlaten. Een ander voorbeeld is het schillenmodel. Daarin doet men alsof de elektronen zich in van elkaar gescheiden schillen om de kern bevinden; dit onder meer ter verklaring van het feit dat edelgassen (helium, neon, enzovoort) geen scheikundige reacties lijken aan te gaan (vóór 1962 dacht men dat; sindsdien weten we dat edelgassen wel scheikundig reageren, maar zeer met mate).

Is zoiets integratiefs als in het atoommodel van jaar X of zoiets verklarends en metaforisch als in het druppelmodel en het schillenmodel ook met de ongeveer 30 formules van het JADE-model aan de hand? Als we alleen al naar de consumptievergelijking kijken, krijgt men bepaald niet de indruk. Zoals we in §§4-6 hebben gezien noemt Don geen concrete literatuur over onderzoek naar de economische geldigheid van de consumptievergelijking. Als we bovendien naar de cijfers kijken in zijn brief van 26 april 1999, dan wordt op slag duidelijk dat we iedere hoop op zo'n aangetoonde houdbaarheid maar beter kunnen laten varen. Voor elk van de zeven termen uit de consumptievergelijking krijgt het CPB namelijk deze gemiddeldes en spreidingen en/of onzekerheidsmarges (de relatieve onzekerheidsmarges, uitgerekend met de algemene meetfoutentheorie, staan tussen haakjes): AP =  $0,0000348 \pm 0,0082764$  (23.782%), BQ =  $-0,0000569 \pm 0,0033015$  (4.634%), CR =  $0,0195582 \pm 0,0227112$  (116%), DS =  $0,0145582 \pm 0,0101143$  (69%), ET =  $0,0032625 \pm 0,0228451$  (700%), FU =  $-0,0002314 \pm 0,0015811$  (683%) en

$GV = -0,0001076 \pm 0,0071274$  (6.624%). Met relatieve onzekerheidsmarges tussen 69% en 23.782% volstaat dit korte commentaar: die aangenomen verbanden hebben geen geldigheid.

En wat het metaforische betreft, in het JADE-model wordt niet een bepaald economisch verschijnsel met een eigenschap van de edelsteen jade vergeleken. JADE staat slechts voor *Joint Analysis of Dynamics and Equilibrium* en bestaat, als gezegd, slechts uit zo'n 30 econometrische formules.

Anders dan econometrische formules vertonen begrotingskundige formules wel gelijkenis met de werkwijze in de natuurkunde. Zo kan men een formule opstellen voor de loonkosten per functionaris over N jaar, waar in §2 een concreet voorbeeld van is uitgewerkt. Als de loonkosten nu S bedragen en de loonkostenstijging per jaar is s%, dan zijn de loonkosten per functionaris over N jaar,  $S_N: S \times (1 + s/100)^N$ . ( $S_1 = f 30.600,-$ ;  $S_2 = f 31.212,-$ ;  $S_3 = f 31.836,24$ ;  $S_4 = f 32.472,96$ .)

Die formule kan worden uitgebreid, precies zoals een formule in het atoommodel van 1999 ten opzichte van die in het atoommodel van 1990 kan zijn uitgebreid. Als er nu F functionarissen in dienst zijn en er komen er per jaar f bij, dan is het aantal functionarissen over N jaar,  $F_N: F + Nf$ . (Voor partij P:  $F_1 = 12.000$ ;  $F_2 = 14.000$ ;  $F_3 = 16.000$ ;  $F_4 = 18.000$ . En voor partij Q:  $F_1 = 9.000$ ;  $F_2 = 8.000$ ;  $F_3 = 7.000$ ;  $F_4 = 6.000$ .) Dus, over N jaar bedragen de totale loonkosten voor deze sector,  $L_N: S_N \times F_N = \{S \times (1 + s/100)^N\} \times \{F + Nf\}$ .

In de econometrie redeneert en werkt men echter anders dan in de natuurkunde en in de begrotingskunde. Econometristen poneren één of meer formules tussen economische grootheden, zoals in de consumptievergelijking. Hun empirische werk bestaat er dan vooral uit 'constanten' te bepalen en onzekerheidsmarges op papier te verkleinen, maar niet uit het natrekken van de gedaante van die formule(s) zelf. P in de consumptievergelijking zou immers ook  $\Delta$  of  $\Delta^3$  kunnen bevatten en in plaats van  $\ln w$  zou het ook  $(w - 1)/w$ ,  $1/(w^2 + 1)$  of nog wat anders kunnen zijn. In de econometrie staat de eenmaal geponeerde formule dan ook niet ter discussie, terwijl zij juist voorwerp van voortdurend onderzoek (geweest) zou moeten zijn. Een econometrische formule wordt hooguit na verloop van tijd ineens vervangen door een geheel andere econometrische formule waarvoor dezelfde bezwaren gelden.

