STOCHASTISCH DENKEN

door Prof. Dr. J. Koerts

1 Determinisme en stochastiek

De vraag, welke methodologie men bij het economisch onderzoek dient toe te passen, is nog steeds niet volledig beantwoord! Typerend voor deze situatie is wellicht de opmerking die T. C. Koopmans maakt in zijn bekende tweede essay:¹⁾ „After more than a century of intensive activity in scientific economics, two economists who have made outstanding contributions to our science (Lionel Robbins en Milton Friedman), seek the ultimate basis of economic knowledge in considerations which (a) contradict each other and (b) are subject to strong objections. One is led to the conclusion that economics as a scientific discipline is somewhat hanging in the air.”

Met name de vraag hoe men uit empirisch materiaal moet leren, vooral indien men maar over beperkt cijfermateriaal beschikt, is nog steeds niet afdoende beantwoord. Al is het dus nog niet geheel duidelijk hoe een econo- mist tot de constructie van zijn theorieën moet komen, over enkele be- langrijke uitgangspunten heeft men overeenstemming bereikt. Deze luiden:

- 1 Al de veronderstellingen (postulaten) die men gebruikt moeten expliciet worden geformuleerd. Alle conclusies waartoe een theorie aanleiding geeft moeten uit de postulaten kunnen worden afgeleid.
- 2 Een theorie moet logisch-consistent zijn. Dat wil zeggen: het moet niet mogelijk zijn tegenstrijdige conclusies uit eenzelfde verzameling postula- ten af te leiden.
- 3 Theorieën dienen aan de hand van cijfermateriaal te kunnen worden getoetst.²⁾ Anders geformuleerd: de mogelijkheid tot verwerping van de theorie dient aanwezig te zijn. Theorieën die in strijd blijken te zijn met de feiten, dienen te worden herzien.

Vooral het gestelde onder punt 3, waarover overigens veel strijd is gevoerd, heeft grote invloed gehad op de formulering van economische theorieën. In feite is het dit empirische uitgangspunt geweest dat de economen genoopt heeft het waarschijnlijkheidsbegrip en daarmee de gehele waarschijnlijkheids- rekening in de economische analyse te introduceren. Om dit in te zien moet men zich realiseren, dat economen altijd op zoek zijn geweest naar wet- matigheden in het economische leven. Om het wat nauwkeuriger te stellen: het is altijd een basishypothese van economen geweest dat er exacte rela- ties tussen economische variabelen zouden bestaan. Deze gedachte komt wellicht voort uit een deterministische wereldvisie. Echter, zodra naar de feiten wordt overgestapt doemt er voor de econo- mist de volgende moeilij- kheid op. De theorieën die hij wil toetsen gaan uit van exacte relaties, doch de feiten zullen iedere exacte a priori uitspraak die uit een theorie wordt verkre- gen tegenspreken. Anders geformuleerd: exacte theorieën zijn altijd in strijd

¹⁾ Koopmans, T. C. (1957), *Three Essays on the State of Economic Science*, New York.

²⁾ We gaan hier niet in op de vraag „wat dan wel moet worden getoetst”. Zie hiervoor J. Koerts en A. P. J. Abrahamse „On the Theory and Application of the General Linear Model”, Rotterdam University Press (1969).

met de feiten, hetgeen impliceert dat iedere theorie in het licht van de cijfers moet worden verworpen. Beschouw bijvoorbeeld een gedragsrelatie in de vorm van een eenvoudige vraagcurve, die de verbruikte hoeveelheid van een goed (q) afhankelijk stelt van zijn prijs (p). In de praktijk van het economisch onderzoek blijkt nu, dat men er niet in slaagt q exact te verklaren uit p . Dit betekent, dat de berekende waarde van q volgens de door de econoom gespecificeerde relatie, afwijkt van de werkelijk waargenomen waarde van q . Dit verschijnsel treedt op onafhankelijk van de gekozen functionele relatie en blijft ook bestaan indien men nog andere variabelen in de analyse introduceert.

Men kan zich afvragen of alleen economen door dit verschijnsel worden geplaagd. Deze vraag moet stellig ontkennend worden beantwoord. In geen enkele wetenschap heeft men het bestaan van exacte functionele verbanden kunnen aantonen. H. Jeffreys wijst hier in zijn bekende boek *Theory of Probability* (1961) nog eens expliciet op als hij stelt: A physical law, for practical use, cannot be merely a statement of exact prediction, if it was it would invariably be wrong and would be rejected at the next trial! Quantitative prediction must always be prediction within a margin of uncertainty.³⁾

Bovengenoemde stand van zaken had een tweetal gevolgen. In de eerste plaats werden de economen gedwongen na te denken over de vraag waar de afwijkingen vandaan kwamen en in de tweede plaats, wilden zij hun theorieën aan de hand van cijfermateriaal kunnen toetsen dan moesten deze op zodanige wijze worden uitgebreid, dat met de mogelijkheid van afwijkingen rekening kon worden gehouden. Met andere woorden: theorieën moesten het bestaan van afwijkingen toelaten of zoals Haavelmo het eens heeft geformuleerd: „we need theories that, without involving us in direct logical contradictions, state that observations will as a rule cluster in a limited subset of the set of all conceivably observations, while it is consistent with the theory that an observation falls outside this subset „now and then”.

In de loop der tijden zijn verschillende argumenten naar voren gebracht die een verklaring voor het ontstaan van afwijkingen beogen te geven:⁴⁾

- 1 De onderzoeker maakt *meetfouten*.
- 2 Niet alle relevante variabelen zijn in de analyse betrokken. (verwaarloosde factoren)
- 3 *Overige specificatiefouten*.

Ad 1. Dit argument is wel het meest voor de hand liggende. Het betekent, dat de onderzoeker fouten maakt zodra hij de numerieke waarde van een economische variabele vaststelt. Het behoeft geen betoog dat econometristen meetfouten maken, het is haast ondoenlijk ze te vermijden. Dit argument is dan ook zeer acceptabel en kan zeer goed als verklaring voor niet te grote afwijkingen dienen. Vooral in de natuurwetenschappen heeft het verschijnsel

³⁾ De eerlijkheid gebiedt echter te zeggen dat de afwijkingen in sommige gevallen bijzonder klein zijn. De onderzoeker handelt dan alsof hij te maken heeft met een exact functioneel verband.

⁴⁾ Het onderscheid tussen 1, 2 en 3 is niet erg principieel. Men kan (1) en (2) ook als specificatiefouten behandelen. Wij sluiten hier gemakshalve maar aan bij deze gebruikelijke indeling.

„meetfouten” de aandacht getrokken en is daar onderwerp van nauwkeurige studie geworden. We komen hierop nog terug.

Ad 2. Het spreekt vanzelf dat verwaarloosde factoren aanleiding geven tot afwijkingen doch men kan zich afvragen waarom de econoom niet alle variabelen in zijn analyse betreft. Een aantal redenen zijn hiervoor aan te geven:

a. De econoom weet dat een zeer groot aantal factoren - die ieder een geringe invloed op de te verklaren variabele bezitten - een rol spelen. Het is echter bij de huidige stand van de wetenschap ondenkbaar dat ieder van deze „kleine” factoren afzonderlijk in de analyse kunnen worden opgenomen. De econoom beperkt zich dan tot enkele „grote” factoren en laat *noodgedwongen* de „kleinere” factoren buiten beschouwing.

b. De econoom beschikt niet over de juiste gegevens met betrekking tot een bepaalde variabele. Helaas komt het in de economie nog veelvuldig voor dat men niet over het juiste cijfermateriaal beschikt.

c. Het is helaas niet altijd mogelijk een relevante variabele te kwantificeren, d.w.z. zijn waarde in een getal vast te leggen. Het kwantificeren van begrippen is een geheel aparte kunst. Lang niet altijd is het duidelijk of men wel een bevredigend kwantitatief explicatum voor een bepaald begrip zal weten te vinden. Een voorbeeld van een begrip dat nog steeds niet bevredigend is gekwantificeerd lijkt mij het begrip „nut” zoals dat wordt gehanteerd in de theorie van het consumentengedrag.⁵⁾

d. Men vermoedt wel dat er nog verscheidene andere variabelen in het spel zijn doch weet niet precies welke.