De grondslag van de econometrie is diverse keren aan bod geweest in onze briefwisseling. Op 29 april 1998 (punt 13) bijvoorbeeld schrijf ik over het FREIA-KOMPAS-model van het CPB dat zelfs ongeveer 1.300 vergelijkingen gebruikt (Don, *Forecast uncertainty in economics* 1994, p.571): 'Deze veelheid aan vergelijkingen acht ik buitengewoon merkwaardig in het licht van het feit dat de econometrie qua wetenschappelijke werkwijze op de natuurkunde zou lijken, zoals onder meer Tinbergen in *Meten* (p.1-21) stelt. Immers, in de natuurkunde is er over één geldige en betrouwbare formule vaak door tientallen mensen tientallen jaren gewerkt'. Nadat ik heb laten zien dat Tinbergens voorstelling van zaken omtrent de wijze waarop natuurkundige kennis verworven zou zijn, voor Newtons zwaartekrachttheorie *niet* klopt, besluit ik: 'Kunt u of kan een andere econometrist (liefst van het CPB) die thuis is in de grondslag van de econometrie, mijn voorstelling van zaken ontcrachten? Zo ja, dan is die grondslag wellicht in orde. Zo nee, dan klopt ze niet: welke consequenties, wetenschappelijke en praktische (ik denk aan de huidige doorrekeningen), trekt u dan daaruit?'

In zijn brief van 15 juni 1998 antwoordt Don onder meer: 'Overigens is Tinbergens visie uit de jaren vijftig en zestig niet maatgevend voor het actuele denken over de grondslagen van de econometrie' (punt 13). Daarmee verdrievoudigt Don de problemen: (i) Hoe en om welke redenen is afscheid genomen van de grondslag van de econometrie, waarin naar de natuurwetenschappelijke methode wordt verwezen? (ii) Welke is de huidige grondslag van de econometrie volgens Don en/of het CPB zelf? (iii) Wat pleit er voor die kennelijk niet-natuurwetenschappelijke grondslag?

In mijn brief van 3 december 1998 vraag ik het zo: 'a. Wat was/waren de reden(en) om afstand te nemen van Tinbergens visie op de econometrie? b. Welke is de huidige grondslag van de econometrie en waar berust die op? Bij de beantwoording van deze vraag mag u zich wat mij betreft beperken tot de grondslag van de econometrie zoals die door het CPB wordt gebezigd, bijvoorbeeld ten aanzien van de consumptie-vergelijking' (punt 22).

In zijn reactie van 19 januari 1999 gaat Don hier niet op in en neemt hij slechts uitdrukkelijk

afstand van de oorspronkelijke grondslag van de econometrie, wat duidt op een terugtrekbeweging: 'Persoonlijk heb ik grote moeite met vergelijkingen tussen natuurwetenschappen, zoals de natuurkunde, en sociale wetenschappen, zoals de economie' (punt 22).

Wat we al eerder zagen keert ook hier weer terug: Don geeft noch een concreet antwoord noch literatuur. De econometrie is daarom niet alleen een kaartenhuis (§§4-11), het is ook nog eens op drijfzand gebouwd. Mijn vierde conclusie, vanuit welke de eerste drie te begrijpen zijn, luidt daarom:

**De reden dat het econometrische voorspellen faalt is dat de grondslag van de econometrie goddeels afwezig is en volstrekt ondeugdelijk is, voorzover hij toch bestaat.**

### **13. Versterkt of verzwakt de econometrie de democratie? Een vijfde conclusie**

Geen economische betekenis van modellen en formules achter de econometrische voorspellingen van het CPB, slechte voorspellingen, grote onzekerheidsmarges, te vaak geen publiceren van onzekerheidsmarges, geen deugdelijke grondslag voorzover er een is: is het wel verantwoord onze democratie daar deels op te laten steunen bij het nemen van beleidsbeslissingen?

Waarschijnlijk vanuit een optimistisch geloof in de maakbaarheid van de samenleving zoals dat in de jaren dertig en vijftig breed heerste, hebben Tinbergen en Frisch mijns inziens oprecht gemeend dat de econometrie een goede bijdrage zou kunnen zijn aan de democratie in het algemeen en aan haar economie in het bijzonder.

Frisch bijvoorbeeld zag in plannen op grond van de econometrie 'de basis voor efficiëntie en een levende democratie' (*Van utopische theorie naar praktische toepassingen*, 1970, p.27). En Tinbergen was van mening dat de 'optimale sociale orde' meetbaar is en vond daarom dat verschillen van mening en/of belang soms een ongewenst effect hadden, dat door de econometrie uitgesloten zou kunnen worden (*Metten in de menswetenschappen*, p.27, p.31-34, p.38, p.44v en p.46).

Ik twijfel niet aan Frisch' en Tinbergens democratische instelling. In *Metten* bijvoorbeeld komt Tinbergen me als een oprechte democraat voor: 'Het parlement oordeelt, door een meerderheidsstemming' (p.31). De vraag is echter of een oprechte instelling ook altijd het gewenste doel bevordert. We kunnen die vraag nu beter overzien dan in de jaren dertig: toen moesten allerlei pogingen om de samenleving deels langs wetenschappelijke weg te helpen maken nog gestalte krijgen, maar in 1999 kunnen we terugkijken op een overmaat aan falende voorspellingen van het CPB (§11).