Ad 3. Onder overige specificatiefouten zullen wij verstaan alle afwijkingen van het model t.o.v. de werkelijkheid die ontstaan tengevolge van ontbrekende verbanden, de introductie van *verkeerde* functionele relaties, wiskundige benaderingen, het introduceren van variabelen die niet in de analyse thuis horen en substitueert variabelen. Bij dit laatste begrip moet men denken aan het feit dat men bij gebrek aan cijfers, niet de „juiste” variabele in de analyse heeft opgenomen doch een substitueert variabele waarvan men hoopt dat deze de invloed van de „juiste” variabele gedeeltelijk zal weergeven.

Voor wij verder gaan is het van belang op te merken dat in het voorafgaande verklaringen voor het ontstaan van afwijkingen zijn gegeven zonder dat het hierbij nodig was de hypothese van exact functionele verbanden te verwerpen. Wij blijven dus voorlopig nog uitgaan van een *deterministische* visie.

Laten we nu ingaan op de vraag hoe men de economische theorieën kan formuleren zodat confrontatie met cijfermateriaal mogelijk wordt. Bij de oplossing van dit probleem hebben de economen gebruik gemaakt van de ervaring die onderzoekers in andere takken van wetenschap al hadden opgedaan met een bepaald type verschijnsel, een zogenaamd „*kansverschijnsel*”.

Definitie 1: Een kansverschijnsel is een verschijnsel, dat indien het zich

⁵⁾ Recente pogingen om dit begrip te meten zijn ondernomen door Van Praag.

onder zoveel mogelijk gelijkblijvende omstandigheden afspeelt niet steeds leidt tot eenzelfde uitkomst, doch tot verschillende uitkomsten, echter zodanig dat er stabiliteit optreedt in de relatieve frequentie waarmee de mogelijke uitkomsten worden waargenomen indien het kansverschijnsel zich een groot aantal malen afspeelt.

Als voorbeeld van een kansverschijnsel kunnen wij het werpen met een munt nemen. Dit is, ik geef het toe, een zeer eenvoudig „experiment” maar wij kunnen er de definitie goed mee illustreren. In de eerste plaats heeft het experiment twee mogelijke uitkomsten, het is dus geen deterministisch experiment, de uitkomst is onzeker. In de tweede plaats mag het waar zijn, dat de uitkomst onzeker is, toch hebben wij allen het vermoeden dat, indien de munt vaak wordt opgegooid, de beide uitkomsten ongeveer even vaak zullen voorkomen. Dit vermoeden wordt bevestigd indien men een groot aantal malen gaat werpen. Met andere woorden: de relatieve frequentie van het aantal malen dat kruis boven komt wordt *stabiel* en is ongeveer gelijk aan 0.5. Deze merkwaardige eigenschap - het stabiel worden van de relatieve frequentie bij een toenemend aantal waarnemingen - vindt men op talrijke gebieden. Om een enkel voorbeeld te geven: de relatieve frequentie van het aantal jongensgeboorten is ongeveer stabiel en schommelt dicht rond de waarde 0.51. Deze empirisch gevonden regelmaat is van groot belang; zij geeft de mogelijkheid tot een bepaalde interpretatie van het begrip waarschijnlijkheid en geeft ons bovendien de mogelijkheid het gedrag van de munt op een geheel nieuwe wijze te beschrijven, n.l. met behulp van een waarschijnlijkheidsverdeling!