Gaan we eerst op Frisch' woorden in. In wezen zijn econometrische voorspellingen een vorm van schijnkennis en die kan nooit efficiënt zijn. Hooguit geven hard ogende getallen beleidsmakers een prettig maar bedrieglijk gevoel van zekerheid. En het CPB doodt de democratie eerder dan dat ze haar verlevendigt. De doorrekeningen bijvoorbeeld halen allerlei keuzemomenten weg. We kunnen niet kiezen tussen zon en regen, maar wel tussen een werkweek van 32, 37 of 42 uur, tussen verschillende groottes voor Schiphol, tussen verschillende lokaties voor de nationale vlieghaven (Schiphol, voor de kust, Markerwaard), enzovoort. Door de econometrisering van de politiek en van het beleid zijn die keuzemomenten drastisch afgenomen en/of impliciet al door het CPB ingevuld. Mijns inziens verklaart dat een deel van de geringer geworden ledenaantallen van de politieke partijen en van de teruglopende opkomstcijfers bij verkiezingen. Het gewone partijlid kan zich het vuur uit de sloffen lopen, maar uiteindelijk geeft de uitkomst van een grotendeels ondoorzichtig blijvend econometrisch model de doorslag. Bovendien is er vlak vóór het verschijnen van de doorrekeningen druk overleg tussen het CPB en de programmacommissies van de betrokken partijen. Niet alleen is dat vanuit wetenschappelijk standpunt bekeken een rare zaak (zijn de uitkomsten nu resultaat van objectief onderzoek of van onderhandelingen?), maar bovendien worden de partijleden en het door hen vastgestelde verkiezingsprogramma op die manier gepasseerd. En uiteindelijk valt er voor kiezers en gekozen volksvertegenwoordigers minder te kiezen.

Dan Tinbergens standpunt. In §11 hebben we gezien dat sociale systemen zich aan objectieve meetprocedures onttrekken. Geen enkele 'sociale orde' is daarom meetbaar en al helemaal niet de 'optimale sociale orde', zo die al als idee bestaat (waar het ene volk zich overwegend prettig bij voelt, daar gruwet het andere volk in meerderheid van). Bovendien zijn verkiezingen er nu juist om bepaalde beslissingen aan het kiezersvolk zelf voor te leggen. Verschillen van mening en/of van belang zijn daarbij een onomzeilbaar gegeven en voeden juist de discussie en de democratie. Het zijn dus juist geen ongewenste elementen met ongewenste effecten. Natuurlijk zijn er ongewenste effecten van allerlei aard, ook in de economie. Gezien §§4-12 zie ik echter niet goed in welke rol de econometrie zou kunnen spelen in het uitsluiten daarvan. Mijns inziens kunnen en moeten ongewenste effecten zo veel mogelijk worden voorkomen met geschikte wetten en door ingebouwde controlemomenten, maar hoe die wetten en controles er uit moeten zien is andermaal een kwestie van democratisch beslissen en geen kwestie van meten.

Laat ik daarom een vijfde conclusie formuleren:

**Econometrische berekeningen van het CPB in het algemeen en die van de doorrekeningen in het bijzonder verzwakken in plaats van versterken de democratie.**

#### **14. Slotconclusie**

Dat de politieke partijen bij het presenteren van hun verkiezingsprogramma's van eenzelfde behoedzame basisscenario uitgaan en dat ze begrotingskundige berekeningen uitvoeren, is nuttig. Beide hebben echter niets econometrisch en kunnen ook zonder het CPB gedaan worden. Voor het overige zijn er slechts bezwaren tegen de doorrekeningen omdat de econometrie er een doorslaggevende rol in speelt. Ik vat de vijf conclusies kort samen:

1. Econometrische formules verwijzen niet naar de economische werkelijkheid.
2. Onzekerheidsmarges ontbreken ten onrechte.
3. Onzekerheidsmarges zijn (te) groot, veelal zelfs groter dan 100%.
4. De econometrie heeft geen duidelijke grondslag en voorzover hij er wel is, deugt hij niet; dit verklaart de vele missers in haar voorspellingen.
5. Econometrie, die van het CPB en anderszins, doet de democratie geen goed.

Mijn los geformuleerde slotconclusie luidt dan ook:

**De doorrekeningen van het CPB zijn een kaartenhuis dat op drijfzand is gebouwd. Voor zover onze democratie in dat kaartenhuis haar intrek heeft genomen, stelt ze haar voortbestaan op het spel. De politieke partijen en de overheid dienen er zo spoedig mogelijk uit te verhuizen. Dat wil zeggen, ze dienen beleid te formuleren en uit te voeren zonder voorspellingen die mede zijn verkregen met econometrische formules en berekeningen.**

Ewald Vervaet  
Amsterdam, 20 mei 1999