Als wij stabiele frequenties waarnemen, dan ligt het voor de hand om het bestaan van een getal p te postuleren, dan kan worden gezien als de wiskundige idealisering van een stabiele frequentie. De waargenomen relatieve frequentie is dan te beschouwen als een afspiegeling van dit onbekende getal. Het getal p noemt men per definitie de *waarschijnlijkheid* van de desbetreffende gebeurtenis. Merk op, dat het begrip waarschijnlijkheid op deze manier gebruikt, via de stabiele frequentie gekoppeld is aan het begrip kansverschijnsel. Men spreekt in dit verband wel van *objectieve* waarschijnlijkheid omdat deze afhangt van de objectieve omstandigheden waaronder het experiment zich afspeelt en niet van de persoon, die het kansverschijnsel waarneemt of doet afspelen. We zijn nu in staat het begrip „waarschijnlijkheidsverdeling” vast te leggen.

Definitie 2: Een waarschijnlijkheidsfunctie(verdeling) is de specificatie van een rij van mogelijke uitkomsten waartoe een kansverschijnsel aanleiding geeft tezamen met de bijbehorende kansen waarmee een aantal mogelijke uitkomsten zich zullen voordoen.⁶⁾

Veronderstel, dat wij voor het probleem worden geplaatst het experiment te beschrijven dat bestaat uit het werpen van een munt. We zouden dan kunnen trachten de uitkomst van dit experiment te voorspellen aan de hand van de volgende variabelen: de snelheid waarmee de munt wordt opgegooid, het gewicht van de munt, de luchtweerstand, de conditie van het tafelblad

⁶⁾ Om redenen van eenvoud bezien we hier alleen discrete verdelingen.

waarop de munt valt, etc. Bij een dergelijke aanpak treden echter direct de volgende bezwaren naar voren:

- 1 Het aantal factoren dat van beslissende invloed is op de uitslag van het experiment blijkt reeds bij enig nadenken bijzonder groot.
- 2 We zijn er niet eens zeker van dat wij alle factoren die de uitslag van het experiment bepalen ook inderdaad kennen.
- 3 Bij de huidige stand van onze kennis ziet de structuur van ons experiment (kansverschijnsel) er zo gecompliceerd uit dat individuele voorspelbaarheid op praktische gronden onmogelijk lijkt.

Wij komen aldus tot de conclusie, dat ofschoon het denkbaar is dat in de verre toekomst een theorie wordt ontworpen die al de relevante variabelen incorporeert en die dan ook tot exacte voorspelling aanleiding geeft, wij heden ten dage een stap terug moeten doen en een andere oplossing zullen moeten gaan zoeken.

Een oplossing van ons probleem wordt gevonden in de beschrijving van het experiment met behulp van een waarschijnlijkheidsverdeling. Dat betekent, dat wij de exacte voorspelling van een individuele worp opgeven en erkennen dat wij niet weten *wanneer* er kruis boven zal komen. Wij stellen slechts vast dat de kans op de uitkomst kruis gelijk is aan 0.5, hetgeen impliceert dat wij wel weten *hoe vaak* er ongeveer kruis boven zal komen indien wij een groot aantal malen met de munt werpen. Deze aanpak is natuurlijk een stapje terug doch het is veel beter dan „niets”.

De econometist past nu, om uit zijn problemen te komen, een analoge redenering toe. Ook hij wordt regelmatig geconfronteerd met situaties zoals aangegeven onder 1, 2 en 3 in de voorgaande alinea en daarom doet ook hij een stapje terug en tracht gebruik te maken van waarschijnlijkheidsverdelingen in zijn analyse. Wij zullen nu een tweetal methoden (modellen) kort bespreken. Het eerste model is volkomen te vergelijken met de situatie „het opgooien van een munt”. Dat wil zeggen: de econometist bestudeert een verschijnsel zoals bijvoorbeeld: het aantal auto's dat in een bepaald uur van de dag een tunnelingang passeert, de lengte van interlokale telefoongesprekken, de tijd die verstrijkt tussen twee achtereenvolgende aankomsten van binnenkomende vliegtuigen op Schiphol, etc. In al deze gevallen wijst hij, net als bij de munt, niet eens meer de factoren (variabelen) aan die voor het verschijnsel in kwestie van belang zijn. Hij beschrijft het verschijnsel met een eenvoudige waarschijnlijkheidsverdeling! Exacte uitspraken zijn dus niet meer mogelijk, alleen de kans op bepaalde gebeurtenissen wordt aangegeven.

Het tweede en in feite zeer veel gebruikte model is van een gemengd type. Hier wijst de econometist wel degelijk een aantal variabelen aan die hij van groot belang acht (de z.g. „grote” factoren). Daarnaast onderkent hij nog het bestaan van een veelheid van „kleine” factoren die hij niet meer afzonderlijk specificeert, hun invloed beschrijft hij met behulp van een waarschijnlijkheidsverdeling. De gang van zaken valt als volgt in symbolen uit te drukken. Volgens de econometist wordt bijvoorbeeld de variabele y : stel de uitgave aan biefstuk per maand van een bepaald gezin, bepaald door een groot aantal variabelen $x_1, \dots, x_n$, zoals de prijs per kg van biefstuk, de prijzen van

andere vleessoorten, de prijzen van verschillende vissoorten, de gezinssamenstelling, etc. In formule

$$(a) \quad y = f(x_1, \dots, x_n)$$

Daar het technisch onmogelijk is om alle variabelen in de analyse te betrekken beperkt hij zich tot de naar zijn mening belangrijkste twee, bijvoorbeeld het gezinsinkomen (x_1) en de prijs/kg van biefstuk (x_2). De invloed van de overige „kleine” variabelen wordt nu beschreven met behulp van een waarschijnlijkheidsverdeling. De specificatie zoals hierboven onder (a) aangegeven wordt nu als volgt omgezet:

$$(b) \quad y = g(x_1, x_2) + u$$

waarin u een kansvariabele voorstelt die de invloed van de „kleine” factoren moet weergeven. Ook in dit model zijn exacte uitspraken niet langer meer mogelijk. Stel, dat het gezinsinkomen x_1 f 2.000,— per maand bedraagt en dat de prijs van biefstuk x_2 , f₁/ons is, dan stelt het model ons alleen maar in staat tot waarschijnlijkheidsuitspraken over de uitgave aan biefstuk. Wat de uitgave in een specifiek geval zal zijn kunnen wij niet aangeven, de specificatie onder (b) verschaft ons alleen een rij van mogelijke uitkomsten, ieder met zijn bijbehorende waarschijnlijkheid.

Wat zijn nu de voordelen van een dergelijke „stochastische” specificatie? In de eerste plaats bereikt de econometrist een logisch gesloten probleemstelling,⁷⁾ zodat toetsing van zijn theorie nu op een niet triviale wijze kan plaatsvinden. Zijn theorie voldoet nu aan de eis zoals die werd geformuleerd door Haavelmo en die wij in het voorafgaande hebben geciteerd. In de tweede plaats komt nu het gehele arsenaal van methodieken zoals die in de wiskundige statistiek is ontwikkeld, tot zijn beschikking. Met name de begrippen zoals het „schatten van onbekende parameters” en het op statistische wijze toetsen van bepaalde hypothesen krijgen nu een helder gedefinieerde inhoud. Door deze werkwijze is een economische wet niet langer een exacte voorspelling maar een uitspraak over de kansen waarmee verschillende uitkomsten zich zullen voordoen.

2 Het verlaten van een deterministische wereldvisie

In het voorafgaande ontstond de noodzaak tot stochastische formuleringen uit menselijke tekortkomingen die, zo werd betoogd, wellicht in de toekomst te vermijden zouden zijn. Men kan zich afvragen of dit uitgangspunt wel houdbaar is. Met name de vraag of *meetfouten* wel te vermijden zijn heeft in de natuurkunde veel aandacht gekregen. Tegenwoordig wordt deze vraag door de moderne natuurkundigen ontkennend beantwoord. Wil men de baan van een deeltje, bijvoorbeeld van een elektron voorspellen, dan moet men op een zeker ogenblik zijn plaats q en impuls p (snelheid maal massa) kennen. Echter, zodra men de plaats nauwkeurig wil vaststellen dan moet men een lichtstraal van korte golflengte op het deeltje afsturen, doch door de botsing

⁷⁾ Hiertoe moet hij wel voldoende veronderstellingen omtrent de waarschijnlijkheidsverdeling opnemen.

met het deeltje wijzigt zich de impuls. Met andere woorden: hoe nauwkeuriger de plaats wordt bepaald (hoe groter het lichtkwantum dat men naar het deeltje stuurt) des te onnauwkeuriger de impuls gemeten kan worden! Dit effect, dat men door meting het milieu verstoort waarin gemeten moet worden is theoretisch gefundeerd in de onzekerheidsrelatie van Heisenberg. Deze zegt, dat bij elke waarneming de fout Δq op de plaats en de fout Δp op de impuls als volgt verbonden zijn

$$\Delta p \cdot \Delta q \sim k/2\pi$$

waarin k het werkingskwantum. Met andere woorden: volgens de moderne theorie is het *principiële* onmogelijk de baan van het deeltje exact te voorspellen omdat wij *nooit* de plaats q en de impuls p exact zullen kennen. Anders gezegd, wij zullen nooit over de noodzakelijke gegevens kunnen beschikken vanwege de meetfouten. Daardoor zijn alleen waarschijnlijkheidsuitspraken mogelijk.

In deze gedachtengang wordt de waarschijnlijkheidsrekening dus een essentieel hulpmiddel van natuurbeschrijving. Wellicht is het interessant nog te vermelden, dat men alleen door volledige verwerping van de bestaande theorie onder de relatie van Heisenberg uit kan. Oppoetsen van de oude theorie zodanig dat de relatie niet meer logisch volgt is uitgesloten.

Met overplanten van ideeën die in andere takken van wetenschap zijn ontwikkeld moet men voorzichtig zijn. Misschien verstoort de econometist eveneens door meting het milieu waarin gemeten moet worden. Wellicht zijn daardoor ook in de economie meetfouten essentieel onvermijdelijk. Men zou zich kunnen indenken dat op micro-niveau een dergelijk effect bestaat. Dit probleem is echter onvoldoende onderzocht; er bestaat dan ook geen analoog voor de onzekerheidsrelatie in de economie. Toch leeft ook bij economen de gedachte dat het idee van volledig determinisme verlaten moet worden, hetgeen wel blijkt uit een uitspraak van T. C. Koopmans als hij stelt: „In dynamic economics, the phenomenon itself is either *essentially* a stochastic process or needs to be treated as such because of the great number of factors at work. Hence the analysis and interpretation of economic data call for the application of the method of statistical inference.”

Of in de economie het nieuwe „stochastische” wereldbeeld inderdaad volledig de plaats van het vertrouwde determinisme zal gaan innemen kan alleen de toekomst ons leren, het heeft er wel alle schijn van, doch om beter zicht op deze vraag te krijgen zal nog veel economisch onderzoek nodig zijn. Voor de praktisch ingestelde econometist is deze vraag gelukkig niet van groot belang want determinisme of niet, hij zal in beide gevallen een stochastisch model dienen te gebruiken. Alleen een dergelijk model verschaft ons heden ten dage een logisch-consistent systeem waarin wij onze ideeën kunnen formaliseren in niet-strijdige veronderstellingen en het dwingt ons om alle acceptatie criteria die wij aanleggen expliciet te formuleren.